

COMMON GENETIC VARIANTS  
UNDERLYING  
COGNITIVE ABILITY

María Florencia Gosso

**Leescommissie:**

dr. W Crusio

prof.dr. NG Martin

dr. GJA Ramakers

dr. S v/d Sluis

prof.dr. CM van Duijn

prof.dr. M Verhage

This thesis was supported by the Universitair Stimulerings Fonds (grant number 96/22), the Human Frontiers of Science Program (grant number rg0154/1998-B), the Netherlands Organization for Scientific Research (NWO) grants 904-57-94 and NWO/SPI 56-464-14192, and by the Centre for Medical Systems Biology (CMSB), a center of excellence approved by the Netherlands Genomics Initiative/Netherlands Organization for Scientific Research (NWO).

The printing of this thesis was supported by, the Center for Neuroscience and Cognitive Research (CNCR), by the “Internationale Stichting Alzheimer Onderzoek” and by Lundbeck B.V.



ISBN: 9789086591664

Printed by: PrintPartners Ipskamp

Cover Design: María Florencia Gosso

Layout: María Florencia Gosso

Copyright <sup>®</sup> María Florencia Gosso, 2007, Amsterdam

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, without permission of the author or, when appropriate, of the scientific journal in which parts of this book have been published.

VRIJE UNIVERSITEIT

COMMON GENETIC VARIANTS  
UNDERLYING  
COGNITIVE ABILITY

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan  
de Vrije Universiteit Amsterdam,  
op gezag van de rector magnificus  
prof.dr. L.M. Bouter,  
in het openbaar te verdedigen  
ten overstaan van de promotiecommissie  
van de faculteit der Psychologie en Pedagogiek  
op maandag 7 januari 2008 om 15.45 uur  
in de aula van de universiteit,  
De Boelelaan 1105

door

María Florencia Gosso

geboren te Paraná, provincie Entre Ríos, Argentinië

|             |                       |
|-------------|-----------------------|
| promotoren: | prof.dr. D.I. Boomsma |
|             | prof.dr. P. Heutink   |
| copromotor: | dr. D. Posthuma       |

## TABLE OF CONTENTS

|                      |   |     |
|----------------------|---|-----|
| CHAPTER 1            | General Introduction  | 1   |
| CHAPTER 2            | Association between the <i>CHRM2</i> gene and intelligence  | 39  |
| CHAPTER 3            | Exploring the functional role of the <i>CHRM2</i> gene and cognition  | 55  |
| CHAPTER 4            | The <i>SNAP-25</i> gene: synaptic plasticity and cognitive ability  | 73  |
| CHAPTER 5            | The <i>SNAP-25</i> gene: regulatory variants and cognition  | 89  |
| CHAPTER 6            | A functional polymorphism in the $\beta 2$ -adrenergic receptors at an amino acid which differs between humans and chimpanzee explains phenotypic differences in intelligence | 109 |
| CHAPTER 7            | <i>COMT</i> and <i>DRD2</i> gene variants: evidence of positive heterosis and gene-gene interaction on working memory functioning   | 129 |
| CHAPTER 8            | Summary and Discussion  | 147 |
| SAMENVATTING         |   | 167 |
| RESUMEN              |   | 177 |
| APPENDICES           |   |     |
| I                    | Descriptives and QTDT association results for <i>DBH</i> , <i>DRD2</i> , <i>DRD3</i> , <i>HTR2A</i> , <i>SERT</i> , and <i>TH</i>   | 188 |
| II                   | Labwork Protocols   | 194 |
| LIST OF PUBLICATIONS |   | 197 |
| ACKNOWLEDGEMENTS     |   | 201 |



# CHAPTER 1

## GENERAL INTRODUCTION





## INTRODUCTION

### *Cognitive ability*

Intelligence has been one of the most studied quantitative behavioral traits for more than 100 years. Historically two main, contrasting concepts about the nature of intelligence have been formulated. The first concept, advocated by the so-called “*g*-theorists”, encompasses the idea of a single general factor called ‘*g*’, which accounts for the variance in test scores that is shared among subtests (Carroll 1993; Humphreys 1985; Jensen 1998; Spearman 1904). This general factor of intelligence (*g*), and the specific factors are represented by Spearman’s two-factor theory of abilities (Spearman 1904). Contrary to Spearman’s two-factor theory, Thurstone (1938) advocated his multiple factor analysis theory, from which relatively independent sub-components of intelligence, so-called Primary Mental Abilities (PMA’s), were obtained. However, intelligent behavior can not be explained by just these PMA’s, and also evidence for *g* was found by Thurstone. Thurstone’s final model therefore takes into account the presence of a general *g* factor, PMA’s, as well as test-specific factors (Thurstone 1947).

Psychometric intelligence tests consist of a number of component subtests that taken together are used to infer a general IQ (intelligence quotient) score. Intelligence tests such as the Revised Amsterdam Child Intelligence Test (RAKIT, (Bleichrodt *et al.* 1984), the Wechsler Intelligence Scale for Children Revised (WISC-R, Dutch version, (Van Haassen *et al.* 1986) and the Wechsler Adult Intelligence Scale (WAIS, (Wechsler 1997)) are theoretically based on Thurstone’s factor analysis theory (Thurstone 1938) and provide an index of general IQ and primary abilities such as word fluency, verbal comprehension, spatial visualization, number facility, associative memory, reasoning, and perceptual speed.

The Wechsler IQ tests have been broadly used for measuring intelligence quotient. First published in 1939 as the Wechsler-Bellevue Scale, Wechsler IQ tests have become among the most widely used tests to assess psychometric intelligence. Both the WISC and WAIS are standardized and frequently revised among different age strata and populations, making them a particularly useful psychometric tool to be used when comparing intellectual abilities among different age and population cohorts. Standardized IQ scores typically have a mean of 100 and a standard deviation of 15 IQ points. The Wechsler Adult Intelligence Scale (WAIS) and Wechsler Intelligence Scale for Children (WISC) consist of several sub-tests, each comprising a number of different items. Three main scores, namely Verbal IQ (VIQ), Performance IQ (PIQ) and Full IQ (FIQ) can be generated; as well as

different IQ dimensions (verbal comprehension, perceptual organization, processing speed, working memory).

Previous twin studies have established that general IQ is influenced by genetic factors at all ages. Heritability estimates increase from around 30% in preschool children to 80% in early adolescence and adulthood (Ando *et al.* 2001; Bartels *et al.* 2002; Boomsma & van Baal 1998; Bouchard & McGue 1981; Luciano *et al.* 2001; Petrill *et al.* 2004; Plomin 1999; Posthuma *et al.* 2001a). The stability of IQ performance during childhood is mainly driven by genetic influences. Bartels *et al.* (2002) and Petrill *et al.* (2004) showed in longitudinal designs that one common factor influenced IQ performance from early childhood to adolescence, and that the influence of this genetic factor is amplified when children grow older. If a trait is heritable, then the next step is to identify what causes this heritability. I will first provide a general introduction into how genetic variation in the human genome can lead to individual differences in trait values, and will then continue with summarizing the current efforts for identifying genetic variants for cognitive ability.

### *How does genetic variation lead to individual trait differences:*

#### *From DNA to phenotype*

The classical dogma of molecular biology is that RNA is transcribed from DNA and subsequently translated into proteins. DNA is made of a long sequence of smaller units strung together (nucleotides). The genetic code consists of 64 triplets of nucleotides (codons). With three exceptions, each codon encodes for one of the 20 amino acids used in the synthesis of proteins (Crick 1968). The human genome is the term used to describe the total genetic information in human cells. DNA located in the nucleus (genomic DNA – gDNA) accounts for the vast majority of the total genetic material (>99.99%); while the remaining genetic fraction is localized at the mitochondrial level (mitochondrial DNA - mDNA). The nuclear genome encompasses about 3,300 Mb (Lander et al. 2001).

#### *Genomic variation in the human genome*

Back in the early 1980s, functional as well as positional cloning were the only approaches available to link simple Mendelian traits to genes. Positional cloning, however, was a difficult and time-consuming approach and more systematic coverage of the human genome was required. The idea of having free access to an accurate human genome map was in its early developmental stage.

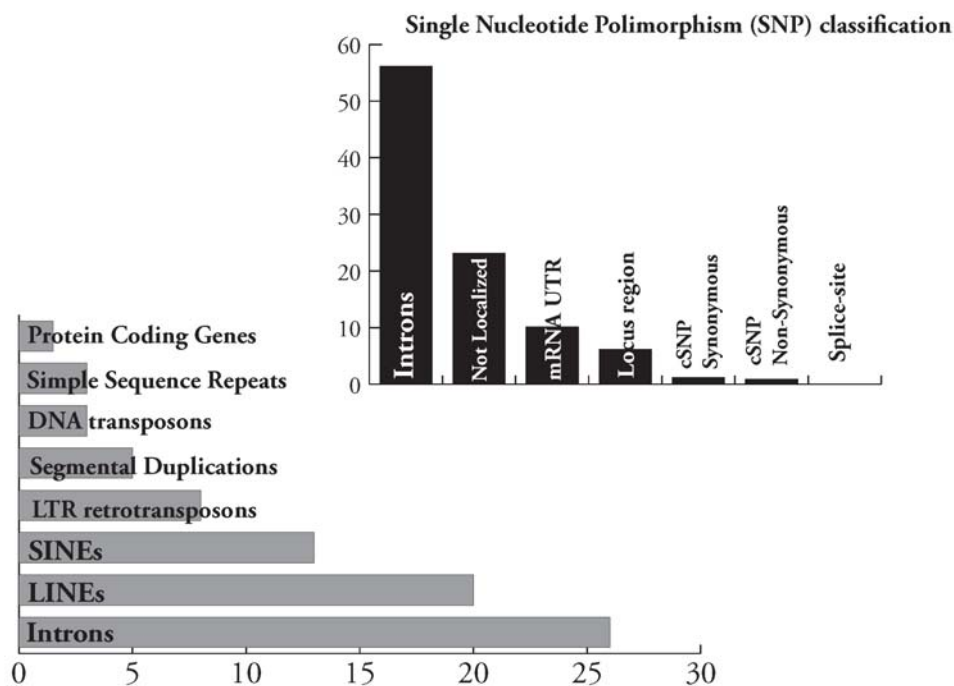
The mesmerizing idea of being able to obtain a global and more systematic view of genomes was the next goal for the scientific community; not only for the purpose of identifying all human genes and their mutations responsible for causing monogenic diseases, but more importantly: investigating genomic variation underlying complex (common) traits as well as obtaining a better understanding of complex regulatory networks.

With an exhaustive scientific worldwide collaboration the completion of the first sequence draft of the human genome was available in 2001 (Lander *et al.* 2001), and completed in 2003, 50 years after Watson & Crick's breakthrough (Watson & Crick 1953). At the beginning of the International Human Genome Project (IHGP), about 80,000 -100,000 genes were predicted to be encoded in our genome (Schuler *et al.* 1996). Posterior analysis of the gathered data revealed that this estimate was 5-fold larger than the actual number of protein-encoding genes observed (25,000-32,000) (Lander *et al.* 2001). Compared to other organisms humans have relatively little protein-encoding genes: rice for example has an estimated number of 50,000 genes. However transcription of genes is found to be more complex in humans, mainly as a result of alternative splicing (AS), which allows each gene to code for multiple proteins, depending on alternative splicing of coding exons.

AS allows different coding exon combinations to be included in transcripts from the same gene during RNA processing, therefore allowing single genes to be processed differently, conferring a broader number of functionally distinct proteins that can be obtained from a finite gene pool. More than 74% of multi-exon genes in humans are alternatively spliced (Johnson *et al.* 2003). This number however, is likely to be an underestimate since expressed sequence tags (ESTs) (they consist of short sequences of about 300-500 base pairs (bp) mRNA transcripts obtained from different tissues) only cover a portion of the transcript (preferentially 5' and 3' ends). In addition, it is likely that the number of original transcripts is underestimated, since AS can produce variants with different exon combinations within regions not tagged by ESTs. Even more important, rare but fundamental splicing events that occur in specific cell tissues at specific developmental stages may not be represented in ESTs collections (<http://www.ncbi.nlm.nih.gov/dbEST>) (Adams 1991). AS can occur in both different tissues and at different stages of tissue development (Black 2003; Lopez 1998). Strikingly, after correcting for differences in EST coverage among tissues, AS has been shown to occur at a significantly higher level in the brain compared with other tissues (Yeo *et al.* 2004), highlighting the important role of such a mechanism for neuronal signaling modulation (Lee & Irizarry 2003). It is

worth noting that despite being a fundamental mechanism for generating protein diversity, AS is just one of several regulatory processes eukaryotic organisms will apply during gene expression regulation (see Box 1.1).

Notably, the protein-coding DNA fraction is encompassed in less than 1.5 % of the total genome sequence (Lander *et al.* 2001; Venter *et al.* 2001). The 98.5 % (non)coding fraction, consists of a variety of genetic elements among which introns and transposable elements constitute the vast majority, followed by simple sequence repeats and segmental duplications (Lander *et al.* 2001) (see Figure 1.1).



**Figure 1.1** | Genomic variation in the Human genome. Percentages of genomic structures that are part of our genome. Note that the relatively low percentage of actively transcribed sequences (coding protein sequences, only to 1.5% of the total genetic sequence) compared to for example intronic sequences (~26%)

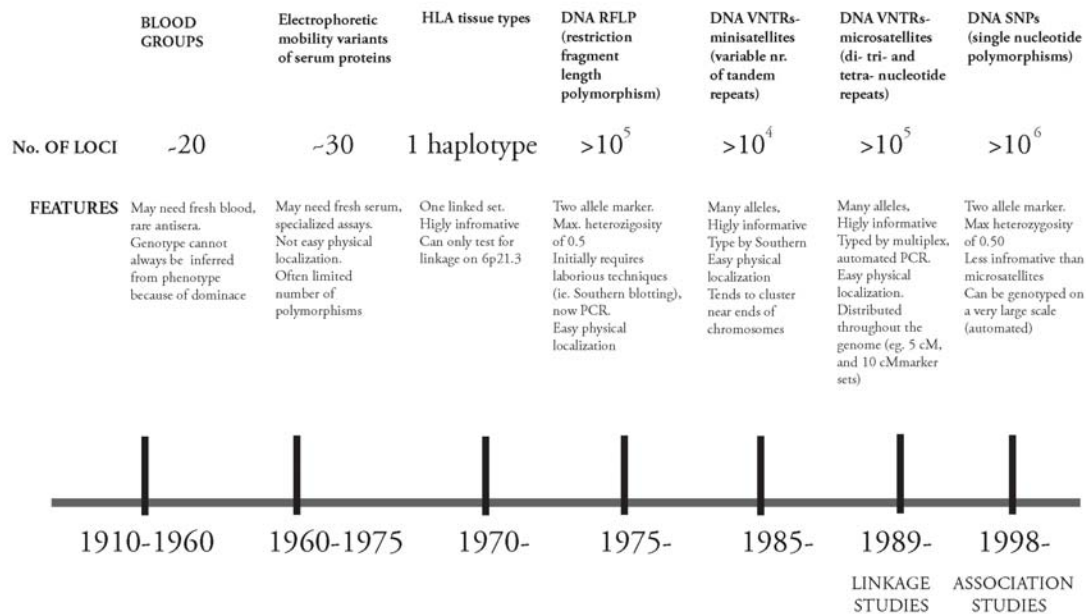
In addition to the relatively small fraction of coding DNA, a skewed gene distribution is observed at both intra- and inter-chromosomal level. High gene density regions (ridges), with relatively rich CG content and short introns have been identified; in contrast to low gene density regions, so-called “genomic deserts” (anti-ridges). Gene deserts constitute approximately 25% of the human genome and have been defined as regions containing non-protein-coding sequences without obvious biological functions (Venter *et al.* 2001). Despite the lack of functionality initially predicted, subsequent studies have shown evidence of rather contrasting patterns: gene desert regions were seen to contain regulatory sequences (enhancers, repressors, insulators) acting at large distances to control the expression of

neighboring genes (Nobrega *et al.* 2003), as well as non-essential coding-gene depleted regions which can be deleted without significant phenotypic effects (Nobrega *et al.* 2004).

Ridges and their counterparts might then represent a higher order structure of the genome. This particular arrangement might be strongly linked to complex gene expression patterns (Ovcharenko *et al.* 2005; Versteeg *et al.* 2003). From the evolutionary perspective, the existence of co-regulated gene clusters is not unique to humans, and has been observed in *Drosophila* (Boutanaev *et al.* 2002), mice (Li *et al.* 2005) and, although to a lower extent, in simpler organisms like yeast (Cohen *et al.* 2000). Ridge-like organizations, however, have not yet been identified in other organisms (Li *et al.* 2005; Velculescu *et al.* 1997).

Genetic variation in the human genome can take many forms, ranging from point mutation changes, simple sequences repeats, and transposable elements, to larger segmental re-arrangements. Point mutations involve substitutions, insertions or losses of a single base pair (bp) at the DNA level. Single nucleotide polymorphisms (SNPs) are the most common type of point mutation. For a variation to be considered a SNP, it must occur in at least 1% of the population. SNPs occur on average every 100 to 300 bp among the entire human genome, representing up to 90% of total human genetic variation (<http://www.hapmap.org>) (The International HapMap Consortium 2005). Changes in a single nucleotide can either lead to no changes at the aminoacid level (synonymous SNP), or changes in the coded aminoacid (non-synonymous SNP). At the beginning of the International Human Genome Project (IHGP), more than 1.4 million SNPs were identified, prioritizing such “SNP discovery” towards non-synonymous coding SNPs. Currently, there are more than 11,500,000 SNPs recorded in the SNP database (dbSNP - build 127), the public repository for DNA variation (<http://www.ncbi.nlm.nih.gov/SNP/index.html>). Notably, only 2% of SNPs are located within coding sequences, while the majority are located in both intronic (56%) and untranslated regions (UTRs) (10%).

Simple repeat polymorphisms can be grouped into two classes in relation to their location in the genome. These are tandem repeats or interspersed repeats, which account for ~3% of the human genome, and can be subsequently classified into three different groups: satellites, minisatellites and microsatellites DNA (see Figure 1.2).



**Figure 1.2** | Techniques to study genetic variation in the human genome

Transposable elements constitute the largest class of interspersed repeats that are found in our genome, encompassing almost 45% of the total sequence, and are divided into two main classes, according to their structure and mode of transposition. Class I elements transpose through a RNA intermediate (retrotransposable elements). This class comprises long and short interspersed nuclear elements (SINEs and LINEs) (> 33%) (Singer 1982) and long terminal repeats (LTR) retroposons (<8%). Class II elements (<3%) transpose directly from DNA to DNA, and include DNA transposons and miniature inverted repeats transposable elements (MITEs). All eukaryotic genomes contain mobile elements, although the proportion and activity of the classes of elements varies widely between genomes (Kidwell & Lisch 2000).

From the evolutionary perspective, due to highly intra-species specificity of repetitive sequences, *Alu*-insertions, the most abundant repeat among SINEs (Mighell *et al.* 1997), have been used as phylogenetic tool to aid the study of human population genetics and primate comparative genomics (Batzer & Deininger 2002). Although the function of retroposons remains contentious, recent research revealed their potential functionality in a gene-regulatory context (Bejerano *et al.* 2006; Han & Boeke 2005; Hellmann-Blumberg *et al.* 1993; Muratani *et al.*

1991; Yang *et al.* 1998). Such evidence proved to be enough to revise the classical conception of “molecular fossils”.

*Copy number variation* (CNV) (also known as large-copy number variations or LCVs), contrary to single base changes, encompasses a wider range of variation in terms of DNA content, ranging from kilobases (Kb) to megabases (Mb) in size (Iafrate *et al.* 2004; Sebat *et al.* 2004). A CNV can be defined as a copy number change involving a DNA fragment that is  $\sim 1$  kilobase (kb) or larger. They can be micro or sub-microscopic, and can include deletions, insertions and duplications.

*Segmental duplications* which constitute  $\sim 5\%$  of the genome, frequently comprise complex blocks of repetitive DNA that can be duplicated in tandem or transposed to new locations in the genome (transchromosomal duplicons).

Such level of complexity has not been adequately considered and this has been the main source of potential misassembling in genomic sequences projects (Cheung *et al.* 2003). Two independent surveys were conducted recently among HapMap populations. The goal of the first study was the quantification of the extent of CNVs in the genome (Redon *et al.* 2006); the second survey aimed to quantify the amount of genetic variation detected in gene expression by both SNPs and CNVs (Stranger *et al.* 2007). The results were striking. CNVs encompassed about 12% of the genome, and estimates of the amount of genetic variation explaining gene expression were about 80% and 18% for SNPs and CNVs, respectively. In addition, signals from these two type of variants showed non-significant overlap (Redon *et al.* 2006; Stranger *et al.* 2007). Interestingly, a strategy was developed lately, which utilizes the large and widely available SNP databases to discover relatively small deletions ( $>8\text{kb}$  in length), usually reflected as Mendelian inconsistencies, null genotypes and Hardy-Weinberg disequilibrium patterns (McCarroll *et al.* 2006). It is worth mentioning, that the aforementioned inconsistencies are indeed quite likely to occur as a result of technical artifacts and genotyping errors. Different technical approaches as well as the use of related individuals, were combined (i.e. FISH, qRT-PCR) to filter out real clustered deletion patterns of artifact inconsistencies (McCarroll *et al.* 2006). As a result, the first CNV database was generated (<http://humanparalogy.gs.washington.edu>) (Conrad *et al.* 2006).

The next step will be to integrate both databases in order to see whether significant correlations exist between LD patterns and CNVs among the genome; since, for example, both single point mutations (SNPs) and CNVs have lately been associated with susceptibility to a wide variety of cognitive disabilities ranging from mild learning disabilities to autistic disorders (Deffenbacher *et al.* 2004; Sebat *et al.* 2007).

### *Searching for genetic variation underlying cognitive ability*

The information gathered via the Human Genome Project (HGP) together with the information available from the International HapMap Project (The International HapMap Consortium 2005), provides us with the unique opportunity to study in a more systematic way, human genome diversity and its relation to complex traits such as cognition.

### *Mental retardation*

So far, successful identification of genes underlying genetic variation in human cognitive ability has been mainly limited to relatively rare mutations with considerably severe cognitive effects, in which mental retardation or milder forms of cognitive disability are part of a heterogeneous phenotype

*Chromosomal rearrangements* are often cytogenetically detectable, and involved segmental duplicons, in which resulting phenotypes are due to dosage imbalance of one or more genes instead of mutation of the gene itself (i.e. Smith-Magenis Syndrome (Greenberg *et al.* 1991), DiGeorge Syndrome, Velo-Cardio Facial Syndrome (Edelmann *et al.* 1999) Down Syndrome involving Robertsonian translocations (Petersen *et al.* 1991)).

*Hypervariation of repetitive sequences* (hypervariable tandem repeats) within a critical region of a gene can cause instability of the region which in turn becomes prone to methylation and subsequent inactivation of gene(s) transcription (i.e. Fragile X Syndrome) (Fu *et al.* 1991; Verkerk *et al.* 1991).

*Single point mutations* can cause either a significant decrease in expression, or production of abnormal mutant proteins. Although they are circumscribed to single amino acid changes, they occur in proteins whose functions are fundamental for orchestrating general neuronal signaling pathways. These point mutations are usually affecting genes located upstream on a signaling cascade and may give rise to syndromic phenotypes with a wide range of cognitive impairments (i.e. Apert Syndrome (Ibrahimi *et al.* 2005), Rett Syndrome (Neul & Zoghbi 2004), Frontotemporal dementia linked to chromosome 17 (FTD-17) (Forman *et al.* 2000)).

### *Phenotypes for cognitive abilities within the normal range:*

#### *Linkage and association studies*

It is now generally assumed that cognitive ability can be considered as a polygenic trait influenced by many genes of moderate to small effect that in turn may interact with each other and with environmental factors (Butcher *et al.* 2006; Plomin & Spinath 2004; Savitz *et al.* 2006). Finding putative candidate genes



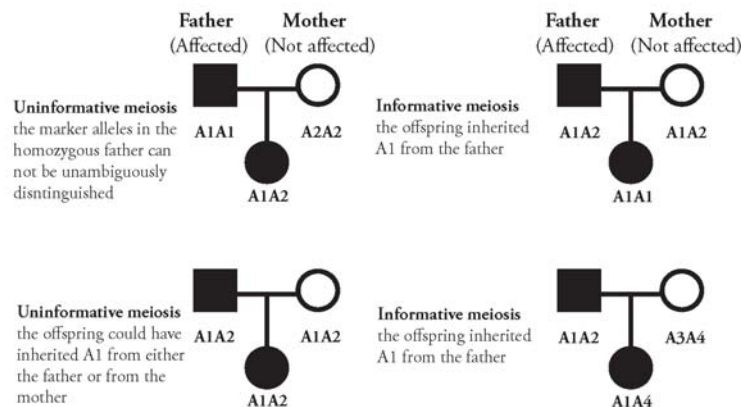
underlying normal cognitive variation, however, proves to be a daunting task, mainly due to its polygenic nature and the need for reliable study designs that account for the detection of such relatively small genetic effects. Two different and somewhat complementary approaches can be used for identifying putative candidate genes underlying variation in cognitive abilities: whole-genome linkage studies (linkage studies) and whole-genome association studies (association studies). Advantages and caveats of each approach will be discussed in the following section as well as recent reports and findings involving both strategies.

### *Whole-genome linkage studies*

Linkage analysis is the method traditionally used to identify monogenic disease genes, but can also be used for polygenic traits. They can be characterized, depending on their mode of inheritance, by observation of their pattern of segregation, as described by Gregor Mendel over 100 years ago (Bateson 1909) (see Figure 1.3). The goal of linkage analysis is to determine whether two loci (one of which is the unknown disease gene and the other being an observed marker gene) tend to co-segregate together more often than they should if they were not physically close together on the same chromosome. When conducting linkage studies, a family or group of families are gathered (variation of the classical linkage analysis exists in which only data from affected sib-pairs is used, often in the context of late onset diseases); phenotypic information is gathered and DNA from all available family members is isolated and genotyped for a determined set of markers – generally microsatellite markers consisting of tri and tetra nucleotides CA repeats evenly spread along the entire set of chromosomes.

In complex human pedigrees an overall calculation of the pedigree likelihood can be obtained under, the alternative assumptions that the loci are linked ( $\theta$ ) or not ( $\theta = 1/2$ ). The ratio of these two likelihoods gives the odds of linkage, and the logarithm of the odds is the lod score (Morton 1955). The threshold for a genome-wide analysis to be considered significant (p-value of 0.05) is given by a lod score of 3.3 or higher (Lander & Schork 1994). Of particular interest has been the use of linkage studies among isolated populations. Studies of such populations, consisting of families originated from a limited pool of individuals, have proved to be advantageous for genetic mapping of Mendelian disease (for review see Heutink & Oostra 2002, and Peltonen 2000). Identification of disease loci tends to be less cumbersome, since they are expected to share relatively larger DNA segments (so-called haplotypes) derived from a common ancestor (founder effect) harbouring the disease loci. After conducting a successful linkage study, the regions under the

significant lod score peaks tend to be relatively large (usually 10-20 Mb in size), harbouring too many potential candidate genes. The next step, is either to perform a finemapping of the linked genomic regions or, if there is enough evidence, just select within this region those genes with unquestionable evidence that might be linked with the aetiology of the disease under study.



**Figure 1.3** Mendelian Traits | During meiosis, crossing-over events occur between homologous chromosomes, resulting in what is referred to as *recombination*. Such recombination processes give rise to an extra source of genetic variation, since the resulting chromosomes will contain portions of the maternal homolog recombined with its paternal counterpart (Moore & Orr-Weaver 1998). Recombination events between two genetic loci (e.g. gene, DNA marker) located on the same chromosome will occur at a rate that is a function of the distance between them. This rate or probability is called the *recombination fraction* ( $q$ ). In other words, loci that are very close to each other tend to be inherited together more often than not. Recombination fractions range from  $q = 0$  for loci right next to each other, to  $q = \frac{1}{2}$  for widely spaced loci – or loci on different chromosomes – (Terwilliger & Ott 1994). The larger the family the more informative meioses one can expect. Recombination events (meiosis) are estimated to occur at least once per pair of homologous chromosomes per generation (Alberts *et al.* 2002). Meiosis is informative for linkage when it can be unambiguously identified as to whether or not a recombination process has taken place, and maternal and paternal loci can be unambiguously assigned to the offspring

Although linkage studies are the method of choice when searching for Mendelian diseases or polygenic traits influenced by rare alleles, linkage analysis is less powerful than association studies for detecting modest to small genetic effects exerted by common loci. This disadvantage resides mainly in the fact that a prohibitive number of families would be required to map the genes of minor to moderate effect that may underlie complex (common) traits (Risch & Merikangas 1996) together with the relatively low resolution of genetic linkage markers (10-20 cM). Alternative methods have been developed in order to overcome the need of large pedigrees when complex (common) traits are investigated. The original sib-pair method was originally developed based on the idea that linkage is supported if sibling pairs with two affected or two unaffected siblings are significantly more alike

in terms of allele sharing at a given marker locus compared with siblings pairs with just one affected member (Penrose 1953). In the early '70 , Haseman and Elston (1972), proposed a method (HE) that has been widely used in linkage for QTLs. Evidence that a given marker is linked to the trait is obtained by regressing the square sib-pair trait difference between phenotypes of siblings on the proportion of maker alleles shared identical-by-descent (IBD) ( $\pi$ ). Variations of the original method have been developed like for example, revisited HE that includes information available in the sib-pair traits sum (Elston *et al.* 2000) or a “weighted HE” method that takes into account the presence of strong residual sib correlations (Forrest 2001; Xu *et al.* 2000).

Nevertheless, when there is allelic heterogeneity (*several* deleterious mutations within a single gene), frequently observed in metabolic disorders with Mendelian inheritance, linkage's power to detect them will always prevail over association's (i.e. Lysosomal storage diseases (Reuser & Drost 2006). Previous epidemiological and linkage studies have provided evidence of genetic contribution to diverse cognitive disabilities (e.g. attention deficit hyperactivity disorder (ADHD), autism, reading disabilities, and dyslexia) (Cardon *et al.* 1994; Galaburda *et al.* 2006; Plomin 2001; The Autism Genome Project Consortium 2007; Wassink *et al.* 2004). The first genome-wide linkage study for intelligence was published recently as a result of a collaborative study using a total 429 families with Australian and Dutch ancestry, (Posthuma *et al.* 2005). Subsequently, four more studies were reported (Buyske *et al.* 2006; Dick *et al.* 2006; Luciano *et al.* 2006; Wainwright *et al.* 2006). The first study reported significant regions of linkage at chromosomes 2 (2q24.1-31.1, LOD score 4.42 for Full-scale IQ) and 6 (6p25.3-22.3, LOD score 3.20 for full-scale IQ) regions previously linked to dyslexia and reading disabilities (Cardon *et al.* 1994; Wassink *et al.* 2004). Several other regions showing suggestive linkage (4p, 7q, 20p, 21p) were also reported (Posthuma *et al.* 2005).

### *Whole-genome association studies*

Due to the expense and labor involved, the extension of candidate-gene studies to a genome-wide approach has not, until now, been feasible, except for worldwide- collaborative studies (e.g. GAIN – Genetic Association Information Network). The next section will be mainly related to candidate-gene association strategies from which study design criteria can also be applied to whole-genome association studies. In candidate-genes association studies, genes are selected based on prior knowledge of biological pathways, which are known to be involved in the expression of the traits, or as a result of significant previous linkage studies.

Contrary to linkage studies, association studies rely on population samples consisting of unrelated individuals. It is important however, to keep in mind that both association and linkage rely on the co-inheritance of adjacent DNA segments. While linkage capitalizes on this by identifying haplotypes that are inherited intact over several generations (few cM in average), association relies on the retention of adjacent DNA variants (alleles) located too close to each other to be split up by recombination events. This type of non-random association of alleles is known as linkage disequilibrium (LD). LD is a function of both allele frequencies and genetic distance and it is generally assumed to operate along DNA regions no longer than 10 kb, although some exceptions have been reported (Buzas et al. 2004).

So far, one of the advantages of association studies over linkage is the fact that they can be conducted among population samples. This advantage is not trivial when for example affected individuals are needed to be collected to identify susceptibility determinants for late-onset diseases. It is, however, quite burdensome in terms of the assumptions made and the way in which outcome results can be interpreted because spurious associations may arise in a population that is a mix of two or more genetically distinct subpopulations. Any trait that is more frequent in one of the subpopulation(s) (e.g. because of assortative mating or cultural differences) will show a statistical association with any allele that has a different frequency across those subpopulation(s) (e.g., as a result of different ancestors or genetic drift). In practice, more than two subpopulations may have combined and it will not be obvious from the combined populations whether or not and in which way the sample is stratified. Such bias due to mix of different subpopulations is known as population stratification or ethnic admixture and may lead to false positive genetic associations. It is likely that, at least two strategies exist to avoid population stratification or at least control it.

The first one involves the genotyping of unlinked markers among population-based samples (Pritchard & Rosenberg 1999). The second one involves a more sophisticated approach, using family-based samples. Family-based designs have the advantage of being robust against population stratification (spurious associations), and also that significant findings using family-based association studies always imply both linkage and association (Abecasis *et al.* 2000). Within this model, the trait means per genotype are compared, and also the residual sib-correlation is modeled as a function of polygenic or environmental factors. The only disadvantage is that a relatively larger number of families are usually needed, since they tend to be less informative (e.g. among twin studies, families consisting

of only MZ will not be informative to the within-family association analysis, unless they are paired with non-twin siblings).

Nevertheless, success on identifying quantitative loci with modest genetic effects has been also achieved, although not without some special study design considerations. For example, the selection of individuals whose phenotypes are in the extremes of the normal distribution curve (e.g. high IQ vs. subjects with normal IQ scores) (Chorney *et al.* 1998). Chorney and co-workers investigated the possible association between the insulin-like growth factor-2 receptor (*IGF2R*) candidate-gene located at 6q26 and cognitive ability. The study tested 37 DNA (microsatellite) markers spanning approximately 110 cM. Although the sample size among high-IQ and control-IQ groups (age of testing between 6-15 years of age) were relatively small ( $N = 50$ ), association results from a marker located on the 3'UTR of the gene evidence, maybe for the first time, (1) the relatively small effect size of associations among QTLs, (2) the presence of (non)coding polymorphisms as causal variants or just close by to the functional variant.

More than a decade ago, an extreme-sample-pooling approach so-called DNA pooling, was firstly documented for homozygosity mapping of candidate regions in autosomal recessive disorders in isolated populations (Carmi *et al.* 1995; Nystuen *et al.* 1996; Scott *et al.* 1996). In line with this, DNA pooling has been utilized as screening device for the identification of QTLs for cognitive abilities. DNA pooling greatly reduces the need for genotyping by pooling DNA from all individuals in a group and genotyping the pooled groups. Unlike DNA chips and other high-throughput approaches, genotyping costs for DNA pooling are independent of sample size (Daniels *et al.* 1998). DNA pooling consists in a three-stage screening design (i.e. original pooling, replication pooling, individual genotyping). The first association studies that carried out DNA pooling involved 66 and 147 (microsatellite) markers for chromosomes 22 (Hill *et al.* 1999), and 4 (Fisher *et al.* 1999), respectively. The study of Hill and co-workers (1999) failed to identify QTLs associated with cognition, mainly due to its lack of power given the reduced sample size ( $N = 50$  for high-IQ cases and normal-IQ controls, respectively) to detect small genetic effects. The Fisher study (1999), on the other hand, successfully confirmed the pooled association (original pooling) at a single marker step (individual genotyping), with further finemapping required given the low resolution of the markers used (spaced about 10 cM).

The first genome-wide association (GWA) using DNA pooling was reported in the early 2000 (Plomin *et al.* 2001). The study implemented the three-step design and used a set of about 1800 simple sequence repeats (SSR) markers among

two independent case-control samples consisting in 200 individuals each ((N= 100 for high-IQ cases and normal-IQ controls, respectively). Although the study failed to replicate findings between samples, it evidenced the need of more markers to be genotyped, especially when association studies instead of relying on direct approaches -non-synonymous polymorphisms (nsSNPs)-; they rely on indirect approaches based on LD between the QTL and the genotyped marker.

In line with the previous studies, Plomin and colleagues (2004) reported association of a functional polymorphism on the *succinate semialdehyde dehydrogenase* (*SSADH*; *ALDH5A1*) gene (c.T538C) and cognition. A significant effect size of 1.5 IQ points was reported. The *ALDH5A1* gene is of particular interest, since it codes for a mitochondrial NAD(+)-dependent succinic semialdehyde dehydrogenase, which represents the last enzyme in the –amino butyric acid (GABA) catabolism and irreversibly oxidizes succinate semialdehyde (SSA) to succinate (Blasi *et al.* 2004). Interestingly, in humans, *ALDH5A1* deficiency results in 4-hydroxybutyric aciduria, an autosomal recessive disorder due to an accumulation of GABA and 4-hydroxybutyric acid in the CNS mainly characterized by developmental delay of motor, mental, and language skills, behavioral problems, and EEG abnormalities (for a review see Gordon 2004 ).

### *Association studies: Direct or Indirect approaches?*

Association studies can be conducted using either the “direct” or the “indirect” approach. The *direct* approach is the methodology of choice when prior knowledge exists in relation to a *known* genetic (functional) polymorphism/s (SNP) (i.e. allelic variants of SNPs, usually on coding regions of a gene, affecting splicing sites or enzymatic activity of the transcript protein). Genotypic variation on these functional polymorphisms can then be compared, as in case-control study designs (Hariri *et al.* 2003; Malhotra *et al.* 2002; Plomin *et al.* 2004; Weinberger *et al.* 2001). In the *indirect* approach, functional variants in a gene are *unknown*, and instead a so-called *tagging* approach can be used. This approach depends on linkage disequilibrium (LD) between a genetic marker (e.g. a SNP) and quantitative loci (QTL) that influences the trait under study. In other words, the *causal* (unknown) variant(s) will not be directly genotyped but instead, *tagged by tag-SNPs*, which in turn are expected to provide information on the actual functional variant when they are in LD with it. In practice, both direct and indirect approaches are frequently combined. As postulated previously, LD is the non-random association of alleles at adjacent loci.

LD can be the result of both molecular and population factors interactions (i.e. recombination fraction variation, genetic drift due to founder effect and/or population admixture) (Ardlie *et al.* 2002). LD patterns within given genetic regions can therefore be used to identify so-called tag-SNPs. These tag-SNPs capture information of multiple surrounding SNPs constituting “haplotype blocks” (Cardon & Abecasis 2003). From a practical perspective, genotyping *tag*-SNPs is generally more efficient and more cost-effective than genotyping all available SNPs within a genomic region.

LD estimates are usually represented by two statistical measurements,  $D'$  and  $r^2$ . Both LD measurements are calculated based on the basic pairwise-disequilibrium coefficient,  $D$ , which reflects the difference between the probability of observing two marker alleles on the same haplotype and observing them independently in a given population (see equation 1) (Lewontin 1964).  $A$  and  $B$  indicate two genetic markers (e.g. an observed marker and a QTL), each with alleles defined as “1” and “2”, and with  $f$  being their frequency.

$$D = f(A_1B_1) - f(A_1)f(B_1) \quad (1)$$

The  $D'$ , or  $|D/D_{\text{MAX}}|$  varies between 0.0 and 1.0. A value of 0.0 means that the observed alleles segregate independently, whereas 1.0 implies that all copies of the *causal* allele occur exclusively with one of the two possible alleles at its counterpart genetic marker. Although values of  $r^2$  also vary from 0.0 to 1.0; they bear an additional feature in relation to population genetics with a more stringent meaning which is that an  $r^2$  value of 1.0 only occurs when the marker allele frequency (MAF) is identical to the *causal* allele. Contrarily, a  $D'$  value = 1.0 reflects the genetic correlation only since the most recent mutation occurs; allele frequencies might vary broadly and also be prone to extensive fluctuations if sample size is small (Weiss & Clark 2002).

### *Criteria for tag-SNP selection: population, allele frequencies and LD*

Surely, a pivotal decision with respect to further interpretation of results is the selection of the population from which *tag*-SNPs will be chosen. One possibility is to re-sequence candidate genes from a representative subgroup of samples to obtain all possible variants, and calculate the LD among them and then subsequently select the most optimal tagging *SNPs*. Another option is to use

information already gathered by The HapMap Project (The International HapMap Consortium 2005) or ENCODE project (The Encode Project Consortium 2004), which is also mirrored by other databases (i.e. Celera Database, Perlegene, SNP Browser). Four different populations with defined historical ancestry have been sequenced and genotyped within The HapMap Project (i.e. Yorubian and Nigerian: YRI; Japanese: JPT; Chinese: CHB; and Americans with ancestry from northern and western Europe: CEU) and so data from a diverse range of populations with different recombination ancestry is available. Ancient African populations (and those with very recent origins in Africa) might require a larger amount of *tag-SNPs* to explain the same amount of variation as compared to a relatively younger population, for example that of western Europe, since in the former more recombination events are likely to have occurred than in the latter (Gabriel *et al.* 2002). As a consequence of this, populations with African ancestry will present a characteristic LD pattern consisting of shorter haplotype blocks.

Several studies have been conducted to validate to what extent power to detect an association is compromised when *tag-SNPs* are drawn from data from any HapMap population (genetically characterized sample) and subsequently used in another sample (phenotypically characterized). The results concluded that power is not significantly compromised and that HapMap populations can be broadly used to select *tag-SNPs* for a broad range of populations (de Bakker *et al.* 2006; Service *et al.* 2007). The rationale behind selecting *tag-SNP* before conducting a genetic association study is straightforward, and most algorithms use  $r^2$  as a metric by which to choose *tag-SNPs* (Carlson *et al.* 2004). Selecting *tagging* polymorphisms in “perfect” LD ( $r^2 = 1.0$ ) will give an extremely low number of *tag-SNPs* to test (since this scenario is not so plausible to occur among human populations due to recombination events). Setting too low values of  $r^2$  ( $< 0.50$ ) will contrarily, result in redundant genotyping. A balanced tradeoff therefore, between the amount of genotyping and the information that can be obtained has to be accomplished. An  $r^2$  value between 0.80 - 0.85 is within a common range for selecting *tag-SNPs* so that all variation is either typed directly as a tag, or in LD with a *tag-SNP* at a level of  $r^2 = 0.85$  or greater (Carlson *et al.* 2004). Also minor allele frequencies (MAFs), are selected to be above  $>0.10$ , not only in order to reach enough power to detect common variants but also to be able to observe all three possible bi-allelic combinations.

At this point the question may arise of *why rare genetic variants might be so “intentionally” left out of our study design*. The answer, although purely theoretical, is simple to grasp, and it mainly relies on the so-called Common Disease - Common



Variant (CD/CV) hypothesis (Reich & Lander 2001). In opposition to what is expected of Mendelian traits where rare loci are the “causal suspects”, complex (common) traits are by definition influenced by several, maybe even hundreds, of common loci, each one exerting a small but fundamental genetic effect. Furthermore, the single quantitative influence of each QTL might be neither indispensable, nor sufficient to triggers what we observe, and measure, as an outcome cognitive phenotype. From an adaptive evolutionary perspective, human-specific change in the genome might have functional implications; such adaptive loci must have been under strong selection and as a result, greatly prevailed over loci with rather neutral effects. Such a strong assumption relies on what is already predicted about modern human population history. It has been estimated that the human population experienced a bottleneck about 70 thousand years ago and its population number was severely reduce up to several thousand individuals (~10,000). Such drastic reduction swept away the majority of the allele diversity acquired until then (Ambrose 1998) - and as a consequence, chances of natural selection were greatly reduced. Subsequently, the prevailing population started to expand again, and this augmentation is believed to have occurred in an exponential manner. The result of this fast expansion, originated from a relatively small number of common ancestors, supports the common disease/ common variant (CD/CV) hypothesis (Reich & Lander 2001) postulating that common variants were mainly “survivors” after the first bottleneck, and expanded boundlessly afterwards; with rare genetic variants on the other hand, having not enough time and opportunity to arise.

Although appealing in many ways, there is still not enough empirical data to completely support such a hypothesis. The replication of candidate-gene association studies amongst different populations has been scarce, mainly due to the small sample size used to detect small genetic effects, as well as genetic *heterogeneity*. Genetic heterogeneity can be defined as the production of identical or similar phenotypes by different genetic mechanisms, obscuring the elucidation of effects exerted by *causal* genetic variants. Detractors of the CD/CV hypothesis postulate environmental factors to have a more important role in complex traits variation, than genetic variants. They suggest that rare alleles might indeed have been selected as a result of environmental pressure during population expansion, whereas common alleles tended to be lost or undergo neutral fixation in response to natural selection (Pritchard 2001). Family-based candidate-gene association studies are particularly biased against finding rare genetic variants (MAFs < 0.1) since their power to detect association relies on relatively high minor allele frequencies (MAFs

$\geq 0.10$ ) and sample sizes are generally not so large as to obtain a significant number of rare heterozygous.

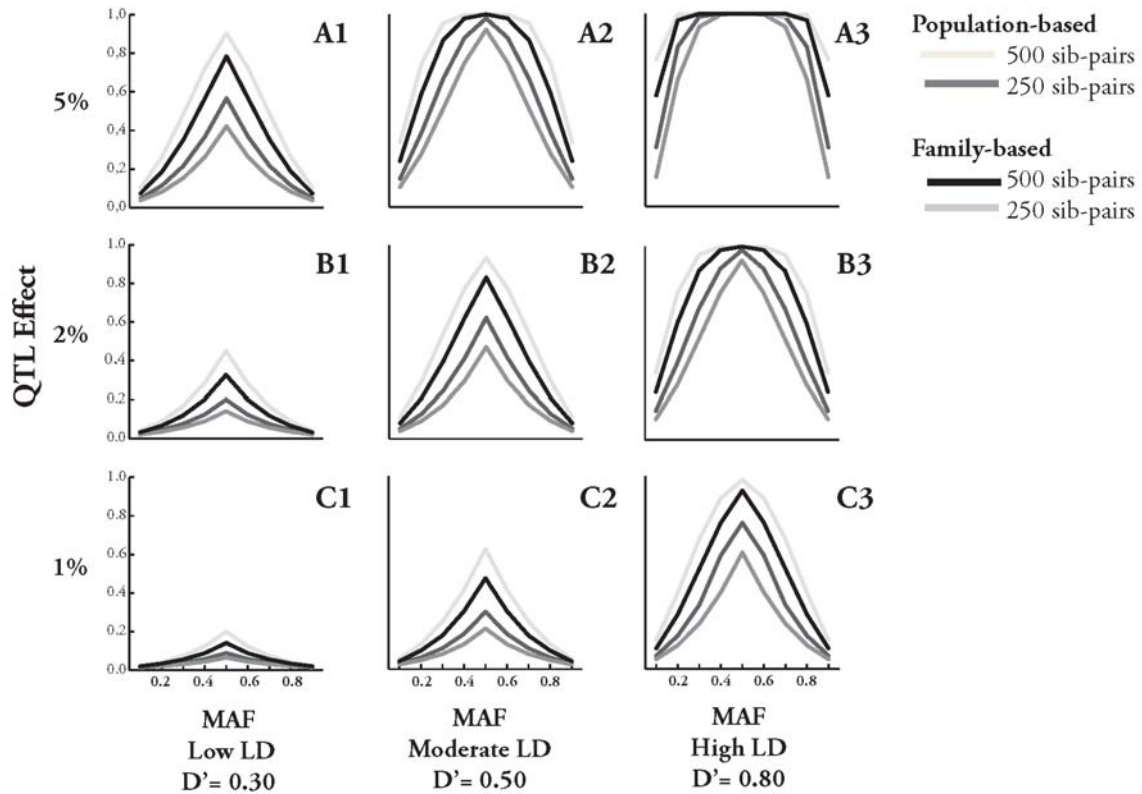
### *Multiple testing corrections*

When evaluating potentially interesting results from a number of statistical tests, it is necessary to determine how often a ‘significant’ p-value would arise by chance if the study were repeated under the hypothesis of no genetic association. Bonferroni Correction has proved to be too conservative especially when non-independent phenotypes are used in the context of association studies. Permutation tests can then be implemented to calculate empirical p-values based on the observed data. To correct for multiple testing a Monte-Carlo permutation framework can be used to calculate empirical p-values (Abecasis *et al.* 2000). Empirical p-values provide an indication of the empirical levels of type I and type II errors, conditional on the observed data. The trait values are shuffled  $N$  times (typically  $N = 1000$ ) among the  $n$  individuals to create permuted data sets that have only random genotype-phenotype associations. The permuted data sets are a representative sample from an appropriate ( $H_0^1$ ). An empirical threshold value for detecting a major QTL effect is obtained by computing the  $(1-\alpha)$  percentile from the  $N$  permuted data sets of the maximum test statistic value over the genetic map.

### *Power to detect QTLs based on linkage disequilibrium*

In terms of sample size, for a given  $r^2$  between the marker and *causal* variant, the power to detect small QTL effects will increase with increase of LD. Although  $r^2$  is preferred as a measure of LD, for power calculation purposes  $D'$  in combination with a relatively large sample size will be selected (500 sib-pairs) (see Figure 2). As observed in Figure 1.4, at low  $D'$  values only relatively moderate (5%) QTL effects will be picked up with power given by the population-based approach, although prone to false positives due to population stratification effects (*A1*). The next scenario, and maybe more realistic, is the presence of moderate LD ( $D' = 0.50$ ) between the *tag* and the *causal* allele in which finding moderate effects (*A2*) might not be an obstacle even after controlling for stratification effects within a family-based frame. Detection of relatively small QTL effects is, however, more cumbersome at a moderate  $D'$  value (*B2*, *C2*). Realistic chances of detecting small effects may arise at relatively high  $D'$  values (*B3*) whereas 1% QTL effects might still remain undetected (*C2*). A possible solution to circumvent such lack of power is the use of extreme samples (Zhang & Sun 2005). Association studies using samples with extreme phenotypic values (e.g. high vs. low IQ) have been reported

(Chorney *et al.* 1998), however, posterior validation of this type of studies is afterwards, even more difficult to obtain (Hill *et al.* 2002).



**Figure 1.4** | Power to detect small QTL effects in the context of variable linkage disequilibrium (LD) patterns and minor allele frequencies (MAFs). Five hundred sibship pairs ( $N = 1000$ ) were chosen to mitigate the bias of  $D'$  owing to small sample size. Power estimates ( $p$ -values of 0.01) to detect small QTL effects (1%, 2%, 5%) with a trait heritability ( $h^2$ ) of 0.60, were calculated at three different LD values ( $D' = 0.30, 0.50$ , and  $0.80$ ) for a broad range of minor allele frequencies (MAFs between 0.10 and 0.90). Family and Population-based power estimated for 500 sib-pairs are plotted, as well as Family and Population-based power estimated for 250 sib-pairs (approximately the sample size used in the present study)

### *Candidate gene selection criteria*

The delineation and subsequent selection of candidate genes to conduct any association study for complex (common) traits is by far one of the most cumbersome tasks. Candidate gene selection is generally based on animal models since although they can be considered “simpler” organisms than humans, they are fully genetically characterized and genetic manipulation for research purposes is ethically possible. As a result, pharmacological and genetic manipulation can be achieved with relatively high reliability (e.g. mouse models for diabetes mellitus type 2, rare Mendelian forms of obesity). Extrapolation of such models to study cognitive ability traits in a more systematic manner, however, already bears a

burden difficult to overcome: how is cognitive (dis)ability defined in animal models and how reliable can this be used to extrapolate to human cognitive traits? This was partially resolved with the development of animal cognitive behavioral tasks and the use of fully genetically characterized inbred mouse lines. Subsequently, the next burden to overcome, and this time related to the organ under study, is the enormous amount of genes expressed in the nervous system (50% of the total genomic pool) in combination with the high degree of alternative splicing (AS) present in the brain. In addition, results from pharmacological challenged animal models (i.e. agonists and antagonists agents used among knock-out mice for a particular receptor. gene), might be inconclusive, due for example, to the lack of specificity in relation to receptor subunits isoforms.

Human linkage studies on the other hand, may give good hints, but at the time of starting the present study (2003), linkage studies were conducted only for cognitive disabilities (dyslexia, autism) and common variants were not likely to play a role, at least among the population selected for conducting such studies. Together, two main criteria were taken into account when selecting candidate genes for cognition in the present thesis. First, the existence of strong evidence for putative candidate genes reported via animal and human studies, which in turn suggests a role in normal cognitive variation. Secondly, a good SNP coverage availability at the HapMap Project database. Based on these two criteria, we selected a list of putative candidate genes from which an initial minimum set of tag-SNPs explaining a large proportion of genetic variation was available.

At the start of the present study in 2003, nine putative candidate genes, based on their suggested role in normal cognitive variation were genotyped for later testing for genetic association with cognition in a family based design (see Table 1.1).

**Table 1.1** Candidate genes list

| Gene                                | Name  | Chromosome | Location    |
|-------------------------------------|---|------------|-------------|
| <sup>(1)</sup> <i>CHRM2</i>         | cholinergic receptor, muscarinic 2          | 7          | 7q31-q35    |
| <sup>(3)</sup> <i>COMT</i>          | catechol-O-methyltransferase                | 22         | 22q11.21    |
| <sup>(A)</sup> <i>DBH</i>           | dopamine beta-hydroxylase                   | 9          | 9q34        |
| <sup>(3,A)</sup> <i>DRD2</i>        | dopamine receptor D2                        | 11         | 11q23       |
| <sup>(A)</sup> <i>DRD3</i>          | dopamine receptor D3                        | 3          | 3q13.3      |
| <sup>(A)</sup> <i>HTR2A</i>         | 5-hydroxytryptamine (serotonin) receptor 2A | 13         | 13q14-q21   |
| <sup>(A)</sup> <i>SERT (SLC6A4)</i> | solute carrier family 6, member 4           | 17         | 17q11-q12   |
| <sup>(2)</sup> <i>SNAP25</i>        | synaptosomal-associated protein, 25kDa      | 20         | 20p12-p11.2 |
| <sup>(A)</sup> <i>TH</i>            | tyrosine hydroxylase                        | 11         | 11p15.5     |

Note: <sup>(1)</sup> Chapters 2 and 3, <sup>(2)</sup> Chapters 4 and 5, <sup>(3)</sup> Chapter 7, and <sup>(All)</sup> Appendix II

These putative candidate-genes can be divided in several domains:

*Cholinergic modulation.* The cholinergic muscarinic receptor family ( $M_1$ - $M_5$ ) belongs to the superfamily of G-protein-coupled receptors. These receptors activate a multitude of signaling pathways important for modulating neuronal excitability, synaptic plasticity and feedback regulation of acetylcholine (ACh) release and cognitive processes, including learning and memory (Hulme 1990; Volpicelli & Levey 2004; Wess 1996).  $M_2$  receptors are selectively coupled to G-proteins of the  $G_i$ / $G_o$  family, which mediate the inhibition of voltage-sensitive  $Ca^{+2}$  channels. Furthermore, the  $M_2$  receptor subtype is likely to have an additional role in cholinergic modulation of excitatory and inhibitory hippocampal circuits acting as autoreceptor (Akam *et al.* 2001; Kitaichi *et al.* 1999b; Kitaichi *et al.* 1999a; Rouse *et al.* 2000; Shapiro *et al.* 1999; Zhang *et al.* 2002), inhibiting ACh release from cholinergic terminals. Comings *et al.* (2003) (Comings *et al.* 2003) reported an association between the *CHRM2* gene and IQ with a SNP in the 3'UTR of the gene. Transcription of the *CHRM2* gene is complex, Krejci *et al.*, (2004)(Krejci *et al.* 2004) determined that the 5'UTR of *CHRM2* consists of five (non)coding regions and at least three different promoters whose different combinations give rise to eight splice variants.

*Dopaminergic functioning.* The catechol-O-methyltransferase (*COMT*) gene has been one of the most extensively studied candidate genes in relation to individual differences in cognitive abilities (i.e. working memory performance). The *COMT* gene is located on chromosome 22q11, and contains six exons (Grossman *et al.* 1992). It is involved in enzymatic activity that degrades DA, norepinephrine, and epinephrine (Axelrod 1957). A functional polymorphism at exon 4 (rs4680) has been extensively studied. The human *COMT* gene contains a common functional single nucleotide polymorphism (SNP) (rs4680) that substitutes a Valine (Val) for a Metionine (Met) residue at codon 158 in the membrane-bound enzyme isoform (MB-*COMT*), codon 108 on its soluble counterpart (S-*COMT*) (Lotta *et al.* 1995). Decrease in enzyme activity present in *Met/Met* individuals, compared to individuals homozygous for the *Val* allele, leads to a relatively higher DA baseline, whereas *Met/Val* heterozygous displays an intermediate enzyme activity. Decreased *COMT* activity might be beneficial from a functional perspective.

Dopaminergic receptor D3 subtype (*DRD3*) is a member of the “D<sub>2</sub>-like” receptors subfamily (*D<sub>2</sub>*, *D<sub>3</sub>*, *D<sub>4</sub>*) (Sibley & Monsma 1992). In animal studies, using in situ hybridization techniques, it has been shown the *DRD3* gene to be predominantly expressed in several limbic areas (Bouthenet *et al.* 1991), regions that are strongly suggested to be involved in regulation of motivation and emotions (Sokoloff *et al.*

1990). Recent studies, using D2/D3-like agonists, showed enhancement in cognitive ability in healthy old adults (Peretti *et al.* 2004), while mutant mice lacking the D3 receptor subtype display deficits in spatial working memory tasks (Glickstein *et al.* 2002).

The *DBH* gene maps on chromosome 9q34. Dopamine  $\beta$ -Hydroxylase, catalyses the conversion of dopamine (DA) to norepinephrine (NE) (Lagercrantz 1976). Interestingly, molecular studies evidence a strong linkage to ABO as well as a direct association between *DBH* locus and plasma *DBH* activity (Perry *et al.* 1991; Wilson *et al.* 1988) making it a potential quantitative trait locus (QTL) to be studied in relation to cognitive ability and in a more narrow sense, executive attention (Faraone & Khan 2006).

The tyrosine hydroxylase (*TH*) enzyme is the rate-limiting step in the synthesis of catecholamines, converting L-tyrosine to L-3,4-dihydroxyphenilalanine (Nagatsu *et al.* 1964). The *TH* gene is located on chromosome 11p15.5. Mutant mice lacking the *TH* gene are not viable and die at the perinatal stage (Kobayashi *et al.* 1995) while their heterozygous counterparts evidenced a decline *TH* activity as a consequence of a gene dosage effect, and display moderate learning and memory deficits (Kobayashi *et al.* 2000).

*Serotonergic modulation.* The human serotonin transporter (*hSERT*) gene is located on chromosome 17q11.1-17q12. Diverse polymorphisms on the *hSERT* gene have been extensively studied in relation to mood disorders and depression (Caspi *et al.* 2003; Ogilvie & Harmar 1997), with contradictory results (Hoehe *et al.* 1998; Kunugi *et al.* 1996; Ohara *et al.* 1998). Serotonin has been shown to play a fundamental role in cognitive functioning (van Kesteren & Spencer 2003).

The *5-Hydroxytryptophan Receptor 2A* (*5-HTR<sub>2A</sub>*) belong to an extense pool of receptors that modulate serotonergic activity. They have been studied in relation to normal cognitive (De Quervain *et al.* 2003) and psychiatric manifestations (Anguelova *et al.* 2003). It has been suggested that supranormal levels of serotonin may be deleterious for spatial working memory, a fundamental component of cognitive performance (Luciana *et al.* 1998). Furthermore, pharmacological studies performed in primates showed the beneficial role of *5-HT<sub>2A</sub>* in cognition processes, particularly in WM tasks (Williams *et al.* 2002). Recent studies, involving a functional polymorphism at the *5-HTR<sub>2A</sub>* gene receptor, showed poorer performance in cognitive tasks in subjects homozygous for the rare allele (De Quervain *et al.* 2003).

*Synaptic plasticity.* Synaptosomal associated protein of 25 kD (*SNAP-25*) is a presynaptic terminal protein encoded by the *SNAP-25* gene which has been

implicated, together with synaptobrevin and syntaxin, in complex formation and synaptic vesicles docking at the presynaptic membrane (Horikawa *et al.* 1993; Seagar & Takahashi 1998), as well as neurite outgrowth (Bark & Wilson 1994). The *SNAP-25* gene located on 20p12-20p11.2, is absent in the *Coloboma* mutation (*Cm*), a chromosomal deletion on the homologous gene on chromosome 2 (78.2 cM) in mice. Homozygous (*Cm/Cm*) mice die early in embryogenesis; while hemizygous (*Cm/+*) mice, exhibit a wide spectrum of phenotypic abnormalities, including attention deficit related symptoms (Snell & Bunker 1967; Heyser *et al.* 1995; Hess *et al.*, 1996).

## METHODOLOGY

### *Sample description*

For this thesis, data from two different age cohorts were available: a young cohort and an adult cohort. Both cohorts consist of twins and their siblings and were recruited from the Netherlands Twin Registry (Boomsma *et al.* 2002; Boomsma *et al.* 2006). The young cohort consisted of 177 twin pairs born between 1990 and 1992, and 55 siblings (Polderman *et al.* 2006a; Polderman *et al.* 2006b), of which 371 subjects were available for genotyping. The genotyped twins were 12.4 (SD = 0.9) years of age and the siblings were between 8 and 15 years old at the time of testing. There were 35 monozygotic male twin pairs (MZM), 28 dizygotic male twin pairs (DZM), 48 monozygotic female twin pairs (MZF), 23 dizygotic female twin pairs (DZF), 26 dizygotic opposite-sex twin pairs (DOS), 24 male siblings and 24 female siblings, and 3 subjects from incomplete twin pairs (1 male, 2 females). Participation in this study included a voluntary agreement to provide buccal swabs for DNA extraction. The adult cohort encompassed a total of 793 family members from 317 extended twin families that participated in the adult cognition study (Posthuma *et al.* 2001c). Participation in this study did not automatically include DNA collection, however, part of the sample (391 subjects from 195 families) returned to the lab to provide blood (70%) or buccal swabs (30%) for DNA extraction. Mean age was 36.2 years (SD = 12.6). There were 25 monozygotic male twin pairs (MZM), 15 dizygotic male twin pairs (DZM), 1 DZM triplet, 20 monozygotic female twin pairs (MZF), 28 dizygotic female twin pairs (DZF) and 23 dizygotic opposite-sex twin pairs (DOS), 29 female siblings and 28 male siblings, and 109 subjects from incomplete twin pairs (41 males, 68 females) (see Table 1.2).

**Table 1.2** Means and standard deviations of Performance IQ (PIQ), Verbal IQ (VIQ) and Full Scale IQ (FSIQ) (corrected for age and sex effects) in the young and adult cohorts

|              | Young Cohort        |                  | Adult Cohort        |                  |
|--------------|---------------------|------------------|---------------------|------------------|
|              | <i>Total sample</i> | <i>Genotyped</i> | <i>Total sample</i> | <i>Genotyped</i> |
| N            | 407                 | 371              | 793                 | 391              |
| Gender (M/F) | 191/216             | 176/195          | 348/445             | 175/216          |
| Age (SD)     | 12.37 (0.93)        | 12.37 (0.92)     | 37.60 (13.00)       | 36.25 (12.64)    |
| PIQ (SD)     | 94.57 (18.93)       | 94.85 (19.14)    | 104.49 (12.34)      | 104.30 (11.64)   |
| VIQ(SD)      | 102.56 (12.74)      | 102.64 (12.92)   | 103.69 (12.26)      | 104.23 (12.15)   |
| FSIQ (SD)    | 98.65 (15.06)       | 98.84 (15.24)    | 103.56 (11.49)      | 103.81 (11.16)   |

### *Cognitive phenotypes*

The WISC-R (1986) (Wechsler 1986) was used to assess psychometric IQ in the young cohort. Six subtests were used; similarities, arithmetic, vocabulary, digit span, object assembly and block design. Psychometric IQ within the adult cohort was assessed with the Dutch version of the WAIS-III (1997) (Wechsler 1997). Eleven subtests were used, Information, Similarities, Arithmetic, Vocabulary, Digit–symbol pairing, Digit-symbol coding, Digit-symbol free recall, Picture completion, Block design, Matrix reasoning and Letter-number sequencing. Heritability estimates for Full Scale IQ (FSIQ), Verbal IQ (VIQ) and Performance IQ (PIQ) were calculated for both young and adult cohorts.

Heritabilities of FSIQ, VIQ, and PIQ, and were 0.80 (95% CI 0.72-0.85), 0.70 (95% CI 0.59-0.78), and 0.73 (95% CI 0.63-0.80), respectively in the young cohort. Using the complete adult cohort the respective heritabilities for FSIQ, VIQ, and PIQ, were 0.78 (95% CI 0.72 – 0.83), 0.78 (95% CI 0.72-0.83), and 0.71 (95% CI 0.62-0.77). These heritability estimates are comparable to those reported previously for these age-cohorts in the Dutch population (Bartels *et al.* 2002; Posthuma *et al.* 2001b).

### *Zygosity*

For all same-sex twin pairs, zygosity was assessed using 11 polymorphic microsatellite markers (Heterozygosity > 0.80). Genotyping was performed blind to familial status and phenotypic data (see tables 1.3a and 1.3b).



**Table 1.3a.** Zygosity status among individuals for *young* cohort

|              |                  | Number of additional siblings |    | Total sample |          | Genotyped sample |          |
|--------------|------------------|-------------------------------|----|--------------|----------|------------------|----------|
|              |                  | 0                             | 1  | Families     | Subjects | Families         | Subjects |
| MZM          | <i>twin pair</i> | 28                            | 11 | 39           | 89       | 35               | 80       |
| MZF          | <i>twin pair</i> | 34                            | 22 | 56           | 134      | 28               | 65       |
| DZM          | <i>twin pair</i> | 19                            | 11 | 30           | 71       | 48               | 116      |
| DZF          | <i>twin pair</i> | 22                            | 3  | 25           | 53       | 23               | 47       |
| DOS          | <i>twin pair</i> | 19                            | 8  | 27           | 62       | 26               | 60       |
| Single twin  |                  |                               |    |              |          | 3                | 3        |
| Total cohort |                  |                               |    | 177          | 409      | 163              | 371      |

**Table 1.3b.** Zygosity status among individuals for *adult* cohort

|              |                    | Number of additional siblings |    |   |   |   |   |   | Total sample |          | Genotyped sample |          |
|--------------|--------------------|-------------------------------|----|---|---|---|---|---|--------------|----------|------------------|----------|
|              |                    | 0                             | 1  | 2 | 3 | 4 | 5 | 6 | Families     | Subjects | Families         | Subjects |
| MZM          | <i>twin pair</i>   | 36                            | 17 | 4 | 1 |   |   | 1 | 59           | 152      | 25               | 66       |
|              | <i>single twin</i> | 1                             | 2  |   |   | 1 |   |   | 4            | 10       | 15               | 17       |
| MZF          | <i>twin pair</i>   | 42                            | 24 | 3 | 2 |   |   |   | 71           | 179      | 20               | 46       |
|              | <i>single twin</i> |                               | 5  |   |   |   |   |   | 5            | 10       | 15               | 16       |
| DZM          | <i>twin pair</i>   | 19                            | 10 | 4 |   |   |   |   | 33           | 85       | 15               | 40       |
|              | <i>single twin</i> | 3                             | 3  | 2 |   |   |   |   | 8            | 15       | 14               | 17       |
| DZF          | <i>twin pair</i>   | 32                            | 29 | 4 |   |   |   |   | 65           | 167      | 28               | 71       |
|              | <i>single twin</i> | 1                             | 2  | 1 |   |   |   |   | 4            | 8        | 28               | 45       |
| DOS          | <i>twin pair</i>   | 23                            | 21 | 4 | 1 | 1 |   |   | 50           | 137      | 23               | 59       |
|              | <i>single twin</i> | 6                             | 4  | 1 |   | 1 |   |   | 12           | 22       | -                | -        |
| no twins     |                    |                               | 4  | 2 |   |   |   |   | 6            | 8        | 12               | 14       |
| Total cohort |                    |                               |    |   |   |   |   |   | 317          | 793      | 195              | 391      |

*Note:* for Tables 2.2a and 2.2b: MZM and MZF= monozygotic male and female twins, respectively. DZM and DZF= dizygotic male and female twins, respectively, and DOS= dizygotic opposite sex twins

## PRACTICAL WORKFLOW

### *DNA collection and isolation*

Genomic DNA in children was obtained from buccal swabs, while DNA in adults was collected from blood samples (70 %) or from buccal swabs (30 %). The DNA isolation from buccal swabs was performed using a chloroform/isopropanol extraction (Meulenbelt *et al.* 1995). Genomic DNA was extracted from blood samples using the salting out protocol (Miller *et al.* 1988) (see Appendix II).

## *Genotyping*

Nine putative candidate genes previously described, were selected to be included in the initial stage of genotyping in 2003, using a combination of single and multiple SNP genotyping techniques. Two main SNP genotyping platforms were used in this study, SNPlex<sup>TM</sup> and TaqMan<sup>®</sup> assays. Assays for the SNPlex<sup>TM</sup> Genotyping System are designed by an automated high-throughput pipeline which combines SNP specific assays into compatible multiplex pools. Although the design pipeline ensures a successful design rate (>93%) (e.g. those SNPs predicted to have a low likelihood of providing valid data are rejected by the assay design pipeline based on a given population, as well as a successful performance under multiplexing conditions), these nominal design rate may slightly decrease as a result of population's characteristics under study and the fact that each unique multiplex design has been only validated *in silico*.

TaqMan<sup>®</sup> assays, on the other hand are in a vast majority available as validated SNPs (i.e., SNPs from which allele frequency information have been determined in samples of African, western and northern European, Chinese and Japanese ancestry) which in turn ensure the successful genotyping of such validated SNPs. Both SNPlex<sup>TM</sup> and Taqman<sup>®</sup> assays were conducted following the manufacturer's recommendations (Applied Biosystems, Foster city, CA, USA). Blind controls of duplicate samples as well as MZ controls, showed an average genotyping error rate of 3.6%.

## *Statistical Tests*

Genetic association tests were conducted using the program QTDT which implements the orthogonal model proposed by Abecasis *et al.*, 2000 (see also Fulker *et al.*, 1999; Posthuma *et al.*, 2004). This model allows to decompose the genotypic effect into orthogonal between- ( $\beta_b$ ) and within- ( $\beta_w$ ) family components, and also models the residual sib-correlation as a function of polygenic or environmental factors. MZ twins can be included and are modelled as such, by adding zygosity status to the datafile. They are not informative to the within family component (unless they are paired with non-twin siblings), but are informative for the between family component. The between-family association component is sensitive to population admixture, whereas the within-family component is significant only in the presence of LD due to close linkage. Spurious associations may arise in a population that is a mix of two or more genetically distinct subpopulations. Any trait that is more frequent in one of the subpopulation(s) (e.g., because of assortative mating or cultural differences) will show a statistical association with any allele that has a different frequency across those subpopulation(s) (e.g., as a result of different ancestors or genetic drift).

In practice, more than two subpopulations may have combined and it will not be obvious from the combined populations whether or not the sample is stratified and in what way. If population stratification acts to create a false association, the test for association using the within family component is still valid, and provides a valid test of association. More importantly, if population stratification acts to hide a genuine association, the test for association using the within family component has more power to detect this association than a population-based association test. In order to correct for multiple testing a Monte-Carlo permutation framework was used to calculate empirical p-values (Abecasis *et al.* 2000). As we tested multiple SNPs, a significance level of 0.01 was kept.

### *Aims and outline of this thesis*

The general aim of this thesis is to find genetic variants that underlie individual differences in intelligence, using a family-based candidate gene indirect association approach. Instead of merely reporting statistical associations, a second aim of this thesis is to further explore initially significant associations, narrowing genetic areas and unraveling the functional variant that explains the statistical association.

Initially a list of nine putative candidate genes was genotyped. Four of the genes on this list are included in more detail in the present thesis: the *CHRM2* gene, the *SNAP-25* gene, the *DRD2* gene and the *COMT* gene. **Chapter 2** and **3** present results of genetic family-based association study involving the *CHRM2* gene. Subsequently, **chapters 4** and **5** present results of genetic family based association study involving the *SNAP-25* gene. **Chapter 6** reports association results between coding functional variants on the *ADRB2* receptor gene and cognitive phenotypes. **Chapter 7** explores the potential gene-gene interaction between two other putative candidate genes included in this study: *COMT* and *DRD2* genes. Finally, **chapter 8** concludes this work by a discussion of the findings of this thesis in a broader perspective, as well as future strategies to investigate the molecular mechanisms underlying the complex genetic and phenotypic variation in cognitive ability.

**Box 1.1** | regulatory processes during eukariotic gene expression regulation**Primary  
regulatory  
control of  
gene  
expression**

**Transcription** DNA is copied into primary RNA transcript consisting of intron-exon segments

**Post-transcriptional events**

*Alternative RNA splicing (AS)* differential exonic transcription of a single gene into different tissues

*Capping* Addition of a modified Guanine (G) residue

*Poliadenylation* Addition of Adenosine (A) residues. It is believe that the primary role of polyadenilation is protecting the nascense mRNA from degradation

*RNA editing* Similar to alternative splicing, it can occur simultaneously or posterior to AS

*RNA interference (RNAi)* Short RNA molecules (19-25 bp) which interact with specific mRNAs to inhibit their transcription

*RNA splicing* intronic sequences are spliced-out in a pre-mRNA molecule giving rise to a mature messenger RNA (mRNA)

*RNA stability* Related to specific RNA sequences at the 3'UTR which determine the rate of turnover of mature mRNA

**Tissue-  
specific  
regulatory  
processes****Translation**

*Promoters* Combination of short sequence elements to which RNA polymerase binds in order to initiate transcription of a gene

*Alternative Promoters* usage of differential promoters in relation to specific tissue requeriements

*Repressors* Any mechanisms that inhibits DNA transcription (e.g. Methylation, Acetylation)

*Enhancers* Short regulatory nucleotide sequences located distant from the promoter with capacity to increase promoter's activity

*Insulators* DNA sequence elements capable of prevent interactions between adjacent chromatine (functional) domains (i.e. gene clusters)

**Post-translational** Histones are essential structures of the chromatine structure. Alteration of histones can make the difference between potentially active and inactive chromatin.

*Acetylation* Adding of *Acetyl* group to specific histone's aminoacid (Serine). Associated with inactivation of DNA trasncription

*Phosphorilation* Adding of *Phosphate* group to specific histone's aminoacid (Serine). Associated with activation of DNA trasncription

*Methylation* Adding of *Methyl* group to specific histone's aminoacid (Lysine). Associated with inactivation of DNA transcription

## Electronic Databases (db)

|                            |   |
|----------------------------|---|
| dbEST:                     | <a href="http://www.ncbi.nlm.nih.gov/dbEST/">http://www.ncbi.nlm.nih.gov/dbEST/</a>   |
| dbSNP:                     | <a href="http://www.hapmap.org/thehapmap.html">http://www.hapmap.org/thehapmap.html</a>   |
| ENCODE Project:            | <a href="http://www.genome.gov/10005107">http://www.genome.gov/10005107</a>   |
| dbSegmental Duplications:  | <a href="http://humanparalogy.gs.washington.edu">http://humanparalogy.gs.washington.edu</a>   |
| SNP Browser Software:      | <a href="http://www.appliedbiosystems.com/support/software/snpexpl/">http://www.appliedbiosystems.com/support/software/snpexpl/</a> |
| The Netherlands Brain Bank | <a href="http://www.brainbank.nl/">http://www.brainbank.nl/</a>   |

## References

- Abecasis, G.R., Cardon, L.R. & Cookson, W.O. (2000) A general test of association for quantitative traits in nuclear families. *Am J Hum Genet* 66 279-292.
- Adams, M.D. (1991) Complementary DNA sequencing: Expressed sequence tags and human genome project. *Science* 252 1651.
- Akam, E., Challiss, R. & Nahorski, S. (2001) G(q/11) and G(i/o) activation profiles in CHO cells expressing human muscarinic acetylcholine receptors: dependence on agonist as well as receptor-subtype. *Br J Pharmacol* 132 950-958.
- Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K. & Watson, J.D. (2002) Molecular biology of the cell, 4th ed, New York.
- Ambrose, S.H. (1998) Late Pleistocene human population bottlenecks, volcanic winter, and differentiation of modern humans. *J Hum Evol* 34 623-651.
- Ando, J., Ono, Y. & Wright, M.J. (2001) Genetic structure of spatial and verbal working memory. *Behav Genet* 31 615-624.
- Angelova, M., Benkelfat, C. & Turecki, G. (2003) A systematic review of association studies investigating genes coding for serotonin receptors and the serotonin transporter: II. Suicidal behavior. *Mol Psychiatry* 8 646-653.
- Ardlie, K.G., Kruglyak, L. & Seielstad, M. (2002) Patterns of linkage disequilibrium in the human genome. *Nat Rev Genet* 3 299-309.
- Axelrod, J. (1957) O-Methylation of Epinephrine and Other Catechols in vitro and in vivo. *Science* 126 400-401.
- Bark, I.C. & Wilson, M.C. (1994) Human cDNA clones encoding two different isoforms of the nerve terminal protein SNAP-25. *Gene* 139 291-292.
- Bartels, M., Rietveld, M.J., Van Baal, G.C. & Boomsma, D.I. (2002) Genetic and environmental influences on the development of intelligence. *Behav Genet* 32 237-249.
- Bateson, W. (1909) *Mendel's principles of heredity*. University Press, Cambridge.
- Batzler, M.A. & Deininger, P.L. (2002) *Alu* repeats and human genomic diversity. *Nature Reviews Genetics* 3 370.
- Bejerano, G., Lowe, C.B., Ahituv, N., King, B., Siepel, A., Salama, S.R., Rubin, E.M., Kent, W.J. & Haussler, D. (2006) A distal enhancer and an ultraconserved exon are derived from a novel retroposon. *Nature* 441 87-90.
- Black, D.L. (2003) Mechanisms of alternative pre-messenger RNA splicing. *Annu Rev Biochem* 72 291-336.
- Blasi, P., Boyle, P.P., Ledda, M., Novelletto, A., Gibson, K.M. & Jakobs, C. (2004) Structure of human succinic semialdehyde dehydrogenase gene: identification of promoter region and alternatively processed isoforms. *Mol Psych* 76 348-362.
- Bleichrodt, N., Drenth, P.J.D., Zaal, J.N. & Resing, W.C.M. (1984) *Revisie Amsterdams Kinder Intelligentie Test*. Lisse: Swets and Zeitlinger B.V.
- Boomsma, D.I., Busjahn, A. & Peltonen, L. (2002) Classical twin studies and beyond. *Nat Rev Genet* 3 872-882.
- Boomsma, D.I., de Geus, E.J., Vink, J.M., Stubbe, J.H., Distel, M.A., Hottenga, J.J., Posthuma, D., van Beijsterveldt, T.C., Hudziak, J.J., Bartels, M. & Willemsen, G. (2006) Netherlands Twin Register: from twins to twin families. *Twin Res Hum Genet* 9 849-857.
- Boomsma, D.I. & van Baal, G.C.M. (1998) Genetic influences on childhood IQ in 5- and 7-year-old Dutch twins. *Dev Neuropsychol* 14 115-126.
- Bouchard, T.J., Jr. & McGue, M. (1981) Familial studies of intelligence: a review. *Science* 212 1055-1059.
- Boutanaev, A.M., Kalmykova, A.I., Shevelyov, Y.Y. & Nurminsky, D.I. (2002) Large clusters of co-expressed genes in the *Drosophila* genome. *Nature* 420 666-669.
- Bouthenet, M.L., Souil, E., Martres, M.P., Sokoloff, P., Giros, B. & Schwartz, J.C. (1991) Localization of dopamine D3 receptor mRNA in the rat brain using in situ hybridization histochemistry: comparison with dopamine D2 receptor mRNA. *Brain Res* 564 203-219.
- Butcher, L.M., Kennedy, J.K. & Plomin, R. (2006) Generalist genes and cognitive neuroscience. *Curr Opin Neurobiol* 16 145-151.
- Buyske, S., Bates, M.E., Gharani, N., Matise, T.C., Tischfield, J.A. & Manowitz, P. (2006) Cognitive traits link to human chromosomal regions. *Behav Genet* 36 65-76.
- Buzas, B., Belfer, I., Hipp, H., Lorincz, I., Evans, C., Phillips, G., Taubman, J., Max, M.B. & Goldman, D. (2004) Haplotype block and superblock structures of the alpha1-adrenergic receptor genes reveal echoes from the chromosomal past. *Mol Genet Genomics* 272 519-529.
- Cardon, L.R. & Abecasis, G.R. (2003) Using haplotype blocks to map human complex trait loci. *Trends Genet* 19 135-140.

- Cardon, L.R., Smith, S.D., Fulker, D.W., Kimberling, W.J., Pennington, B.F. & DeFries, J.C. (1994) Quantitative trait locus for reading disability on chromosome 6. *Science* 266 276-279.
- Carlson, C.S., Eberle, M.A., Rieder, M.J., Yi, Q., Kruglyak, L. & Nickerson, D.A. (2004) Selecting a maximally informative set of single-nucleotide polymorphisms for association analyses using linkage disequilibrium. *Am J Hum Genet* 74 106-120.
- Carmi, R., Rokhlina, T., Kwitek-Black, A.E., Elbedour, K., Nishimura, D., Stone, E.M. & Sheffield, V.C. (1995) Use of a DNA pooling strategy to identify a human obesity syndrome locus on chromosome 15. *Hum Mol Genet* 4 9-13.
- Carroll, J.B. (1993) *Human cognitive abilities: A survey of factor analytic studies*. New York: Cambridge University Press.
- Caspi, A., Sugden, K., Moffitt, T.E., Taylor, A., Craig, I.W., Harrington, H., McClay, J., Mill, J., Martin, J., Braithwaite, A. & Poulton, R. (2003) Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. *Science* 301 386-389.
- Cheung, J., Estivill, X., Khaja, R., MacDonald, J.R., Lau, K., Tsui, L.C. & Scherer, S.W. (2003) Genome-wide detection of segmental duplications and potential assembly errors in the human genome sequence. *Genome Biol* 4 R25.
- Chorney, M.J., Chorney, K., Seese, N., Owen, M.J., Daniels, J., McGuffin, P., Thompson, L.A., Detterman, D.K., Benbow, C., Lubinski, D., Eley, T. & Plomin, R. (1998) A quantitative trait locus associated with cognitive ability in children. *Psychol Sci* 9 159-166.
- Cohen, B.A., Mitra, R.D., Hughes, J.D. & Church, G.M. (2000) A computational analysis of whole-genome expression data reveals chromosomal domains of gene expression. *Nat Genet* 26 183.
- Comings, D.E., Wu, S., Rostamkhani, M., McGue, M., Lacono, W.G., Cheng, L.S. & MacMurray, J.P. (2003) Role of the cholinergic muscarinic 2 receptor (CHRM2) gene in cognition. *Mol Psychiatry* 8 10-11.
- Conrad, D.F., Andrews, T.D., Carter, N.P., Hurles, M.E. & Pritchard, J.K. (2006) A high-resolution survey of deletion polymorphism in the human genome. *Nat Genet* 38 75-81.
- Crick, F.H. (1968) The origin of the genetic code. *J Mol Biol* 38 367-379.
- Daniels, J., Holmans, P., Williams, N., Turic, D., McGuffin, P., Plomin, R. & Owen, M.J. (1998) A simple method for analyzing microsatellite allele image patterns generated from DNA pools and its application to allelic association studies. *Am J Hum Genet* 62 1189-1197.
- de Bakker, P.I., Burtt, N.P., Graham, R.R., Guiducci, C., Yelensky, R., Drake, J.A., Bersaglieri, T., Penney, K.L., Butler, J., Young, S., Onofrio, R.C., Lyon, H.N., Stram, D.O., Haiman, C.A., Freedman, M.L., Zhu, X., Cooper, R., Groop, L., Kolonel, L.N., *et al* (2006) Transferability of tag SNPs in genetic association studies in multiple populations. *Nat Genet* 38 1298-1303.
- De Quervain, D.J., Henke, K., Aerni, A., Coluccia, D., Wollmer, M.A., Hock, C., Nitsch, R.M. & Papassotiropoulos, A. (2003) A functional genetic variation of the 5-HT<sub>2a</sub> receptor affects human memory. *Nat Neurosci* 6 1141-1142.
- Deffenbacher, K.E., Kenyon, J.B., Hoover, D.M., Olson, R.K., Pennington, B.F., DeFries, J.C. & Smith, S.D. (2004) Refinement of the 6p21.3 quantitative trait locus influencing dyslexia: linkage and association analyses. *Hum Genet* 115 128-138.
- Dick, D.M., Aliev, F., Bierut, L., Goate, A., Rice, J., Hinrichs, A., Bertelsen, S., Wang, J.C., Dunn, G., Kuperman, S., Schuckit, M., Nurnberger, J., Jr., Porjesz, B., Beglieter, H., Kramer, J. & Hesselbrock, V. (2006) Linkage Analyses of IQ in the Collaborative Study on the Genetics of Alcoholism (COGA) Sample. *Behav Genet* 36 77-86.
- Edelmann, L., Pandita, R.K., Spiteri, E., Funke, B., Goldberg, R., Palanisamy, N., Chaganti, R.S., Magenis, E., Shprintzen, R.J. & Morrow, B.E. (1999) A common molecular basis for rearrangement disorders on chromosome 22q11. *Hum Mol Genet* 8 1157-1167.
- Elston, R.C., Buxbaum, S., Jacobs, K.B. & Olson, J.M. (2000) Haseman and Elston revisited. *Genet Epidemiol* 19 1-17.
- Faraone, S.V. & Khan, S.A. (2006) Candidate gene studies of attention-deficit/hyperactivity disorder. *J Clin Psychiatry* 67 Suppl 8 13-20.
- Fisher, S.E., Marlow, A.J., Lamb, J., Maestrini, E., Williams, D.F., Richardson, A.J., Weeks, D.E., Stein, J.F. & Monaco, A.P. (1999) A quantitative-trait locus on chromosome 6p influences different aspects of developmental dyslexia. *Am J Hum Genet* 64 146-156.
- Forman, M.S., Lee, V.M. & Trojanowski, J.Q. (2000) New insights into genetic and molecular mechanisms of brain degeneration in tauopathies. *J Chem Neuroanat* 20 225-244.
- Forrest, W.F. (2001) Weighting improves the "new Haseman-Elston" method. *Hum Hered* 52 47-54.
- Fu, Y.H., Kuhl, D.P., Pizzuti, A., Pieretti, M., Sutcliffe, J.S., Richards, S., Verkerk, A.J., Holden, J.J., Fenwick, R.G., Jr., Warren, S.T. & *et al*. (1991) Variation of the CGG repeat at the fragile X site results in genetic instability: resolution of the Sherman paradox. *Cell* 67 1047-1058.
- Fulker, D.W., Cherny, S.S., Sham, P.C. & Hewitt, J.K. (1999) Combined linkage and association sib-pair analysis for quantitative traits. *Am J Hum Genet* 64 259-267.
- Gabriel, S.B., Schaffner, S.F., Nguyen, H., Moore, J.M., Roy, J., Blumenstiel, B., Higgins, J., DeFelice, M., Lochner, A., Faggart, M., Liu-Cordero, S.N., Rotimi, C., Adeyemo, A., Cooper, R., Ward, R., Lander, E.S., Daly, M.J. & Altshuler, D. (2002) The structure of haplotype blocks in the human genome. *Science* 296 2225-2229.
- Galaburda, A.M., LoTurco, J., Ramus, F., Fitch, R.H. & Rosen, G.D. (2006) From genes to behavior in developmental dyslexia. *Nat Neurosci* 9 1213-1217.
- Glickstein, S.B., Hof, P.R. & Schmauss, C. (2002) Mice Lacking Dopamine D2 and D3 Receptors Have Spatial Working Memory Deficits. *J Neurosci* 22 5619-5629.

- Greenberg, F., Guzzetta, V., Montes de Oca-Luna, R., Magenis, R.E., Smith, A.C., Richter, S.F., Kondo, I., Dobyns, W.B., Patel, P.I. & Lupski, J.R. (1991) Molecular analysis of the Smith-Magenis syndrome: a possible contiguous-gene syndrome associated with del(17)(p11.2). *Am J Hum Genet* 49 1207-1218.
- Grossman, M.H., Emanuel, B.S. & Budarf, M.L. (1992) Chromosomal mapping of the human catechol-O-methyltransferase gene to 22q11.1-q11.2. *Genomics* 12 822-825.
- Han, J.S. & Boeke, J.D. (2005) LINE-1 retrotransposons: modulators of quantity and quality of mammalian gene expression? *Bioessays* 27 775-784.
- Hariri, A.R., Goldberg, T.E., Mattay, V.S., Kolachana, B.S., Callicott, J.H., Egan, M.F. & Weinberger, D.R. (2003) Brain-derived neurotrophic factor val66met polymorphism affects human memory-related hippocampal activity and predicts memory performance. *J Neurosci* 23 6690-6694.
- Haseman, J.K. & Elston, R.C. (1972) The investigation of linkage between a quantitative trait and a marker locus. *Behav Genet* 2 3-19.
- Hellmann-Blumberg, U., Hintz, M.F., Gatewood, J.M. & Schmid, C.W. (1993) Developmental differences in methylation of human Alu repeats. *Mol Cell Biol* 13 4523-4530.
- Hess, E.J., Collins, K.A. & Wilson, M.C. (1996) Mouse model of hyperkinesis implicates SNAP-25 in behavioral regulation. *J Neurosci* 16 3104-3111.
- Heutink, P. & Oostra, B.A. (2002) Gene finding in genetically isolated populations. *Hum Mol Genet* 11 2507-2515.
- Heyser, C.J., Wilson, M.C. & Gold, L.H. (1995) Coloboma hyperactive mutant exhibits delayed neurobehavioral developmental milestones. *Brain Res Dev Brain Res* 89 264-269.
- Hill, L., Chorney, M.J., Lubinski, D., Thompson, L.A. & Plomin, R. (2002) A quantitative trait locus not associated with cognitive ability in children: a failure to replicate. *Psychol Sci* 13 561-562.
- Hill, L., Craig, I.W., Asherson, P., Ball, D., Eley, T., Ninomiya, T., Fisher, P.J., Turic, D., McGuffin, P., Owen, M.J., Chorney, K., Chorney, M.J., Benbow, C.P., Lubinski, D., Thompson, L.A. & Plomin, R. (1999) DNA pooling and dense marker maps: a systematic search for genes for cognitive ability. *Neuroreport* 10 843-848.
- Hoehe, M.R., Wendel, B., Grunewald, I., Chiaroni, P., Levy, N., Morris-Rosendahl, D., Macher, J.P., Sander, T. & Crocq, M.A. (1998) Serotonin transporter (5-HTT) gene polymorphisms are not associated with susceptibility to mood disorders. *Am J Med Genet* 81 1-3.
- Horikawa, H.P., Saisu, H., Ishizuka, T., Sekine, Y., Tsugita, A., Odani, S. & Abe, T. (1993) A complex of rab3A, SNAP-25, VAMP/synaptobrevin-2 and syntaxins in brain presynaptic terminals. *FEBS Lett* 330 236-240.
- Hulme, E.C. (1990) Muscarinic acetylcholine receptors: typical G-coupled receptors. *Symp Soc Exp Biol* 44 39-54.
- Humphreys, L.G. (1985) *General Intelligence: An integration of factor, test, and simplex theory. Handbook of intelligence: Theories, measurements, and applications. B. Wolman, New York: . Wiley.*
- Iafrate, A.J., Feuk, L., Rivera, M.N., Listewnik, M.L., Donahoe, P.K., Qi, Y., Scherer, S.W. & Lee, C. (2004) Detection of large-scale variation in the human genome. *Nat Genet* 36 949-951.
- Ibrahimi, O.A., Chiu, E.S., McCarthy, J.G. & Mohammadi, M. (2005) Understanding the molecular basis of Apert syndrome. *Plast Reconstr Surg* 115 264-270.
- Jensen, A.R. (1998) *The g Factor: The science of mental ability*. Westport, CT: Praeger.
- Johnson, J.M., Castle, J., Garrett-Engle, P., Kan, Z., Loerch, P.M., Armour, C.D., Santos, R., Schadt, E.E., Stoughton, R. & Shoemaker, D.D. (2003) Genome-wide survey of human alternative pre-mRNA splicing with exon junction microarrays. *Science* 302 2141-2144.
- Kidwell, M.G. & Lisch, D.R. (2000) Transposable elements and host genome evolution. *Trends In Ecology And Evolution* 15 95-99.
- Kitaichi, K., Day, J.C. & Quirion, R. (1999b) A novel muscarinic M(4) receptor antagonist provides further evidence of an autoreceptor role for the muscarinic M(2) receptor sub-type. *Eur J Pharmacol* 383 53-56.
- Kitaichi, K., Hori, T., Srivastava, L.K. & Quirion, R. (1999a) Antisense oligodeoxynucleotides against the muscarinic m2, but not m4, receptor supports its role as autoreceptors in the rat hippocampus. *Brain Res Mol Brain Res* 67 98-106.
- Kobayashi, K., Morita, S., Sawada, H., Mizuguchi, T., Yamada, K., Nagatsu, I., Hata, T., Watanabe, Y., Fujita, K. & Nagatsu, T. (1995) Targeted disruption of the tyrosine hydroxylase locus results in severe catecholamine depletion and perinatal lethality in mice. *J Biol Chem* 270 27235-27243.
- Kobayashi, K., Noda, Y., Matsushita, N., Nishii, K., Sawada, H., Nagatsu, T., Nakahara, D., Fukabori, R., Yasoshima, Y., Yamamoto, T., Miura, M., Kano, M., Mamiya, T., Miyamoto, Y. & Nabeshima, T. (2000) Modest neuropsychological deficits caused by reduced noradrenaline metabolism in mice heterozygous for a mutated tyrosine hydroxylase gene. *J Neurosci* 20 2418-2426.
- Krejci, A., Bruce, A.W., Dolezal, V., Tucek, S. & Buckley, N.J. (2004) Multiple promoters drive tissue-specific expression of the human M muscarinic acetylcholine receptor gene. *J Neurochem* 91 88-98.
- Kunugi, H., Tatsumi, M., Sakai, T., Hattori, M. & Nanko, S. (1996) Serotonin transporter gene polymorphism and affective disorder. *Lancet* 347 1340.
- Lagercrantz, H. (1976) On the composition and function of large dense cored vesicles in sympathetic nerves. *Neuroscience* 1 81-92.

- 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. & Lehoczký, J. & LeVine, R. & McEwan, P., *et al* (2001) Initial sequencing and analysis of the human genome. *Nature* 409 860.
- Lander, E.S. & Schork, N.J. (1994) Genetic dissection of complex traits. *Science* 265 2037-2048.
- Lee, C.J. & Irizarry, K. (2003) Alternative splicing in the nervous system: an emerging source of diversity and regulation. *Biological Psychiatry* 54 771.
- Lewontin, R.C. (1964) The interaction of selection and linkage. I. General considerations; heterotic models. *Genetics* 49 49-67.
- Li, Q., Lee, B.T. & Zhang, L. (2005) Genome-scale analysis of positional clustering of mouse testis-specific genes. *BMC Genomics* 6 7.
- Lopez, A.J. (1998) Alternative splicing of pre-mRNA: developmental consequences and mechanisms of regulation. *Annu Rev Genet* 32 279-305.
- Lotta, T., Vidgren, J., Tilgmann, C., Ulmanen, I., Melen, K., Julkunen, I. & Taskinen, J. (1995) Kinetics of human soluble and membrane-bound catechol O-methyltransferase: a revised mechanism and description of the thermolabile variant of the enzyme. *Biochemistry* 34 4202-4210.
- Luciana, M., Collins, P.F. & Depue, R.A. (1998) Opposing roles for dopamine and serotonin in the modulation of human spatial working memory functions. *Cereb Cortex* 8 218-226.
- Luciano, M., Wright, M., Smith, G.A., Geffen, G.M., Geffen, L.B. & Martin, N.G. (2001) Genetic covariance among measures of information processing speed, working memory, and IQ. *Behav Genet* 31 581-592.
- Luciano, M., Wright, M.J., Duffy, D.L., Wainwright, M.A., Zhu, G., Evans, D.M., Geffen, G.M., Montgomery, G.W. & Martin, N.G. (2006) Genome-wide Scan of IQ Finds Significant Linkage to a Quantitative Trait Locus on 2q. *Behav Genet* 36 45-55.
- Malhotra, A.K., Kestler, L.J., Mazzanti, C., Bates, J.A., Goldberg, T. & Goldman, D. (2002) A functional polymorphism in the COMT gene and performance on a test of prefrontal cognition. *Am J Psychiatry* 159 652-654.
- McCarroll, S.A., Hadnott, T.N., Perry, G.H., Sabeti, P.C., Zody, M.C., Barrett, J.C., Dallaire, S., Gabriel, S.B., Lee, C., Daly, M.J. & Altshuler, D.M. (2006) Common deletion polymorphisms in the human genome. *Nat Genet* 38 86-92.
- Meulenbelt, I., Droog, S., Trommelen, G.J., Boomsma, D.I. & Slagboom, P.E. (1995) High-yield noninvasive human genomic DNA isolation method for genetic studies in geographically dispersed families and populations. *Am J Hum Genet* 57 1252-1254.
- Mighell, A.J., Markham, A.F. & Robinson, P.A. (1997) Alu sequences. *FEBS Lett* 417 1-5.
- Miller, S.A., Dykes, D.D. & Polesky, H.F. (1988) A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* 16 1215.
- Morton, N.E. (1955) Sequential tests for the detection of linkage. *Am J Hum Genet* 7 277-318.
- Muratani, K., Hada, T., Yamamoto, Y., Kaneko, T., Shigeto, Y., Ohue, T., Furuyama, J. & Higashino, K. (1991) Inactivation of the cholinesterase gene by Alu insertion: possible mechanism for human gene transposition. *Proc Natl Acad Sci U S A* 88 11315-11319.
- Nagatsu, T., Levitt, M. & Udenfriend, S. (1964) Tyrosine Hydroxylase. The Initial Step In Norepinephrine Biosynthesis. *J Biol Chem* 239 2910-2917.
- Neul, J.L. & Zoghbi, H.Y. (2004) Rett syndrome: a prototypical neurodevelopmental disorder. *Neuroscientist* 10 118-128.
- Nobrega, M.A., Ovcharenko, I., Afzal, V. & Rubin, E.M. (2003) Scanning human gene deserts for long-range enhancers. *Science* 302 413.
- Nobrega, M.A., Zhu, Y., Plajzer-Frick, I., Afzal, V. & Rubin, E.M. (2004) Megabase deletions of gene deserts result in viable mice. *Nature* 431 988.
- Nystuen, A., Benke, P.J., Merren, J., Stone, E.M. & Sheffield, V.C. (1996) A cerebellar ataxia locus identified by DNA pooling to search for linkage disequilibrium in an isolated population from the Cayman Islands. *Hum Mol Genet* 5 525-531.
- Ogilvie, A.D. & Harmar, A.J. (1997) Association between the serotonin transporter gene and affective disorder: the evidence so far. *Mol Med* 3 90-93.
- Ohara, K., Nagai, M., Tsukamoto, T., Tani, K., Suzuki, Y. & Ohara, K. (1998) Functional polymorphism in the serotonin transporter promoter at the SLC6A4 locus and mood disorders. *Biol Psychiatry* 44 550-554.
- Ovcharenko, I., Loots, G.G., Nobrega, M.A., Hardison, R.C., Miller, W. & Stubbs, L. (2005) Evolution and functional classification of vertebrate gene deserts. *Genome Res* 15 137-145.
- Peltonen, L. (2000) Positional cloning of disease genes: advantages of genetic isolates. *Hum Hered* 50 66-75.
- Penrose, L.S. (1953) The general purpose sibpair linkage test. *Ann Eugen* 18 120-124.
- Peretti, C.S., Gierski, F. & Harrois, S. (2004) Cognitive skill learning in healthy older adults after 2 months of double-blind treatment with piribedil. *Psychopharmacology (Berl)* 176 175-181.
- Perry, S.E., Summar, M.L., Phillips, J.A., 3rd & Robertson, D. (1991) Linkage analysis of the human dopamine beta-hydroxylase gene. *Genomics* 10 493-495.
- Petersen, M.B., Adelsberger, P.A., Schinzel, A.A., Binkert, F., Hinkel, G.K. & Antonarakis, S.E. (1991) Down syndrome due to de novo Robertsonian translocation t(14q;21q): DNA polymorphism analysis suggests that the origin of the extra 21q is maternal. *Am J Hum Genet* 49 529-536.
- Petrill, S.A., Lipton, P.A., Hewitt, J.K., Plomin, R., Cherny, S.S., Corley, R. & DeFries, J.C. (2004) Genetic and environmental contributions to general cognitive ability through the first 16 years of life. *Dev Psychol* 40 805-812.



- Plomin, R. (1999) Genetics and general cognitive ability. *Nature* 402 C25-29.
- Plomin, R. (2001) Genetic factors contributing to learning and language delays and disabilities. *Child Adolesc Psychiatr Clin N Am* 10 259-277, viii.
- Plomin, R., Hill, L., Craig, I.W., McGuffin, P., Purcell, S., Sham, P., Lubinski, D., Thompson, L.A., Fisher, P.J., Turic, D. & Owen, M.J. (2001) A genome-wide scan of 1842 DNA markers for allelic associations with general cognitive ability: a five-stage design using DNA pooling and extreme selected groups. *Behav Genet* 31 497-509.
- Plomin, R. & Spinath, F.M. (2004) Intelligence: genetics, genes, and genomics. *J Pers Soc Psychol* 86 112-129.
- Plomin, R., Turic, D.M., Hill, L., Turic, D.E., Stephens, M., Williams, J., Owen, M.J. & O'Donovan, M.C. (2004) A functional polymorphism in the succinate-semialdehyde dehydrogenase (aldehyde dehydrogenase 5 family, member A1) gene is associated with cognitive ability. *Molecular Psychiatry* 9 582-586.
- Polderman, T.J.C., Gosso, M.F., Posthuma, D., Van Beijsterveldt, T.C., Heutink, P., Verhulst, F.C. & Boomsma, D.I. (2006a) A longitudinal twin study on IQ, executive functioning, and attention problems during childhood and early adolescence. *Acta Neurol Belg* 106 191-207.
- Polderman, T.J.C., Stins, J.F., Posthuma, D., Gosso, M.F., Verhulst, F.C. & Boomsma, D.I. (2006b) The phenotypic and genotypic relation between working memory speed and capacity. *Intelligence* 34 549-560.
- Posthuma, D., de Geus, E.J. & Boomsma, D.I. (2001a) Perceptual speed and IQ are associated through common genetic factors. *Behav Genet* 31 593-602.
- Posthuma, D., de Geus, E.J.C. & Boomsma, D.I. (2001b) Perceptual speed and IQ are associated through common genetic factors. *Behav Genet* 31 593-602.
- Posthuma, D., de Geus, E.J.C., Boomsma, D.I. & Neale, M.C. (2004) Combined linkage and association tests in mx. *Behav Genet* 34 179-196.
- Posthuma, D., Luciano, M., de Geus, E.J.C., Wright, M.J., Slagboom, P.E., Montgomery, G.W., Boomsma, D.I. & Martin, N.G. (2005) A Genomewide Scan for Intelligence Identifies Quantitative Trait Loci on 2q and 6p. *Am J Hum Genet* 77 318-326.
- Posthuma, D., Neale, M.C., Boomsma, D.I. & de Geus, E.J.C. (2001c) Are smarter brains running faster? Heritability of alpha peak frequency, IQ, and their interrelation. *Behav Genet* 31 567-579.
- Pritchard, J.K. (2001) Are rare variants responsible for susceptibility to complex diseases? *Am J Hum Genet* 69 124-137.
- Pritchard, J.K. & Rosenberg, N.A. (1999) Use of unlinked genetic markers to detect population stratification in association studies. *Am J Hum Genet* 65 220-228.
- Redon, R., Ishikawa, S., Fitch, K.R., Feuk, L., Perry, G.H., Andrews, T.D., Fiegler, H., Shapero, M.H., Carson, A.R., Chen, W., Cho, E.K., Dallaire, S., Freeman, J.L., Gonzalez, J.R., Gratacos, M., Huang, J., Kalaitzopoulos, D., Komura, D., MacDonald, J.R., *et al* (2006) Global variation in copy number in the human genome. *Nature* 444 444-454.
- Reich, D.E. & Lander, E.S. (2001) On the allelic spectrum of human disease. *Trends Genet* 17 502-510.
- Reuser, A.J. & Drost, M.R. (2006) Lysosomal dysfunction, cellular pathology and clinical symptoms: basic principles. *Acta Paediatr Suppl* 95 77-82.
- Risch, N. & Merikangas, K. (1996) The future of genetic studies of complex human diseases. *Science* 273 1516-1517.
- Rouse, S.T., Edmunds, S.M., Yi, H., Gilmor, M.L. & Levey, A.I. (2000) Localization of M(2) muscarinic acetylcholine receptor protein in cholinergic and non-cholinergic terminals in rat hippocampus. *Neurosci Lett* 284 182-186.
- Savitz, J., Solms, M. & Ramesar, R. (2006) The molecular genetics of cognition: dopamine, COMT and BDNF. *Gene Brain Behav* 5 311-328.
- Schuler, G.D. & Boguski, M.S. & Stewart, E.A. & Stein, L.D. & Gyapay, G. & Rice, K. & White, R.E. & Rodriguez-Tome, P. & Aggarwal, A. & Bajorek, E. & Bentolila, S. & Birren, B.B. & Butler, A. & Castle, A.B. & Chiannikulchai, N. & Chu, A. & Clee, C. & Cowles, S. & Day, P.J., *et al* (1996) A gene map of the human genome. *Science* 274 540-546.
- Scott, D.A., Carmi, R., Elbedour, K., Yosefsberg, S., Stone, E.M. & Sheffield, V.C. (1996) An autosomal recessive nonsyndromic-hearing-loss locus identified by DNA pooling using two inbred Bedouin kindreds. *Am J Hum Genet* 59 385-391.
- Seagar, M. & Takahashi, M. (1998) Interactions between presynaptic calcium channels and proteins implicated in synaptic vesicle trafficking and exocytosis. *J Bioenerg Biomembr* 30 347-356.
- Sebat, J., Lakshmi, B., Malhotra, D., Troge, J., Lese-Martin, C., Walsh, T., Yamrom, B., Yoon, S., Krasnitz, A., Kendall, J., Leotta, A., Pai, D., Zhang, R., Lee, Y.H., Hicks, J., Spence, S.J., Lee, A.T., Puura, K., Lehtimaki, T., *et al* (2007) Strong Association of De Novo Copy Number Mutations with Autism. *Science*.
- Sebat, J., Lakshmi, B., Troge, J., Alexander, J., Young, J., Lundin, P., Maner, S., Massa, H., Walker, M., Chi, M., Navin, N., Lucito, R., Healy, J., Hicks, J., Ye, K., Reiner, A., Gilliam, T.C., Trask, B., Patterson, N., *et al* (2004) Large-scale copy number polymorphism in the human genome. *Science* 305 525-528.
- Service, S., Sabatti, C. & Freimer, N. (2007) Tag SNPs chosen from HapMap perform well in several population isolates. *Genet Epidemiol* 31 189-194.
- Shapiro, M.S., Loose, M.D., Hamilton, S.E., Nathanson, N.M., Gomeza, J., Wess, J. & Hille, B. (1999) Assignment of muscarinic receptor subtypes mediating G-protein modulation of Ca(2+) channels by using knockout mice. *Proc Natl Acad Sci U S A* 96 10899-10904.
- Sibley, D.R. & Monsma, F.J., Jr. (1992) Molecular biology of dopamine receptors. *Trends Pharmacol Sci* 13 61-69.

- Singer, M.F. (1982) SINEs and LINEs: highly repeated short and long interspersed sequences in mammalian genomes. *Cell* 28 433-434.
- Snell, G.D. & Bunker, H.P. (1967) Coloboma Linkage. *Mouse News Lett* 37 34.
- Sokoloff, P., Giros, B., Martres, M.P., Bouthenet, M.L. & Schwartz, J.C. (1990) Molecular cloning and characterization of a novel dopamine receptor (D3) as a target for neuroleptics. *Nature* 347 146-151.
- Spearman, C. (1904) "General Intelligence" objectively determined and measured. *Am J Psychol* 15 201-293.
- Stranger, B.E., Forrest, M.S., Dunning, M., Ingle, C.E., Beazley, C., Thorne, N., Redon, R., Bird, C.P., de Grassi, A., Lee, C., Tyler-Smith, C., Carter, N., Scherer, S.W., Tavare, S., Deloukas, P., Hurles, M.E. & Dermitzakis, E.T. (2007) Relative impact of nucleotide and copy number variation on gene expression phenotypes. *Science* 315 848-853.
- The Autism Genome Project Consortium (2007) Mapping autism risk loci using genetic linkage and chromosomal rearrangement. *Nat Genet* 39 319-328.
- The Encode Project Consortium (2004) The ENCODE (ENCyclopedia Of DNA Elements) Project. *Science* 306 636-640.
- The International HapMap Consortium (2005) A haplotype map of the human genome. *Nature* 437 1299-1320.
- Thurstone, L.L. (1938) Primary Mental Abilities. *Psychometric Monographs*.
- Thurstone, L.L. (1947) *Multiple factor analysis*. Chicago: University of Chicago Press.
- Van Haassen, P.P., De Bruyn, E.E.J., Pijl, Y.L., Poortinga, Y.H., Lutjespelberg, H.C., van der Steene, G., Coetsier, P., Speolders-Claes, R. & Stinissen, J. (1986) *Intelligence Scale for Children-Revised, Dutch Version*. Lisse: Swets and Zetlinger B.V.
- van Kesteren, R.E. & Spencer, G.E. (2003) The role of neurotransmitters in neurite outgrowth and synapse formation. *Rev Neurosci* 14 217-231.
- Velculescu, V.E., Zhang, L., Zhou, W., Vogelstein, J., Basrai, M.A., Bassett, D.E., Jr., Hieter, P., Vogelstein, B. & Kinzler, K.W. (1997) Characterization of the yeast transcriptome. *Cell* 88 243-251.
- Venter, J.C. & Adams, M.D. & Myers, E.W. & Li, P.W. & Mural, R.J. & Sutton, G.G. & Smith, H.O. & Yandell, M. & Evans, C.A. & Holt, R.A. & Gocayne, J.D. & Amanatides, P. & Ballew, R.M. & Huson, D.H. & Wortman, J.R. & Zhang, Q. & Kodira, C.D. & Zheng, X.H. & Chen, L., *et al* (2001) The sequence of the human genome. *Science* 291 1304-1351.
- Verkerk, A.J., Pieretti, M., Sutcliffe, J.S., Fu, Y.H., Kuhl, D.P., Pizzuti, A., Reiner, O., Richards, S., Victoria, M.F., Zhang, F.P. & *et al*. (1991) Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome. *Cell* 65 905-914.
- Versteeg, R., van Schaik, B.D.C., van Batenburg, M.F., Roos, M., Monajemi, R., Caron, H., Bussemaker, H.J. & van Kampen, A.H.C. (2003) Gene Density, Intron Length, GC Content, and Repeat Pattern for Domains of Highly and Weakly Expressed Genes. *Genome Res* 13 1998-2004.
- Volpicelli, L.A. & Levey, A.I. (2004) Muscarinic acetylcholine receptor subtypes in cerebral cortex and hippocampus. *Prog Brain Res* 145 59-66.
- Wainwright, M.A., Wright, M.J., Luciano, M., Montgomery, G.W., Geffen, G.M. & Martin, N.G. (2006) A linkage study of academic skills defined by the Queensland core skills test. *Behav Genet* 36 56-64.
- Wassink, T.H., Brzustowicz, L.M., Bartlett, C.W. & Szatmari, P. (2004) The search for autism disease genes. *Ment Retard Dev Disabil Res Rev* 10 272-283.
- Watson, J.D. & Crick, F.H. (1953) The structure of DNA. *Cold Spring Harb Symp Quant Biol* 18 123-131.
- Wechsler, D. (1986) Wechsler Intelligence Scale for Children-Revised (WISC-R). Lisse: Swets & Zeitlinger.
- Wechsler, D. (1997) WAIS-III Wechsler Adult Intelligence Scale San Antonio, Texas: Psychological Corporation.
- Weinberger, D.R., Egan, M.F., Bertolino, A., Callicott, J.H., Mattay, V.S., Lipska, B.K., Berman, K.F. & Goldberg, T.E. (2001) Prefrontal neurons and the genetics of schizophrenia. *Biol Psychiatry* 50 825-844.
- Weiss, K.M. & Clark, A.G. (2002) Linkage disequilibrium and the mapping of complex human traits. *Trends Genet* 18 19-24.
- Wess, J. (1996) Molecular biology of muscarinic acetylcholine receptors. *Crit Rev Neurobiol* 10 69-99.
- Williams, G.V., Rao, S.G. & Goldman-Rakic, P.S. (2002) The Physiological Role of 5-HT<sub>2A</sub> Receptors in Working Memory. *J Neurosci* 22 2843-2854.
- Wilson, A.F., Elston, R.C., Siervogel, R.M. & Tran, L.D. (1988) Linkage of a gene regulating dopamine-beta-hydroxylase activity and the ABO blood group locus. *Am J Hum Genet* 42 160-166.
- Xu, X., Weiss, S., Xu, X. & Wei, L.J. (2000) A unified Haseman-Elston method for testing linkage with quantitative traits. *Am J Hum Genet* 67 1025-1028.
- Yang, Z., Boffelli, D., Boonmark, N., Schwartz, K. & Lawn, R. (1998) Apolipoprotein(a) gene enhancer resides within a LINE element. *J Biol Chem* 273 891-897.
- Yeo, G., Holste, D., Kreiman, G. & Burge, C. (2004) Variation in alternative splicing across human tissues. *Genome Biology* 5 R74.
- Zhang, K. & Sun, F. (2005) Assessing the power of tag SNPs in the mapping of quantitative trait loci (QTL) with extremal and random samples. *BMC Genetics* 6.
- Zhang, W., Basile, A.S., Gomeza, J., Volpicelli, L.A., Levey, A.I. & Wess, J. (2002) Characterization of central inhibitory muscarinic autoreceptors by the use of muscarinic acetylcholine receptor knock-out mice. *J Neurosci* 22 1709-1717.

# CHAPTER 2

## Association between the *CHRM2* gene and intelligence

This chapter is published as:

MF Gosso, M van Belzen, EJC de Geus, JC Polderman, P Heutink, DI Boomsma, D Posthuma  
(2006) Association between the *CHRM2* gene and intelligence in a sample of 304 Dutch families.  
*Genes, Brain and Behavior* 5: 577-584

*Abstract*

The *CHRM2* gene is thought to be involved in neuronal excitability, synaptic plasticity and feedback regulation of acetylcholine release, and has previously been implicated in higher cognitive processing. In a sample of 667 individuals from 304 families, we genotyped three SNPs in the *CHRM2* gene on 7q31-35. From all individuals standardized intelligence measures were available. Using a test of within-family association, which controls for possible effects of population stratification, a highly significant association was found between the *CHRM2* gene and intelligence. The strongest association was between rs324650 and performance IQ, where the T allele was associated with an increase of 4.6 PIQ points. In parallel to a large family based association, we observed an attenuated - although still significant - population based association, illustrating that population stratification may decrease our chances of detecting allele-trait associations. Such a mechanism has been predicted earlier, and this paper is one of the first to empirically show that family based association methods are not only needed to guard against false positives, but are also invaluable in guarding against false negatives.

## INTRODUCTION

Individual performance across a single aspect of cognitive ability is highly predictive of performance on other aspects of cognitive ability. Indeed, about 40% of the population variance of measures of these individual cognitive processes can be accounted for by a single general intelligence factor (Plomin & Spinath 2004). Multivariate genetic analyses indicate that this general intelligence factor is highly heritable (Boomsma & van Baal 1998; Cherny & Cardon 1994; Plomin *et al.* 1994; Posthuma *et al.* 2001), and that there is substantial overlap in the genes influencing different aspects of cognitive ability. This implies that genes associated with one aspect of cognitive ability are likely to be associated with other aspects as well. As noted by Plomin *et al.* (2004), these quantitative genetic findings make general intelligence an excellent target for molecular genetic research.

In spite of its high heritability, reports on the actual genes influencing intelligence are scarce. Recently, Comings *et al.* (2003) reported an association between a variant of the cholinergic muscarinic receptor 2 (*CHRM2*) gene explaining 1% of the variance in scores on full-scale IQ (FSIQ), and years of education. We recently conducted an autosomal genome scan for intelligence using two independent, unselected samples consisting of 329 Australian families and 100 Dutch families, totalling 625 sib-pairs (Posthuma *et al.* 2005). Although the most promising regions were 2q and 6p, we also found modest evidence for linkage with performance IQ at 7q31-36, right above the *CHRM2* gene.

The cholinergic muscarinic receptor family (M<sub>1</sub>-M<sub>5</sub>) belongs to the superfamily of G-protein-coupled receptors. These receptors activate a multitude of signaling pathways important for modulating neuronal excitability, synaptic plasticity and feedback regulation of acetylcholine (ACh) release and cognitive processes, including learning and memory (Hulme 1990; Volpicelli & Levey 2004; Wess 1996). On the basis of its putative role in cognitive ability, we genotyped three tagging SNPs in the *CHRM2* gene in a sample of 667 Dutch individuals from 304 twin families. This current sample overlaps only marginally (17%) with the sample used in the linkage analysis. A family based genetic association test was used, which allows evaluating evidence for association which is free from spurious effects of population stratification (Abecasis *et al.* 2000; Fulker *et al.* 1999; Posthuma *et al.* 2004).

## MATERIALS AND METHODS

### *Subjects*

All twins and their siblings were part of two larger cognitive studies and were recruited from the Netherlands Twin Registry, which ensures population based sampling (Boomsma 1998). We have previously shown such a population to be representative of the total population with respect to intelligence (Posthuma *et al.* 2000). Informed consent was obtained from the participants (adult cohort) or from their parents if they were under 18 (young cohort). The current study was approved by the institutional review board of the VU University Medical Center. None of the individuals tested suffered from severe physical or mental handicaps, as assessed through surveys sent out to participants or their parents every two years.

### *Young Cohort*

The young cohort consisted of 177 twin pairs born between 1990 and 1992, and 55 siblings. The twins were 12 (mean= 12.42, SD= 0.16) years of age and the siblings were between 8 and 15 (mean= 12.08, SD= 2.56) years old at the time of testing. There were 41 monozygotic male twin pairs (MZM), 28 dizygotic male twin pairs (DZM), 56 monozygotic female twin pairs (MZF), 25 dizygotic female twin pairs (DZF), 27 dizygotic opposite-sex twin pairs (DOS), 28 male siblings and 27 female siblings. Participation in this study included a voluntary agreement to provide buccal swabs for DNA extraction.

### *Adult Cohort*

A total of 793 family members from 317 extended twin families participated in the adult cognition study (Posthuma *et al.* 2001). Participation in this study did not automatically include DNA collection, however, part of the sample (276 subjects) returned to the lab to provide blood for DNA extraction. Mean age was 37.3 years (SD = 12.5). There were 20 monozygotic male twin pairs (MZM), 11 dizygotic male twin pairs (DZM), 1 DZM triplet, 14 monozygotic female twin pairs (MZF), 22 dizygotic female twin pairs (DZF) and 17 dizygotic opposite-sex twin pairs (DOS), 23 female siblings and 23 male siblings, and 59 nine subjects from incomplete twin pairs (18 males, 41 females).

### *Cognitive testing*

In the young cohort, cognitive ability was assessed with the Dutch adaptation of the WISC-R (Wechsler 1986), and consisted of four verbal subtests (similarities, vocabulary, arithmetic, and digit span) and two performance subtests (block design, and object assembly). In the adult cohort, the Dutch adaptation of

the WAISIII-R (Wechsler 1997), assessed IQ and consisted of four verbal subtests (VIQ: information, similarities, vocabulary, and arithmetic) and four performance subtests (PIQ: picture completion, block design, matrix reasoning, and digit-symbol substitution). The correlation between verbal IQ and performance IQ is usually around 0.50 (0.53 in our data), implying that only 25% of the variance in PIQ and VIQ is shared. Thus, a substantial part of the variance in these two measures is non-overlapping, and theoretically they are expected to capture different aspects of cognitive ability. We therefore included VIQ and PIQ as measures of the two different aspects of intelligence as well as Full scale IQ (FSIQ) as a general measure of intelligence. In both cohorts, VIQ, PIQ and FSIQ were normally distributed (see Table 2.1).

**Table 2.1** Means and standard deviations of IQ in the Young and Adult cohorts

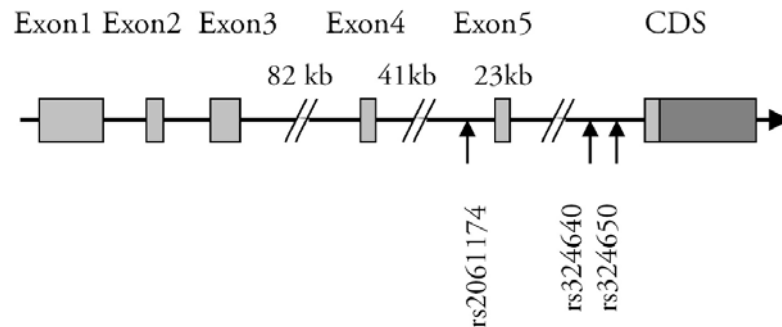
|                | Young Cohort        |                  | Adult Cohort        |                  |
|----------------|---------------------|------------------|---------------------|------------------|
|                | <i>Total sample</i> | <i>Genotyped</i> | <i>Total sample</i> | <i>Genotyped</i> |
| N              | 409                 | 391              | 793                 | 276              |
| Age (SD)       | 12.37 (0.95)        | 12.36 (0.90)     | 37.60 (13.00)       | 37.40 (12.42)    |
| Mean PIQ (SD)  | 101.40 (12.85)      | 101.66 (12.96)   | 100.96 (12.50)      | 100.04 (12.40)   |
| Mean VIQ(SD)   | 98.42 (19.04)       | 98.90 (19.02)    | 92.78 (13.83)       | 93.03 (14.36)    |
| Mean FSIQ (SD) | 99.81 (15.20)       | 100.21 (15.21)   | 95.74 (11.62)       | 95.59 (12.04)    |

### *DNA collection and Genotyping*

Buccal swabs were collected from 391 children; DNA in adults was collected from blood samples in 276 adults. The DNA isolation from buccal swabs was performed using a chloroform/isopropanol extraction (Meulenbelt *et al.* 1995). DNA was extracted from blood samples using the salting out protocol (Miller *et al.* 1988). Zygosity was assessed using 11 polymorphic microsatellite markers (Het > 0.80). Eight subjects were not included in further association analyses due to genotypic incompatibilities on the marker alleles (i.e. more than four different marker alleles observed within one family).

SNPs were selected based on their minor allele frequency (MAF) and genotypic correlation ( $\rho$ ) as obtained from a randomly selected Caucasian sample ([http://www.celeradiagnostics.com/cdx/applera\\_genomics](http://www.celeradiagnostics.com/cdx/applera_genomics)). MAF had to be > 0.10 in order to avoid the rare heterozygous genotypes and SNPs with a above > 0.85 with any of the other SNPs were not selected, to avoid redundancy. SNP genotyping was performed blind to familial status and phenotypic data. Three

tagging SNPs were selected, rs2061174, rs324640, and rs324650, using SNP Browser v2.0.4, (<http://www.appliedbiosystems.com/support/software/snplex/>) (NCBI build 34) (see Figure 2.1).



**Figure 2.1** | Location of single nucleotide polymorphisms (SNPs) within the *CHRM2* gene on chromosome 7

SNP genotyping was performed as part of a SNPLex assay. We here focus on the SNPs in the CHRM2 only, as this gene was selected based on its putative role in cognition and its position under one of our linkage peaks. The SNPLex assay was conducted following the manufacturer's recommendations (Applied Biosystems, Foster city, CA, USA). Stock genomic DNA (gDNA) solutions (50ng/ul) were diluted after fragmentation to a final concentration of 18.5ng/μl. Diluted gDNA aliquots (2μl) were spotted and dried down in 384 well plates (Bioplastics, Landgraaf The Netherlands). Previous to the oligo ligation assay (OLA) reaction, reagents were phosphorylated and diluted (1:1). The OLA reaction was performed in a total volume of 5 μl, which contained 37 ng of gDNA, 0.5μl of ligation buffer, 0.025 μl of 48-SNPLex ligase and 1 μl of activated ligation probe pool. The PCR conditions were 3 minutes at 90 °C, 30 cycles of 15 seconds at 90 °C, 30 seconds at 60 °C and 30 seconds at 51 °C (2% ramp), followed by denaturation at 99 °C for 10 minutes. After this step, a purification step was conducted after which the OLA products were ready to be amplified in a final volume of 10 ul. Exonuclease I 0.1 μl and lambda exonuclease 0.2 μl (Applied Biosystems, Foster city, CA, USA) were added, and incubated at 37°C for 90 minutes followed by a deactivation step at 80°C for 10 minutes. The purified OLA products were diluted (2:3) for further amplification. Amplification of OLA products was performed in 10μl, which contained 2μl of diluted OLA reaction, [1x] SNPLex amplification master mix and



[20x] SNplex amplification primers. The OLA amplification conditions were: 95 °C for 10 minutes, followed by 95 °C for 15 seconds, 63 °C for 60 seconds for 30 cycles. After the hybridization step, analysis of the fluorescence intensity was performed in an aliquot (7.5ul) using ABI Sequencer 3730 (Applied Biosystems, Foster city, CA, USA). All pre-PCR steps were performed on a cooled block. Reactions were carried out in Gene Amp 9700 Thermocycler (Applied Biosystems, Foster city, CA, USA). Data was analyzed using Genemapper v3.7 (Applied Biosystems, Foster city, CA, USA).

### *Statistical analyses*

Allele frequencies of the three selected SNPs were estimated in both Young and Adult cohorts using PEDSTATS (<http://www.sph.umich.edu/csg/abecasis/PedStats>) in which a Hardy-Weinberg test is implemented, based on an exact calculation of the probability of observing a certain number of the minor SNP allele. We first determined the heritability of IQ scores in the young and adult cohorts using the standard strategy of comparing MZ and DZ twin resemblance. Specifically we used the variance decomposition framework implemented in Mx (Posthuma *et al.* 2003).

Genetic association tests were conducted using the program QTDT which implements the orthogonal model proposed by Abecasis *et al.*, 2000 (see also Fulker *et al.*, 1999; Posthuma *et al.*, 2004) This model allows to decompose the genotypic effect into orthogonal between- ( $\beta_b$ ) and within- ( $\beta_w$ ) family components, and also models the residual sib-correlation as a function of polygenic or environmental factors. MZ twins can be included and are modelled as such, by adding zygosity status to the datafile. They are not informative to the within family component (unless they are paired with non-twin siblings), but are informative for the between family component. The between-family association component is sensitive to population admixture, whereas the within-family component is significant only in the presence of LD due to close linkage. Spurious associations may arise in a population that is a mix of two or more genetically distinct subpopulations. Any trait that is more frequent in one of the subpopulation(s) (e.g., because of assortative mating or cultural differences) will show a statistical association with any allele that has a different frequency across those subpopulation(s) (e.g., as a result of different ancestors or genetic drift). In practice, more than two subpopulations may have combined and it will not be obvious from the combined populations whether or not the sample is stratified and in what way.

If population stratification acts to create a false association, the test for association using the within family component is still valid. More importantly, if population stratification acts to hide a genuine association, the test for association using the within family component has more power to detect this association than a population based association test. To correct for multiple testing a Monte-Carlo permutation framework was used to calculate empirical p-values (Abecasis *et al.* 2000). Empirical p-values provide an indication of the empirical levels of type I and type II errors, conditional on the observed data.

### *Results*

Comparison of MZ and DZ twin similarities for IQ measures showed that the observed variation in IQ could be attributed to additive genetic variance and unique environmental variance, and not to shared environmental variance. Heritabilities of PIQ, VIQ and FSIQ were 0.73 (95% CI 0.63-0.80), 0.70 (95% CI 0.59-0.78) and 0.80 (95% CI 0.72-0.85) respectively in the young cohort. Using the complete adult cohort the respective heritabilities for PIQ, VIQ, and FSIQ were 0.71 (95% CI 0.62-0.77), 0.78 (95% CI 0.72-0.83), and 0.78 (95% CI 0.72 – 0.83). These heritability estimates are comparable to those reported previously for these age-cohorts in the Dutch population (Bartels *et al.* 2002; Posthuma *et al.* 2001).

In total, 667 subjects were available for SNP genotyping. On the basis of blind controls and MZ checks, no genotyping errors were found. For SNP rs2061174, 2.8% of the genotypes could not be called (648 genotypes succeeded), for SNP rs324640, 4.0% of the genotypes could not be called (640 genotypes succeeded), and for SNP rs324650, 5.7% of the genotypes could not be called (629 genotypes succeeded). SNP rs2061174 (A/G) in intron 4 had a MAF of 0.34. Two SNPs in intron 5, rs324640 (A/G) and rs324650 (A/T) both had similar MAFs between 0.48-0.49. Observed haplotype frequencies were estimated using Haploview 3.11 which implements the EM-algorithm (<http://www.broad.mit.edu/mpg/haploview>). Only one twin from each MZ pair was included. LD was calculated from the estimated haplotype frequencies. The two SNPs in intron 5, lying 4 kb apart are in very strong LD ( $r^2 > 0.90$ ). The SNP in intron 4, lies about 28 kb apart from the SNPs in the downstream region and shows lower LD with the two SNPs in intron 5 ( $r^2 < 0.35$ ).

Genotypic means per cohort are given in Tables 2.2a and 2.2b. The three SNPs were in Hardy-Weinberg equilibrium in both cohorts as well as in the combined cohort. As the heritabilities were comparable across cohorts, as well as the

allele frequencies and the directions of the genotypic effects, we pooled the two cohorts for the association tests. The models used in QTDT included effects of age and sex on the means and modeled additive allelic between- and within family effects. When testing for the presence of population stratification (i.e. equivalence of the between and within family effects), we found significant evidence for the presence of population stratification in the association between rs324650 and both PIQ and FSIQ ( $p < 0.05$ ), indicating the association effects across the total population are not equal to the association effects as found within families.

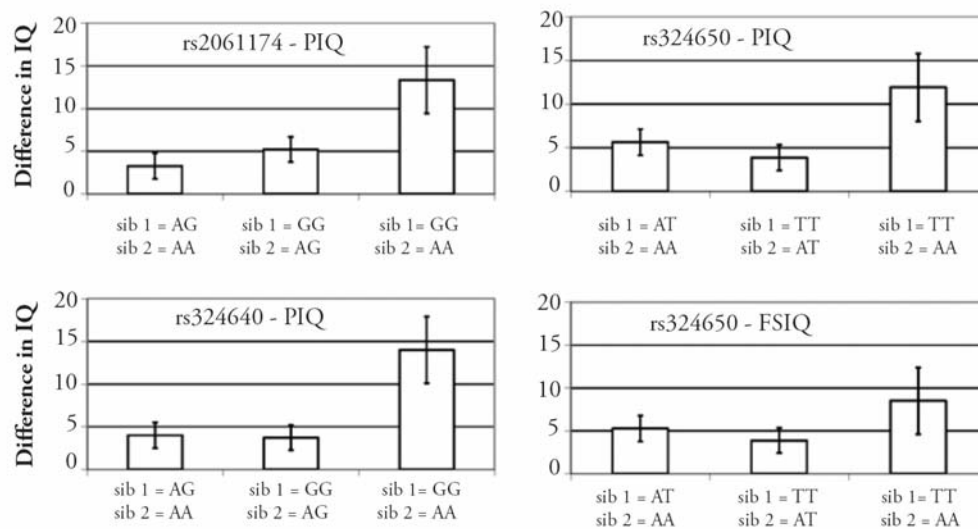
**Table 2.2a** Means (SD) per genotype for performance IQ (PIQ), verbal IQ (VIQ) and full-scale IQ (FSIQ) in the young cohort

|           |                  | Genotype       |                |                 | Total N |
|-----------|------------------|----------------|----------------|-----------------|---------|
|           |                  | AA             | AG             | GG              |         |
| rs2061174 | <i>Frequency</i> | <i>0.43</i>    | <i>0.45</i>    | <i>0.12</i>     |         |
|           | Mean PIQ (SD)    | 100.16 (13.75) | 102.20 (12.55) | 105.26 (11.57)  | 381     |
|           | Mean VIQ (SD)    | 99.81 (18.71)  | 97.56 (19.75)  | 101.00 (18.32)  | 382     |
|           | Mean FSIQ (SD)   | 99.64 (15.52)  | 99.93 (15.40)  | 103.62 (14.16)  | 381     |
| rs324640  |                  | <b>AA</b>      | <b>AG</b>      | <b>GG</b>       |         |
|           | <i>Frequency</i> | <i>0.28</i>    | <i>0.50</i>    | <i>0.22</i>     |         |
|           | Mean PIQ (SD)    | 100.60 (14.02) | 101.81 (12.82) | 102.67 (12.27)  | 381     |
|           | Mean VIQ (SD)    | 98.78 (18.82)  | 98.41 (18.82)  | 100.38 (18.96)  | 382     |
|           | Mean FSIQ (SD)   | 99.18 (16.35)  | 100.10 (14.67) | 101.88 (15.42)  | 381     |
| rs324650  |                  | <b>AA</b>      | <b>AT</b>      | <b>TT</b>       |         |
|           | <i>Frequency</i> | <i>0.29</i>    | <i>0.50</i>    | <i>0.21</i>     |         |
|           | Mean PIQ (SD)    | 100.50 (14.15) | 101.76 (12.74) | 102.96 (12.051) | 379     |
|           | Mean VIQ (SD)    | 97.67 (19.47)  | 98.94 (18.93)  | 100.05 (18.81)  | 380     |
|           | Mean FSIQ (SD)   | 98.56 (16.11)  | 100.35 (14.83) | 101.89 (15.05)  | 379     |

**Table 2.2b** Means (SD) per genotype for PIQ, VIQ and FSIQ in the adult cohort

|           |                  | Genotype      |                |                | Total N |
|-----------|------------------|---------------|----------------|----------------|---------|
|           |                  | AA            | AG             | GG             |         |
| rs2061174 | <i>Frequency</i> | <i>0.45</i>   | <i>0.44</i>    | <i>0.11</i>    |         |
|           | Mean PIQ (SD)    | 98.59 (12.55) | 100.95 (12.79) | 102.76 (9.21)  | 259     |
|           | Mean VIQ (SD)    | 91.94 (14.08) | 94.15 (15.71)  | 93.62 (9.69)   | 260     |
|           | Mean FSIQ (SD)   | 94.06 (12.08) | 97.27 (12.67)  | 96.03 (8.12)   | 256     |
| rs324640  |                  | <b>AA</b>     | <b>AG</b>      | <b>GG</b>      |         |
|           | <i>Frequency</i> | <i>0.26</i>   | <i>0.44</i>    | <i>0.30</i>    |         |
|           | Mean PIQ (SD)    | 98.17 (13.39) | 100.68 (12.64) | 101.09 (11.5)  | 252     |
|           | Mean VIQ (SD)    | 91.26 (16.77) | 93.28 (13.76)  | 93.559 (13.33) | 251     |
|           | Mean FSIQ (SD)   | 94.60 (12.23) | 95.71 (12.92)  | 96.17 (10.78)  | 248     |
| rs324650  |                  | <b>AA</b>     | <b>AT</b>      | <b>TT</b>      |         |
|           | <i>Frequency</i> | <i>0.27</i>   | <i>0.45</i>    | <i>0.29</i>    |         |
|           | Mean PIQ (SD)    | 98.84 (13.07) | 100.41 (12.78) | 100.07 (12.51) | 238     |
|           | Mean VIQ (SD)    | 92.33 (16.13) | 93.22 (13.23)  | 93.39 (13.70)  | 239     |
|           | Mean FSIQ (SD)   | 95.33 (11.72) | 95.48 (12.48)  | 95.94 (11.09)  | 235     |

As it seems obvious that population stratification is caused by pooling the two age cohorts, we also tested for population stratification in each cohort separately and found evidence for population stratification for the same SNPs, in the same direction (i.e. between effects smaller than within), in the young cohort (rs324650 with PIQ: rs324650 with FSIQ:  $p < 0.05$ ), but not in the adult cohort. The within family effects were comparable across both cohorts. Using the within-family association, a significant association of PIQ with all three SNPs was found. The strongest effect was seen with the T allele of rs324650 ( $p < 0.001$ ), which was associated with an increase of 4.6 IQ points among family members (see Table 2.3). Put otherwise, the difference between AT and TT genotypes or AT and AA is 4.6, while the difference between AA and TT genotypes is estimated at 9.2. Notably, effects sizes of the increaser alleles were all reduced, although still significant, in the total association test as compared with the effect sizes based solely on the within family association. As within family associations are not sensitive to spurious associations due to population stratification, whereas between family associations are, this means that stratification acted to hide a true association. Figure 2.2 shows the observed mean difference in IQ-points between different genotypes for individuals within families, for associations significant at the 0.01 level.



**Figure 2.2** | Observed mean difference in IQ scores between siblings (i.e. within family pairs) with different genotypes for those single-nucleotide polymorphisms (SNPs) in the CHRM2 gene that shows a significant association. Sibling pairs include dizygotic pairs and non-twin sibling pairs. The number of pairs on which the difference scores are based is: rs2061174: 97 (AG-AA), 64 (GG-AG) and 15 (GG-AA); rs324640: 87 (AG-AA), 97 (GG-AG) and 9 (GG-AA); rs324650: 75 (AT-AA), 98 (TT-AT) and 11 (TT-AA). FSIQ, full-scale IQ; PIQ, performance IQ

**Table 2.3** Tests for genetic association between the *CHRM2* gene and intelligence (combined cohort)

|           |      | Within family association |                      |                  | Population based association |                    |                  |
|-----------|------|---------------------------|----------------------|------------------|------------------------------|--------------------|------------------|
|           |      | N                         | $\chi^2$             | Genotypic effect | N                            | $\chi^2$           | Genotypic effect |
| rs2061174 | PIQ  | 175                       | 7.3 ( $p < .01$ )    | 3.7 (G)          | 648                          | 9.0 ( $p < .01$ )  | 2.4 (G)          |
|           | VIQ  | 175                       | 0.4                  | 1.2 (G)          | 648                          | 0.0                | 0.2 (G)          |
|           | FSIQ | 174                       | 2.4                  | 2.3 (G)          | 644                          | 2.3                | 1.4 (G)          |
| rs324640  | PIQ  | 209                       | 7.7 ( $p < .01$ )*   | 3.7 (G)          | 640                          | 5.2 ( $p < .05$ )* | 1.8 (G)          |
|           | VIQ  | 209                       | 1.9                  | 2.5 (G)          | 640                          | 1.2                | 1.1 (G)          |
|           | FSIQ | 207                       | 4.6 ( $p < .01$ )    | 3.1 (G)          | 636                          | 2.9                | 1.4 (G)          |
| rs324650  | PIQ  | 193                       | 12.1 ( $p < .001$ )* | 4.6 (T)          | 629                          | 6.0 ( $p < .05$ )* | 1.9 (T)          |
|           | VIQ  | 193                       | 3.9 ( $p < .05$ )    | 3.6 (T)          | 629                          | 1.8                | 1.4 (T)          |
|           | FSIQ | 191                       | 8.0 ( $p < .01$ )*   | 4.1 (T)          | 625                          | 4.0 ( $p < .05$ )* | 1.7 (T)          |

\* statistically significant based on 1000 Monte-Carlo permutations

*Note:* N denotes the number of individuals. For the within family association test it denotes the number of individuals informative for the within family association, i.e. those individuals that occur in families with more than one genotype. The N for the between family association differs slightly from the added totals of Tables 2.2a and 2.2b as QTDT assumes equal genotypes for MZ twins and includes non-typed MZ co-twins with IQ scores

## Discussion

To investigate the possible role of the *CHRM2* gene in intelligence, we employed a family based genetic association test. Significant evidence was found for an association between the *CHRM2* gene and IQ, showing an effect size of 4.6 IQ points for the increaser allele of SNP rs324650 ( $p < 0.001$ ). We also found that the effect sizes based on the within family effects were 1.5 to 2.5 times as large as the population-based effect sizes, suggesting that population stratification resulted in an underestimation of the genuine allelic effect. The attenuation of allelic effects due to population stratification occurs when across subpopulations higher trait values tend to go together with a lower frequency of the increaser allele, or vice versa. This is consistent with findings from mouse model systems in which it has been shown that the same allele at the same locus may cause a major disease in one mouse strain, but no disease phenotype in a strain with a different genetic background (e.g. Linder, 2001; Liu *et al.*, 2001; Montgutelli, 2000). The same has been reported for effects on gene expression in different environmental backgrounds (Cabib *et al.* 2000; Crabbe *et al.* 1999). In humans, the presence of different genetic (or environmental) backgrounds that derive from mixed strata may differentially affect the expression of gene variants (GxE interaction). As non-Mendelian traits are likely to be influenced by multiple (risk-) factors which in turn are likely to interact with each other, neglecting the presence of population stratification may realistically hide genuine allele-trait associations, and may be

responsible in part for the difficulties in replicating reported associations. We previously predicted this phenomenon based on theory and simulations (Posthuma *et al.* 2004) and now show it to occur in practice as well.

The most significant association was seen with rs324650 where the presence of the T allele was associated with an increase of 4.6 points for performance IQ. This extends earlier findings of Comings *et al.* (2003) who found a weaker association between the *CHRM2* gene and both total IQ score and years of education, which was only significant after stratifying on parental origin of transmission. The current study therefore looked at different SNPs than the SNP that was used by Comings *et al.*, (2003). Their SNP was in the 3' UTR region and is not classified as a tagging SNP. The SNPs used in the current sample are tagging SNPs, two of which (rs324640 and rs324650) are in LD with the SNP used in the study by Comings *et al.* (2003). These two SNPs are also the ones that show a significant association with IQ in the current study, albeit stronger than in the study by Comings *et al.*, (2003). This study thus provides further evidence of a role of the cholinergic muscarinic receptor in cognition.

The M2 subtype cholinergic muscarinic receptor is, like the M1 and M4-subtypes, predominantly expressed in the CNS (Volpicelli & Levey 2004). The M2 receptors are predominantly located at the pre-synaptic level (Levey *et al.* 1991; Mrzljak *et al.* 1993), spread throughout the brain but with the highest levels in the cerebral cortical, forebrain cholinergic nuclei, cervical spinal cord region, cerebellum and thalamus (Flynn & Mash 1993; Piggott *et al.* 2002; Spencer *et al.* 1986; Wei *et al.* 1994). M2 receptors are selectively coupled to G-proteins of the Gi/Go family, which mediate the inhibition of voltage-sensitive  $\text{Ca}^{+2}$  channels. Furthermore, the M2 receptor subtype is likely to have an additional role in cholinergic modulation of excitatory and inhibitory hippocampal circuits acting as autoreceptor (Akam *et al.* 2001; Kitaichi *et al.* 1999a; Kitaichi *et al.* 1999b; Rouse *et al.* 2000; Shapiro *et al.* 1999; Zhang *et al.* 2002), inhibiting ACh release from cholinergic terminals. It is well known that exposure to a novel environment causes pronounced ACh release at the level of the neocortex and hippocampus, and that these high levels of ACh are necessary for memory formation (Miranda *et al.* 2000; Pepeu & Giovannini 2004; Ramirez-Lugo *et al.* 2003). Many studies in animals confirm the importance of cholinergic activity for acquisition and retrieval of several learning tasks (Orsetti *et al.* 1996; Vannucchi *et al.* 1997). Importantly, M2 knock-out animal models and studies using selective receptor antagonist agents have shown enhancement in performance in several tasks with cognitive components (Carey *et al.* 2001; Quirion *et al.* 1995; Seeger *et al.* 2004). Finally,

higher M2 distribution volumes have been found in post-mortem and in vivo studies in AD patients compared to healthy controls (Cohen *et al.* 2003).

The association found in the present study as well as by Comings and co-workers were all found with SNPs in (non)coding regions of the gene. We found association with SNPs located in intron 4 (rs2061174), and intron 5 (rs324640, rs324650) of the *CHRM2* gene. Comings *et al.* (2003) found an association between the *CHRM2* gene and IQ with a SNP in the 3'UTR of the gene. Transcription of the *CHRM2* gene is complex. Krejci *et al.*, (2004) determined that the 5'UTR of *CHRM2* consists of four (non)coding regions whose different combinations give rise to eight splice variants. In addition, expression is regulated by two promoters. One promoter regulates expression at the cardiac cell level, whereas the second promoter could be considered neuron specific. Experiments using reporter genes demonstrated that additional regulatory sequences are present further upstream of the proximal promoter(s) and, even more interesting, within the intronic regions (Krejci *et al.* 2004). On the basis of the findings from animal and functional studies (Carey *et al.* 2001; Miranda *et al.* 2000; Pepeu & Giovannini 2004; Quirion *et al.* 1995; Ramirez-Lugo *et al.* 2003; Seeger *et al.* 2004), we hypothesize that a (non)coding polymorphism might be involved in regulation of expression or alternative splicing of the *CHRM2* gene. This polymorphism may subsequently affect mAChR2 transcription, as well as the fine-tuning negative feedback of this particular receptor, making it less reactive to ACh increases during cognitive processing. Another possibility is that SNP rs324650 is in strong LD with the causative regulatory variant in the *CHRM2* gene.

Identifying genes for variation in the range of normal intelligence could provide important clues to the genetic etiology of disturbed cognition in e.g. autism, reading disorder, and ADHD. It is worth to mention the first genome-wide linkage screen in Autism, performed by the International Molecular Genetic Study of Autism Consortium (IMGSAC), involving sib pairs from the United Kingdom. Interestingly, the strongest linkage signal for autism occurred at 7q near the *CHRM2* gene (for a review on linkage scans for autism see Wassink *et al.*, 2004). With regard to attention problems it is of note that a number of recent findings clearly implicate deviant acetylcholinergic neurotransmission in attentional processing (Akari Hagiwara 2005; Beane & Marrocco 2004). Thus far, candidate gene approaches for attention disorder have focused only on genetic variation in nicotinic receptors (Sacco *et al.* 2004; Todd *et al.* 2003). The results in the current study tentatively suggest that muscarinic signaling may be involved as well.

## References

- Abecasis, G.R., Cardon, L.R. & Cookson, W.O. (2000) A general test of association for quantitative traits in nuclear families. *Am J Hum Genet* 66 279-292.
- Akam, E.C., Challiss, R.A. & Nahorski, S.R. (2001) G(q/11) and G(i/o) activation profiles in CHO cells expressing human muscarinic acetylcholine receptors: dependence on agonist as well as receptor-subtype. *Br J Pharmacol* 132 950-958.
- Akari Hagiwara, Y.F.M.D.-T.T.O.R.S. (2005) Differential distribution of release-related proteins in the hippocampal CA3 area as revealed by freeze-fracture replica labeling. *The Journal of Comparative Neurology* 489 195-216.
- Bartels, M., Rietveld, M.J., Van Baal, G.C. & Boomsma, D.I. (2002) Genetic and environmental influences on the development of intelligence. *Behav Genet* 32 237-249.
- Beane, M. & Marrocco, R.T. (2004) Norepinephrine and acetylcholine mediation of the components of reflexive attention: implications for attention deficit disorders. *Prog Neurobiol* 74 167-181.
- Boomsma, D.I. (1998) Twin registers in Europe: an overview. *Twin Res* 1 34-51.
- Boomsma, D.I. & van Baal, G.C.M. (1998) Genetic influences on childhood IQ in 5- and 7-year-old Dutch twins. *Dev Neuropsychol* 14 115-126.
- Cabib, S., Orsini, C., Le Moal, M. & Piazza, P.V. (2000) Abolition and reversal of strain differences in behavioral responses to drugs of abuse after a brief experience. *Science* 289 463-465.
- Carey, G.J., Billard, W., Binch, H., 3rd, Cohen-Williams, M., Crosby, G., Grzelak, M., Guzik, H., Kozlowski, J.A., Lowe, D.B., Pond, A.J., Tedesco, R.P., Watkins, R.W. & Coffin, V.L. (2001) SCH 57790, a selective muscarinic M(2) receptor antagonist, releases acetylcholine and produces cognitive enhancement in laboratory animals. *Eur J Pharmacol* 431 189-200.
- Cherny, S.S. & Cardon, L.R. (1994) General cognitive ability. . In de Fries, J., Plomin, R. & Fulker, D. (eds), *Nature and nurture during middle childhood*. Blackwell Publishers, Oxford, pp. 46-56.
- Cohen, R.M., Podruchny, T.A., Bokde, A.L., Carson, R.E., Herscovitch, P., Kiesewetter, D.O., Eckelman, W.C. & Sunderland, T. (2003) Higher in vivo muscarinic-2 receptor distribution volumes in aging subjects with an apolipoprotein E-epsilon4 allele. *Synapse* 49 150-156.
- Comings, D.E., Wu, S., Rostamkhani, M., McGue, M., Lacono, W.G., Cheng, L.S. & MacMurray, J.P. (2003) Role of the cholinergic muscarinic 2 receptor (CHRM2) gene in cognition. *Mol Psychiatry* 8 10-11.
- Crabbe, J.C., Wahlsten, D. & Dudek, B.C. (1999) Genetics of mouse behavior: interactions with laboratory environment. *Science* 284 1670-1672.
- Flynn, D.D. & Mash, D.C. (1993) Distinct kinetic binding properties of N-[3H]-methylscopolamine afford differential labeling and localization of M1, M2, and M3 muscarinic receptor subtypes in primate brain, *Synapse*, Vol 14, pp 283-296.
- Fulker, D.W., Cherny, S.S., Sham, P.C. & Hewitt, J.K. (1999) Combined linkage and association sib-pair analysis for quantitative traits. *Am J Hum Genet* 64 259-267.
- Hulme, E.C. (1990) Muscarinic acetylcholine receptors: typical G-coupled receptors. *Symp Soc Exp Biol* 44 39-54.
- Kitaichi, K., Day, J.C. & Quirion, R. (1999a) A novel muscarinic M(4) receptor antagonist provides further evidence of an autoreceptor role for the muscarinic M(2) receptor sub-type. *Eur J Pharmacol* 383 53-56.
- Kitaichi, K., Hori, T., Srivastava, L.K. & Quirion, R. (1999b) Antisense oligodeoxynucleotides against the muscarinic m2, but not m4, receptor supports its role as autoreceptors in the rat hippocampus. *Brain Res Mol Brain Res* 67 98-106.
- Krejci, A., Bruce, A.W., Dolezal, V., Tucek, S. & Buckley, N.J. (2004) Multiple promoters drive tissue-specific expression of the human M muscarinic acetylcholine receptor gene. *J Neurochem* 91 88-98.
- Levey, A.I., Kitt, C.A., Simonds, W.F., Price, D.L. & Brann, M.R. (1991) Identification and localization of muscarinic acetylcholine receptor proteins in brain with subtype-specific antibodies. *J Neurosci* 11 3218-3226.
- Linder, C.C. (2001) The influence of genetic background on spontaneous and genetically engineered mouse models of complex diseases. *Lab Anim (NY)* 30 34-39.
- Liu, J., Corton, C., Dix, D.J., Liu, Y., Waalkes, M.P. & Klaassen, C.D. (2001) Genetic background but not metallothionein phenotype dictates sensitivity to cadmium-induced testicular injury in mice. *Toxicol Appl Pharmacol* 176 1-9.
- Meulenbelt, I., Droog, S., Trommelen, G.J., Boomsma, D.I. & Slagboom, P.E. (1995) High-yield noninvasive human genomic DNA isolation method for genetic studies in geographically dispersed families and populations. *Am J Hum Genet* 57 1252-1254.
- Miller, S.A., Dykes, D.D. & Polesky, H.F. (1988) A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* 16 1215.
- Miranda, M.I., Ramirez-Lugo, L. & Bermudez-Rattoni, F. (2000) Cortical cholinergic activity is related to the novelty of the stimulus. *Brain Res* 882 230-235.
- Montagutelli, X. (2000) Effect of the genetic background on the phenotype of mouse mutations. *J Am Soc Nephrol* 11 Suppl 16 S101-105.
- Mrzljak, L., Levey, A.I. & Goldman-Rakic, P.S. (1993) Association of m1 and m2 muscarinic receptor proteins with asymmetric synapses in the primate cerebral cortex: morphological evidence for cholinergic modulation of excitatory neurotransmission. *Proc Natl Acad Sci USA* 90 5194-5198.



- Orsetti, M., Casamenti, F. & Pepeu, G. (1996) Enhanced acetylcholine release in the hippocampus and cortex during acquisition of an operant behavior. *Brain Res* 724 89-96.
- Pepeu, G. & Giovannini, M.G. (2004) Changes in acetylcholine extracellular levels during cognitive processes. *Learn Mem* 11 21-27.
- Piggott, M., Owens, J., O'Brien, J., Paling, S., Wyper, D., Fenwick, J., Johnson, M., Perry, R. & Perry, E. (2002) Comparative distribution of binding of the muscarinic receptor ligands pirenzepine, AF-DX 384, (R,R)-I-QNB and (R,S)-I-QNB to human brain. *J Chem Neuroanat* 24 211-223.
- Plomin, R., Pedersen, N.L., Lichtenstein, P. & McClearn, G.E. (1994) Variability and stability in cognitive abilities are largely genetic later in life. *Behav Genet* 24 207-215.
- Plomin, R. & Spinath, F.M. (2004) Intelligence: genetics, genes, and genomics. *J Pers Soc Psychol* 86 112-129.
- Posthuma, D., Beem, A.L., de Geus, E.J., van Baal, G.C., von Hjelmberg, J.B., Iachine, I. & Boomsma, D.I. (2003) Theory and practice in quantitative genetics. *Twin Res* 6 361-376.
- Posthuma, D., De Geus, E.J., Bleichrodt, N. & Boomsma, D.I. (2000) Twin-singleton differences in intelligence? *Twin Res* 3 83-87.
- Posthuma, D., de Geus, E.J. & Boomsma, D.I. (2001) Perceptual speed and IQ are associated through common genetic factors. *Behav Genet* 31 593-602.
- Posthuma, D., de Geus, E.J., Boomsma, D.I. & Neale, M.C. (2004) Combined linkage and association tests in mx. *Behav Genet* 34 179-196.
- Posthuma, D., Luciano, M., de Geus, E.J.C., Wright, M.J., Slagboom, P.E., Montgomery, G.W., Boomsma, D.I. & Martin, N.G. (2005) A Genomewide Scan for Intelligence Identifies Quantitative Trait Loci on 2q and 6p. *Am J Hum Genet* 77 318-326.
- Quirion, R., Wilson, A., Rowe, W., Aubert, I., Richard, J., Doods, H., Parent, A., White, N. & Meaney, M.J. (1995) Facilitation of acetylcholine release and cognitive performance by an M(2)-muscarinic receptor antagonist in aged memory-impaired. *J Neurosci* 15 1455-1462.
- Ramirez-Lugo, L., Miranda, M.I., Escobar, M.L., Espinosa, E. & Bermudez-Rattoni, F. (2003) The role of cortical cholinergic pre- and post-synaptic receptors in taste memory formation. *Neurobiol Learn Mem* 79 184-193.
- Rouse, S.T., Edmunds, S.M., Yi, H., Gilmor, M.L. & Levey, A.I. (2000) Localization of M(2) muscarinic acetylcholine receptor protein in cholinergic and non-cholinergic terminals in rat hippocampus. *Neurosci Lett* 284 182-186.
- Sacco, K.A., Bannon, K.L. & George, T.P. (2004) Nicotinic receptor mechanisms and cognition in normal states and neuropsychiatric disorders. *J Psychopharmacol* 18 457-474.
- Seeger, T., Fedorova, I., Zheng, F., Miyakawa, T., Koustova, E., Gomeza, J., Basile, A.S., Alzheimer, C. & Wess, J. (2004) M2 muscarinic acetylcholine receptor knock-out mice show deficits in behavioral flexibility, working memory, and hippocampal plasticity. *J Neurosci* 24 10117-10127.
- Shapiro, M.S., Loose, M.D., Hamilton, S.E., Nathanson, N.M., Gomeza, J., Wess, J. & Hille, B. (1999) Assignment of muscarinic receptor subtypes mediating G-protein modulation of Ca(2+) channels by using knockout mice. *Proc Natl Acad Sci U S A* 96 10899-10904.
- Spencer, D.G., Jr., Horvath, E. & Traber, J. (1986) Direct autoradiographic determination of M1 and M2 muscarinic acetylcholine receptor distribution in the rat brain: relation to cholinergic nuclei and projections. *Brain Res* 380 59-68.
- Todd, R.D., Lobos, E.A., Sun, L.W. & Neuman, R.J. (2003) Mutational analysis of the nicotinic acetylcholine receptor alpha 4 subunit gene in attention deficit/hyperactivity disorder: evidence for association of an intronic polymorphism with attention problems. *Mol Psychiatry* 8 103-108.
- Vannucchi, M.G., Scali, C., Kopf, S.R., Pepeu, G. & Casamenti, F. (1997) Selective muscarinic antagonists differentially affect in vivo acetylcholine release and memory performances of young and aged rats. *Neuroscience* 79 837-846.
- Volpicelli, L.A. & Levey, A.I. (2004) Muscarinic acetylcholine receptor subtypes in cerebral cortex and hippocampus. *Prog Brain Res* 145 59-66.
- Wassink, T.H., Brzustowicz, L.M., Bartlett, C.W. & Szatmari, P. (2004) The search for autism disease genes. *Ment Retard Dev Disabil Res Rev* 10 272-283.
- Wechsler, D. (1986) Wechsler Intelligence Scale for Children-Revised (WISC-R). Lisse: Swets & Zeitlinger.
- Wechsler, D. (1997) WAIS-III Wechsler Adult Intelligence Scale San Antonio, Texas: Psychological Corporation.
- Wei, J., Walton, E.A., Milici, A. & Buccafusco, J.J. (1994) m1-m5 muscarinic receptor distribution in rat CNS by RT-PCR and HPLC. *J Neurochem* 63 815-821.
- Wess, J. (1996) Molecular biology of muscarinic acetylcholine receptors. *Crit Rev Neurobiol* 10 69-99.
- Zhang, W., Basile, A.S., Gomeza, J., Volpicelli, L.A., Levey, A.I. & Wess, J. (2002) Characterization of central inhibitory muscarinic autoreceptors by the use of muscarinic acetylcholine receptor knock-out mice. *J Neurosci* 22 1709-1717.



# CHAPTER 3

## Exploring the functional role of the *CHRM2* gene in human cognition

This chapter is published as:

MF Gosso, EJC de Geus, TJC Polderman, DI Boomsma, D Posthuma, and P Heutink (2007)  
Exploring the functional role of the *CHRM2* gene in human cognition: results from a dense  
genotyping and brain expression study. *BMC Med Genet*, 8:66

### *Abstract*

The *CHRM2* gene located on the long arm of chromosome 7 (7q31-35), is involved in neuronal excitability, synaptic plasticity and feedback regulation of acetylcholine release, and has been implicated in higher cognitive processing. Using two independent Dutch cohorts, we previously reported an association between intronic polymorphisms on the *CHRM2* gene and intelligence (Gosso *et al.* 2006) validating its putative role on cognitive variation. This finding was rapidly confirmed in an independent study by Dick and co-workers (2007). Converging evidence now points to two different 5'UTR regions of the *CHRM2* gene to be associated with intelligence. However; no functional variants within this area have yet been reported. In order to identify the relevant functional variant(s), a second round of genetic analyses, using a denser coverage of SNPs, in the *CHRM2* gene was conducted using the original two family-based samples, which consisted of a children's sample (mean age 12.4) and an adult sample (mean age 37.6). An additional 115 subjects were added to the adult sample. For all individuals standardized intelligence measures were available. Using a test of within-family association, which controls for possible effects of population stratification, two of the previous associated variants rs2061174, and rs324650, still showed the strongest association with intelligence ( $P < 0.01$ ). Interestingly, one of these variants, rs324650, is located within a short interspersed repeat (SINE), namely a MIRb (mammalian-wide interspersed repeat) repeat of 160 bp long, which is conserved within the primate lineage. Although the function of short interspersed repeats remains contentious, recent research revealed potential functionality of SINE repeats in a gene-regulatory context. We therefore investigated genotype-dependent *CHRM2* gene expression levels in the brain, to explore putative enhancer/inhibition activity exerted by variants within the muscarinic acetylcholinergic receptor.

## INTRODUCTION

Identifying genes for variation in the range of normal intelligence could provide important clues to the genetic etiology of disturbed cognition in e.g. autism, reading disorder, and ADHD. Since the earliest 90's several groups have focussed on the identification – and subsequent replication – of common genetic polymorphisms underlying normal variation in cognitive abilities (Buyske *et al.* 2006; Cherny & Cardon 1994; Dick *et al.* 2006; Plomin *et al.* 2004; Posthuma *et al.* 2005). Among a handful of candidate genes that have been investigated in relation to normal cognitive variation as summarized in Posthuma & De Geus (2006), the muscarinic 2 cholinergic receptor gene (*CHRM2*) has been consistently found to be associated with cognitive ability, and currently is the best replicated gene associated with general intelligence. A population-based association study conducted by Comings *et al.* (2003) reported an association between a 3'UTR variant of the *CHRM2* gene explaining 1% of the variance in scores on full-scale IQ (FSIQ), and years of education. Subsequently, we found suggestive evidence for linkage with performance IQ at 7q31-36, in the vicinity of the *CHRM2* gene in a genome scan for intelligence based on 329 Australian families and 100 Dutch families, totalling 625 sib-pairs (Posthuma *et al.* 2005). We subsequently reported association between genetic variants within the *CHRM2* gene and intelligence quotient (IQ) using two independent Dutch cohorts (Gosso *et al.* 2006). This finding was then replicated by Dick and colleagues (Dick *et al.* 2007). All three association studies (Comings *et al.*, 2003; Gosso *et al.*, 2006; Dick *et al.*, 2007) report significant association with IQ and non-coding regions within in the *CHRM2* gene (rs8191992 located in the 3' untranslated region (UTR) (Comings *et al.* 2003), and rs2061174 (Dick *et al.* 2007), and rs324650 (Gosso *et al.* 2006b) in introns 4 and 5, respectively).

The *CHRM2* gene belongs to the superfamily of G-protein-coupled receptors (GPCRs). Muscarinic acetylcholine receptors (M<sub>1</sub>-M<sub>5</sub>) activate a multitude of signaling pathways important for modulating neuronal excitability, synaptic plasticity and feedback regulation of acetylcholine (ACh) release (Levey *et al.* 1991; Volpicelli & Levey 2004). Combined behavioral and pharmacological animal studies involving M<sub>2</sub> antagonists have shown the importance of cholinergic receptor activity for acquisition and retrieval of several learning tasks (Accili *et al.* 1996; Carey *et al.* 2001; Gautam *et al.* 2006; Orsetti *et al.* 1996; Seeger *et al.* 2004). Despite its confirmed putative role in cognitive processes, further evidence for genetic regulatory variants on the *CHRM2* gene have been difficult to assess, mainly due to its complex transcriptional expression patterns. Three different *CHRM2* promoters have been reported based on work performed on different human cell lines (Fenech *et al.* 2004). In combination with *alternative* splicing patterns, this results in, at least, 6 different mRNA transcripts encoding for the same receptor protein (isoforms

A till F) (Fenech *et al.* 2004; Krejci *et al.* 2004). Promoter activity for the *CHRM2* gene was postulated to be tissue specific. The first promoter located upstream of exon 1, was preferentially used in cardiac cells (isoforms A and B); promoter 2 on intron 1 alternatively expressed in brain (isoforms C and D); and a third promoter located on intro 2, non-tissue specific (isoforms E and F). Independently, Zhou and coworkers (Zhou *et al.* 2001) reported a fourth putative promoter region on intron 5, but this last result was not independently confirmed yet (Fenech *et al.* 2004). Although *CHRM2* promoter usage is believed to be tissue specific, a single protein receptor is encoded. The functional significance of these transcripts is still unknown. To fine-map the *CHRM2* gene and to detect its functional role in cognitive ability, we genotyped a dense set of *tag*-SNPs within and flanking the *CHRM2* gene in a sample of 762 Dutch individuals from 358 twin families belonging to two different age cohorts (mean ages 12.4 and 37.6). A family-based genetic association test was used, which allows evaluating evidence for association free from spurious effects of population stratification (Abecasis *et al.* 2000; Fulker *et al.* 1999; Posthuma *et al.* 2004). In addition, gene expression assays were performed on brain controls to determine whether a significant correlation exists between the associated SNPs and *CHRM2* gene expression levels.

## MATERIALS AND METHODS

### *Subjects*

All young and adult twins and their siblings were part of two larger cognitive studies and were recruited from the Netherlands Twin Registry (Bartels *et al.* 2002; Boomsma *et al.* 2006). We have shown previously that the adult participants are representative of the Dutch population with respect to intelligence (Posthuma *et al.* 2001). Informed consent was obtained from the participants (adult cohort) or from their parents if they were under 18 (young cohort). The study was approved by the institutional review board of the VU University Medical Center. None of the individuals tested suffered from severe physical or mental handicaps, as assessed through surveys sent out to participants or their parents every two years.

### *Young Cohort*

The young cohort consisted of 177 twin pairs born between 1990 and 1992, and 55 siblings (Polderman *et al.* 2006a; Polderman *et al.* 2006b), of which 371 were available for genotyping. Mean age of the genotyped twins was 12.4 (SD = 0.9) years and the siblings were between 8 and 15 years old at the time of testing. There were 35 monozygotic male twin pairs (MZM), 28 dizygotic male twin pairs (DZM), 48 monozygotic female twin pairs (MZF), 23 dizygotic female twin pairs (DZF), 26 dizygotic opposite-sex twin pairs (DOS), 24 male siblings and 24 female siblings, and 3 subjects form incomplete twin

pairs (1 male, 2 females). Participation in this study included a voluntary agreement to provide buccal swabs for DNA extraction. This sample is similar to the sample used in our initial analyses, except for twenty individuals that were deleted from analyses in the current sample due to a more stringent threshold of genotyping failure per individual.

### *Adult Cohort*

A total of 793 family members from 317 extended twin families participated in the adult cognition study (Posthuma *et al.* 2005). Participation in this study did not automatically include DNA collection, however, part of the sample, 276 subjects returned to the lab to provide blood samples. The sample characteristics have been described elsewhere (Gosso *et al.* 2006). One hundred fifteen additional individuals provided buccal swabs via the NTR Biobank project (Hoekstra *et al.* 2004) for DNA extraction. Mean age of the total genotyped sample was 36.2 years (SD= 12.6). There were 25 monozygotic male twin pairs (MZM), 15 dizygotic male twin pairs (DZM), 1 DZM triplet, 20 monozygotic female twin pairs (MZF), 28 dizygotic female twin pairs (DZF) and 23 dizygotic opposite-sex twin pairs (DOS), 29 female siblings and 28 male siblings, and 109 subjects from incomplete twin pairs (41 males, 68 females).

### *Cognitive testing*

In the young cohort, cognitive ability was assessed with the Dutch adaptation of the WISC-R (Wechsler 1986), and consisted of four verbal subtests (similarities, vocabulary, arithmetic, and digit span) and two performance subtests (block design, and object assembly). In the adult cohort, the Dutch adaptation of the WAISIII-R (Wechsler 1997), assessed IQ and consisted of four verbal subtests (VIQ: information, similarities, vocabulary, and arithmetic) and four performance subtests (PIQ: picture completion, block design, matrix reasoning, and digit-symbol substitution). The correlation between verbal IQ and performance IQ is usually around 0.50 (0.53 in our data), implying that only 25% of the variance in PIQ and VIQ is shared. Thus, a substantial part of the variance in these two measures is non-overlapping, and theoretically they are expected to capture different aspects of cognitive ability. We therefore included VIQ and PIQ as measures of the two different aspects of intelligence as well as Full scale IQ (FSIQ) as a general measure of intelligence. In both cohorts, VIQ, PIQ and FSIQ were normally distributed, (see Table 3.1). For both cohorts IQ scores standardized for the effects of age and sex were calculated. These were then z-transformed within cohorts to allow easy comparison across cohorts and across different tests.

**Table 3.1** Means and standard deviations of IQ (corrected for age and sex effects) in the young and adult cohorts

|              | Young Cohort   |                      |                |                      | Adult Cohort   |                      |                |                      |
|--------------|----------------|----------------------|----------------|----------------------|----------------|----------------------|----------------|----------------------|
|              | Total sample   | Skewness<br>Kurtosis | Genotyped      | Skewness<br>Kurtosis | Total sample   | Skewness<br>Kurtosis | Genotyped      | Skewness<br>Kurtosis |
| N            | 407            |                      | 371            |                      | 793            |                      | 391            |                      |
| Gender (M/F) | 191/216        |                      | 176/195        |                      | 348/445        |                      | 175/216        |                      |
| Age (SD)     | 12.37 (0.93)   |                      | 12.37 (0.92)   |                      | 37.60 (13.00)  |                      | 36.25 (12.64)  |                      |
| PIQ (SD)     | 94.57 (18.93)  | 0.165<br>-0.308      | 94.85 (19.14)  | 0.175<br>-0.304      | 104.49 (12.34) | 0.197<br>0.099       | 104.30 (11.64) | 0.135<br>0.312       |
| VIQ (SD)     | 102.56 (12.74) | 0.121<br>0.242       | 102.64 (12.92) | -0.080<br>-0.332     | 103.69 (12.26) | -0.308<br>-0.005     | 104.23 (12.15) | -0.410<br>0.256      |
| FSIQ (SD)    | 98.65 (15.06)  | -0.042<br>-0.252     | 98.84 (15.24)  | -0.03<br>-0.254      | 103.56 (11.49) | 0.087<br>0.167       | 103.81(11.16)  | 0.073<br>0.512       |

### *DNA collection and isolation*

Buccal swabs were collected from 371 children; DNA in adults was collected from blood samples (276 subjects) and buccal swabs (115 subjects). The DNA isolation from buccal swabs was performed using a chloroform/isopropanol extraction (Meulenbelt *et al.* 1995; Min *et al.* 2006). DNA was extracted from blood samples using the salting out protocol described elsewhere (Miller *et al.* 1988). Zygosity was assessed using 11 highly polymorphic microsatellite markers (Heterozygosity > 0.80). Genotyping was performed blind to familial status and phenotypic data.

### *DNA and RNA extraction from tissue homogenates*

Control brains from 50 individuals, 23 males with a mean age of 70.3 years (SD= 9.38), and 27 females with a mean age of 73.3 years (SD=10.50) were obtained at autopsy from The Netherlands Brain Bank (<http://www.brainbank.nl>). This material comes mainly from the superior and inferior parietal lobe. DNA isolation from 0.20 gram of frozen tissue was performed using the Puregene™ Kit (Gentra Systems, USA) according to protocol and doubled volume of all reagents per tissue weight. To verify DNA isolation, products were run on a 1% agarose gel.

Total RNA was isolated from 0.10 gram of frozen brain tissue with RNA-Bee™ following the manufacturer's recommendations (Isotex Diagnostics, Inc., USA). RNA was purified using the Qiagen RNeasy Mini kit (Qiagen Benelux B.V., The Netherlands) according to protocol and verified on a 2% agarose gel. Five µg RNA was used to make cDNA using 200 U of Superscript™ III Reverse Transcriptase (Invitrogen, The Netherlands) in First Strand Buffer (Invitrogen, The Netherlands),  $3.4 \times 10^{-2}$  µg/µl random hexamer oligo,  $3.4 \times 10^{-2}$  µg/µl poly d(T) 12-18, 1.3 mM dNTPs, 1.1 µM DTT (Invitrogen, The Netherlands), 10 U RNaseOUT™ Ribonuclease Inhibitor Recombinant (Invitrogen, The Netherlands) and incubated two hours at 50°C. Subsequently, 20 U RNase H (Invitrogen, The Netherlands) was added and incubated 30 minutes at 37°C. Products were run on a 1% agarose gel to examine the quality.



### Genotyping

Single nucleotide polymorphisms (SNPs) were selected using the information available from the International HapMap Project. SNP selection was based on a randomly selected population with northern and western European ancestry by the Centre d'Etude du polymorphisme Humain (CEPH) (<http://www.hapmap.org/thehapmap.html.en>). Allele Frequency MAF had to be  $> 0.05$  in order to exclude rare homozygous genotypes. Forty-two SNPs within the *CHRM2* gene were thus selected from the CEPH population using Haploview version 3.32 (NCBI build 36.1).

SNP genotyping was performed using the SNPLEX<sup>®</sup> assay platform. The SNPLEX assay was conducted following the manufacturer's recommendations (Applied Biosystems, Foster city, CA, USA). All pre-PCR steps were performed on a cooled block. Reactions were carried out in Gene Amp 9700 Thermocycler (Applied Biosystems, Foster city, CA, USA). Data was analyzed using Genemapper v3.7 (Applied Biosystems, Foster city, CA, USA).

### *CHRM2* transcripts at brain level

Three different primer combinations were used to investigate the presence of *CHRM2* transcript variants in normal brain controls. Forward primers F<sub>A&B</sub>GAGGCATCCAGGTCTCCAT, F<sub>C&D</sub>CGCAGCTCTCGCCA-GAGCCTT, and F<sub>E&F</sub>AAAGGACTCCTCGCTCCTTC were used in combination with a unique reverse primer R<sub>A-F</sub>CCCGATAATGGTCACCAAAC in order to tag isoforms A till F. PCR was performed at 94°C for 30 sec, 55°C for 30 sec, and 72°C for 1:30 min, for 40 cycles, followed by a 7 min extension at 72°C. To verify primers specificity PCR products were run on a 2% agarose gel.

### Gene expression assay

RT-PCR was performed using specific primers encompassing the untranslated exon 5 (the last untranslated exon), which is present in all mRNA transcripts, and the coding sequence (CDS) of the *CHRM2* gene; F-GAAACCAGCGACAGGTTTAAATG, R-GCTATTGTTAGAGGAGTTTGTTGAGTTATTC. PCR was carried out at 94°C for 1 min, 64°C for 1 min, and 72°C for 1 min, for 40 cycles, followed by a 10 min extension at 72°C. Optimization of primer concentration and cDNA input was performed and dissociation curves for the selected primers obtained. Two housekeeping genes - *β-actin* and *HPRT*- were used as internal controls. RT-PCR reactions were performed twice independently, each time in duplicate.

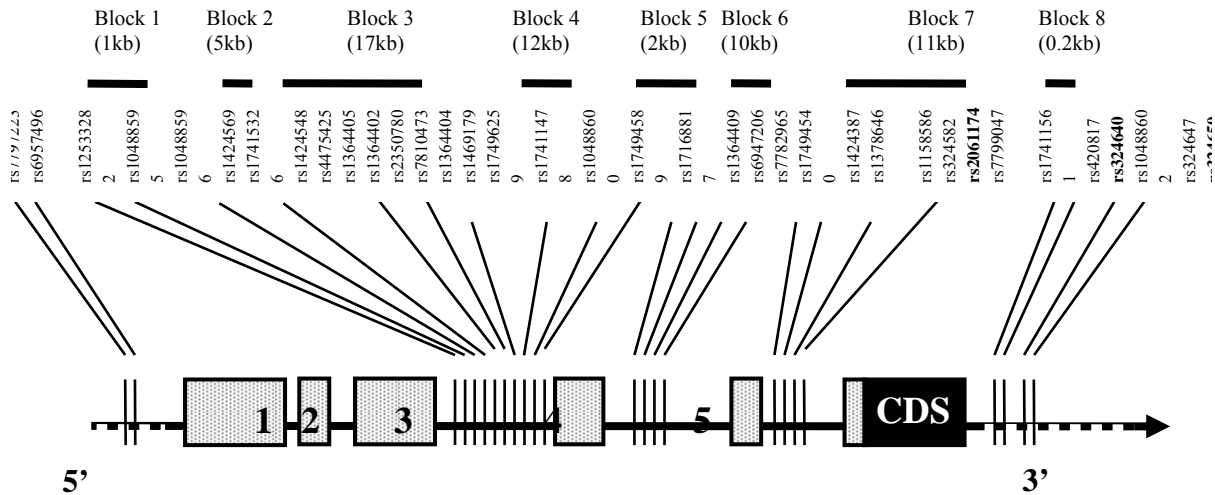
### *Statistical analyses*

Allele frequencies of all SNPs were estimated in both cohorts using Haploview (<http://www.broad.mit.edu/mpg/haploview/>) in which a Hardy-Weinberg test is implemented, based on an exact calculation of the probability of observing a certain number of heterozygotes conditional on the number of copies of the minor SNP allele. Genetic association tests were conducted using the program QTDT which implements the orthogonal model proposed by Abecasis *et al.*, 2000 (see also Fulker *et al.*, 1999; Posthuma *et al.*, 2004). This model allows one to decompose the genotypic effect into orthogonal between- ( $\beta_b$ ) and within- ( $\beta_w$ ) family components, and also models the residual sib-correlation as a function of polygenic or environmental factors. MZ twins can be included and are modelled as such, by adding zygosity status to the datafile. They are not informative to the within family component (unless they are paired with non-twin siblings), but are informative for the between family component.

The between-family association component is sensitive to population admixture, whereas the within-family component is significant only in the presence of LD due to close linkage. The models used in QTDT take into account additive allelic between- and within family effects. It is worth noting that, if population stratification acts to create a false association, the test for association using the within family component is still valid. More importantly, if population stratification acts to hide a genuine association, the test for association using the within family component has more power to detect this association than a population based association test. A significance level  $\alpha$  of 0.01 was chosen.

### *Results*

Genotyping success rate was 95.36 (SD = 3.80) among both cohorts. Six *tag*-SNPs, (rs6957496, rs1424569, rs10488600, rs17494540, rs324582, and rs11773032), although with high success genotyping rate, deviated from HWE ( $P < 0.05$ ). One *tag*-SNP, rs11773032 showed no variation in our population and were thus deleted from further analysis. LD parameters  $D'$  and  $r^2$  were obtained for all successfully genotyped SNPs. LD blocks were generated applying the algorithm defined by Gabriel *et al.*, (2002), in which confidence bounds on  $D'$  are generated if 95% of the information shows “strong LD”. By default, this method ignores markers with MAF  $< 0.05$  (see Figure 3.1 and Table 3.2).



**Figure 3.1** | Location of single nucleotide polymorphisms (SNPs) within the *CHRM2* gene on chromosome 7 and LD blocks defined by them, respectively. Coding sequence (CDS) is depicted in black. Untranslated exons (Exon 1 till Exon 5) are depicted in grey. SNPs already reported in our previous study (Gosso *et al.*, 2006) are in **bold**

Two 5'UTR SNPs previously reported showed the strongest association with IQ, rs2061174 (intron 4) in the *adult* cohort and rs324650 (intron 5) in the *young* cohort (Gosso *et al.* 2006b) (see Figure 3.2). Within-family genetic effects were reflected in an increased IQ of 6.89 (PIQ) points for those individuals bearing the “A” allele (rs2061174) within the *adult* cohort, whereas individuals in the *young* cohort bearing the “T” allele (rs324650), showed an increment of 5.30 IQ (VIQ) points (see Table 3.3). Interestingly, the most significant variant in the *young* cohort, rs324650, is part of a short interspersed repeat (SINE), namely a MIRb (mammalian-wide interspersed repeat) repeat of 160 bp long. The derived “T” allele contained in this repeat seems to be human-specific. In addition, this MIRb repeat is also present in no-human primates lineages –rhesus (*macaca mulatto*) and chimpanzee (*pan troglodytes*)-, but not in other mammalian lineages. Such an allele-specific effect may reflect that the variant being in LD with the *causal* allele, or that the “T” allele is, directly modifying binding-properties of transcription starting sites (TSS) (Mei & Yan 1985).

**Table 3.2** SNPs descriptives for young, adult and combined cohorts

|    | rs#               | Position <sup>a</sup> | Tagged SNP     | LD <sub>BLOCK</sub> | M.A. | MAF Young | MAF Adult | HWpval | %Geno |
|----|-------------------|-----------------------|----------------|---------------------|------|-----------|-----------|--------|-------|
| 1  | rs7797223         | 136198443             |                |                     | T    | 0.25      | 0.26      | 0.77   | 95.6  |
| 2  | <b>rs6957496</b>  | 136202377             | 1              |                     | G    | 0.09      | 0.11      | 0.02   | 96.6  |
| 3  | <b>rs12533282</b> | 136207518             | 1, 4, 5,       | <b>1 (1 kb)</b>     | G    | 0.18      | 0.17      | 0.40   | 98.6  |
| 4  | rs10488595        | 136208970             |                |                     | A    | 0.18      | 0.17      | 0.71   | 97.7  |
| 5  | rs10488596        | 136209134             |                |                     | T    | 0.18      | 0.16      | 0.37   | 97.2  |
| 6  | <b>rs1424569</b>  | 136211219             |                |                     | A    | 0.44      | 0.47      | 0.02   | 94.8  |
| 7  | <b>rs17415326</b> | 136214872             |                |                     | C    | 0.02      | 0.05      | 0.48   | 95.1  |
| 8  | rs1424548         | 136219956             |                |                     | T    | 0.37      | 0.36      | 0.53   | 98.2  |
| 9  | <b>rs4475425</b>  | 136225739             |                | <b>2 (5kb)</b>      | A    | 0.21      | 0.24      | 0.87   | 94.8  |
| 10 | <b>rs1364405</b>  | 136231025             | 41             |                     | A    | 0.35      | 0.33      | 0.08   | 97.9  |
| 11 | <b>rs1364402</b>  | 136234903             |                | <b>3 (17 kb)</b>    | C    | 0.12      | 0.11      | 1.00   | 98.7  |
| 12 | <b>rs2350780</b>  | 136243509             |                |                     | G    | 0.40      | 0.39      | 0.59   | 98.8  |
| 13 | <b>rs7810473</b>  | 136246997             |                |                     | G    | 0.42      | 0.42      | 0.30   | 98.9  |
| 14 | rs1364404         | 136248827             |                |                     | T    | 0.31      | 0.32      | 0.62   | 98.1  |
| 15 | <b>rs1469179</b>  | 136251497             | 22             |                     | A    | 0.44      | 0.46      | 0.18   | 97.3  |
| 16 | rs17496259        | 136251883             |                |                     | A    | 0.31      | 0.31      | 0.48   | 95.8  |
| 17 | rs17411478        | 136251909             |                |                     | T    | 0.31      | 0.32      | 0.37   | 99.1  |
| 18 | <b>rs10488600</b> | 136255998             |                |                     | T    | 0.10      | 0.13      | 0.00   | 98.0  |
| 19 | <b>rs17494589</b> | 136256129             | 26             |                     | A    | 0.20      | 0.18      | 0.07   | 94.9  |
| 20 | <b>rs17168817</b> | 136258808             |                |                     | T    | 0.08      | 0.06      | 0.87   | 99.2  |
| 21 | rs1364409         | 136262573             |                | <b>4 (12 kb)</b>    | T    | 0.32      | 0.35      | 0.13   | 96.4  |
| 22 | rs6947206         | 136265651             |                |                     | C    | 0.46      | 0.48      | 0.12   | 94.0  |
| 23 | <b>rs7782965</b>  | 136274673             | 21, 26, 27     |                     | T    | 0.32      | 0.35      | 0.45   | 90.4  |
| 24 | rs17494540        | 136277380             |                |                     | C    | 0.20      | 0.18      | 0.01   | 96.3  |
| 25 | rs1424387         | 136282543             |                |                     | C    | 0.31      | 0.31      | 0.39   | 99.0  |
| 26 | rs1378646         | 136285541             |                | <b>5 (2kb)</b>      | G    | 0.35      | 0.37      | 0.32   | 98.8  |
| 27 | rs1158586         | 136287676             |                |                     | G    | 0.34      | 0.40      | 0.42   | 93.0  |
| 28 | <b>rs324582</b>   | 136301147             |                |                     | G    | 0.07      | 0.10      | 0.02   | 96.4  |
| 29 | <b>rs2061174</b>  | 136311940             | 30             | <b>6 (10 kb)</b>    | G    | 0.34      | 0.35      | 0.93   | 84.6  |
| 30 | rs7799047         | 136322098             |                |                     | G    | 0.34      | 0.35      | 1.00   | 93.5  |
| 31 | <b>rs17411561</b> | 136332728             | 14, 16, 17, 25 |                     | C    | 0.32      | 0.25      | 0.25   | 87.7  |
| 32 | rs420817          | 136337943             |                | <b>7 (11kb)</b>     | C    | 0.48      | 0.47      | 0.21   | 95.7  |
| 33 | <b>rs324640</b>   | 136339536             | 32             |                     | G    | 0.46      | 0.50      | 0.17   | 86.2  |
| 34 | <b>rs10488602</b> | 136341043             |                |                     | C    | 0.22      | 0.23      | 0.43   | 98.1  |
| 35 | rs324647          | 136343292             |                |                     | C    | 0.14      | 0.15      | 0.13   | 95.9  |
| 36 | <b>rs324650</b>   | 136344201             |                |                     | T    | 0.47      | 0.48      | 0.08   | 85.2  |
| 37 | <b>rs324651</b>   | 136349801             | 35             |                     | T    | 0.14      | 0.14      | 0.13   | 93.1  |
| 38 | <b>rs8191992</b>  | 136351848             |                | <b>8 (0.2 kb)</b>   | T    | 0.45      | 0.48      | 0.60   | 96.5  |
| 39 | <b>rs8191993</b>  | 136352103             |                |                     | G    | 0.35      | 0.35      | 0.93   | 94.9  |
| 40 | <b>rs7780181</b>  | 136357075             |                |                     | G    | 0.42      | 0.44      | 0.83   | 98.7  |
| 41 | rs1424543         | 136360300             |                |                     | C    | 0.36      | 0.32      | 0.01   | 95.4  |
| 42 | <b>rs11971309</b> | 136362695             | 8              |                     | T    | 0.38      | 0.37      | 0.57   | 90.0  |
| 43 | <b>rs11773032</b> | 136391582             |                |                     | A    | 0.00      | 0.01      | 1.00   | 98.1  |

<sup>a</sup> Chromosomal single nucleotide position (SNP) position based on *Build 36.1*. Tag-SNPs are depicted in **bold**

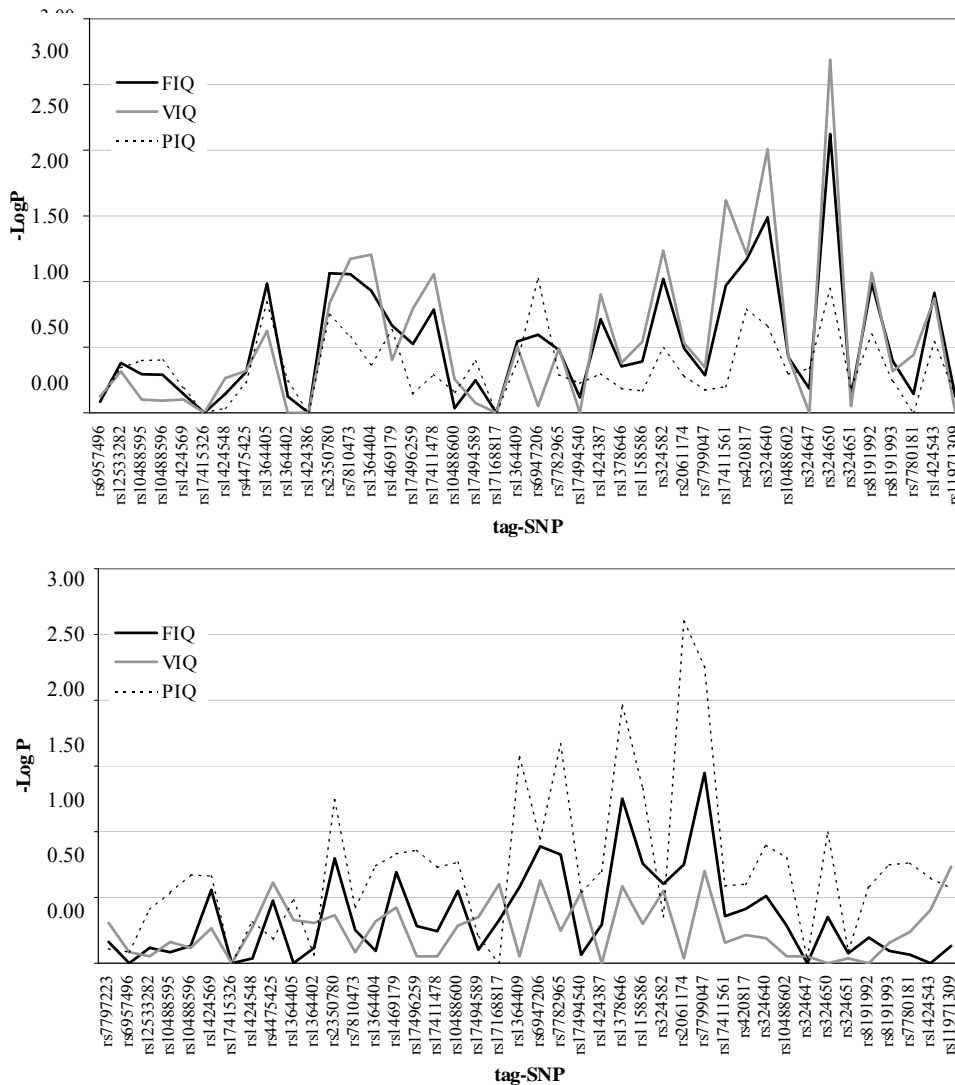
**Table 3.3** QTDT results for the within-family analysis for *young* and *adult* cohorts for the most significant variants among *CHRM2* gene

| <i>position (bp)</i>      | <b>Phenotype</b> | <i>Young cohort</i> |                            |              |                         | <i>Adult cohort</i> |                            |              |                         |
|---------------------------|------------------|---------------------|----------------------------|--------------|-------------------------|---------------------|----------------------------|--------------|-------------------------|
|                           |                  | <b>N</b>            | <b><math>\chi^2</math></b> | <b>P</b>     | <b>Genotypic Effect</b> | <b>N</b>            | <b><math>\chi^2</math></b> | <b>P</b>     | <b>Genotypic Effect</b> |
| rs2350780<br>(136243509)  | PIQ              | 95                  | 1.81                       | 0.179        | 3.63 (A)                | 95                  | 3.62                       | 0.057        | 3.31 (A)                |
|                           | VIQ              | 95                  | 2.11                       | 0.147        | 2.47 (A)                | 95                  | 0.62                       | 0.431        | 1.26 (A)                |
|                           | FSIQ             | 95                  | 2.94                       | 0.086        | 3.48 (A)                | 95                  | 1.98                       | 0.159        | 2.22 (A)                |
| rs1364409<br>(136262573)  | PIQ              | 96                  | 0.67                       | 0.413        | 2.33 (A)                | 92                  | 4.90                       | 0.027        | 3.13 (A)                |
|                           | VIQ              | 96                  | 1.02                       | 0.313        | 1.84 (A)                | 92                  | 0.02                       | 0.888        | 1.05 (A)                |
|                           | FSIQ             | 96                  | 1.14                       | 0.286        | 2.23 (A)                | 92                  | 1.25                       | 0.264        | 0.72 (A)                |
| rs7782965<br>(136274673)  | PIQ              | 85                  | 0.180                      | 0.671        | 2.00 (C)                | 91                  | 5.29                       | 0.021        | 3.36 (C)                |
|                           | VIQ              | 85                  | 0.43                       | 0.512        | 1.74 (C)                | 91                  | 0.33                       | 0.566        | 0.16 (C)                |
|                           | FSIQ             | 85                  | 0.94                       | 0.332        | 2.05 (C)                | 91                  | 2.08                       | 0.149        | 1.60 (C)                |
| rs1378646<br>(136214872)  | PIQ              | 98                  | 0.20                       | 0.655        | 1.26 (A)                | 90                  | <b>6.48</b>                | <b>0.011</b> | <b>3.77 (A)</b>         |
|                           | VIQ              | 98                  | 0.66                       | 0.417        | 1.39 (A)                | 90                  | 1.27                       | 0.260        | 1.10 (A)                |
|                           | FSIQ             | 98                  | 0.59                       | 0.442        | 1.55 (A)                | 90                  | 3.65                       | 0.056        | 2.36 (A)                |
| rs2061174<br>(136311940)  | PIQ              | 111                 | 0.41                       | 0.522        | 1.69 (A)                | 101                 | <b>9.14</b>                | <b>0.003</b> | <b>6.89 (A)</b>         |
|                           | VIQ              | 111                 | 1.10                       | 0.294        | 1.68 (A)                | 101                 | 0.01                       | 0.920        | 1.78 (A)                |
|                           | FSIQ             | 111                 | 0.98                       | 0.322        | 1.91 (A)                | 101                 | 1.82                       | 0.177        | 3.76 (A)                |
| rs17411561<br>(136332728) | PIQ              | 85                  | 0.23                       | 0.632        | 1.47 (C)                | 79                  | 1.28                       | 0.589        | 0.69 (C)                |
|                           | VIQ              | 85                  | 5.09                       | 0.024        | 4.35 (C)                | 79                  | 0.15                       | 0.699        | 0.44 (T)                |
|                           | FSIQ             | 85                  | 2.59                       | 0.108        | 3.61 (C)                | 79                  | 0.60                       | 0.439        | 0.08 (C)                |
| rs324640<br>(136339536)   | PIQ              | 105                 | 1.51                       | 0.219        | 3.45 (A)                | 123                 | 2.36                       | 0.126        | 3.05 (A)                |
|                           | VIQ              | 105                 | <b>6.67</b>                | <b>0.010</b> | <b>4.59 (A)</b>         | 123                 | 0.21                       | 0.647        | 1.57 (A)                |
|                           | FSIQ             | 105                 | 4.57                       | 0.033        | 4.42 (A)                | 123                 | 1.04                       | 0.308        | 2.22 (A)                |
| rs324650<br>(136344201)   | PIQ              | 100                 | 2.51                       | 0.113        | 4.40 (T)                | 117                 | 2.69                       | 0.101        | 1.69 (T)                |
|                           | VIQ              | 100                 | <b>9.50</b>                | <b>0.002</b> | <b>5.30 (T)</b>         | 117                 | 0.00                       | 1.000        | 0.78 (T)                |
|                           | FSIQ             | 100                 | <b>7.14</b>                | <b>0.008</b> | <b>5.35 (T)</b>         | 117                 | 0.58                       | 0.446        | 0.77 (T)                |

*Note:* N denotes the number of individuals informative for the within family association test, i.e. those individuals that occur in families with more than one genotype. QTDT assumes equal genotypes for MZ twins and includes non-typed MZ co-twins with IQ scores

### *Correlations with CHRM2 tag-SNPs and CHRM2 transcripts expression at brain level*

Previous studies have shown that of the six known isoforms of *CHRM2* only C and D are expressed in the brain (Fenech *et al.*, 2004; Krejci *et al.*, 2004). In contrast to this, we observed all six *CHRM2* transcripts isoforms in brain material (data not shown). After normalizing raw gene expression data to expression level of the housekeeping genes, no correlation between gene expression and *CHRM2* gene genotypes for SNPs rs2061174, rs324640 or rs324650 was observed (data not shown).



**Figure 3.2** QTD family-based results for tag-SNPs plotted against FSIQ, VIQ, and PIQ for young (top panel) and adult (bottom panel) cohorts

## DISCUSSION

Converging evidence from previous studies (Comings *et al.* 2003; Dick *et al.* 2007; Gosso *et al.* 2006) has pointed to a role of the *CHRM2* gene in intelligence. None of these studies, however, have identified the functional polymorphism explaining its role at a molecular level. The present study aimed to zoom in on the functional variants already associated with human cognition, by fine-mapping the most significant areas within this gene and also investigating differential brain expression as a function of different genotypes on the SNPs most strongly related to intelligence.

A total of 42 SNPs within the *CHRM2* gene were genotyped, after which we determined the LD pattern among these SNPs in the *CHRM2* gene in a *young* and *adult* cohort. Association analysis was conducted separately among both age-related cohorts to detect possible age-dependent gene effects. Associations were found in different regions of the gene for each age cohort. Our current analyses showed that the same SNPs that were associated previously with intelligence, were again most significant, whereas a new SNP (rs2350780) showed a trend towards significance. Because of the dense coverage of SNPs used in this study, this confirms the importance of intron 4 and intron 5 regions, but rules out association with SNPs located elsewhere in the gene. Four new SNPs in the intron 3 region, (rs2350780, rs1364409, rs7782965, and rs1378646) showed association with PIQ in the *adult* cohort. These SNPs are in high LD ( $r^2$  between 0.58 – 0.72) between the most significant SNPs. SNP rs2350780 and rs2061174 were also found to be associated with intelligence by Dick and co-workers (Dick *et al.*, 2007). These intronic SNPs are located 68 kb apart in introns 3 and 4, respectively. In our cohort, LD between these two variants is 0.58. Therefore, it will be of great interest to try to define whether these two variants can be considered QTLs reflecting independent effects among different populations or they are indeed tagging the causal variant. A similar scenario occurs among the *young* cohort with rs324640, which is almost in complete LD ( $r^2 = 0.93$ ) with rs324650. This could be explained by assuming that more than one variant exists which in turn may be biologically important for IQ phenotypes outcomes at different developmental stages in life.

We found the most significant associations with PIQ in the adult cohort (rs2061174,  $\chi^2 = 9.14$ ;  $P = 0.003$ ) and with VIQ in the children cohort (rs323650,  $\chi^2 = 9.50$ ;  $P = 0.002$ ). Because only part of the variance in PIQ and VIQ is shared, and these results might reflect brain maturation processes and age-related genetic effects. Alternatively, the results could point to and potentially explain the genetic overlap between PIQ and VIQ, in which common genetic variants do not only interact modulating hippocampal neurotransmitter activity, but also and even more interesting, from the epigenetic point of view, they might modulate priming and dendritic outgrowth underlying synaptic plasticity during embryogenesis (Niculescu *et al.* 2006) during post-natal stages (Olivera *et al.* 2003), reflecting phenotypic variation at different IQ domains across the lifespan.

From a developmental perspective, brain maturation can be considered the most complex and dynamic lifelong process taking place in humans. Neuronal plasticity patterns (e.g. dendritic “pruning”, synapse elimination, myelination) have been shown to vary significantly across life and among diverse brain structures (for a review see Toga *et al.*, 2006). Variation in cognitive phenotypes may be the result of diverse allele-dependent effects that, although small in

effect size, may contribute to cognitive phenotype outcomes across life. Gene expression analyses showed that, in contrast to previously reported findings (Fenech *et al.* 2004; Krejci *et al.* 2004), all six currently known transcripts (isoforms A till F) of the *CHRM2* gene were present in brain tissue. Genotype-dependent *CHRM2* expression, however, did not reveal functional significance of any of the SNPs that were significantly related to intelligence.

*In situ* hybridization experiments on mammals (e.g. mice) (Lein *et al.* 2007) have been of great utility to aid specific localization and further interpretation of gene expression patterns, however, localization of *CHRM2* receptor transcripts have been conducted using probe sequences that did not distinguish between alternatively spliced transcripts. Because various cDNA isoforms may have the same annotation, there has been concern about the reproducibility of data, for example from different microarray platforms, for which conflicting results were obtained for the same transcript (Marshall 2004). Alternative splicing of 5'UTR regions, which may be of fundamental importance for regulating tissue-specific transcription, are not selectively tagged. Although brain expression analysis could not reveal differential expression of *CHRM2*, our study further zooms in on the *CHRM2* gene, clearly confirming two regions of most importance to intelligence on introns 4 and 5. These regions are poorly conserved regions among relatively distant species, they were highly conserved among primate species. Interestingly, the variant associated in the *young* cohort (rs324650) is located within a SINE repeat (MIRb). SINE repeats belongs to a wide family of transposable elements, which constitute the largest class of interspersed repeats that are found in our genome (12%) together with long interspersed repeats (LINE) and long terminal repeats (LTRs) (Smit 1999). SINEs repeats transpose through an RNA intermediate (reverse transcription process). All eukaryotic genomes contain mobile elements (retrotransposable elements), although the proportion and activity of the classes of elements varies widely between genomes (Kidwell & Lisch 2000). Nevertheless, the *CHRM2* gene, like its G-protein receptor counterparts share the interesting feature, at least from a functional perspective, of being an *intronless* gene (Fredriksson *et al.* 2003), which is also observed among dopamine receptors (O'Dowd 1993), widely studied in relation to attention deficits.

Recent research has revealed the potential functionality of retroposons in a gene-regulatory context (Bejerano *et al.* 2006; Han & Boeke 2005; Hellmann-Blumberg *et al.* 1993; Mei & Yan 1985; Muratani *et al.* 1991). It has been postulated that retroposon insertion processes may favour generation of intronless proteins (for a review see Flavell 1995 and Brosius 2003). If this hypothesis holds, the resulting intronless proteins are expected to contain exons among their 5'UTR. Not surprisingly, among G-proteins with intronless open reading frames (ORFs), about 18% have been reported the presence of



untranslated exons on their 5'UTR (O'Dowd 1993; Rouquier *et al.* 1998). The majority of mammalian GPCRs are related to central nervous system activity, which often requires high and differential expression of many genes (Lee & Irizarry 2003; Rouquier *et al.* 1998). Indeed, multiple promoters and transcripts have been reported for *CHRM2* gene suggesting that the associated regions we identified harbour functional elements involved in regulation of transcription and/or alternative splicing (Fenech *et al.* 2004; Krejci *et al.* 2004; Zhou *et al.* 2001). Further investigation involving functional assays and non-coding polymorphisms may aid the search and subsequent identification of regulatory variants underlying normal cognitive variation.

## REFERENCES

- Abecasis, G., Cardon, L. & Cookson, W. (2000) A general test of association for quantitative traits in nuclear families. *Am J Hum Genet* 66 279-292.
- Accili, D., Fishburn, C.S., Drago, J., Steiner, H., Lachowicz, J.E., Park, B.H., Gauda, E.B., Lee, E.J., Cool, M.H., Sibley, D.R., Gerfen, C.R., Westphal, H. & Fuchs, S. (1996) A targeted mutation of the D3 dopamine receptor gene is associated with hyperactivity in mice. *Proc Natl Acad Sci U S A* 93 1945-1949.
- Bartels, M., Rietveld, M.J., Van Baal, G.C. & Boomsma, D.I. (2002) Genetic and environmental influences on the development of intelligence. *Behav Genet* 32 237-249.
- Bejerano, G., Lowe, C.B., Ahituv, N., King, B., Siepel, A., Salama, S.R., Rubin, E.M., Kent, W.J. & Haussler, D. (2006) A distal enhancer and an ultraconserved exon are derived from a novel retroposon. *Nature* 441 87-90.
- Boomsma, D.I., de Geus, E.J., Vink, J.M., Stubbe, J.H., Distel, M.A., Hottenga, J.J., Posthuma, D., van Beijsterveldt, T.C., Hudziak, J.J., Bartels, M. & Willemsen, G. (2006) Netherlands Twin Register: from twins to twin families. *Twin Res Hum Genet* 9 849-857.
- Brosius, J. (2003) The contribution of RNAs and retroposition to evolutionary novelties. *Genetica* 118 99-116.
- Buyske, S., Bates, M.E., Gharani, N., Matise, T.C., Tischfield, J.A. & Manowitz, P. (2006) Cognitive traits link to human chromosomal regions. *Behav Genet* 36 65-76.
- Carey, G.J., Billard, W., Binch, H., 3rd, Cohen-Williams, M., Crosby, G., Grzelak, M., Guzik, H., Kozlowski, J.A., Lowe, D.B., Pond, A.J., Tedesco, R.P., Watkins, R.W. & Coffin, V.L. (2001) SCH 57790, a selective muscarinic M(2) receptor antagonist, releases acetylcholine and produces cognitive enhancement in laboratory animals. *Eur J Pharmacol* 431 189-200.
- Cherny, S.S. & Cardon, L.R. (1994) General cognitive ability. . In de Fries, J., Plomin, R. & Fulker, D. (eds), *Nature and nurture during middle childhood*. Blackwell Publishers, Oxford, pp. 46-56.
- Comings, D.E., Wu, S., Rostamkhani, M., McGue, M., Lacono, W.G., Cheng, L.S. & MacMurray, J.P. (2003) Role of the cholinergic muscarinic 2 receptor (CHRM2) gene in cognition. *Mol Psychiatry* 8 10-11.
- Dick, D.M., Aliev, F., Bierut, L., Goate, A., Rice, J., Hinrichs, A., Bertelsen, S., Wang, J.C., Dunn, G., Kuperman, S., Schuckit, M., Nurnberger, J., Jr., Porjesz, B., Beglieter, H., Kramer, J. & Hesselbrock, V. (2006) Linkage Analyses of IQ in the Collaborative Study on the Genetics of Alcoholism (COGA) Sample. *Behav Genet* 36 77-86.
- Dick, D.M., Aliev, F., Kramer, J., Wang, J.C., Hinrichs, A., Bertelsen, S., Kuperman, S., Schuckit, M., Nurnberger, J., Jr., Edenberg, H.J., Porjesz, B., Begleiter, H., Hesselbrock, V., Goate, A. & Bierut, L. (2007) Association of CHRM2 with IQ: Converging Evidence for a Gene Influencing Intelligence. *Behav Genet* 37 265-272.
- Fenech, A.G., Billington, C.K., Swan, C., Richards, S., Hunter, T., Ebejer, M.J., Felice, A.E., Ellul-Micallef, R. & Hall, I.P. (2004) Novel polymorphisms influencing transcription of the human CHRM2 gene in airway smooth muscle. *Am J Respir Cell Mol Biol* 30 678-686.
- Flavell, A.J. (1995) Retroelements, reverse transcriptase and evolution. *Comp Biochem Physiol B Biochem Mol Biol* 110 3-15.
- Fredriksson, R., Lagerstrom, M.C., Lundin, L.-G. & Schioth, H.B. (2003) The G-Protein-Coupled Receptors in the Human Genome Form Five Main Families. Phylogenetic Analysis, Paralogue Groups, and Fingerprints. *Mol Pharmacol* 63 1256-1272.
- Fulker, D.W., Cherny, S.S., Sham, P.C. & Hewitt, J.K. (1999) Combined linkage and association sib-pair analysis for quantitative traits. *Am J Hum Genet* 64 259-267.
- Gabriel, S.B., Schaffner, S.F., Nguyen, H., Moore, J.M., Roy, J., Blumenstiel, B., Higgins, J., DeFelice, M., Lochner, A., Faggart, M., Liu-Cordero, S.N., Rotimi, C., Adeyemo, A., Cooper, R., Ward, R., Lander, E.S., Daly, M.J. & Altshuler, D. (2002) The structure of haplotype blocks in the human genome. *Science* 296 2225-2229.
- Gautam, D., Duttaroy, A., Cui, Y., Han, S., Deng, C., Seeger, T., Alzheimer, C. & Wess, J. (2006) M1-M3 muscarinic acetylcholine receptor-deficient mice: novel phenotypes. *J Mol Neur* 30 157-160.
- Gosso, M.F., van Belzen, M.J., de Geus, E.J.C., Polderman, J.C., Heutink, P., Boomsma, D.I. & Posthuma, D. (2006a) Association between the CHRM2 gene and intelligence in a sample of 304 Dutch families. . *Genes, Brain and Behavior* 5 557-584.
- Han, J.S. & Boeke, J.D. (2005) LINE-1 retrotransposons: modulators of quantity and quality of mammalian gene expression? *Bioessays* 27 775-784.
- Hellmann-Blumberg, U., Hintz, M.F., Gatewood, J.M. & Schmid, C.W. (1993) Developmental differences in methylation of human Alu repeats. *Mol Cell Biol* 13 4523-4530.
- Hoekstra, C., Meijer, P., Kluft, C., Heutink, P., Smit, G., de Geus, E., Smit, J.H., van Bruggen, A., Montgomery, G.W. & Boomsma, D.I. (2004) Genetics of dizygotic twinning: a feasibility study for a biobank. *Twin Res* 7 556-563.
- Kidwell, M.G. & Lisch, D.R. (2000) Transposable elements and host genome evolution. *Trends In Ecology And Evolution* 15 95-99.
- Krejci, A., Bruce, A.W., Dolezal, V., Tucek, S. & Buckley, N.J. (2004) Multiple promoters drive tissue-specific expression of the human M muscarinic acetylcholine receptor gene. *J Neurochem* 91 88-98.
- Lee, C.J. & Irizarry, K. (2003) Alternative splicing in the nervous system: an emerging source of diversity and regulation. *Biological Psychiatry* 54 771.
- Lein, E.S. & Hawrylycz, M.J. & Ao, N. & Ayres, M. & Bensinger, A. & Bernard, A. & Boe, A.F. & Boguski, M.S. & Brockway, K.S. & Byrnes, E.J. & Chen, L. & Chen, L. & Chen, T.-M. & Chi Chin, M. & Chong, J. & Crook, B.E. & Czaplinska, A. & Dang, C.N. & Datta, S., *et al* (2007) Genome-wide atlas of gene expression in the adult mouse brain. *Nature* 445 168.

- Levey, A.I., Kitt, C.A., Simonds, W.F., Price, D.L. & Brann, M.R. (1991) Identification and localization of muscarinic acetylcholine receptor proteins in brain with subtype-specific antibodies. *J Neurosci* 11 3218-3226.
- Marshall, E. (2004) Getting the noise out of gene arrays. *Science* 306 630-631.
- Mei, L. & Yan, J.C. (1985) [Advances in subclassifications of muscarinic acetylcholine receptors]. *Sheng Li Ke Xue Jin Zhan* 16 248-252.
- Meulenbelt, I., Droog, S., Trommelen, G.J., Boomsma, D.I. & Slagboom, P.E. (1995) High-yield noninvasive human genomic DNA isolation method for genetic studies in geographically dispersed families and populations. *Am J Hum Genet* 57 1252-1254.
- Miller, S.A., Dykes, D.D. & Polesky, H.F. (1988) A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* 16 1215.
- Min, J.L., Lakenberg, N., Bakker-Verweij, M., Suchiman, E., Boomsma, D.I., Slagboom, P.E. & Meulenbelt, I. (2006) High microsatellite and SNP genotyping success rates established in a large number of genomic DNA samples extracted from mouth swabs and genotypes. *Twin Res Hum Genet* 9 501-506.
- Muratani, K., Hada, T., Yamamoto, Y., Kaneko, T., Shigeto, Y., Ohue, T., Furuyama, J. & Higashino, K. (1991) Inactivation of the cholinesterase gene by Alu insertion: possible mechanism for human gene transposition. *Proc Natl Acad Sci U S A* 88 11315-11319.
- Niculescu, M.D., Craciunescu, C.N. & Zeisel, S.H. (2006) Dietary choline deficiency alters global and gene-specific DNA methylation in the developing hippocampus of mouse fetal brains. *Faseb J* 20 43-49.
- O'Dowd, B.F. (1993) Structures of dopamine receptors. *J Neurochem* 60 804-816.
- Olivera, S., Henley, J.M. & Rodriguez-Ithurralde, D. (2003) AMPA receptor potentiation by acetylcholinesterase is age-dependently upregulated at synaptogenesis sites of the rat brain. *Int J Dev Neurosci* 21 49-61.
- Orsetti, M., Casamenti, F. & Pepeu, G. (1996) Enhanced acetylcholine release in the hippocampus and cortex during acquisition of an operant behavior. *Brain Res* 724 89-96.
- Plomin, R., Turic, D.M., Hill, L., Turic, D.E., Stephens, M., Williams, J., Owen, M.J. & O'Donovan, M.C. (2004) A functional polymorphism in the succinate-semialdehyde dehydrogenase (aldehyde dehydrogenase 5 family, member A1) gene is associated with cognitive ability. *Mol Psychiatry* 9 582-586.
- Polderman, T.J.C., Gosso, M.F., Posthuma, D., Van Beijsterveldt, T.C., Heutink, P., Verhulst, F.C. & Boomsma, D.I. (2006a) A longitudinal twin study on IQ, executive functioning, and attention problems during childhood and early adolescence. *Acta Neurol Belg* 106 191-207.
- Polderman, T.J.C., Stins, J.F., Posthuma, D., Gosso, M.F., Verhulst, F.C. & Boomsma, D.I. (2006b) The phenotypic and genotypic relation between working memory speed and capacity. *Intelligence* 34 549-560.
- Posthuma, D., de Geus, E.J. & Boomsma, D.I. (2001) Perceptual speed and IQ are associated through common genetic factors. *Behav Genet* 31 593-602.
- Posthuma, D., de Geus, E.J., Boomsma, D.I. & Neale, M.C. (2004) Combined linkage and association tests in mx. *Behav Genet* 34 179-196.
- Posthuma, D. & De Geus, E.J.C. (2006) Progress in the molecular-genetic study of intelligence. *Curr Dir Psych Sci* 15 151-155.
- Posthuma, D., Luciano, M., Geus, E.J., Wright, M.J., Slagboom, P.E., Montgomery, G.W., Boomsma, D.I. & Martin, N.G. (2005) A Genomewide Scan for Intelligence Identifies Quantitative Trait Loci on 2q and 6p. *Am J Hum Genet* 77 318-326.
- Rouquier, S., Taviaux, S., Trask, B.J., Brand-Arpon, V., van den Engh, G., Demaille, J. & Giorgi, D. (1998) Distribution of olfactory receptor genes in the human genome. *Nat Genet* 18 243-250.
- Seeger, T., Fedorova, I., Zheng, F., Miyakawa, T., Koustova, E., Gomez, J., Basile, A.S., Alzheimer, C. & Wess, J. (2004) M2 muscarinic acetylcholine receptor knock-out mice show deficits in behavioral flexibility, working memory, and hippocampal plasticity. *J Neurosci* 24 10117-10127.
- Smit, A.F. (1999) Interspersed repeats and other mementos of transposable elements in mammalian genomes. *Curr Opin Genet Dev* 9 657-663.
- Toga, A.W., Thompson, P.M. & Sowell, E.R. (2006) Mapping brain maturation. *Trends Neurosci* 29 148-159.
- Volpicelli, L.A. & Levey, A.I. (2004) Muscarinic acetylcholine receptor subtypes in cerebral cortex and hippocampus. *Prog Brain Res* 145 59-66.
- Wechsler, D. (1986) Wechsler Intelligence Scale for Children-Revised (WISC-R). Lisse: Swets & Zeitlinger.
- Wechsler, D. (1997) WAIS-III Wechsler Adult Intelligence Scale San Antonio, Texas: Psychological Corporation.
- Zhou, C., Fryer, A.D. & Jacoby, D.B. (2001) Structure of the human M(2) muscarinic acetylcholine receptor gene and its promoter. *Gene* 271 87-92.



# CHAPTER 4

The *SNAP-25* gene is associated  
with cognitive ability

This chapter is published as:

M.F. Gosso, E.J.C. de Geus, M.J. van Belzen, T.J.C. Polderman, P. Heutink, D.I. Boomsma, and D. Posthuma (2006). The *SNAP-25* gene is associated with cognitive ability: Evidence from a family based study in two independent Dutch cohorts. *Mol Psychiatry* 11: 878-886

*Abstract*

The synaptosomal associated protein of 25 kD (*SNAP-25*) gene plays an integral role in synaptic transmission, and is differentially expressed in the mammalian brain in the neocortex, hippocampus, anterior thalamic nuclei, substantia nigra, and cerebellar granular cells. Recent studies have suggested a possible involvement of *SNAP-25* in learning and memory, both of which are key components of human intelligence. In addition, the *SNAP-25* gene lies in a linkage area implicated previously in human intelligence. In two independent family based Dutch samples of 391 (mean age 12.4) and 276 (mean age 37.3) subjects respectively, we genotyped twelve SNPs in the *SNAP-25* gene on 20p12-20p11.2. From all individuals standardized intelligence measures were available. Using a family based association test, a strong association was found between three SNPs in the *SNAP-25* gene and intelligence, two of which showed association in both independent samples. The strongest, replicated association was found between SNP rs363050 and performance IQ, where the A allele was associated with an increase of 2.84 PIQ points ( $P=0.0002$ ). Variance in this single nucleotide polymorphism accounts for 3.4 % of the phenotypic variance in performance IQ (PIQ).

## INTRODUCTION

Intelligence is one of the most heritable traits in humans, with heritability estimates ranging from 25 to 40% in early childhood (Bartels *et al.* 2002) to 80% in adulthood (Posthuma *et al.* 2001). Recently the first genome-wide scan for intelligence was published, identifying two regions on chromosome 2q and 6p that showed significant linkage to intelligence, and several other regions showing suggestive linkage (4p, 7q, 20p, 21p) (Posthuma *et al.* 2005). Other scans followed shortly, replicating the 6p region, and also pointing to other regions (e.g. 14q) (Buyske *et al.* 2006; Dick *et al.* 2006; Luciano *et al.* 2006; Wainwright *et al.* 2006). An alternative approach to gene finding is to perform genetic association tests with candidate genes that are selected based on prior knowledge of biochemical functioning. We followed the latter approach and selected a putative candidate gene that was recently shown to be involved in learning and memory, which are two major components of intelligence. Several studies have demonstrated that the hippocampus plays a central role in learning and memory (Buyske *et al.* 2006; Dick *et al.* 2006; Luciano *et al.* 2006; Squire & Kandel 1999; Wainwright *et al.* 2006). Damage to the hippocampus selectively impairs the ability to learn and remember (Kim *et al.* 1993; Logue *et al.* 1997; Morris *et al.* 1982; Phillips & LeDoux 1995; Sandin *et al.* 2004; Selden *et al.* 1991; Squire & Kandel 1999; Sutherland *et al.* 1982). The synaptosomal associated protein of 25 kD (*SNAP-25*) gene lies in an area of previous suggestive linkage to intelligence (20p12 – p11.2) (Posthuma *et al.* 2005). and is highly expressed by neurons in the hippocampus (Frassoni *et al.* 2005; Geddes *et al.* 1990; Oyler *et al.* 1989). The *SNAP-25* gene product is a presynaptic plasma membrane protein which is an integral component of the vesicle docking and fusion machinery that regulates neurotransmitter release (Horikawa *et al.* 1993; Oyler *et al.* 1989; Seagar & Takahashi 1998). It is also implicated in axonal growth and synaptic plasticity (Osen-Sand *et al.* 1993). Three lines of evidence suggest a major role of *SNAP-25* in learning and memory in humans. Firstly, selective inhibition of *SNAP-25* expression prevents axonal elongation and the transformation of growth cones to synaptic terminals (Osen-Sand *et al.* 1993), especially in hippocampal neurons (Grosse *et al.* 1999). Such remodeling of nerve terminals in the adult brain may serve as a morphological substrate of learning and memory (Hou *et al.* 2004; Osen-Sand *et al.* 1993). Secondly, mRNA levels of *SNAP-25* are increased after the induction of long-term potentiation (LTP) in granule cells of the dentate gyrus (Roberts *et al.* 1998). Hippocampal LTP is thought to be a form of synaptic plasticity that underlies memory and learning (Bliss & Collingridge 1993; Martin *et al.* 2000; Morris 1989; Morris *et al.* 1986). Thirdly, inhibition of hippocampal *SNAP-25* leads to impaired long-term

contextual fear memory, spatial memory, as well as decreased LTP (Hou *et al.* 2004). The suggestive (according to the Lander and Kruglyak guidelines) (Lander & Kruglyak 1995) linkage finding of general intelligence (20p12 – p11.2) (Posthuma *et al.* 2005) to the area containing *SNAP-25* renders this gene a putative candidate gene for human intelligence.

The present study aims to investigate whether *SNAP-25* gene plays a role in human intelligence. To this end a family-based association approach is used in two independent cohorts of children (mean age 12.4) and adults (mean age 37.3).

## MATERIALS AND METHODS

### *Subjects*

All twins and their siblings were part of two larger cognitive studies and were recruited from the Netherlands Twin Registry (Boomsma 1998). Informed consent was obtained from the participants (Adult Cohort) or from their parents if they were under 18 (Young Cohort). The current study was approved by the institutional review board of the VU University Medical Center. None of the individuals tested suffered from severe physical or mental handicaps, as assessed through standard questionnaire.

### *Young Cohort*

The young cohort consisted of 177 twin pairs born between 1990 and 1992, and 55 siblings (Polderman *et al.* 2006). The twins were 12 (mean= 12.4, SD= 0.95) years of age and the siblings were between 8 and 15 years old at the time of testing. There were 41 monozygotic male twin pairs (MZM), 28 dizygotic male twin pairs (DZM), 56 monozygotic female twin pairs (MZF), 25 dizygotic female twin pairs (DZF), 27 dizygotic opposite-sex twin pairs (DOS), 28 male siblings and 27 female siblings. Participation in this study included a voluntary agreement to provide buccal swabs for DNA extraction.

### *Adult Cohort*

A total of 793 family members from 317 extended twin families participated in the adult cognition study (Posthuma *et al.* 2001). Participation in this study did not automatically include DNA collection, however, part of the sample (276 subjects) returned to the lab to provide blood for DNA extraction. Mean age was 37.3 years (SD= 12.50). There were 20 monozygotic male twin pairs (MZM), 11 dizygotic male twin pairs (DZM), 1 DZM triplet, 14 monozygotic female twin pairs (MZF), 22 dizygotic female twin pairs (DZF) and 17 dizygotic opposite-sex



twin pairs (DOS), 23 female siblings and 23 male siblings, and 59 nine subjects from incomplete twin pairs (18 males, 41 females).

### *Cognitive testing*

In the young cohort, cognitive ability was assessed with the Dutch adaptation of the Wechsler Intelligence Scale for Children-Revised (Wechsler 1986), and consisted of four verbal subtests (similarities, vocabulary, arithmetic, and digit span) and two performance subtests (block design, and object assembly). In the adult cohort, the Dutch adaptation of the Wechsler Adult Intelligence Scale III-Revised (Wechsler 1997), assessed IQ and consisted of four verbal subtests (information, similarities, vocabulary, and arithmetic) and four performance subtests (picture completion, block design, matrix reasoning, and digit-symbol substitution). In both cohorts, Verbal IQ (VIQ), performance IQ (PIQ) and full scale IQ (FSIQ) were normally distributed. Correlations between FSIQ/VIQ, FSIQ/PIQ, and PIQ/VIQ were 0.89, 0.81 and 0.45, respectively in the young cohort, and 0.90, 0.84, and 0.55 respectively in the adult cohort. Means and standard deviations of the full and genotyped cohorts are provided in Table 4.1.

**Table 4.1.** Means and standard deviations of Performance IQ (PIQ), Verbal IQ (VIQ) and Full Scale IQ (FSIQ) in the young and adult cohorts

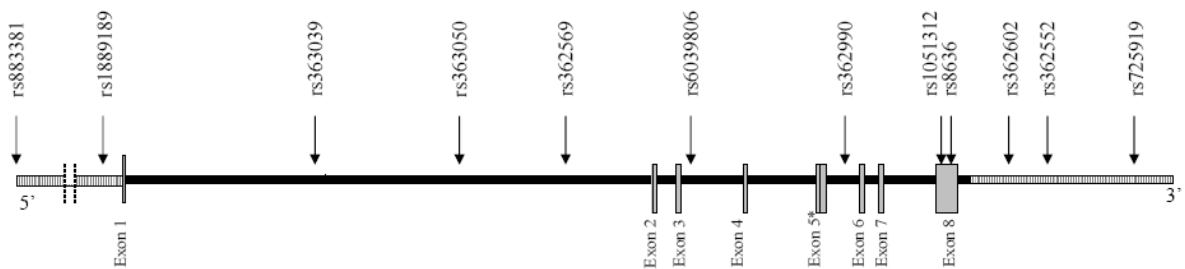
|                | Young Cohort        |                  | Adult Cohort        |                  |
|----------------|---------------------|------------------|---------------------|------------------|
|                | <i>Total sample</i> | <i>Genotyped</i> | <i>Total sample</i> | <i>Genotyped</i> |
| N              | 409                 | 391              | 793                 | 276              |
| Age (SD)       | 12.37 (0.95)        | 12.36 (0.90)     | 37.60 (13.00)       | 37.40 (12.42)    |
| Mean PIQ (SD)  | 101.40 (12.85)      | 101.66 (12.96)   | 100.96 (12.50)      | 100.04 (12.40)   |
| Mean VIQ(SD)   | 98.42 (19.04)       | 98.90 (19.02)    | 92.78 (13.83)       | 93.03 (14.36)    |
| Mean FSIQ (SD) | 99.81 (15.20)       | 100.21 (15.21)   | 95.74 (11.62)       | 95.59 (12.04)    |

### *DNA collection and Genotyping*

Buccal swabs were obtained from 391 children; blood was obtained from 276 adults. The DNA isolation from buccal swabs was performed using a chloroform/isopropanol extraction (Meulenbelt *et al.* 1995). DNA was extracted from blood samples using the salting out protocol (Miller *et al.* 1988).

Zygoty was assessed using 11 polymorphic microsatellite markers (Het > 0.80). Tagging single nucleotide polymorphisms (*tag*-SNPs) selection criteria were defined as SNPs with a minor allele frequency (MAF) above 0.10 and genotypic correlation ( $\rho$ ) across the genotypes of maximal 0.85 as obtained from a randomly selected

Caucasian sample ([http://www.celeradiagnostics.com/cdx/applera\\_genomics](http://www.celeradiagnostics.com/cdx/applera_genomics)). MAF had to be  $> 0.10$  in order to avoid the rare heterozygous genotypes and SNPs with a  $\rho$  above 0.85 with any of the other SNPs were not selected, to avoid redundancy. Twelve *tag*-SNPs in the *SNAP-25* gene were selected according to these criteria (<http://www.appliedbiosystems.com/support/software/snplex/>) using SNP Browser version 2.0.4, (NCBI build 34). Ranging from the 5'UTR to 3'UTR region within the *SNAP-25* gene, the following SNPs were selected: rs883381, rs1889189, rs363039, rs363050, rs362569, rs6039806, rs362990, rs1051312, rs8636, rs362602, rs362552, and rs725919 (see Figure 4.1). Genotyping was performed blind to familial status and phenotypic data. Both MZ twins of a pair were included in genotyping serving as additional controls.



**Figure 4.1** | Location of *tagging* single nucleotide polymorphisms (*tag*-SNPs) selected within the *SNAP-25* gene on chromosome 20 p12-p11.2

The SNplex assay was conducted following the manufacturer's recommendations (Applied Biosystems, Foster city, CA, USA). All pre-PCR steps were performed on a cooled block. Reactions were carried out in Gene Amp 9700 Thermocycler (Applied Biosystems, Foster city, CA, USA). PCR products were analyzed with ABI3730 Sequencer (Applied Biosystems, Foster city, CA, USA). Data was analyzed using Genemapper v3.7 (Applied Biosystems, Foster city, CA, USA).

### *Statistical analyses*

Allele frequencies of the twelve selected *tag*-SNPs were estimated in both cohorts using Pedstats (<http://www.sph.umich.edu/csg/abecasis/PedStats>) in which a Hardy-Weinberg test is implemented, based on an exact calculation of the probability of observing a certain number of heterozygotes conditional on the

number of copies of the minor SNP allele. MZ twins were considered as one genotype, when estimating allele frequencies.

Linkage disequilibrium (LD) parameters ( $D'$  and  $r^2$ ) were calculated from the haplotypes frequencies estimates using Haploview 3.2 (<http://www.broad.mit.edu/Mpg/haploview>).  $D'=1$  if, and only if, two SNPs have not been separated by recombination (or recurrent mutation). This LD parameter is sensitive to sample size, especially when SNPs with rare allele frequencies are considered. The value of  $r^2=1$  if, and only if, the SNPs have not been separated by recombination and have the same allele frequency. For quantifying and comparing LD in the context of mapping,  $r^2$  is slightly preferred (Ardlie *et al.* 2002). Values of  $r^2$  ranged from 0.001 to 0.680 in our sample, conforming relatively low LD between the separate *tag*-SNPs (see Table 4.2).

**Table 4.2** Estimates of linkage disequilibrium (LD) parameters  $r^2$  (lower) and  $D'$ (upper) for *tag*-SNPs within the *SNAP-25* gene

|           | rs883381 | rs1889189 | rs363039 | rs363050 | rs362569 | rs6039806 | rs362990 | rs1051312 | rs8636 | rs362602 | rs362552 | rs725919 |
|-----------|----------|-----------|----------|----------|----------|-----------|----------|-----------|--------|----------|----------|----------|
| rs883381  | -        | 0.885     | 0.495    | 0.491    | 0.012    | 0.014     | 0.144    | 0.154     | 0.021  | 0.428    | 0.149    | 0.133    |
| rs1889189 | 0.223    | -         | 1.000    | 0.467    | 0.641    | 0.491     | 0.428    | 0.173     | 0.116  | 0.342    | 0.268    | 0.326    |
| rs363039  | 0.211    | 0.242     | -        | 0.828    | 0.072    | 0.071     | 0.195    | 0.184     | 0.066  | 0.316    | 0.177    | 0.234    |
| rs363050  | 0.192    | 0.078     | 0.459    | -        | 0.276    | 0.146     | 0.059    | 0.171     | 0.318  | 0.265    | 0.017    | 0.135    |
| rs362569  | 0.000    | 0.124     | 0.004    | 0.039    | -        | 0.909     | 0.938    | 0.078     | 0.453  | 0.036    | 0.234    | 0.177    |
| rs6039806 | 0.000    | 0.098     | 0.002    | 0.015    | 0.630    | -         | 0.971    | 0.099     | 0.394  | 0.037    | 0.248    | 0.165    |
| rs362990  | 0.012    | 0.031     | 0.027    | 0.001    | 0.486    | 0.397     | -        | 0.226     | 0.536  | 0.122    | 0.230    | 0.187    |
| rs1051312 | 0.009    | 0.019     | 0.012    | 0.014    | 0.003    | 0.006     | 0.013    | -         | 0.095  | 0.710    | 0.900    | 0.917    |
| rs8636    | 0.000    | 0.010     | 0.004    | 0.074    | 0.078    | 0.079     | 0.062    | 0.002     | -      | 0.098    | 0.811    | 0.888    |
| rs362602  | 0.070    | 0.085     | 0.033    | 0.034    | 0.001    | 0.001     | 0.003    | 0.496     | 0.003  | -        | 1.000    | 0.936    |
| rs362552  | 0.015    | 0.013     | 0.024    | 0.000    | 0.032    | 0.028     | 0.048    | 0.240     | 0.155  | 0.258    | -        | 0.977    |
| rs725919  | 0.008    | 0.014     | 0.030    | 0.0070   | 0.013    | 0.009     | 0.027    | 0.176     | 0.134  | 0.158    | 0.682    | -        |

Haplotypes were estimated using SNPs that showed a significant association with IQ in both samples, using the Expectation-Maximization (EM) algorithm to obtain the maximum likelihood estimates of haplotype frequencies in each sample (Excoffier & Slatkin 1995), as implemented in the Allegro software package (Gudbjartsson *et al.* 2005). The EM algorithm allows for missing data and can be applied when no parental genotypes are available.

Genetic association tests were conducted using the program QTDT which implements the orthogonal association model proposed by Abecasis *et al.*, 2000 (see also Fulker *et al.*, 1999; extended by Posthuma *et al.*, 2004). This model allows the decomposition of the genotypic association effect into orthogonal between- ( $\beta_b$ ) and within- ( $\beta_w$ ) family components, can incorporate fixed effects of covariates and can also model the residual sib-correlation as a function of polygenic or environmental factors. MZ twins can be included and are modeled as such, by adding zygosity status to the datafile. They are not informative to the within-family association component (unless they are paired with non-twin siblings), but are informative for the between-family component. The between-family association component is sensitive to population admixture, whereas the within-family component is significant only in the presence of LD due to close linkage. If population stratification acts to create a false association, the test for association using the within-family component is still valid, and provides a conservative test of association. Testing for the equality of the  $\beta_b$  and  $\beta_w$  effects, serves as a test of population stratification. If this test is not significant, the between and within-family effects are equal and a total association test that uses the whole population at once can be applied. It should be noted, however, that given the relatively modest sample size, both the within-family test and the population stratification test are not as powerful as the 'total' association test. As we tested multiple SNPs, a significance level of 0.01 was kept.

## RESULTS

### *Single SNP analysis*

In total 391 subjects for the young cohort and 276 subjects for the adult cohort were available for SNP genotyping. Based on blind controls and MZ checks no genotyping errors were found. Eight SNPs out of the twelve selected were in Hardy-Weinberg equilibrium (HWE) in both cohorts. SNPs not in HWE (rs362990, rs6039806, rs362569 and rs1051312) were not included in further analyses. SNP rs883381 had a success rate of 80%, success in the young cohort; for all other SNPs in HWE, success rates were between 96.0 – 98.0% (see Table 4.3). The models used in QTDT included effects of age and sex on the means and modeled additive allelic between- and within-family effects. Residual sib correlations were modeled as a function of polygenic additive effects and non-shared environmental effects. Tests for the presence of population stratification were all non-significant indicating that genotypic effects within families were not

significantly different from those observed between families, suggesting that the more powerful total association test can be interpreted.

**Table 4.3** List of selected *tag*-SNP within the *SNAP-25* gene with their estimated heterozygosity Rates for the young/adult cohort

| Name             | Position        | Gene Location   | Obs HET          | Pred HET         | Succes Rate         | MAF              | HWE                          |
|------------------|-----------------|-----------------|------------------|------------------|---------------------|------------------|------------------------------|
| rs883381         | 10160727        | 5'UTR           | 0.41/0.54        | 0.47/0.46        | 80.10/94.80         | 0.38/0.37        | OK/OK                        |
| rs1889189        | 10192086        | 5'UTR           | 0.40/0.37        | 0.43/0.44        | 100.00/99.30        | 0.32/0.33        | OK/OK                        |
| rs363039         | 10215496        | intron 1        | 0.42/0.50        | 0.42/0.47        | 99.00/100.00        | 0.30/0.37        | OK/OK                        |
| rs363050         | 10229257        | intron 1        | 0.52/0.47        | 0.49/0.50        | 100.00/100.00       | 0.42/0.47        | OK/OK                        |
| <i>rs362569</i>  | <i>10241733</i> | <i>intron 1</i> | <i>0.37/0.50</i> | <i>0.46/0.48</i> | <i>90.70/91.00</i>  | <i>0.36/0.41</i> | <i>Not in HWE/OK</i>         |
| <i>rs6039806</i> | <i>10253654</i> | <i>intron 3</i> | <i>0.38/0.49</i> | <i>0.50/0.50</i> | <i>92.80/88.00</i>  | <i>0.46/0.47</i> | <i>Not in HWE/OK</i>         |
| <i>rs362990</i>  | <i>10271221</i> | <i>intron 5</i> | <i>0.30/0.41</i> | <i>0.39/0.40</i> | <i>100.00/96.30</i> | <i>0.27/0.27</i> | <i>Not in HWE/OK</i>         |
| <i>rs1051312</i> | <i>10282088</i> | <i>exon 8</i>   | <i>0.22/0.21</i> | <i>0.46/0.47</i> | <i>77.80/71.50</i>  | <i>0.37/0.38</i> | <i>Not in HWE/Not IN HWE</i> |
| rs8636           | 10282742        | exon 8          | 0.44/0.51        | 0.45/0.49        | 98.20/98.90         | 0.34/0.42        | OK/OK                        |
| rs362602         | 10288528        | 3'UTR           | 0.42/0.47        | 0.47/0.47        | 100.00/98.50        | 0.39/0.38        | OK/OK                        |
| rs362552         | 10291217        | 3'UTR           | 0.42/0.39        | 0.43/0.39        | 99.20/97.40         | 0.29/0.26        | OK/OK                        |
| rs725919         | 10298094        | 3'UTR           | 0.34/0.34        | 0.36/0.33        | 99.00/98.90         | 0.23/0.21        | OK/OK                        |

*Note:* *Tag*-SNPs were selected if allele frequency was >10% (18.0 % heterozygosity) and a genotypic correlation ( ) < 0.85. SNPs not in HWE are in italics

*Abbreviations:* **Obs HET** = Observed Heterozygosity. **Pred HET** = Predicted Heterozygosity. **MAF** = Minor Allele Frequency. **HWE** = Hardy Weinberg Equilibrium

Three SNPs (rs363039, rs363050, rs362602) showed significant associations with IQ. Two of these SNPs (rs363039, rs363050) were associated with IQ in both the young cohort and the independent adult cohort, showing association in the same direction and the same order of magnitude. The third SNP (rs362602) was seen as a trend to significant association only in the adult cohort. When we combined the two cohorts, the strongest association was seen between PIQ and rs363050, which is located on the 5'UTR of the *SNAP-25* gene ( $\chi^2=13.56$ ,  $P=0.0002$ ) The increaser allele of this SNP was associated with an increase of 2.84 IQ points (see Tables 4.4, 4.5 and Figure 4.2).

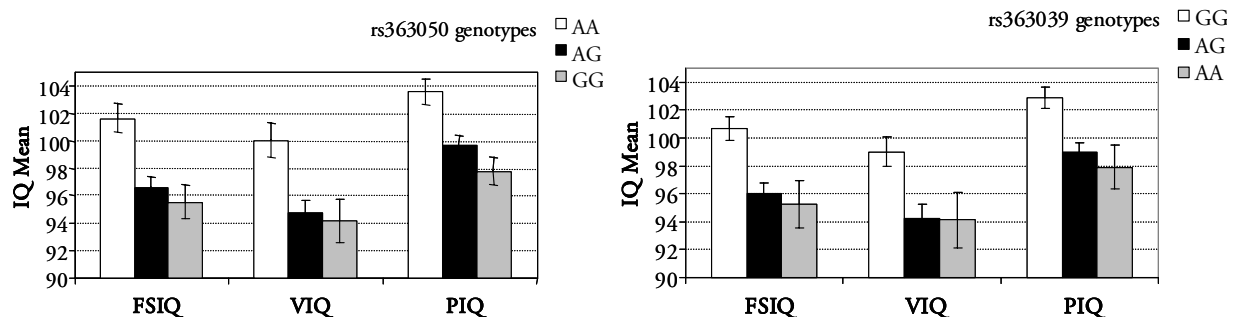
Within-family association tests are based on all siblings that are part of pairs with contrasting genotypes within a family and are thus less powerful than total association tests. The latter is preferred if there is no evidence of population stratification. It is, however, interesting to check whether the significant associations observed in the total association test are also present when looking only at the within family association. In the within-association test, for SNP rs363039 a trend was seen in both cohorts separately, while the G allele was suggestive of association ( $p < 0.05$ ) in the combined cohort. For SNP rs363050 the within

family association with the A allele was suggestive ( $p=0.06$ ) in the combined cohort. SNPs rs8636 and rs362602 were significant in the adult cohort ( $p < 0.01$ ) when only considering the within family test. These results support the results as found using the more powerful total association test.

### *Haplotype analysis*

The two SNPs that showed a significant association with IQ in both cohorts were 13 kB apart. Because these SNPs are in LD with each other ( $r^2= 0.46$ ) these SNPs were used to estimate haplotypes within each sample. Haplotype analysis of SNPs that are in LD with each other is more powerful than single SNP analysis because the combination of SNPs into a haplotype can be considered as a multiallelic marker that is more informative than a bi-allelic marker.

Non-significant SNPs were not used for further haplotype analysis, as all SNPs were selected on the basis of being tagging SNPs. From Table 4.2 it can be seen that indeed LD among the tag-SNPs flanking the two most significantly SNPs is very low, which would also be expected given the lack of association with these flanking SNPs and IQ.



**Figure 4.2** | IQ Means and Standard error for the means for Full scale IQ (FSIQ), Verbal IQ (VIQ) and Performance IQ (PIQ) for the combined cohort are plotted against the two most significant replicated *tag*-SNPs rs363050 (left) and rs363039 (right) genotypes

Haplotypes were estimated using the two SNPs rs363039 and rs363050 that were associated with psychometric IQ scores. Four possible haplotypes were *G-A*, *A-G*, *G-G*, and *A-A*, with haplotype frequencies 0.55, 0.29, 0.13, and 0.03 respectively in the young cohort and 0.54, 0.31, 0.13, and 0.03 respectively in the adult sample. Significant associations were found in both samples. When the data were combined, highly significant associations were observed with the *G-A* haplotype with FSIQ ( $\chi^2_{(1)}=11.14$ ,  $P= 0.0008$ ), VIQ ( $\chi^2_{(1)}=7.15$ ,  $P= 0.0074$ ) and

PIQ ( $\chi^2_{(1)}=10.61$ ,  $P= 0.0011$ ) (see Table 4.6). These results confirm the single SNP association results.

**Table 4.4** Family based association analysis for *SNAP-25* tagging SNPs for young, adult, and combined cohorts

| <i>tag-SNP</i>          |           | Young Cohort |             |                |                         | Adult Cohort |             |                |                         | Combined |              |                |                         |
|-------------------------|-----------|--------------|-------------|----------------|-------------------------|--------------|-------------|----------------|-------------------------|----------|--------------|----------------|-------------------------|
| position (bp)           | Phenotype | <i>N</i>     | $\chi^2$    | <i>P-value</i> | <i>Genotypic effect</i> | <i>N</i>     | $\chi^2$    | <i>P-value</i> | <i>Genotypic effect</i> | <i>N</i> | $\chi^2$     | <i>P-value</i> | <i>Genotypic effect</i> |
| rs883381<br>(10160727)  | PIQ       | 308          | 2.05        | 0.152          | 1.59 (G)                | 253          | 0.95        | 0.329          | 1.24 (G)                | 561      | 3.07         | 0.080          | 1.46 (G)                |
|                         | VIQ       | 306          | 0.00        | 0.981          | 0.04 (T)                | 254          | 2.24        | 0.134          | 2.00 (G)                | 560      | 0.55         | 0.457          | 0.80 (G)                |
|                         | FSIQ      | 308          | 0.30        | 0.584          | 0.70 (G)                | 252          | 1.88        | 0.170          | 1.62 (G)                | 560      | 1.50         | 0.221          | 1.09 (G)                |
| rs1889189<br>(10192086) | PIQ       | 385          | 0.07        | 0.788          | 0.30 (T)                | 268          | 1.13        | 0.287          | 1.22 (T)                | 653      | 0.63         | 0.427          | 0.63 (T)                |
|                         | VIQ       | 383          | 2.54        | 0.111          | 2.56 (C)                | 269          | 1.88        | 0.170          | 1.68 (T)                | 652      | 0.59         | 0.444          | 0.81 (C)                |
|                         | FSIQ      | 385          | 0.94        | 0.332          | 1.25 (T)                | 267          | 1.91        | 0.166          | 1.49 (T)                | 652      | 0.03         | 0.861          | 0.15 (C)                |
| rs363039<br>(10215496)  | PIQ       | 381          | <b>6.52</b> | <b>0.010</b>   | <b>2.99 (G)</b>         | 271          | 3.28        | 0.070          | 2.08 (G)                | 652      | <b>9.21</b>  | <b>0.002</b>   | <b>2.51 (G)</b>         |
|                         | VIQ       | 379          | 3.21        | 0.073          | 3.09 (G)                | 272          | 6.37        | 0.012          | 3.14 (G)                | 651      | <b>7.88</b>  | <b>0.005</b>   | <b>3.12 (G)</b>         |
|                         | FSIQ      | 381          | 5.83        | 0.016          | 3.34 (G)                | 270          | 6.18        | 0.013          | 2.68 (G)                | 651      | <b>10.88</b> | <b>0.001</b>   | <b>2.98 (G)</b>         |
| rs363050<br>(10229257)  | PIQ       | 385          | <b>7.88</b> | <b>0.005</b>   | <b>3.01 (A)</b>         | 267          | 5.44        | 0.020          | 2.55 (A)                | 652      | <b>13.56</b> | <b>0.0002</b>  | <b>2.84 (A)</b>         |
|                         | VIQ       | 383          | 1.94        | 0.164          | 2.21 (A)                | 268          | 6.03        | 0.014          | 2.89 (A)                | 651      | <b>5.91</b>  | <b>0.015</b>   | <b>2.52 (A)</b>         |
|                         | FSIQ      | 385          | 5.47        | 0.019          | 2.96 (A)                | 266          | <b>6.90</b> | <b>0.009</b>   | <b>2.69(A)</b>          | 651      | <b>11.48</b> | <b>0.0007</b>  | <b>2.86 (A)</b>         |
| rs8636<br>(10282742)    | PIQ       | 378          | 0.21        | 0.643          | 0.50 (T)                | 264          | 2.95        | 0.086          | 1.99 (C )               | 642      | 2.22         | 0.137          | 1.18 (T)                |
|                         | VIQ       | 376          | 0.30        | 0.587          | 0.86 (T)                | 265          | 1.92        | 0.166          | 1.71 (C)                | 641      | 0.07         | 0.785          | 0.29 (C)                |
|                         | FSIQ      | 378          | 0.00        | 0.978          | 0.03 (T)                | 261          | 3.39        | 0.065          | 1.99 (C )               | 641      | 1.08         | 0.299          | 0.90 (C)                |
| rs362602<br>(10288528)  | PIQ       | 385          | 0.33        | 0.567          | 0.59 (A)                | 265          | 3.01        | 0.082          | 1.97(G)                 | 650      | 0.30         | 0.583          | 0.42 (G)                |
|                         | VIQ       | 383          | 1.40        | 0.236          | 1.80 (A)                | 266          | <b>7.91</b> | <b>0.005</b>   | <b>3.49 (G)</b>         | 649      | 0.06         | 0.811          | 0.24 (G)                |
|                         | FSIQ      | 385          | 1.16        | 0.281          | 1.31 (A)                | 264          | <b>7.57</b> | <b>0.006</b>   | <b>2.96 (G)</b>         | 649      | 0.22         | 0.639          | 0.40 (G)                |
| rs362552<br>(10291217)  | PIQ       | 382          | 1.01        | 0.314          | 1.09 (G)                | 260          | 0.30        | 0.587          | 0.73 (A)                | 642      | 0.32         | 0.572          | 0.47 (G)                |
|                         | VIQ       | 380          | 0.31        | 0.575          | 0.90 (G)                | 261          | 1.20        | 0.273          | 1.59 (A)                | 641      | 0.01         | 0.924          | 0.10 (G)                |
|                         | FSIQ      | 382          | 0.67        | 0.415          | 1.04 (G)                | 259          | 1.01        | 0.315          | 1.27 (A)                | 641      | 0.04         | 0.832          | 0.19 (G)                |
| rs725919<br>(10298094)  | PIQ       | 381          | 1.87        | 0.172          | 1.63 (A)                | 267          | 0.45        | 0.505          | 0.94 (G)                | 648      | 0.60         | 0.439          | 0.70 (A)                |
|                         | VIQ       | 379          | 0.14        | 0.712          | 0.65 (A)                | 268          | 1.21        | 0.271          | 1.70 (G)                | 647      | 0.55         | 0.457          | 0.26 (A)                |
|                         | FSIQ      | 381          | 0.65        | 0.420          | 1.13 (A)                | 266          | 1.11        | 0.292          | 1.40 (G)                | 647      | 0.05         | 0.831          | 0.21 (A)                |

*Note:* The genotypic effect is the increase in IQ points associated with the increaser allele P-values below <.01 are in bold. Sex and age were included as covariates, residual variance was modelled as a function of polygenic effects and non-shared environmental effects

**Table 4.5** Means (SD) per genotype for PIQ, VIQ and FSIQ for young and adult cohorts in the four *tag*-SNPs within the *SNAP-25* gene that show association with a significant association

| <i>tag</i> -SNP      |                  | Young cohort       |                |                |         | Adult cohort       |               |                |         |
|----------------------|------------------|--------------------|----------------|----------------|---------|--------------------|---------------|----------------|---------|
| <i>position (bp)</i> | Phenotype        | Genotype Frequency |                |                | Total N | Genotype Frequency |               |                | Total N |
|                      |                  | <b>GG</b>          | <b>AG</b>      | <b>AA</b>      |         | <b>GG</b>          | <b>AG</b>     | <b>AA</b>      |         |
|                      | <i>Frequency</i> | <b>0.49</b>        | <b>0.42</b>    | <b>0.09</b>    |         | <b>0.38</b>        | <b>0.50</b>   | <b>0.12</b>    |         |
| rs363039             | PIQ (SD)         | 103.68 (13.72)     | 100.58 (12.30) | 97.06 (10.79)  | 381     | 101.57 (12.46)     | 96.89 (11.69) | 98.96 (13.71)  | 271     |
| (10168496)           | VIQ (SD)         | 101.13 (19.28)     | 97.22 (19.14)  | 97.79 (16.19)  | 379     | 95.53 (14.86)      | 90.62 (13.66) | 89.71 (14.15)  | 272     |
|                      | FSIQ (SD)        | 102.45 (16.09)     | 98.47 (14.63)  | 96.94 (13.54)  | 381     | 97.73 (12.66)      | 92.98 (11.28) | 93.21 (13.27)  | 270     |
|                      |                  | <b>AA</b>          | <b>AG</b>      | <b>GG</b>      |         | <b>AA</b>          | <b>AG</b>     | <b>GG</b>      |         |
|                      | <i>Frequency</i> | <b>0.32</b>        | <b>0.52</b>    | <b>0.16</b>    |         | <b>0.29</b>        | <b>0.47</b>   | <b>0.24</b>    |         |
| rs363050             | PIQ (SD)         | 104.26 (13.26)     | 101.11 (13.08) | 98.24 (11.29)  | 385     | 102.50 (13.58)     | 97.60 (11.95) | 97.34 (11.01)  | 267     |
| (10182257)           | VIQ (SD)         | 103.12 (18.82)     | 96.52 (19.08)  | 98.68 (17.92)  | 383     | 95.26 (15.65)      | 92.26 (13.05) | 89.20 (15.12)  | 268     |
|                      | FSIQ (SD)        | 103.98 (15.33)     | 98.36 (15.44)  | 98.03 (13.92)  | 385     | 98.08 (13.72)      | 93.98 (11.42) | 92.71 (11.45)  | 266     |
|                      |                  | <b>AA</b>          | <b>AG</b>      | <b>GG</b>      |         | <b>AA</b>          | <b>AG</b>     | <b>GG</b>      |         |
|                      | <i>Frequency</i> | <b>0.40</b>        | <b>0.42</b>    | <b>0.18</b>    |         | <b>0.39</b>        | <b>0.47</b>   | <b>0.14</b>    |         |
| rs362602             | PIQ (SD)         | 102.45 (11.74)     | 101.45 (13.11) | 100.44 (15.34) | 385     | 100.23 (12.45)     | 99.54 (12.82) | 100.11 (11.06) | 265     |
| (10241528)           | VIQ (SD)         | 99.97 (16.86)      | 99.50 (19.62)  | 95.82 (21.78)  | 383     | 92.45 (14.98)      | 91.11 (14.40) | 97.63 (12.40)  | 266     |
|                      | FSIQ (SD)        | 101.28 (12.81)     | 100.37 (15.77) | 97.10 (19.01)  | 385     | 94.91 (12.59)      | 94.08 (12.43) | 98.16 (10.99)  | 264     |

*Note:* N denotes the number of individuals. For the within family association test it denotes the number of individuals informative for the within family association, i.e. those individuals that occur in families with more than one genotype. The N for the between family association differs slightly from the added totals of Tables 4.4 as QTDT assumes equal genotypes for MZ twins and includes non-typed MZ co-twins with available IQ scores

**Table 4.6** Family based association analysis for *SNAP-25* *tagging* Haplotype for Young, Adult, and Combined cohorts

| Haplotype Frequency (Young/Adult) | Phenotype | Young N | $\chi^2$ | <i>P-value</i> | <i>Genotypic effect</i> | Adult N | $\chi^2$ | <i>P-value</i> | <i>Genotypic effect</i> | Combined N | $\chi^2$ | <i>P-value</i> | <i>Genotypic effect</i> |
|-----------------------------------|-----------|---------|----------|----------------|-------------------------|---------|----------|----------------|-------------------------|------------|----------|----------------|-------------------------|
| <i>G-A</i>                        | PIQ       | 380     | 6.53     | 0.011          | 2.34                    | 263     | 3.82     | 0.051          | 1.86                    | 643        | 10.61    | 0.001          | 2.18                    |
| 0.55 / 0.54                       | VIQ       | 378     | 2.98     | 0.084          | 2.33                    | 264     | 5.47     | 0.019          | 2.44                    | 642        | 7.15     | 0.007          | 2.41                    |
|                                   | FSIQ      | 380     | 5.91     | 0.015          | 2.64                    | 262     | 5.75     | 0.016          | 2.15                    | 642        | 11.14    | 0.001          | 2.44                    |
| <i>A-G</i>                        | PIQ       | 380     | 2.54     | 0.111          | -1.98                   | 263     | 4.19     | 0.041          | -2.50                   | 643        | 6.45     | 0.011          | -2.24                   |
| 0.29 / 0.31                       | VIQ       | 378     | 0.82     | 0.365          | -1.65                   | 264     | 5.63     | 0.018          | -3.17                   | 642        | 4.24     | 0.040          | -2.44                   |
|                                   | FSIQ      | 380     | 1.98     | 0.159          | -2.07                   | 262     | 6.56     | 0.010          | -2.96                   | 642        | 6.81     | 0.009          | -2.53                   |
| <i>G-G</i>                        | PIQ       | 380     | 1.23     | 0.267          | -1.48                   | 263     | 0.17     | 0.680          | -0.65                   | 643        | 1.49     | 0.222          | -1.24                   |
| 0.13 / 0.13                       | VIQ       | 378     | 0.16     | 0.689          | -0.79                   | 264     | 0.09     | 0.764          | -0.51                   | 642        | 0.23     | 0.632          | -0.66                   |
|                                   | FSIQ      | 380     | 0.78     | 0.377          | -1.32                   | 262     | 0.06     | 0.806          | -0.36                   | 642        | 0.78     | 0.377          | -0.98                   |
| <i>A-A</i>                        | PIQ       | 380     | 1.27     | 0.260          | -1.97                   | 263     | 0.07     | 0.791          | -0.65                   | 643        | 1.14     | 0.286          | -1.50                   |
| 0.03 / 0.03                       | VIQ       | 378     | 2.13     | 0.144          | -3.74                   | 264     | 0.58     | 0.446          | -1.96                   | 642        | 2.65     | 0.104          | -3.07                   |
|                                   | FSIQ      | 380     | 2.31     | 0.129          | -3.17                   | 262     | 0.42     | 0.517          | -1.46                   | 642        | 2.58     | 0.108          | -2.48                   |



## DISCUSSION

To investigate the possible role of the *SNAP-25* gene in intelligence, we employed a family based genetic association test in two independent cohorts of 391 children (mean age 12.4), and 276 adults (mean age 37.3). Replicated association was found in the two cohorts for two SNPs in the *SNAP-25* gene. Strongest evidence was found for SNP rs363050 in intron 1 at the 5'UTR, showing an effect size of 2.84 IQ points ( $P = 0.0002$ ) for the increaser allele. Haplotype analyses confirmed the region containing these two SNPs to be strongly associated with IQ. The *SNAP-25* gene, located on chromosome 20 p12-12p11.2 encodes a presynaptic terminal protein. In the mature brain, the *SNAP-25* gene product forms a complex with syntaxin and the synaptic vesicle proteins (synaptobrevin and synaptotagmin) that mediates exocytosis of neurotransmitter from the synaptic vesicle into the synaptic cleft (see Horikawa et al., 1993; Seagar et al., 1998, Bark et al., 1994; Low et al., 1999). During development *SNAP-25* is also involved in synaptogenesis, forming presynaptic sites and neuritic outgrowth (Osen-Sand *et al.* 1993; Oyler *et al.* 1989). *SNAP-25* is thought to be differentially expressed in the brain, and is primarily present in the neocortex, hippocampus, anterior thalamic nuclei, substantia nigra, and cerebellar granular cells. In the mature brain, expression is mainly seen at presynaptic terminals (Oyler *et al.* 1989).

*SNAP-25* exists in two splicing variants in relation to exon 5, *SNAP-25a* and *SNAP-25b*. Both isoforms differ in only 9 out of 39 amino acids encoded by the alternative spliced exons (Bark & Wilson 1991), resulting in a differentiated membrane anchoring relatively to cysteine residues involved in posttranscriptional fatty acylation (Hess *et al.* 1992). Both isoforms are thought to be equally important but at different time points for both neuronal maturation and neurotransmitter release (Bark *et al.* 1995; Bark & Wilson 1994; Grosse *et al.* 1999; Osen-Sand *et al.* 1993). Roberts and colleagues (1998) demonstrated that mRNA levels of both isoforms are elevated after induction of LTP, suggesting a role of *SNAP-25* in synaptic plasticity. A recent study involving antisense oligonucleotides against *SNAP-25* at the hippocampal CA1 region, reported the possible involvement of *SNAP-25* in learning and memory, particularly memory consolidation (Hou *et al.* 2004). Steffensen et al (1996) found that hippocampal LTP is attenuated in hemizygous mice from the *Coloboma* mice strain. The *Coloboma* mice strain is characterized by a 2-cM deletion on the mouse homolog of chromosome 2, in a region containing the mouse *SNAP-25* gene. Mice hemizygous for this deletion exhibit a wide spectrum of phenotypic and neurological abnormalities such as ophthalmic deformation, head bobbing, circling,

hyperactivity and small body size ( Snell & Bunker 1967; Hess *et al.* 1992; Heyser *et al.* 1995; Hess *et al.*, 1996). Because of the observed increased hyperactivity of hemizygous *Coloboma* mice, the role of *SNAP-25* in Attention Deficit Hyperactivity Disorder (ADHD) has been tested in several studies (Brophy *et al.* 2002; Feng *et al.* 2005; Kustanovich *et al.* 2003; Mill *et al.* 2002; Xu *et al.* 2005). All, except one (Xu *et al.*, 2005), report a significant association of *SNAP-25* with ADHD in humans. The exact role of *SNAP-25* in ADHD, however, remains unknown. ADHD is a neuropsychiatric condition characterized by hyperactive behavior and impaired attentive ability, resulting in both social and academic dysfunction. The present study suggests that involvement of *SNAP-25* may not be specific to the hyperactivity component of ADHD, but plays a more general role in learning and memory, through its effect on LTP and synaptic plasticity.

Both individual and haplotype analyses were conducted with two SNPs that showed significant association with intelligence in our study, tagging the 5'UTR region of *SNAP-25* gene. Genetic (non)coding variants lying within this non-coding region might be regulating this protein expression. These variants may influence regulatory binding sites, which in turn may modify gene expression and consequently neurotransmitter release regulation. Subtle changes in the fine-tuning at the neurotransmitter release machinery level, as well as in the interaction between neurotransmitter receptors subtypes might be manifest in substantial differences when LTP is being achieved. This complex fine-tuning may be reflected as individual differences in memory and learning, two fundamental aspects of human intelligence. Future functional studies will provide the insight needed in order to disentangle the complex interplay among *SNAP-25* gene (non)coding variants and cognitive ability.

## REFERENCES

- Abecasis, G.R., Cardon, L.R. & Cookson, W.O. (2000) A general test of association for quantitative traits in nuclear families. *Am J Hum Genet* 66 279-292.
- Ardlie, K.G., Kruglyak, L. & Seielstad, M. (2002) Patterns of linkage disequilibrium in the human genome. *NatRevGenet* 3 299-309.
- Bark, C. & Wilson, M.C. (1991) Alternative splicing generates a variant SNAP-25 protein during development. *Soc Neurosci Abstr* 17 531.
- Bark, I.C., Hahn, K.M., Ryabinin, A.E. & Wilson, M.C. (1995) Differential expression of SNAP-25 protein isoforms during divergent vesicle fusion events of neural development. *Proc Natl Acad Sci U S A* 92 1510-1514.
- Bark, I.C. & Wilson, M.C. (1994) Human cDNA clones encoding two different isoforms of the nerve terminal protein SNAP-25. *Gene* 139 291-292.
- Bartels, M., Rietveld, M., Van Baal, G.C.M. & Boomsma, D.I. (2002) Genetic and environmental influences on the development of intelligence. *Behav Genet* 32 237-249.
- Bliss, T.V. & Collingridge, G.L. (1993) A synaptic model of memory: long-term potentiation in the hippocampus. *Nature* 361 31-39.
- Boomsma, D.I. (1998) Twin registers in Europe: an overview. *Twin Res* 1 34-51.
- Brophy, K., Hawi, Z., Kirley, A., Fitzgerald, M. & Gill, M. (2002) Synaptosomal-associated protein 25 (SNAP-25) and attention deficit hyperactivity disorder (ADHD): evidence of linkage and association in the Irish population. *Mol Psychiatry* 7 913-917.
- Buyske, S., Bates, M.E., Gharani, N., Matise, T.C., Tischfield, J.A. & Manowitz, P. (2006) Cognitive traits link to human chromosomal regions. *Behav Genet* 36 65-76.
- Dick, D.M., Aliev, F., Bierut, L., Goate, A., Rice, J., Hinrichs, A., Bertelsen, S., Wang, J.C., Dunn, G., Kuperman, S., Schuckit, M., Nurnberger, J., Jr., Porjesz, B., Beglieter, H., Kramer, J. & Hesselbrock, V. (2006) Linkage Analyses of IQ in the Collaborative Study on the Genetics of Alcoholism (COGA) Sample. *Behav Genet* 36 77-86.
- Excoffier, L. & Slatkin, M. (1995) Maximum-likelihood estimation of molecular haplotype frequencies in a diploid population. *Mol Biol Evol* 12 921 - 927.
- Feng, Y., Crosbie, J., Wigg, K., Pathare, T., Ickowicz, A., Schachar, R., Tannock, R., Roberts, W., Malone, M., Swanson, J., Kennedy, J.L. & Barr, C.L. (2005) The SNAP25 gene as a susceptibility gene contributing to attention-deficit hyperactivity disorder. 10 998.
- Frassoni, C., Inverardi, F., Coco, S., Ortino, B., Grumelli, C., Pozzi, D., Verderio, C. & Matteoli, M. (2005) Analysis of SNAP-25 immunoreactivity in hippocampal inhibitory neurons during development in culture and in situ. *Neuroscience* 131 813.
- Fulker, D.W., Cherny, S.S., Sham, P.C. & Hewitt, J.K. (1999) Combined linkage and association sib-pair analysis for quantitative traits. *Am J Hum Genet* 64 259-267.
- Geddes, J.W., Hess, E.J., Hart, R.A., Kesslak, J.P., Cotman, C.W. & Wilson, M.C. (1990) Lesions of hippocampal circuitry define synaptosomal-associated protein-25 (SNAP-25) as a novel presynaptic marker. *Neuroscience* 38 515-525.
- Grosse, G., Grosse, J., Tapp, R., Kuchinke, J., Gorsleben, M., Fetter, I., Hohne-Zell, B., Gratzl, M. & Bergmann, M. (1999) SNAP-25 requirement for dendritic growth of hippocampal neurons. *J Neurosci Res* 56 539-546.
- Gudbjartsson, D.F., Thorvaldsson, T., Kong, A., Gunnarsson, G. & A, I. (2005) Allegro version 2. *Nature Genetics* 37 1015-1016.
- Hess, D.T., Slater, T.M., Wilson, M.C. & Skene, J.H. (1992) The 25 kDa synaptosomal-associated protein SNAP-25 is the major methionine-rich polypeptide in rapid axonal transport and a major substrate for palmitoylation in adult CNS. *J Neurosci* 12 4634-4641.
- Hess, E.J., Collins, K.A. & Wilson, M.C. (1996) Mouse model of hyperkinesis implicates SNAP-25 in behavioral regulation. *J Neurosci* 16 3104-3111.
- Heyser, C.J., Wilson, M.C. & Gold, L.H. (1995) Coloboma hyperactive mutant exhibits delayed neurobehavioral developmental milestones. *Brain Res Dev Brain Res* 89 264-269.
- Horikawa, H.P., Saisu, H., Ishizuka, T., Sekine, Y., Tsugita, A., Odani, S. & Abe, T. (1993) A complex of rab3A, SNAP-25, VAMP/syntaxin-2 and syntaxins in brain presynaptic terminals. *FEBS Lett* 330 236-240.
- Hou, Q., Gao, X., Zhang, X., Kong, L., Wang, X., Bian, W., Tu, Y., Jin, M., Zhao, G., Li, B., Jing, N. & Yu, L. (2004) SNAP-25 in hippocampal CA1 region is involved in memory consolidation. *Eur J Neurosci* 20 1593-1603.
- Kim, J.J., Rison, R.A. & Fanselow, M.S. (1993) Effects of amygdala, hippocampus, and periaqueductal gray lesions on short- and long-term contextual fear. *Behav Neurosci* 107 1093-1098.
- Kustanovich, V., Merriman, B., McGough, J., McCracken, J.T., Smalley, S.L. & Nelson, S.F. (2003) Biased paternal transmission of SNAP-25 risk alleles in attention-deficit hyperactivity disorder. *Mol Psychiatry* 8 309-315.
- Lander, E. & Kruglyak, L. (1995) Genetic dissection of complex traits: guideline for interpreting and reporting linkage results. *Nat Genet* 11 241 - 247.
- Logue, S.F., Paylor, R. & Wehner, J.M. (1997) Hippocampal lesions cause learning deficits in inbred mice in the Morris water maze and conditioned-fear task. *Behav Neurosci* 111 104-113.

- Low, P., Norlin, T., Risinger, C., Larhammar, D., Pieribone, V.A., Shupliakov, O. & Brodin, L. (1999) Inhibition of neurotransmitter release in the lamprey reticulospinal synapse by antibody-mediated disruption of SNAP-25 function. *Eur J Cell Biol* 78 787-793.
- Luciano, M., Wright, M.J., Duffy, D.L., Wainwright, M.A., Zhu, G., Evans, D.M., Geffen, G.M., Montgomery, G.W. & Martin, N.G. (2006) Genome-wide Scan of IQ Finds Significant Linkage to a Quantitative Trait Locus on 2q, *Behavior Genetics*.
- Martin, S.J., Grimwood, P.D. & Morris, R.G. (2000) Synaptic plasticity and memory: an evaluation of the hypothesis. *Annu Rev Neurosci* 23 649-711.
- Meulenbelt, I., Droog, S., Trommelen, G.J., Boomsma, D.I. & Slagboom, P.E. (1995) High-yield noninvasive human genomic DNA isolation method for genetic studies in geographically dispersed families and populations. *Am J Hum Genet* 57 1252-1254.
- Mill, J., Curran, S., Kent, L., Gould, A., Hockett, L., Richards, S., Taylor, E. & Asherson, P. (2002) Association study of a SNAP-25 microsatellite and attention deficit hyperactivity disorder. *Am J Med Genet* 114 269-271.
- Miller, S.A., Dykes, D.D. & Polesky, H.F. (1988) A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* 16 1215.
- Morris, R.G. (1989) Synaptic plasticity and learning: selective impairment of learning rats and blockade of long-term potentiation in vivo by the N-methyl-D-aspartate receptor antagonist AP5. *J Neurosci* 9 3040-3057.
- Morris, R.G., Anderson, E., Lynch, G.S. & Baudry, M. (1986) Selective impairment of learning and blockade of long-term potentiation by an N-methyl-D-aspartate receptor antagonist, AP5. *Nature* 319 774-776.
- Morris, R.G., Garrud, P., Rawlins, J.N. & O'Keefe, J. (1982) Place navigation impaired in rats with hippocampal lesions. *Nature* 297 681-683.
- Osen-Sand, A., Catsicas, M., Staple, J.K., Jones, K.A., Ayala, G., Knowles, J., Grenningloh, G. & Catsicas, S. (1993) Inhibition of axonal growth by SNAP-25 antisense oligonucleotides in vitro and in vivo. *Nature* 364 445-448.
- Oyler, G.A., Higgins, G.A., Hart, R.A., Battenberg, E., Billingsley, M., Bloom, F.E. & Wilson, M.C. (1989) The identification of a novel synaptosomal-associated protein, SNAP-25, differentially expressed by neuronal subpopulations. *J Cell Biol* 109 3039-3052.
- Phillips, R.G. & LeDoux, J.E. (1995) Lesions of the fornix but not the entorhinal or perirhinal cortex interfere with contextual fear conditioning. *J Neurosci* 15 5308-5315.
- Polderman, T.J.C., Stins, J.F., Posthuma, D., Gosso, M.F., Verhulst, F.C. & Boomsma, D.I. (2006) The phenotypic and genotypic relation between working memory speed and capacity, pp Submitted.
- Posthuma, D., de Geus, E.J. & Boomsma, D.I. (2001) Perceptual speed and IQ are associated through common genetic factors. *Behav Genet* 31 593-602.
- Posthuma, D., de Geus, E.J., Boomsma, D.I. & Neale, M.C. (2004) Combined linkage and association tests in mx. *Behav Genet* 34 179-196.
- Posthuma, D., Luciano, M., Geus, E.J., Wright, M.J., Slagboom, P.E., Montgomery, G.W., Boomsma, D.I. & Martin, N.G. (2005) A Genomewide Scan for Intelligence Identifies Quantitative Trait Loci on 2q and 6p. *Am J Hum Genet* 77 318-326.
- Roberts, L.A., Morris, B.J. & O'Shaughnessy, C.T. (1998) Involvement of two isoforms of SNAP-25 in the expression of long-term potentiation in the rat hippocampus. *Neuroreport* 9 33-36.
- Sandin, J., Ogren, S.O. & Terenius, L. (2004) Nociceptin/orphanin FQ modulates spatial learning via ORL-1 receptors in the dorsal hippocampus of the rat. *Brain Research* 997 222.
- Seagar, M. & Takahashi, M. (1998) Interactions between presynaptic calcium channels and proteins implicated in synaptic vesicle trafficking and exocytosis. *J Bioenerg Biomembr* 30 347-356.
- Selden, N.R., Everitt, B.J., Jarrard, L.E. & Robbins, T.W. (1991) Complementary roles for the amygdala and hippocampus in aversive conditioning to explicit and contextual cues. *Neuroscience* 42 335-350.
- Snell, G.D. & Bunker, H.P. (1967) Coloboma Linkage. *Mouse News Lett* 37 34.
- Squire, L.R. & Kandel, E.R. (1999) Memory: From Mind to Molecules. New York: Scientific American Library.
- Steffensen, S.C., Wilson, M.C. & Henriksen, S.J. (1996) Coloboma contiguous gene deletion encompassing Snap alters hippocampal plasticity. *Synapse* 22 281-289.
- Sutherland, R.J., Kolb, B. & Whishaw, I.Q. (1982) Spatial mapping: definitive disruption by hippocampal or medial frontal cortical damage in the rat. *Neurosci Lett* 31 271-276.
- Wainwright, M., Wright, M., Luciano, M., Montgomery, G., Geffen, G. & Martin, N. (2006) A Linkage Study of Academic Skills Defined by the Queensland Core Skills Test, *Behavior Genetics*.
- Wechsler, D. (1986) Wechsler Intelligence Scale for Children-Revised (WISC-R). Lisse: Swets & Zeitlinger.
- Wechsler, D. (1997) WAIS-III Wechsler Adult Intelligence Scale San Antonio, Texas: Psychological Corporation.
- Xu, X., Knight, J., Brookes, K., Mill, J., Sham, P., Craig, I., Taylor, E. & Asherson, P. (2005) DNA pooling analysis of 21 norepinephrine transporter gene SNPs with attention deficit hyperactivity disorder: no evidence for association. *Am J Med Genet B Neuropsychiatr Genet* 134 115-118.

# CHAPTER 5

## The *SNAP-25* gene: regulatory variants and cognition

This chapter is published as:

This chapter is published as: MF Gosso, EJC de Geus, TJC Polderman, DI Boomsma, P Heutink, and D Posthuma (2007) Common variants underlying cognitive ability: further evidence for association between the SNAP-25 gene and cognition using a family-based study in two independent Dutch cohorts. *Genes, Brain and Behavior* (*in press*)

*Abstract*

The *synaptosomal associated protein of 25 kD (SNAP-25)* gene, located on chromosome 20 p12-12p11.2 encodes a presynaptic terminal protein. *SNAP-25* is differentially expressed in the brain, and primarily present in the neocortex, hippocampus, anterior thalamic nuclei, substantia nigra, and cerebellar granular cells. Recently, a family-based genetic association was reported between variation in IQ phenotypes and two intronic variants on the *SNAP-25* gene. The present study is a follow-up association study in two Dutch cohorts of 371 children (mean age 12.4) and 391 adults (mean age 36.2). It examines the complete genomic region of the *SNAP-25* gene to narrow-down the location of causative genetic variant underlying the association. Two new variants on intron 1, (rs363043 and rs353016) close to the two previously reported variants (rs363039 and rs363050) showed association with variation in IQ phenotypes across both cohorts. All four SNPs were located on intron 1, within a region of about 13.8 kb, and are known to affect transcription factor binding sites. Although single SNP analyses showed similar allelic effects across cohorts different haplotypes were observed to be significantly associated with IQ variation in the young and the adult cohorts. Haplotype *G-T-T* was the strongest associated haplotype (PIQ  $\chi^2(1)= 9.36$   $P= 0.002$ ) in the *young* cohort; whereas *G-C-T* showed the strongest association among the *adult* cohort (FSIQ  $\chi^2(1)=10.08$   $P= 0.001$ ). Such heterogeneous genetic association might indicate that the causal variant is older than the SNPs that have been tested in this study and is present on both haplotypes.

## INTRODUCTION

Cognitive ability is currently considered as a polygenic trait influenced by many genes of moderate to small effect that in turn may interact with each other and with environmental factors (Butcher et al. 2006; Plomin & Spinath 2004; Savitz et al. 2006). Identifying the actual genes underlying normal cognitive variation has proven to be a daunting task, mainly due to this polygenic nature. So far, successful identification of genes underlying genetic variation in human cognitive ability has been mainly limited to mutations for relatively rare neurological disorders with considerably severe cognitive effects in which mental retardation or milder forms of cognitive disability are part of a syndromic phenotype (i.e. Fragile X Syndrome (Verkerk *et al.* 1991), Apert Syndrome (Ibrahimi *et al.* 2005), Rett Syndrome (Neul & Zoghbi 2004)). These mutations occur generally in key regulatory proteins within general neuronal signaling pathways.

Among diverse brain structures, the hippocampus is a critical part of the central nervous system (CNS) associated with learning and memory processes (Squire & Kandel 1999). Hippocampal Long Term Potentiation (LTP) is thought to be a form of synaptic plasticity that underlies memory and learning (Bliss & Collingridge 1993; Martin *et al.* 2000; Morris 1989; Morris *et al.* 1986). The *SNAP-25* gene product is a presynaptic plasma membrane protein which is an integral component of the vesicle docking and fusion machinery that regulates neurotransmitter release (Horikawa *et al.* 1993; Oyler *et al.* 1989; Seagar & Takahashi 1998). The *SNAP-25* gene is highly expressed by neurons in the hippocampus (Frassoni *et al.* 2005; Geddes *et al.* 1990; Oyler *et al.* 1989) and its expression has been significantly correlated with LTP formation (Hou *et al.* 2006). It is also implicated in axonal growth and synaptic plasticity (Osen-Sand *et al.* 1993). Current evidence suggests a major role of *SNAP-25* in learning and memory in humans, based on pharmacological, animal (Grosse et al. 1999; Hou et al. 2004; Osen-Sand et al. 1993) and human studies (Greber-Platzer *et al.* 2003; Thompson *et al.* 2003).

We recently conducted a family-based association study using an indirect (*tagging*) approach that involved the *SNAP-25* gene and psychometric intelligence scores as a measure of cognitive ability in humans (Gosso *et al.* 2006). Psychometric intelligence tests consist of a number of component subtests that taken together are used to infer a general IQ (intelligence quotient) score. Two single nucleotide polymorphisms (SNPs) in the *SNAP-25* gene showed a highly

significant association with IQ. Both were (non)coding variants. Associations in a (non)coding region of SNAP-25 can arise from variants in intronic and untranslated regions that influence gene expression (e.g. variants located on promoter regions, transcription starting sites, and 3'UTR miRNA target sites) which in turn might result in individual variation among IQ phenotypes.

The initial analyses (Gosso *et al.* 2006) were based on a tagging approach. We here perform follow-up analyses to (1) narrow-down the location of causative genetic variant underlying the association on intron 1, and (2) identify extra regions on *SNAP-25* gene not *tagged* during the previous analyses.

Two independent extended cohorts of children (mean age 12.4) and adults (mean age 36.2) were used in order to identify these putative regulatory genomic variants underlying variation among IQ phenotypes.

## MATERIALS AND METHODS

### *Subjects*

All twins and their siblings were part of two larger cognitive studies and were recruited from the Netherlands Twin Registry (Boomsma 1998; Boomsma *et al.* 2006). Informed consent was obtained from the participants (adult cohort) or from their parents if they were under 18 (young cohort). The current study was approved by the institutional review board of the VU University Medical Center. None of the individuals tested suffered from severe physical or mental handicaps, as assessed through standard questionnaire.

### *Young Cohort*

The young cohort consisted of 177 twin pairs born between 1990 and 1992, and 55 siblings (Polderman *et al.* 2006a; Polderman *et al.* 2006b), of which 371 were available for genotyping. The genotyped twins were 12.4 (SD= 0.9) years of age and the siblings were between 8 and 15 years old at the time of testing. There were 35 monozygotic male twin pairs (MZM), 28 dizygotic male twin pairs (DZM), 48 monozygotic female twin pairs (MZF), 23 dizygotic female twin pairs (DZF), 26 dizygotic opposite-sex twin pairs (DOS), 24 male siblings and 24 female siblings, and 3 subjects form incomplete twin pairs (1 male, 2 females). Participation in this study included a voluntary agreement to provide buccal swabs for DNA extraction.



This sample is similar to the sample used in our initial analyses, except for twenty individuals that were deleted from analyses in the current sample due to a more stringent threshold of genotyping failure per individual.

### *Adult Cohort*

A total of 793 family members from 317 extended twin families participated in the adult cognition study (Posthuma *et al.* 2005). Participation in this study did not automatically include DNA collection, however, part of the sample (276 subjects) returned to the lab to provide blood for DNA extraction or participated in NTR Biobank project (115 subjects) (Hoekstra *et al.* 2004). Mean age was 36.25 years (SD= 12.60). There were 25 monozygotic male twin pairs (MZM), 15 dizygotic male twin pairs (DZM), 1 DZM triplet, 20 monozygotic female twin pairs (MZF), 28 dizygotic female twin pairs (DZF) and 23 dizygotic opposite-sex twin pairs (DOS), 29 female siblings and 28 male siblings, and 109 subjects from incomplete twin pairs (41 males, 68 females).

### *Cognitive testing*

In the *young* cohort, cognitive ability was assessed with the Dutch adaptation of the Wechsler Intelligence Scale for Children-Revised (Wechsler 1986), and consisted of four verbal subtests (similarities, vocabulary, arithmetic, and digit span) and two performance subtests (block design, and object assembly).

In the *adult* cohort, the Dutch adaptation of the Wechsler Adult Intelligence Scale III-Revised (Wechsler 1997), assessed IQ and consisted of four verbal subtests (information, similarities, vocabulary, and arithmetic) and four performance subtests (picture completion, block design, matrix reasoning, and digit-symbol substitution).

In both cohorts, Verbal IQ (VIQ), performance IQ (PIQ) and full scale IQ (FSIQ) were normally distributed. Correlations between FSIQ/VIQ, FSIQ/PIQ, and PIQ/VIQ were 0.89, 0.81 and 0.45, respectively in the young cohort, and 0.90, 0.84, and 0.55 respectively in the adult cohort. Means and standard deviations of the full and genotyped cohorts are provided in Table 5.1.

**Table 5.1** Means and standard deviations of Performance IQ (PIQ), Verbal IQ (VIQ) and Full Scale IQ (FSIQ) in the young and adult cohorts.

|              | Young Cohort        |                  | Adult Cohort        |                  |
|--------------|---------------------|------------------|---------------------|------------------|
|              | <i>Total sample</i> | <i>Genotyped</i> | <i>Total sample</i> | <i>Genotyped</i> |
| N            | 407                 | 371              | 793                 | 391              |
| Gender (M/F) | 191/216             | 176/195          | 348/445             | 175/216          |
| Age (SD)     | 12.37 (0.93)        | 12.37 (0.92)     | 37.60 (13.00)       | 36.25 (12.64)    |
| SPIQ (SD)    | 94.57 (18.93)       | 94.85 (19.14)    | 104.49 (12.34)      | 104.30 (11.64)   |
| SVIQ(SD)     | 102.56 (12.74)      | 102.64 (12.92)   | 103.69 (12.26)      | 104.23 (12.15)   |
| SFSIQ (SD)   | 98.65 (15.06)       | 98.84 (15.24)    | 103.56 (11.49)      | 103.81 (11.16)   |

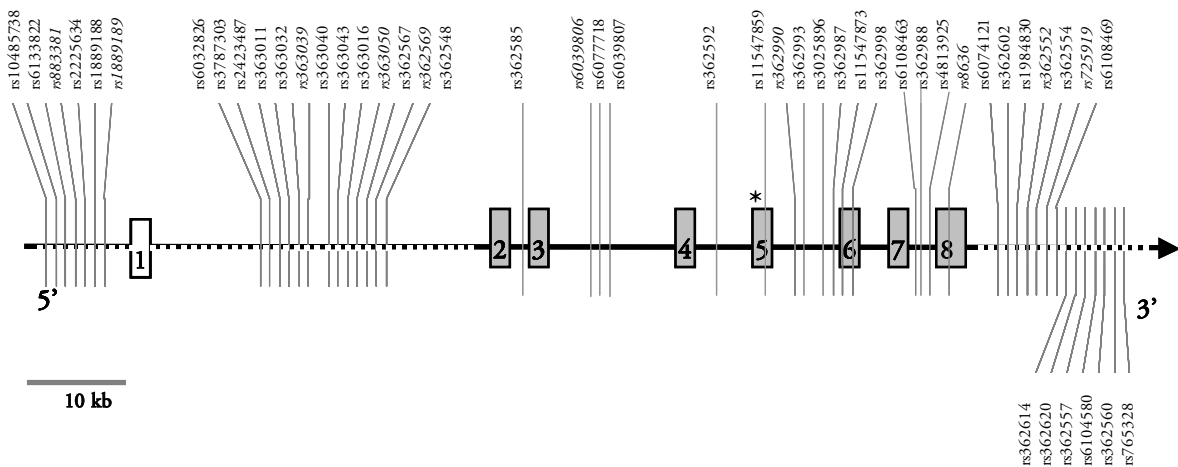
### *DNA collection and genotyping*

Buccal swabs were obtained from 371 children; DNA in adults was collected from blood samples (276 subjects) and buccal swabs (115 subjects). The DNA isolation from buccal swabs was performed using a chloroform/isopropanol extraction (Meulenbelt *et al.* 1995). DNA was extracted from blood samples using the salting out protocol (Miller *et al.* 1988).

Zygosity was assessed using 11 polymorphic microsatellite markers (Het > 0.80). Single nucleotide polymorphisms (SNPs) were selected based on their minor allele frequency (MAF) as obtained from a randomly selected population with northern and western European ancestry by the Centre d'Etude du polymorphisme Humain (CEPH) (<http://www.hapmap.org/thehapmap.html.en>). MAF had to be > 0.10 in order to reach enough power to detect common variants and also be able to observe all three possible bi-allelic combinations.

Forty-nine SNPs were selected using Haploview version 3.32 (<http://www.broad.mit.edu/mpg/haploview>) (NCBI build 36.1) to be genotyped in both cohorts (see Figure 5.1). Genotyping was performed blind to familial status and phenotypic data. Both MZ twins of a pair were included in genotyping and served as additional controls.

The SNPlex assay was conducted following the manufacturer's recommendations (Applied Biosystems, Foster city, CA, USA). All pre-PCR steps were performed on a cooled block. Reactions were carried out in Gene Amp 9700 Thermocycler (Applied Biosystems, Foster city, CA, USA). PCR products were analyzed with ABI3730 Sequencer (Applied Biosystems, Foster city, CA, USA). Data was analyzed using Genemapper v3.7 (Applied Biosystems, Foster city, CA, USA).



**Figure 5.1** | Location of the selected single nucleotide polymorphisms (SNPs) within the *SNAP-25* gene on chromosome 20 p12-p11.2. In *italics*, SNPs previously reported.

### Bioinformatics

Transcription factors (TFs) are proteins that recognize specific DNA sequences so-called transcription factor binding sites (TFBS), their interaction is fundamental for regulation of gene expression (for review see Garvie and Wolberg, 2001). Physical TF binding sites are constitutive DNA sequences found every 10 to 15 bps throughout the genome, and weight matrices are usually used to accurately predict them (Cartharius *et al.* 2005).

In order to identify gain or loss of physical transcription factor (TF) binding sites, SNP variation was analyzed using MatInspector and SNPInspector (both programs are available via Genomatix browser (<http://www.genomatix.de.html>)). While the former aids in the identification of physical TF binding sites, the latter identifies physical TF binding sites affected by SNPs (SNPs with putative regulatory activity). A random expectation value (the program assigns an expectation value for the number of transcription factor binding site matches per 1000 bp of random DNA sequence) is assigned to each transcription factor, as well as a percentage of vertebrate promoters containing the TF binding site.

### Statistical analyses

Allele frequencies of selected SNPs were estimated in both *young* and *adult* cohorts using Haploview v3.32 (<http://www.broad.mit.edu/mpg/haploview/>).

Hardy-weinberg equilibrium p-values were estimated for each variant, which is the probability that its deviation from H-W equilibrium could be explained by chance. Only one member of a twin pair was selected for HWE calculations. Linkage disequilibrium (LD) parameter ( $r^2$ ) was calculated from the haplotypes frequencies estimates using Haploview 3.32 ([http://www.broad.mit.edu/mpg/haplo view](http://www.broad.mit.edu/mpg/haplo_view)). Haplotypes were estimated using SNPs that showed a significant association with IQ in both samples, using the Expectation-Maximization (EM) algorithm to obtain the maximum likelihood estimates of haplotype frequencies in each sample (Excoffier & Slatkin 1995), as implemented in the Allegro software package version 2 (Gudbjartsson *et al.* 2005).

The EM algorithm allows for missing data and can be applied when no parental genotypes are available. Genetic association tests were conducted using the program QTDIT which implements the orthogonal association model proposed by Abecasis *et al.*, 2000 (see also Fulker *et al.*, 1999; extended by Posthuma *et al.*, 2004). This model allows the decomposition of the genotypic association effect into orthogonal between- ( $\beta_b$ ) and within- ( $\beta_w$ ) family components, can incorporate fixed effects of covariates and can also models the residual sib-correlation as a function of polygenic or environmental factors.

MZ twins can be included and are modelled as such, by adding zygosity status to the datafile. They are not informative to the within-family association component (unless they are paired with non-twin siblings), but are informative for the between family component. The between-family association component is sensitive to population admixture, whereas the within-family component is significant only in the presence of LD due to close linkage.

Testing for the equality of the  $\beta_b$  and  $\beta_w$  effects, serves as a test of population stratification. If population stratification acts to create a false association, the test for association using the within family component is still valid, and provides a conservative test of association. If this test is not significant, the between and within family effects are equal and total association test that uses the whole population at once can be applied. It should be noted, however, that given the relatively modest sample size, both the within family test and the population stratification test are not as powerful as the 'total' association test. As we tested multiple SNPs, a significance level of 0.01 was kept.

## RESULTS

### *Single SNP analysis*

Genotyping of 7 out of the 49 SNPs failed in both cohorts (rs1889188, rs2423487, rs363040, rs362548, rs11547873, rs11547859, and rs3025896). The LD structure for the remaining 42 SNPs is given in Table 5.2 for the young and adult cohort separately (see also Figure 5.2). Three out of the remaining 42 SNPs were not in Hardy-Weinberg equilibrium (HWE) in both cohorts. Further analyses will focus on the 39 variants that were in HWE). LD values for the young and adult cohort are given in Table 5.3. QTDT modelled additive allelic between- and within-family effects. Residual sib correlations were modelled as a function of polygenic additive effects and non-shared environmental effects.

Tests for the presence of population stratification were all non-significant indicating that genotypic effects within families were not significantly different from those observed between families, suggesting that the more powerful total association test can be interpreted. Four SNPs (rs363039, rs363043, rs363016, and rs363050) located on intron 1, showed significant association with IQ phenotypes.

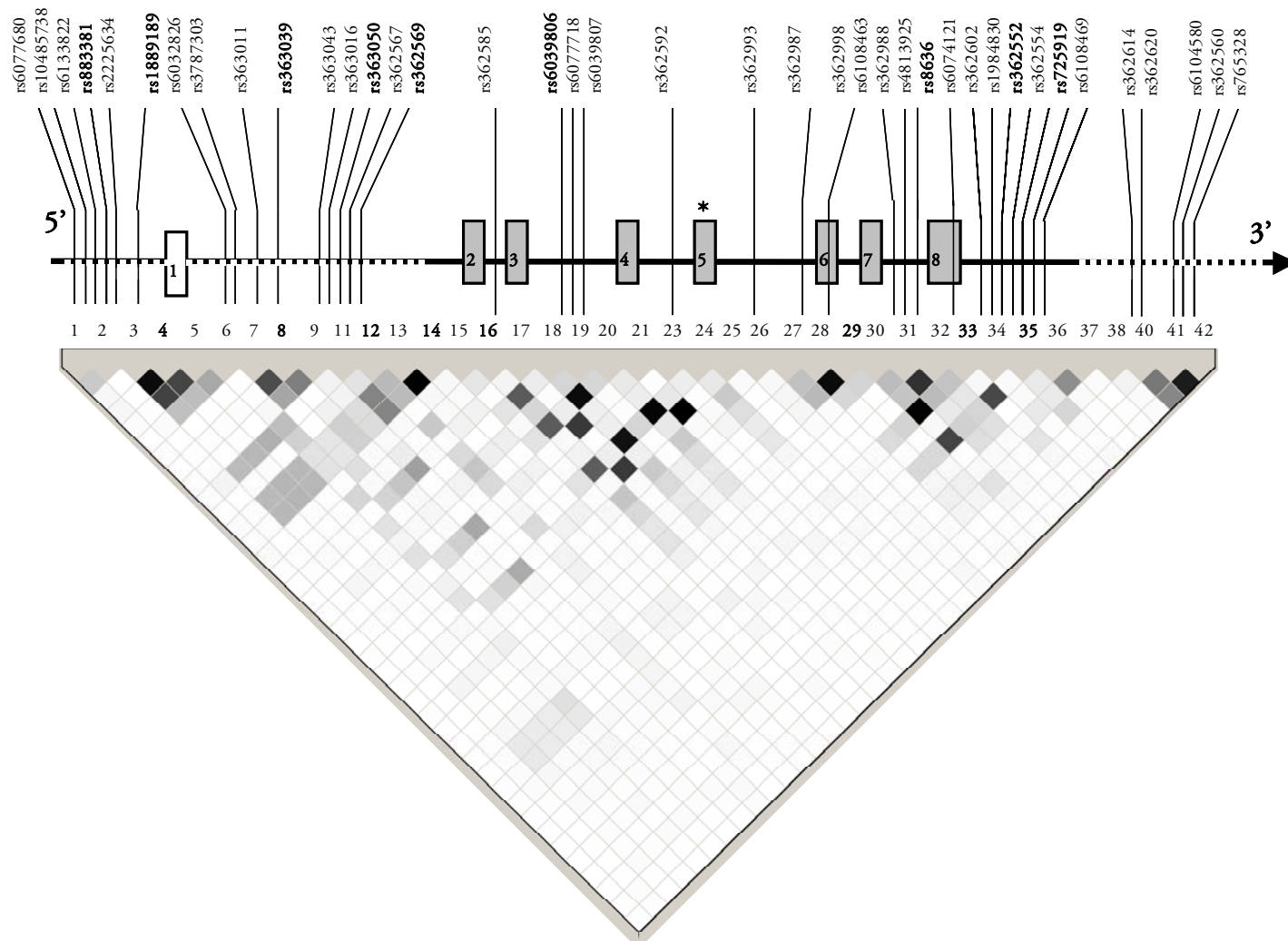
Significance was strongest in the young cohort ( $p < 0.01$ ) and showed a trend in the same direction in the older cohort ( $p < 0.10$ ) (see Tables 5.4 and 5.5). Analyses of the combined sample resulted in highly significant associations for these four SNPs. Two of these SNPs (rs363039, rs363050) were previously associated with IQ variation in these same young and adult cohorts (Gosso *et al.* 2006), whereas two new SNPs that were found significant, were added in the current follow-up analyses (rs363043, rs363016). All four SNPs showed association in the same direction and the same order of magnitude. As can be observed from table 5.3, rs363050 and rs363016 are in complete LD ( $r^2 = 0.98$ ), and as expected, the similarity of association results reflects the high LD between them.

The strongest association among this intron 1 region was observed between rs363016 and FSIQ ( $\chi^2 = 15.99$ ,  $P = 0.0001$ ). The increaser allele of this SNP was associated with an increase of 3.28 IQ points (see Table 5.5, and Figure 5.3). Subsequently, further haplotype analyses were conducted with only three variants out of the four significantly associated SNPs (rs363039, rs363043, and rs363016) (see *haplotype analysis* section below).

**Table 5.2** SNP descriptives for *young* and *adult* cohorts.

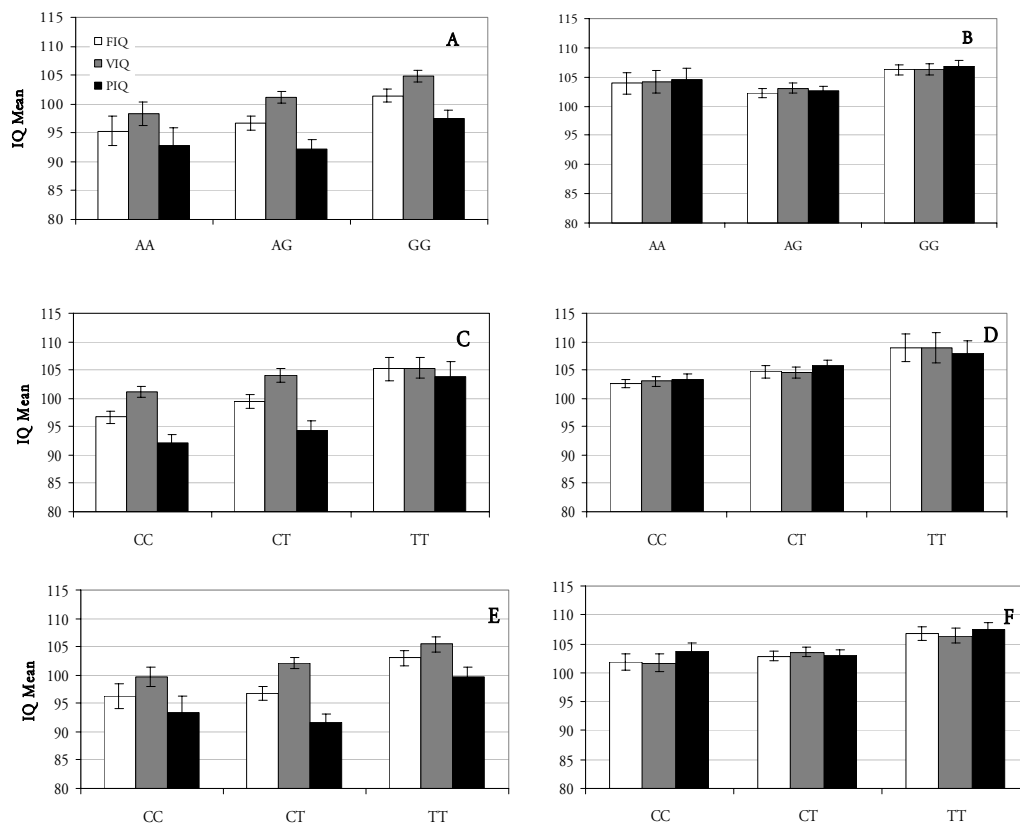
|    | SNP              | Position <sup>a</sup> | Tagged-SNP <sup>b</sup> | LD Block <sup>c</sup> | M.A. | HWE       | MAF       | Genotyping Success Rate |
|----|------------------|-----------------------|-------------------------|-----------------------|------|-----------|-----------|-------------------------|
| 1  | rs6077680        | 10100059              |                         |                       | A    | 0.65/0.84 | 0.35/0.42 | 0.94/0.86               |
| 2  | rs10485738       | 10100526              |                         |                       | G    | 0.70/0.41 | 0.29/0.34 | 0.95/0.83               |
| 3  | rs6133822        | 10107081              |                         | Block1 (6 kb)         | G    | 0.88/0.20 | 0.37/0.37 | 0.95/0.90               |
| 4  | <b>rs883381</b>  | 10113727              | 3                       |                       | T    | 0.71/0.23 | 0.38/0.37 | 0.82/0.93               |
| 5  | rs2225634        | 10126022              |                         | Block 2 (19 kb)       | T    | 0.89/0.99 | 0.44/0.38 | 0.98/0.93               |
| 6  | <b>rs1889189</b> | 10145086              | 12, 13                  |                       | T    | 0.51/0.51 | 0.33/0.37 | 0.99/0.99               |
| 7  | rs6032826        | 10151817              |                         | Block3 (13kb)         | G    | 0.82/0.83 | 0.24/0.23 | 0.83/0.85               |
| 8  | rs3787303        | 10156748              |                         |                       | C    | 0.98/0.85 | 0.17/0.17 | 0.94/0.91               |
| 9  | rs363011         | 10165336              | 10                      |                       | C    | 0.46/0.61 | 0.12/0.10 | 0.97/0.70               |
| 10 | rs363032         | 10166644              |                         |                       | C    | 0.27/0.00 | 0.06/0.07 | 0.98/0.93               |
| 11 | <b>rs363039</b>  | 10168496              |                         |                       | A    | 0.90/0.08 | 0.29/0.36 | 0.98/0.99               |
| 12 | rs363043         | 10174146              |                         | Block4 (8 kb)         | T    | 0.91/0.89 | 0.31/0.25 | 0.93/0.93               |
| 13 | rs363016         | 10179174              |                         |                       | C    | 0.35/0.96 | 0.41/0.43 | 0.92/0.93               |
| 14 | <b>rs363050</b>  | 10182257              | 12,13                   |                       | G    | 0.83/0.99 | 0.41/0.43 | 0.98/0.99               |
| 15 | rs362567         | 10194091              |                         | Block5 (31kb)         | A    | 0.07/0.97 | 0.13/0.11 | 0.97/0.96               |
| 16 | <b>rs362569</b>  | 10194733              |                         |                       | C    | 0.05/0.64 | 0.36/0.39 | 0.89/0.92               |
| 17 | rs362585         | 10204252              |                         |                       | T    | 0.30/0.30 | 0.07/0.10 | 0.92/0.94               |
| 18 | <b>rs6039806</b> | 10206654              | 17, 20, 24              |                       | A    | 0.06/0.95 | 0.47/0.47 | 0.93/0.89               |
| 19 | rs6077718        | 10209142              | 28                      |                       | G    | 0.83/0.80 | 0.12/0.14 | 0.97/0.96               |
| 20 | rs6039807        | 10211576              |                         |                       | G    | 0.07/0.92 | 0.48/0.49 | 0.93/0.92               |
| 21 | rs362592         | 10220492              |                         |                       | A    | 0.98/0.08 | 0.08/0.08 | 0.94/0.90               |
| 22 | <b>rs362990</b>  | 10224221              |                         |                       | T    | 0.01/0.92 | 0.26/0.30 | 0.98/0.97               |
| 23 | rs362993         | 10224716              |                         |                       | T    | 0.58/0.05 | 0.07/0.06 | 0.91/0.92               |
| 24 | rs362987         | 10225452              |                         |                       | C    | 0.21/0.84 | 0.46/0.48 | 0.87/0.84               |
| 25 | rs362998         | 10225621              | 17, 21                  |                       | T    | 0.99/0.57 | 0.08/0.09 | 0.89/0.84               |
| 26 | rs6108463        | 10228505              |                         |                       | C    | 0.42/0.98 | 0.17/0.18 | 0.98/0.95               |
| 27 | rs362988         | 10229370              |                         |                       | A    | 0.20/0.60 | 0.49/0.41 | 0.94/0.88               |
| 28 | rs4813925        | 10234313              |                         | Block6 (1kb)          | A    | 0.46/0.49 | 0.30/0.43 | 0.85/0.87               |
| 29 | <b>rs8636</b>    | 10235742              | 28                      |                       | T    | 0.78/0.71 | 0.34/0.41 | 0.97/0.99               |
| 30 | rs6074121        | 10238703              |                         | Block7 (12 kb)        | C    | 0.80/0.99 | 0.31/0.27 | 0.98/0.93               |
| 31 | rs362602         | 10241528              |                         |                       | G    | 0.26/1.00 | 0.38/0.38 | 0.99/0.97               |
| 32 | rs1984830        | 10242596              |                         |                       | G    | 0.50/1.00 | 0.34/0.36 | 0.97/0.94               |
| 33 | <b>rs362552</b>  | 10244217              | 30                      |                       | G    | 0.84/1.00 | 0.31/0.28 | 0.98/0.99               |
| 34 | rs362554         | 10248961              |                         |                       | T    | 0.73/0.93 | 0.05/0.07 | 0.93/0.93               |
| 35 | <b>rs725919</b>  | 10251094              |                         |                       | A    | 0.44/0.82 | 0.24/0.24 | 0.98/0.99               |
| 36 | rs6108469        | 10252374              |                         |                       | G    | 0.84/0.95 | 0.29/0.26 | 0.94/0.92               |
| 37 | rs362614         | 10258449              |                         |                       | T    | 0.84/0.99 | 0.23/0.22 | 0.89/0.87               |
| 38 | rs362620         | 10262383              |                         | Block8 ( kb)          | A    | 0.18/0.18 | 0.20/0.16 | 0.91/0.94               |
| 39 | rs362557         | 10263041              | 40                      |                       | G    | 0.88/0.03 | 0.43/0.47 | 0.98/0.96               |
| 40 | rs6104580        | 10267359              |                         | Block9 ( kb)          | T    | 0.63/0.21 | 0.22/0.21 | 0.94/0.95               |
| 41 | rs362560         | 10277347              | 40, 41                  |                       | C    | 0.62/0.85 | 0.33/0.31 | 0.98/0.93               |
| 42 | rs765328         | 10281596              |                         |                       | G    | 0.74/0.78 | 0.34/0.33 | 0.91/0.96               |

*Note:* SNPs were selected if allele frequency was >10% (18.0 % heterozygosity) and a genotypic correlation ( ) < 0.85. SNPs already reported in our previous study (Gosso *et al.*, 2006) are in **bold**. <sup>a</sup> Chromosomal location in base pairs based on Build 36.1. <sup>b</sup> SNP tagged with selection criteria. <sup>c</sup> LD block defined using the algorithm proposed by Gabriel *et al.*, 2000, based in D'. **MA/MAF** = Minor Allele/Frequency, **HWE** = Hardy Weinberg Equilibrium test (p-values).



**Figure 5.2** | Location of the selected single nucleotide polymorphisms (SNPs) within the *SNAP-25* gene on chromosome 20 p12-p11.2 and their LD pattern ( $r^2$ )  
In bold SNPs previously reported. \*Denotes splicing site on exon 5

A few other significant p-values (i.e.  $\leq 0.01$ ) were observed that may suggest a second and new association peak located within a region of 4.9 kb in the 3'UTR of *SNAP-25*. This involved two untranslated variants located only 658 base pairs apart (rs362620 and rs362557) in the young cohort. The A allele of rs362620 was associated with an increase of 1.29 VIQ points ( $\chi^2=6.56$ ,  $P=0.01$ ), whereas the C allele of rs362557 was associated with an increase of 2.84 in VIQ ( $\chi^2=7.02$ ,  $P=0.008$ ). LD patterns between these two SNP were extremely low ( $r^2=0.18$ ). In addition, within the same region, association results for the adult cohort were observed with a neighbour variant (rs6104580) located 4.32 kb away from the variants associated on the young cohort. In the old cohort the T allele of rs6104580 was associated with an increase of 3.26 PIQ points ( $\chi^2=8.36$ ,  $P=0.004$ ). LD between the rs6104580 and the two SNPs associated in the young cohort (rs362620 and rs362557) was relatively low ( $r^2$  between 0.04 and 0.34) and variants rs363039, rs363043, and rs363016 were considered for subsequent haplotype analysis (see *haplotype analysis* section below).



**Figure 5.3** | IQ means for *young* and *adult* cohorts plotted against the genotypes for rs363039 for *young* (A) and *adult* (B), rs363043 for *young* (C) and *adult* (D), and rs363016 for *young* (E) and *adult* (F), respectively.



**Table 5.4** Means (SD) per genotype for PIQ, VIQ and FIQ for *young* and *adult* cohorts for the four *tag*-SNPs within the *SNAP-25* gene that show association with a significant association on intron 1.

| SNP<br>position (bp)   | Phenotype | Young Cohort<br>Genotype Frequency |                          |                          | Total<br>N | Adult Cohort<br>Genotype Frequency |                          |                          | Total<br>N |
|------------------------|-----------|------------------------------------|--------------------------|--------------------------|------------|------------------------------------|--------------------------|--------------------------|------------|
|                        |           | <b>AA</b><br><i>0.09</i>           | <b>AG</b><br><i>0.43</i> | <b>GG</b><br><i>0.48</i> |            | <b>AA</b><br><i>0.09</i>           | <b>AG</b><br><i>0.51</i> | <b>GG</b><br><i>0.40</i> |            |
| rs363039<br>(10168496) | FIQ (SD)  | 95.29 (14.35)                      | 96.71 (14.63)            | 101.43 (15.55)           | 362        | 103.92 (11.13)                     | 102.26 (10.85)           | 106.25 (11.29)           | 391        |
|                        | VIQ (SD)  | 98.37 (11.34)                      | 101.16 (12.49)           | 104.88 (13.37)           | 363        | 104.22 (11.71)                     | 103.04 (12.14)           | 106.23 (12.03)           | 391        |
|                        | PIQ (SD)  | 92.86 (16.66)                      | 92.27 (19.54)            | 97.49 (18.92)            | 362        | 104.48 (11.95)                     | 102.58 (11.33)           | 106.86 (11.65)           | 391        |
|                        |           | <b>CC</b><br><i>0.50</i>           | <b>CT</b><br><i>0.37</i> | <b>TT</b><br><i>0.13</i> |            | <b>CC</b><br><i>0.55</i>           | <b>CT</b><br><i>0.37</i> | <b>TT</b><br><i>0.07</i> |            |
| rs363043<br>(10174146) | FIQ (SD)  | 96.66 (15.64)                      | 99.38 (14.20)            | 105.17 (14.32)           | 346        | 102.65 (10.93)                     | 104.68 (11.85)           | 108.92 (12.74)           | 364        |
|                        | VIQ (SD)  | 101.25 (12.83)                     | 104.03 (12.76)           | 105.32 (12.34)           | 347        | 103.05 (12.07)                     | 104.61 (11.93)           | 108.90 (13.73)           | 364        |
|                        | PIQ (SD)  | 92.24 (19.01)                      | 94.31 (18.60)            | 103.80 (17.46)           | 346        | 103.41 (10.96)                     | 105.69 (13.30)           | 107.86 (11.90)           | 364        |
|                        |           | <b>CC</b><br><i>0.15</i>           | <b>CT</b><br><i>0.51</i> | <b>TT</b><br><i>0.34</i> |            | <b>CC</b><br><i>0.17</i>           | <b>CT</b><br><i>0.51</i> | <b>TT</b><br><i>0.32</i> |            |
| rs363016<br>(10179174) | FIQ (SD)  | 96.28 (15.24)                      | 96.82 (15.26)            | 103.03 (15.05)           | 340        | 101.92 (11.39)                     | 102.82 (10.94)           | 106.73 (12.07)           | 362        |
|                        | VIQ (SD)  | 99.66 (12.09)                      | 102.06 (12.92)           | 105.40 (13.21)           | 341        | 101.72 (12.65)                     | 103.54 (11.36)           | 106.42 (12.83)           | 362        |
|                        | PIQ (SD)  | 93.49 (19.26)                      | 91.59 (19.09)            | 99.66 (18.63)            | 340        | 103.64 (10.94)                     | 103.11 (12.07)           | 107.44 (12.14)           | 362        |
|                        |           | <b>AA</b><br><i>0.34</i>           | <b>AG</b><br><i>0.50</i> | <b>GG</b><br><i>0.16</i> |            | <b>AA</b><br><i>0.32</i>           | <b>AG</b><br><i>0.50</i> | <b>GG</b><br><i>0.18</i> |            |
| rs363050<br>(10182257) | FIQ (SD)  | 102.54 (15.00)                     | 96.94 (15.33)            | 96.52 (14.63)            | 363        | 106.44 (11.94)                     | 102.94 (10.57)           | 102.25 (11.08)           | 391        |
|                        | VIQ (SD)  | 105.18 (13.04)                     | 102.04 (13.16)           | 99.27 (11.70)            | 364        | 105.99 (12.72)                     | 104.01 (11.43)           | 102.02 (12.65)           | 391        |
|                        | PIQ (SD)  | 99.04 (18.58)                      | 91.91 (19.42)            | 94.40 (18.59)            | 363        | 107.43 (12.04)                     | 102.79 (11.59)           | 103.90 (10.32)           | 391        |

**Table 5.5** Family based association analysis for *SNAP-25 tag-SNPs* for *young*, *adult*, and *combined* cohorts

|                        |           | Young Cohort |             |                  |                 | Adult Cohort |                  |              |                 | Combined Cohort  |              |               |                 |
|------------------------|-----------|--------------|-------------|------------------|-----------------|--------------|------------------|--------------|-----------------|------------------|--------------|---------------|-----------------|
|                        | Phenotype | N            | P           | Genotypic Effect | N               | P            | Genotypic Effect | N            | P               | Genotypic Effect |              |               |                 |
| rs363039<br>(10168496) | FIQ       | 362          | <b>6.43</b> | <b>0.011</b>     | <b>3.61 (G)</b> | 391          | 5.56             | 0.018        | 2.12 (G)        | 753              | <b>11.79</b> | <b>0.0006</b> | <b>2.75 (G)</b> |
|                        | VIQ       | 363          | 5.64        | 0.018            | 2.97 (G)        | 391          | 2.83             | 0.093        | 1.70 (G)        | 754              | <b>7.47</b>  | <b>0.0063</b> | <b>2.13 (G)</b> |
|                        | PIQ       | 362          | 4.63        | 0.031            | 3.79 (G)        | 391          | <b>7.05</b>      | <b>0.008</b> | <b>2.45 (G)</b> | 753              | <b>11.53</b> | <b>0.0007</b> | <b>3.21 (G)</b> |
| rs363043<br>(10174146) | FIQ       | 346          | <b>7.48</b> | <b>0.006</b>     | <b>3.61 (T)</b> | 364          | 4.24             | 0.039        | 2.12 (T)        | 710              | <b>11.25</b> | <b>0.0008</b> | <b>2.75 (T)</b> |
|                        | VIQ       | 347          | 3.58        | 0.058            | 2.20 (T)        | 364          | 3.14             | 0.076        | 2.07 (T)        | 711              | <b>6.57</b>  | <b>0.0104</b> | <b>2.01 (T)</b> |
|                        | PIQ       | 346          | <b>7.46</b> | <b>0.006</b>     | <b>4.54 (T)</b> | 364          | 2.73             | 0.098        | 1.75 (T)        | 710              | <b>9.86</b>  | <b>0.0017</b> | <b>3.06 (T)</b> |
| rs363016<br>(10179174) | FIQ       | 340          | <b>8.71</b> | <b>0.003</b>     | <b>4.07 (T)</b> | 362          | <b>7.42</b>      | <b>0.006</b> | <b>2.57 (T)</b> | 702              | <b>15.99</b> | <b>0.0001</b> | <b>3.28 (T)</b> |
|                        | VIQ       | 341          | <b>7.02</b> | <b>0.008</b>     | <b>3.10 (T)</b> | 362          | <b>6.08</b>      | <b>0.014</b> | <b>2.43 (T)</b> | 703              | <b>13.01</b> | <b>0.0003</b> | <b>2.76 (T)</b> |
|                        | PIQ       | 340          | <b>6.08</b> | <b>0.014</b>     | <b>4.16 (T)</b> | 362          | 5.30             | 0.021        | 2.21 (T)        | 702              | <b>11.22</b> | <b>0.0008</b> | <b>3.21 (T)</b> |
| rs363050<br>(10182257) | FIQ       | 363          | <b>5.97</b> | <b>0.015</b>     | <b>3.16 (A)</b> | 391          | 5.57             | 0.018        | 2.01 (G)        | 754              | <b>11.36</b> | <b>0.0008</b> | <b>2.49 (A)</b> |
|                        | VIQ       | 364          | <b>7.49</b> | <b>0.006</b>     | <b>3.10 (A)</b> | 391          | 4.08             | 0.043        | 1.94 (G)        | 755              | <b>11.11</b> | <b>0.0009</b> | <b>2.38 (A)</b> |
|                        | PIQ       | 363          | 2.59        | 0.108            | 2.65 (A)        | 391          | 4.71             | 0.030        | 1.86 (G)        | 754              | <b>7.18</b>  | <b>0.0074</b> | <b>2.29 (A)</b> |

Abbreviations: FSIQ, full-scale IQ; PIQ, performance IQ; SNAP-25, synaptosomal-associated protein of 25 kDa; VIQ, verbal IQ. Note: The genotypic effect is the increase in IQ points associated with the increaser allele. P values below < 0.01 are in bold. Residual variance was modelled as a function of polygenic effects and non-shared environmental effects.

### *Haplotype analysis*

Based on LD patterns among the four (non)coding SNPs significantly associated with IQ phenotypes, only three were selected to conduct further haplotype analysis. The selected SNPs encompassed a genomic region of about 10.7 kb (rs363039, and two variants within LD block 4, rs363043, and rs363016). LD ( $r^2$ ) among these variants ranged between 0.10 and 0.54. These SNPs were used to estimate haplotypes within each sample. Haplotype analysis of SNPs with a relatively low LD is more powerful than single SNP analysis because the combination of SNPs into a haplotype can be considered as a multiallelic marker that is more informative than a bi-allelic marker when the causal variant(s) are not genotyped. Five possible haplotypes were observed in our samples (*A-C-C*, *A-T-T*, *G-C-C*, *G-C-T*, and *G-T-T*). Haplotypes *A-T-C* and *G-T-C* were not observed and *A-C-T* was only observed in the adult cohort at a very low frequency (see Table 5.6). Significant associations were found in both samples; however, it is worth noting that different allelic combinations were associated across cohorts. Within the *young* cohort, *G-T-T* was the strongest associated haplotype (PIQ  $\chi^2(1)=9.36$ ,  $P=0.002$ ); whereas *G-C-T* showed the strongest association among the *adult* cohort (FSIQ  $\chi^2(1)=10.08$ ,  $P=0.001$ ).

When the data were combined, highly significant associations were observed among all IQ phenotypes for both, *G-T-T* (PIQ ( $\chi^2(1)=8.27$ ,  $P=0.004$ ) and *G-C-T* (VIQ ( $\chi^2(1)=8.61$ ,  $P=0.003$ ) haplotypes (see Table 5.6), confirming the single SNP association results. Haplotype analysis was also conducted for the rs362620 and rs362557 haplotype in the second region. Significant association was found between the *A-C* haplotype and VIQ ( $\chi^2(1)=6.09$ ,  $P=0.01$ ) with an increase of 2.65 IQ points, corroborating the single SNP analysis in the *young* cohort. This association should be interpreted with more than the usual caution since the single SNP analyses had not shown replication between cohorts or within cohort across different IQ phenotypes.

### *Bioinformatics*

A search with MatInspector and SNPInspector revealed that all three SNPs affected TF binding sites (see Table 5.7). Although there is no hard evidence that these TF binding sites are functional, it at least allows for the possibility that the (non)coding variants identified could affect regulatory gene expression.

**Table 5.6** | Family based association analysis for *tagging* haplotypes (rs363039,rs363043,rs363016) on intron 1 of the *SNAP-25* gene for *young*, *adult*, and *combined* cohorts

| Haplotype                   |           | Young Cohort (328) |              |                  | Adult Cohort (325) |              |                  | Combined Cohort (653) |              |                  |
|-----------------------------|-----------|--------------------|--------------|------------------|--------------------|--------------|------------------|-----------------------|--------------|------------------|
| (Young/Adult)               | Phenotype | P                  |              | Genotypic Effect | P                  |              | Genotypic Effect | P                     |              | Genotypic Effect |
| <i>A-C-C</i><br>(0.30/0.31) | FSIQ      | 3.89               | 0.049        | -3.46            | 4.74               | 0.029        | -9.64            | <b>6.88</b>           | <b>0.009</b> | <b>-4.07</b>     |
|                             | VIQ       | <b>6.54</b>        | <b>0.011</b> | <b>-3.75</b>     | 4.97               | 0.026        | -8.53            | <b>10.20</b>          | <b>0.001</b> | <b>-4.13</b>     |
|                             | PIQ       | 0.87               | 0.351        | -2.08            | 2.90               | 0.089        | -9.65            | 1.96                  | 0.162        | -2.65            |
| <i>A-C-T</i><br>(NA/0.02)   | FSIQ      | NA                 | NA           | NA               | 2.98               | 0.084        | -3.61            | 2.18                  | 0.140        | 1.51             |
|                             | VIQ       | NA                 | NA           | NA               | 2.40               | 0.121        | -4.52            | 1.57                  | 0.210        | 0.66             |
|                             | PIQ       | NA                 | NA           | NA               | 1.99               | 0.158        | 0.00             | 1.77                  | 0.183        | 2.65             |
| <i>A-T-T</i><br>(0.03/0.01) | FSIQ      | 0.24               | 0.624        | 1.20             | 2.83               | 0.093        | -3.61            | 0.76                  | 0.383        | -1.36            |
|                             | VIQ       | 1.14               | 0.286        | 2.33             | 2.12               | 0.145        | -2.58            | 0.06                  | 0.806        | -0.39            |
|                             | PIQ       | 0.04               | 0.841        | -0.76            | 2.30               | 0.129        | -3.98            | 1.63                  | 0.202        | -2.65            |
| <i>G-C-C</i><br>(0.13/0.11) | FSIQ      | 2.38               | 0.123        | -2.64            | 0.92               | 0.337        | -2.71            | 3.26                  | 0.071        | -2.41            |
|                             | VIQ       | 0.61               | 0.435        | -1.16            | 1.04               | 0.308        | -2.45            | 1.76                  | 0.185        | -1.55            |
|                             | PIQ       | 3.47               | 0.062        | -3.98            | 0.29               | 0.590        | -1.89            | 2.91                  | 0.088        | -2.84            |
| <i>G-C-T</i><br>(0.28/0.29) | FSIQ      | 0.54               | 0.462        | 1.05             | <b>10.08</b>       | <b>0.001</b> | <b>4.37</b>      | <b>7.33</b>           | <b>0.007</b> | <b>2.71</b>      |
|                             | VIQ       | 2.34               | 0.126        | 1.94             | <b>7.26</b>        | <b>0.007</b> | <b>3.23</b>      | <b>8.61</b>           | <b>0.003</b> | <b>2.58</b>      |
|                             | PIQ       | 0.01               | 0.920        | 0.18             | <b>7.79</b>        | <b>0.005</b> | <b>4.92</b>      | 4.04                  | 0.044        | 2.65             |
| <i>G-T-T</i><br>(0.26/0.26) | FSIQ      | <b>6.97</b>        | <b>0.008</b> | <b>3.76</b>      | 0.94               | 0.332        | 1.36             | <b>6.57</b>           | <b>0.010</b> | <b>2.56</b>      |
|                             | VIQ       | 1.94               | 0.164        | 1.68             | 0.74               | 0.390        | 0.78             | 2.39                  | 0.122        | 1.29             |
|                             | PIQ       | <b>9.36</b>        | <b>0.002</b> | <b>5.49</b>      | 1.00               | 0.317        | 1.70             | <b>8.27</b>           | <b>0.004</b> | <b>3.60</b>      |

**Table 5.7** | List of putative transcription factor (TF) binding sites modified by common SNPs within intron 1 on the *SNAP-25* gene

| SNP<br>Position (bp)   | Change | Gain<br>Loss | Vertebrate Matrix | Position <sup>a</sup>  | Strand | Name   | Random<br>expectation <sup>b</sup> | Promoter<br>matches <sup>c</sup> |
|------------------------|--------|--------------|-------------------|------------------------|--------|--|------------------------------------|----------------------------------|
| rs363039<br>(10168496) | G->A   | lost         | V\$P53F/P53.02    | 10168485 -<br>10168507 | -      | tumor suppressor p53   | 4.06                               | 14.4                             |
|                        | G->A   | lost         | V\$ZNFP/SZF1.01   | 10168493 -<br>10168517 | +      | SZF1, hematopoietic progenitor<br>-restricted KRAB-zinc finger protein                     | 0.17                               | 13.1                             |
| rs363043<br>(10174146) | C->T   | new          | V\$COMP/COMP1.01  | 10174136 -<br>10174158 | +      | COMP1, cooperates with myogenic<br>proteins in multicomponent complex                      | 0.02                               | 31.7                             |
|                        | C->T   | new          | V\$EVII/EVII.04   | 10174146 -<br>10174162 | -      | Ecotropic viral integration site 1<br>encoded factor, amino-terminal<br>zinc finger domain | 0.01                               | 23.2                             |
|                        | C->T   | lost         | V\$HOXC/PBX1.01   | 10174139 -<br>10174155 | -      | Homeo domain factor Pbx-1  | 0.05                               | 21.8                             |
| rs363016<br>(10179174) | C->T   | new          | V\$HOXF/HOXB9.01  | 10179168 -<br>10179184 | +      | Abd-B-like homeodomain<br>protein Hoxb-9   | 1.41                               | 33.0                             |
|                        | C->T   | new          | V\$TBPF/ATATA.01  | 10179167 -<br>10179183 | -      | Avian C-type LTR TATA box  | 0.04                               | 41.2                             |

Note: <sup>a</sup>Based on Build 36.1, <sup>b</sup>Matches per 1000 base pairs (bp), <sup>c</sup>Percentage (%) of vertebrate promoters containing the matched matrix.

Abbreviations SZF1: Krüppel-like zinc finger protein, KRAB: Krüppel-associated box, COMP1: Myogenic factor 4, Pbx-1: pre B-cell leukemia transcription factor 1, Hoxb-9: Homeobox B9, LTR: Lont terminal repeat

## DISCUSSION

To continue our investigation of the possible role of the *SNAP-25* gene in intelligence, we employed a family-based genetic association test in two independent cohorts of 371 children (mean age 12.42), and 391 adults (mean age 36.25). The selected SNPs gave a dense coverage of the first intron of SNAP 25, which was previously reported to be associated with intelligence (Gosso *et al.* 2006). Single and haplotype analysis was conducted in the present study in order to (1) narrow-down the location of causative genetic variant underlying the association on intron 1, and (2) identify extra regions on *SNAP-25* gene not *tagged* during the previous analyses.

Four SNPs (rs363039, rs363043, rs363016, and rs363050) located on intron 1, showed significant association with IQ phenotypes. Haplotype analysis confirmed the single association results. Combined data across age cohorts showed highly significant associations among all IQ phenotypes for both *G-T-T* (PIQ ( $\chi^2(1)=8.27$ ,  $P=0.004$ ) and *G-C-T* (VIQ ( $\chi^2(1)=8.61$ ,  $P=0.003$ ) haplotypes. Interestingly, two haplotypes were independently found associated to IQ phenotypes among *young* and *adult* cohorts. Within the *young* cohort, *G-T-T* was the strongest associated haplotype (PIQ  $\chi^2(1)=9.36$ ,  $P=0.002$ ); whereas *G-C-T* showed the strongest association among the *adult* cohort (FSIQ  $\chi^2(1)=10.08$ ,  $P=0.001$ ). Such differential genotypic effects might be possibly explained within a heterogeneous genomic context. For example, although physical TF binding sites are a constitutive portion of the genome, the cellular and genomic context will determine whether a given TF sequence(s) become functional or not. Transcription factors are differentially expressed in response to developmental requirements, and even more important, like QTLs, single TF will not be sufficient to trigger a regulatory response, but they will rather interact in a collaborative manner. Likely, IQ can be considered a truly polygenic complex trait, and as such, not a single common allelic variant might be involved in variation among IQ phenotypes, but rather, similar genetic effects might be exerted by diverse allelic variants. Alternatively our results could indicate that the causal variant is older than the SNPs that have been tested and in fact is present on both haplotypes.

The *SNAP-25* gene, located on chromosome 20 p12-12p11.2 encodes a presynaptic terminal protein. *SNAP-25* is thought to be differentially expressed in the brain, and is primarily present in the neocortex, hippocampus, anterior thalamic nuclei, substantia nigra, and cerebellar granular cells. In the mature brain, expression is mainly seen at presynaptic terminals (Oyler *et al.* 1989). Two splicing

variant of the *SNAP-25* exist, *SNAP-25a* and *SNAP-25b* isoforms (Bark & Wilson 1991). During development, *SNAP-25a* isoform is involved in synaptogenesis, forming presynaptic sites and neuritic outgrowth (Osen-Sand *et al.* 1993; Oyler *et al.* 1989), whereas in the mature brain, the *SNAP-25b* isoform forms a complex with syntaxin and the synaptic vesicle proteins (synaptobrevin and synaptotagmin) that mediates exocytosis of neurotransmitter from the synaptic vesicle into the synaptic cleft (see Horikawa *et al.*, 1993; Seagar *et al.*, 1998, Bark *et al.*, 1994 ; Low *et al.*, 1999). *SNAP-25* isoforms (*SNAP-25a*, and *SNAP-25b*) are fundamental for keeping a balanced trade-off between synaptic formation and neurotransmitter vesicle release, however, evolutionary (comparative genomics) analysis of the coding sequence showed no selection in favour of any of the genes' coding variants on *SNAP-25*. If variation of coding variants may not *per se* be associated to phenotypic variation, then, the next possible scenario might be the presence of regulatory effects exerted by variants on (non)coding regions. Regulatory (non)coding variants may interact in a concerted manner rather than in isolation, with the capacity to regulate gene expression. Genetic (non)coding variants present within intron 1 might be involved in regulation of protein isoforms expression. All associated SNPs were involved in transcription factor (TF) binding sites changes (gain/loss of TF binding sites). Furthermore, because functional TF binding sites are predicted to interact in a cooperative manner rather than in isolation, a global overview might be required to (1) identify known and unknown TF binding sites and (2) putative functional (non)coding polymorphisms that may affect spatial and temporal regulation of gene expression.

Contrary to what is expected in Mendelian traits, subtle changes are postulated to influence the phenotypic outcome of complex (common) traits. Further functional studies may aid identification of functional polymorphisms that may affect functional TF binding sites, which in turn may be utilized to uncover genetic regulatory interactions underlying normal cognitive variation.

## REFERENCES

- Abecasis, G.R., Cardon, L.R. & Cookson, W.O. (2000) A general test of association for quantitative traits in nuclear families. *Am J Hum Genet* 66 279-292.
- Bark, C. & Wilson, M.C. (1991) Alternative splicing generates a variant SNAP-25 protein during development. *Soc Neurosci Abstr* 17 531.
- Bark, I.C. & Wilson, M.C. (1994) Human cDNA clones encoding two different isoforms of the nerve terminal protein SNAP-25. *Gene* 139 291-292.
- Bliss, T.V. & Collingridge, G.L. (1993) A synaptic model of memory: long-term potentiation in the hippocampus. *Nature* 361 31-39.
- Boomsma, D.I. (1998) Twin registers in Europe: an overview. *Twin Res* 1 34-51.
- Boomsma, D.I., de Geus, E.J., Vink, J.M., Stubbe, J.H., Distel, M.A., Hottenga, J.J., Posthuma, D., van Beijsterveldt, T.C., Hudziak, J.J., Bartels, M. & Willemsen, G. (2006) Netherlands Twin Register: from twins to twin families. *Twin Res Hum Genet* 9 849-857.
- Butcher, L.M., Kennedy, J.K. & Plomin, R. (2006) Generalist genes and cognitive neuroscience. *Curr Opin Neurobiol* 16 145-151.
- Cartharius, K., Frech, K., Grote, K., Klocke, B., Haltmeier, M., Klingenhoff, A., Frisch, M., Bayerlein, M. & Werner, T. (2005) MatInspector and beyond: promoter analysis based on transcription factor binding sites *Bioinformatics* 21 2933-2942.
- Excoffier, L. & Slatkin, M. (1995) Maximum-likelihood estimation of molecular haplotype frequencies in a diploid population. *Mol Biol Evol* 12 921 - 927.
- Frassoni, C., Inverardi, F., Coco, S., Ortino, B., Grumelli, C., Pozzi, D., Verderio, C. & Matteoli, M. (2005) Analysis of SNAP-25 immunoreactivity in hippocampal inhibitory neurons during development in culture and in situ. *Neuroscience* 131 813.
- Fulker, D.W., Cherny, S.S., Sham, P.C. & Hewitt, J.K. (1999) Combined linkage and association sib-pair analysis for quantitative traits. *Am J Hum Genet* 64 259-267.
- Gabriel, S.B., Schaffner, S.F., Nguyen, H., Moore, J.M., Roy, J., Blumenstiel, B., Higgins, J., DeFelice, M., Lochner, A., Faggart, M., Liu-Cordero, S.N., Rotimi, C., Adeyemo, A., Cooper, R., Ward, R., Lander, E.S., Daly, M.J. & Altshuler, D. (2002) The structure of haplotype blocks in the human genome. *Science* 296 2225-2229.
- Garvie, C.W. & Wolberger, C. (2001) Recognition of specific DNA sequences. *Mol Cell* 8 937-946.
- Geddes, J.W., Hess, E.J., Hart, R.A., Kesslak, J.P., Cotman, C.W. & Wilson, M.C. (1990) Lesions of hippocampal circuitry define synaptosomal-associated protein-25 (SNAP-25) as a novel presynaptic marker. *Neuroscience* 38 515-525.
- Gosso, M.F., de Geus, E.J., van Belzen, M.J., Polderman, T.J., Heutink, P., Boomsma, D.I. & Posthuma, D. (2006) The SNAP-25 gene is associated with cognitive ability: evidence from a family-based study in two independent Dutch cohorts. *Mol Psychiatry* 11 878-886.
- Greber-Platzer, S., Fleischmann, C., Nussbaumer, C., Cairns, N. & Lubec, G. (2003) Increased RNA levels of the 25 kDa synaptosomal associated protein in brain samples of adult patients with Down Syndrome. *Neurosci Lett* 336 77-80.
- Grosse, G., Grosse, J., Tapp, R., Kuchinke, J., Gorsleben, M., Fetter, I., Hohne-Zell, B., Gratzl, M. & Bergmann, M. (1999) SNAP-25 requirement for dendritic growth of hippocampal neurons. *J Neurosci Res* 56 539-546.
- Gudbjartsson, D.F., Thorvaldsson, T., Kong, A., Gunnarsson, G. & A, I. (2005) Allegro version 2. *Nature Genetics* 37 1015-1016.
- Hoekstra, C., Meijer, P., Kluff, C., Heutink, P., Smit, G., de Geus, E., Smit, J.H., van Bruggen, A., Montgomery, G.W. & Boomsma, D.I. (2004) Genetics of dizygotic twinning: a feasibility study for a biobank. *Twin Res* 7 556-563.
- Horikawa, H.P., Saisu, H., Ishizuka, T., Sekine, Y., Tsugita, A., Odani, S. & Abe, T. (1993) A complex of rab3A, SNAP-25, VAMP/synaptobrevin-2 and syntaxins in brain presynaptic terminals. *FEBS Lett* 330 236-240.
- Hou, Q., Gao, X., Zhang, X., Kong, L., Wang, X., Bian, W., Tu, Y., Jin, M., Zhao, G., Li, B., Jing, N. & Yu, L. (2004) SNAP-25 in hippocampal CA1 region is involved in memory consolidation. *Eur J Neurosci* 20 1593-1603.
- Hou, Q.L., Gao, X., Lu, Q., Zhang, X.H., Tu, Y.Y., Jin, M.L., Zhao, G.P., Yu, L., Jing, N.H. & Li, B.M. (2006) SNAP-25 in hippocampal CA3 region is required for long-term memory formation. *Biochem Biophys Res Commun* 347 955-962.
- Ibrahimi, O.A., Chiu, E.S., McCarthy, J.G. & Mohammadi, M. (2005) Understanding the molecular basis of Apert syndrome. *Plast Reconstr Surg* 115 264-270.
- Low, P., Norlin, T., Risinger, C., Larhammar, D., Pieribone, V.A., Shupliakov, O. & Brodin, L. (1999) Inhibition of neurotransmitter release in the lamprey reticulospinal synapse by antibody-mediated disruption of SNAP-25 function. *Eur J Cell Biol* 78 787-793.
- Martin, S.J., Grimwood, P.D. & Morris, R.G. (2000) Synaptic plasticity and memory: an evaluation of the hypothesis. *Annu Rev Neurosci* 23 649-711.
- Meulenbelt, I., Droog, S., Trommelen, G.J., Boomsma, D.I. & Slagboom, P.E. (1995) High-yield noninvasive human genomic DNA isolation method for genetic studies in geographically dispersed families and populations. *Am J Hum Genet* 57 1252-1254.
- Miller, S.A., Dykes, D.D. & Polesky, H.F. (1988) A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* 16 1215.

- Morris, R.G. (1989) Synaptic plasticity and learning: selective impairment of learning rats and blockade of long-term potentiation in vivo by the N-methyl-D-aspartate receptor antagonist AP5. *J Neurosci* 9 3040-3057.
- Morris, R.G., Anderson, E., Lynch, G.S. & Baudry, M. (1986) Selective impairment of learning and blockade of long-term potentiation by an N-methyl-D-aspartate receptor antagonist, AP5. *Nature* 319 774-776.
- Neul, J.L. & Zoghbi, H.Y. (2004) Rett syndrome: a prototypical neurodevelopmental disorder. *Neuroscientist* 10 118-128.
- Osen-Sand, A., Catsicas, M., Staple, J.K., Jones, K.A., Ayala, G., Knowles, J., Grenningloh, G. & Catsicas, S. (1993) Inhibition of axonal growth by SNAP-25 antisense oligonucleotides in vitro and in vivo. *Nature* 364 445-448.
- Oyler, G.A., Higgins, G.A., Hart, R.A., Battenberg, E., Billingsley, M., Bloom, F.E. & Wilson, M.C. (1989) The identification of a novel synaptosomal-associated protein, SNAP-25, differentially expressed by neuronal subpopulations. *J Cell Biol* 109 3039-3052.
- Plomin, R. & Spinath, F.M. (2004) Intelligence: genetics, genes, and genomics. *J Pers Soc Psychol* 86 112-129.
- Polderman, T.J.C., Gosso, M.F., Posthuma, D., Van Beijsterveldt, T.C., Heutink, P., Verhulst, F.C. & Boomsma, D.I. (2006a) A longitudinal twin study on IQ, executive functioning, and attention problems during childhood and early adolescence. *Acta Neurol Belg* 106 191-207.
- Polderman, T.J.C., Stins, J.F., Posthuma, D., Gosso, M.F., Verhulst, F.C. & Boomsma, D.I. (2006b) The phenotypic and genotypic relation between working memory speed and capacity. *Intelligence* 34 549-560.
- Posthuma, D., de Geus, E.J., Boomsma, D.I. & Neale, M.C. (2004) Combined linkage and association tests in mx. *Behav Genet* 34 179-196.
- Posthuma, D., Luciano, M., Geus, E.J., Wright, M.J., Slagboom, P.E., Montgomery, G.W., Boomsma, D.I. & Martin, N.G. (2005) A Genomewide Scan for Intelligence Identifies Quantitative Trait Loci on 2q and 6p. *Am J Hum Genet* 77 318-326.
- Savitz, J., Solms, M. & Ramesar, R. (2006) The molecular genetics of cognition: dopamine, COMT and BDNF. *Gene Brain Behav* 5 311-328.
- Seagar, M. & Takahashi, M. (1998) Interactions between presynaptic calcium channels and proteins implicated in synaptic vesicle trafficking and exocytosis. *J Bioenerg Biomembr* 30 347-356.
- Squire, L.R. & Kandel, E.R. (1999) Memory: From Mind to Molecules. New York: Scientific American Library.
- Thompson, P.M., Egbofoama, S. & Vawter, M.P. (2003) SNAP-25 reduction in the hippocampus of patients with schizophrenia. *Prog Neuropsychopharmacol Biol Psychiatry* 27 411-417.
- Verkerk, A.J., Pieretti, M., Sutcliffe, J.S., Fu, Y.H., Kuhl, D.P., Pizzuti, A., Reiner, O., Richards, S., Victoria, M.F., Zhang, F.P. & et al. (1991) Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome. *Cell* 65 905-914.
- Wechsler, D. (1986) Wechsler Intelligence Scale for Children-Revised (WISC-R). Lisse: Swets & Zeitlinger.
- Wechsler, D. (1997) WAIS-III Wechsler Adult Intelligence Scale San Antonio, Texas: Psychological Corporation.





# CHAPTER 6

A functional non-synonymous  
polymorphism under positive evolutionary  
selection in the *ADRB2* gene  
explains phenotypic differences in  
intelligence among humans

Z Bochdanovits, MF Gosso, L van den Berg, P Rizzu, T JC Polderman, LM Pardo, JCN de Geus, DI Boomsma, P Heutink, and D Posthuma (*submitted*) A functional non-synonymous polymorphism under positive evolutionary selection in the beta-2 adrenergic receptor explains phenotypic differences in intelligence among humans

*Abstract*

Comparative genomics offers a novel approach to unravel the genetic basis of complex traits. We propose a two stage design where genes ascertained for enhanced protein evolution in primates are subsequently searched for the presence of non-synonymous coding SNPs in the current human population at amino acid sites that differ between humans and chimpanzee. Positively selected genes among primates are generally presumed to determine phenotypic differences between humans and chimpanzee, such as the enhanced cognitive ability of our species. Amino acid substitutions segregating in humans at positively selected amino acid sites are expected to affect phenotypic differences among humans, therefore we conducted an association study between cognitive ability and the most likely candidate gene among the five that harboured more than one such polymorphism. The derived, human specific, allele of the beta-2 adrenergic receptor was found to be the increaser allele and explained 1% of the phenotypic variation in performance IQ. The polymorphism is known to affect signaling activity and modulation of beta-2 adrenergic signaling has been shown to adjust memory consolidation, a trait related to cognition. This result illustrates the validity of our approach to detect functional variation and demonstrates for the first time a direct link between quantitative genetics and the genetics of species differences.

## INTRODUCTION

Human genetics has traditionally been involved in unraveling the genetic basis of phenotypic differences between individuals, mostly for disease related traits. This research program has proven to be very productive for detecting the genetic basis of monogenetic disorders. Progress in understanding the heritability of complex traits has proven more difficult and is faced with the problem of ascertaining the functional variant following the Linkage Disequilibrium based mapping (Cardon and Bell 2001). Given the need for novel methods to identify functional genetic variants on a genome-wide scale, we focused on a relatively recent and very interesting addition to the field of human genetics and propose a two stage approach to ascertain amino acid substitutions that affect a complex trait in humans based on a comparative genomic analysis.

In evolutionary genetic terms, human genetic studies aimed at detecting disease mutations operate in the realm of microevolution (within species) with the main focus on describing genetic variation that inflicts undesired vs. desired, i.e. (mal-)adaptive, phenotypes. Comparative genomics effectively studies the macroevolution of our species searching for the genetic basis of the adaptive phenotypic divergence that uniquely identifies *Homo sapiens*. Assuming that the between and within species differences in a given trait are at least in part caused by the same genes, there is an obvious added value in trying to merge these two fields. Traits that have presumably been subject of adaptive divergence during human evolution, such as reproduction (Torgerson *et al.* 2002), immune system (Hughes 1997), social behavior (Sterck *et al.* 1997) and cognition, are still highly relevant for maintaining the desired quality of life of modern humans in an every day setting. Consequently, the collection of genes that can be linked to the adaptive evolution of our species are bound to be enriched for genes that affect these highly relevant phenotypes. We assume that if such genes still harbor functional genetic variation within our species then these different alleles are bound to explain current phenotypic (possibly pathological) differences between humans in reproductive disorders, (auto-)immune disease and all behavioral/neurological syndromes, including natural variation in cognitive ability.

Comparative genomic approaches have been used before to identify genes that may have contributed to the evolution of “human specific traits” (Clark *et al.* 2003; Nielsen *et al.* 2005). These studies have successfully compared the rate of protein evolution among primates and rodents, to identify fast evolving genes in the lineage leading to modern humans. These genes can be expected to affect traits that quantitatively differ between humans and other primates. Alternatively, brain related genes have been shown to have a higher rate of divergence within primates

compared to rodents (Dorus *et al.* 2004), confirming the general belief that becoming human has involved phenotypic changes to our brain. Recently, it has been suggested that evolutionary changes to gene expression might contribute to the phenotypic differences between humans and apes. Again, evidence has been put forward to suggest positive selection in the human brain (Khaitovich *et al.* 2006). These studies together suggest that during human evolution easily detectable genetic changes have occurred and some of them affect our brain. If so, strategies aimed at ascertaining fast evolving genes/genomic regions should allow the identification of novel genes with important functions in the central nervous system (CNS). Indeed, recently a novel RNA gene involved in cortical development has been described following an analysis of genomic regions subject to accelerated evolution in humans (Pollard *et al.* 2006). However, as yet none of the identified fast evolving gene has been linked directly to the adaptive phenotype it presumably affects.

The comparative genomic approach applied here was based on data currently available from the NCBI Homologene database. In the first stage the analysis was aimed at detecting genes that had been subject of enhanced protein evolution specifically in the primate lineage leading to modern humans. Such genes may be expected to affect an adaptive phenotypic trait that should be quantitatively different between humans and our closest primate relatives, i.e. “human specific”. The rate of protein evolution can be (crudely) measured as the ratio of non-synonymous vs. synonymous substitution rates among pair-wise comparisons of species, commonly referred to as  $K_a/K_s$ . It has been argued before that if a gene exhibits an enhanced rate of protein evolution among primates but not among rodents that can be considered evidence for positive selection specific to the lineage leading to modern humans as opposed to diversifying selection among mammals in general (Clark *et al.* 2003; Dorus *et al.* 2004). Being a summary statistic across the entire coding sequence, a high  $K_a/K_s$  ratio is known to be a rather conservative method for assessing biologically relevant protein evolution (MacCallum and Hill 2006). However, the aim of the present study is not to detect *all* genes involved in the evolution of the human brain, but to find a limited number of loci that could be unequivocally linked to phenotypic variation. In other words, our main methodological concern is to restraint the false positive rate. Following the identification of positively selected genes we conducted a second step of the analysis where non-synonymous coding SNPs were ascertained at positions that differ between the human and chimpanzee reference sequences. This approach has not been applied before and identifies amino acid changes in humans that are obvious candidates to have a functional effect on a phenotype that has diverged during

human speciation (Figure 6.1). To confirm the assumption that such positively selected amino acid variants affect phenotypic differences among humans a genetic association study has been performed with cognitive ability based on two independent family based Dutch samples.

## MATERIALS AND METHODS

### *Data and analysis*

The NCBI Homologene database build 39.1 has been downloaded and the relevant information on the rate of nucleotide substitutions between humans and chimpanzee and mouse and rat has been extracted using custom made scripts in a Linux environment. The ratio between the non-synonymous vs. synonymous substitution rate has been calculated using Jukes and Cantor's method (Weir 1983). Because the distribution of Ka/Ks values was different between primates and rodents, the data were standard normalized before comparing the ratio between lineages. All statistical analyses have been performed in SPSS 11.5.

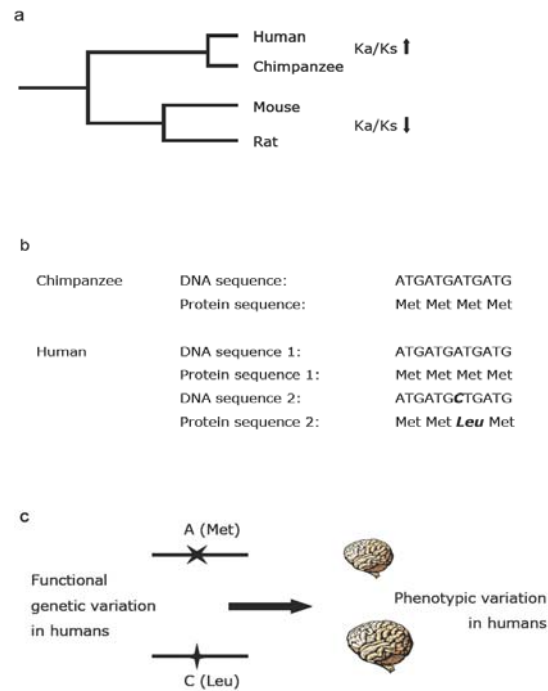
### *Simulation study*

The evolver and codeml programs from the PAML package have been used. Evolver can simulate the evolution of coding sequences between the four species conditioned on the known topology of the phylogenetic tree connecting the species and the observed average Ka/Ks values. Codeml subsequently analyses the data to generate the ration of non-synonymous vs. synonymous substitution rate. These values were than subjected to a similar analysis as the real data.

### *Genetic association*

The QTDT software package was used to perform a family based association analysis in two Dutch cohorts phenotyped for several measures of cognitive ability (for details see Gosso *et al.* 2006). IQ data were corrected for age and sex, and all analyses were performed while modeling the environmental and (poly-) genetic components of variance.

**Figure 6.1** | Template of a comparative genomic approach to ascertain functional genetic variants in the human genome. **a**, genome-wide comparison of Ka/Ks values among primates vs. rodents. Including two non-primate species allows for excluding genes with divergent protein evolution in mammals in general. Genes with enhanced protein evolution specific to the primate lineage are followed up in the next step. **b** the coding sequence of genes that show accelerated protein evolution between humans and chimpanzee are searched for non-synonymous mutations present in the current human population at amino acid positions that differ between the two species. **c**, the functional genetic variants identified in the comparative genomic analysis are subjected to genetic association analysis in human samples



## RESULTS

### *Simulation study to estimate false positive rate*

To estimate false positive rate from our comparative genomic analysis, the evolution of 25000 protein coding DNA sequences each 1800 bps long has been simulated using the PAML software package (Yang 1997). The simulations were conditioned on the known topology of the phylogenetic tree connecting the four species and on the observed average Ka/Ks ratio in the real dataset. The accuracy of the simulation for recreating realistic DNA sequence evolution was confirmed by considering the average nucleotide divergence between the four simulated species. The “human”-“chimp”, “rat”-“mouse” and “human”-“mouse” divergence was about 96%, 93% and 85% respectively, closely matching the real data. The simulated dataset was subsequently used to estimate false positive rate, as the percentage of the 25,000 “genes” that would have been called positive in the real dataset. The variability in the estimated false positive rate was very small based on 15 replicates, with an average of 1.5% +/- 0.1%.

### *Positively selected genes in primates*

The NCBI Homologene database (build 39.1) contained information on the number of synonymous and non-synonymous substitutions between humans and chimpanzee for more than 13,000 genes. Six hundred and twenty-six genes had a

Ka/Ks ratio higher than 1. For 311 of these genes data were available from the mouse-rat comparison as well. Because the aim of this analysis was to identify enhanced protein evolution specific to the primate lineage, genes were considered only if the normalized Ka/Ks ratio in mouse-rat was more than 2x the standard deviation lower compared to the human – chimp comparison. Two hundred and two out of 311 genes fulfilled this criterion (Supplementary Table 6.1). Because data on all four species was available for only 7,080 genes, the percentage positive genes is 2.9%. This is 2-fold higher than the expected false positive rate ( $p=2.57 \times 10^{-19}$ , one-sample t-test).

### *Positively selected genes are predominantly expressed in the human brain*

The number of genes found to have enhanced rate of protein evolution specifically among primates exceeds what could be expected by chance, therefore this set of genes can be expected to contain genes that have contributed to the phenotypic divergence between the two species. This divergence has undoubtedly involved brain related phenotypes, consequently we searched the NCBI UniGene database and listed for all genes whether or not they were expressed in the central nervous system. According to this database 32% (28,149 out of 86,810 entries) of the human genome is expressed in the brain. Among the 334 genes with only human vs. chimpanzee data available this percentage was 60%, but among the 202 genes subject of primate specific positive selection 79% was expressed in brain ( $p=2 \times 10^{-40}$ , one sample t-test).

This result shows that the criteria we applied for the comparative genomic analysis indeed significantly enriches the sample for brain related genes. This finding is in line with previous results (Dorus *et al.* 2004) and shows that becoming human has significantly coincided with the phenotypic evolution of our brain.

### *Non-synonymous coding SNPs at amino acids that differ between humans and chimpanzee*

Given that the genes identified in this analysis are indeed predominantly brain related we hypothesized that among these genes there are candidate genes for behavioral traits. Clearly, a wide range of phenotypes are potential “candidate traits”, however, the most obvious trait that humans like to think of as “human specific” is our cognitive ability. Among other equally feasible traits, intelligence must surely be one that has been subject to adaptive evolution. Hence positively selected human proteins with expression in the central nervous system (CNS) can be expected to affect cognition. Among the 202 “confirmed” positively selected genes, 160 (79%) are expressed in the CNS and among these 22 (14%) contain

non-synonymous coding SNPs, i.e. amino acid changing mutations, at amino acids that differ between the human and chimpanzee reference sequences. Effectively, the current human population segregates “derived”, “human specific” and “ancestral”, “chimpanzee specific” variants of these proteins (see Table 6.1). Assuming that these fast evolving genes with expression on the brain have indeed contributed to the phenotypic divergence between the two species, it is tempting to assume that the ancestral vs. the derived alleles that are still present in the human population might contribute to phenotypic differences among modern humans for the same traits that have diverged during our speciation. From the 22 genes identified here, 5 harbor more than one amino acid changing polymorphisms and one is a very good candidate gene for cognitive ability. To address this issue, we performed a genetic association study for IQ using these 2 non-synonymous coding SNPs (rs1042713 and rs1042714) in the beta-2 adrenergic receptor as markers in two independent family based Dutch samples.

### *Genetic association between the beta-2 adrenergic receptor and cognition*

Three measures of cognitive ability have been analyzed: full scale, verbal and performance IQ. Two independent family based Dutch samples of 391 (mean age 12.4 years) and 409 (mean age 36.7 years) subjects respectively have been phenotyped for these traits. The Dutch adaptations of the children and adult versions of the Wechsler Intelligence Scale were used. For a detailed description of the cohorts and the measurement instruments see Gosso *et al.* (2006), barring the inclusion of 133 additional adult samples. The QTDT software package (Abecasis *et al.* 2000) was used to first check for inconsistencies in the genotypes and test for Hardy & Weinberg equilibrium. No inconsistencies were found which suggests that no systematic genotyping error has occurred, but *ADRB2* rs1042714 was not in H&W equilibrium. The basic assumption of the current approach is that the polymorphisms under study have been subject to directional selection, and this might be a reason for failure to be in HW equilibrium. However, to avoid ambiguity in the interpretation of the results of the association study this marker was not followed up. Before performing the association analysis first a test for population stratification was carried out with the remaining SNP. *ADRB2* rs1042713 showed significant stratification, consequently only the family based evidence for association will be interpreted. *ADRB2* rs1042713 showed evidence for family based association with performance IQ. The association was only present in the young cohort ( $p=0.0113$ ) (see Table 6.2), , this might be explained by the observation that different age classes are known to differ in the genetic architecture of cognitive ability reflected in different heritability estimates (see Gosso *et al.* 2006). The derived, human specific A allele was the increaser allele and explained



1% of the phenotypic variation in IQ. It should be noted however that effect sizes are usually overestimated in sample sizes of this order.

**Table 6.1** Twenty two fast evolving, brain expressed genes that segregate “derived”, “human specific” and “ancestral”, “chimpanzee specific” variants of these proteins in the extant human population

| Gene         | rs #       | SNP status | Allele | Allele frequency CEU | Protein Residue | Amino Acid position |
|--------------|------------|------------|--------|----------------------|-----------------|---------------------|
| HEXB         | rs11556045 | Derived    | A      | 0.788                | Lys [K]         | 121                 |
|              |            | Ancestral  | G      | 0.212                | Arg [R]         |                     |
| KAZALD1      | rs807037   | Derived    | G      | 0.650                | Gly [G]         | 255                 |
|              |            | Ancestral  | C      | 0.350                | Ala [A]         |                     |
| TMEM86A      | rs7945285  | Derived    | T      | 1.000                | Val [V]         | 215                 |
|              |            | Ancestral  | C      | 0.000                | Ala [A]         |                     |
| FLJ38725     | rs3764147  | Derived    | G      | 0.246                | Val [V]         | 254                 |
|              |            | Ancestral  | A      | 0.754                | Ile [I]         |                     |
| PLA2G4B      | rs3816533  | Derived    | C      | 0.839                | Arg [R]         | 422                 |
|              |            | Ancestral  | T      | 0.161                | Cys [C]         |                     |
| USP8         | rs11638390 | Derived    | A      | 0.825                | Thr [T]         | 739                 |
|              |            | Ancestral  | G      | 0.175                | Ala [A]         |                     |
| MGC14151     | rs8522     | Derived    | T      | NA                   | Leu [L]         | 13                  |
|              |            | Ancestral  | C      | NA                   | Pro [P]         |                     |
| SCRN2        | rs17856536 | Derived    | G      | NA                   | Arg [R]         | 103                 |
|              |            | Ancestral  | A      | NA                   | Lys [K]         |                     |
|              | rs17856535 | Derived    | G      | NA                   | Gly [G]         | 411                 |
|              |            | Ancestral  | A      | NA                   | Ser [S]         |                     |
| STXBP4       | rs1156287  | Derived    | G      | NA                   | Gly [G]         | 92                  |
|              |            | Ancestral  | A      | NA                   | Arg [R]         |                     |
| ZNRF4        | rs17304380 | Derived    | A      | 0.142                | His [H]         | 163                 |
|              |            | Ancestral  | G      | 0.858                | Arg [R]         |                     |
|              | rs8103406  | Derived    | T      | 0.340                | Ser [S]         | 157                 |
|              |            | Ancestral  | G      | 0.660                | Ala [A]         |                     |
|              | rs2240744  | Derived    | A      | 0.150                | Gln [Q]         | 78                  |
|              |            | Ancestral  | G      | 0.850                | Arg [R]         |                     |
| ARTN         | rs2242637  | Derived    | A      | NA                   | Gln [Q]         | 19                  |
|              |            | Ancestral  | G      | NA                   | Arg [R]         |                     |
| SH2D2A       | rs926103   | Derived    | G      | 0.258                | Ser [S]         | 52                  |
|              |            | Ancestral  | A      | 0.742                | Asn [N]         |                     |
| ETNK2        | rs3737655  | Derived    | A      | NA                   | Gln [Q]         | 10                  |
|              |            | Ancestral  | C      | NA                   | Pro [P]         |                     |
| GTSE1        | rs6008600  | Derived    | A      | NA                   | Thr [T]         | 181                 |
|              |            | Ancestral  | G      | NA                   | Ala [A]         |                     |
| IL8RA        | rs16858811 | Derived    | T      | NA                   | Met [M]         | 31                  |
|              |            | Ancestral  | G      | NA                   | Arg [R]         |                     |
| DKFZP564J102 | rs4862653  | Derived    | A      | 0.083                | Lys [K]         | 146                 |
|              |            | Ancestral  | G      | 0.917                | Glu [E]         |                     |
|              | rs4862650  | Derived    | A      | 0.083                | Lys [K]         | 41                  |
|              |            | Ancestral  | G      | 0.917                | Glu [E]         |                     |
| FLJ23577     | rs6897513  | Derived    | A      | 0.517                | Asn [N]         | 71                  |
|              |            | Ancestral  | C      | 0.483                | His [H]         |                     |
| CAST         | rs754615   | Derived    | G      | 0.583                | Cys [C]         | 408                 |
|              |            | Ancestral  | C      | 0.417                | Ser [S]         |                     |

**Table 6.1** Twenty two fast evolving, brain expressed genes that segregate “derived”, “human specific” and “ancestral”, “chimpanzee specific” variants of these proteins in the extant human population (*continuation*)

| Gene      | rs #       | SNP status | Allele | Allele frequency CEU | Protein Residue | Amino Acid position |
|-----------|------------|------------|--------|----------------------|-----------------|---------------------|
| ADRB2     | rs1042714  | Derived    | C      | 0.533                | Gln [Q]         | 27                  |
|           |            | Ancestral  | G      | 0.467                | Glu [E]         |                     |
|           | rs1042713  | Derived    | A      | 0.325                | Arg [R]         | 16                  |
|           |            | Ancestral  | G      | 0.675                | Gly [G]         |                     |
| UBD       | rs2076487  | Derived    | C      | 1.000                | Ala [A]         | 99                  |
|           |            | Ancestral  | G      | 0.000                | Gly [G]         |                     |
|           | rs2076484  | Derived    | T      | 1.000                | Leu [L]         | 51                  |
|           |            | Ancestral  | C      | 0.000                | Ser [S]         |                     |
| LOC441376 | rs16889283 | Derived    | G      | NA                   | Gly [G]         | 96                  |
|           |            | Ancestral  | C      | NA                   | Arg [R]         |                     |
| ZFP37     | rs2282076  | Derived    | T      | 0.417                | Val [V]         | 7                   |
|           |            | Ancestral  | A      | 0.583                | Asp [D]         |                     |

**Table 6.2** Association study in two independent family based Dutch cohorts between three measures of cognitive ability and the ancestral vs. derived allele of a protein that undergone recent positive selection in primates

| Cohort | SNP       | Phenotype | N   | Stratification |         | Family-based |         | Population-based |         |
|--------|-----------|-----------|-----|----------------|---------|--------------|---------|------------------|---------|
|        |           |           |     | $\chi^2$       | p-value | $\chi^2$     | p-value | $\chi^2$         | p-value |
| Young  | rs1042713 | FSIQ      | 359 | 2.31           | 0.129   | 3.20         | 0.073   | 0.94             | 0.333   |
|        |           | VIQ       | 360 | 0.00           | 0.993   | 0.12         | 0.726   | 0.35             | 0.553   |
|        |           | PIQ       | 359 | 6.37           | 0.012   | 6.42         | 0.011   | 0.75             | 0.387   |
| Adult  | rs1042713 | FSIQ      | 350 | 0.60           | 0.438   | 1.20         | 0.273   | 0.62             | 0.432   |
|        |           | VIQ       | 350 | 0.10           | 0.756   | 0.28         | 0.597   | 0.21             | 0.646   |
|        |           | PIQ       | 350 | 1.05           | 0.305   | 1.35         | 0.246   | 0.33             | 0.567   |

## DISCUSSION

Ascertaining the genetic variants that affect complex traits has proven to be more difficult compared to detecting mutations for monogenetic disorders. Currently, genome-wide association studies are often considered the most promising search strategy, but require genotyping a very large number of neutral polymorphisms on very large sample sets. The cost of such an enterprise and the collection of a sufficiently large sample can still be prohibitive while the functional variant would still remain undetected. Alternative strategies to identify potentially functional genetic variants are therefore clearly needed. Because speciation involves phenotypic divergence in adaptive, usually complex, traits it is tempting to hypothesize that the genetic changes during speciation should involve loci that remain to control phenotypic variation in the relevant phenotypes within the newly emerged species.

Next to the obvious implications to evolutionary genetics, this assumption opens up the intriguing possibility to ascertain the genetic basis of phenotypic variation in humans from a comparative genomic analysis. An intrinsic advantage of such an approach is that genes to be revealed are bound to affect highly relevant traits, as inter species genetic divergence must relate to adaptive phenotypes that have been under recent selection, such as reproduction, immune system and behavioral traits. In this study we aimed to present a case study of this approach. It should be emphasized that any complex phenotype that has plausibly diverged between related species could be subject of a similar analysis. In this case the choice has been made to focus on a trait which undoubtedly has played an important role in becoming human. Our cognitive ability exceeds that of our closest primate relatives and some of the genetic changes during human speciation must have affected our intelligence. At the same time IQ is a highly heritable trait in humans with estimates of broad-sense heritability up to 80% (McGue M. 1993). Genetic linkage and association studies have shown before that segregating variants influence between individual differences in cognitive ability (Posthuma *et al.* 2005; Buyske *et al.* 2006; Luciano *et al.* 2006). As such, human intelligence is a good “candidate trait” to be investigated in the context of the present approach. Choosing a brain related phenotype also seems warranted by the vast overrepresentation of genes expressed in the central nervous system among those that exhibit accelerated protein evolution specific to the primate lineage. This result on its own confirms the common assumption that the phenotypic evolution of our brain has significantly coincided with human speciation (Dorus *et al.* 2004). Fourteen percent of the fast evolving, brain related human genes harbored potentially functional polymorphisms suggesting that on a genome wide scale at least several hundred such “derived” vs. “ancestral” variants should be present in the human genome at positively selected amino acid sites. All of these can be expected to affect phenotypic variation in an adaptive, complex trait.

Clearly, the choice of the right phenotype to test with the right candidate gene is crucial for this method to be useful from a “gene hunting” perspective. However, the present results suggest that this should be feasible in general. An educated guess on only one candidate gene resulted in detecting a significant genetic association with human intelligence. A huge advantage of selecting amino acid changing mutations as markers is that any significant association directly links the trait to a plausible causative variant.

The beta-adrenergic receptors belong to the G-protein-coupled receptor superfamily and mediate some of the physiological actions of catecholamines

(noradrenaline and adrenaline) in a variety of tissues (Liggett 2000). Different beta-receptor subtypes have been characterized (Liggett 2000). The beta-2 receptor is expressed in the smooth muscle of both the airways and blood vessels (to a lesser extent), and also can be found in the CNS (Hillman *et al.* 2005b). There is evidence that the beta adrenergic receptors might have a role in memory and learning formation. Noradrenalin (NA) which is the ligand for the adrenergic receptors exerts an ample range of functions in brain affecting cognition, behavior and emotion (Kobayashi and Kobayashi 2001). NA projections have been shown to extend to brain regions including hippocampus, amygdala, cerebral cortex and thalamus (Kobayashi and Kobayashi 2001). In particular, the hippocampus has been implicated in certain aspects of memory and in learning (Bliss and Collingridge 1993; Kobayashi and Kobayashi 2001). Furthermore, neurons from the hippocampus have been shown to exhibit activity dependent synaptic enhancement. This long-term potentiation (LTP) is the proposed model for memory and learning processes (Bliss and Collingridge 1993). These studies are supported by a few functional animal studies, in which deficits in memory in learning have been related to impaired NA synthesis (Kobayashi and Kobayashi 2001). In addition, animal studies, show that the activation of the beta-2 receptors induce LTP in neurons in hippocampus (Hillman *et al.* 2005a) and increase performance in tasks evaluating long-term memory and learning with the infusion of specific beta-2- agonists (Gibbs and Summers 2000), while beta-2- antagonists impaired memory consolidation (Gibbs and Summers 2005). More specifically, it has been shown that noradrenergic activation of the basolateral complex of the amygdala, involving  $\alpha$ -adrenoreceptors, selectively enhances memory consolidation for emotionally arousing experiences (Rooszendaal *et al.* 2006). The consensus of these animal studies is that higher beta-2 adrenergic receptor activity enhances memory and learning.

Genetic variation in the human beta-2 adrenergic receptor includes non-synonymous polymorphisms and the molecular function of the two missense mutations evaluated here has been studied before (Green *et al.* 1994). The substitutions of Gly for Arg at amino acid 16 (Arg16  $\rightarrow$  Gly; A $\rightarrow$ G in base pairs, rs1042713), Glu for Gln at amino acid 27 (Gln27 $\rightarrow$ Glu; C $\rightarrow$ G in base pairs, rs1042714), and a combination of both substitutions were considered. All three receptors variants displayed normal agonist binding and signaling activity. However, the two mutations differed markedly in the degree of agonist-promoted downregulation of receptor expression. The A $\rightarrow$ G, but not the C $\rightarrow$ G polymorphism enhanced the baseline level of isoproterenol induced down regulation of receptor density from 26% to 41%. Consequently, the Gly16 protein

(the G allele) has lower receptor density (Snyder *et al.* 2006) and presumably lower total signaling activity. Given that lower beta-2 adrenergic receptor activity seems to reduce performance in learning tasks, our finding that the G allele is the performance IQ decreaser allele might plausibly be explained by the effect of the mutation on beta-2 adrenergic receptor signaling. The other (Gln27→Glu; C→G) mutation did not enhance receptor down regulation. This polymorphism was included in the present study but because the genotypes were not in Hardy & Weinberg equilibrium the results of the association study were not reported. However, it is interesting to note that this SNP showed no evidence for association with IQ which might be explained by a lack of any effect of this mutation on the molecular function of the beta-2 adrenergic receptor.

The present approach involves a plausible assumption on the genetic architecture of adaptive traits, namely that the loci involved in the adaptive phenotypic divergence during speciation would be the same genes that affect variation in the phenotype within the new species. As such, the present results also provide a first attempt to merge genetic analysis at the level of macro- and microevolution. Focusing on non-synonymous variation at amino acid sites that differ between species in genes previously ascertained for increased rate of molecular divergence identifies a class of functional mutations that is bound to affect highly important traits in humans. The reproductive and immune systems as well as a broad range of behavioral phenotypes are thought to be under strong selection during speciation (Hughes 1997; Sterck *et al.* 1997; Torgerson *et al.* 2002). Any human (disease) phenotype related to these processes is likely to be controlled by genes that have been subject of recent directional selection. In this study we focused on an obvious “human specific” trait that has undergone strong phenotypic divergence during human speciation and is heritable in our species.

The comparative genomic analysis resulted in a set of candidate genes with an acceptable false discovery rate. Subsequent ascertainment of “ancestral” vs. “human specific” protein variants in genes expressed in the CNS resulted in only one gene to be followed up. The derived, human specific allele in the beta-2 adrenergic receptor was found to convey a significant increase in performance IQ. Compared to recent findings on the same cohort, the size of this effect is considerable (Gosso *et al.* 2006) and it might be explained by the reduction of agonist induced down-regulation of receptor density associated with the derived allele.

Considering both the evidence from animal studies on the function of the beta-2 adrenergic receptor and the evidence of the functional effect of the rs1042713 polymorphism, the identified genetic association is very plausible and therefore this study provides a strong case for finding genes for complex human traits starting from a comparative genomic analysis. This study demonstrates the feasibility to use interspecies comparisons to ascertain functional mutations that affect complex human traits. Given the increasing availability of full genome sequences of (closely) related species, similar approaches are bound to significantly contribute to the understanding of the genetic of complex traits in general and to the ascertainment of genes that affect traits with high societal relevance.

**Supplementary Table 6.1** Two hundred two genes with accelerated protein evolution in primates

| GENE ID                 | GENE NAME AND/OR BRIEF DESCRIPTION                          |
|-------------------------|---|
| OR9Q2                   | olfactory receptor, family 9, subfamily Q, member           |
| OR4D6                   | olfactory receptor, family 4, subfamily D, member           |
| OR4K14                  | olfactory receptor, family 4, subfamily K, member           |
| OR6K6                   | olfactory receptor, family 6, subfamily K, member           |
| OR2A12                  | olfactory receptor, family 2, subfamily A, member           |
| LOC441549               | similar to hypothetical protein 9330140G23                  |
| CDX1                    | caudal type homeo box transcription factor 1                |
| ELOVL3                  | elongation of very long chain fatty acids (FEN1/Elo2,       |
| LOC91561                | similar to ribosomal protein                                |
| PAMCI                   | peptidylglycine alpha-amidating monooxygenase COOH-terminal |
| PKLR                    | pyruvate kinase, liver and RBC                              |
| IL24                    | interleukin 24  |
| LOC440737               | similar to ribosomal protein                                |
| CD8B1                   | CD8 antigen, beta polypeptide 1 (p37)                       |
| GDDR                    | down-regulated in gastric cancer GDDR                       |
| MYL1                    | myosin, light polypeptide 1, alkali; skeletal,              |
| CYP39A1                 | cytochrome P450, family 39, subfamily A, polypeptide        |
| FRK                     | fyn-related kinase  |
| MCMDC1                  | minichromosome maintenance deficient domain containing      |
| SFTPC                   | surfactant, pulmonary-associated protein C                  |
| PRRX2                   | paired related homeobox 2                                   |
| TNMD                    | tenomodulin   |
| LOC347512               | similar to Heat shock transcription factor, Y-linked (Heat  |
| OR1N1                   | olfactory receptor, family 1, subfamily N, member           |
| OR52B4                  | olfactory receptor, family 52, subfamily B, member          |
| MRGPRD                  | MAS-related GPR, member D                                   |
| MCART2                  | mitochondrial carrier triple repeat 2                       |
| GUCA2B                  | guanylate cyclase activator 2B (uroguanylin)                |
| LOC391013               | similar to Group IIC secretory phospholipase A2 precursor   |
| OR5BF1                  | olfactory receptor, family 5, subfamily BF, member          |
| KRTAP15-1               | keratin associated protein 15-1                             |
| LOC344760               | similar to seven transmembrane helix receptor               |
| GSTA5                   | glutathione S-transferase A5                                |
| LOC442206               | similar to putative G-protein coupled receptor              |
| LOC402286               | similar to RIKEN cDNA 4930511M11                            |
| similar to FoxB2 rotein |   |
| FOXR2                   | forkhead box R2   |
| LOC392451               |   |
| FLJ25831                | HIN-6 protease  |
| CDX4                    | caudal type homeo box transcription factor 4                |
| CD40LG                  | CD40 ligand (TNF superfamily, member 5, hyper-IgM           |
| LHPP                    | phospholysine phosphohistidine inorganic pyrophosphate      |
| SLC43A1                 | solute carrier family 43, member 1                          |
| FLJ90119                | hypothetical protein FLJ90119                               |
| MRGX2                   | G protein-coupled receptor MRGX2                            |
| ZDHHC13                 |   |
| C11orf15                | chromosome 11 open reading frame 15                         |
| FLJ14966                | hypothetical protein FLJ14966                               |

**Supplementary Table 6.1** (*continuation*)

| GENE ID   | GENE NAME AND/OR BRIEF DESCRIPTION   |
|-----------|--|
| MS4A3     | membrane-spanning 4-domains, subfamily A, member 3   |
| FADS3     | fatty acid desaturase 3  |
| MAP4K2    | mitogen-activated protein kinase kinase kinase kinase  |
| PPP1R14B  | protein phosphatase 1, regulatory (inhibitor) subunit  |
| C12orf11  | chromosome 12 open reading frame 11  |
| MGC13183  | hypothetical protein MGC13183  |
| CHST11    | carbohydrate (chondroitin 4) sulfotransferase 11   |
| CACNB3    | calcium channel, voltage-dependent, beta 3 subunit   |
| SENP1     | SUMO1/sentrin specific protease 1  |
| STAC3     | SH3 and cysteine rich domain 3   |
| CDK4      | cyclin-dependent kinase 4  |
| LOC387870 | similar to protein tyrosine phosphatase, receptor type, Q; glomerular mesangial cell receptor protein-tyrosine phosphatase precursor |
| SGCG      | sarcoglycan, gamma (35kDa dystrophin-associated  |
| FLJ20449  | hypothetical protein FLJ20449  |
| FLJ38725  | hypothetical protein FLJ38725  |
| PPP1R3E   | protein phosphatase 1, regulatory (inhibitor) subunit  |
| C14orf104 | chromosome 14 open reading frame 104   |
| NGB       | neuroglobin  |
| ACYP1     | acylphosphatase 1, erythrocyte (common) type   |
| C14orf130 | chromosome 14 open reading frame 130   |
| PLA2G4B   | phospholipase A2, group IVB (cytosolic)  |
| USP8      | ubiquitin specific protease 8  |
| MNS1      | meiosis-specific nuclear structural protein 1  |
| CDR2      | cerebellar degeneration-related protein 2, 62kDa   |
| LOC283816 | similar to RIKEN cDNA 1810036I24   |
| E2F4      | E2F transcription factor 4, p107/p130-binding  |
| SNAI3     | snail homolog 3 (Drosophila)   |
| LOC388339 | similar to ribosomal protein   |
| MGC14151  | hypothetical protein MGC14151  |
| FAM64A    | family with sequence similarity 64, member A   |
| SCRN2     | secernin 2   |
| FLJ35808  | hypothetical protein FLJ35808  |
| STXBP4    | syntaxin binding protein 4   |
| SITPEC    | likely ortholog of mouse signaling intermediate in Toll  |
| JUNB      | jun B proto-oncogene   |
| LOC148066 | zinc and ring finger 4   |
| GAMT      | guanidinoacetate N-methyltransferase   |
| LENG1     | leukocyte receptor cluster (LRC) member 1  |
| ADSS      | adenylosuccinate synthase  |
| MAN1A2    | mannosidase, alpha, class 1A, member 2   |
| NEGR1     | neuronal growth regulator 1  |
| ARTN      | artemin  |
| FLJ11838  | hypothetical protein FLJ11838  |
| GNL2      | guanine nucleotide binding protein-like 2  |
| MGC4796   | Ser/Thr-like kinase  |
| C1QG      | complement component 1, q subcomponent, gamma  |
| TCEA3     | transcription elongation factor A (SII), 3   |



**Supplementary Table 6.1** (*continuation*)

| GENE ID   | GENE NAME AND/OR BRIEF DESCRIPTION                                   |
|-----------|--|
| C1orf33   | chromosome 1 open reading frame 33                                   |
| SH2D2A    | SH2 domain protein 2A  |
| CTMP      | C-terminal modulator protein   |
| NIT1      | nitrilase 1  |
| CD244     | CD244 natural killer cell receptor 2B4                               |
| C1orf9    | chromosome 1 open reading frame 9                                    |
| NMNAT2    | nicotinamide mononucleotide adenylyltransferase                      |
| PRDX6     | peroxiredoxin 6  |
| LGTN      | ligatin  |
| IPO9      | importin 9   |
| ETNK2     | ethanolamine kinase 2  |
| CGI-115   | CGI-115 protein  |
| TLR5      | toll-like receptor 5   |
| C20orf7   | chromosome 20 open reading frame 7                                   |
| ANKRD5    | ankyrin repeat domain 5  |
| PROCR     | protein C receptor, endothelial (EPCR)                               |
| DPM1      | dolichyl-phosphate mannosyltransferase polypeptide 1, catalytic      |
| FLJ20406  | Lck interacting transmembrane adaptor 1                              |
| C21orf70  | chromosome 21 open reading frame 70                                  |
| LOC150223 | hypothetical protein LOC150223                                       |
| HMG2L1    | high-mobility group protein 2-like 1                                 |
| SEPT3     | neuronal-specific septin 3   |
| ST13      | suppression of tumorigenicity 13 (colon carcinoma) (Hsp70)           |
| GTSE1     | G-2 and S-phase expressed 1  |
| ADRA2B    | adrenergic, alpha-2B-, receptor                                      |
| SLC3A1    | solute carrier family 3 (cystine, dibasic and neutral amino)         |
| LOC220717 | similar to ribosomal protein P0                                      |
| SFRS7     | splicing factor, arginine/serine-rich 7, 35kDa                       |
| KCNK3     | potassium channel, subfamily K, member 3                             |
| CHST10    | carbohydrate sulfotransferase 10                                     |
| ATP5G3    | ATP synthase, H <sup>+</sup> transporting, mitochondrial F0 complex, |
| HSPE1     | heat shock 10kDa protein 1 (chaperonin 10)                           |
| IL8RA     | interleukin 8 receptor, alpha  |
| ABHD5     | abhydrolase domain containing 5                                      |
| MLH1      | mutL homolog 1, colon cancer, nonpolyposis type 2 (E.                |
| ADPRTL3   | poly (ADP-ribose) polymerase family, member 3                        |
| GPR27     | G protein-coupled receptor 27  |
| ARPC4     | actin related protein 2/3 complex, subunit 4,                        |
| CPOX      | coproporphyrinogen oxidase   |
| PVRL3     | poliovirus receptor-related 3  |
| DIRC2     | disrupted in renal carcinoma 2                                       |
| GPR175    | G protein-coupled receptor 175                                       |
| GMPS      | guanine monophosphate synthetase                                     |
| MLF1      | myeloid leukemia factor 1  |
| TIPARP    | TCDD-inducible poly(ADP-ribose) polymerase                           |
| IRA1      | transducin (beta)-like 1X-linked receptor 1                          |
| MRPL47    | mitochondrial ribosomal protein L47                                  |
| ALB       | albumin  |

**Supplementary Table 6.1** (*continuation*)

| GENE ID          | GENE NAME AND/OR BRIEF DESCRIPTION                                  |
|------------------|---|
| TPARL            | TPA regulated locus   |
| MAP2K1IP1        | mitogen-activated protein kinase kinase 1 interacting protein       |
| SCYE1            | small inducible cytokine subfamily E, member 1 (endothelial         |
| DKFZP566M114     | hypothetical protein DKFZp566M114                                   |
| GPM6A            | glycoprotein M6A  |
| DKFZP564J102     | DKFZP564J102 protein  |
| HCN1             | hyperpolarization activated cyclic nucleotide-gated potassium       |
| FLJ25422         | hypothetical protein FLJ25422                                       |
| FLJ23577         | KPL2 protein  |
| HTR1A            | 5-hydroxytryptamine (serotonin) receptor 1A                         |
| LOC91942         | Myc-induced mitochondria protein                                    |
| POLK             | polymerase (DNA directed) kappa                                     |
| CAST             | calpastatin   |
| NDUFA2           | NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 2,              |
| PPP2CA           | protein phosphatase 2 (formerly 2A), catalytic subunit, alpha       |
| ADRB2            | adrenergic, beta-2-, receptor, surface                              |
| C6orf109         | chromosome 6 open reading frame 109                                 |
| UBD              | ubiquitin D   |
| E2F3             |   |
| LOC346171        | similar to bA145L22.2 (novel KRAB box containing C2H2 type          |
| CD83             | CD83 antigen (activated B lymphocytes, immunoglobulin               |
| EDN1             | endothelin 1  |
| CGA              | glycoprotein hormones, alpha polypeptide                            |
| SLC16A10         | solute carrier family 16 (monocarboxylic acid transporters), member |
| SMPDL3A          | sphingomyelin phosphodiesterase, acid-like 3A                       |
| BCLAF1           | BCL2-associated transcription factor 1                              |
| CBX3             | chromobox homolog 3 (HP1 gamma homolog,                             |
| INMT             | indolethylamine N-methyltransferase                                 |
| LOC402281        | similar to opposite strand transcription unit to Stag3; Gat         |
| CYP51A1          | cytochrome P450, family 51, subfamily A, polypeptide                |
| LOC441376        | similar to alanine and arginine rich domain containing              |
| SPATC1           | spermatogenesis and centriole associated                            |
| IL11RA           | interleukin 11 receptor, alpha                                      |
| C9orf24          | chromosome 9 open reading frame 24                                  |
| HNRPK            | heterogeneous nuclear ribonucleoprotein K                           |
| HBLD2            | HESB like domain containing 2                                       |
| TRIM14           | tripartite motif-containing 14                                      |
| ZFP37            | zinc finger protein 37 homolog (mouse)                              |
| ZMYND19          | zinc finger, MYND domain containing 19                              |
| UBE1             | ubiquitin-activating enzyme E1 (A1S9T and BN75 temperature          |
| MAGEB2           | melanoma antigen family B, 2  |
| NAP1L2           | nucleosome assembly protein 1-like 2                                |
| KIF4A            | kinesin family member 4A  |
| IL13RA2          | interleukin 13 receptor, alpha 2                                    |
| GLA gala         | ctosidase, alpha  |
| GASP2            | G protein-coupled receptor associated sorting protein               |
| TKTL1 transketol | ase-like 1  |
| FUNDC2           | FUN14 domain containing 2   |
| alpha NAC        |   |

## REFERENCES

- Abecasis GR, Cardon LR, Cookson WOC (2000) A general test of association for quantitative traits in nuclear families. *Am J Hum Genet* 66:279-292
- Bliss TV, Collingridge GL (1993) A synaptic model of memory: long-term potentiation in the hippocampus. *Nature* 361:31-9
- Buyske S, Bates ME, Gharani N, Matise TC, Tischfield JA, Manowitz P (2006) Cognitive traits link to human chromosomal regions. *Behav Genet* 36:65-76
- Cardon LR, Bell JI (2001) Association study designs for complex diseases. *Nat Rev Genet* 2:91-99
- Clark AG, Glanowski S, Nielsen R, Thomas PD, Kejariwal A, Todd MA, Tanenbaum DM, Civello D, Lu F, Murphy B, Ferreira S, Wang G, Zheng X, White TJ, Sninsky JJ, Adams MD, Cargill M (2003) Inferring Nonneutral Evolution from Human-Chimp-Mouse Orthologous Gene Trios. *Science* 302:1960-1963
- Dorus S, Vallender EJ, Evans PD, Anderson JR, Gilbert SL, Mahowald M, Wyckoff GJ, Malcom CM, Lahn BT (2004) Accelerated evolution of nervous system genes in the origin of Homo sapiens. *Cell* 119:1027-1040
- Gibbs ME, Summers RJ (2000) Separate roles for beta2- and beta3-adrenoceptors in memory consolidation. *Neuroscience* 95:913-22
- Gibbs ME, Summers RJ (2005) Contrasting roles for [beta]1, [beta]2 and [beta]3-adrenoceptors in memory formation in the chick. *Neuroscience* 131:31-42
- Gosso MF, de Geus EJC, van Belzen MJ, Polderman TJC, Heutink P, Boomsma DI, Posthuma D (2006) The SNAP-25 gene is associated with cognitive ability: evidence from a family-based study in two independent Dutch cohorts. *Mol Psychology* 11:878-886
- Green SA, Turki J, Innis M, Liggett SB (1994) Amino-terminal polymorphisms of the human beta 2-adrenergic receptor impart distinct agonist-promoted regulatory properties. *Biochemistry* 33:9414-9
- Hillman KL, Doze VA, Porter JE (2005a) Functional characterization of the beta-adrenergic receptor subtypes expressed by CA1 pyramidal cells in the rat hippocampus. *J Pharmacol Exp Ther* 314:561-7
- Hillman KL, Knudson CA, Carr PA, Doze VA, Porter JE (2005b) Adrenergic receptor characterization of CA1 hippocampal neurons using real time single cell RT-PCR. *Brain Res Mol Brain Res* 139:267-76
- Hughes A (1997) Rapid evolution of immunoglobulin superfamily C2 domains expressed in immune system cells. *Mol Biol Evol* 14:1-5
- Khaitovich P, Enard W, Lachmann M, Paabo S (2006) Evolution of primate gene expression. *Nat Rev Genet* 7:693-702
- Kobayashi K, Kobayashi T (2001) Genetic evidence for noradrenergic control of long-term memory consolidation. *Brain Dev* 23 Suppl 1:S16-23
- Liggett SB (2000) Pharmacogenetics of beta-1- and beta-2-adrenergic receptors. *Pharmacology* 61:167-73
- Luciano M, Wright MJ, Duffy DL, Wainwright MA, Zhu G, Evans DM, Geffen GM, Montgomery GW, Martin NG (2006) Genome-wide scan of IQ finds significant linkage to a quantitative trait locus on 2q. *Behav Genet* 36:45-55
- MacCallum C, Hill E (2006) Being positive about selection. *PloS Biol* 4:293-295
- McGue M, BTJ, Jr., Iacono, W.G., & Lykken, D.T. (1993) Behavioral genetics of cognitive ability: a life-span perspective. In: Plomin R, MGE (ed) *Nature, Nurture and Psychology*. American Psychological Association, Washington DC, pp 59-76
- Nielsen R, Bustamante C, Clark AG, Glanowski S, Sackton TB, Hubisz MJ, Fledel-Alon A, Tanenbaum DM, Civello D, White TJ, J. Sninsky J, Adams MD, Cargill M (2005) A Scan for Positively Selected Genes in the Genomes of Humans and Chimpanzees. *PLoS Biology* 3:e170
- Pollard KS, Salama SR, Lambert N, Lambot M-A, Coppens S, Pedersen JS, Katzman S, King B, Onodera C, Siepel A, Kern AD, Dehay C, Igel H, Ares M, Vanderhaeghen P, Haussler D (2006) An RNA gene expressed during cortical development evolved rapidly in humans. *Nature* 443: 167-172
- Posthuma D, Luciano M, Geus EJCd, Wright MJ, Slagboom PE, Montgomery GW, Boomsma DI, Martin NG (2005) A Genomewide Scan for Intelligence Identifies Quantitative Trait Loci on 2q and 6p. *Am J Hum Genet* 77:318-326
- Roozendaal B, Okuda S, Van der Zee EA, McGaugh JL (2006) Glucocorticoid enhancement of memory requires arousal-induced noradrenergic activation in the basolateral amygdala. *Proc Natl Acad Sci USA* 103:6741-6746
- Snyder EM, Hulsebus ML, Turner ST, Joyner MJ, Johnson BD (2006) Genotype related differences in beta(2) adrenergic receptor density and cardiac function. *Med Sci Sports Exerc* 38:882-886
- Sterck EHM, Watts DP, van Schaik CP (1997) The evolution of female social relationships in nonhuman primates *Behavioral ecology and sociobiology*. 41: 291-309
- Torgerson DG, Kulathinal RJ, Singh RS (2002) Mammalian Sperm Proteins Are Rapidly Evolving: Evidence of Positive Selection in Functionally Diverse Genes. *Mol Biol Evol* 19:1973-1980
- Weir BS (1983) *Statistical Analysis of DNA Sequence Data*. Dekker, New York
- Yang ZH (1997) PAML: a program package for phylogenetic analysis by maximum likelihood. *Computer Applications in Biosciences* 13:555-556



# CHAPTER 7

*COMT* and *DRD2* gene variants:  
evidence of positive heterosis and gene-gene  
interaction on working memory functioning

This chapter is under revision as:

MF Gosso, EJC de Geus, TJC Polderman, DI Boomsma, P Heutink, and D Posthuma. Catechol O-methyl transferase and dopamine D2 receptor gene polymorphisms: evidence of positive heterosis and gene-gene interaction on working memory functioning. *EJHG*, *under revision*.

*Abstract*

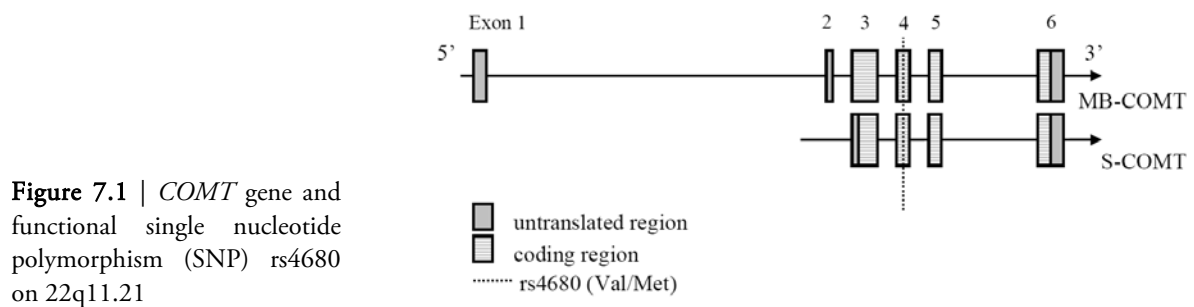
A common functional single nucleotide polymorphism (rs4680) that substitutes a Valine for a Metionine residue at codon 158 in the catechol-O-methyltransferase (*COMT*) gene has been extensively studied in relation to individual differences in working memory (WM) performance. The present study tested the association of the *COMT* Val<sup>108/158</sup>Met polymorphism with WM performance in two independent family based Dutch samples: 371 children (mean age 12.4) and 391 adults (mean age 36.2). A significant association was found between *COMT* polymorphism and WM scores in the *adult* but not in the *young* cohort. The association reflected positive heterosis such that the Met/Met and Val/Val homozygotes performed less well than the Met/Val heterozygotes on the WM tasks. A secondary analysis was conducted in which a *DRD2* tagging SNP (rs2075654) was tested for an interactive effect with the *COMT* polymorphism on WM performance. A significant interactive effect of the *DRD2* and *COMT* genes was found in the *adult* cohort, such that heterosis was present only in the *DRD2* genotype that has been linked to lower receptor density. Our results support previous findings that working memory performance needs an optimal level of dopamine signaling within the PFC. This optimum depends on enzymatic activity controlling dopamine level as well as dopamine receptor sensitivity, both of which may differ as a function of age and genotype. As a consequence the effects of a single polymorphism in a dopaminergic gene on a well-defined cognitive trait may easily remain hidden if the interaction with age and other genes in the pathway are not taken into account.

## INTRODUCTION

Working memory (WM) is a fundamental component of human intelligence (Colom *et al.* 2004; Conway *et al.* 2002). It refers to those processes that support the short-term maintenance or manipulation of relevant information in the presence of distracting irrelevant information. Nonhuman primates physiology and human functional imaging studies support a critical involvement of the prefrontal cortex (PFC) in WM (Curtis & D'Esposito 2003; Fuster 1997). The PFC is the primary target of extensive dopamine (DA) projections from the midbrain and several lines of evidence suggest that the frontal DA level is a critical modulator of WM performance (Conway *et al.* 2002; Gibbs & D'Esposito 2005). Hence, genes involved in dopaminergic pathway metabolism have been of particular interest to explain individual differences in WM performance (Brunner *et al.* 1993; Harmer *et al.* 2001; Malhotra *et al.* 2002; Reuter & Hennig 2005; Savitz *et al.* 2006; Williams & Castner 2006). Amongst these, the catechol-O-methyltransferase (*COMT*) gene has been studied the most extensively.

The *COMT* gene is located on chromosome 22q11, and contains six exons (Grossman *et al.* 1992). It is involved in enzymatic activity that degrades DA, norepinephrine, and epinephrine (Axelrod 1957). Two promoters encoding different isoforms, a membrane-bound *COMT* (MB-*COMT*) and a soluble *COMT* (S-*COMT*) are known. Both transcripts start at exon 3 (Tenhunen *et al.* 1994). The two *COMT* isoforms differ in their kinetic characteristics; MB-*COMT* presents a higher substrate affinity in combination with a lower catalytic activity than its soluble S-*COMT* counterpart (Lotta *et al.* 1995). Differential expression and activity profiles are also well characterized, MB-*COMT* is predominantly expressed in brain neurons (Matsumoto *et al.* 2003; Reenilä & Mannisto 2001), whereas S-*COMT* is predominantly expressed in other tissues, such as liver, blood and kidney (Hong *et al.* 1998; Lundstrom *et al.* 1995; Tenhunen *et al.* 1994).

The human MB-*COMT* plays a crucial role in regulation of dopamine signaling at the PFC level. It contains a common functional single nucleotide polymorphism (SNP) (rs4680) that substitutes a Valine for a Metionine residue at codon 158 (see Figure 7.1).

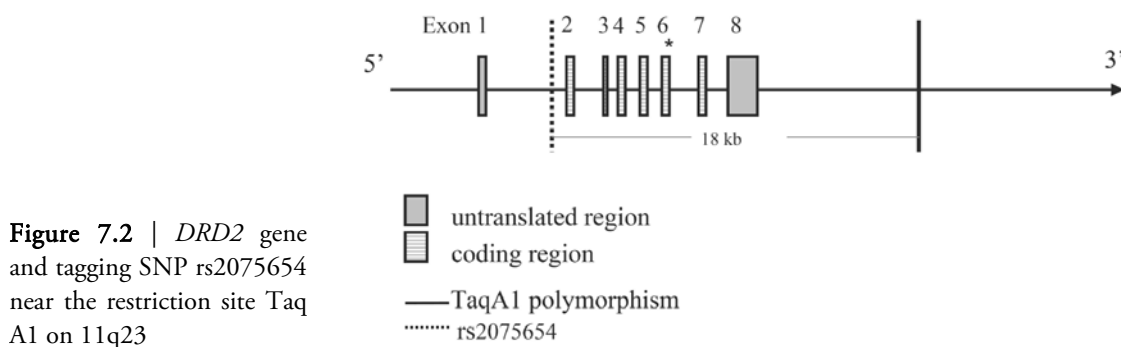


The *Met* allele encodes an enzyme isoform with relatively lower activity (Lotta *et al.* 1995), and is thought to be specific to humans; since no equivalent polymorphism has been found in any other species (Palmatier *et al.* 1999). Decrease in enzyme activity present in *Met/Met* individuals, compared to individuals homozygous for the *Val* allele, leads to a relatively higher DA level, whereas *Met/Val* heterozygous display an intermediate enzyme activity (Boudikova *et al.* 1990; Weinshilboum & Raymond 1977). Because frontal DA level is a critical modulator of WM processes, the decreased *COMT* activity of *Met* carriers might be beneficial to their cognitive performance. In line with this, Savitz *et al.* (2006) found that 20 of the 26 studies on the association between the *COMT* Val<sup>108/158</sup>*Met* polymorphism and cognitive function reported a significant association. All but two of these studies suggested that the low-activity *Met* allele yields better performance on cognitive tasks that have a working memory component. However, these studies were often based on small and/or clinical samples (ADHD in children; schizophrenia in adults) and no significant association to WM was found in a much larger sample of healthy adult males (Stefanis *et al.* 2004).

However, Mattay and colleagues (2003), have convincingly shown that the role of the *COMT* Val<sup>108/158</sup>*Met* polymorphism in PFC function, particularly in WM performance may be less straightforward. Although homozygous individuals for the *Met* allele perform significantly better than individuals homozygous for the *Val* allele, when the *Met* homozygotes are given DA agonists their response actually deteriorates. In contrast, the response of the ‘dopamine-poor’ *Val* homozygotes improves with DA agonists. This suggests that the relation between DA availability at the PFC level as indexed by *COMT* activity and WM performance is not linear, but instead follows an inverted-U shape. Dopamine signaling, furthermore, is not only dependent on the availability of dopamine, but also on the efficiency of the dopamine receptor and its downstream signaling cascade. Because of their



importance in reward processing, dopaminergic receptors, in particular the dopamine D2 receptor gene (*DRD2*), have been studied extensively in addiction research (Comings & Blum 2000). The *DRD2* gene is located on chromosome 11 at q22-q23 (see Figure 7.2). A *DRD2* *Taq* IA variant, a restriction fragment length polymorphism (RFLP), located on the 3'-untranslated region (3'UTR) of the *DRD2* gene, is associated with altered receptor density (Thompson *et al.* 1997). Individuals with the A1 allele show a 30-40% reduction in D2 dopamine receptor density compared with those homozygous for the A2 allele (Jonsson *et al.* 1999; Pohjalainen *et al.* 1998; Ritchie & Noble 1996, 2003; Thompson *et al.* 1997).



**Figure 7.2** | *DRD2* gene and tagging SNP rs2075654 near the restriction site *Taq* A1 on 11q23

Recently, Reuter *et al.* (2005) conducted an association analysis using an adult cohort enriched for *COMT* and *DRD2* homozygotes. They found a significant interaction between *DRD2* *Taq* IA, and *COMT* Val<sup>108/158</sup>Met polymorphisms and performance on response interference on the Stroop color-word conflict task. *Met* homozygotes performed better than *Val* allele carriers, but only if they had the *DRD2* genotype associated with low receptor density. In fact, *Met* homozygotes also bearing two *DRD2* A2 alleles showed a significantly worse performance compared to all other genotypes. Although response interference and working memory are not unitary constructs, we showed a significant correlation ( $r = -0.26$ ,  $p < 0.05$ ) between these two measures of PFC function (Stins *et al.* 2005). Hence, we hypothesize that a *COMT* by *DRD2* interaction may also be found for WM performance.

In the present study, which included 762 genotyped subjects, from two independent family based Dutch samples of 371 (mean age 12.4) and 391 (mean age 36.2) subjects respectively, our principal goal was to test for association of the *COMT* Val<sup>108/158</sup>Met polymorphism with WM performance. The use of a family based sample made it possible to test for association in a combined within- and between-family design to estimate genetic effects which are free from spurious

effects of population stratification (Abecasis *et al.* 2000; Boomsma 1998; Fulker *et al.* 1999; Laird & Lange 2006; Posthuma *et al.* 2004). As a secondary analysis we tested for an interactive effect of the *COMT* Val<sup>108/158</sup>Met polymorphism and genetic variation in the *DRD2* gene on WM performance comparable to the effect reported by Reuter *et al.* (2005) for Stroop interference. Because dopamine receptor sensitivity has been shown to decline with ageing in both animal (Lee *et al.* 2001) and human studies (Kaasinen & Rinne 2002) our analyses will allow this interaction to be different in children and adults.

## MATERIALS AND METHODS

### *Subjects*

All twins and their siblings were part of two larger cognitive studies and were recruited from the Netherlands Twin Registry (Boomsma 1998; Boomsma *et al.* 2006). Informed consent was obtained from the participants (adult cohort) or from their parents if they were under 18 (young cohort). The current study was approved by the institutional review board of the VU University Medical Center. None of the individuals tested suffered from severe physical or mental handicaps, as assessed through standard questionnaire.

### *Young Cohort*

The young cohort consisted of 177 twin pairs born between 1990 and 1992, and 55 siblings (Polderman *et al.* 2006a; 2006b), of which 371 were available for genotyping. The genotyped twins were 12.4 (SD= 0.9) years of age and the siblings were between 8 and 15 years old at the time of testing. There were 35 monozygotic male twin pairs (MZM), 28 dizygotic male twin pairs (DZM), 48 monozygotic female twin pairs (MZF), 23 dizygotic female twin pairs (DZF), 26 dizygotic opposite-sex twin pairs (DOS), 24 male siblings and 24 female siblings, and 3 subjects form incomplete twin pairs (1 male, 2 females). Participation in this study included a voluntary agreement to provide buccal swabs for DNA extraction.

### *Adult Cohort*

A total of 793 family members from 317 extended twin families participated in the adult cognition study (Posthuma *et al.* 2005). Participation in this study did not automatically include DNA collection, however, part of the sample, 276 returned to the lab to provide blood samples, 115 provided buccal swabs via the NTR Biobank project (Hoekstra *et al.* 2004) for DNA extraction. Mean age of the

genotyped sample was 36.2 years (SD= 12.6). There were 25 monozygotic male twin pairs (MZM), 15 dizygotic male twin pairs (DZM), 1 DZM triplet, 20 monozygotic female twin pairs (MZF), 28 dizygotic female twin pairs (DZF) and 23 dizygotic opposite-sex twin pairs (DOS), 29 female siblings and 28 male siblings, and 109 subjects from incomplete twin pairs (41 males, 68 females).

### *Cognitive testing*

WM tasks were assessed in the young cohort, using the Dutch adaptation of the Wechsler Intelligence Scale for Children-Revised (WISC-R) (Wechsler 1986) consisting of two subtests: arithmetic, and digit span. Arithmetic involves numerical reasoning, attention and short-term memory for meaningful information. Unlike paper-and-pencil tests of arithmetic problem solving, this test measures verbal mathematical reasoning skills by giving the child oral problems to solve. Digit Span measures short-term auditory memory for non-meaningful information and concentration. It measures a child's ability to remember a sequence of numbers (both backwards and forwards). WM performance was indexed as the sum score of the two subtests and corrected for age and sex. The Dutch adaptation of the Wechsler Adult Intelligence Scale III-Revised (WAIS-III) (Wechsler 1997) was used to assess WM performance in the adult cohort and consisted of two subtests taxing WM (arithmetic, letter-number sequencing). The *Arithmetic* subtest is similar to the *Arithmetic* subtest in the WISC, except that it is adjusted for an adult age. In *Letter-number sequencing*, subjects have to repeat a series of alternate letters and numbers the examiner reads aloud. Numbers should be given first than letters, in numerical/alphabetical order. *Letter-number sequencing* and *Digit Span* are thought to measure the same underlying construct and can replace each other in the calculation of the WM dimension, following the WAIS guidelines (Wechsler 1997). WM was indexed as the sum score of *Arithmetic* and *Letter-number sequencing* and corrected for age and sex.

### *DNA collection and Genotyping*

DNA isolation from buccal swabs was performed using a chloroform/isopropanol extraction (Meulenbelt *et al.* 1995). DNA was extracted from blood samples using the salting out protocol (Miller *et al.* 1988). Zygosity was assessed using 11 polymorphic microsatellite markers (Het > 0.80). Genotyping was performed blind to familial status and phenotypic data. Both MZ twins of a pair were included, serving as additional quality control on genotyping. *COMT* genotyping was performed using fluorogenic probes in the high-throughput 5' nuclease assay (TaqMan, PE Applied Biosystems, Foster city, CA), which

combines polymerase chain reaction amplification and detection into a single step. The assay requires two allele-specific probes, which labelled two alleles with different fluorescent reporter dyes for discrimination. Following allele-specific hybridization, the detection probe is cleaved during each amplification cycle by the 5'-exonuclease activity of Taq DNA polymerase if the probe's target sequence is present.

For *DRD2*, instead of the A1 allele of the Taq IA polymorphism, a *tag*-SNP (rs2075654) lying 18 kb downstream of the Taq IA variant was genotyped. In view of the LD between the tag-SNP and *Taq* IA ( $r^2 = 0.65$ ) we will refer to the T allele as 'A1'. *DRD2* SNP genotyping was performed as part of a SNPLEX assay, that included multiple other genes, following a *tagging* approach (Hirschhorn & Daly 2005). The SNPLEX assay was conducted following the manufacturer's recommendations (Applied Biosystems, Foster city, CA, USA). Here we focus on the *DRD2* gene only. Results on cognitive effects of two other genes are described elsewhere (Gosso *et al.* 2006a; 2006b).

### *Statistical analyses*

Allele frequencies of the *COMT* Val<sup>108/158</sup>Met and *DRD2* A1/A2 polymorphisms were estimated in both the *young* and the *adult* cohorts using Haploview (<http://www.broad.mit.edu/mpg/haploview>) in which a Hardy-Weinberg test is implemented, based on an exact calculation of the probability of observing a certain number of heterozygotes conditional on the number of copies of the minor SNP allele. Family based genetic association tests were conducted using the program QTDT (<http://www.sph.umich.edu/csg/abecasis/QTDT/>) which implements the orthogonal association model proposed by Abecasis *et al.*, 2000 (see also Fulker *et al.*, 1999; extended by Posthuma *et al.*, 2004) This model allows one the decomposition of the genotypic association effect into orthogonal between- ( $\beta_b$ ) and within- ( $\beta_w$ ) family components, and can incorporate fixed effects of covariates and can also model the residual sib-correlation as a function of polygenic or environmental factors. MZ twins can be included and are modeled as such, by adding zygosity status to the datafile. They are not informative to the within-family association component (unless they are paired with non-twin siblings), but are informative for the between family component. The between-family association component is sensitive to population admixture, whereas the within-family component is significant only in the presence of true association. If population stratification acts to create a false association, the test for association using the within family component is still valid, and provides a conservative test of association. Testing for the equality of the  $\beta_b$  and  $\beta_w$  effects, serves as a test of population stratification. If this test is not significant, the between and within

family effects are equal and the more powerful association test that uses the whole population at once can be applied. Both main genetic effects and interaction effects on the trait values were evaluated by taking into account the dependency of the data. The residual sib correlation was modeled as a function of residual genetic variance and non-shared environmental variance. The DRD2 genotype was recoded into carriers (A1+) versus non-carriers (A1 -). We used one-sided hypothesis testing for the interaction effects since our hypotheses specify the direction of genetic effects. The two-locus LD calculator program - 2LD - (Zhao 2004) was used to calculate the genotype distribution for both gene loci in the *young* and *adult* cohort.

## RESULTS

In total 762 subjects were available for SNP genotyping. Based on blind controls and intrapair MZ comparisons a low genotyping error rate was found (0.015%). For the total sample, the success rate was 98.5% and 77.1% for the *COMT* Val<sup>108/158</sup>Met polymorphism in the young and adult cohort, respectively. For *DRD2* rs2075654, success rates were 97%% and 100%. 301 adults and 360 children had genotype data for both *COMT* and *DRD2*. The distribution of genotype and allele frequencies of the *COMT* and *DRD2* polymorphisms are shown in columns 3 and 4 of Table 7.1. Means, standard deviations, and standard errors for WM performance are provided in columns 5 to 7, for the full sample, the total genotyped sample and each of the *COMT* and *DRD2* genotype groups. Two-locus LD calculation showed that the *COMT* and *DRD2* genotypes were not linked among the *young* ( $\chi^2 = 0.08$ ,  $p = 0.783$ ), and *adult* cohort ( $\chi^2 = 1.17$ ,  $p = 0.279$ ), respectively.

### *Stratification*

Tests for the presence of population stratification were not significant at the 0.05 level, indicating that genotypic effects within families were not significantly different from those observed between families, suggesting that the more powerful population-based association test can be meaningfully interpreted for both *COMT* and *DRD2*.

### *COMT polymorphism*

WM performance was plotted against the three possible genotype groups (Met/Met, Met/Val, and Val/Val) suggesting a positive heterosis pattern in both the *young* and the *adult cohort* (see Figure 7.3). Heterosis refers to a situation in which a given trait is significantly greater (or lesser) in individuals heterozygous at a specific gene marker than those homozygous for either allele. We tested for heterosis by adding a non-additive (dominance) genetic component to the population based analysis in QTDT (see Table 7.2a). A significant heterosis effect was found for the association between *COMT* Val<sup>108/158</sup>Met polymorphism and WM in the *adult cohort* ( $\chi^2 = 5.02$ ,  $p = 0.025$ ) even after correction for multiple testing. In the *young cohort* no-significant association was found for heterosis ( $\chi^2 = 1.54$ ,  $p = 0.215$ ), although the effects were in the same direction as in the adult cohort and the combined sample showed the strongest effect ( $\chi^2 = 6.30$ ,  $p = 0.012$ ).

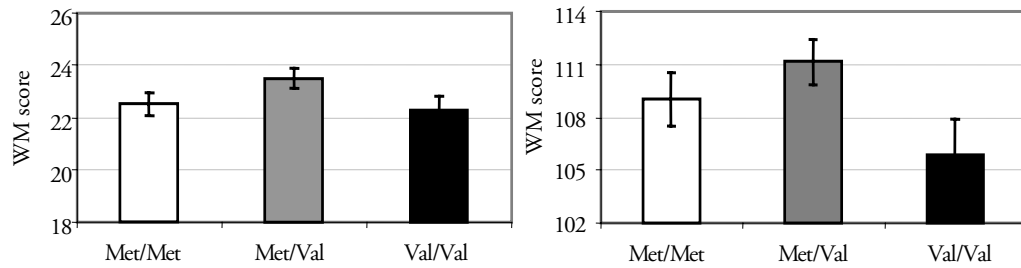
**Table 7.1** Means and standard deviations WM scores for *COMT* and *DRD2* genotypes

| Cohort           | Genotype | N   | Frequency | WM Mean | Std Deviation | Std Error of Mean |
|------------------|----------|-----|-----------|---------|---------------|-------------------|
| <i>Young</i>     |          |     |           |         |               |                   |
| Total sample     |          | 407 |           | 22.88   | 4.85          | 0.24              |
| Genotyped sample |          | 371 |           | 22.90   | 4.89          | 0.25              |
| <i>COMT</i>      | Met/Met  | 117 | 0.32      | 22.51   | 4.96          | 0.46              |
| N=365            | Met/Val  | 175 | 0.48      | 23.49   | 5.06          | 0.38              |
|                  | Val/Val  | 73  | 0.20      | 22.27   | 4.33          | 0.51              |
| <i>DRD2</i>      | A1A1     | 9   | 0.03      | 23.98   | 2.08          | 0.69              |
| N=360            | A1A2     | 94  | 0.26      | 23.12   | 5.10          | 0.53              |
|                  | A2A2     | 257 | 0.71      | 22.78   | 4.88          | 0.30              |
| <i>Adult</i>     |          |     |           |         |               |                   |
| Total sample     |          | 793 |           | 109.91  | 15.89         | 0.56              |
| Genotyped sample |          | 391 |           | 109.92  | 15.80         | 0.80              |
| <i>COMT</i>      | Met/Met  | 100 | 0.33      | 109.04  | 15.34         | 1.53              |
| N=301            | Met/Val  | 138 | 0.46      | 111.15  | 15.24         | 1.30              |
|                  | Val/Val  | 63  | 0.21      | 105.86  | 16.09         | 2.03              |
| <i>DRD2</i>      | A1A1     | 16  | 0.04      | 104.96  | 14.66         | 3.67              |
| N=391            | A1A2     | 109 | 0.27      | 110.76  | 14.25         | 1.36              |
|                  | A2A2     | 266 | 0.69      | 110.36  | 16.32         | 1.00              |

*Note:* the WM score of children and adults are based on different subtests. Although these index the same theoretical construct, the WM scores cannot be compared directly across age cohorts.

### *DRD2 polymorphism*

No significant association was found between the *DRD2* rs2075654 *tagging*-SNP and WM performance (see Table 7.2b).



**Figure 7.3** | Working memory (WM) means plotted for the three genotypes of the *COMT* Val<sup>108/156</sup>Met polymorphism for *young* (left panel) and *adult* (right panel) cohort. Bars denote standard error.

**Table 7.2a** Results of genetic association analysis for the *COMT* gene

| Cohort       | Model                  | -2LnLk  | df  | vs model | Δdf | χ <sup>2</sup> | p-value |
|--------------|------------------------|---------|-----|----------|-----|----------------|---------|
| Young Sample | 1. Dominance+ Additive | 977.72  | 360 |          |     |                |         |
|              | 2. Additive            | 979.26  | 361 | 1        | 1   | 1.54           | 0.215   |
|              | 3. No Genetic effect   | 979.52  | 362 | 2        | 1   | 0.24           | 0.621   |
| Adult Sample | 1. Dominance+ Additive | 802.48  | 296 |          |     |                |         |
|              | 2. Additive            | 807.50  | 297 | 1        | 1   | 5.02           | 0.025   |
|              | 3. No Genetic effect   | 808.62  | 298 | 2        | 1   | 1.13           | 0.287   |
| Combined     | 1. Dominance+ Additive | 1783.34 | 661 |          |     |                |         |
|              | 2. Additive            | 1789.64 | 662 | 1        | 1   | 6.30           | 0.012   |
|              | 3. No Genetic effect   | 1790.76 | 663 | 2        | 1   | 1.11           | 0.292   |

**Table 7.2b** Results of genetic association analysis for the *DRD2* gene

| Cohort       | Model      | χ <sup>2</sup> (1) | p-value |
|--------------|------------|--------------------|---------|
| Young Sample | A1+ vs A1- | 0.42               | 0.517   |
| Adult Sample | A1+ vs A1- | 0.04               | 0.842   |
| Combined     | A1+ vs A1- | 0.10               | 0.752   |

*Note:* For *DRD2* it was tested whether carriers of the A1 + allele and non-carriers showed different trait value.

### *COMT and DRD2 interaction*

As population stratification was not significant for the main effects of *COMT* and *DRD2*, suggesting the population based association is valid - we conducted a population based test for the interaction effects as well. For this purpose, two dummy variables were created; an additive *COMT* x carrier vs

noncarrier *DRD2* effect (Met/Met;A-, Val/Met;A-, Val/Val;A- , Val/Met;A+ = 0; Met/Met;A+ = -1 ; Val/Val;A+ = 1) and a dominance deviation of *COMT* x carrier vs noncarrier *DRD2* effect (Met/Met;A-, Val/Met;A-, Val/Val;A-, Met/Met;A+ , Val/Val;A+ = 0; Val/Met;A+ = 1). The dummy variables were analysed in QTDIT using the -cu option. Figure 7.4 plots WM performance against six possible combined genotype groups (Met/Met;A-, Val/Met;A-, Val/Val;A-; Met/Met;A+, Val/Met;A+, Val/Val;A+). The figure suggests that the heterosis found for the *COMT* gene is entirely limited to subjects with the *DRD2* A+ genotype, the genotype previously linked to reduced receptor density. In support of this, a significant interaction effect was detected between the *DRD2* and *COMT* polymorphisms in the combined sample ( $p = 0.042$ ), which seemed confined to the *adult cohort* (see Table 7.3).

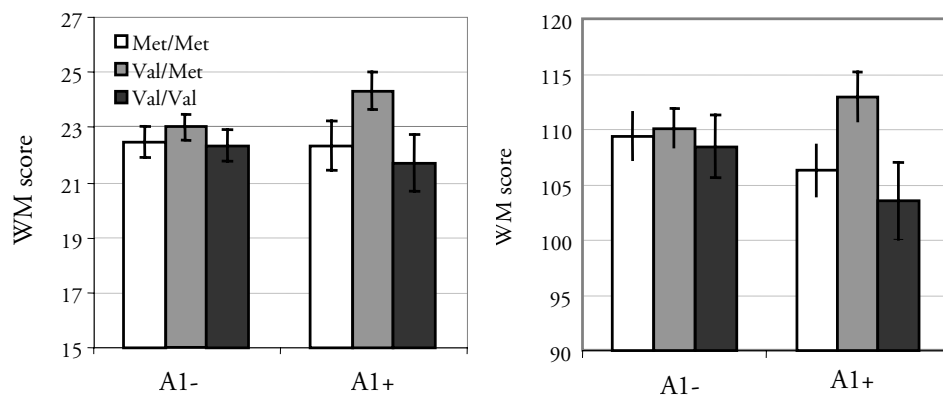
**Table 7.3** Results of population-based gene-gene interaction analysis – for *COMT* al<sup>108/158</sup>Met, and *DRD2* rs2075654 the *adult* and *young cohorts*

| Cohort | Model   | -2LnLk  | df  | vs<br>model | $\chi^2$ | p-value* |
|--------|---|---------|-----|-------------|----------|----------|
| Young  | 1. Full model   | 952.64  | 350 |             |          |          |
|        | 2. no dominance interaction <i>COMT</i> x <i>DRD2</i>             | 953.70  | 351 | 1           | 1.06     | 0.152    |
|        | 3. no dominance or additive interaction <i>COMT</i> x <i>DRD2</i> | 954.32  | 352 | 2           | 0.62     | 0.216    |
| Adult  | 1. Full model   | 757.94  | 280 |             |          |          |
|        | 2. no dominance interaction <i>COMT</i> x <i>DRD2</i>             | 759.60  | 281 | 1           | 1.66     | 0.099    |
|        | 3. no dominance or additive interaction <i>COMT</i> x <i>DRD2</i> | 759.70  | 282 | 2           | 0.10     | 0.376    |
| All    | 1. Full model   | 1714.04 | 638 |             |          |          |
|        | 2. no dominance interaction <i>COMT</i> x <i>DRD2</i>             | 1717.02 | 639 | 1           | 2.98     | 0.042    |
|        | 3. no dominance or additive interaction <i>COMT</i> x <i>DRD2</i> | 1717.74 | 640 | 2           | 0.72     | 0.198    |

\*one-sided p-value

*Note:* Model 1. = full model, includes a grand mean, main additive and dominance effects of *COMT*, a main effect of A1 carriers vs. non carriers of *DRD2*, and additive and dominance interaction of *COMT* genotypes X *DRD2* A1 carriers vs. non-carriers. Model 2. = model 1, except the dominance interaction. Model 3. = model 2, except the additive interaction.





**Figure 7.4** | Working memory (WM) means plotted for the six combined genotype groups of the *COMT* and *DRD2* genes for *young* (left panel) and *adult* (right panel) cohorts. Bars denote standard error.

*Note:* Due to relative small sample size, the less frequent homozygous genotypes (A1/A1) were pooled with the heterozygous A1/A2. A+ denotes the combined A1/A1 and A1/A2 genotypes; A- denotes the A2/A2 genotype.

## DISCUSSION

In the present study, we tested the association of the *COMT* Val<sup>108/158</sup>Met polymorphism with WM performance. A significant association was found in the *adult* but not in the *young sample*. The association reflected positive heterosis such that the *Met/Val* heterozygotes performed better than both *Met/Met* and *Val/Val* homozygotes on the WM tasks. An age-dependent positive heterosis pattern has previously been reported in a longitudinal study by Harris and colleagues (2005). The *COMT* genotype was not associated with childhood intelligence measured at age 11 in the Scottish Mental Survey of 1932. At age 79, *COMT* genotype was significantly related to differences in verbal declarative memory and to scores on the personality traits of intellect/imagination. For both traits the elderly Val/Met heterozygotes had higher scores than both homozygous groups, which echoes the pattern of heterosis on WM found in our adult sample. Because the *COMT* polymorphism has been hypothesized to have a non-linear effect on DA availability in the prefrontal cortex (Mattay *et al.* 2003) the finding of heterosis is in keeping with the idea that the relationship between DA signaling and cognitive performance follows an inverted U-shaped curve, with both suboptimal and supra-optimal DA activity impairing prefrontal function (Cools & Robbins 2004). Burst firing of VTA neurons causes synaptic DA release into pyramidal cells in the PFC. Because these cells contain little dopamine transporter (*DAT*), most DA diffuses out of the synaptic cleft to bind to extrasynaptic D1 receptors where it is inactivated by *COMT* (Bilder *et al.* 2004). The higher activity Val allele decreases extrasynaptic DA levels and therefore *D1* activation, shifting the balance in favor of intrasynaptic

D2 receptor activation (Bilder *et al.* 2004; Winterer & Weinberger 2004). Cognitive performance may be critically dependent on the D1/D2 binding ratio, with a relative lack of D1-signaling causing impulsivity, distractibility and poor working memory performance with schizophrenia at the extreme end (Winterer & Weinberger 2004). A relative lack of D2-signaling, on the other hand, may fail to signal the presence of reward information, a signal that is required to engage the PFC in updating its working memory system (Weinberger *et al.* 2001).

The above suggests that the optimum level of DA signaling depends not simply on frontal DA availability, but on its combination with D2 receptor sensitivity. Therefore, individual differences in DA availability as well as D2 receptor sensitivity may come into play during the performance of WM tasks. We tested this expectation in a secondary analysis in which a *DRD2* tagging SNP (rs2075654) was tested for an interactive effect with the *COMT* polymorphism. No significant main effect on WM was found for the rs2075654 *tag*-SNP in *DRD2*, neither for the *young sample* nor for the *adult cohort*. However, the *DRD2* and *COMT* polymorphisms had a significant interactive effect on WM performance in the combined cohort. The interaction suggested that the *Met/Val* heterozygotes perform better than both *Met/Met* and *Val/Val* homozygotes only when they carry one or two A1 alleles. The A1 alleles have been associated with lower receptor density, suggesting that the U-curve shaped effect of DA availability on WM performance disappears when receptor density is high. Such a pattern has been previously reported by Reuter et al (Reuter *et al.* 2005) who reported a significant interactive effect between the *DRD2 Taq IA* and the *COMT* polymorphisms on the amount of response interference in the Stroop color-word conflict task.

Inspection of figure 7.3 suggests that the interaction is stronger in the *adult* than in the *young* cohort. Although *p*-values in neither cohort reach formal significance levels, this age difference may be real. Evidence for age-related changes regarding DA metabolism within the PFC has been postulated in both animal (Lee *et al.* 2001) and human studies (Kaasinen & Rinne 2002), with increased DA metabolism (e.g. *MAO*, *COMT*) thought to be present at a more mature age (Gottfries 1990). Furthermore, several lines of evidence showed a decrease of DA receptors with age (Suhara *et al.* 1991; Volkow *et al.* 1996; Wong *et al.* 1984). When age-related changes in overall levels of DA availability and DA receptor sensitivity are superimposed on the influence of genetic polymorphisms on these levels, a different change in overall DA signaling may occur with age in the various *COMT-DRD2* haplotypes. When we add to this that there may be an inverted U-curved relation between DA signaling and WM performance, interactions between age, *DRD2*, and *COMT* genes should be the rule rather than the exception.

Clearly, full genetic contribution to dopaminergic variation in frontal executive function will rely on far more complex interactions between multiple receptor (e.g. *DRD1*, *DRD2*, *DRD4*), transporters and enzymatic polymorphisms (e.g. *DAT*, *COMT*, *MAO*) (Berman & Noble 1995; Bertolino *et al.* 2006; Tsai *et al.* 2002; Williams & Castner 2006). Further studies systematically involving such interactions are needed in order to obtain a clearer overview of the dopaminergic pathway. At the same time denser SNP coverage of the area under study is needed in genes like *DRD2* to reveal the true functional variants, which while tagged, are still undiscovered.

In summary, our results are in keeping with previous findings suggesting that working memory performance needs an optimal level of dopamine signaling within the PFC. This optimum depends on enzymatic activity controlling dopamine level as well as dopamine receptor sensitivity, both of which may differ as a function of age and genotype. As a consequence the effects of a single polymorphism in a dopaminergic gene on a well-defined cognitive trait may easily remain hidden if the interaction with age and other genes in the pathway are not taken into account.

## REFERENCES

- Abecasis, G.R., Cardon, L.R. & Cookson, W.O. (2000) A general test of association for quantitative traits in nuclear families. *Am J Hum Genet* 66 279-292.
- Axelrod, J. (1957) O-Methylation of Epinephrine and Other Catechols in vitro and in vivo. *Science* 126 400-401.
- Berman, S.M. & Noble, E.P. (1995) Reduced visuospatial performance in children with the D2 dopamine receptor A1 allele. *Behav Genet* 25 45-58.
- Bertolino, A., Blasi, G., Latorre, V., Rubino, V., Rampino, A., Sinibaldi, L., Caforio, G., Petruzzella, V., Pizzuti, A., Scarabino, T., Nardini, M., Weinberger, D.R. & Dallapiccola, B. (2006) Additive effects of genetic variation in dopamine regulating genes on working memory cortical activity in human brain. *J Neurosci* 26 3918-3922.
- Bilder, R.M., Volavka, J., Lachman, H.M. & Grace, A.A. (2004) The catechol-O-methyltransferase polymorphism: relations to the tonic-phasic dopamine hypothesis and neuropsychiatric phenotypes. *Neuropsychopharmacology* 29 1943-1961.
- Boomsma, D.I. (1998) Twin registers in Europe: an overview. *Twin Res* 1 34-51.
- Boomsma, D.I., de Geus, E.J., Vink, J.M., Stubbe, J.H., Distel, M.A., Hottenga, J.J., Posthuma, D., van Beijsterveldt, T.C., Hudziak, J.J., Bartels, M. & Willemsen, G. (2006) Netherlands Twin Register: from twins to twin families. *Twin Res Hum Genet* 9 849-857.
- Boudikova, B., Szumlanski, C.L., Maidak, B. & Weinshilboum, R.M. (1990) Human liver catechol-O-methyltransferase pharmacogenetics. *Clinical pharmacology and therapeutics* 48 381-389.
- Brunner, H.G., Nelen, M., Breakefield, X.O., Ropers, H.H. & van Oost, B.A. (1993) Abnormal behavior associated with a point mutation in the structural gene for monoamine oxidase A. *Science* 262 578-580.
- Colom, R., Rebollo, I., Palacios, A., Juan-Espinosa, M. & Kyllonen, P.C. (2004) Working memory is (almost) perfectly predicted by *g*. *Intelligence* 32 277-296.
- Comings, D.E. & Blum, K. (2000) Reward deficiency syndrome: genetic aspects of behavioral disorders. *Prog Brain Res* 126 325-341.
- Conway, A.R.A., Cowan, N., Bunting, M.F., Theriault, D.J. & Minkhoff, S.R.B. (2002) A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence. *Intelligence* 30 163-183.
- Cools, R. & Robbins, T.W. (2004) Chemistry of the adaptive mind. *Philos Transact A Math Phys Eng Sci* 362 2871-2888.
- Curtis, C.E. & D'Esposito, M. (2003) Persistent activity in the prefrontal cortex during working memory. *Trends Cogn Sci* 7 415-423.
- Fulker, D.W., Cherny, S.S., Sham, P.C. & Hewitt, J.K. (1999) Combined linkage and association sib-pair analysis for quantitative traits. *Am J Hum Genet* 64 259-267.
- Fuster, J.M. (1997) Network memory. *Trends Neurosci* 20 451-459.
- Gibbs, S.E. & D'Esposito, M. (2005) Individual capacity differences predict working memory performance and prefrontal activity following dopamine receptor stimulation. *Cogn Affect Behav Neurosci* 5 212-221.
- Gosso, M.F., de Geus, E.J., van Belzen, M.J., Polderman, T.J., Heutink, P., Boomsma, D.I. & Posthuma, D. (2006a) The SNAP-25 gene is associated with cognitive ability: evidence from a family-based study in two independent Dutch cohorts. *Mol Psychiatry* 11 878-886.
- Gosso, M.F., van Belzen, M., de Geus, E.J., Polderman, J.C., Heutink, P., Boomsma, D.I. & Posthuma, D. (2006b) Association between the CHRM2 gene and intelligence in a sample of 304 Dutch families. *Genes Brain Behav* 5 577-584.
- Gottfries, C.G. (1990) Neurochemical aspects on aging and diseases with cognitive impairment. *J Neurosci Res* 27 541-547.
- Grossman, M.H., Emanuel, B.S. & Budarf, M.L. (1992) Chromosomal mapping of the human catechol-O-methyltransferase gene to 22q11.1-q11.2. *Genomics* 12 822-825.
- Harmer, C.J., McTavish, S.F., Clark, L., Goodwin, G.M. & Cowen, P.J. (2001) Tyrosine depletion attenuates dopamine function in healthy volunteers. *Psychopharmacology (Berl)* 154 105-111.
- Harris, S.E., Wright, A.F., Hayward, C., Starr, J.M., Whalley, L.J. & Deary, I.J. (2005) The functional COMT polymorphism, Val 158 Met, is associated with logical memory and the personality trait intellect/imagination in a cohort of healthy 79 year olds. *Neurosci Lett* 385 1-6.
- Hirschhorn, J.N. & Daly, M.J. (2005) Genome-wide association studies for common diseases and complex traits. *Nat Rev Genet* 6 95-108.
- Hoekstra, C., Meijer, P., Kluft, C., Heutink, P., Smit, G., de Geus, E., Smit, J.H., van Bruggen, A., Montgomery, G.W. & Boomsma, D.I. (2004) Genetics of dizygotic twinning: a feasibility study for a biobank. *Twin Res* 7 556-563.
- Hong, J., Shu-Leong, H., Tao, X. & Lap-Ping, Y. (1998) Distribution of catechol-O-methyltransferase expression in human central nervous system. *Neuroreport* 9 2861-2864.
- Jonsson, E.G., Nothen, M.M., Grunhage, F., Farde, L., Nakashima, Y., Propping, P. & Sedvall, G.C. (1999) Polymorphisms in the dopamine D2 receptor gene and their relationships to striatal dopamine receptor density of healthy volunteers. *Mol Psychiatry* 4 290-296.
- Kaasinen, V. & Rinne, J.O. (2002) Functional imaging studies of dopamine system and cognition in normal aging and Parkinson's disease. *Neurosci Biobehav Rev* 26 785-793.
- Laird, N.M. & Lange, C. (2006) Family-based designs in the age of large-scale gene-association studies. *Nat Rev Genet* 7 385-394.

- Lee, J.J., Chang, C.K., Liu, I.M., Chi, T.C., Yu, H.J. & Cheng, J.T. (2001) Changes in endogenous monoamines in aged rats. *Clin Exp Pharmacol Physiol* 28 285-289.
- Lotta, T., Vidgren, J., Tilgmann, C., Ulmanen, I., Melen, K., Julkunen, I. & Taskinen, J. (1995) Kinetics of human soluble and membrane-bound catechol O-methyltransferase: a revised mechanism and description of the thermolabile variant of the enzyme. *Biochemistry* 34 4202-4210.
- Lundstrom, K., Tenhunen, J., Tilgmann, C., Karhunen, T., Panula, P. & Ulmanen, I. (1995) Cloning, expression and structure of catechol-O-methyltransferase. *Biochim Biophys Acta* 1251 1-10.
- Malhotra, A.K., Kestler, L.J., Mazzanti, C., Bates, J.A., Goldberg, T. & Goldman, D. (2002) A functional polymorphism in the COMT gene and performance on a test of prefrontal cognition. *Am J Psychiatry* 159 652-654.
- Matsumoto, M., Weickert, C.S., Akil, M., Lipska, B.K., Hyde, T.M., Herman, M.M., Kleinman, J.E. & Weinberger, D.R. (2003) Catechol O-methyltransferase mRNA expression in human and rat brain: evidence for a role in cortical neuronal function. *Neuroscience* 116 127-137.
- Mattay, V.S., Goldberg, T.E., Fera, F., Hariri, A.R., Tessitore, A., Egan, M.F., Kolachana, B., Callicott, J.H. & Weinberger, D.R. (2003) Catechol O-methyltransferase val158-met genotype and individual variation in the brain response to amphetamine. *Proc Natl Acad Sci U S A* 100 6186-6191.
- Meulenbelt, I., Droog, S., Trommelen, G.J., Boomsma, D.I. & Slagboom, P.E. (1995) High-yield noninvasive human genomic DNA isolation method for genetic studies in geographically dispersed families and populations. *Am J Hum Genet* 57 1252-1254.
- Miller, S.A., Dykes, D.D. & Polesky, H.F. (1988) A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* 16 1215.
- Palmatier, M.A., Kang, A.M. & Kidd, K.K. (1999) Global variation in the frequencies of functionally different catechol-O-methyltransferase alleles. *Biol Psychiatry* 46 557-567.
- Pohjalainen, T., Rinne, J.O., Nagren, K., Lehtikoinen, P., Anttila, K., Syvalahti, E.K. & Hietala, J. (1998) The A1 allele of the human D2 dopamine receptor gene predicts low D2 receptor availability in healthy volunteers. *Mol Psychiatry* 3 256-260.
- Polderman, T.J.C., Gosso, M.F., Posthuma, D., Van Beijsterveldt, T.C., Heutink, P., Verhulst, F.C. & Boomsma, D.I. (2006a) A longitudinal twin study on IQ, executive functioning, and attention problems during childhood and early adolescence. *Acta Neurol Belg* 106 191-207.
- Polderman, T.J.C., Stins, J.F., Posthuma, D., Gosso, M.F., Verhulst, F.C. & Boomsma, D.I. (2006b) The phenotypic and genotypic relation between working memory speed and capacity. *Intelligence* 34 549-560.
- Posthuma, D., de Geus, E.J., Boomsma, D.I. & Neale, M.C. (2004) Combined linkage and association tests in mx. *Behav Genet* 34 179-196.
- Posthuma, D., Luciano, M., Geus, E.J., Wright, M.J., Slagboom, P.E., Montgomery, G.W., Boomsma, D.I. & Martin, N.G. (2005) A Genomewide Scan for Intelligence Identifies Quantitative Trait Loci on 2q and 6p. *Am J Hum Genet* 77 318-326.
- Reenilä, I. & Mannisto, P.T. (2001) Catecholamine metabolism in the brain by membrane-bound and soluble catechol-O-methyltransferase (COMT) estimated by enzyme kinetic values. *Med Hypotheses* 57 628-632.
- Reuter, M. & Hennig, J. (2005) Association of the functional catechol-O-methyltransferase VAL158MET polymorphism with the personality trait of extraversion. *Neuroreport* 16 1135-1138.
- Reuter, M., Peters, K., Schroeter, K., Koebeke, W., Lenardon, D., Bloch, B. & Hennig, J. (2005) The influence of the dopaminergic system on cognitive functioning: A molecular genetic approach. *Behav Brain Res* 164 93-99.
- Ritchie, T. & Noble, E.P. (1996) [3H]naloxone binding in the human brain: alcoholism and the TaqI A D2 dopamine receptor polymorphism. *Brain Res* 718 193-197.
- Ritchie, T. & Noble, E.P. (2003) Association of Seven polymorphisms of the D2 dopamine Receptor Gene with Brain Receptor-Binding Characteristics. *Neurochemical Research* 28 73-82.
- Savitz, J., Solms, M. & Ramesar, R. (2006) The molecular genetics of cognition: dopamine, COMT and BDNF. *Genes Brain Behav* 5 311-328.
- Stefanis, N.C., Van Os, J., Avramopoulos, D., Smyrnis, N., Evdokimidis, I., Hantoumi, I. & Stefanis, C.N. (2004) Variation in catechol-O-methyltransferase val158 met genotype associated with schizotypy but not cognition: a population study in 543 young men. *Biol Psychiatry* 56 510-515.
- Stins, J.F., Polderman, J.C., Boomsma, D.I. & de Geus, E.J. (2005) Response interference and working memory in 12-year-old children. *Child Neuropsychol* 11 191-201.
- Suhara, T., Fukuda, H., Inoue, O., Itoh, T., Suzuki, K., Yamasaki, T. & Tateno, Y. (1991) Age-related changes in human D1 dopamine receptors measured by positron emission tomography. *Psychopharmacology (Berl)* 103 41-45.
- Tenhunen, J., Salminen, M., Lundstrom, K., Kiviluoto, T., Savolainen, R. & Ulmanen, I. (1994) Genomic organization of the human catechol O-methyltransferase gene and its expression from two distinct promoters. *Eur J Biochem* 223 1049-1059.
- Thompson, J., Thomas, N., Singleton, A., Piggott, M., Lloyd, S., Perry, E.K., Morris, C.M., Perry, R.H., Ferrier, I.N. & Court, J.A. (1997) D2 dopamine receptor gene (DRD2) TaqI A polymorphism: reduced dopamine D2 receptor binding in the human striatum associated with the A1 allele. *Pharmacogenetics* 7 479-484.
- Tsai, S.J., Yu, Y.W., Lin, C.H., Chen, T.J., Chen, S.P. & Hong, C.J. (2002) Dopamine D2 receptor and N-methyl-D-aspartate receptor 2B subunit genetic variants and intelligence. *Neuropsychobiology* 45 128-130.

- Volkow, N.D., Ding, Y.S., Fowler, J.S., Wang, G.J., Logan, J., Gatley, S.J., Hitzemann, R., Smith, G., Fields, S.D. & Gur, R. (1996) Dopamine transporters decrease with age. *J Nucl Med* 37 554-559.
- Wechsler, D. (1986) Wechsler Intelligence Scale for Children-Revised (WISC-R). Lisse: Swets & Zeitlinger.
- Wechsler, D. (1997) WAIS-III Wechsler Adult Intelligence Scale San Antonio, Texas: Psychological Corporation.
- Weinberger, D.R., Egan, M.F., Bertolino, A., Callicott, J.H., Mattay, V.S., Lipska, B.K., Berman, K.F. & Goldberg, T.E. (2001) Prefrontal neurons and the genetics of schizophrenia. *Biol Psychiatry* 50 825-844.
- Weinshilboum, R.M. & Raymond, F.A. (1977) Inheritance of low erythrocyte catechol-o-methyltransferase activity in man. *American Journal of Human Genetics* 29 125-135.
- Williams, G.V. & Castner, S.A. (2006) Under the curve: critical issues for elucidating D1 receptor function in working memory. *Neuroscience* 139 263-276.
- Winterer, G. & Weinberger, D.R. (2004) Genes, dopamine and cortical signal-to-noise ratio in schizophrenia. *Trends in neurosciences* 27 683-690.
- Wong, D.F., Wagner, H.N., Jr., Dannals, R.F., Links, J.M., Frost, J.J., Ravert, H.T., Wilson, A.A., Rosenbaum, A.E., Gjedde, A., Douglass, K.H. & et al. (1984) Effects of age on dopamine and serotonin receptors measured by positron tomography in the living human brain. *Science* 226 1393-1396.
- Zhao, J.H. (2004) 2LD, Genecounting and Hap: computer programs for linkage disequilibrium analysis. *Bioinformatics* 20 1325-1326.

# CHAPTER 8

## SUMMARY AND DISCUSSION

## SUMMARY

In order to identify common DNA polymorphisms explaining variation in cognitive ability, a family-based association study was conducted among a set of putative candidate genes (see Table 1.1 in Chapter 1). The use of two age-dependent cohorts allowed us to detect genetic effects exerted by genetic variants located on putative candidate genes, and more specifically, those polymorphisms that might be involved in variation across different stages of life. Finally, the results of the present work as well as the possibilities of further research are discussed.

*Sample description*

For this thesis, data from two different age cohorts was available: a young cohort (Polderman *et al.*, 2006a; 2006b) and an adult cohort (Posthuma *et al.*, 2005). Both cohorts consist of twins and their siblings and were recruited from the Netherlands twin Register (Boomsma *et al.* 2006; Boomsma *et al.* 2002) (see Table 8.1). Participation in this study included a voluntary agreement to provide buccal swabs for DNA extraction. For all same-sex twin pairs, zygosity was assessed using 11 polymorphic microsatellite markers (Heterozygosity > 0.80). Genotyping was performed blind to familial status and phenotypic data.

**Table 8.1** Zygosity status among individuals for young and adult cohort participating in this study

|              |             | Young Cohort |          | Adult Cohort |          |
|--------------|-------------|--------------|----------|--------------|----------|
|              |             | Families     | Subjects | Families     | Subjects |
| MZM          | twin pair   | 35           | 80       | 25           | 66       |
|              | single twin | -            | -        | 15           | 17       |
| MZF          | twin pair   | 28           | 65       | 20           | 46       |
|              | single twin | -            | -        | 15           | 16       |
| DZM          | twin pair   | 48           | 116      | 15           | 40       |
|              | single twin | -            | -        | 14           | 17       |
| DZF          | twin pair   | 23           | 47       | 28           | 71       |
|              | single twin | -            | -        | 28           | 45       |
| DOS          | twin pair   | 26           | 60       | 23           | 59       |
|              | single twin | -            | -        | -            | -        |
| Single twin  |             | 3            | 3        | 12           | 14       |
| Total cohort |             | 163          | 371      | 195          | 391      |

*IQ Phenotypes*

The WISC-R (1986) (Wechsler 1986) was used to assess psychometric IQ in the young cohort. Six subtests were used: similarities, arithmetic, vocabulary, digit



span, object assembly and block design. Psychometric IQ within the adult cohort was assessed with the Dutch version of the WAIS-III (1997) (Wechsler 1997). Eleven subtests were used: Information, Similarities, Arithmetic, Vocabulary, Digit–symbol pairing, Digit-symbol coding, Digit-symbol free recall, Picture completion, Block design, Matrix reasoning and Letter-number sequencing (see Table 8.2a). Heritability estimates for Full Scale IQ (FSIQ), Verbal IQ (VIQ) and Performance IQ (PIQ) are given in Table 8.2b. These heritability estimates are comparable to those reported previously for the young and adult cohorts in the Dutch population (Bartels *et al.* 2002; Posthuma *et al.* 2001).

**Table 8.2a** Subtests comprising the WAIS<sup>A</sup> and WISC<sup>B</sup>

|             | Subtest                            | Description  |
|-------------|------------------------------------|--|
| Verbal      | 1 Information <sup>A</sup>         | General factual knowledge, long term memory              |
|             | 2 Similarities <sup>A, B</sup>     | Abstract reasoning, categories, relationships            |
|             | 3 Arithmetic <sup>A, B</sup>       | Attention, concentration, numerical reasoning            |
|             | 4 Vocabulary <sup>A, B</sup>       | Word knowledge, verbal fluency                           |
|             | 5 Comprehension <sup>A</sup>       | Social judgment, common sense reasoning                  |
|             | 6 Digit Span <sup>A, B</sup>       | Short term auditory memory, concentration                |
| Performance | 7 Picture Completion <sup>A</sup>  | Alertness to essential detail                            |
|             | 8 Coding <sup>A</sup>              | Visual motor co-ordination, speed, concentration         |
|             | 9 Picture Arrangement <sup>A</sup> | Sequential, logical thinking                             |
|             | 10 Block Design <sup>B</sup>       | Spatial, abstract visual problem solving                 |
|             | 11 Object Assembly <sup>B</sup>    | Visual analysis, construction of objects                 |
|             | 12 Symbol Search <sup>A</sup>      | Speed of processing novel information                    |
|             | 13 Mazes <sup>A</sup>              | Fine motor co-ordination, planning, following directions |

**Table 8.2b** Heritability estimates ( $h^2$ ) for Full-Scaled IQ (FSIQ), Verbal IQ (VIQ), and Performance IQ (PIQ) for young and adult cohorts

| Cognitive Phenotypes (subtests) | Subtests               | Heritability estimates ( $h^2$ ) |
|---------------------------------|------------------------|----------------------------------|
| FSIQ Young                      | 2, 3, 4, 6, 10, 11     | 0.80 (0.72-0.85)                 |
| FSIQ Adult                      | 1, 2, 3, 4, 10, 12, 13 | 0.78 (0.72-0.83)                 |
| VIQ Young                       | 2, 3, 4, 6             | 0.70 (0.59-0.78)                 |
| VIQ Adult                       | 1, 2, 3, 4             | 0.78 (0.72-0.83)                 |
| PIQ Young                       | 10, 11                 | 0.73 (0.63-0.80)                 |
| PIQ Adult                       | 10,12,13               | 0.71 (0.62-0.77)                 |

*The Muscarinic acetylcholine receptor type 2 (CHRM2) gene: genetic variation and cognitive ability*

Muscarinic acetylcholine receptors (mAChR) belong to a group of seven transmembrane-spanning receptors that includes the adrenergic receptors, whose signals are transduced across membranes via interaction with GTP-binding proteins. Several macromolecular interactions are involved in the response triggered by activation of muscarinic receptors (Hulme 1990), ranging from inhibition of adenylyl cyclase, stimulation of phosphoinositide hydrolysis and regulation of potassium channels (Caulfield 1993). It is also known that promoters of neuronal proteins contain consensus sequences for various transcription factors whose expression is increased by mAChR activation in different neuronal cells (Nitsch *et al.* 1998; Von der Kammer *et al.* 1999). Genotypic variation within the muscarinic acetylcholine receptors type 2 (*CHRM2*) gene was investigated in relation to variation in cognitive phenotypes (chapters 2 and 3). What makes this putative candidate gene particularly interesting is its autoreceptor activity at brain structures such as the hippocampal formation, which is fundamentally related to memory and learning processes (Iannazzo & Majewski 2000). The *CHRM2* gene may be involved in fine-tuning of feedback inhibition after memory formation (Miranda *et al.* 2000; Orsetti *et al.* 1996).

The first putative region we found associated with IQ variation (i.e. Performance IQ (PIQ):  $\chi^2=9.14$ ,  $P=0.003$ ; 6.89 IQ points increase in PIQ) among the adult cohort is located in intron 4. The second region encompassing intron 5 – the last intron before the coding sequence – was found associated with variation in IQ (i.e. Full scaled IQ (FIQ):  $\chi^2=7.14$ ,  $P=0.008$ ; 5.35 IQ points increase in FIQ, Verbal IQ (VIQ):  $\chi^2=9.50$ ,  $P=0.002$ ; 5.30 IQ points increase in VIQ) among the young cohort. Although preliminary expression analysis of *CHRM2* in relation to genotypic variation did not reveal differential transcript expression, our association results confirmed previous independent results that have evidenced a putative role of the *CHRM2* gene in cognition (Comings *et al.* 2003; Dick *et al.* 2006; Dick *et al.* 2007). Because association results were found among young and adult cohorts in different regions of the *CHRM2* gene; one could propose that both associated regions contain regulatory elements that are differently used during early and adult life. *How can these differential associations can fit in a plausible biological model?* A large body of literature has shown the importance of choline availability during brain development, and more specifically during hippocampal development (for review see Blusztajn and Wurtman 1983, Glenn *et al.*, 2007; Mellot *et al.*, 2007). In line with this, Cermak and colleagues (Cermak *et al.* 1998; Cermak *et al.*, 1999)

reported a significant association between choline availability and performance in hippocampal-related tasks. Moreover, the rate of Ach turnover (i.e. synthesis, degradation and choline utilization) underlying this association seemed to be programmed early in life in relation to the amounts of Ach availability during critical developmental stages. For example, the hippocampus of prenatally choline-deficient animals is characterized by accelerated Ach turnover, possibly indicating an adaptative response to the reduce availability of choline *in utero* (Cermak *et al.*, 1998). Such a “metabolic imprinting” hypothesis, is attractive from the epigenetic point of view since it takes into account environmental factors occurring prenatally that may explain the lack of replication of association studies when only genetic factors are considered (Niculescu *et al.* 2006).

*The role of the synaptosomal protein of 25 kDa (SNAP-25) gene: synaptic plasticity underlying cognitive ability*

The synaptosomal protein of 25 kDa (*SNAP-25*) gene was one of the putative candidate genes that was investigated in the present study (chapters 4 and 5). Several studies have associated this gene to a wide range of cognitive impairment disorders, ranging from ADHD phenotypes to schizophrenia. The activity of *SNAP-25* is exerted across different organs and tissues. At the brain level, its expression seems to be critical in the hippocampus, in relation to long term potentiation (LTP) and memory consolidation (Hou *et al.* 2004; Hou *et al.* 2006). Within the brain alternative splicing of exon 5 gives rise to two protein isoforms: *SNAP-25a* and *SNAP-25b* (Bark & Wilson 1991). During development, the *SNAP-25a* isoform is known to be the main isoform present, which, in turn, is involved in synaptogenesis. In the adult brain, however, *SNAP-25b* is the predominant isoform which forms a fusion machinery complex (SNARE) together with syntaxin and the synaptic vesicle proteins (synaptobrevin and synaptotagmin). This complex mediates exocytosis of neurotransmitters from the synaptic vesicle into the synaptic cleft. These isoforms are fundamental for keeping a balanced trade-off between synaptic formation and neurotransmitter vesicle release. Our results (chapters 4 and 5) showed genomic variation in intron 1 of the *SNAP-25* gene to be associated with variation in IQ phenotypes (i.e. Full scaled IQ (FIQ):  $\chi^2 = 15.99$   $P = 0.0001$ , 3.28 IQ points increase in FIQ, Verbal IQ (VIQ):  $\chi^2 = 13.01$   $P = 0.0003$ , 2.76 IQ points increase in VIQ, and Performance IQ (PIQ):  $\chi^2 = 11.22$   $P = 0.0008$ , 3.21 IQ points increase in PIQ).

These reported genetic (non)coding variants present in intron 1 might be involved in regulation of protein isoform expression, since all associated SNPs were

located in putative transcription factor binding sites (TFBS). In line with the idea that quantitative trait loci (QTL) with small effect sizes can be expected to explain complex traits, it is likely that (non)coding variants affecting regulatory sequences (e.g. TFBS, promoters, enhancers) may regulate gene expression and function in a more subtle way than do polymorphisms in coding regions. Although the associated variants reported could indeed be the biologically relevant variants, another possibility is that they might be associated variants, in high LD with the causal variant(s). Therefore, further functional studies are required in order to corroborate these polymorphisms as functional variants underlying variation within cognitive phenotypes.

*Functional non-synonymous polymorphisms under positive evolutionary selection underlying phenotypic differences in cognitive abilities*

Using a comparative genomics approach, a two stage design was proposed in which genes ascertained for enhanced protein evolution in primates were subsequently searched for the presence of non-synonymous coding SNPs in the current human population at amino acid sites that differ between humans and chimpanzees (Chapter 6). In this study, as our primary interest was to find genes involved in phenotypic differences in cognitive abilities, we focused on genes expressed in the central nervous system (CNS). Because positively selected genes among primates are generally presumed to determine phenotypic differences between humans and non-human primates, amino acid substitutions segregating in humans at positively selected amino acid sites are expected to affect phenotypic differences among humans. After selecting candidate genes harboring such amino acid substitutions, an association study between cognitive ability and the  $\beta$ -2 adrenergic receptor (*ADRB2*) gene was performed. The  $\beta$ -adrenergic receptors belong to the G-protein-coupled receptor superfamily and mediate some of the physiological actions of catecholamines (noradrenaline and adrenaline) in a variety of tissues (Liggett 2000) via activation exerted by noradrenaline (NA) (Kobayashi & Kobayashi 2001).

Several lines of evidence postulate the *ADRB2* gene as having a fundamental role in memory and learning formation. Hippocampal long-term potentiation (LTP) has been shown to increase after the activation of the  $\beta$ -2 receptor, that in turn, induces LTP in neurons in the hippocampus (Hillman *et al.* 2005). In line with this, the use of  $\beta$ 2- agonists (Gibbs & Summers 2000) has been shown to increase performance related to memory and learning tasks, while  $\beta$ 2- antagonists have been shown to impair memory consolidation (Gibbs & Summers 2005). Two non-synonymous coding SNPs under positive selection (rs1042713 and

rs1042714) in the  $\beta$ -2 adrenergic receptor were selected to conduct a genetic association study for IQ in two independent family-based Dutch cohorts. Interestingly, the derived, human-specific, allele of the beta-2 adrenergic receptor (rs1042713) conferred an 8 IQ point increase in verbal IQ. However, a word of caution must be given due to possible inflation of the estimated genetic effect size due to the relatively small sample size.

### *Positive heterosis and gene-gene interaction on working memory functioning*

In chapter 7, we tested the association of the *Catechol-O-methyl transferase* (*COMT*) Val<sup>108/158</sup>Met polymorphism with working memory (WM) performance. The association reflected positive heterosis such that the Met/Val heterozygotes performed better than each homozygous counterparts on the WM tasks. The association was found in the adult but not in the young cohort. Interestingly, an age-dependent positive heterosis pattern has previously been reported in a longitudinal study by Harris and colleagues (Harris *et al.* 2005). Because the *COMT* polymorphism has been hypothesized to have a non-linear effect on DA availability in the prefrontal cortex (PFC) (Mattay *et al.* 2003), the finding of heterosis is in keeping with the idea that the relationship between DA signalling and cognitive performance follows an inverted U-shaped curve, with both suboptimal and supra-optimal DA activity impairing prefrontal function (Cools & Robbins 2004).

Nevertheless, DA activity is not only dependent on its catabolism rate (*COMT*) but also on the presence of DA receptors. In line with this, cognitive performance may be also critically dependent on the *D1/D2* binding ratio, with a relative lack of *D1*-signalling causing impulsivity, distractibility and poor working memory performance with schizophrenia at the extreme end (Winterer *et al.* 2004). On the other hand, a relative lack of *D2*-signalling may fail to signal the presence of reward information, a signal that is required to engage the PFC in updating its working memory system (Weinberger *et al.* 2001). Consequently, individual differences in DA availability as well as *D2* receptor sensitivity may come into play during the performance of WM tasks. This expectation was tested in a secondary analysis in which a *DRD2* tagging SNP (rs2075654) was tested for an interactive effect with the *COMT* polymorphism. Our results suggest that the interaction is stronger in the adult than in the young cohort. Although p-values in neither cohort reach formal significance levels, this age difference may be real. Although no significant main effect on WM was found for the rs2075654 tag-SNP, either for the young cohort or for the adult cohort, the *DRD2* and *COMT* polymorphisms had a significant interactive effect on WM performance when both samples are

combined. The interaction suggested that the Met/Val heterozygotes perform better than both Met/Met and Val/Val homozygotes only when they carry at least one A1 allele. Evidence for age-related changes regarding DA metabolism within the PFC has been reported in both animal (Lee *et al.* 2001) and human studies (Kaasinen & Rinne 2002), with increased DA metabolism (e.g. *MAO*, *COMT*) thought to be present at a more mature age (Gottfries 1990) as well as a decrease of DA receptors with age (Suhara *et al.* 1991; Volkow *et al.* 1996; Wong *et al.* 1984). Such a pattern has been previously described by Reuter *et al.* (2005), who reported a significant interactive effect between the *DRD2* Taq IA and the *COMT* polymorphisms on the amount of response interference in the Stroop color-word conflict task. Nevertheless, full genetic contribution to dopaminergic variation in frontal executive function will rely on far more complex interactions between multiple receptors (e.g. *DRD1*, *DRD2*, *DRD4*), transporters and enzymatic polymorphisms (e.g. *DAT*, *COMT*, *MAO*) (Berman & Noble 1995; Bertolino *et al.* 2006; Tsai *et al.* 2002; Williams & Castner 2006). Further studies involving such interactions are needed in order to obtain a clearer overview of the dopaminergic pathway underlying working memory tasks.

### *Dopaminergic and serotonergic candidate genes*

Common variants among other dopaminergic (*DBH*, *DRD2*, *DRD3*, and *TH*) and serotonergic (*HTR2A* and *SERT*) system were also investigated in the context of a family-based association study for cognitive ability. Since the preliminary results were not significant after multiple testing correction, no further investigation was performed within these candidate genes. It is worth noting that preliminary results based on these candidate genes may not be conclusive and a more extensive coverage of these genes may be required (see Tables in Appendix I).

In summary, two types of polymorphisms within a candidate-gene design were found associated with variation among cognitive ability phenotypes: coding or so-called functional polymorphisms (*COMT* and *ADRB2* genes) and non-coding polymorphisms (*CHRM2*, *SNAP-25*, and *DRD2* genes). While effects of the former type imply changes that can be observed at the protein level (i.e. enzymatic activity), the later type of polymorphisms should be considered as part of a wider range of regulatory elements, whose genetic contribution, although small in size, is necessary for the final phenotypic outcome. Future functional studies in combination with analysis of gene expression profiles at different brain regions may aid in understanding the role of genetic variants and their relation with synaptic plasticity underlying cognition, learning and attention traits.

## DISCUSSION

### *Linkage and candidate gene studies for cognitive traits*

Clearly, the full genetic contribution to cognitive variation within a normal range may rely on far more complex interactions between common genetic variants influencing neurotransmitter systems activity and specific neuronal networks. Furthermore, due to the relatively modest genetic effects exerted by such common genetic variants, linkage studies have proven not to have the required power to detect them. Conversely, association studies, and in particular candidate-gene studies, were claimed to be the ultimate answer to linkage's pitfalls. Taking into account that QTL underlying normal cognitive variation may interact in a concerted manner rather than in isolation, earlier attempts to identify QTLs underlying complex (common) traits were mainly focussed on coding variants and their functional role in gene expression – i.e. variation in enzymatic activity, receptor density alteration, truncation of protein expression - (for a review see Savitz *et al.* 2006). Even for extensively characterized common coding variants (e.g. *COMT*, *BDNF*), the need for a more comprehensive coverage of genomic sequences (e.g. promoter sequences, introns, and untranslated regions) has been pointed out (Blasi *et al.* 2002; Meyer-Lindenberg *et al.* 2006).

In general, one of the drawbacks of candidate-gene studies is that they are difficult to replicate. The general lack of replication for published association studies most likely reflects the complex interplay between genetic variants whose properties may differ across different populations and non-genetic factors (pre- and postnatal), that together, may model our cognitive profiles. Nevertheless, if a stringent definition of “replication” is taken, i.e. find a significant association with the same SNP, the same statistical test, using the same subtests-phenotypes and identical sources of DNA material (for a review see Sullivan 2007), then relatively few studies can be considered as a true replication. However, these type of replication studies are not always possible nor should their findings be considered conclusive unless functional studies also support a biological role for the associated variants.

Despite the difficulties, some success in identifying QTLs for cognitive abilities has been achieved. For example, in humans, a rare recessive autosomal condition, 4-hydroxybutyric aciduria, caused by deficiency in the *aldehyde dehydrogenase 5 family, member A1* (*ALDH5A1*, *SSADH*) gene, results in accumulation of  $\gamma$ -aminobutyric acid (GABA) and 4-hydroxybutyric acid in the CNS mainly characterized by developmental delay of motor, mental, and language

skills, behavioral problems, and EEG abnormalities (for a review see Gordon 2004). Interestingly, several mutations on the *ALDH5A1* gene have been identified (i.e. deletions leading to frameshift mutations, and altered splicing site mutations) in relation to this rare deficiency, but with a relatively poor genotype-phenotype correlation (Bekri *et al.* 2004). This could be explained by the observation that there is a broad range of *ALDH5A1* activity among both normal control subjects and parents of patients with *ALDH5A1* deficiency (Gibson *et al.* 1991). This is in line with the fact that *ALDH5A1* deficiency can be caused by different gene mutations (allelic heterogeneity). Furthermore, Plomin and colleagues (2004) reported a functional polymorphism (non-synonymous coding SNP, T538C in exon 3) in the *ALDH5A1* gene to be associated with cognition. These results were based on selected samples consisting of subjects with high IQs vs. subjects with normal IQs. Thus, if these results were extrapolated to a normal (unselected) population, a rather small effect size of about 1.5 IQ points might be observed. However, independent replications using unselected samples are needed to determine the role of the *ALDH5A1* in cognition.

A subsequent worldwide survey showed changes in which the ancestral coding variant (T) in the *ALDH5A1* gene shared with primates is being replaced with its derived coding counterpart (C), the same variant that was found associated with an increase in IQ among high-IQ individuals in the Plomin study (2004). Interestingly, it was reported that preliminary results also showed a significant correlation between the derived T allele in *ALDH5A1* and the Microcephaly (*MCPH1*) gene (Leone *et al.* 2006), at least within a restricted asian population. Further studies involving a wider range of populations may be required.

Nevertheless, independently of the previous reported correlations, microcephaly genes (*MCPH1-6*) have been widely shown to be involved in brain development. Several recessive mutations in this gene are known to cause primary microcephaly (Garshasbi *et al.* 2006; Jackson *et al.* 2002; Trimborn *et al.* 2004), a rare clinical condition characterized by marked reduction in brain size coupled with mental retardation. Based on a comparative genomics approach, microcephaly genes haven been postulated to contribute to the evolution of the human brain (Dorus *et al.* 2004), however, association studies involving these genes did not reveal an association with cognition (Mekel-Bobrov *et al.* 2007; Woods *et al.* 2006). So far, these studies failed to show a role of these genes within the normal population (e.g. genotype-phenotype correlation between coding variants and cognitive phenotypes). In line with this, the observed mutations affecting individuals with *ALDH5A1* deficiency or microcephaly mutations may still reflect an extreme of a continuum in genotype-phenotype variation. While rare mutations



affecting coding regions may be partially responsible for the disease phenotypes observed, effects exerted by (non)coding polymorphisms on the other hand, are likely to be affecting the phenotypic variation among the normal expected range.

Polymorphisms within (non)coding regions (e.g. introns, promoters, enhancers, transcription factor (TF) binding sites sequences, splicing regulatory sequences) have been suggested to influence gene expression in a more subtle manner (Knight 2003, 2005), and may even explain the lack of conclusive results when only coding variants affecting for example, enzymatic activity, were taken into account.

It is worth noting that these types of mechanisms, applicable to QTLS, exert their effects in a interactive manner and are likely to be influenced by several common genetic variants, which individually will not *per se* exert variation at the phenotypic level. For example, transcription factors (TFs) are proteins that recognize specific DNA sequences or transcription factor binding sites (TFBS). Their interaction is therefore fundamental for regulation of gene expression ( for review see Garvie and Wolberg, 2001). More importantly, their functionality (e.g. tissue-specific promoters), depend also on their surrounding genomic and cellular context. Therefore, the relatively high heritabilities found for cognitive traits (i.e. VIQ, PIQ, FSIQ) (Bartels *et al.* 2002; Posthuma *et al.* 2001) will not *per se* result in positive association findings, since the QTLs underlying these phenotypic traits may range from a few major factors each with a large effect to dozens of QTLs with a modest to small genetic effect. It is further worth noting that given the relatively high sib/dz correlations for IQ, it is unlikely that the high heritability is caused by gene-gene interactions, since the resemblance for first degree relatives would drop very quickly if gene-gene interactions exist. Thus, a possible scenario might be the existence of different allelic combinations underlying variation among cognitive phenotypes, or the same combination of variants might cause extremely different phenotypic outcomes across different populations due to for example, different linkage disequilibrium (LD) patterns among populations, or even more plausible, relatively small genetic effect sizes, and gene-environmental background interactions. In relation to relative small genetic effect sizes, an old discussion going on since the beginning of association studies may arise: *Can positive, non-replicated findings from independent samples be considered false-positives?*

If a well designed association study “survives” the threshold imposed for multiple testing corrections, or even more reliably, if it “survives” permutation analysis conditional to the empirical genotype-phenotype distribution observed, it

may be very real. Fundamentally, however, as mentioned before, functional studies are always needed in order to show meaningful biological changes exerted by the associated polymorphism(s).

### *Brain development and morphological substrates for learning processes*

Morphological changes in vertebrate development during brain maturation are by far, the most complex events that ever occurred. From a developmental perspective, brain maturation can be considered the most complex and dynamic lifelong process taking place in humans (Toga *et al.* 2006). Neuronal plasticity patterns, as well as basal activity of neurotransmitter systems, have been shown to vary significantly across life and among diverse brain structures (Gogtay *et al.* 2004; Mattay *et al.* 2003). The interaction of a wide variety of neuronal factors such as cell adhesion molecules (Bonfanti 2006; Edelman & Jones 1998), hormones (Aberg *et al.* 2006), and transcription factors (Takahashi & Liu 2006) are fundamental for establishing a mature and dynamic neuronal architecture. Interestingly, neurotransmitters such as acetylcholine (ACh), adrenaline (A), dopamine (DA), gamma-aminobutyric acid (GABA), noradrenaline (NA), and serotonin (5-HT), besides having their well known synaptic neurotransmission function, have been shown to be multifunctional substances participating in developmental processes in all animal species so far investigated (Lauder 1993). These substances, referred to as “prenervous” neurotransmitters, play regulatory roles throughout ontogenesis, including stages prior to development of the nervous system, via their receptors – generally G-coupled protein receptors whose effects are exerted via second messenger molecules - (for review see Lauder 1993). For example, a large body of literature has evidenced the fundamental role of the cholinergic system during brain development, in particular at the level of the hippocampus and related brain structures involved in memory and learning processes (for reviews see Biagioni *et al.* 2000; Olivera *et al.* 2003b). Studies using animal models have shown that the availability of choline during critical periods of foetal development alters hippocampal development and affects memory function throughout life (Blusztajn *et al.* 1998; Cermak *et al.* 1998). Muscarinic acetylcholine receptors (mAChR), beside their well-known role of neurotransmitter release, are also involved in early stages in development, via the activation of the early growth response (EGR) transcription factor family. Transcription of EGR-dependent target genes, have been shown to be under the control of extracellular and intracellular signals coupled to mAChR (von der Kammer *et al.* 1998).

While neurogenesis in the mammalian brain can be considered as a programmed process confined to the prenatal period, at the hippocampal level (i.e. dentate gyrus), neurogenesis has been shown to occur in the adult brain of rodents (Altman

& Das 1965; Cameron *et al.* 1993), primates (Gould *et al.* 1999a; Kornack & Rakic 1999), and humans (Eriksson *et al.* 1998). Based on animal models, it has been proposed that adult-generated neurons are originated as a response to fulfill neuronal requirements encompassing novel learning processes; more specifically, those related to associative learning (Gould *et al.* 1999b). Furthermore, it was shown that an enriched environment triggers this neuronal generation in normal cognitive phenotypes (Kempermann *et al.* 1997b) as well as in animal models with cognitive impairment (Meredith *et al.* 2007; Restivo *et al.* 2005), and that this response has a genetic component (Kempermann *et al.* 1997a).

It has been observed in rodents (Miranda *et al.* 2000; Orsetti *et al.* 1996) that during acquisition of novel tasks with a reward component (rule learning), ACh levels increase at the hippocampus and related brain structures. On the other hand, reduced ACh levels have been shown to be present when retrieval of rewarded acquired tasks is needed (pair learning). Such ACh fluctuations during acquirement of novel learning tasks, where the *CHRM2* gene may play a critical role (Carey *et al.* 2001), are likely to be involved in synaptic plasticity underlying memory formation (Miranda *et al.* 2003).

Related to age-dependent changes, several structural imaging studies have mapped the neuroanatomical course of human brain development and have shown that the most critical structural changes occur over the periods of childhood and adolescence (Gogtay *et al.* 2004; Sowell *et al.* 2004). In summary, regions subserving primary functions, such as motor and sensory systems, mature earliest, with temporal and parietal cortices associated with basic language skills and spatial attention maturing next; followed by maturation of higher-order association areas, such as the prefrontal and lateral temporal cortices (Gogtay *et al.* 2004; Sowell *et al.* 2004). Cross-sectional studies have shown gray matter loss during childhood and adolescence in parallel with increased myelination, in line with the idea of an ongoing sculpting process of the immature brain into the fully functioning one (Giedd *et al.* 1996; Pfefferbaum *et al.* 1994). Developmental changes in cortical development have been found to correlate with behavioral performance measurements (Reiss *et al.* 1996; Sowell *et al.* 2004). Notably, Schlaggar *et al.*, (2002) found that adults (but not children) significantly activated the dorsal prefrontal cortex (PFC) on a single-word processing task. Although other age-group differences in PF activation were also found, these were related to accuracy of task performance independent of age.

The presence of postnatal neuronal maturation within the hippocampal formation was evidenced in different primate studies, particularly at the dentate

gyrus (DG) (Giedd *et al.* 1996; Gould *et al.* 1999b; Kornack & Rakic 1999). Conversely, while the general patterns of projections from related structures to the DG in infant monkeys resembles those observed in the adult (mossy fiber projections that are already established at birth have mature-looking synapses), the dendritic arborisation and the synaptic organization in the molecular layer of the DG demonstrate significant postnatal maturation (for review see Lavenex *et al.* 2007). Therefore, it may be crucial to determine when particular developmental milestones are reached in order to suggest ages at which specific hippocampal circuits become capable of subserving age-related functional processes. The work performed in murine models in relation to hippocampal-dependent learning tasks and different neurotransmitter systems, (e.g. cholinergic-glutamatergic neurotransmitter systems) has been of critical importance. Murine models have shown mAChR activity to regulate promoter activation of genes involved in synaptogenesis such as the acetylcholinesterase (AChE) gene, and glutamatergic receptors (i.e.  $\alpha$ -amino-3-hydroxy-5-methylisoxazole-4-propionic acid (AMPA) receptor), which in turn have been associated with higher brain functions as LTP, memory and behavior (Olivera *et al.* 2003a; Olivera *et al.* 2003b).

#### *Follow up research:*

##### *Candidate pathways or genomic-whole association studies ?*

In the present study, the association between common variants on candidate-genes underlying variation in cognitive ability was analysed. The use of families consisting of twin pairs and extra siblings encompassing two age-cohorts allowed us to conduct both population and family-based association analyses – controlling with the former for possible stratification effects -. Significant association was found between (non)coding variants for the *CHRM2* (chapters 2 and 3) and the *SNAP-25* gene (chapters 4 and 5), respectively.

None of the regions associated with variation in cognitive phenotypes were highly conserved, at least not between organisms than may serve as animal models in further studies like for example mouse or fruitfly. Nevertheless, these regions were highly conserved among the primate lineage. This suggests a possible primate specific development role of these variants, or causal variants tagged by them, in neuronal networks involved in highly cognitive processes. In line with this, it has been recently shown that about two-thirds of fundamental transcription factors (TF) and their target sequences, or TF bindings sites (TFBS), despite their highly conserved function among human and rodents, are not aligned between mouse and human genomes (Odom *et al.* 2007). Therefore, these findings may have important implications for the use of the mouse as a model organism, since depending on the

phenotype under study, susceptibility due to polymorphisms around TF or TFBS may be species specific.

Furthermore, age-dependent effects were observed in relation to both candidate-gene and gene x gene interaction analyses presented in chapters 2 till 5, and 6 till 7, respectively. These effects may reflect the ongoing developmental changes occurring at young age and adulthood. However, functional imaging studies in combination with gene expression analysis may be required for identifying the specific brain structures involved, as well as functional studies in order to unambiguously link the observed cognitive variation with the associated genomic variants.

Association studies on candidate-genes or candidate-pathways are limited to a few genes that, although with a putative effect among cognitive phenotypes, will give us still a very narrow view of our possible genomic interactions underlying cognition. On the other hand, genome-wide association (GWA) studies may provide us with a broader and unbiased overview of common genetic variants. Due to the expense and labour involved, genome-wide association studies have not been feasible till now (Syvanen 2005).

Recently, several genome-wide association studies were published within a large collaborative project frame, one in relation to the identification of new loci conferring susceptibility to breast cancer (Easton *et al.* 2007), another involving association between specific genetic variation and levels of gene expression (Spielman *et al.* 2007), and the a third study involving a wide range of clinical phenotypes with major public health importance (i.e. bipolar disorder, coronary artery disease, Crohn's disease, rheumatoid arthritis, type-1 and type-2 diabetes) (The Wellcome Trust Case Control Consortium 2007) (WTCCC). All studies pointed to (i) the need for large numbers of cases and controls, generally achieved within a multicenter-collaborative setup; (ii) the presence of common variants lying on (non)coding sequences with a putative functionality (TFBS, promoter sequences, CNVs, retrosposons). In particular, the work presented by Spielman and co-workers (2007), showed how complementary association studies in combination with gene expression studies (i.e. genomic comparative hybridization (CGH) studies) might be, given the fact that gene expression phenotypes can also be taken into account together with clinical disease status and/or behavioral traits. In the study conducted by the WTCCC, several loci were found associated with disease phenotypes, and confirmed afterwards. Despite the large sample size used, association signals for many other potential candidate variants remain to be confirmed. Nonetheless, the identification of causal gene variants have yet to be unraveled mainly due to: (i) incomplete coverage of common-variation genome-

wide; (ii) reduced power to detect rare variants and/or variants with relatively small effect sizes; (iii) difficulties in gene-boundary definitions.

In relation to gene expression (endo)phenotypes, the potential influence of common copy number variation (CNV), which can modify patterns of gene regulation underlying QTLs was recently proposed (McCarroll *et al.* 2006; Redon *et al.* 2006). This phenomenon might be of particular interest for cognitive traits. Interestingly, linkage disequilibrium (LD) patterns between SNPs and CNVs have been shown to be relatively modest (Locke *et al.* 2006), possibly due to a high rate of spontaneous recurrence of CNVs. Notably, CNVs have been shown to encompass more than 12% of our genome, and this number may still be an underestimate as the resolution of current studies is still limited. Around 80% of SNPs and 18% of CNVs might be involved in variation of gene expression. In addition, signals from these two type of variants showed non-significant overlap (Redon *et al.* 2006; Stranger *et al.* 2007). This observation opens exciting new possibilities for genetic and cognitive research.

It will be of great interest to determine common variation among common CNVs in relation to human phenotypes (Wirtenberger *et al.* 2006), i.e. cognitive phenotypes among a normal cognitive range, as well as phenotypes encompassing different levels of mental retardation manifestations. CNVs may indeed represent another level of genomic organization, where alteration of gene copy number and gene dosage may, in combination with SNPs, sculpt our final cognitive profile, determining in combination with environmental factors, our ultimate cognitive phenotypic outcome.

## REFERENCES

- Aberg, N.D., Brywe, K.G. & Isgaard, J. (2006) Aspects of growth hormone and insulin-like growth factor-I related to neuroprotection, regeneration, and functional plasticity in the adult brain. *ScientificWorldJournal* 6 53-80.
- Altman, J. & Das, G.D. (1965) Autoradiographic and histological evidence of postnatal hippocampal neurogenesis in rats. *J Comp Neurol* 124 319-335.
- Bark, C. & Wilson, M.C. (1991) Alternative splicing generates a variant SNAP-25 protein during development. *Soc Neurosci Abstr* 17 531.
- Bartels, M., Rietveld, M.J., Van Baal, G.C. & Boomsma, D.I. (2002) Genetic and environmental influences on the development of intelligence. *Behav Genet* 32 237-249.
- Bekri, S., Fossoud, C., Plaza, G., Guenne, A., Salomons, G.S., Jakobs, C. & Van Obberghen, E. (2004) The molecular basis of succinic semialdehyde dehydrogenase deficiency in one family. *Mol Genet Metab* 81 347-351.
- Berman, S.M. & Noble, E.P. (1995) Reduced visuospatial performance in children with the D2 dopamine receptor A1 allele. *Behav Genet* 25 45-58.
- Bertolino, A., Blasi, G., Latorre, V., Rubino, V., Rampino, A., Sinibaldi, L., Caforio, G., Petruzzella, V., Pizzuti, A., Scarabino, T., Nardini, M., Weinberger, D.R. & Dallapiccola, B. (2006) Additive effects of genetic variation in dopamine regulating genes on working memory cortical activity in human brain. *J Neurosci* 26 3918-3922.
- Biagioni, S., Tata, A.M., De Jaco, A. & Augusti-Tocco, G. (2000) Acetylcholine synthesis and neuron differentiation. *Int J Dev Biol* 44 689-697.
- Blasi, P., Boyl, P.P., Ledda, M., Novelletto, A., Gibson, K.M., Jakobs, C., Hogema, B., Akaboshi, S., Loreni, F. & Malaspina, P. (2002) Structure of human succinic semialdehyde dehydrogenase gene: identification of promoter region and alternatively processed isoforms. *Mol Psych* 76 348-362.
- Blusztajn, J.K. & Wurtman, R.J. (1983) Choline and cholinergic neurons. *Science* 221 614-620.
- Blusztajn, J.K., Cermak, J.M., Holler, T. & Jackson, D.A. (1998) Imprinting of hippocampal metabolism of choline by its availability during gestation: implications for cholinergic neurotransmission. *J Physiol (Paris)* 92 199-203.
- Bonfanti, L. (2006) PSA-NCAM in mammalian structural plasticity and neurogenesis. *Prog Neurobiol* 80 129-164.
- Boomsma, D.I., de Geus, E.J., Vink, J.M., Stubbe, J.H., Distel, M.A., Hottenga, J.J., Posthuma, D., van Beijsterveldt, T.C., Hudziak, J.J., Bartels, M. & Willemsen, G. (2006) Netherlands Twin Register: from twins to twin families. *Twin Res Hum Genet* 9 849-857.
- Boomsma, D.I., Vink, J.M., van Beijsterveldt, T.C., de Geus, E.J.C., Beem, A.L., Mulder, E.J., Derks, E.M., Riese, H., Willemsen, G.A., Bartels, M., van den Berg, M., Kupper, N.H., Polderman, T.J.C., Posthuma, D., Rietveld, M.J., Stubbe, J.H., Knol, L.I., Stroet, T. & van Baal, G.C. (2002) Netherlands Twin Register: a focus on longitudinal research. *Twin Res* 5 401-406.
- Cameron, H.A., Woolley, C.S., McEwen, B.S. & Gould, E. (1993) Differentiation of newly born neurons and glia in the dentate gyrus of the adult rat. *Neuroscience* 56 337.
- Carey, G.J., Billard, W., Binch, H., 3rd, Cohen-Williams, M., Crosby, G., Grzelak, M., Guzik, H., Kozlowski, J.A., Lowe, D.B., Pond, A.J., Tedesco, R.P., Watkins, R.W. & Coffin, V.L. (2001) SCH 57790, a selective muscarinic M(2) receptor antagonist, releases acetylcholine and produces cognitive enhancement in laboratory animals. *Eur J Pharmacol* 431 189-200.
- Caulfield, M.P. (1993) Muscarinic receptors--characterization, coupling and function. *Pharmacol Ther* 58 319-379.
- Cermak, J.M., Blusztajn, J.K., Meck, W.H., Williams, C.L., Fitzgerald, C.M., Rosene, D.L. & Loy, R. (1999) Prenatal availability of choline alters the development of acetylcholinesterase in the rat hippocampus. *Dev Neurosci* 21 94-104.
- Cermak, J.M., Holler, T., Jackson, D.A. & Blusztajn, J.K. (1998) Prenatal availability of choline modifies development of the hippocampal cholinergic system. *Faseb J* 12 349-357.
- Comings, D.E., Wu, S., Rostamkhani, M., McGue, M., Lacono, W.G., Cheng, L.S. & MacMurray, J.P. (2003) Role of the cholinergic muscarinic 2 receptor (CHRM2) gene in cognition. *Mol Psychiatry* 8 10-11.
- Cools, R. & Robbins, T.W. (2004) Chemistry of the adaptive mind. *Philos Transact A Math Phys Eng Sci* 362 2871-2888.
- Dick, D.M., Aliev, F., Bierut, L., Goate, A., Rice, J., Hinrichs, A., Bertelsen, S., Wang, J.C., Dunn, G., Kuperman, S., Schuckit, M., Nurnberger, J., Jr., Porjesz, B., Beglieter, H., Kramer, J. & Hesselbrock, V. (2006) Linkage Analyses of IQ in the Collaborative Study on the Genetics of Alcoholism (COGA) Sample. *Behav Genet* 36 77-86.
- Dick, D.M., Aliev, F., Kramer, J., Wang, J.C., Hinrichs, A., Bertelsen, S., Kuperman, S., Schuckit, M., Nurnberger, J., Jr., Edenberg, H.J., Porjesz, B., Begleiter, H., Hesselbrock, V., Goate, A. & Bierut, L. (2007) Association of CHRM2 with IQ: Converging Evidence for a Gene Influencing Intelligence. *Behav Genet* 37 265-272.
- Dorus, S., Vallender, E.J., Evans, P.D., Anderson, J.R., Gilbert, S.L., Mahowald, M., Wyckoff, G.J., Malcom, C.M. & Lahn, B.T. (2004) Accelerated evolution of nervous system genes in the origin of Homo sapiens. *Cell* 119 1027-1040.
- Easton, D.F. & Pooley, K.A. & Dunning, A.M. & Pharoah, P.D. & Thompson, D. & Ballinger, D.G. & Struwing, J.P. & Morrison, J. & Field, H. & Luben, R. & Wareham, N. & Ahmed, S. & Healey, C.S. & Bowman, R. & Luccarini, C. & Conroy, D. & Shah, M. & Munday, H. & Jordan, C., et al (2007) Genome-wide association study identifies novel breast cancer susceptibility loci. *Nature*.

- Edelman, G.M. & Jones, F.S. (1998) Gene regulation of cell adhesion: a key step in neural morphogenesis. *Brain Res Brain Res Rev* 26 337-352.
- Eriksson, P.S., Perfilieva, E., Bjork-Eriksson, T., Alborn, A.-M., Nordborg, C., Peterson, D.A. & Gage, F.H. (1998) Neurogenesis in the adult human hippocampus. *Nat Med* 4 1313.
- Garshasbi, M., Motazacker, M.M., Kahrizi, K., Behjati, F., Abedini, S.S., Nieh, S.E., Firouzabadi, S.G., Becker, C., Ruschendorf, F., Nurnberg, P., Tzschach, A., Vazifehmand, R., Erdogan, F., Ullmann, R., Lenzner, S., Kuss, A.W., Ropers, H.H. & Najmabadi, H. (2006) SNP array-based homozygosity mapping reveals MCPH1 deletion in family with autosomal recessive mental retardation and mild microcephaly. *Hum Genet* 118 708-715.
- Garvie, C.W. & Wolberger, C. (2001) Recognition of specific DNA sequences. *Mol Cell* 8 937-946.
- Gibbs, M.E. & Summers, R.J. (2000) Separate roles for beta2- and beta3-adrenoceptors in memory consolidation. *Neuroscience* 95 913-922.
- Gibbs, M.E. & Summers, R.J. (2005) Contrasting roles for beta1, beta2 and beta3-adrenoceptors in memory formation in the chick. *Neuroscience* 131 31-42.
- Gibson, K.M., Lee, C.F., Chambliss, K.L., Kamali, V., Francois, B., Jaeken, J. & Jakobs, C. (1991) 4-Hydroxybutyric aciduria: application of a fluorometric assay to the determination of succinic semialdehyde dehydrogenase activity in extracts of cultured human lymphoblasts. *Clin Chim Acta* 196 219-221.
- Giedd, J.N., Snell, J.W., Lange, N., Rajapakse, J.C., Casey, B.J., Kozuch, P.L., Vaituzis, A.C., Vauss, Y.C., Hamburger, S.D., Kaysen, D. & Rapoport, J.L. (1996) Quantitative magnetic resonance imaging of human brain development: ages 4-18. *Cereb Cortex* 6 551-560.
- Glenn, M.J., Gibson, E.M., Kirby, E.D., Mellott, T.J., Blusztajn, J.K. & Williams, C.L. (2007) Prenatal choline availability modulates hippocampal neurogenesis and neurogenic responses to enriching experiences in adult female rats. *Eur J Neurosci* 25 2473-2482.
- Gogtay, N., Giedd, J.N., Lusk, L., Hayashi, K.M., Greenstein, D., Vaituzis, A.C., Nugent, T.F., 3rd, Herman, D.H., Clasen, L.S., Toga, A.W., Rapoport, J.L. & Thompson, P.M. (2004) Dynamic mapping of human cortical development during childhood through early adulthood. *Proc Natl Acad Sci U S A* 101 8174-8179.
- Gordon, N. (2004) Succinic semialdehyde dehydrogenase deficiency (SSADH) (4-hydroxybutyric aciduria, gamma-hydroxybutyric aciduria). *Eur J Paediatr Neurol* 8 261-265.
- Gottfries, C.G. (1990) Neurochemical aspects on aging and diseases with cognitive impairment. *J Neurosci Res* 27 541-547.
- Gould, E., Reeves, A.J., Fallah, M., Tanapat, P., Gross, C.G. & Fuchs, E. (1999a) Hippocampal neurogenesis in adult Old World primates. *Proc Natl Acad Sci U S A* 96 5263-5267.
- Gould, E., Tanapat, P., Hastings, N.B. & Shors, T.J. (1999b) Neurogenesis in adulthood: a possible role in learning. *Trends Cogn Sci* 3 186-192.
- Harris, S.E., Wright, A.F., Hayward, C., Starr, J.M., Whalley, L.J. & Deary, I.J. (2005) The functional COMT polymorphism, Val 158 Met, is associated with logical memory and the personality trait intellect/imagination in a cohort of healthy 79 year olds. *Neurosci Lett* 385 1-6.
- Hillman, K.L., Doze, V.A. & Porter, J.E. (2005) Functional characterization of the beta-adrenergic receptor subtypes expressed by CA1 pyramidal cells in the rat hippocampus. *J Pharmacol Exp Ther* 314 561-567.
- Hou, Q., Gao, X., Zhang, X., Kong, L., Wang, X., Bian, W., Tu, Y., Jin, M., Zhao, G., Li, B., Jing, N. & Yu, L. (2004) SNAP-25 in hippocampal CA1 region is involved in memory consolidation. *Eur J Neurosci* 20 1593-1603.
- Hou, Q.L., Gao, X., Lu, Q., Zhang, X.H., Tu, Y.Y., Jin, M.L., Zhao, G.P., Yu, L., Jing, N.H. & Li, B.M. (2006) SNAP-25 in hippocampal CA3 region is required for long-term memory formation. *Biochem Biophys Res Commun* 347 955-962.
- Hulme, E.C. (1990) Muscarinic acetylcholine receptors: typical G-coupled receptors. *Symp Soc Exp Biol* 44 39-54.
- Iannazzo, L. & Majewski, H. (2000) M(2)/M(4)-muscarinic receptors mediate automodulation of acetylcholine outflow from mouse cortex. *Neurosci Lett* 287 129-132.
- Jackson, A.P., Eastwood, H., Bell, S.M., Adu, J., Toomes, C., Carr, I.M., Roberts, E., Hampshire, D.J., Crow, Y.J., Mighell, A.J., Karbani, G., Jafri, H., Rashid, Y., Mueller, R.F., Markham, A.F. & Woods, C.G. (2002) Identification of microcephalin, a protein implicated in determining the size of the human brain. *Am J Hum Genet* 71 136-142.
- Kaasinen, V. & Rinne, J.O. (2002) Functional imaging studies of dopamine system and cognition in normal aging and Parkinson's disease. *Neurosci Biobehav Rev* 26 785-793.
- Kempermann, G., Kuhn, H.G. & Gage, F.H. (1997a) Genetic influence on neurogenesis in the dentate gyrus of adult mice. *Proc Natl Acad Sci U S A* 94 10409-10414.
- Kempermann, G., Kuhn, H.G. & Gage, F.H. (1997b) More hippocampal neurons in adult mice living in an enriched environment. *Nature* 386 493-495.
- Knight, J.C. (2003) Functional implications of genetic variation in non-coding DNA for disease susceptibility and gene regulation. *Clin Sci (Lond)* 104 493-501.
- Knight, J.C. (2005) Regulatory polymorphisms underlying complex disease traits. *J Mol Med* 83 97-109.
- Kobayashi, K. & Kobayashi, T. (2001) Genetic evidence for noradrenergic control of long-term memory consolidation. *Brain Dev* 23 Suppl 1 S16-23.
- Kornack, D.R. & Rakic, P. (1999) Continuation of neurogenesis in the hippocampus of the adult macaque monkey. *PNAS* 96 5768-5773.



- Lauder, J.M. (1993) Neurotransmitters as growth regulatory signals: role of receptors and second messengers. *Trends Neurosci* 16 233-240.
- Lavenex, P., Banta Lavenex, P. & Amaral, D.G. (2007) Postnatal development of the primate hippocampal formation. *Dev Neurosci* 29 179-192.
- Lee, J.J., Chang, C.K., Liu, I.M., Chi, T.C., Yu, H.J. & Cheng, J.T. (2001) Changes in endogenous monoamines in aged rats. *Clin Exp Pharmacol Physiol* 28 285-289.
- Leone, O., Blasi, P., Palmerio, F., Kozlov, A.I., Malaspina, P. & Novelletto, A. (2006) A human derived SSADH coding variant is replacing the ancestral allele shared with primates. *Ann Hum Biol* 33 593-603.
- Liggett, S.B. (2000) Pharmacogenetics of beta-1- and beta-2-adrenergic receptors. *Pharmacology* 61 167-173.
- Locke, D.P., Sharp, A.J., McCarroll, S.A., McGrath, S.D., Newman, T.L., Cheng, Z., Schwartz, S., Albertson, D.G., Pinkel, D., Altshuler, D.M. & Eichler, E.E. (2006) Linkage disequilibrium and heritability of copy-number polymorphisms within duplicated regions of the human genome. *Am J Hum Genet* 79 275-290.
- Mattay, V.S., Goldberg, T.E., Fera, F., Hariri, A.R., Tessitore, A., Egan, M.F., Kolachana, B., Callicott, J.H. & Weinberger, D.R. (2003) Catechol O-methyltransferase val158-met genotype and individual variation in the brain response to amphetamine. *Proc Natl Acad Sci U S A* 100 6186-6191.
- McCarroll, S.A., Hadnott, T.N., Perry, G.H., Sabeti, P.C., Zody, M.C., Barrett, J.C., Dallaire, S., Gabriel, S.B., Lee, C., Daly, M.J. & Altshuler, D.M. (2006) Common deletion polymorphisms in the human genome. *Nat Genet* 38 86-92.
- Mekel-Bobrov, N., Posthuma, D., Gilbert, S.L., Lind, P., Gossio, M.F., Luciano, M., Harris, S.E., Bates, T.C., Polderman, T.J.C., Whalley, L.J., Fox, H., Starr, J.M., Evans, P.D., Montgomery, G.W., Fernandes, C., Heutink, P., Martin, N.G., Boomsma, D.I., Deary, I.J., et al (2007) The ongoing adaptive evolution of ASPM and Microcephalin is not explained by increased intelligence. *Hum Mol Genet* 16 600-608.
- Mellott, T.J., Follettie, M.T., Diesl, V., Hill, A.A., Lopez-Coviella, I. & Blusztajn, J.K. (2007) Prenatal choline availability modulates hippocampal and cerebral cortical gene expression. *FASEB J* 21 1311-1323.
- Meredith, R.M., Holmgren, C.D., Weidum, M., Burnashev, N. & Mansvelder, H.D. (2007) Increased threshold for spike-timing-dependent plasticity is caused by unreliable calcium signaling in mice lacking fragile X gene FMR1. *Neuron* 54 627-638.
- Meyer-Lindenberg, a., Nichols, T., Callicott, J.H., Ding, J., Kolachana, B., Buckholtz, J., Mattay, V.S., Egan, M. & Weinberger, D.R. (2006) Impact of complex genetic variation in COMT on human brain function. *Mol Psychiatry* 11 867-877.
- Miranda, M.I., Ferreira, G., Ramirez-Lugo, L. & Bermudez-Rattoni, F. (2003) Role of cholinergic system on the construction of memories: taste memory encoding. *Neurobiol Learn Mem* 80 211-222.
- Miranda, M.I., Ramirez-Lugo, L. & Bermudez-Rattoni, F. (2000) Cortical cholinergic activity is related to the novelty of the stimulus. *Brain Res* 882 230-235.
- Niculescu, M.D., Craciunescu, C.N. & Zeisel, S.H. (2006) Dietary choline deficiency alters global and gene-specific DNA methylation in the developing hippocampus of mouse fetal brains. *FASEB J* 20 43-49.
- Nitsch, R.M., Rossner, S., Albrecht, C., Mayhaus, M., Enderich, J., Schliebs, R., Wegner, M., Arendt, T. & von der Kammer, H. (1998) Muscarinic acetylcholine receptors activate the acetylcholinesterase gene promoter. *J Physiol Paris* 92 257-264.
- Odom, D.T., Dowell, R.D., Jacobsen, E.S., Gordon, W., Danford, T.W., MacIsaac, K.D., Rolfe, P.A., Conboy, C.M., Gifford, D.K. & Fraenkel, E. (2007) Tissue-specific transcriptional regulation has diverged significantly between human and mouse. *Nat Genet* 39 730.
- Olivera, S., Henley, J.M. & Rodriguez-Ithurralde, D. (2003a) AMPA receptor potentiation by acetylcholinesterase is age-dependently upregulated at synaptogenesis sites of the rat brain. *Int J Dev Neurosci* 21 49-61.
- Olivera, S., Rodriguez-Ithurralde, D. & Henley, J.M. (2003b) Acetylcholinesterase promotes neurite elongation, synapse formation, and surface expression of AMPA receptors in hippocampal neurones. *Mol Cell Neurosci* 23 96-106.
- Orsetti, M., Casamenti, F. & Pepeu, G. (1996) Enhanced acetylcholine release in the hippocampus and cortex during acquisition of an operant behavior. *Brain Res* 724 89-96.
- Pfefferbaum, A., Mathalon, D.H., Sullivan, E.V., Rawles, J.M., Zipursky, R.B. & Lim, K.O. (1994) A quantitative magnetic resonance imaging study of changes in brain morphology from infancy to late adulthood. *Arch Neurol* 51 874-887.
- Plomin, R., Turic, D.M., Hill, L., Turic, D.E., Stephens, M., Williams, J., Owen, M.J. & O'Donovan, M.C. (2004) A functional polymorphism in the succinate-semialdehyde dehydrogenase (aldehyde dehydrogenase 5 family, member A1) gene is associated with cognitive ability. *Molecular Psychiatry* 9 582-586.
- Polderman, T.J.C., Gossio, M.F., Posthuma, D., Van Beijsterveldt, T.C., Heutink, P., Verhulst, F.C. & Boomsma, D.I. (2006a) A longitudinal twin study on IQ, executive functioning, and attention problems during childhood and early adolescence. *Acta Neurol Belg* 106 191-207.
- Polderman, T.J.C., Stins, J.F., Posthuma, D., Gossio, M.F., Verhulst, F.C. & Boomsma, D.I. (2006b) The phenotypic and genotypic relation between working memory speed and capacity. *Intelligence* 34 549-560.
- Posthuma, D., de Geus, E.J.C. & Boomsma, D.I. (2001) Perceptual speed and IQ are associated through common genetic factors. *Behav Genet* 31 593-602.
- Posthuma, D., Luciano, M., Geus, E.J., Wright, M.J., Slagboom, P.E., Montgomery, G.W., Boomsma, D.I. & Martin, N.G. (2005) A Genomewide Scan for Intelligence Identifies Quantitative Trait Loci on 2q and 6p. *Am J Hum Genet* 77 318-326.

- Redon, R., Ishikawa, S., Fitch, K.R., Feuk, L., Perry, G.H., Andrews, T.D., Fiegler, H., Shapero, M.H., Carson, A.R., Chen, W., Cho, E.K., Dallaire, S., Freeman, J.L., Gonzalez, J.R., Gratacos, M., Huang, J., Kalaitzopoulos, D., Komura, D., MacDonald, J.R., et al (2006) Global variation in copy number in the human genome. *Nature* 444 444-454.
- Reiss, A.L., Abrams, M.T., Singer, H.S., Ross, J.L. & Denckla, M.B. (1996) Brain development, gender and IQ in children. A volumetric imaging study. *Brain* 119 ( Pt 5) 1763-1774.
- Restivo, L., Ferrari, F., Passino, E., Sgobio, C., Bock, J., Oostra, B.A., Bagni, C. & Ammassari-Teule, M. (2005) Enriched environment promotes behavioral and morphological recovery in a mouse model for the fragile X syndrome. *Proc Natl Acad Sci U S A* 102 11557-11562.
- Reuter, M., Peters, K., Schroeter, K., Koebke, W., Lenardon, D., Bloch, B. & Hennig, J. (2005) The influence of the dopaminergic system on cognitive functioning: A molecular genetic approach. *Behav Brain Res* 164 93-99.
- Savitz, J., Solms, M. & Ramesar, R. (2006) The molecular genetics of cognition: dopamine, COMT and BDNF. *Gene Brain Behav* 5 311-328.
- Schlaggar, B.L., Brown, T.T., Lugar, H.M., Visscher, K.M., Miezin, F.M. & Petersen, S.E. (2002) Functional neuroanatomical differences between adults and school-age children in the processing of single words. *Science* 296 1476-1479.
- Sowell, E.R., Thompson, P.M. & Toga, A.W. (2004) Mapping changes in the human cortex throughout the span of life. *Neuroscientist* 10 372-392.
- Spielman, R.S., Bastone, L.A., Burdick, J.T., Morley, M., Ewens, W.J. & Cheung, V.G. (2007) Common genetic variants account for differences in gene expression among ethnic groups. *Nat Genet* 39 226-231.
- Stranger, B.E., Forrest, M.S., Dunning, M., Ingle, C.E., Beazley, C., Thorne, N., Redon, R., Bird, C.P., de Grassi, A., Lee, C., Tyler-Smith, C., Carter, N., Scherer, S.W., Tavare, S., Deloukas, P., Hurles, M.E. & Dermitzakis, E.T. (2007) Relative impact of nucleotide and copy number variation on gene expression phenotypes. *Science* 315 848-853.
- Suhara, T., Fukuda, H., Inoue, O., Itoh, T., Suzuki, K., Yamasaki, T. & Tateno, Y. (1991) Age-related changes in human D1 dopamine receptors measured by positron emission tomography. *Psychopharmacology (Berl)* 103 41-45.
- Sullivan, P.F. (2007) Spurious genetic associations. *Biol Psychiatry* 61 1121-1126.
- Syvanen, A.C. (2005) Toward genome-wide SNP genotyping. *Nat Genet* 37 Suppl S5-10.
- Takahashi, H. & Liu, F.C. (2006) Genetic patterning of the mammalian telencephalon by morphogenetic molecules and transcription factors. *Birth Defects Res C Embryo Today* 78 256-266.
- The Wellcome Trust Case Control Consortium (2007) Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 447 661-678.
- Toga, A.W., Thompson, P.M. & Sowell, E.R. (2006) Mapping brain maturation. *Trends Neurosci* 29 148-159.
- Trimborn, M., Bell, S.M., Felix, C., Rashid, Y., Jafri, H., Griffiths, P.D., Neumann, L.M., Krebs, A., Reis, A., Sperling, K., Neitzel, H. & Jackson, A.P. (2004) Mutations in microcephalin cause aberrant regulation of chromosome condensation. *Am J Hum Genet* 75 261-266.
- Tsai, S.J., Yu, Y.W., Lin, C.H., Chen, T.J., Chen, S.P. & Hong, C.J. (2002) Dopamine D2 receptor and N-methyl-D-aspartate receptor 2B subunit genetic variants and intelligence. *Neuropsychobiology* 45 128-130.
- Volkow, N.D., Ding, Y.S., Fowler, J.S., Wang, G.J., Logan, J., Gatley, S.J., Hitzemann, R., Smith, G., Fields, S.D. & Gur, R. (1996) Dopamine transporters decrease with age. *J Nucl Med* 37 554-559.
- Von der Kammer, H., Albrecht, C., Mayhaus, M., Hoffmann, B., Stanke, G. & Nitsch, R.M. (1999) Identification of genes regulated by muscarinic acetylcholine receptors: application of an improved and statistically comprehensive mRNA differential display technique. *Nucl Acids Res* 27 2211-2218.
- von der Kammer, H., Mayhaus, M., Albrecht, C., Enderich, J., Wegner, M. & Nitsch, R.M. (1998) Muscarinic acetylcholine receptors activate expression of the EGR gene family of transcription factors. *J Biol Chem* 273 14538-14544.
- Wechsler, D. (1986) Wechsler Intelligence Scale for Children-Revised (WISC-R). Lisse: Swets & Zeitlinger.
- Wechsler, D. (1997) WAIS-III Wechsler Adult Intelligence Scale. San Antonio, Texas: Psychological Corporation.
- Weinberger, D.R., Egan, M.F., Bertolino, A., Callicott, J.H., Mattay, V.S., Lipska, B.K., Berman, K.F. & Goldberg, T.E. (2001) Prefrontal neurons and the genetics of schizophrenia. *Biol Psychiatry* 50 825-844.
- Williams, G.V. & Castner, S.A. (2006) Under the curve: critical issues for elucidating D1 receptor function in working memory. *Neuroscience* 139 263-276.
- Winterer, G., Coppola, R., Goldberg, T.E., Egan, M.F., Jones, D.W., Sanchez, C.E. & Weinberger, D.R. (2004) Prefrontal broadband noise, working memory, and genetic risk for schizophrenia. *Am J Psychiatry* 161 490-500.
- Wirtenberger, M., Hemminki, K. & Burwinkel, B. (2006) Identification of frequent chromosome copy-number polymorphisms by use of high-resolution single-nucleotide-polymorphism arrays. *Am J Hum Genet* 78 520-522.
- Wong, D.F., Wagner, H.N., Jr., Dannals, R.F., Links, J.M., Frost, J.J., Ravert, H.T., Wilson, A.A., Rosenbaum, A.E., Gjedde, A., Douglass, K.H. & et al. (1984) Effects of age on dopamine and serotonin receptors measured by positron tomography in the living human brain. *Science* 226 1393-1396.
- Woods, R.P., Freimer, N.B., De Young, J.A., Fears, S.C., Sicotte, N.L., Service, S.K., Valentino, D.J., Toga, A.W. & Mazziotta, J.C. (2006) Normal variants of Microcephalin and ASPM do not account for brain size variability. *Hum Mol Genet* 15 2025-2029.

# SAMENVATTING

## SAMENVATTING

Het doel van het onderzoek zoals beschreven in dit proefschrift was het identificeren van DNA polymorfismes die een gedeelte van de variatie in cognitief vermogen tussen mensen kunnen verklaren. Met behulp van familiale data uit twee leeftijdscohorten werden effecten van genetische varianten in vantevoren geselecteerde kandidaat genen onderzocht. Tevens werd onderzocht of de geselecteerde genetische varianten in verschillende leeftijdscohorten differentiele effecten laten zien.

### *Beschrijving van de onderzoeksgroep*

Voor het onderzoek dat beschreven wordt in dit proefschrift waren twee cohorten beschikbaar: een jong cohort (Polderman *et al.*, 2006a; 2006b) en een volwassen cohort (Posthuma *et al.*, 2005). Beide cohorten bestaan uit tweelingen die opgenomen zijn in het Nederlands Tweelingen Register en hun broers of zussen (Boomsma *et al.* 2006; Boomsma *et al.* 2002) (zie Tabel 8.1). Deelnemers van het Nederlands Tweelingen Register kunnen vrijwillig toestemming geven voor het verzamelen van DNA uit wangslimvlies. Bij alle unisex-tweelingen werd de zygositeit vastgesteld door het genotyperen van 11 polymorfe microsatelliet markers (heterozygositeit > 0.80). Genotyperingen werden uitgevoerd zonder kennis van de familiale status en fenotypische data .

**Tabel 1** Zygositeit van de individuen in het jonge en het volwassen cohort uit deze studie

|                |                  | Jonge Cohort |          | Volwassen Cohort |          |
|----------------|------------------|--------------|----------|------------------|----------|
|                |                  | Gezinnen     | Personen | Gezinnen         | Personen |
| MZM            | Tweelingparen    | 35           | 80       | 25               | 66       |
|                | Incomplete paren | -            | -        | 15               | 17       |
| MZF            | Tweelingparen    | 28           | 65       | 20               | 46       |
|                | Incomplete paren | -            | -        | 15               | 16       |
| DZM            | Tweelingparen    | 48           | 116      | 15               | 40       |
|                | Incomplete paren | -            | -        | 14               | 17       |
| DZF            | Tweelingparen    | 23           | 47       | 28               | 71       |
|                | Incomplete paren | -            | -        | 28               | 45       |
| DOS            | Tweelingparen    | 26           | 60       | 23               | 59       |
|                | Incomplete paren | -            | -        | -                | -        |
| Enkel tweeling |                  | 3            | 3        | 12               | 14       |
| Totaal cohort  |                  | 163          | 371      | 195              | 391      |

## IQ Phenotypes

Psychometrisch IQ werd in het jonge cohort gemeten met zes subtesten van de WISC-R (Wechsler, 1986), namelijk Rekenen, Cijferreeksen, Blokpatronen, Figuur Leggen, Overeenkomsten en Woordenschat. In het volwassen cohort werd IQ gemeten met 11 subtesten van de WAIS-III (Wechsler, 1997): Onvolledige Tekeningen, Rekenen, Blokpatronen, Overeenkomsten, Cijfers en Letters Nazeggen, Informatie, Matrix Redeneren, Symbool Substitutie – Coderen, Symbool Substitutie – Paren Associëren, Symbool Substitutie – Vrije Reproductie en Woordenschat (zie tabel 8.1a). Erfelijkheidsschattingen voor Totaal IQ (TIQ), Verbaal IQ (VIQ) en Performaal IQ (PIQ) staan in Tabel 8.2b. De gevonden erfelijkheidsschattingen zijn vergelijkbaar met eerdere genetische studies naar IQ in jonge en volwassen cohorten in de Nederlands populatie (Bartels *et al.* 2002; Posthuma *et al.* 2001).

**Table 8.2a** Subtesten van de WAIS<sup>A</sup> en de WISC<sup>B</sup>

|            | Subtest                         | Description   |
|------------|---------------------------------|---|
| Verbaal    | 1 Informatie <sup>A</sup>       | Algemene feitelijke kennis, lange termijn geheugen      |
|            | 2 Overeenkomsten <sup>A,B</sup> | Abstract redeneren, categoriseren, relaties leggen      |
|            | 3 Rekenen <sup>A,B</sup>        | Aandacht, concentratie, numeriek redeneren              |
|            | 4 Woordenschat <sup>A,B</sup>   | Woordenkennis, verbale taalbeheersling                  |
|            | 5 Begrijpen <sup>A</sup>        | Sociale regels en concepten begrijpen                   |
|            | 6 Cijferreeksen <sup>A,B</sup>  | Auditief korte termijn geheugen, concentratie           |
| Performaal | 7 Onvolledige                   | Aandacht voor essentiële details                        |
|            | 8 Symbool Substitutie           | -Visuele motorische coördinatie, snelheid, concentratie |
|            | 9 Plaatjes Ordenen <sup>A</sup> | Sequentieel, logisch nadenken                           |
|            | 10 Blokpatronen <sup>B</sup>    | Ruimtelijk inzicht, abstract visuele problemen oplossen |
|            | 11 Figuur Leggen <sup>B</sup>   | Visuele analyse, samenstelling van objecten             |
|            | 12 Symbool Zoeken <sup>A</sup>  | Verwerking van nieuwe informatie                        |
|            | 13 Mazes <sup>A</sup>           | Fijne motorische coördinatie, planning van richtingen   |

**Table 8.2b** Erfelijkheidsschattingen ( $h^2$ ) voor Totaal IQ (TIQ), Verbaal IQ (VIQ) en Performaal IQ (PIQ) voor het jonge en volwassen cohort

| Eigenschap           | Subtest(en)            | Erfelijkheidsschattingen ( $h^2$ ) |
|----------------------|------------------------|------------------------------------|
| TIQ jonge cohort     | 2, 3, 4, 6, 10, 11     | 0.80 (0.72-0.85)                   |
| TIQ volwassen cohort | 1, 2, 3, 4, 10, 12, 13 | 0.78 (0.72-0.83)                   |
| VIQ jonge cohort     | 2, 3, 4, 6             | 0.70 (0.59-0.78)                   |
| VIQ volwassen cohort | 1, 2, 3, 4             | 0.78 (0.72-0.83)                   |
| PIQ jonge cohort     | 10, 11                 | 0.73 (0.63-0.80)                   |
| PIQ volwassen cohort | 10,12,13               | 0.71 (0.62-0.77)                   |

### *Het muscarine acetylcholine receptor type 2 (CHRM2) gen en variatie in cognitief vermogen*

Muscarine acetylcholine receptoren (mAChR) behoren tot een groep van receptoren met zeven transmembraandomeinen waar adrenerge receptoren ook toe behoren. Deze receptoren geven signalen van buiten de cel door via interactie met intracellulaire GTP-bindende eiwitten. Er zijn verschillende macromoleculaire interacties betrokken bij de respons op activatie van muscarinerge receptoren (Hulme 1990), variërend van de remming van adenylyl cyclase tot stimulatie van hydrolyse van phosphoinositide en de regulatie van kaliumkanalen (Caulfield 1993). Promotors van neuronale eiwitten bevatten een bindingsplaats voor diverse transcriptie factoren waarvan de expressie verhoogd wordt door mAChR activatie (Nitsch *et al.* 1998; Von der Kammer *et al.* 1999).

Het *CHRM2* gen is bijzonder interessant voor ons onderzoek omdat het een autoreceptor activiteit heeft in hersenstructuren zoals de hippocampus, die van fundamenteel belang is voor geheugen en leerprocessen (Iannazzo & Majewski 2000). Het gen zou betrokken kunnen zijn bij de fijne afstemming van de terugkoppeling van inhibitie na geheugenvorming (Miranda *et al.* 2000; Orsetti *et al.* 1996). In het onderzoek beschreven in dit proefschrift is onderzocht of er een verband bestaat tussen genotypische variatie in het *CHRM2* gen en variatie in cognitieve fenotypes (Hoofdstukken 2 en 3).

In het volwassen cohort vonden we in intron 4 van het *CHRM2* gen een associatie met variatie in per formaal IQ (PIQ) ( $\chi^2=9.14$ ,  $P=0.003$ ; 6.89 IQ punten bewerkstelligt in per formaal IQ). In het jonge cohort vonden we een associatie van een gedeelte van het gen dat intron 5 bevat (het laatste intron voorafgaand aan de coderende sequentie) met variatie in totaal IQ (TIQ) ( $\chi^2=7.14$ ,  $P=0.008$ ; 5.35 IQ punten bewerkstelligt in totaal IQ) en verbaal IQ (VIQ) ( $\chi^2=9.50$ ,  $P=0.002$ ; 5.30 IQ punten bewerkstelligt in verbaal IQ). Deze associaties bevestigen de resultaten van andere onafhankelijke studies waarin ook werd gevonden dat *CHRM2* een rol speelt in cognitie (Comings *et al.* 2003; Dick *et al.* 2006; Dick *et al.* 2007). Omdat in het jonge cohort een associatie werd gevonden met een ander gedeelte van het *CHRM2* gen dan in het volwassen cohort, zou men kunnen veronderstellen dat de beide regio's regulatoire elementen bevatten die verschillend gebruikt worden in verschillende levensstadia. In een eerste gen-expressie studie (hoofdstuk 3) werd vooralsnog geen verschil gevonden tussen de gen expressie van de verschillende genetische varianten van *CHRM2*.

Het belang van de beschikbaarheid van choline tijdens de hersenontwikkeling en in het bijzonder de ontwikkeling van de hippocampus, wordt ondersteund door een grote hoeveelheid literatuur (Blusztajn and Wurtman 1983; Blusztajn *et al.* 1998;

Glenn *et al.*, 2007; Mellot *et al.*, 2007). Cermak en collega's (Cermak *et al.* 1998) vonden bijvoorbeeld bij muizen een significante associatie tussen de beschikbaarheid van choline en de prestatie in hippocampus-gerelateerde taken. De snelheid van turnover van ACh (d.w.z. synthese, afbraak en choline gebruik) die ten grondslag ligt aan deze associatie leek vroeg in het leven geprogrammeerd te zijn door de hoeveelheid beschikbare ACh tijdens kritieke ontwikkelingsstadia. De hippocampus van muizen met een prenatale choline deficiëntie werd bijvoorbeeld gekarakteriseerd door een versnelde ACh turnover, wat mogelijk wijst op een adaptieve respons op de verlaagde beschikbaarheid van ACh in utero (Cermak *et al.*, 1998). Een dergelijke "metabool imprinting" hypothese is aantrekkelijk vanuit het oogpunt van de epigenetica omdat de hypothese rekening houdt met omgevingsinvloeden die in utero hun uitwerking hebben en die een verklaring kunnen bieden voor het gebrek aan replicatie van associatie studies waarin uitsluitend genetische factoren bestudeerd worden (Niculescu *et al.* 2006).

### *De rol van het "synaptosomal protein of 25 kDa" (SNAP-25) gen in synaptische plasticiteit en cognitieve vaardigheden*

Het "synaptosomal protein of 25 kDa" (SNAP-25) gen was één van de kandidaat genen die in deze studie is onderzocht (Hoofdstuk 4 en 5). Verschillende studies hebben een associatie gevonden tussen dit gen en verstoord cognitief functioneren, variërend van ADHD tot schizofrenie. SNAP-25 is actief in verschillende organen en weefsels. In de hersenen, en met name in de hippocampus, is SNAP-25 noodzakelijk voor 'long-term potentiation' (LTP) en het geheugen (Hou *et al.* 2004; Hou *et al.* 2006). In de hersenen komen de isoformen SNAP-25a en SNAP-25b voor, deze zijn ontstaan door alternatieve splicing van exon 5 (Bark & Wilson 1991). Tijdens de ontwikkeling is SNAP-25a de meest voorkomende isoform. Deze isoform is onder andere betrokken bij synaptogenese. Echter, in het volwassen brein is SNAP-2b de meest voorkomende isoform. Dit eiwit vormt een zogenaamd "fusion machinery complex" (SNARE) samen met syntaxine en de twee synaptische vesicle proteïnen synaptobrevine en synaptotagmine. Deze eiwitten faciliteren exocytose van neurotransmitters uit synaptische vesicles in de synaptische spleet. Beide isoformen zijn noodzakelijk om de balans tussen de vorming van synapsen en het vrijkomen van neurotransmitter vesicles te bewaren. In de hoofdstukken 4 en 5 wordt een genomische variatie geschreven in intron 1 van het SNAP-25 gen dat geassocieerd wordt met een variatie in het IQ fenotype (d.w.z. Totaal IQ (TIQ):  $\chi^2 = 15.99$   $P = 0.0001$ , 3.28 IQ punten bewerkstelligt in totaal IQ, Verbaal IQ (VIQ):  $\chi^2 = 13.01$   $P = 0.0003$ , 2.76 IQ punten bewerkstelligt in verbaal IQ, en Performa IQ (PIQ):  $\chi^2 = 11.22$   $P = 0.0008$ , 3.21 IQ punten bewerkstelligt in performa IQ).

Deze (niet-)coderende variant in intron 1 is mogelijk betrokken bij de regulatie van de expressie van de verschillende isoformen, omdat alle geassocieerde SNPs gelokaliseerd zijn in mogelijke “transcription factor binding sites” (TFBS). *Quantitative trait loci* (QTL) met klein effect kunnen complexe aandoeningen verklaren. Het is hierdoor ook mogelijk dat variaties in niet-coderende gebieden een effect hebben op regulatieve sequenties (bv. TFBS, promoters, enhancers). De regulatie van gen expressie kan op deze manier subtieler beïnvloed worden dan door polymorfismen in coderende gebieden.

Naast de verklaring dat deze geassocieerde varianten biologisch relevant zijn, is er ook een andere mogelijkheid. Namelijk, dat deze geassocieerde varianten in “linkage disequilibrium” (LD) zijn met de *causale* polymorfisme(n). Functionele studies zijn noodzakelijk om uit te zoeken of deze polymorfismen inderdaad functionele varianten zijn die variatie in cognitie kunnen verklaren.

### *Functionele niet-synonieme polymorfisme onder positieve evolutionaire selectie ligt aan basis van fenotypische verschillen in cognitief vermogen*

Op basis van een vergelijkende genoom analyse werd een tweedelig studieopzet voorgesteld waarbij in genen, die een verhoogde snelheid van eiwit evolutie laten zien in primaten, gezocht werd naar het voorkomen van niet-synonieme coderende SNPs die verschillen tussen de huidige menselijke populatie en de chimpansee (Hoofdstuk 6). Deze studie was gericht op genen die tot expressie komen in het centraal zenuwstelsel, omdat onze belangstelling uitging naar het vinden van genen die betrokken zijn bij fenotypische verschillen in cognitief vermogen. Omdat er in het algemeen wordt aangenomen dat genen die onder positief selectieve druk staan tussen primaten de verschillen tussen mens en chimpansee bewerkstelligen ligt het voor de hand om te veronderstellen dat amino zuur substituties op deze positief geselecteerde eiwitposities fenotypische verschillen tussen mensen zouden kunnen verklaren. Na dergelijke kandidaatgenen geselecteerd te hebben werd een associatie studie tussen cognitief vermogen en de  $\beta$ -2 adrenerge receptor uitgevoerd. De  $\beta$ -2 adrenerge receptor behoort tot de G-eiwit gekoppelde receptor superfamilie en bemiddelt in verschillende weefsels (Liggett 2000) enkele fysiologische effecten van catecholamines (noradrenaline en adrenaline) als gevolg van activatie door noradrenaline (NA) (Kobayashi & Kobayashi 2001). Verschillende studies hebben de betrokkenheid van ADRB2 bij geheugenformatie en leervermogen vastgesteld. “long-term potentiation” (LTP) in de hippocampus neemt toe na activatie van de  $\beta$ -2 receptor, en deze activatie induceert op zijn beurt LTP (Hillman *et al.* 2005). In overeenstemming met deze bevinding is aangetoond dat het gebruik van  $\beta$ 2- agonisten (Gibbs & Summers 2000) prestaties in geheugen- en leervermogen toetsen verhoogd, terwijl  $\beta$ 2-



antagonisten het vasthouden van geheugen verstoren (Gibbs & Summers 2005). Voor deze studie zijn twee niet-synonieme coderende SNPs (rs1042713 and rs1042714) in de  $\beta$ -2 adrenerge receptor gekozen om een genetische associatiestudie met IQ uit te voeren in twee onafhankelijke Nederlandse cohorten van verwanten. Het is treffend dat werd gevonden dat juist de mens-specifieke variant van de receptor (rs1042713) een toename van 8 IQ punten bewerkstelligt in verbaal IQ. Enige terughoudendheid bij deze vinding is wel op zijn plaats aangezien effect groottes in relatief kleine sample vaak worden overschat.

### *Positieve heterosis en gen-gen interactie bij het functioneren van het werkgeheugen*

In Hoofdstuk 7 hebben we getest of er een associatie van het Val<sup>108/158</sup>Met polymorfisme in catechol O-methyl transferase (COMT) met werkgeheugen (WG) was. Met/Val heterozygoten bleken beter te presteren in de WG taken dan homozygoten. Met andere woorden: we vonden aanwijzingen voor positieve heterosis. De associatie werd gevonden in het volwassen cohort, maar niet in het jonge cohort. In een longitudinale studie van Harris en collega's (2005) werden leeftijdsafhankelijke effecten van positieve heterosis gevonden. In de literatuur is gesuggereerd dat het COMT polymorfisme een niet-lineair effect heeft op de beschikbaarheid van DA in de prefrontale cortex (PFC) (Mattay *et al.* 2003). Daarnaast is gevonden dat het verband tussen DA signalering en cognitief vermogen een omgekeerde U-vormige curve volgt, waarbij zowel suboptimale als supra-optimale DA activiteit het functioneren van de prefrontale cortex belemmert (Cools & Robbins 2004). Onze bevinding van heterosis is in overeenstemming met deze ideeën. De DA activiteit hangt echter niet alleen af van de afbraaksnelheid (COMT), maar ook van de aanwezigheid van DA receptoren. Cognitief vermogen zou daarom ook sterk af kunnen hangen van de D1/D2 bindingsratio, waarbij een relatief tekort aan D1 signalering impulsiviteit, aandachtstekort en slecht werkgeheugen veroorzaakt, met schizofrenie aan het extreme einde van het spectrum (Winterer *et al.* 2004). Een relatief tekort aan D2 signalering zou het verwerken van informatie over beloning kunnen verstoren. De PFC heeft dit signaal nodig om het werkgeheugen bij te kunnen werken (Weinberger *et al.* 2001). Individuele verschillen in zowel DA beschikbaarheid als D2 receptor gevoeligheid kunnen daarom een rol kunnen spelen bij de prestatie op WG taken. Deze hypothese hebben wij getest met een tweede analyse, waarin we onderzochten of er interactie was tussen een DRD2 tagging SNP (rs2075654) en het COMT polymorfisme. Er werd geen significant hoofdeffect gevonden van de rs2075654 tag-SNP op WG, noch in het jonge cohort, noch in het volwassen cohort. Er was

echter wel een significante interactie tussen het DRD2 genotype en het effect van het COMT polymorfisme op WG wanneer de jonge en volwassen groep werden gecombineerd. In het onderzoek beschreven in Hoofdstuk 7 werd gevonden dat de Met/Val heterozygoten uitsluitend beter presteerden dan Met/Met en Val/Val homozygoten wanneer ze tenminste één A1 allel droegen. Een dergelijk patroon is eerder beschreven door Reuter *et al.* (2005), die een significante interactie vonden tussen het DRD2 Taq IA polymorfisme en het COMT polymorfisme bij de mate van respons interferentie in de Stroop kleur-woord test. Zowel bij dieren (Lee *et al.* 2001) als bij mensen (Kaasinen & Rinne 2002) is bewijs gevonden voor leeftijdsafhankelijke veranderingen in DA metabolisme in de PFC, waarbij verondersteld wordt dat er op latere leeftijd een verhoogd DA metabolisme is (bv. MAO, COMT; (Gottfries 1990), terwijl het aantal DA receptoren afneemt met de leeftijd (Suhara *et al.* 1991; Volkow *et al.* 1996; Wong *et al.* 1984). De volledige genetische bijdrage aan dopaminerge variatie in frontale executieve functies zal afhangen van veel complexere interacties tussen polymorfismen in genen die coderen voor receptoren (bv. DRD1, DRD2, DRD4), transporters en enzymen (bv. DAT, COMT, MAO) (Berman & Noble 1995; Bertolino *et al.* 2006; Tsai *et al.* 2002; Williams & Castner 2006). Zulke interacties dienen verder bestudeerd te worden om een duidelijker overzicht te krijgen van de rol van de dopaminerge pathway bij het presteren in geheugen taken.

### *Dopaminerge en serotonerge kandidaatgenen*

Veel voorkomende varianten in andere vermeende kandidaatgenen uit het dopaminerge (*DBH*, *DRD2*, *DRD3* en *TH*) en serotonerge (*HTR2A* en *SERT*) systeem werden ook onderzocht in dit proefschrift. Wegens het uitvoeren van meerdere statistische testen en de vergrootte kans die daardoor ontstaat op het vinden van vals-positieve uitkomsten, is een strenge correctie op de statistische significantie uitgevoerd. Geen van bovengenoemde genen liet na deze correctie een significant effect zien. Een complete tabel met uitkomsten van deze analyse op opgenomen in de bijlage zie Tabellen in bijlage I).

Samenvattend zijn er in een kandidaatgenstudie twee typen polymorfismen gevonden die associatie vertonen met variatie binnen verschillende cognitieve vaardigheden: coding oftewel functionele polymorfismen (in *COMT* en *ADRB2*) en (non)coding polymorfismen (in *CHRM2*, *SNAP-25* en *DRD2*). Effecten van het eerste type impliceren veranderingen die op eiwit niveau geobserveerd kunnen worden (zoals b.v. enzymactiviteit), terwijl het andere type beschouwd moet worden als onderdeel van een groter spectrum van regulatoire elementen waarvan de genetische contributie, alhoewel klein in omvang, noodzakelijk is voor het

uiteindelijke fenotype. Toekomstige functionele studies, in combinatie met analyse van gen expressie profielen van verschillende hersengebieden, kunnen helpen in het verkrijgen van inzicht in de rol van genetische varianten en hun relatie met synaptische plasticiteit die ten grondslag ligt aan cognitie, leer- en attentieproblemen.



## RESUMEN

## RESUMEN

En el presente trabajo, un estudio de asociación fue realizado utilizando una serie de genes candidatos (ver Tabla 1.1, capítulo 1) con el fin de indentificar polimorfismos en el ADN capaces de explicar variancia en habilidad cognitiva. La utilización de dos cohortes de edades diferentes, permitió la detección de efectos ejercidos por polimorfismos localizados en genes candidatos, y de manera más específica, de polimorfismos que se encontrarían ejerciendo efectos durante diferentes etapas en la vida.

*Descripción de cohortes*

En la presente tesis, información sobre dos grupos de individuos de edades diferentes fue utilizada: una cohorte de niños (Polderman *et al.*, 2006a; 2006b) y una cohorte de adultos (Posthuma *et al.*, 2005). Ambos cohortes constituídos por gemelos y hermanos fraternos fueron recrutados de el Registro de Gemelos Holandés (Boomsma *et al.* 2006; Boomsma *et al.* 2002) (ver Tabla 1.1). La participación en el presente estudio incluyó un acuerdo voluntario para proveer hisopados bucales y/o muestras de sangre venosa para la subsecuente extracción de ADN. La prueba de zigosidad de gemelos fue realizada en gemelos de igual sexo mediante el uso de 11 marcadores genéticos polimórficos (Heterozygosidad > 0.80). El genotipado fue llevado a cabo independientemente de el conocimiento de el estatus familiar o de la información fenotípica disponible (ver Tabla 8.1).

**Tabla 8.1** Zigosidad de individuos para cohortes de niños y adulta utilizados en el presente estudio

|               |                         | Cohorte de niños |         | Cohorte de Adultos |         |
|---------------|-------------------------|------------------|---------|--------------------|---------|
|               |                         | Familias         | Sujetos | Familias           | Sujetos |
| MZM           | <i>pares de gemelos</i> | 35               | 80      | 25                 | 66      |
|               | <i>gemelo único</i>     | -                | -       | 15                 | 17      |
| MZF           | <i>pares de gemelos</i> | 28               | 65      | 20                 | 46      |
|               | <i>gemelo único</i>     | -                | -       | 15                 | 16      |
| DZM           | <i>pares de gemelos</i> | 48               | 116     | 15                 | 40      |
|               | <i>gemelo único</i>     | -                | -       | 14                 | 17      |
| DZF           | <i>pares de gemelos</i> | 23               | 47      | 28                 | 71      |
|               | <i>gemelo único</i>     | -                | -       | 28                 | 45      |
| DOS           | <i>pares de gemelos</i> | 26               | 60      | 23                 | 59      |
|               | <i>gemelo único</i>     | -                | -       | -                  | -       |
| Gemelo único  |                         | 3                | 3       | 12                 | 14      |
| Cohorte total |                         | 163              | 371     | 195                | 391     |

### *Fenotipos de cociente intelectual (CI)*

Para determinar el coeficiente intelectual (CI) en la cohorte de niños, fue utilizado el test WISC-R (1986) (Wechsler 1986). Seis subtests fueron usados: Similitudes, Aritmética, Vocabulario, Serie de dígitos, Objetos, y Diseño con cubos. El CI fue determinado en la cohorte adulta mediante la utilización de la versión holandesa del test WAIS-III (1997) (Wechsler 1997). Once subtests fueron utilizados: Información, Similitudes, Aritmética, Vocabulario, Búsqueda de símbolos, Codificación, Codificación libre, Rompecabezas, Completamiento de figuras, Ordenamiento de matrices y Ordenamiento de números-letras (ver Tabla 8.2a). Las estimaciones de heritabilidad para CI Total (CIT), CI verbal (CIV), y CI de ejecución (CIE) son dados en la Tabla 8.2b. Estas estimaciones de heredabilidad son comparables con aquellas reportadas previamente para cohortes consistentes en niños y adultos en la población holandesa (Bartels *et al.* 2002; Posthuma *et al.* 2001).

**Tabla 8.2a** Subtests comprendidos en los tests WAIS<sup>A</sup> and WISC<sup>B</sup>

|                    | <b>Subtest</b>                           | <b>Descripción</b>  |
|--------------------|--|---|
| <b>Verbal</b>      | 1 Información <sup>A</sup>               | <i>Conocimiento general, memoria a largo plazo</i>              |
|                    | 2 Similitudes <sup>A,B</sup>             | <i>Razonamiento abstracto, categorías, relaciones</i>           |
|                    | 3 Aritmética <sup>A,B</sup>              | <i>Atención, concentración, razonamiento numérico</i>           |
|                    | 4 Vocabulario <sup>A,B</sup>             | <i>Conocimiento de palabras, fluidez verbal</i>                 |
|                    | 5 Comprensión <sup>A</sup>               | <i>Razonamiento, sentido común</i>                              |
|                    | 6 Digit Span <sup>A,B</sup>              | <i>Memoria auditiva a corto plazo, concentración</i>            |
| <b>Performance</b> | 7 Completamiento de figuras <sup>A</sup> | <i>Alertness to essential detail</i>                            |
|                    | 8 Codificación <sup>A</sup>              | <i>Visual motor co-ordination, speed, concentración</i>         |
|                    | 9 Rompecabezas <sup>A</sup>              | <i>Razonamiento lógico y secuencial</i>                         |
|                    | 10 Diseño con cubos <sup>B</sup>         | <i>Resolución de problemas abstractos, espacio-visuales</i>     |
|                    | 11 Object Assembly <sup>B</sup>          | <i>Análisis visual, construcción de objetos</i>                 |
|                    | 12 Búsqueda de símbolos <sup>A</sup>     | <i>Velocidad de procesamiento de información</i>                |
|                    | 13 Ordenamiento de matrices <sup>A</sup> | <i>Fine motor co-ordination, planning, following directions</i> |

**Tabla 8.2b** Estimaciones de heredabilidad ( $h^2$ ) para CI Total (CIT), CI Verbal (CIV), y CI de Ejecución (CIE) para las cohortes de niños y adultos

| <b>Fenotipos Cognitivos (subtests)</b> | <b>Subtests</b>        | <b>Heritabilidad (<math>h^2</math>) (95% CI)</b> |
|--|------------------------|--|
| CIT niños                              | 2, 3, 4, 6, 10, 11     | 0.80 (0.72-0.85)                                 |
| CIT adultos                            | 1, 2, 3, 4, 10, 12, 13 | 0.78 (0.72-0.83)                                 |
| CIV niños                              | 2, 3, 4, 6             | 0.70 (0.59-0.78)                                 |
| CIV adultos                            | 1, 2, 3, 4             | 0.78 (0.72-0.83)                                 |
| CIE niños                              | 10, 11                 | 0.73 (0.63-0.80)                                 |
| CIE adultos                            | 10, 12, 13             | 0.71 (0.62-0.77)                                 |

### *Gen de receptor muscarínico colinérgico tipo 2 (CHRM2) y variancia en habilidad cognitiva*

Los receptores muscarínicos colinérgicos (mAChR) son receptores responsables de la activación de numerosas señales celulares, mediadas vía diferentes tipos de proteínas-G intracelulares (Hulme 1990), desde inhibición de la adenilciclase, estimulación de la hidrólisis del fosfatidil inositol y regulación de canales de potasio (Caulfield 1993). También es sabido que promotores de proteínas neuronales contienen secuencias consensuales para diversos factores de transcripción cuya expresión es aumentada por la activación de mAChR en diferentes células neuronales (Nitsch et al. 1998; Von der Kammer et al. 1999). La variabilidad genotípica a nivel de receptores muscarínicos colinérgicos tipo 2 (*CHRM2*) fue investigada en relación a la variancia de fenotipos cognitivos (Capítulos 2 y 3). La razón por la cual este gen candidato es de particular interés radica en su propiedad como autoreceptor a nivel de estructuras cerebrales, en particular a nivel de el hipocampo, estructura fundamentalmente involucrada en funciones de memoria y procesos de aprendizaje (Iannazzo & Majewski 2000). El gen *CHRM2* estaría involucrado en el ajuste final de la retro-inhibición luego de la haberse producido la formación de memorias (Miranda *et al.* 2000; Orsetti *et al.* 1996).

La primera región putativa que se halló asociada con variancia a nivel de CI (i.e. Coeficiente intelectual de Ejecución (CIE):  $\chi^2=9.14$ ,  $P=0.003$ ; confiriendo un incremento de 6.89 puntos en CIE) en la cohorte consistente en adultos se localiza en el intron 4. La segunda región, localizada el intrón 5 – el último intrón antes del comienzo de la secuencia de codificación del gen – se halló asociada con variancia en CI (i.e. CI Total (CIT): ( $\chi^2=7.14$ ,  $P=0.008$ ; confiriendo un incremento en 5.35 puntos en CIT , y CI Verbal (CIV):  $\chi^2=9.50$ ,  $P=0.002$ ; confiriendo un incremento de 5.30 puntos en CIV) en la cohorte de niños. Si bien nuestros resultados en referencia a la expresión genética de dicho gen no pudieron ser confirmados, diversos estudios independientes previos sí han podido evidenciar un rol fundamental de el gen *CHRM2* en habilidad cognitiva (Comings *et al.* 2003; Dick *et al.* 2006; Dick *et al.* 2007). Debido a que los resultados de asociación para las cohortes de niños y adultos fueron hallados en diferentes regiones del gen *CHRM2*, se podría proponer la existencia de elementos con actividad regulatoria en ambas regiones asociadas los cuales serían diferencialmente utilizados durante diferentes etapas de la vida. *De qué manera estas diferentes asociaciones podrían ubicarse en el contexto de un posible modelo biológico?* Tendiendo en cuenta, por ejemplo, el rol esencial de la colina en el desarrollo cerebral, y más específicamente durante el



desarrollo del hipocampo (Blusztajn and Wurtman 1983, Blusztajn *et al.* 1998 Glenn *et al.*, 2007; Mellot *et al.*, 2007). Siguiendo la misma línea de pensamiento, Cermak y colaboradores (Cermak *et al.* 1998,; Cermak *et al.*, 1999) reportaron una asociación significativa entre la disponibilidad de colina y la capacidad en tareas hipocampo-dependientes. Particularmente, la tasa de conversión de Ach (i.e. síntesis, degradación y utilización de colina) relacionada con dicha asociación podría encontrarse programada de manera temprana en la vida en relación con la cantidad de Ach disponible durante estadios críticos en el desarrollo. Por ejemplo, el hipocampo de aquellos animales que recibieron durante períodos prenatales cantidades insuficientes de colina se ha caracterizado por una tasa elevada de conversión, indicando posiblemente, una respuesta de tipo adaptativa a la disponibilidad reducida de colina *in utero* (Cermak *et al.*, 1998). Esta hipótesis de imprinting metabólico es atractiva desde el punto de vista epigenético, ya que con la misma se considerarían factores medio-ambientales ocurridos *in útero* (i.e. regulación de la expresión genética a través de metilación de ADN), la cual explicaría la carencia de replicación en estudios de asociación cuando sólo cuando factores genéticos son considerados (Niculescu *et al.* 2006).

### *El rol del gen de la proteína sinaptosómica de 25 kDa (SNAP-25) en plasticidad sináptica y su relación con actividad cognitiva*

El gen de la proteína sinaptosómica de 25 kDa (SNAP-25) fue también considerado como un gen candidato en investigado en el presente estudio (Capítulos 4 y 5). Diferentes estudios han asociado a este gen con un amplio espectro de desórdenes cognitivos, desde fenotipos relacionados con ADHD hasta esquizofrenia. La actividad del gen de SNAP-25 es ejercida a lo largo de diferentes órganos y tejidos. A nivel cerebral, los efectos de SNAP-25 podrían ser de fundamental importancia a nivel de el hipocampo, más específicamente en relación con potenciación a largo plazo (PLP) y consolidación de memorias (Hou *et al.* 2004; Hou *et al.* 2006). El ajuste alternativo (*alternative splicing*) del exón 5 da lugar a dos isoformas proteicas: SNAP-25a y SNAP-25b (Bark & Wilson 1991). Durante el desarrollo embrionario, SNAP-25a constituye la mayor isoforma presente, la cual se encuentra íntimamente relacionada con el proceso de sinaptogénesis. Contrariamente, en el cerebro adulto, SNAP-25b es la isoforma predominante, esta isoforma constituye la maquinaria de fusión de el complejo denominado SNARE (por sus siglas en inglés de *soluble NSF attachment receptor*) junto con syntaxina y proteínas vesiculares sinápticas (synaptobrevina y synaptogamina), las cuales median la exocitosis de neurotransmisores en vesículas sinápticas hacia el botón sináptico. Ambas isoformas son fundamentales para mantener el delicado balance necesario para la formación sináptica y la liberación

de vesículas transportadoras de neurotransmisores. Nuestros resultados (Capítulos 4 y 5) evidenciaron variación genómica en el intrón 1 del gen *SNAP-25* asociada con variancia en fenotipos de CI (i.e. Cociente Intelectual Total (CIT):  $\chi^2 = 15.99$   $P = 0.0001$ , incremento de 3.28 puntos en CIT, Cociente Intelectual Verbal (CIV):  $\chi^2 = 13.01$   $P = 0.0003$ , incremento de 2.76 puntos en CIV y Cociente Intelectual de Ejecución (CIE):  $\chi^2 = 11.22$   $P = 0.0008$ , incremento de 3.21 puntos en CIE).

Variantes genéticas (no)codificantes presentes en intrón 1 serían en parte responsables de la regulación de la expresión de dichas isoformas proteicas. Todos los SNPs asociados se hallan localizados en secuencias relacionadas con factores de transcripción de genes (TFBS, por sus siglas en inglés).

Teniendo en mente que estas características complejas podrían estar influenciadas por QTLs con un efecto relativamente pequeño, variantes (no)codificantes que se encuentren afectando secuencias con actividad regulatoria (e.g. TFBS, promotores, facilitadores) podrían afectar la expresión y función genética de una manera mucho más sutil en comparación con efectos ejercidos por polimorfismos ubicados en regiones codificantes.

Las variantes asociadas que han sido reportadas pueden tratarse, por cierto, de variantes con un rol biológico fundamental. Sin embargo, otra posibilidad que debe tenerse también en cuenta, es la existencia de un elevado grado de desequilibrio de ligamiento entre dichas variantes y las mutaciones causales. Es por eso que subsecuentes estudios funcionales son necesarios para corroborar, y ser capaces de catalogar a estos polimorfismos como variantes funcionales con capacidad de regular la variancia a nivel de fenotipos cognitivos.

### *Polimorfismos (no)codificantes funcionales bajo selección evolutiva positiva como base de diferencias fenotípicas en habilidades cognitivas*

Utilizando un estudio de comparación de genomas, se propuso y realizó un diseño en dos etapas en el cual aquellos genes seleccionados para la evolución proteica en primates fueron subsecuentemente investigados en función a la presencia de polimorfismos codificantes no-sinónimos (SNP, por sus siglas en inglés) en la actual población humana en aquellas posiciones aminoacídicas que difieren entre humanos y chimpancé (Capítulo 6). En este estudio, debido a que nuestro interés primordial fue enfocado en hallar aquellos genes relacionados con diferencias fenotípicas a nivel de habilidades cognitivas, la búsqueda se concentró por lo tanto, en genes expresados a nivel de el sistema nervioso central (SNC). Debido a que genes seleccionados de manera positiva en primates son considerados presumiblemente como genes capaces de determinar diferencias fenotípicas entre

humanos y primates no-humanos, es de esperar que sustituciones aminoacídicas que segregen en humanos en posiciones aminoacídicas seleccionadas positivamente sean determinantes capaces de afectar diferencias fenotípicas en estos últimos.

Luego de seleccionar genes candidatos conteniendo dichas substituciones, un estudio de asociación fue realizado entre habilidad cognitiva y el gen *ADRB2* (*receptor adrenérgico beta tipo 2*). Los receptores  $\beta$ -adrenérgicos pertenecen a la superfamilia de receptores ligados a proteínas-G los cuales a su vez median a funciones fisiológicas de catecolaminas (noradrenalina y adrenalina) en una gran variedad de tejidos (Liggett 2000) mediante la activación mediada por noradrenalina (NA) (Kobayashi & Kobayashi 2001). Diversas líneas de investigación han postulado un rol fundamental del gen *ADRB2* en procesos que involucran la formación de memorias y aprendizaje. Se ha demostrado un incremento en la potenciación a largo-plazo a nivel del hipocampo luego de la activación de receptores-beta2 (Hillman *et al.* 2005). En línea con este concepto, el uso de  $\beta$ 2-agonistas (Gibbs & Summers 2000) ha demostrado un incremento en la performance relacionada con tareas de memoria y aprendizaje, mientras que ciertas alteraciones en la consolidación de memorias fueron observadas ante el uso de  $\beta$ 2-antagonistas (Gibbs & Summers 2005). Dos polimorfismos no-sinónimos codificantes bajo selección positiva (rs1042713 and rs1042714) en el gen del receptor adrenérgico beta2 fueron seleccionados para realizar un estudio de asociación basado en familias en relación con CI en dos cohortes independientes holandesas. Como resultado se pudo demostrar que el alelo derivado humano-específico de el gen del receptor adrenérgico  $\beta$ 2 (rs1042713) confiere un incremento en el CIV de 8 puntos, de todas maneras, este incremento debe de ser tomado con precaución, debido a inflaciones de el efecto genético debido a un número de muestras relativamente pequeño.

### *Heterosis positiva e interacción gen-gen en el funcionamiento de memoria de trabajo*

En el capítulo 7, la asociación del polimorfismo de *COMT* Val108/158Met fue estudiada en relación con el funcionamiento de trabajo de memoria (WM, por sus siglas en inglés). La asociación reflejó heterosis positiva en la cual los heterozygotos Met/Val demostraron una mejor capacidad de trabajo comparados con ambos grupos de homozygotos en las tareas de WM. La asociación fue encontrada en la cohorte de adultos pero no en la cohorte joven. Análogamente, un patrón positivo de heterosis edad-dependiente fue reportado en un estudio longitudinal por Harris y colaboradores (2005). Debido a que la hipótesis de un efecto no lineal en la disponibilidad de DA en la corteza prefrontal ha sido

propuesto en relación a el polimorfismo en el gen de *COMT* (Mattay *et al.* 2003), el hallazgo de heterosis se encuentra en línea con la idea de que la relación entre la señalización de DA y la performance cognitiva seguiría un patrón de U-invertida, en donde ambos extremos de actividad dopaminérgica (sub-óptimal y supra-óptimal) dan lugar a alteraciones en la función prefrontal (Cools & Robbins 2004). Sin embargo, la actividad de DA no es dependiente sólo de su tasa catabólica sino que también depende de la presencia de receptores de DA. En línea con este concepto, la capacidad cognitiva también podría depender en forma crítica de la capacidad de ligamiento entre receptores D1/D2, en la cual, una carencia relativa en la señal de D1 causaría impulsividad, distracción y trabajo de memoria escaso, y en donde manifestaciones como, por ejemplo, esquizofrenia, podría encontrarse formando parte de uno de los fenotipos más extremos (Winterer *et al.* 2004). De manera opuesta, una escasez relativa en la señal de D2 podría comprometer señales relacionadas con información de recompensa, señales necesarias para comprometer a la corteza prefrontal (CPF) en la actualización de su sistema de memoria de trabajo (Weinberger *et al.* 2001). Consecuentemente, las diferencias individuales con respecto a la disponibilidad de DA, como también la sensibilidad experimentada por receptores tipo D2 podrían entrar en juego durante la performance de tareas que involucren memoria de trabajo. Esta expectativa fue examinada en un análisis secundario en el cual el SNP rs2075654 (*DRD2*) fue analizado bajo un modelo interactivo junto con polimorfismo de *COMT*. Nuestros resultados sugieren que la interacción sería más fuerte en la cohorte adulta. A pesar que no se obtuvieron valores significativos en ninguna de las cohortes, esta diferencia de edades debería ser particularmente considerada. A pesar que no pudo observarse un efecto significativo en relación a memoria de trabajo para el polimorfismo rs2075654 en ninguna de las cohortes, el análisis combinado de ambas cohortes para los polimorfismos de *DRD2* y *COMT* mostraron un efecto interactivo significativo para la realización de tareas que involucren memoria de trabajo. Dicha interacción sugiere que heterocigotes *Met/Val* tendrían una mejor capacidad de respuesta para tareas de memoria que ambos homocigotos *Met/Met* y *Val/Val*, únicamente cuando estos son portadores de por lo menos un alelo A1.

La evidencia para cambios relacionados con la edad, en función a el metabolismo de DA en la corteza prefrontal ha sido reportada en estudios realizados en animales (Lee *et al.* 2001) y en humanos (Kaasinen & Rinne 2002), en la cual se postula un incremento de el metabolismo dopaminérgico (e.g *MAO*, *COMT*), a edades más avanzadas (Gottfries 1990), como también una relación inversamente proporcional de densidad de receptores de DA con el aumento de la edad (Suhara *et al.* 1991; Volkow *et al.* 1996; Wong *et al.* 1984). Este patrón ha sido reportado previamente por Reuter *et al.* (2005) quien describió un efecto interactivo

significante entre *DRD2 Taq IA* y el polimorfismo de *COMT* en relación a la respuesta de interferencia en la prueba de conflicto de Stroop (del inglés *Stroop color-word conflict task*). Sin embargo, la contribución genética total para la variación dopaminérgica en funciones frontales de ejecución se encuentran con toda probabilidad basadas en interacciones mucho más complejas entre numerosos receptores (e.g. *DRD1*, *DRD2*, *DRD4*), transportadores y polimorfismos localizados en proteínas con capacidad enzimática (e.g. *DAT*, *COMT*, *MAO*) (Berman & Noble 1995; Bertolino *et al.* 2006; Tsai *et al.* 2002; Williams & Castner 2006). Estudios futuros en los cuales se tengan en cuenta dichas interacciones son necesarios para lograr obtener una visión más precisa del metabolismo dopaminérgico en relación con tareas que involucren memoria de trabajo

### *Genes candidates dopaminérgicos y serotoninérgicos*

Otras variantes genómicas comunes dentro de tanto genes candidatos dopaminérgicos (*DBH*, *DRD2*, *DRD3*, and *TH*) como serotoninérgicos (*HTR2A* and *SERT*) fueron analizados en el marco de un estudio de asociación familiar con habilidad cognitiva. Los resultados preliminares no fueron significativos tendiendo en cuenta el umbral de corrección considerando pruebas múltiples. Investigaciones futuras que incluyan un mayor número de variantes genómicas en éstos genes candidatos serían necesarias ya que los resultados presentados en este trabajo no pueden considerarse conclusivos (ver Tablas en Appendix I).

En resumen, dos tipos de polimorfismos en el marco de un diseño utilizando genes candidatos fueron hallados asociados con variancia a nivel de fenotipos de habilidad cognitiva: polimorfismos codificantes, o también llamados “funcionales” (genes de *COMT* y *ADRB2*) y polimorfismos no-codificantes (genes de *CHRM2*, *SNAP-25* y *DRD2*). Mientras que los efectos del primer tipo de polimorfismos implican cambios que son observados a nivel proteico (i.e. actividad enzimática), cambios en el segundo tipo deben de ser considerados como parte de un espectro más amplio de elementos regulatorios, en los cuales la contribución genética, a pesar de ser pequeña en cuanto a su efecto, es necesaria para el fenotipo final resultante. En el futuro, estudios funcionales en combinación con el análisis de perfiles de expresión genética en diferentes regiones cerebrales, podrán servir de ayuda para interpretar el rol de variantes genéticas a nivel de plasticidad sináptica y su influencia en características de cognición, aprendizaje y atención.



## APPENDICES

**Table A1.1a** Mean IQ and Standard deviation (SD) for *DBH* gene for young and adult cohorts

| <i>Tag-SNP</i>            |                  | Young Cohort   |                |                |         | Adult Cohort   |                |                |         |
|---------------------------|------------------|----------------|----------------|----------------|---------|----------------|----------------|----------------|---------|
| position (bp)             | Phenotype        | Genotype       |                |                | Total N | Genotype       |                |                | Total N |
|                           |                  | AA             | AG             | GG             |         | AA             | AG             | GG             |         |
|                           | <i>Frequency</i> | 0.29           | 0.46           | 0.25           |         | 0.31           | 0.46           | 0.23           |         |
| rs1029372<br>(135468263)  | FIQ (SD)         | 103.84 (11.41) | 104.51 (9.69)  | 102.65 (12.99) | 367     | 99.36 (14.64)  | 100.53 (15.03) | 93.94 (15.98)  | 361     |
|                           | VIQ (SD)         | 104.09 (12.12) | 104.70 (10.57) | 103.64 (14.34) | 367     | 103.51 (12.66) | 103.26 (13.18) | 99.66 (12.37)  | 362     |
|                           | PIQ (SD)         | 104.32 (12.47) | 105.53 (10.88) | 102.27 (11.92) | 367     | 94.88 (18.61)  | 97.06 (19.14)  | 89.57 (19.42)  | 361     |
|                           | <i>Frequency</i> | 0.60           | 0.36           | 0.04           |         | 0.62           | 0.34           | 0.04           |         |
| rs1108581<br>(135495062)  | FIQ (SD)         | 104.47 (10.61) | 103.01 (11.97) | 102.99 (11.20) | 368     | 97.10 (15.68)  | 100.61 (14.81) | 105.25 (11.63) | 349     |
|                           | VIQ (SD)         | 105.06 (11.44) | 103.12 (13.12) | 103.45 (12.18) | 368     | 101.16 (13.20) | 104.22 (12.69) | 109.46 (9.30)  | 350     |
|                           | PIQ (SD)         | 104.67 (11.67) | 103.88 (11.51) | 103.82 (11.62) | 368     | 93.39 (19.97)  | 96.25 (18.11)  | 98.76 (15.42)  | 349     |
|                           | <i>Frequency</i> | 0.70           | 0.29           | 0.01           |         | 0.64           | 0.33           | 0.03           |         |
| rs11794996<br>(135444784) | FIQ (SD)         | 103.51 (11.76) | 104.12 (8.96)  | 115.31 (18.52) | 369     | 99.15 (14.64)  | 97.71 (16.99)  | 100.81 (13.68) | 361     |
|                           | VIQ (SD)         | 103.93 (12.66) | 104.91 (10.51) | 111.52 (19.02) | 369     | 102.94 (12.83) | 101.60 (13.59) | 103.72 (9.37)  | 362     |
|                           | PIQ (SD)         | 103.89 (12.26) | 104.48 (9.71)  | 117.07 (14.86) | 369     | 94.94 (18.58)  | 94.34 (20.34)  | 97.50 (24.61)  | 361     |
|                           | <i>Frequency</i> | 0.35           | 0.55           | 0.11           |         | 0.37           | 0.47           | 0.15           |         |
| rs129900<br>(135541826)   | FIQ (SD)         | 103.41 (11.03) | 103.99 (10.86) | 103.46 (12.97) | 367     | 96.06 (14.91)  | 102.28 (15.50) | 96.26 (14.04)  | 363     |
|                           | VIQ (SD)         | 103.81 (12.63) | 104.83 (11.41) | 101.70 (13.59) | 367     | 100.07 (12.86) | 105.42 (12.69) | 101.51 (12.71) | 364     |
|                           | PIQ (SD)         | 103.86 (11.24) | 104.03 (11.75) | 106.77 (12.64) | 367     | 92.69 (18.96)  | 97.95 (19.13)  | 91.55 (19.18)  | 363     |
|                           | <i>Frequency</i> | 0.57           | 0.29           | 0.14           |         | 0.30           | 0.47           | 0.22           |         |
| rs1611125<br>(135499133)  | FIQ (SD)         | 102.41 (9.34)  | 104.91 (12.48) | 103.37 (9.52)  | 364     | 98.55 (15.71)  | 99.50 (14.94)  | 98.03 (15.59)  | 351     |
|                           | VIQ (SD)         | 103.25 (11.54) | 104.97 (13.01) | 104.35 (10.19) | 364     | 103.14 (12.62) | 102.86 (12.96) | 101.64 (13.32) | 352     |
|                           | PIQ (SD)         | 103.06 (9.76)  | 105.47 (12.73) | 103.18 (10.97) | 364     | 93.90 (18.79)  | 95.90 (18.93)  | 94.20 (20.57)  | 351     |
|                           | <i>Frequency</i> | 0.43           | 0.48           | 0.09           |         | 0.51           | 0.41           | 0.08           |         |
| rs2073837<br>(135512749)  | FIQ (SD)         | 103.12 (11.07) | 104.17 (11.51) | 105.44 (9.18)  | 369     | 98.67 (14.87)  | 99.52 (15.98)  | 96.68 (14.27)  | 345     |
|                           | VIQ (SD)         | 103.77 (12.46) | 104.26 (12.23) | 106.25 (9.37)  | 369     | 102.59 (12.76) | 102.83 (13.41) | 102.70 (12.48) | 346     |
|                           | PIQ (SD)         | 103.35 (11.13) | 105.08 (12.43) | 105.55 (9.73)  | 369     | 94.75 (19.07)  | 95.57 (18.59)  | 91.12 (22.24)  | 345     |
|                           | <i>Frequency</i> | 0.23           | 0.53           | 0.23           |         | 0.31           | 0.42           | 0.27           |         |
| rs2519148<br>(135484387)  | FIQ (SD)         | 101.67 (12.59) | 104.81 (10.22) | 103.95 (11.08) | 358     | 98.44 (14.56)  | 98.99 (15.67)  | 98.75 (16.38)  | 336     |
|                           | VIQ (SD)         | 102.14 (14.28) | 105.38 (10.41) | 104.01 (13.04) | 358     | 102.59 (13.66) | 102.96 (13.11) | 102.50 (12.26) | 337     |
|                           | PIQ (SD)         | 101.97 (12.57) | 105.16 (11.50) | 105.04 (10.48) | 358     | 93.92 (18.57)  | 94.99 (18.72)  | 94.81 (21.80)  | 336     |



**Table A1.1b** Mean IQ and Standard deviation (SD) for *DRD2* gene for young and adult cohorts

| <i>Tag-SNP</i>            |                  | Young Cohort       |                |                |         | Adult Cohort       |                |                |         |
|---------------------------|------------------|--------------------|----------------|----------------|---------|--------------------|----------------|----------------|---------|
| position (bp)             | Phenotype        | Genotype Frequency |                |                | Total N | Genotype Frequency |                |                | Total N |
|                           |                  | GG                 | GT             | TT             |         | GG                 | GT             | TT             |         |
|                           | <i>Frequency</i> | 0.68               | 0.25           | 0.07           |         | 0.71               | 0.27           | 0.03           |         |
| rs10891556<br>(112857971) | FIQ (SD)         | 103.64 (11.44)     | 103.92 (11.43) | 104.95 (9.87)  | 362     | 98.67 (15.18)      | 99.20 (15.69)  | 98.49 (15.19)  | 376     |
|                           | VIQ (SD)         | 104.11 (12.79)     | 104.01 (11.68) | 106.25 (9.12)  | 362     | 102.29 (13.03)     | 103.42 (12.93) | 103.76 (13.38) | 376     |
|                           | PIQ (SD)         | 104.12 (11.90)     | 104.78 (11.93) | 103.92 (10.86) | 362     | 94.75 (19.49)      | 95.13 (18.76)  | 92.49 (16.89)  | 376     |
|                           |                  | CC                 | CG             | GG             |         | CC                 | CG             | GG             |         |
|                           | <i>Frequency</i> | 0.43               | 0.45           | 0.12           |         | 0.47               | 0.46           | 0.07           |         |
| rs1984739<br>(112890119)  | FIQ (SD)         | 103.59 (11.64)     | 104.63 (10.67) | 102.05 (10.62) | 366     | 97.87 (15.86)      | 98.52 (14.77)  | 105.68 (14.41) | 361     |
|                           | VIQ (SD)         | 103.73 (12.79)     | 105.15 (11.27) | 103.08 (11.78) | 366     | 101.77 (13.13)     | 102.80 (13.11) | 106.26 (10.99) | 362     |
|                           | PIQ (SD)         | 104.37 (11.77)     | 104.91 (11.69) | 102.37 (10.86) | 366     | 94.05 (19.25)      | 94.08 (19.21)  | 103.60 (17.90) | 361     |
|                           |                  | CC                 | CT             | TT             |         | CC                 | CT             | TT             |         |
|                           | <i>Frequency</i> | 0.67               | 0.29           | 0.04           |         | 0.73               | 0.24           | 0.04           |         |
| rs2075654<br>(112794276)  | FIQ (SD)         | 103.64 (11.19)     | 105.28 (11.19) | 100.87 (10.23) | 363     | 98.63 (15.98)      | 98.89 (14.12)  | 96.84 (5.93)   | 358     |
|                           | VIQ (SD)         | 104.09 (12.60)     | 105.67 (11.06) | 102.74 (11.54) | 363     | 102.27 (13.33)     | 102.88 (12.49) | 107.06 (7.89)  | 359     |
|                           | PIQ (SD)         | 104.16 (11.50)     | 105.36 (12.16) | 100.28 (11.20) | 363     | 94.91 (20.14)      | 94.63 (17.22)  | 86.08 (7.82)   | 358     |
|                           |                  | AA                 | AG             | GG             |         | AA                 | AG             | GG             |         |
|                           | <i>Frequency</i> | 0.31               | 0.40           | 0.29           |         | 0.33               | 0.51           | 0.17           |         |
| rs6589377<br>(112860946)  | FIQ (SD)         | 102.87 (10.72)     | 103.86 (12.28) | 104.38 (10.61) | 344     | 98.96 (14.96)      | 98.93 (15.18)  | 95.87 (16.19)  | 349     |
|                           | VIQ (SD)         | 103.28 (11.48)     | 104.52 (13.04) | 104.77 (12.50) | 344     | 103.69 (12.96)     | 102.59 (12.97) | 99.78 (12.59)  | 349     |
|                           | PIQ (SD)         | 103.70 (11.24)     | 103.76 (12.85) | 104.96 (11.00) | 344     | 93.65 (17.82)      | 95.49 (19.38)  | 91.84 (21.94)  | 349     |

**Table A1.1c** Mean IQ and Standard deviation (SD) for *DRD3* gene for young and adult cohorts

| <i>tag-SNP</i>            |                  | Young Cohort       |                |                |     | Adult Cohort       |                |                |     |
|---------------------------|------------------|--------------------|----------------|----------------|-----|--------------------|----------------|----------------|-----|
| position (bp)             | Phenotype        | Genotype Frequency |                |                | N   | Genotype Frequency |                |                | N   |
|                           |                  | AA                 | AG             | GG             |     | AA                 | AG             | GG             |     |
|                           | <i>Frequency</i> | 0.33               | 0.46           | 0.21           |     | 0.32               | 0.43           | 0.25           |     |
| rs2087017<br>(115324703)  | FIQ (SD)         | 103.34 (11.17)     | 103.66 (11.59) | 105.12 (10.11) | 360 | 99.66 (14.68)      | 97.09 (15.80)  | 99.88 (14.89)  | 386 |
|                           | VIQ (SD)         | 104.19 (12.26)     | 103.94 (12.64) | 105.78 (10.61) | 360 | 103.81 (13.31)     | 101.21 (13.10) | 103.32 (12.57) | 386 |
|                           | PIQ (SD)         | 103.45 (11.83)     | 104.15 (12.28) | 105.33 (10.50) | 360 | 95.20 (18.91)      | 93.15 (19.53)  | 95.98 (18.88)  | 386 |
|                           |                  | AA                 | AG             | GG             |     | AA                 | AG             | GG             |     |
|                           | <i>Frequency</i> | 0.93               | 0.07           | 0.00           |     | 0.93               | 0.07           | 0.00           |     |
| rs35375733<br>(115299109) | FIQ (SD)         | 103.82 (11.19)     | 104.26 (9.43)  | -              | 366 | 98.79 (15.31)      | 98.97 (15.09)  | -              | 364 |
|                           | VIQ (SD)         | 104.21 (12.24)     | 105.18 (10.22) | -              | 366 | 102.74 (12.91)     | 101.48 (13.72) | -              | 365 |
|                           | PIQ (SD)         | 104.38 (11.64)     | 103.88 (10.74) | -              | 366 | 94.78 (19.08)      | 94.94 (19.68)  | -              | 364 |
|                           |                  | GG                 | AG             | AA             |     | GG                 | AG             | AA             |     |
|                           | <i>Frequency</i> | 0.78               | 0.21           | 0.01           |     | 0.80               | 0.16           | 0.04           |     |
| rs3773678<br>(10183426)   | FIQ (SD)         | 102.36 (11.72)     | 106.84 (10.20) | 110.41 (NA)    | 275 | 99.29 (15.52)      | 96.98 (14.76)  | 92.52 (13.44)  | 300 |
|                           | VIQ (SD)         | 102.80 (12.87)     | 107.83 (11.04) | 108.43 (NA)    | 275 | 102.39 (13.13)     | 104.33 (12.33) | 95.95 (10.89)  | 300 |
|                           | PIQ (SD)         | 102.96 (12.02)     | 106.39 (11.31) | 113.84 (NA)    | 275 | 95.96 (19.03)      | 89.72 (20.00)  | 90.26 (17.72)  | 300 |

**Table A1.1d** Mean IQ and Standard deviation (SD) for *HTR2A* gene for young and adult cohorts

| <i>tag</i> -SNP<br>position (bp) | Phenotype        | Young Cohort       |                |                |         | Adult Cohort       |                |                |         |
|----------------------------------|------------------|--------------------|----------------|----------------|---------|--------------------|----------------|----------------|---------|
|                                  |                  | Genotype Frequency |                |                | Total N | Genotype Frequency |                |                | Total N |
|                                  |                  | CC                 | CG             | GG             |         | CC                 | CG             | GG             |         |
|                                  | <i>Frequency</i> | 0.77               | 0.22           | 0.01           |         | 0.75               | 0.24           | 0.01           |         |
| rs1216970                        | FIQ (SD)         | 103.62 (10.65)     | 104.94 (12.13) | 93.53 (14.48)  | 370     | 97.82 (15.20)      | 101.81 (14.41) | 107.42 (24.89) | 365     |
| (46272013)                       | VIQ (SD)         | 104.15 (11.87)     | 105.34 (12.52) | 90.77 (15.16)  | 370     | 101.81 (12.89)     | 104.99 (12.33) | 109.56 (19.76) | 366     |
|                                  | PIQ (SD)         | 104.05 (11.27)     | 105.17 (12.65) | 99.76 (15.27)  | 370     | 93.94 (18.83)      | 97.67 (19.74)  | 102.11 (25.82) | 365     |
|                                  |                  | CC                 | CT             | TT             |         | CC                 | CT             | TT             |         |
|                                  | <i>Frequency</i> | 0.75               | 0.22           | 0.03           |         | 0.71               | 0.28           | 0.01           |         |
| rs1923884                        | FIQ (SD)         | 103.94 (10.77)     | 103.92 (11.83) | 101.73 (10.98) | 370     | 97.46 (15.36)      | 100.99 (13.95) | 105.25 (25.12) | 317     |
| (46319837)                       | VIQ (SD)         | 104.36 (11.83)     | 104.42 (12.43) | 102.30 (13.34) | 370     | 102.00 (12.93)     | 105.04 (11.92) | 108.23 (20.09) | 318     |
|                                  | PIQ (SD)         | 104.46 (11.30)     | 104.21 (12.49) | 102.52 (10.05) | 370     | 93.16 (19.24)      | 95.47 (18.30)  | 99.61 (25.80)  | 317     |
|                                  |                  | CC                 | AC             | AA             |         | CC                 | AC             | AA             |         |
|                                  | <i>Frequency</i> | 0.27               | 0.53           | 0.20           |         | 0.25               | 0.57           | 0.17           |         |
| rs9534512                        | FIQ (SD)         | 105.13 (10.87)     | 102.83 (11.12) | 103.96 (11.44) | 368     | 104.34 (14.67)     | 96.08 (14.35)  | 98.95 (16.55)  | 362     |
| (46371922)                       | VIQ (SD)         | 105.18 (11.44)     | 103.66 (12.34) | 103.83 (12.72) | 368     | 106.89 (13.31)     | 100.95 (12.17) | 101.81 (13.37) | 363     |
|                                  | PIQ (SD)         | 106.10 (11.25)     | 102.90 (11.99) | 105.02 (10.72) | 368     | 100.67 (18.72)     | 91.74 (18.36)  | 95.15 (19.96)  | 362     |

**Table A1.1e** Mean IQ and Standard deviation (SD) for *SERT* gene for young and adult cohorts

| <i>tag</i> -SNP<br>position (bp) | Phenotype        | Young Cohort       |                |                |     | Adult Cohort       |                |                |     |
|----------------------------------|------------------|--------------------|----------------|----------------|-----|--------------------|----------------|----------------|-----|
|                                  |                  | Genotype Frequency |                |                | N   | Genotype Frequency |                |                | N   |
|                                  |                  | AA                 | AG             | GG             |     | AA                 | AG             | GG             |     |
|                                  | <i>Frequency</i> | 0.65               | 0.29           | 0.06           |     | 0.69               | 0.28           | 0.03           |     |
| rs4436830                        | FIQ (SD)         | 104.04 (11.35)     | 104.11 (11.72) | 102.96 (7.70)  | 363 | 98.29 (14.91)      | 99.95 (15.25)  | 98.75 (19.22)  | 366 |
| (25407835)                       | VIQ (SD)         | 104.59 (12.29)     | 103.90 (12.14) | 104.47 (8.08)  | 363 | 102.31 (12.96)     | 103.22 (12.60) | 102.96 (15.39) | 366 |
|                                  | PIQ (SD)         | 104.21 (11.48)     | 105.58 (12.79) | 102.58 (9.94)  | 363 | 94.17 (18.85)      | 95.96 (19.65)  | 97.86 (17.77)  | 366 |
|                                  |                  | TT                 | CT             | CC             |     | TT                 | CT             | CC             |     |
|                                  | <i>Frequency</i> | 0.68               | 0.30           | 0.01           |     | 0.73               | 0.25           | 0.02           |     |
| rs4494608                        | FIQ (SD)         | 103.61 (10.89)     | 104.20 (11.40) | 100.47 (12.67) | 362 | 98.86 (15.31)      | 98.98 (15.45)  | 101.45 (8.40)  | 389 |
| (25512917)                       | VIQ (SD)         | 104.30 (11.94)     | 103.96 (12.50) | 101.32 (12.14) | 362 | 103.06 (12.95)     | 101.88 (13.26) | 102.21 (7.40)  | 389 |
|                                  | PIQ (SD)         | 103.90 (11.80)     | 105.39 (11.16) | 99.96 (10.59)  | 362 | 94.43 (19.03)      | 95.98 (19.76)  | 100.07 (9.78)  | 389 |
|                                  |                  | AA                 | AT             | TT             |     | AA                 | AT             | TT             |     |
|                                  | <i>Frequency</i> | 0.28               | 0.48           | 0.24           |     | 0.29               | 0.51           | 0.21           |     |
| rs8067576                        | FIQ (SD)         | 103.68 (10.90)     | 103.54 (10.35) | 104.74 (12.64) | 374 | 98.82 (14.78)      | 98.20 (15.31)  | 100.22 (15.89) | 365 |
| (25468380)                       | VIQ (SD)         | 104.12 (12.05)     | 104.21 (11.92) | 104.41 (12.57) | 374 | 102.92 (12.75)     | 101.79 (12.96) | 104.01 (13.16) | 366 |
|                                  | PIQ (SD)         | 103.88 (10.69)     | 103.99 (10.78) | 105.98 (13.72) | 374 | 94.51 (18.66)      | 94.43 (19.74)  | 96.28 (18.75)  | 365 |

**Table A1.1f** Mean IQ and Standard deviation (SD) for *TH* gene for young and adult cohorts

| <i>tag</i> -SNP<br>position (bp) | Phenotype        | Young Cohort       |                |                |         | Adult Cohort       |                |                |         |
|----------------------------------|------------------|--------------------|----------------|----------------|---------|--------------------|----------------|----------------|---------|
|                                  |                  | Genotype Frequency |                |                | Total N | Genotype Frequency |                |                | Total N |
|                                  |                  | AA                 | AG             | GG             |         | AA                 | AG             | GG             |         |
|                                  | <i>Frequency</i> | 0.27               | 0.51           | 0.22           |         | 0.36               | 0.47           | 0.17           |         |
| rs2070762<br>(2142911)           | FIQ (SD)         | 102.74 (10.81)     | 104.88 (11.71) | 101.87 (9.86)  | 351     | 98.17 (13.89)      | 98.63 (16.07)  | 99.42 (15.04)  | 378     |
|                                  | VIQ (SD)         | 102.84 (12.45)     | 105.30 (12.57) | 102.93 (10.95) | 351     | 103.38 (12.53)     | 102.70 (12.72) | 102.07 (13.05) | 379     |
|                                  | PIQ (SD)         | 103.83 (11.14)     | 105.24 (12.35) | 101.93 (10.46) | 351     | 92.74 (17.12)      | 94.38 (19.73)  | 96.58 (20.43)  | 378     |
|                                  | <i>Frequency</i> | 0.33               | 0.53           | 0.14           |         | 0.38               | 0.48           | 0.15           |         |
| rs4074905<br>(2145761)           | FIQ (SD)         | 103.36 (10.09)     | 103.81 (11.73) | 101.92 (11.63) | 360     | 100.94 (15.05)     | 97.49 (15.30)  | 97.91 (15.82)  | 365     |
|                                  | VIQ (SD)         | 104.02 (11.15)     | 104.40 (12.52) | 101.79 (13.65) | 360     | 104.48 (12.48)     | 101.44 (13.18) | 102.23 (13.45) | 364     |
|                                  | PIQ (SD)         | 103.74 (10.60)     | 104.11 (12.39) | 103.19 (11.73) | 360     | 96.84 (19.94)      | 93.65 (18.60)  | 93.42 (19.58)  | 365     |

**Table A1.2a** QTDT results for *DBH* gene for young and adult cohort

| <i>Tag</i> -SNP<br>location (bp) | Phenotype | Young Cohort     |          |         |                  |          |       | Adult Cohort     |          |         |                  |          |       |
|----------------------------------|-----------|------------------|----------|---------|------------------|----------|-------|------------------|----------|---------|------------------|----------|-------|
|                                  |           | Family-based     |          |         | Population-based |          |       | Family-based     |          |         | Population-based |          |       |
|                                  |           | N <sub>Fam</sub> | $\chi^2$ | P-val   | N <sub>Pop</sub> | $\chi^2$ | P-val | N <sub>Fam</sub> | $\chi^2$ | P-val   | N <sub>Pop</sub> | $\chi^2$ | P-val |
| rs1029372<br>(135468263)         | FIQ (SD)  | 124              | 0.01     | 0.920   | 367              | 1.08     | 0.299 | 142              | 2.96     | 0.085   | 361              | 0.54     | 0.462 |
|                                  | VIQ (SD)  | 124              | 0.04     | 0.841   | 367              | 1.36     | 0.244 | 142              | 3.32     | 0.068   | 362              | 0.14     | 0.708 |
|                                  | PIQ (SD)  | 124              | 0.11     | 0.740   | 367              | 0.52     | 0.471 | 142              | 1.32     | 0.251   | 361              | 1.12     | 0.290 |
| rs1108581<br>(135495062)         | FIQ (SD)  | 57               | 0.20     | 0.655   | 368              | 2.66     | 0.103 | 114              | 4.11     | 0.043 * | 349              | 0.00     | 1.000 |
|                                  | VIQ (SD)  | 57               | 0.34     | 0.560   | 368              | 3.44     | 0.064 | 114              | 1.80     | 0.180   | 350              | 0.22     | 0.639 |
|                                  | PIQ (SD)  | 57               | 0.14     | 0.708 * | 368              | 1.02     | 0.313 | 114              | 5.16     | 0.023 * | 349              | 0.18     | 0.671 |
| rs11794996<br>(135444784)        | FIQ (SD)  | 72               | 0.07     | 0.791   | 369              | 0.29     | 0.590 | 94               | 0.05     | 0.823   | 361              | 0.40     | 0.527 |
|                                  | VIQ (SD)  | 72               | 0.08     | 0.777   | 369              | 0.41     | 0.522 | 94               | 0.27     | 0.603   | 362              | 0.26     | 0.610 |
|                                  | PIQ (SD)  | 72               | 0.00     | 1.000   | 369              | 0.03     | 0.862 | 94               | 0.06     | 0.806   | 361              | 0.62     | 0.431 |
| rs129900<br>(135541826)          | FIQ (SD)  | 96               | 1.40     | 0.237   | 367              | 0.06     | 0.806 | 119              | 0.29     | 0.590   | 363              | 0.02     | 0.888 |
|                                  | VIQ (SD)  | 96               | 1.25     | 0.264   | 367              | 0.29     | 0.590 | 119              | 0.24     | 0.624   | 364              | 0.11     | 0.740 |
|                                  | PIQ (SD)  | 96               | 0.37     | 0.543   | 367              | 0.07     | 0.791 | 119              | 0.18     | 0.671   | 363              | 0.56     | 0.454 |
| rs1611125<br>(135499133)         | FIQ (SD)  | 106              | 0.86     | 0.354   | 364              | 0.12     | 0.729 | 135              | 4.97     | 0.026   | 351              | 0.42     | 0.517 |
|                                  | VIQ (SD)  | 106              | 0.09     | 0.764   | 364              | 0.40     | 0.527 | 135              | 3.32     | 0.068   | 352              | 0.16     | 0.689 |
|                                  | PIQ (SD)  | 106              | 2.92     | 0.087   | 364              | 0.91     | 0.340 | 135              | 5.36     | 0.021   | 351              | 0.96     | 0.327 |
| rs2073837<br>(135512749)         | FIQ (SD)  | 92               | 0.11     | 0.740   | 369              | 0.01     | 0.920 | 99               | 0.60     | 0.439   | 345              | 0.56     | 0.454 |
|                                  | VIQ (SD)  | 92               | 0.20     | 0.655   | 369              | 0.08     | 0.777 | 99               | 0.20     | 0.655   | 346              | 0.57     | 0.450 |
|                                  | PIQ (SD)  | 92               | 0.01     | 0.920   | 369              | 0.01     | 0.920 | 99               | 0.54     | 0.462   | 346              | 0.56     | 0.454 |
| rs2519148<br>(135484387)         | FIQ (SD)  | 91               | 0.10     | 0.752   | 358              | 0.00     | 1.000 | 123              | 0.71     | 0.399   | 336              | 2.98     | 0.084 |
|                                  | VIQ (SD)  | 91               | 0.68     | 0.410   | 358              | 0.09     | 0.764 | 123              | 0.24     | 0.624   | 337              | 1.99     | 0.158 |
|                                  | PIQ (SD)  | 91               | 0.02     | 0.888   | 358              | 0.09     | 0.764 | 123              | 1.51     | 0.219   | 336              | 3.59     | 0.058 |

**Table A1.2b** QTDT results for *DRD2* gene for young and adult cohorts

|                           |           | Young Cohort     |          |       |                  |          |       | Adult Cohort     |          |       |                  |          |       |
|---------------------------|-----------|------------------|----------|-------|------------------|----------|-------|------------------|----------|-------|------------------|----------|-------|
| Tag-SNP<br>location (bp)  | Phenotype | Family-based     |          |       | Population-based |          |       | Family-based     |          |       | Population-based |          |       |
|                           |           | N <sub>Fam</sub> | $\chi^2$ | P-val | N <sub>Pop</sub> | $\chi^2$ | P-val | N <sub>Fam</sub> | $\chi^2$ | P-val | N <sub>Pop</sub> | $\chi^2$ | P-val |
| rs10891556<br>(112857971) | FIQ (SD)  | 88               | 0.88     | 0.348 | 362              | 0.06     | 0.806 | 94               | 0.49     | 0.484 | 376              | 0.40     | 0.527 |
|                           | VIQ (SD)  | 88               | 0.09     | 0.764 | 362              | 0.13     | 0.718 | 94               | 0.11     | 0.740 | 376              | 0.30     | 0.584 |
|                           | PIQ (SD)  | 88               | 1.29     | 0.256 | 362              | 0.26     | 0.610 | 94               | 0.57     | 0.450 | 376              | 0.25     | 0.617 |
| rs1984739<br>(112890119)  | FIQ (SD)  | 119              | 0.01     | 0.920 | 363              | 1.11     | 0.292 | 137              | 0.11     | 0.740 | 361              | 0.15     | 0.699 |
|                           | VIQ (SD)  | 119              | 0.03     | 0.862 | 363              | 0.97     | 0.325 | 137              | 0.41     | 0.522 | 362              | 0.01     | 0.920 |
|                           | PIQ (SD)  | 119              | 0.06     | 0.806 | 363              | 0.87     | 0.351 | 137              | 0.03     | 0.862 | 361              | 0.71     | 0.399 |
| rs2075654<br>(112794276)  | FIQ (SD)  | 73               | 0.34     | 0.560 | 363              | 0.21     | 0.647 | 95               | 0.06     | 0.806 | 358              | 0.00     | 1.000 |
|                           | VIQ (SD)  | 73               | 0.43     | 0.512 | 363              | 0.45     | 0.502 | 95               | 0.01     | 0.920 | 359              | 0.15     | 0.699 |
|                           | PIQ (SD)  | 73               | 1.65     | 0.199 | 363              | 1.23     | 0.267 | 95               | 0.19     | 0.663 | 358              | 0.22     | 0.639 |
| rs6589377<br>(112860946)  | FIQ (SD)  | 103              | 1.15     | 0.284 | 344              | 0.11     | 0.740 | 102              | 0.00     | 1.000 | 349              | 1.02     | 0.313 |
|                           | VIQ (SD)  | 103              | 0.01     | 0.920 | 344              | 1.44     | 0.230 | 102              | 0.07     | 0.791 | 349              | 0.70     | 0.403 |
|                           | PIQ (SD)  | 103              | 2.41     | 0.121 | 344              | 0.17     | 0.680 | 102              | 0.64     | 0.424 | 349              | 1.09     | 0.296 |

**Table A1.2c** QTDT results for *DRD3* gene for young and adult cohorts

|                           |           | Young Cohort     |          |       |                  |          |       | Adult Cohort     |          |       |                  |          |       |
|---------------------------|-----------|------------------|----------|-------|------------------|----------|-------|------------------|----------|-------|------------------|----------|-------|
| Tag-SNP<br>location (bp)  | Phenotype | Family-based     |          |       | Population-based |          |       | Family-based     |          |       | Population-based |          |       |
|                           |           | N <sub>Fam</sub> | $\chi^2$ | P-val | N <sub>Pop</sub> | $\chi^2$ | P-val | N <sub>Fam</sub> | $\chi^2$ | P-val | N <sub>Pop</sub> | $\chi^2$ | P-val |
| rs2087017<br>(115324703)  | FIQ (SD)  | 95               | 1.64     | 0.200 | 360              | 0.16     | 0.689 | 142              | 2.00     | 0.157 | 386              | 1.27     | 0.260 |
|                           | VIQ (SD)  | 95               | 0.06     | 0.806 | 360              | 0.11     | 0.740 | 142              | 1.42     | 0.233 | 386              | 0.86     | 0.354 |
|                           | PIQ (SD)  | 95               | 2.35     | 0.125 | 360              | 0.35     | 0.554 | 142              | 1.28     | 0.258 | 386              | 1.02     | 0.313 |
| rs35375733<br>(115299109) | FIQ (SD)  | 20               | NA       | NA    | 366              | 0.33     | 0.566 | 40               | 0.01     | 0.920 | 364              | 0.04     | 0.841 |
|                           | VIQ (SD)  | 20               | NA       | NA    | 366              | 1.40     | 0.237 | 40               | 0.05     | 0.823 | 365              | 0.09     | 0.764 |
|                           | PIQ (SD)  | 20               | NA       | NA    | 366              | 0.01     | 0.920 | 40               | 0.09     | 0.764 | 364              | 0.00     | 1.000 |
| rs3773678<br>(10183426)   | FIQ (SD)  | 43               | 3.17     | 0.075 | 349              | 3.23     | 0.072 | 40               | 0.09     | 0.764 | 300              | 4.80     | 0.028 |
|                           | VIQ (SD)  | 43               | 0.07     | 0.791 | 349              | 0.00     | 1.000 | 40               | 0.02     | 0.888 | 300              | 4.06     | 0.044 |
|                           | PIQ (SD)  | 43               | 5.50     | 0.019 | 349              | 4.60     | 0.032 | 40               | 0.16     | 0.689 | 300              | 4.27     | 0.039 |

**Table A1.2d** QTDT results for *HTR2A* gene for young and adult cohorts

|                          |           | Young Cohort |          |       |                  |          |       | Adult Cohort |          |       |                  |          |       |
|--------------------------|-----------|--------------|----------|-------|------------------|----------|-------|--------------|----------|-------|------------------|----------|-------|
| Tag-SNP<br>location (bp) | Phenotype | Family-based |          |       | Population-based |          |       | Family-based |          |       | Population-based |          |       |
|                          |           | N            | $\chi^2$ | P-val | N                | $\chi^2$ | P-val | N            | $\chi^2$ | P-val | N                | $\chi^2$ | P-val |
| rs1216970<br>(46272013)  | FIQ (SD)  | 27           | NA       | NA    | 370              | 3.34     | 0.068 | 87           | 0.29     | 0.590 | 365              | 0.00     | 1.000 |
|                          | VIQ (SD)  | 27           | NA       | NA    | 370              | 2.49     | 0.115 | 87           | 0.76     | 0.383 | 366              | 0.04     | 0.841 |
|                          | PIQ (SD)  | 27           | NA       | NA    | 370              | 2.19     | 0.139 | 87           | 0.03     | 0.862 | 365              | 0.02     | 0.888 |
| rs1923884<br>(46319837)  | FIQ (SD)  | 37           | 0.36     | 0.549 | 370              | 1.20     | 0.273 | 103          | 0.52     | 0.471 | 317              | 0.00     | 1.000 |
|                          | VIQ (SD)  | 37           | 0.16     | 0.689 | 370              | 1.26     | 0.262 | 103          | 1.12     | 0.290 | 318              | 0.00     | 1.000 |
|                          | PIQ (SD)  | 37           | 0.23     | 0.632 | 370              | 0.54     | 0.462 | 103          | 0.00     | 1.000 | 317              | 0.05     | 0.823 |
| rs9534512<br>(46371922)  | FIQ (SD)  | 111          | 0.86     | 0.354 | 368              | 5.08     | 0.024 | 124          | 2.44     | 0.118 | 362              | 0.76     | 0.383 |
|                          | VIQ (SD)  | 111          | 1.91     | 0.167 | 368              | 5.80     | 0.016 | 124          | 1.22     | 0.269 | 363              | 0.37     | 0.543 |
|                          | PIQ (SD)  | 111          | 0.10     | 0.752 | 368              | 2.50     | 0.114 | 124          | 3.78     | 0.052 | 362              | 1.17     | 0.279 |

*Note:* The standard version of QTDT only attempts to estimate allele effects when at least 30 informative individuals are available.

**Table A1.2e** QTDT results for *SERT* gene for young and adult cohorts

| <i>Tag</i> -SNP<br>location (bp) | Phenotype | Young Cohort |          |       |                  |          |       | Adult Cohort |          |       |                  |          |       |
|----------------------------------|-----------|--------------|----------|-------|------------------|----------|-------|--------------|----------|-------|------------------|----------|-------|
|                                  |           | Family-based |          |       | Population-based |          |       | Family-based |          |       | Population-based |          |       |
|                                  |           | N            | $\chi^2$ | P-val | N                | $\chi^2$ | P-val | N            | $\chi^2$ | P-val | N                | $\chi^2$ | P-val |
| rs4436830<br>(25407835)          | FIQ (SD)  | 84           | 0.97     | 0.325 | 363              | 0.88     | 0.348 | 92           | 1.59     | 0.207 | 366              | 0.37     | 0.543 |
|                                  | VIQ (SD)  | 84           | 0.24     | 0.624 | 363              | 0.35     | 0.554 | 92           | 0.96     | 0.327 | 366              | 0.12     | 0.729 |
|                                  | PIQ (SD)  | 84           | 1.15     | 0.284 | 363              | 1.07     | 0.301 | 92           | 1.21     | 0.271 | 366              | 0.17     | 0.680 |
| rs4494608<br>(25512917)          | FIQ (SD)  | 65           | 2.21     | 0.137 | 362              | 0.60     | 0.439 | 81           | 0.00     | 1.000 | 389              | 0.00     | 1.000 |
|                                  | VIQ (SD)  | 65           | 0.15     | 0.699 | 362              | 0.09     | 0.764 | 81           | 1.28     | 0.258 | 389              | 0.61     | 0.435 |
|                                  | PIQ (SD)  | 65           | 3.13     | 0.077 | 362              | 1.40     | 0.237 | 81           | 1.53     | 0.216 | 389              | 0.59     | 0.442 |
| rs8067576<br>(25468380)          | FIQ (SD)  | 126          | 0.03     | 0.862 | 374              | 0.00     | 1.000 | 136          | 0.32     | 0.572 | 365              | 0.19     | 0.663 |
|                                  | VIQ (SD)  | 126          | 0.06     | 0.806 | 374              | 0.02     | 0.888 | 136          | 1.58     | 0.209 | 366              | 0.00     | 1.000 |
|                                  | PIQ (SD)  | 126          | 0.01     | 0.920 | 374              | 0.01     | 0.920 | 136          | 2.69     | 0.101 | 365              | 1.18     | 0.277 |

**Table A1.2e** QTDT results for *TH* gene for young and adult cohorts

| <i>Tag</i> -SNP<br>location (bp) | Phenotype | Young Cohort |          |       |                  |          |       | Adult Cohort |          |       |                  |          |       |
|----------------------------------|-----------|--------------|----------|-------|------------------|----------|-------|--------------|----------|-------|------------------|----------|-------|
|                                  |           | Family-based |          |       | Population-based |          |       | Family-based |          |       | Population-based |          |       |
|                                  |           | N            | $\chi^2$ | P-val | N                | $\chi^2$ | P-val | N            | $\chi^2$ | P-val | N                | $\chi^2$ | P-val |
| rs2070762<br>(2142911)           | FIQ (SD)  | 100          | 0.00     | 1.000 | 351              | 0.11     | 0.740 | 150          | 0.05     | 0.823 | 378              | 0.16     | 0.689 |
|                                  | VIQ (SD)  | 100          | 0.33     | 0.566 | 351              | 0.38     | 0.538 | 150          | 0.02     | 0.888 | 378              | 0.02     | 0.888 |
|                                  | PIQ (SD)  | 100          | 0.11     | 0.740 | 351              | 0.94     | 0.332 | 150          | 0.00     | 1.000 | 378              | 0.23     | 0.632 |
| rs4074905<br>(2145761)           | FIQ (SD)  | 107          | 0.01     | 0.920 | 360              | 1.66     | 0.198 | 121          | 0.20     | 0.655 | 365              | 0.35     | 0.554 |
|                                  | VIQ (SD)  | 107          | 0.45     | 0.502 | 360              | 1.05     | 0.306 | 121          | 1.50     | 0.221 | 365              | 0.71     | 0.399 |
|                                  | PIQ (SD)  | 107          | 0.25     | 0.617 | 360              | 1.66     | 0.198 | 121          | 0.34     | 0.560 | 365              | 0.10     | 0.752 |

## LABWORK TECHNIQUES

### *DNA collection and isolation*

Genomic DNA in children was obtained from buccal swabs, while DNA in adults was collected from blood samples. The DNA isolation from buccal swabs was performed using a chloroform/isopropanol extraction (Meulenbelt *et al.* 1995). DNA isolation from buccal swabs is a relatively easy lab procedure with the advantage of being a non-invasive technique from which high-yield of high-quality DNA can be obtained (Meulenbelt *et al.* 1995; Min *et al.* 2006). The mouth swab sample should be taken in a mouth without food remains. After rubbing, the cotton swab should be placed in a Falcon tube, containing 0.5 ml of STE buffer (100mM NaCl, 10 mM TrisHCL [pH 8.0] and 10 mM EDTA) with proteinase K (0.2 mg/ml) and SDS (0.5%) per cotton swab. First, a lysis step is performed in which Falcon tubes containing the cotton swabs and STE buffer with proteinase K and SDS, are placed in a 65 °C water bath for 2 hours. After incubation at 65 °C, in order to collect a maximum amount of buffer from the soaked cotton swabs, sticks are placed in a syringe, which in turn is placed up-side down in the Falcon tube. Falcon tubes are then centrifuged for 5 min at 1,000 rpm.

Genomic DNA is subsequently isolated from the collected lysis buffer using phenol:chloroform:isoamylalcohol (24:24:1) and chloroform:isoamylalcohol (24:1) extractions followed by isopropanol precipitation as described by Sambrook *et al.* (Sambrook *et al.* 1990). DNA pellets were finally reconstituted in 400 ul of TE (10 mM TrisHCL [pH 8.0] and 0.1 mM EDTA) and treated with RNase to remove low-molecular-weight RNA and smears.

Genomic DNA was extracted from blood samples using the salting out protocol described by Miller *et al.* (Miller *et al.* 1988). Buffy coats of nucleated cells from anticoagulated blood were suspended in a Falcon tube with 3 volumes of nuclei lysis buffer (10 mM TrisHCL, 400 mM NaCl and 2mM EDTA, pH 8.2). Cell lysates were incubate overnight at 37°C in the presence of 0.2 ml of 10% SDS and 0.5 ml of proteinase K solution (1mg of proteinase K in 1%SDS and 2mM EDTA). After overnight digestion, protein precipitation step is performed in which 1 ml of saturated NaCl solution (approximately 6M) was added into each tube, and vigorously homogenized for 15 seconds, followed by centrifugation at 2,500 rpm for 15 minutes. The supernatant containing the genomic DNA after the protein precipitation step was transferred into a new Falcon tube containing exactly 2 volumes absolute ethanol. Falcon tubes were inverted several times until a precipitated DNA “cloud” was observed. Falcon tubes were centrifuged at 2,500 rpm for about 15 minutes and supernatant was discarded. DNA pellet was then

washed twice with 70% ethanol and the DNA washed pellet was finally dissolved in 100-200  $\mu$ l TE (10 mM TrisHCL [pH 8.0] and 0.1 mM EDTA).

### *DNA and RNA extraction from tissue homogenates*

Control brains from 50 individuals, 23 males with a mean age of 70.3 years (SD= 9.38), and 27 females with a mean age of 73.3 years (SD=10.50) were obtained at autopsy from The Netherlands Brain Bank (NBB) (<http://www.brainbank.nl>). This material comes mainly from the superior and inferior parietal lobe. All tissue donations were approved by an Ethics Committee. DNA isolation from 0.20 gram of frozen tissue was performed using the Puregene™ Kit (Gentra Systems, USA) according to protocol with doubled volume of all reagents per tissue weight. To verify DNA isolation, products were run on a 1% agarose gel. Total RNA was isolated from 0.10 gram of frozen brain tissue with RNA-Bee™ (Isotex Diagnostics, Inc., USA). RNA was purified using the Qiagen RNeasy Mini kit (Qiagen Benelux B.V., The Netherlands) according to protocol and verified on a 2% agarose gel.

Five  $\mu$ g RNA was used to make cDNA using 200 U of Superscript™ III Reverse Transcriptase (Invitrogen, The Netherlands) in First Strand Buffer (Invitrogen, The Netherlands),  $3.4 \times 10^{-2}$   $\mu$ g/ $\mu$ l random hexamer oligo,  $3.4 \times 10^{-2}$   $\mu$ g/ $\mu$ l poly d(T) 12-18, 1.3 mM dNTPs, 1.1  $\mu$ M DTT (Invitrogen, The Netherlands), 10 U RNaseOUT™ Ribonuclease Inhibitor Recombinant (Invitrogen, The Netherlands) and incubated two hours at 50°C. Subsequently, 20 U RNase H (Invitrogen, The Netherlands) was added and incubated 30 minutes at 37°C. Products were run on a 1% agarose gel to examine the quality.

*SNPlex™ assay* was conducted following the manufacturer's recommendations (Applied Biosystems, Foster city, CA, USA). Stock genomic DNA (gDNA) solutions (50ng/ $\mu$ l) were diluted after fragmentation to a final concentration of 18.5ng/ $\mu$ l. Diluted gDNA aliquots (2 $\mu$ l) were spotted and dried down in 384 well plates (Bioplastics, Landgraaf The Netherlands). Previous to the oligo ligation assay (OLA) reaction, reagents were phosphorylated and diluted (1:1).

The OLA reaction was performed in a total volume of 5  $\mu$ l, which contained 37 ng of gDNA, 0.5 $\mu$ l of ligation buffer, 0.025  $\mu$ l of 48-SNPlex ligase and 1  $\mu$ l of activated ligation probe pool. The PCR conditions were 3 minutes at 90 °C, 30 cycles of 15 seconds at 90 °C, 30 seconds at 60 °C and 30 seconds at 51 °C (2% ramp), followed by denaturation at 99 °C for 10 minutes. After this step, a purification step was conducted after which the OLA products were ready to be amplified in a final volume of 10  $\mu$ l. Exonuclease I 0.1  $\mu$ l and lambda exonuclease

0.2 µl (Applied Biosystems, Foster city, CA, USA) were added, and incubated at 37°C for 90 minutes followed by a deactivation step at 80°C for 10 minutes. The purified OLA products were diluted (2:3) for further amplification. Amplification of OLA products was performed in 10µl, which contained 2µl of diluted OLA reaction, [1x] SNplex amplification master mix and [20x] SNplex amplification primers. The OLA amplification conditions were: 95 °C for 10 minutes, followed by 95 °C for 15 seconds, 63 °C for 60 seconds for 30 cycles.

After the hybridization step, analysis of the fluorescence intensity was performed in an aliquot (7.5ul) using ABI Sequencer 3730 (Applied Biosystems, Foster city, CA, USA). All pre-PCR steps were performed on a cooled block. Reactions were carried out in Gene Amp 9700 Thermocycler (Applied Biosystems, Foster city, CA, USA). Data was analyzed using Genemapper v3.7 (Applied Biosystems, Foster city, CA, USA).

*Taqman*® genotyping was performed using specific fluorogenic probes in the high-throughput 5' nuclease assay (TaqMan, PE Applied Biosystems, Foster city, CA), combining polymerase chain reaction amplification and detection into a single step. The assay requires two allele-specific probes, which labelled two alleles with different fluorescent reporter dyes for discrimination. Following allele-specific hybridization, the detection probe is cleaved during each amplification cycle by the 5'-exonuclease activity of Taq DNA polymerase if the probe's target sequence is present.

### *Quantitative Real Time-PCR (qRT-PCR)*

To determine whether SNP polymorphisms that showed the strongest association with cognitive ability, show differential brain expression levels, RT-PCR was performed on cDNA obtained from 50 brain samples from the NBB (<http://www.brainbank.nl/>). In each brain sample, a selected candidate gene expression levels was compared to two housekeeping genes: *β-Actin* and *Hypoxanthine-guanine phosphoribosyltransferase* gene (HPRT).

Amplification characteristics were then compared for all these primer sets. qRT-PCR was performed on TaqMan® (Applied Biosystems, USA) using 0.5 µl cDNA, made as described above, in a volume of 10 µl using 0.5 µl 2xPowerSYBR® Green PCR Master Mix (Applied Biosystems) with 0.3 pM forward and reverse specific primers. Thermal cycling conditions were 40 cycles at 95°C for 15 s and at 60°C for 1 min. Amplification was followed by a hold at 50°C for 2 min and at 95°C for 10 min. Optimization of primer concentration and cDNA input was performed and dissociation curves for the selected primers obtained. qRT-PCR reactions were performed for each sample in duplicate, twice.



## List of Publications



## LIST OF PUBLICATIONS

- Gosso, M.F.**, de Geus, E.J., van Belzen, M.J., Polderman, T.J., Heutink, P., Boomsma, D.I. & Posthuma, D. (2006a) The SNAP-25 gene is associated with cognitive ability: evidence from a family-based study in two independent Dutch cohorts. *Mol Psychiatry* 11 878-886.
- Gosso, M.F.**, van Belzen, M.J., de Geus, E.J.C., Polderman, J.C., Heutink, P., Boomsma, D.I. & Posthuma, D. (2006b) Association between the CHRM2 gene and intelligence in a sample of 304 Dutch families. *Genes, Brain and Behavior* 5 557-584.
- Mekel-Bobrov, N., Posthuma, D., Gilbert, S.L., Lind, P., **Gosso, M.F.**, Luciano, M., Harris, S.E., Bates, T.C., Polderman, T.J.C., Whalley, L.J., Fox, H., Starr, J.M., Evans, P.D., Montgomery, G.W., Fernandes, C., Heutink, P., Martin, N.G., Boomsma, D.I., Deary, I.J., *et al* (2007) The ongoing adaptive evolution of ASPM and Microcephalin is not explained by increased intelligence. *Hum Mol Genet* 16 600-608.
- Polderman, T.J.C., Stins, J.F., Posthuma, D., **Gosso, M.F.**, Verhulst, F.C. & Boomsma, D.I. (2006b) The phenotypic and genotypic relation between working memory speed and capacity. *Intelligence* 34 549-560.
- Polderman, T.J.C., **Gosso, M.F.**, Posthuma, D., Van Beijsterveldt, T.C., Heutink, P., Verhulst, F.C. & Boomsma, D.I. (2006a) A longitudinal twin study on IQ, executive functioning, and attention problems during childhood and early adolescence. *Acta Neurol Belg* 106 191-207.
- Gosso, M.F.**, de Geus, E.J.C., Polderman, T.J.C., Boomsma, D.I., Heutink, P. & Posthuma, D. (2007) Common variants underlying cognitive ability: further evidence for association between the SNAP-25 gene and cognition using a family-based study in two independent Dutch cohorts. *Genes, Brain and Behavior (in press)*.
- Gosso, M.F.**, de Geus, E.J.C., Polderman, T.J.C., Boomsma, D.I., Posthuma, D. & Heutink, P. (2007) Exploring the functional role of the CHRM2 gene in human cognition: results from a dense genotyping and brain expression study. *BMC Med Genet* 8:66.
- Gosso, M.F.**, de Geus, E.J.C., Polderman, T.J.C., Boomsma, D.I., Heutink, P. & Posthuma, D. Catechol O-methyl transferase and dopamine D2 receptor gene polymorphisms: evidence of positive heterosis and gene-gene interaction on working memory functioning. *EJHG, under revision*.
- Bochdanovits Z., **Gosso M.F.**, van den Berg L., Rizzu P., Polderman T. J.C., Pardo L.M., de Geus J.C.N., Boomsma D.I., Heutink P., and Posthuma D. A functional non-synonymous polymorphism under positive evolutionary

selection in the beta-2 adrenergic receptor explains phenotypic differences in intelligence among humans. *Submitted*

## PUBLISHED ABSTRACTS

- Gosso, M.F.**, Posthuma, D., De Geus, E.J.C., Van Belzen, M., Sondervan, D., Heutink, P., Boomsma, D.I. (2004) Mapping genes for memory, attention and cognitive abilities. 3<sup>rd</sup> *ENP Meeting*, 247.
- Gosso, M.F.**, Posthuma, D., Polderman TJC, De Geus, E.J.C., Van Belzen, M., Heutink, P., Boomsma, D.I. (2005) A family-based candidate gene approach to human attention and intelligence". 4<sup>th</sup> *ENP Meeting*, 155.
- Gosso, M.F.**, Van Belzen, M, De Geus, E.J.C., T.J.C. Polderman, , Heutink, P., Boomsma, D.I., Posthuma, D., (2006) A family-based association approach provides further evidence for a role of the *CHRM2* gene in cognition". 5<sup>th</sup> *ENP Meeting*, 259.
- Gosso, M.F.** (2005). SNP Genotyping System: 48-SNPlex Assay. 2ND SNP-NL Workshop
- Gosso, M.F.**, Van Belzen, M, De Geus, E.J.C., Polderman, T.J.C., Heutink, P., Boomsma, D.I. Posthuma, D. (2006). *CHRM2* gene and cognitive ability. 15<sup>th</sup> *ZonMw Genetica Retraite*.
- Gosso, M.F.**, Van Belzen, M, De Geus, E.J.C., Polderman, T.J.C., Heutink, P., Boomsma, D.I. Posthuma, D. (2006). Family-based association approach provides further evidence for a role of the *CHRM2* gene in cognition. *Eur J Hum Genet*, 14, 292.
- Boomsma DI, Polderrman TJC, **Gosso MF**, Verhulst FC, and Posthuma D (2006). Genes and Behavior: a twin study of executive functioning, IQ and attention problems in childhood and early adolescence. *Acta Neurol Belg*, Suppl I3.
- Posthuma D, **Gosso MF**, Luciano M, de Geus EJC, Wright MJ, Montgomery GW, van Belzen M, Heutink P, Slagboom PE, Boomsma DI, and Martin NG (2005) *Gene finding for intelligence*. *BGA*, 10.
- Posthuma D, **Gosso MF**, Altink B, Polderman TJC, Heutink P, van't Ent D, Boomsma DI, and de Geus EJC (2007). Imaging, Genetics & Cognition: a pilot study involving the *CHRM2* and *SNAP-25* genes. *BGA*, 15.

DANKWOORD  
ACKNOWLEDGEMENTS  
AGRADECIMIENTOS

---

Firstly, I would like to thank the generous participation of all the twin families that took part in this study. This thesis will not have been possible without them and of course, all the laborious work collecting their phenotypic information done at the NTR.

My promoters were fundamental, I am glad I was able to be “*in between*” two major fields. Prof.dr. Dorret Boomsma. Dear Dorret, I want to thank you for giving me the opportunity to join the Biological Psychology department. Your advice brought me a different perspective that I will always appreciate.

Prof.dr. Peter Heutink. Dear Peter, I enjoyed a lot doing the practical work at the Medical Genomic section. I think the “mental-coaching-meetings” paid-off ! (*but still, I regret not have been able to conquer for us the espresso-machine*).

I am also very grateful to my co-promoter, Danielle. Dear Danielle, I really do not know *how*, but you were always available for last minute discussions, tips, and support.

Although not as an “official” co-promoter, prof.dr. Eco de Geus. Dear Eco, you were quite a key player specially at the end of this project (*I start to believe you might be able to travel in time in order to check drafts for publication*).

I want to thank also to the reading committee: dr. Wim Crusio, prof.dr. Nick Martin; dr. Ger Ramakers, dr. Sophie van der Sluis, prof.dr. Cornelia van Duijn, and prof.dr. Mathijs Verhage for their valuable comments.

The secretaries from BP, Natascha, Hannah and Michiel, thanks for keeping things smoothly organized for me, in particular Natascha and Hannah during the last phase.

Many people helped me along this journey, my colleagues from BP, although I spent most of my time “*on the other side*”, I am particular grateful to my room mates Annabelle and Dirk, for their cordiality and help during my short stays. Also Tinca, we share a couple of publications and I appreciate and enjoyed a lot working together.

My colleagues from MGA. At the very beginning, when I started the practical work, the MGA section consisted just in a single lab with couple of PCR machines and empty benches. I have the privilege to witness how this original scenario turns into what is now. Also colleagues from the VWM group. I want to thank all for technical tips and practical set-up advice as well as personal support.

My paranimfen, Maria and Luba, thanks for being literarily “side-by-side” in such especial occasion .

---

Finally, to my family and friends from my beautiful and always remembered Argentina. Más que trillado y hasta cliché quizás, pero mi corazoncito late con un dejo de nostalgia cada vez que pienso en mi linda Paraná. Generalmente separados por la distancia, son incontables los momentos en los que estuvieron y seguirán estando junto a mí.

A mis amigos, por las incontables *asados* y tardes de *maté* compartidas (mates-lavados, mi especialidad, no crean, que es todo un arte lograr sebarlos así!). A mis queridos tía Mary y Norman, dos queridas personas que me ayudaron a ver la vida con mayor perspectiva y que siempre voy a llevar en mi corazón.

Y finalmente, a mis padres, por haberme brindado su amor incondicional y la libertad para crecer.

Florencia ,  
Noviembre 2007.