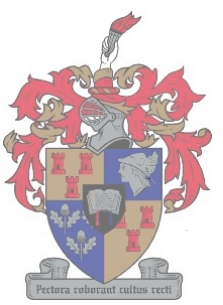
THE SEARCH FOR TOURETTE SYNDROME GENES:

A CONCEPTUAL AND EXPERIMENTAL APPROACH



Ingrid Simonic

Dissertation presented for the Degree of Doctor of Philosophy

at the University of Stellenbosch, South Africa

Promoters: Dr George S Gericke

Prof Andries E Retief

December 1999

I the undersigned hereby declare that the work contained in this dissertation is my own original work and has not previously in its entirety or in part been submitted at any university for a degree.

21-11-99

Ingrid Simonic Date

SUMMARY

Tourette syndrome has been reported in most populations throughout the world. Overall, there appears to be similar clinical phenomenology and psychopathology, which may serve as an indication of the biological nature for the condition.

The diagnosis of Tourette syndrome represents a challenge for physicians because of clinical heterogeneity and often-present comorbidity with other known neurobehavioural conditions. Due to these clinical overlaps Tourette syndrome may serve as a model disorder for investigating the relationship between various neurological and behavioral domains of childhood reflecting either the expression of a common biological pathway or a common genetic background. The understanding of the genetic basis of Tourette syndrome is therefore of special importance, because it may provide useful insights for the study of other developmental disorders. However, the lack of objective biological markers of clinical manifestation together with a possible high phenocopy rate, unclear mode of inheritance, incomplete penetrance, and frequent bilinear transmission of predisposing genes represent major obstacles for those attempting to elucidate the genetic basis of Tourette syndrome.

The research presented in this document is a result of six years' effort of the author and her collaborators to generate cytogenetic and molecular genetic data contributing to a better understanding of genetic and environmental factors affecting the phenotypic expression of Tourette syndrome. Theoretical and experimental results of this collaborative effort are assembled in seven articles (four published, three currently submitted for a publication) and a general introductory section relating to the problems, methods and methodology described and utilized in data collection for the individual papers.

Taken as a whole, while the study of chromosome fragile site expression in Tourette syndrome probands yielded equivocal results leading to a number of rather

speculative but interesting interpretations, the results of subsequent molecular genetic studies are far clearer.

The three most valuable outcomes of these studies for future genetic investigations in Tourette syndrome gene-mapping efforts in the Afrikaner population, and complex genetic traits in general, are:

1. The evidence for association/linkage of at least three genomic regions with Tourette syndrome in the Afrikaner population, with two of the regions (11q23 and 8q22) being suggestively linked to Tourette syndrome by others in different populations and employing different analytical methods.

2. The evidence for extended background linkage disequilibrium in the general Afrikaner population (> 5 cM) which further strengthens existing experimental data demonstrating the suitability of this population for gene-mapping efforts involving complex traits.

3. The proof based on real rather than computer-simulated data that sequential and semiparametric methods of analysis could be sufficiently powerful to generate cumulative evidence for positive linkage with the trait in the regions which repeatedly yielded both highly significant as well as suggestively significant disease-marker associations in the initial set of samples.

OPSOMMING

Tourettesindroom is 'n algemene oorerflike neurobiologiese probleem wat in verskeie bevolkingsgroepe vanoor die wereld beskryf is. As gevolg van identiese fenomenologie en psigopatologie ten spyte van omgewingsverskille, is dit aanduidend van 'n sterk biologiese grondslag vir die toestand.

Die teenwoordigheid van kliniese meersoortigheid en die verhoogde voorkoms van 'n verskeidenheid komorbiede probleme by 'n subgroep van individue met Tourettesindroom, veroorsaak dikwels probleme met die akkurate identifisering hiervan. Dit skep egter ook geleenthede vir die bestudering by kinders, van verskeie neurologiese en gedragsmanifestasies gebaseer op 'n gemene genetiese substraat. Insig in die genetiese-omgewings wisselwerking by Tourettesindroom baan dus die weg vir begrip van ander ontwikkelingsprobleme wat ook by kinders aangetref word. Die afwesigheid van 'n betroubare biologiese merker of merkers vir hierdie kliniese entiteit, die algemene voorkoms van fenokopieë, komplekse oorerwingspatroon, onvolledige penetrasie en algemene verskynsel van oorerwing vanaf beide ouers, verteenwoordig 'n aantal formidabele struikelblokke ten opsigte van die analise van die genetiese basis van Tourettesindroom.

TS word as een van die komplekse oorerflike toestande beskou, wat beteken dat daar duidelike oorerflike faktore by betrokke is, maar dat die oorerwing nie-mendelies van aard is. Die gebruiklike reduksionistiese benaderings wat so suksesvol was vir die analise van die enkelgeentoestande, werk nie meer onder hierdie omstandighede nie, en vir die rede word verskeie nie-parametriese of semiparametriese modelle ingespan.

Die gedokumenteerde resultate verteenwoordig die navorsing uitgevoer tesame met plaaslike en oorsese medewerkers op hierdie gebied gedurende die laaste ses jaar. Die teoretiese en eksperimentele resultate word weergegee in sewe publikasies. Hiertydens is sitogenetiese en molekulêre gegewens versamel in 'n poging om die genetiese en omgewingsfaktore onderliggend tot die ekspressie van Tourettesindroom

te bepaal. Die teoretiese en eksperimentele resultate van hierdie poging word weergegee in sewe publikasies, waarvan vier reeds gepubliseer is, en 'n algemene inleidende afdeling wat die probleme en metodes bespreek soos tydens die versameling en analise van die data ervaar is.

Die resultate word in twee afdelings aangebied: eerstens is daar die teoretisering ten opsigte van die bevinding van chromosomale breekbaarheid, wat aangedui is om verhoog te wees in die Tourette groep. Die betekenis van hierdie bevinding is tans nog onduidelik, en as gevolg van resolusieverskille nie direk met die DNA bevindings korreleerbaar nie. Hierdie merkerareas moet egter deurgaans in gedagte gehou word as moontlik aanwysend van die ligging van kandidaatgene vir Tourettesindroom.

Die belangrikste gedeelte behandel egter die benadering tot die totale genoomsifting, sowel as die veilgheidsmaatreels ingebou deur die heranalise van verskeie subgroepe en gevolglike replisering van resultate.

Die mees waardevolle implikasies van hierdie navorsing ten opsigte van die uitstippeling van die pad vorentoe vir Tourettesindroom geenkartering by die Afrikaner, en komplekse oorerflike toestande in die algemeen, sluit die volgende in:

1. Die bewyse gevind vir die bevestiging van 3 genomiese streke soos oorspronklik deur die eerste fase assosiasiestudies aangetoon by die manifestering van Tourettesindroom in die Afrikaner, en waar ten minste twee van die gebiede (11q23 en 8q22) ook deur ander navorsers in ander bevolkingsgroepe met hierdie toestand gekoppel is;

2. Die kwantifisering van die stand van koppelings-disekwilibrium by 'n aantal lokusse in die Afrikaner genepoel van < 5cM. Hierdie gegewens versterk die gedagtes met betrekking tot die geskiktheid van hierdie bevolkingsgroep vir geenkarteringspogings vir komplekse toestande;

3. Die bewys, gebaseer op reële in stede van gemodelleerde data, dat opeenvolgende, semiparametriese analisemetodes oor voldoende statistiese krag beskik om kumulatiewe getuienis te verskaf vir positiewe koppeling van TS met streke wat ook in die oorspronkilke siektemerker assosiasiestudies betekenisvolle resultate gelewer het.

"The day after I saw Ray, it seemed to me that I noticed three Touretters in the street downtown New York. I was confounded, for Tourette's syndrome was said to be excessively rare.

...Was it possible that I had been overlooking this all the time, either not seeing such patients or vaguely dismissing them as 'nervous', 'cracked', 'twitchy'?

The next day, without specially looking, I saw another two in the street. At this point I conceived a whimsical fantasy or private joke: suppose (I said to myself) that Tourette's is very common but fails to be recognized but once recognized is easily and constantly seen."

Sacks O. (1987)

CONTENTS

ACKNOWLEDGEMENTS

I wish to express my sincere gratitude to:

- My promoter, Dr George S Gericke, MRC Neurogenetics Research Initiative in Pretoria, and co-supervisor, Professor Andries E Retief, University of Stellenbosch, for their guidance and interest during the course of this project.

- Professor Jurg Ott, Department of Statistical Genetics, Rockefeller University, New York. Without his continuous academic, technical and financial support, a considerable portion of the experimental work on this project would have never been accomplished.

- Dr James L. Weber, Medical Research Foundation, Marshfield, for a very productive collaboration, molecular genetic training in his laboratory and for "keeping my spirits up" at all times.

- To my husband, Dr Milo Simonic, and my sons Viktor and Jan, for their patient understanding of my scientific aspirations.

- To the members of the Tourette Syndrome Support Group in Pretoria, particularly the chair of the group, Mrs Cicely van Straaten, for her interest and continuous help during the time-consuming process of the sample collection.

- Additional names of people and organizations to whom I am grateful for technical and financial support are mentioned at the end of each of the included articles.

LIST OF ABBREVIATIONS

ADHD	Attention deficit hyperactivity disorder
AML1	acute myeloid leukemia
APC	aphidicolin
ASP	affected sib pair
BN5IT	Temperature sensitive complementation gene, cell cycle specific
bp	base pair
BRCA1	Breast cancer – 1, early onset
BrDC	5-bromodeoxycitidine
BrDU	5-bromodeoxyuridine
CBL2	Cas-Br-M oncogene
CCM92	Chromosome Coordinating Meeting (1992)
CDR	cyclin-D related
cM	centiMorgan
COGA	Collaborative Study on the Genetics of Alcoholism
CT	Chronic tic disorder
DMSO	dimethyl sulphoxide
DNA	deoxyribonucleic acid
DRD2	dopamine D2 receptor gene
DSMIV	Diagnostic and Statistical Manual of Mental Disorders (1994), 4^{th} edition
DZ	dizygotic
EST	expressed sequence tag
FHIT	fragile histidine triad gene
FMR1/FRAXA	Fragile X mental retardation - 1
FMR2/FRAXE	Fragile X mental retardation, FRAXE - type
FS	fragile site
GEM	GTP-binding protein, over-expressed in skeletal muscle
GHRR	genotype-based haplotype relative risk method
GTP	guanosine triphosphate

GTS	Gilles de la Tourette syndrome
HHRR	haplotype-based haplotype relative risk method
HLA	Human leukocyte antigen
HRR	haplotype relative risk method
HTR-3	5-hydroxytryptamine (serotonin) receptor - 3
http	hypertext transmission protocol
IBD	identical by descent
IBS	identical by state
IGK@	immunoglobulin kappa light chain gene cluster
IgG	immunoglobulin G
kB	kilobase
LCR	locus control region
LD	linkage disequilibrium
LOH	loss of heterozygosity
Mb	megabase
MIM	Mendelian Inheritance in Man
mRNA	messenger ribonucleic acid
MTG8	the chimeric fusion AML1/ETO proto-oncogene
MZ	monozygotic
NIH	National Institutes of Health (USA)
OCD	obsessive compulsive disorder
OMIM	On-line Mendelian Inheritance in Man
PANDAS	Pediatric Autoimmune Neuropsychiatric Disorders Associated with Streptococcal Infections
PCR	polymerase chain reaction
QTL	quantitative trait locus
RF	rheumatic fever
RFLP	restriction fragment length polymorphism
SC	Sydenham chorea
STRP	short tandem repeat polymorphism
TCF9	transcriptional factor-9

TDT	transmission disequilibrium test
TS	Tourette syndrome
TSA	Tourette Syndrome Association
VNTR	variable number of tandem repeats
www	world-wide web

CHAPTER 1

General introduction

1.1. History

1.11. Recognition and clinical description of Tourette syndrome

The movement disorder known as Tourette syndrome has been recognized since 1885, the year in which Georges Gilles de la Tourette (1857-1904) published a two-part article describing the case histories of 9 French men and women with the syndrome (Lajonchere et al. 1996). His description of the disorder, which now carries his name, was based on 'the case of the cursing marquise' reported by Itard (1825) (MIM 137580). The marquise's life history was selected by Gilles de la Tourette, who himself never examined her, as a prototypical example of the syndrome's major features such as involuntary movements and sounds (i.e. barking), markedly enhanced startle reactions, coprolalia (inappropriate and involuntary swearing), and tendency to repeat both vocalizations (echolalia) and movements (echopraxia). In his article, Gilles de la Tourette assumed that the condition manifests itself in childhood, and does not affect the senses or intellect. Finally, he considered the condition to be hereditary with varying severity throughout a person's life-span and incurable.

Another man, who deserves credit for the description and recognition of the syndrome is Jean-Martin Charcot (1825-1893), the leading French neurologist of the time and Tourette's mentor. He urged Gilles de la Tourette to undertake the task of classifying movement disorders, and more importantly, he recognized the disorder, described by

his intern, as different from other movement disorders commonly understood at the time as "hysteria".

For nearly a century after its original description, Tourette syndrome was considered rare with only 485 cases reported worldwide by 1973 (Robertson and Baron-Cohen 1998). The situation changed dramatically mainly as a result of the work of Shapiro et al. (1982, 1989) at the Sinai Medical School in New York, who found, that many patients with Tourette-like symptoms responded to treatment with haloperidol, a dopamine D2 receptor inhibitor. Today, Tourette syndrome has become of great interest to neurologists and medical professionals in general, which led to the realization that the syndrome and related conditions are much more common than had previously been considered.

It is generally accepted that Tourette syndrome can assume many forms. For some people it may involve mild facial tics and odd vocalizations. For others it involves more dramatic uncontrollable movements, often accompanied by additional problems such as hyperactivity, poor attention, obsessions and compulsions (Robertson and Baron-Cohen 1998).

The diagnostic criteria for Tourette syndrome presently included in Diagnostic and Statistical Manual of Mental disorders, 4th ed. (DSM-IV) provide a reasonable basis for the diagnosis, however, they do not provide an adequate description of the numerous symptoms that can co-occur as part of the phenotype. DSM-IV criteria also fail to recognize the developmental course of the disorder typically characterized by varying severity of expression of multiple symptoms, as well as diminution of the symptoms in adolescence (Leckman et al. 1998).

The initial signs of Tourette syndrome are usually involuntary tic-like movements, which may progress in the course of disease to echolalia, grunting, coprolalia, and self-mutilation. Earlier studies have found, that self-mutilation symptoms are present in ~40% of clinic-patient populations (van Woert et al. 1977). Coprolalia, on the other hand, previously thought to be one of the most notorious symptoms of the syndrome,

occurred less frequently (~10%) in studied patient populations (Goldenberg et al. 1994) and may be a culturally related phenomenon due to widely different prevalence in Japanese and US individuals with Tourette syndrome (Nomura and Segawa 1982, Robertson and Stern 1997).

While the diagnosis of Tourette syndrome is quite straightforward for the physician, the degree to which other behaviors are associated and represent the spectrum of Tourette symptoms is not clear. In a controversial presidential address to the American Society of Human Genetics, Comings (1989) extended the phenotypic range of Tourette syndrome to include attention deficit disorder, conduct disorder, major depressive disorder, manic-depressive disorder, panic disorder, schizoid disorders, sleep disorders, specific reading disability, stuttering, male type II alcoholism and a female type of familial obesity. He suggested that the spectrum of behaviors associated with Tourette syndrome could be explained on the basis of a gene causing an imbalance of the mesencephalic-mesolimbic dopamine pathways, resulting in dis-inhibition of the limbic system.

Pauls et al. (1988) criticized the methods and conclusions of the above author. In a family study of 86 probands the authors found only chronic tics (CT), and obsessive-compulsive disorder (OCD) with increased frequency in the first degree relatives. They did not observe an increased frequency of any other behavioral condition suggested by Comings (1989) as part of Tourette syndrome spectrum phenotype by comparing the first degree relatives of Tourette index patients with the relatives of control subjects.

Further investigations of these initial observations led to the conclusion that at least some forms of OCD and CT are etiologically related to Tourette syndrome (Pauls et al 1991). The conclusion was also supported by data from families of OCD probands, where the rates of Tourette syndrome and CT in the relatives were elevated (Leonard et al. 1992, Pauls et al. 1995). In the course of the above studies it became apparent that the nature of obsessions and compulsions that occur among the relatives of Tourette syndrome probands as well as the treatment responses with respect to OCD

are different than those experienced by patients/families with pure OCD (no personal or family history of tics) (Eapen et al. 1997, Leckman et al. 1997, Zohar et al. 1997).

Considerable effort has been made in an attempt to elucidate the relationship between Tourette syndrome and attention deficit hyperactivity disorder (ADHD). While Comings and Comings (1987) proposed ADHD to be a variant expression of the etiologic factors responsible for the manifestation of Tourette syndrome and CT, Pauls et al. (1986a, 1993) found no support for such a proposed relationship in their studies. They suggested that while it is possible for ADHD to be associated with increased clinical severity of Tourette syndrome, it is unlikely that in the absence of tics, ADHD was a variant expression of genetic factors underlying Tourette syndrome.

In 1993, Kiessling et al. reported an increase in tic disorder frequency in children following a community outbreak of streptococcal infection in Providence, Rhode Island. Not only did tics begin abruptly following the infection, antineuronal antibodies directed against human caudate were found in 45% of tic cases (n=30), compared to 20% of controls. The authors speculated that some cases of Tourette syndrome may result from antibodies that cross-react with streptococcal antigens mainly in the basal ganglia, in a process similar to Sydenham Chorea (SC).

In 1995, Allen et al. reported four cases of a new, infection-triggered, autoimmune subtype of Tourette syndrome and pediatric OCD, called pediatric autoimmune neuropsychiatric disorder associated with streptococcal infections (PANDAS). Subsequent reports have indeed suggested that an autoimmune reaction triggered by infection and directed against the brain may contribute to the pathogenesis of tics, GTS and OCD. Moreover, Swedo et al. (1997) reported that 23 of 27 (85%) PANDAS patients, eight of nine (89%) SD patients, and four of 24 (17%) healthy children were positive for a B-cell antigen, known as D8/17, an immune marker for rheumatic fever (RF). The findings were confirmed by Murphy et al. (1997).

Finally, an animal model was constructed to support the role of antibodies in producing tic-like movements and sounds in rats micro-infused with IgG from serum derived from Tourette patients. Brain slides of these rats showed preferential staining of striatal regions with IgG from Tourette patients but not with IgG from controls (Hallett et al. 1996, 1997).

The precise relationship between the core phenotype of the tic disorder and the associated features await the elucidation of the nature of this condition on molecular level.

1.12. The inheritance of Tourette syndrome

The familial nature of Tourette syndrome was first noted and commented on by de la Tourette himself in his 1885 article.

Since a study of Eisenberger et al. was reported (1959), familial aggregation of the syndrome has been confirmed in a large number of published and unpublished Tourette syndrome pedigrees. The results of most family studies were remarkably consistent (for review see Alsbrook and Pauls 1997) with reports of elevated rates of Tourette syndrome and CT among first degree relatives when compared to the rates in control samples or the general population.

To prove the existence of genetic factors in the manifestation of the disorder, results of family studies were followed by the analysis of concordance rates for the affection status in twins. Price et al. (1985) and Hyde et al. (1992) reported significantly higher concordance rates among monozygotic (MZ) twins as compared to dizygotic (DZ) twins when either Tourette syndrome, CT or OCD were considered as affected. The MZ twin data also suggested that non-genetic factors play a role in the manifestation of Tourette syndrome, since the concordance rates in affection status were < 1.0, and severity of affection status varied.

Earlier segregation analyses performed on collected family history data demonstrated that a single-major-locus hypothesis best explained the patterns of observed Tourette syndrome transmission (Comings et al. 1984, Devor et al. 1984, Price et al. 1988) and provided strong evidence for an autosomal dominant model of inheritance (Pauls and Leckman 1986b).

The results of two more recent segregation analyses (Hasstedt et al. 1995, Walkup et al. 1996) provide evidence for a major gene model with a more complex mode of inheritance. According to this mixed model of inheritance, it can be predicted that 0.01% of individuals in the population are homozygous for the susceptibility allele, 1.89% are heterozygous, and 98.1% are homozygous for a normal allele. The placement of the threshold for liability indicates that all individuals homozygous for the susceptibility allele at the major locus are affected, whereas only 2.2% of males and 0.3% of females heterozygous at the major locus are affected. Parameter estimates from the mixed model of Tourette syndrome inheritance predict that 38% of individuals affected are homozygous for the major locus, whereas 62% of affected individuals have only one copy of the susceptibility allele. The contribution of the multifactorial background accounts for an estimated 40%-45% of the phenotypic variance (Walkup et al. 1996).

While the mode of inheritance is not simple, it is clear that Tourette syndrome has a significant genetic basis and that some individuals with Tourette syndrome, CT and OCD manifest variant expression of the same genetic susceptibility factors. The localization and characterization of genetic factors responsible for the expression of Tourette syndrome is of major importance for our understanding of the pathogenesis of this disorder.

Attempts to localize the responsible gene(s) have thus far not yielded consistent positive results. Linkage analysis of data from series of multiply affected families resulted in the exclusion of > 90% of the genome (see Barr and Shandor 1998 for review). These analyses were completed assuming a dominant mode of inheritance

and locus heterogeneity for Tourette syndrome and spectrum disorders, which could have led to false exclusion of relevant genomic region(s).

1.2. The aims of the study

To elucidate the genetic basis of Tourette syndrome, different methods were employed during the course of the study:

1. Increased expression of chromosomal fragile sites has been documented for several psychiatric and neurological disorders including schizophrenia and bipolar disorder. In order to systematically search for subtle chromosomal abnormalities and/or Tourette syndrome specific fragile site expression, cytogenetic investigations have been initiated in a random group of Tourette syndrome index cases. The initial study (Chapter 2) involved the evaluation of spontaneous, rare-folate sensitive and BrDU-inducible FS expression in the Tourette syndrome males as opposed to age-matched controls.

2. Because the rare-folate sensitive and BrDU-inducible fragile sites represent only a small fraction of all fragile sites inducible on human chromosomes, a second study was initiated in order to investigate the expression of common fragile sites in the Tourette syndrome index cases as opposed to the controls (Chapter 3). The aim of this study was to define those common fragile sites, which could serve as discriminatory cytogenetic markers for Tourette syndrome.

3. The overview of literature, spanning two decades of fragility studies was condensed and published with the intention to provide a theoretical background for our chromosomal fragility findings in association with the Tourette syndrome phenotype, as well as the identification of a new group of common aphidicolin-inducible fragile sites, not previously reported (Chapter 4).

4. Our initial idea to follow up a particular subgroup of chromosomal regions characterized by increased fragility in the Afrikaner Tourette syndrome

individuals on the molecular level was eventually substituted by a different method. The new strategy, which represents a unique approach to mapping a common complex trait, consisted of several subsequent steps. The whole genome search for marker-disease association was performed using pooled DNA samples genotyped with > 1,000 genetic markers distributed throughout the genome. Markers with differences in allelic distributions between case and control pools were then subjected to individual typings in two non-overlapping sets of case-control samples and subsequently to the statistical evaluation of marker allele distributions. The published results of the work (Chapter 5) represent an important contribution to current Tourette syndrome gene-mapping efforts.

5. The Afrikaner population of South Africa is regarded as genetic isolate suitable for gene-mapping efforts, particularly because a strong founder effect has been repeatedly documented for several monogenic traits with an increased prevalence rate in the population. For a number of these disorders, extensive haplotype sharing was documented among affected individuals at or near the disease loci. However, until recently no investigations were performed in order to examine the background linkage disequilibrium (LD) distribution in the general Afrikaner population. Our preliminary investigation of the extent of background LD distribution (Chapter 6) represents an important step in paving the way for utilizing the Afrikaner population in whole genome association studies.

6. Significant case-control association findings, even when achieved in young genetically isolated population, may still be a result of population stratification, therefore all association findings require to be confirmed by nuclear-family-based linkage methods such as the transmission disequilibrium test (TDT) and/or haplotype relative risk method (HRR). The results of our marker-disease association findings for Tourette syndrome are regarded as preliminary, until confirmed by semiparametric linkage/LD methods using nuclear-family genotyping data in genomic regions of interest (Chapter 7).

1.3. Introduction to the methods utilized to investigate the genetics of Tourette syndrome during the present study

1.31. Fragile sites on human chromosomes

Agents known to inhibit DNA replication induce the expression of chromosomal fragile sites (FS), which appear as gaps, breaks, or discontinuities in chromosome structure. When examined in metaphase preparations, it is not possible to distinguish between a random gap or break in chromosomal structure and chromosomal FS. Only statistically significant recurrence of gaps, breaks, or lesions at the same chromosomal band (region) and under the same culture conditions delineates FS (Sutherland and Richards 1999).

The majority of culture conditions resulting in FS expression cause inhibition of DNA repair or replication, either due to nutrient deprivation (e.g. perturbed nucleotide pools) or inhibition of DNA replication enzymes (e.g. DNA polymerase alpha). As a consequence, under-replicated DNA sequences, primarily at FS (Hansen et al. 1997, Le Beau et al. 1998) do not package completely before the G2 phase and manifest as discontinuities in chromosomal structure.

1.311. Rare fragile sites

FS were initially classified according to the methods of their induction (Sutherland et al. 1998). The two main classes were "rare" and "common" FS. The common FS appear to be part of normal chromosome structure and are present at all common FS loci in every individual. The frequency of their expression differs among the individuals and is modulated by factors like age, sex, and hormonal status (Tedeschi et al. 1992).

The most widely studied subgroup of the rare FS are so called rare-folate sensitive sites, induced by folic acid and thymidine deprivation in the cell culture media. The frequencies of their expression are relatively low (in ~4%-20% investigated

metaphases) and also the population frequency of individuals carrying one of these sites is low (~ 5%).

Rare folate-sensitive FS comprise the most studied class of FS at the molecular level and their relationship with tandemly repeated sequences of varying complexity has been well established. The first cloned FS site was FRAXA, a rare-folate sensitive FS at the chromosome Xq28 region, cytogenetically expressed in association with the most common form of male mental retardation called Fragile X syndrome. The DNA sequence at the FRAXA locus is characterized by tandemly repeated CCG units with interspersed CCT repeat units (Hirst et al. 1994). Due to increased copy numbers of CCG units alone (>55 repeat units), the sequence becomes prone to the expansion and subsequently a full fragile X mutation. A common haplotype of surrounding sequences has recently been characterized, which can be used to predict which alleles at FRAXA locus are likely to proceed to expansion (Gunter et al. 1998).

Four other rare FS have been cloned: FRAXE and FRAXF at Xq28, FRA11B at 11q23.3 chromosomal region and FRA16A at 16q22. All of these sites are associated with $(CCG)/(CGG)_n$ triplet repeat expansions which become hyper-methylated beyond a critical size/number of repeats. Three of them are associated with clinical problems and in two cases, gene responsible for disease state was identified: *FMR1* in FRAXA (MIM 309550) and *CBL2* (MIM 165360) in FRA11B. FRAXE has also been associated with a mild form of mental retardation, expansion of $(CCG)_n$ arrays (>200 repeats) and hypepermethylation of a CpG island adjacent to a gene called *FMR2* (MIM 309548)(Gecz et al. 1996).

FRA16A is characterized by longer CCG repeat units that lack CCT interruptions and are more prone to expansion (Nancarrow et al. 1994, 1995). FRA11B was found in a mother and brother of a child with Jacobsen syndrome, suggesting that the breakage in FS during early development could have resulted in a chromosome deletion in the patient. This FS has been assigned to an interval of approximately 100 kb containing the 5' end of the *CBL2* gene, which includes a CCG trinucleotide repeat. The chromosomal deletion breakpoint in the patient was mapped within the same interval

(Jones et al. 1994). In later reports the association of the 11q23.3 deletions with CCG expansions in *CBL2* and FRA11B expression was not confirmed (Michaelis et al. 1998).

1.312. Common fragile sites

A large number of common FS are expressed as a result of aphidicolin (APC) treatment of cell cultures 24 hours before harvesting. The most frequently observed (~70% - 80% metaphases) in all individuals is FRA3B, the FS on chromosome 3p14.2. The second most frequently expressed is FRA16B, the FS on chromosome 16q23. The frequencies of expression differ in various reports and are mainly due to different culture conditions and also to APC being dissolved in dimethyl sulphoxide (DMSO) (resulting in lower expression rates at most FS), or in ethanol. This enhances overall fragility caused by APC, even though ethanol itself does not induce FS expression and has no clastogenic effects. The rank orders of expressed FS per metaphase remain relatively consistent between different studies.

Although they comprise the vast majority of fragile sites, much less is known at the molecular level about the "common" fragile sites. These FS sites are seen as a constant feature of all chromosomes and have been shown to display a number of characteristics of unstable, highly recombinogenic DNA in vitro, including chromosome rearrangements, sister chromatid exchange and, more recently, intrachromosomal gene amplification (Glover 1998).

Only one such fragile site, FRA3B at 3p14.2, has been extensively investigated at the molecular level. It extends over a broad region of about 500 kb, and no trinucleotide or other simple repeat motifs have been identified in the region. The FS lies within the *FHIT* gene locus (MIM 601153), which is unstable in a number of tumors and tumor cell lines (Heubner et al. 1997, 1998). It thus appears that genomic instability at common fragile regions has a potential to facilitate chromosome rearrangements associated with cancers (Popescu et al. 1990, Paz-y-Mino et al. 1992, Popescu et al. 1994, Wilke et al. 1996, Huang et al. 1998, Smith et al. 1998, Huang et al. 1999). The

co-occurrence of oncogenes and cancer suppressor genes at the same chromosomal bands as common FS has long been recognized, but only recently, studies of environmental and genetic factors that influence FS expression and instability became an important field of research aiming towards a better understanding of malignancy.

FRA16B, the FS on chromosome 16q23 induced by DNA minor groove-binding agents such as distamycin A and berenil, and FRA10B, induced by BrdU and/or BrdC, have also been cloned (Yu et al. 1997, Hewett et al. 1998). Both FS are caused by highly expanded (up to several thousand copies to yield a fragile site) AT-rich microsatellite repeats, which vary in size and composition due to somatic and intergenerational instability.

Different aspects of chromosomal fragility currently under investigation are: the potential of certain viruses to induce specific FS expression in infected cells (Li et al. 1998) and the association of common fragility with early events of DNA amplification leading to acquired resistance to drugs (Kuo et al. 1998). Both types of observation suggest that high local levels of transcription can interfere with metaphase chromatin packaging and are sufficient to generate fragile chromosome areas. The fact that common FS could represent the cytogenetic expression of transcriptionally active regions was also supported by Sbrana et al (1998) in their study of FS expression modulation by camptothecin, a specific inhibitor of topoisomerase I.

1.32. Genetic mapping

1.321. The principles

Genetic mapping in principle means comparing the inheritance pattern of disease traits with the co-inheritance pattern of certain chromosomal regions. *Mendel's laws* of genetic inheritance, and mathematical formulae developed by *J. B. S. Haldane* that relates map distances to recombination frequencies provided the key elements in early systematic searches for disease-causing genes in experimental crosses. After the

recognition of naturally-occurring DNA sequence variations as a source of genetic markers, it became possible to trace the inheritance in human pedigrees.

The aim of genetic mapping therefore, is to evaluate how often two loci (e.g. disease-marker locus, marker-marker locus) are separated by meiotic recombination. Three common but distinct measures of the separation of loci are used in the process:

1/ *Recombination fraction, θ,* the probability that two loci will be separated by recombination at meiosis.

2/ *Map distance*, measured in *centimorgans, cM*, named after the American geneticist Thomas Hunt Morgan, represents the expected number of recombinations occurring between two loci at meiosis. 1 cM equals a crossover value of 1%.

3/ *Physical distance*, measured in base pairs, *bps*, of DNA.

A single recombination event during meiotic cell division produces two recombinant and two non-recombinant chromatids (progeny). If two loci under investigation are on different chromosomes, they will segregate independently and the chance that a daughter cell will be recombinant or non-recombinant for these loci on particular chromosomes is 50%. In average, two loci cannot produce more than 50% recombinants, not even in the case of double or triple crossover events. Recombination will rarely separate loci, which lie very close together on a chromosome, because only a crossover located precisely in the small space between the two loci will create recombinants. The further apart two loci are on the chromosome, the more likely it is that a crossover will separate them. Thus the recombination fraction is a measure of genetic distance between the two loci.

The relationship between the recombination fraction, genetic map distance and physical distance is non-linear and variable in different parts of the genome and between the sexes. For small distances there is an approximate equivalence between $\theta = 0.01$, map distance = 1 cM and physical distance = 10^6 bp (one Megabase). The

haploid human genome comprises approximately 10^9 Mb and has a sex-averaged map length of approximately 3,300 cM.

The only possible way to recognize recombinants from non-recombinants at two genetic loci is to use loci with more than one sequence variant (allele) in the human population. Such allelic sequence variation is described as a DNA polymorphism if more than one variant at a locus occurs with a frequency greater than 0.01 in a human population. It has been calculated that DNA polymorphisms occur approximately in 1/ 250 to 1/ 300 bases in human genomic DNA.

The identification of different types of DNA polymorphisms made it possible to develop different sets of polymorphic genetic markers. The first generation of DNA markers, were called *restriction fragment length polymorphisms (RFLPs)*, owing their nomenclature to the existence of restriction site polymorphisms. They had only two alleles (the restriction site was either present or absent), which made most of the meioses uninformative for the *RFLPs* (the recombinants could not be distinguished from non-recombinants). *Minisatellite (VNTR)* markers were a great improvement, since large number of alleles became available for which most meioses were informative. Classical minisatellites, however, have been difficult to handle with standard PCR techniques, because they span a large area and often fail to amplify. In addition, they seem to cluster in subtelomeric regions of chromosomes.

The standard tools in current genetic mapping efforts are *microsatellites*, which are moderately sized arrays of tandemly repeated DNA sequences, highly polymorphic and dispersed over considerable portions of the nuclear genome. The bulk of them are $(CA)_n$ repeats (Généthon). The disadvantage of markers based on dinucleotide repeat sequences for large-scale genotyping is that they are prone to replication slippage during PCR amplification. This means, each allele gives a ladder of 'stutter bands' on a gel, which makes genotyping results difficult to read. Tri- and tetranucleotide repeats usually give clearer results with a single band from each allele and are gradually replacing dinucleotide repeats as the markers of choice. Much effort is being devoted to producing compatible sets of microsatellite markers which can be

amplified together in a multiplex PCR reaction, and have allele sizes which allow them to be run in the same gel lane without producing overlapping bands (Center for Medical Genetics, Marshfield).

Typing suitable families with highly polymorphic genetic markers equally spaced throughout the genome and scoring of the genotypes is usually followed by statistical evaluation of the likelihood that disease and marker loci are linked. When large and sufficiently informative families are used in gene mapping efforts, such an analysis is relatively simple. Usually, however, only imperfect family data are available and recombinants cannot be identified unambiguously. Computer-generated lod score methods subsequently have to be applied for extracting linkage information from such 'imperfect' families.

The *lod score, Z,* is the logarithm of the odds that two loci are linked (with recombination fraction θ) rather than unlinked (recombination fraction 0.5). The lod scores are calculated for a range of θ values by looking at each meiosis, marker after marker (two-point linkage analysis), or by analyzing genotyping data for more than two markers simultaneously (multipoint linkage analysis). All lod scores are zero at θ=0.5 since they measure the ratio of two independent probabilities. If there are no recombinants between the disease and marker locus, the lod score will reach a maximum at θ=0. If there are recombinants, the thresholds for a single test are Z=3.0 and Z= -2.0. The threshold for accepting linkage, with a 5% chance of error is Z=3.0 or 1000:1 odds. Linkage can be rejected if Z<-2.0.

Values of Z between –2.0 and 3.0 are inconclusive when applying multiple marker typings (e.g. in whole genome searches) since the chances of spurious positive results are greater when compared to a situation where only one marker was typed. The threshold lod score for a study using *n* markers would then be $3 + log(n)$. In practice, lod scores below 5, whether with one marker or many are regarded as provisional and require to be followed up by confirmatory studies (Ott 1991, Terwilliger and Ott 1994).

Standard lod score analysis is a tremendously powerful method for scanning the genome in 20-30 Mb segments to locate disease genes. Unfortunately, it has some drawbacks, in that it can be effectively applied only for Mendelian monogenic disorders:

- Standard lod score analysis requires specification of a precise genetic model, including the mode of inheritance, the gene frequency and a penetrance of each genotype.
- It has limits with respect to the achievable resolution - not < than 1 Mb - an uncomfortably large genetic region for positional cloning of an unknown disease gene.
- Locus heterogeneity (the disease phenotype is produced by mutations in two or more unlinked genes) can cause a failure to identify linkage in either of the regions.

Not all of the 65,000 – 80,000 human genes will be identified as disease-causing genes. Those genes, which are indispensable to embryonic development, where mutations are mostly lethal, will remain unrecorded in humans. Out of the currently listed 5,000 Mendelian traits, about $1/10^{th}$ have been placed on the human genome map and only about $1/100^{th}$ of the disease genes were identified by positional cloning (OMIM).

1.322. Gene mapping strategies for complex traits

The term '*complex trait*' refers to any phenotype that does not exhibit classic Mendelian recessive or dominant inheritance attributable to a single gene locus. In general, complexities arise when the simple correspondence between genotype and phenotype breaks down, either because the same genotype results in different phenotypes (due to the effects of chance, environment, or interaction with other genes) or different genotypes result in the same phenotype. Most common traits of medical relevance belong to this category, including those responsible for

susceptibility to heart disease, hypertension, diabetes, cancer, infectious diseases, and the majority of neuropsychiatric disorders (Lander and Schork 1994).

After impressive successes in mapping single gene disorders, the attention of many investigators is turning to more challenging problems of the genetic dissection of complex traits. The majority of these 'gene-hunting' efforts were, however, hampered by a fundamental genetic complication such as the 'imperfect co-segregation' of genetic markers with complex traits caused by:

1/ *Incomplete penetrance* and high *phenocopy* rate, which means that the genotype at a given locus may affect the probability of the disease, but not fully determine the outcome. The predisposing allele may than be present in some unaffected individuals and absent in some affected individuals.

2/ *Genetic (locus) heterogeneity,* implying that mutations in any one of several genes at different chromosomal loci may result in identical disease phenotypes. One of the chromosomal loci will then co-segregate with a disease in some families but not in others. Genetic heterogeneity is different from *allelic heterogeneity*, in which one finds multiple disease-causing mutations at a single gene locus. Allelic heterogeneity tends not to interfere with gene mapping.

3/ *Polygenic inheritance,* refer to those traits which require the simultaneous presence of several mutations at multiple loci. Polygenic traits can be classified as discrete traits, measured by a specific outcome, or quantitative traits, measured by a continuous variable. Polygenic inheritance complicates genetic mapping, because no single locus is strictly required to produce a discrete trait or a high value of a quantitative trait.

4/ High frequency of disease-causing alleles in the population, often resulting in *bilineal transmission* of disease alleles in affected pedigrees. This interferes

with traditional linkage analysis and becomes an even greater problem if combined with genetic heterogeneity.

The first challenge in the genetic dissection of a complex phenotype is the identification of a candidate map location of the genes underlying susceptibility/resistance to the trait via conducting a genome-wide search for linked/associated genetic loci. The second challenge is the fine-structure localization of any component gene to physical segments small enough to facilitate positional cloning or recognition of candidate genes (Devlin and Risch 1995).

The methods available for genetic dissection of complex traits fall into four categories: linkage analysis, allele-sharing methods, association studies in human population, and genetic analysis of large crosses in model organisms such as the mouse and rat. The latter will not be discussed in this thesis.

Linkage analysis

If a limited number of loci are major determinants of susceptibility to a complex trait, it should be possible to map such loci by linkage analysis (Risch 1990). Large pedigrees typically contain a broad spectrum of phenotypes for a complex disease, therefore, only families characterized by a strong history of disease, with ostensible mendelian inheritance can be chosen for analysis. This approach has been successfully applied to mapping and later identification of the *BRCA1* (MIM 113705) breast cancer susceptibility gene on chromosome 17, when using age of disease onset as a quantitative trait (QT). By examining the inheritance pattern of pairs of regions, multiple sclerosis has been mapped in large Finnish kindreds to both *HLA* (MIM 142830) on chromosome 6p21.3 and the gene for myelin basic protein on chromosome 18 (Tienari et al. 1992).

Like any model-based method, linkage analysis can be very powerful, if the correct genetic model has been specified. The use of a wrong model, however, can lead to misspecification of true linkages or accepting false ones. For some psychiatric disorders this approach has been a source of false positive findings, e.g. in mapping

schizophrenia to a locus on chromosome 5 (Kennedy et al. 1995), particularly because the model requires persons to be classified clearly as affected or unaffected. Such a classification depends mainly on the age of onset of the disease, on the diagnostic criteria used and their reliability, and also on the validity of diagnostic categories for genetic research (Farmer et al. 1994). For these and other difficulties, model-free, non-parametric methods are preferred in psychiatric genetics. These methods ignore unaffected people, but look for shared chromosomal segments in affected individuals.

Allele sharing methods

Shared segment methods can be used in nuclear families, e.g. sib pair analysis (Fulker and Cardon 1994), within known extended families, or in populations that are descended from a small founder group (Holmas 1993, Holmans and Craddock 1995, Houwen et al. 1994). Pairs of sibs are expected to share 0,1 or 2 parental haplotypes with a frequency of ¼, ½ and ¼, respectively. If both sibs are affected by a genetic disease, they will share a chromosomal region carrying the disease locus with higher frequency than predicted by random segregation. Sib-pair studies require no prior assumptions about parameters such as mode of inheritance, penetrance, phenocopy rate, and disease allele frequency (Kruglyak and Lander 1995a).

The affected sib pair approach has been successfully applied to mapping the non-HLA susceptibility locus for type 1 diabetes on chromosome 11 (Davies et al. 1994), and is strongly pursued with other complex diseases, including schizophrenia, manic depression, alcoholism, and Tourette syndrome (The International Tourette Syndrome Genetic Linkage Consortium).

Because allele-sharing methods are non-parametric (that is, they assume no model for the inheritance of the trait), they tend to be more robust but less powerful than a correctly specified linkage model. The power of allele-sharing methods to demonstrate linkage for a complex trait depends on a number of factors: number of sibships (trios, case-control individuals), degree of genetic heterogeneity, risk ratio for the sibs versus population prevalence, and informativeness of the marker (Göring and Ott 1997). In the next couple of years, the application of non-parametric methods in

genetic mapping is expected to produce a large number of susceptibility loci, with many false positive findings as a tradeoff. The true susceptibility loci will then have to be sorted out by well-designed confirmatory studies.

An important difference between linkage mapping of single and complex disorders is that, whereas for single gene diseases recombination events can define an exact interval in which the disease gene must lie, in complex diseases recombination events can only alter the probability that the susceptibility locus is localized within a particular interval. Fine linkage mapping for complex traits, therefore, requires very large samples. For example, localizing a susceptibility gene to a 1 cM interval requires a median of 200 sib pairs for a locus causing a fivefold increased risk to a first degree relative and 700 sib pairs for a locus causing a 2-fold increased risk. To narrow the candidate chromosomal regions defined by allele-sharing methods, population-based linkage disequilibrium or candidate gene approaches may be applied (Craddock and Owen, 1996).

1.323. Population association studies in mapping susceptibility loci

An alternative to linkage mapping in families is to look for statistical association between a disease and some marker genotype at the population level (Owen and McGuffin 1993, Risch and Merikangas 1996). While linkage implies a relationship between loci, association represents a relationship between alleles, meaning that unrelated people across the whole population, who have a certain allele at one locus have a statistically more than random chance of having some particular allele at a second locus. Linkage is usually necessary, but never sufficient, for allelic association (Hodge 1993).

Disease–marker association studies are based on a comparison of unrelated affected (case) and unaffected (control) individuals with the same population background. The marker–allele at a particular locus of interest is associated with the disease trait if it occurs at a significantly higher frequency in affected as compared to control individuals. In the case of a positive association finding, the associated allele may

directly cause susceptibility to the trait and will be associated with the disease in every human population. Alternatively, a particular marker is in a close proximity to the disease gene, which means that the allele is in linkage disequilibrium (LD) with the disease-causing mutation (Greenberg 1993, Hodge et al. 1981, Hodge 1994, Jorde 1995).

Positive association can also arise as an artifact of population admixture, meaning that affected and control individuals originate predominantly from ethnically different populations with different marker-allele frequencies (Kidd 1993). To prevent spurious association arising from population admixture, association studies should be performed within genetically homogeneous populations or by using 'internal controls' for marker-allele frequencies: a study of affected individuals and their parents, such as the haplotype relative risk method (HRR) (Falk and Rubinstein 1987, Terwilliger and Ott 1992), and transmission disequilibrium test (TDT) (Spielman et al. 1993, Spielman and Ewens 1996).

Genomic searches for association are most meaningful if performed in young, genetically isolated populations in which LD extends over greater genetic distances, and the number of disease-causing mutations is likely to be fewer. Suitability of a population for the localization of disease genes by disequilibrium mapping is usually assessed from the demographic history of particular population, and from the existence of a founder effect for the disease under the study in particular population (Laan and Pääbo 1997).

When positive association findings are reported, various research groups attempt to replicate the original findings in independent studies (Sobell et al. 1993). Each of such confirmatory studies, however, might slightly differ from the original design, e.g. by revising the definition of the disorder. This raises a question, whether each confirmatory step is in fact a valid attempt to replicate the original finding, or whether each step should rather be considered an exploratory analysis generating new hypotheses after failing to support the original hypothesis.

Such 'definition drift' can be minimized if a single group of investigators attempts to replicate their initial positive findings on discrete subsamples using uniform diagnostic criteria and laboratory procedures. Such a sequential approach to the study design and analysis of case-control data includes a reduction of the number of candidates investigated (genomic areas, candidate genes, genetic markers) by testing (analyzing, genotyping) them subsequently and separately on several independent subsamples (discrete or cumulative sample approach). After each stage, significant candidates are tested further, until only true associations are likely to be retained. If 1,000 candidates were tested initially, then only 1 false "positive" result is expected after 3 stages of testing, which minimizes the chance of a type 1 statistical error without seriously decreasing the power of the study (Schaid and Sommer 1994, Sham 1994).

Typing hundreds of markers in order to achieve sufficient coverage of the genome for LD studies, apart from the amount of labor and cost of genotyping, raises a serious problem of multiple hypothesis testing. The threshold for the genome-wide (multiple hypothesis testing) significance is set at p=0.05/n, where *n* is the number of independent potential associations checked (Kruglyak and Lander 1995b, Lander and Kruglyak 1995, Morton 1998, Kruglyak 1999). Such correction of statistical significance may cause important findings to be missed, because only extraordinarily strong association findings would remain significant after the correction (Curtis 1996, Witte et al. 1996). Having to deal with the statistical complications and with high false positive rates, have led most researchers to accept consistent replication as the best evidence for a true association. It is therefore recommended by many, that even the findings which did not produce assigned genome-wise significance, and only achieved point-wise significance level should be followed up in a multiple testing manner.

1.324. DNA pooling and case-control association studies

Determination of genotypes at several hundred polymorphic loci in hundreds of individuals is required for mapping complex traits. The idea of using pooled DNA

samples to reduce the burden of labor and cost intensive genotyping was first suggested by Arnheim et al. (1985) in the context of case-control studies. This author argued that alleles in LD with a disease would be enriched (or deficient) in a pooled DNA sample of affected individuals in comparison with a pooled control DNA sample.

Initially, the pooled DNA sample approach has been applied as a genetic mapping tool in isolated populations with reduced allelic diversity, e.g. in mapping the gene for Bardet-Biedl syndrome (Sheffield et al. 1994), cerebellar ataxia (Nystuen et al. 1996), and autosomal recessive non-syndromic hearing loss (Scott et al. 1996). In all the above studies it was expected that affected individuals would be homozygous for a single marker allele at a locus closely linked to the disease gene. Thus the markers were identifiable by visual examination of either silver-stained or radioactively labeled markers.

For complex phenotypes it is inevitable to quantify the marker allele frequencies, since the prevalence of individuals homozygous for a marker allele linked to the disease locus will not reach or be close to 100%. The quantification of marker allele frequencies is only possible by direct genotyping, but can also be estimated from pooled PCR products (Graff et al. 1997). It has been well documented that the allele frequencies estimated from pooled DNA samples show a correlation with allele frequencies obtained by direct genotyping. Estimations were mostly made using GENESCAN software for quantifying allele amplification at polymorphic markers using 5' fluorescently labeled forward PCR primers.

A good correspondence of the PCR products from the pooled samples with those obtained by direct genotyping was also achieved for shorter alleles, even though it is well known that PCR may be biased towards greater efficiency of amplification of shorter, rather than longer, DNA templates. The trend of frequencies from smaller alleles to be overestimated and larger alleles to be underestimated in pooled PCR products is, however, minor and does not appear to significantly affect overall allele frequency estimations (Pacek et al. 1993, Shaw et al. 1998a).

DNA pooling can be efficiently used as an initial searching tool for a candidate map location of susceptibility/protective genes via a genome-wide screen. The method can also be employed to follow up and confirm regions identified in linkage studies or to investigate candidate disease loci. The experimental designs using a pooled DNA approach should also include application of correction methods for stutter artifact and preferential amplification (Barcellos et al. 1997, Daniels et al. 1998).

When initial identification of relevant loci is followed by individual genotyping, the estimation of actual allele frequencies in pooled samples is not crucial. Rather it is important to recognize the variance in allele distributions between applicable DNA pools at a large number of loci. It is expected, that by employing such a research strategy, subtle differences in allele distributions could be missed in the initial screen without any major consequences, since they are not important and will not give rise to statistically significant differences between case and control groups following genotyping of individual DNA samples. The visual examination of radioactively labeled PCR products is therefore expected to be sufficiently powerful to recognize significant differences between applicable DNA pools.

1.325. Linkage disequilibrium mapping

It has long been recognized that classical linkage methods which had been successfully used for mapping genes with major effects have limited power to detect genes of modest effect, which are more likely to be responsible for complex traits (Risch and Merikangas 1996). Gene-mapping efforts have therefore been redirected towards linkage disequilibrium (LD) mapping which relies on the assumption that a single ancestral mutation is responsible for a large proportion of disease cases in a present day population. Such a mutation is considered to have arisen originally in a chromosomal region carrying a particular set of marker alleles – the ancestral haplotype (Kruglyak 1997).

The size of the preserved original ancestral haplotype in the present day population largely depends on the number of generations since the introduction of the mutation

into the population, and the recombination frequency between loci at a particular chromosome region. The detection of identity by descent (IBD) region (ancestral haplotype) among affected individuals on a population level provides a strong evidence for the presence of a relevant disease gene in the region.

Formal analysis of IBD in population samples is usually based on the evaluation of LD – that is, non-random association between individual marker and disease alleles (Service et al 1999). Associations between any flanking marker alleles in general, can be produced by several factors: recent mutation at one of the loci; population founder effect; admixture between populations with different allele frequencies at the loci; selection; or demographic history of the population (Luo 1998). The magnitude of LD is maintained by the recombination frequency between the loci and dissipates more rapidly with physical distance in telomeric regions of the chromosomes than in centromeric regions (Watkins et al. 1994).

In the human genome, LD has been studied mainly in genetic regions surrounding disease genes on affected chromosomes. LD has been successfully applied for the first time in cloning the cystic fibrosis gene (Riordan et al. 1989) and since then widely used for the fine-mapping stage of the localization of disease genes in single founder populations (Devlin and Risch 1995, Jorde 1995, Peterson et al. 1995), because it incorporates information on recombinations that have occurred during the entire period from the mutational event to the present time (de la Chapelle and Wright 1998).

Although there is growing interest in the employment of LD for initial genome-screening studies of complex diseases, the use of the method has been limited until recently to the mapping of rare monogenic diseases in genetic isolates (Friedman et al. 1995, Houwen et al. 1994, Newport et al. 1996). This limitation has mainly been technological, since in most populations, LD extends over very small genetic distances and to identify shared segments or disease associated genetic regions on the population level requires typing of an impractically dense set of genetic markers.

For example, in mixed populations, such as the American population, LD is not expected to exceed ½ cM, in which case at least 6,000 markers would be required to examine disease–marker associations. The exception are young, genetically isolated populations in which LD extends over greater genetic distances (Peltonen 1996). In these populations, LD is expected to stretch a distance of up to 3 cM around disease loci, in which case testing of 1,100 evenly spaced genetic markers is required for a genome-wide LD study.

Today, the problems with utilization of LD are no longer technological. The key issues are analysis and interpretation of large-scale genotyping data. There is a growing consensus among geneticists that full utilization of current and forthcoming technological tools will require a better understanding of population genetics and the distribution of LD across the human genome (Freimer et al. 1997).

It has been speculated that genes for rare monogenic traits are caused by relatively recent mutations, which are more likely detectable in populations with 'star-like' genealogies and so called founder effects. The reverse situation would prevail in the case of alleles involved in complex traits. Such alleles, which are generally common in the population, are likely to be old. According to computer simulation studies, a large area of LD in populations with a constant size will surround such an old mutation, whereas in recently expanded populations, the area will display very little LD (Slatkin 1994, Terwilliger et al. 1998). Consequently, there is a substantial probability of obtaining significant nonrandom associations between closely spaced neutral loci in populations with a constant size (at equilibrium under mutation and genetic drift), while there is a very little chance of finding nonrandom associations even between completely linked loci if the population growth has been sufficiently rapid.

The conclusions achieved by computer simulations have been supported by studies of Laan and Pääbo (1997) in populations of Saami, Swedish, Finnish and Estonian males. Finns and Saami (formerly known as Lapps) are two populations living in close geographical proximity, but with a completely different demographic history.

While the Finnish population expanded rapidly in the last few thousand years, the Saami do not present evidence of any such expansion. The authors established their genotyping results from 7 dinucleotide repeat loci spanning a region of about 4 cM on chromosome Xq13. While several pairs of loci in the Saami contributed to the finding of non-random allelic associations in the region, only one pair of closely spaced loci displayed significant LD level in the Finish population. This pair of loci shows strong association in two admixed populations from Fenno-Scandinavian region (Swedish and Estonian) used for comparison purposes under the same study circumstances (cryptic duplicate markers? - Center for Medical Genetics, Marshfield).

On the contrary, Peterson et al. (1995), who studied background LD distribution in Finns over several regions on chromosome 4, detected LD between a number of loci separated by more than 1 cM. Freimer et al. (1997) argues that the divergent findings of the two studies are probably caused by two factors, which have been proven to influence the ability to detect LD: variable genomic regions and marker heterozygosity. Freimer et al. (1997) also criticized the conclusion, that rapidly expanding populations will not be useful for mapping genes that contribute to disease susceptibility, mainly because of a difference between methods for quantifying background LD and those that search for LD around a disease locus.

1.325. Linkage analysis in nuclear families.

It has long been recognized that combining association and linkage strategies in mapping genes for complex diseases is inevitable (Rutter 1994). The two methods have different advantages and disadvantages. While the strength of linkage studies has been proven by mapping large number of rare monogenic mendelian disorders, the great strength of association studies lies in their power to detect genes of very small effect not detectable by classical linkage methods (a crucial asset when studying a multiple gene disorder).

The power of association studies is, however, compromised by numerous weaknesses. Firstly they are successful only if the marker itself is of pathological importance, or it

is very close to a susceptibility locus, which usually means impractically dense marker maps have to be typed in order to detect genome-wide associations. Secondly, the use of anonymous markers for association studies is limited by the fact, that if linkage exists between an anonymous marker and a disease locus, yet the alleles at both loci are in linkage equilibrium, then an association between them will not be detected (Schaid and Sommer 1993). Third, an important weakness of conventional association studies is a high rate of spurious associations resulting from inadequate matching of cases and controls, especially when there is ethnic heterogeneity within sampled populations.

The development of family-based association methods helped to overcome high rates of false positive association findings by allowing an artificial well-matched control sample to be constructed from marker data generated from family members of probands in the same study (Craddock and Owen 1996).

Rubinstein et al. (1981) and Falk and Rubinstein (1987) recommended using the affected offspring's genotype (made up of alleles transmitted from parents to the affected child) at the marker locus as the 'case' sample, and an artificial genotype made up of the alleles not transmitted to the child from its parents as the 'control' sample in an association test. This method is sometimes called the affected family-based control or haplotype relative risk (HRR) method and can be applied either to genotypes or alleles. Statistically, HRR is a weighted average of two odds ratios: (1) the odds of the presence of a particular marker among either of the two alleles transmitted to the affected child are divided by the odds (2) that the marker is among either of the two non-transmitted alleles (Schaid and Sommer 1993). The validity of HRR methods as tests for association or linkage depends on the structures of the pedigrees (i.e., simplex or multiplex pedigrees) as well as the populations to which they are applied (Spielman and Ewens 1996).

Ott (1989) and Terwilliger and Ott (1992) further defined the statistical properties of the HRR for diseases with a recessive-like mode of inheritance. They developed a haplotype-based haplotype relative risk method (HHRR) in order to make a better use

of all the information present in the nuclear family data via discrimination between parental homozygotes and heterozygotes for a particular allele. Genotype-based HRR (GHRR) of Falk and Rubinstein (1987) does not treat each of the two parental genotypes as independent observations, which makes the HHRR method more powerful for mapping complex traits with recessive-like inheritance. In the case of a dominant disease with reduced penetrance, it is better not to distinguish between parental homo- and heterozygotes for associated alleles. In this case GHRR is more powerful as a gene-mapping tool.

As an alternative to the HRR method of analysis, Spielman et al. (1993) proposed the transmission/disequilibrium testing (TDT) as means to test for linkage in the presence of association. The TDT method evaluates whether the frequency of transmission of alleles from heterozygous parents to their affected children deviates from 50%, the expected mendelian frequency when there is no linkage. The TDT method can be applied to large pedigrees with many affected subjects as well as to a single disease case per family (simplex family), provided that genotypes for both parents of the affected subjects are available (Schaid 1998). However, a simplex family is informative for linkage only when linkage disequilibrium exists – that is, when the likelihood of the coupling and repulsion linkage phases in the parents are not equal (Spielman et al. 1994).

The TDT methods allow computation of *P* values by means of the χ^2 distribution, which is valid only when the number of informative (i.e., heterozygous) parents in the study is large. Although designed as a linkage test, the TDT is also valid as a test of association in simplex families, even in the presence of admixed population structure. The TDT is not valid as a test for association if the families are multiplex (contain affected sibs) or have affected members in multiple generations (Spielman and Ewens 1996). The reason is that any χ^2 test assumes independent observations for the data, and marker data sampled from related affected individuals are not necessarily independent, even when association is absent (i.e., under the null hypothesis for tests of association).

When the TDT is used to test for linkage, it is valid to use all affected subjects who have parental data: The false positive rate for linkage is not inflated, even if subjects are genetically related, because under the null hypothesis of no linkage, the transmission of parental marker alleles follows Mendel's law of transmission, so that sibs are independent (Spielman and Ewens 1996).

Shortly after the introduction of TDT, the intended use was as a test for linkage with a particular marker – for example, at or very near a candidate gene. Since then, the TDT has also been used as a screening test and applied to data from many markers throughout the genome. This is usually associated with a large increase in type I error (false-positive) rate and as a result of that, for each individual marker, a significance level smaller (more extreme) than the nominal level should be required. While for standard lod score analysis or the affected sib-pair method the maximum lod scores are considered statistically significant if > 3 (which approximately translates to P=.0001), the TDT method is often claimed to be statistically significant if P<.05.

Even though stringent, the suggested per-locus significance levels are still achievable when standard lod score analysis, or affected sib-pair methods are used, because there is a high correlation with the evidence for linkage at closely linked markers. In the case of TDT, scores for closely spaced markers do not necessarily exhibit high correlation (Spielman and Ewens 1996). Therefore, in order to achieve a specific genome-wide significance level, the per-locus significance level must be essentially equal to the genome-wide rate divided by the number of marker loci (standard correction for multiple independent tests derived from the Bonferroni inequalities).

In general, TDT has greater power to detect linkage for a genetic trait with a recessive-type inheritance model than for a dominant-type model. Its power is also higher when there is a greater difference in marker allele frequency between disease and normal chromosomes (the presence of association). On the other hand, factors like increased recombination rate, decreased penetrance, recurrent mutation at the marker or at the disease locus, older age of the mutant disease allele, and incomplete

initial LD between marker and disease locus (Xiong and Guo 1998), all negatively influence the power of TDT.

Apart from the above reasons, collecting DNA from parents of affected probands is also useful for the construction of multimarker haplotypes, which can be much more informative for mapping disease genes than to study single markers at a time (Lander and Schork 1994). As was discussed in the previous chapter (1.324), associations are preserved between alleles at tightly linked loci under conditions of LD. This means that LD can be evaluated through identification of conserved haplotypes among affected individuals, rather than with statistical tests of association (Houwen et al. 1994). The detection of IBD-shared haplotypes (chromosome segments) in the case of rare recessive conditions is possible from marker analysis of a handful of disease chromosomes, provided patients are distantly related. In the case of dominant traits, the search for shared chromosome segments is only feasible with a larger sample size, depending on disease penetrance, and disease heterogeneity (either allelic, locus, or etiological).

In order to distinguish haplotypes 'identical-by- state' (IBS) from those 'identical-by-descent' (IBD), dense maps of highly informative markers have to be typed spanning existing regions of interest. The probability of IBD haplotype sharing among affected individuals on the population level at or near non-disease loci is very small. Such sharing may, however, occur as a result of a background kinship – that is, individuals share an IBD haplotype for a different reason than a common affected ancestor. In heterogeneous populations, the extent of background kinship is very small, characterized by a kinship coefficient < .0001. In isolated populations, on the other hand, the kinship coefficient increases dramatically, as a result of higher consanguinity rates, and is often > .01 (Kruglyak 1997).

CHAPTER 2

American Journal of Medical Genetics (Neuropsychiatric Genetics) 60:444-447 (1995)

Increased Chromosomal Breakage in Tourette syndrome Predicts the Possibility of Variable Multiple Gene Involvement in Spectrum Phenotypes: Preliminary Findings and Hypothesis

George S. Gericke*, Ingrid Simonic*, Elma Cloete*, Cecile Buckle*, and Piet J. Becker**

*Department of Human Genetics and Developmental Biology, University of Pretoria, and **Center for Epidemiological Research in Southern Africa, MRC, Pretoria, South Africa

Increased chromosomal breakage was found in 12 patients with DSM-IV Tourette syndrome (TS) as compared with 10 non-TS control individuals with respect to untreated, modified RPMI, and BrdU treated lymphocyte cultures (P<0.001 in each category). A hypothesis is proposed that a major TS gene is probably connected to genetic instability, and associated chromosomal marker sites may be indicative of the localization of secondary genes whose altered expression could be responsible for associated comorbid conditions. This concept implies that genes influencing higher brain functions may be situated at or near highly recombinogenic areas allowing enhanced amplification, duplication and recombination following chromosomal strand breakage. Further studies on a larger sample size are required to confirm the findings relating to chromosomal breakage and to analyze the possible implications for a paradigmatic shift in linkage strategy for complex disorders by focusing on areas at or near unstable chromosomal marker sites.

KEY WORDS: Tourette syndrome, chromosomal instability, spectrum disorder, behavioral genetics, brain evolution

The extent of comorbid behavioral abnormalities described in association with Tourette syndrome (TS) and related tic disorders (Singer and Rosenberg 1989, Trimble 1989, Sverd 1991, Comings and Comings 1993) complicates both diagnostic precision as well as optimal research design as required for demonstration of linkage or significant association with a defined marker.

While most pedigree analyses predict TS to be due to a single autosomal gene disorder (Baron et al. 1981, Kidd and Pauls 1982, Comings et al. 1984, Devor 1984, Pauls and Leckman 1986) almost the whole autosomal genome has now been screened without linkage being demonstrated (Heutink et al. 1993). The possibility of heterogeneity has been invoked to explain the inability to reproduce linkage findings.

Both linkage and association studies have failed to yield the expected insight into many complex neurobehavioral disorders (Crowe 1993, Kidd 1993).

In this paper we suggest another approach based on the finding of increased spontaneous and chemically induced chromosomal instability in TS patients which could possibly supply marker sites for identification of the genes involved in the heterogeneous presentation of this, and possibly other neurobehavioral spectrum disorders.

The study was initiated because of the finding of increased fragility observed during a fra-X cytogenetic workup of a retarded nondysmorphic male with tics where 14 percent of metaphases showed chromosome breakage at 3p13 and expression of an Xq26 fragile site; and the presence of a 46,XY, t(7;12)(q22q24.1) karyotype in another mildly retarded child with tics and self-injurious behavior. The breakpoint at 7q22 in this patient was similar to that described in several relatives in a TS kindred: 46,XX/XY, t(7;18)(q22;q22.1) by Comings et al. (1986).

Spontaneous, modified RPMI (modRPMI)(Irvine Scientific) and BrdU(Sigma) inducible breaks (Sutherland and Ledbetter 1989, Sutherland 1991) were studied in 12 males with DSMIV Tourette syndrome (7children, 5 adults) and 10 unaffected junior male students. Blood samples were collected after informed consent was obtained.

For patient and control subjects three lymphocyte cultures were set up simultaneously. Investigator who treated the cultures, and analyzed all slides, was blind to the affection status of the individual cases. RPMI 1640 (Highveld Biological) medium supplemented with 10% fetal calf serum (SteriLab) was used in two cultures, and mod RPMI (folic acid and thymidine deficient medium) with the fetal calf serum supplement reduced to 5% was used to establish a third culture. Cultures were maintained for 72 hours. 40mg/L BrdU was added for the last 16 h to one of the cultures growing in RPMI 1640 medium. Colcemid (0.2μg/ml) was added to all cultures 35 min prior to harvesting. Chromosome spreads were obtained according to the standard procedure: 0.075M KCl for 10 min, methanol:acetic acid 3:1 fixation. For the analysis of breakage rate and the location of breakpoints, 100 G-banded metaphases at 400-band level from the spontaneous (untreated) cultures and 50 metaphases from mod RPMI and BrdU-treated cultures were scored for each individual. Chromatid and/or chromosome breaks were assigned as a single event; the few chromatid type of exchanges were considered as double events. After testing for equality of variance in the TS and control groups, these groups were compared with respect to the mean number of aberrations per cell using the appropriate two-sided Student t-test at the 0.05 level of significance. In view of the small sample size, the results from the t-test were confirmed employing non-paramatric Mann-Whitney test.

The only specific rare folate-sensitive fragile site induced in TS cells was fra(10)(q25). For each of the 3 categories-untreated, mod RPMI-, and BrdU-treated categories – the TS and control groups differed significantly ($P<0.001$ in each category) with respect to the mean number of aberrations per cell (Table 1). These results indicate that TS cells manifest chromosomal instability in this preliminary study.

Although an analysis of the relevance of particular breakpoints has not been possible due to the small sample size, breakage sites observed in our patients may be of pathogenetic importance because some of them appear to be at areas previously implicated in linkage studies for neurobehavioral disorders. Some interesting differences in this preliminary study were fragility at 22q12-q13 in all TS patients and

none in the controls; this area has recently been considered with regard to potential linkage with schizophrenia (Pulver et al. 1994); involvement of 18q22 previously considered to be associated with TS (Comings et al. 1986; Donnai 1987) and Xp22/Yq11-12 homologous areas implicated frequently in TS patients indicating a possible involvement of pseudo-autosomal genes.

TABLE 1

Mean number of aberrations per cell for the TS and control groups

		Mean (sd) no. of aberrations/cell		P-values	
Treatment Category	No. of cells analyzed per patient	Patients (n=12)	Controls (n=10)	t-test	Mann-Whitney test
Untreated	100	0.121 (0.034)	0.052 (0.026)	<0.001	<0.001
Mod RPMI	50	0.255 (0.074)	0.086 (0.044)	<0.001	<0.001
BrdU	50	0.163 (0.050)	0.092 (0.033)	<0.001	=0.0024

In addition one patient with TS and 47, XXY with paternal pseudoautosomal nondisjunction as a possible mechanism being responsible for the 47, XXY karyotype and two with fragility at Xq27.3 in association with TS were observed before initiating the study protocol as reported here.

Barletta et al. (1991) reported the association of other fragile sites in three families with the fragile X-syndrome and considered the existence of some underlying genetic mechanism in both autosomal fragile sites and fragile X-syndrome. The degree of fragility in obligate female carriers of the fragile X chromosome was found to be a potentially important predictor of psychopathology among women with normal IQ (Freund et al. 1992). Women with positive fragility scored highest on the "odd

communication" and "peculiar mannerisms" (?tics) items. Increased fragility was found to be associated with schizotypal features. The authors stated that a need exists to clarify the nature of the association between the clinical neuropsychiatric phenotype and fragile site expression.

Garofalo et al. (1992) found increased fragility in association with schizophrenia and considered the possibility that this finding may be useful for the search for the location of major genes for that disease. The existence of significant overlap between schizophrenia and severe TS has been documented (Comings 1990).

Several investigators have published reports of an association between TS and chromosomal abnormalities. At the Sixth Genetic Workshop on TS in the Netherlands in 1990, case studies were reported with TS like symptoms and changes in chromosomes 3, 8, 9, and 13. TS phenotype co-segregating in two instances with an 18q22.1 abnormality was reported by Comings et al. (1986) and Donnai (1987). Co-occurrence of TS with FRAXA (Kerbeshian et al. 1984), trisomy 21 (Barabas et al. 1986), karyotype 47, XYY (Marskey, 1974), 47, XXX and 9p monosomy (Singh et al. 1982), and 9p monosomy alone (Taylor et al. 1991) were also reported.

The high prevalence for tic related disorders could indicate an evolutionary advantage, and a proportion of any population would harbor this "instability" phenomenon and benefit from it by increased recombination, preventing inbreeding depression of genes coding for higher brain functions. In this regard, the term "recombination-enhancing areas" may be more appropriate than the negative connotation ascribed to "instability". Models have been proposed (Windle et al. 1991, Kimmel et al. 1992) indicating the possible importance of chromosome breakage and subsequent gene deletion resulting from resection of broken chromosome ends, as initial events in gene amplification. Are we witnessing genetic changes relevant to the evolution of a more complex nervous system? Although instability makes the genome more sensitive to environmental input, this may, however, render it more vulnerable to injury, for which certain mental disorders could represent the penalty.

Phenotypic heterogeneity could result from environmentally precipitated genetic instability in susceptible individuals, varying degrees of which might predict severity of behavioral dysfunction in TS gene carriers.

Over time, a complex and expanded list of behavioral and medical problems could presumably result from the same underlying phenomenon. In this regard, it would be interesting to consider the possibility of a fundamental interrelationship between TS and other conditions such as fra-X and auto-immune problems, all recorded as problems with an increased prevalence in inbred old Colony Mennonites (Jaworski et al. 1989). This may be informative for the study of possible pathogenic relationships between such conditions in larger outbreed populations.

TS appears to represent a condition where more comorbid disorders cosegregate, and as proposed here, more genes be differently expressed than would be expected in circumscribed psychiatric syndromes. If confirmed, this phenomenon may allow multiple gene involvement in TS to be used as an entry to identification of major genes for other disorders not manifesting as spectrum conditions, but which are usually expressed as comorbid disorders in TS pedigrees.

The need for a heuristic model in complex multifactorial neurobehavioral disorders is evident. In this regard, comments concerning studies on alcoholism (Devor 1993) may be considered relevant for the study of TS as well; namely, that a paradigmatic shift in linkage strategy away from the search for *the* gene should be redirected towards finding regions of interest that can be more fully explored by other molecular techniques. It is not (yet) possible to make this assumption, and for this reason a further study on a larger patient sample is in progress.

ACKNOWLEDGMENTS

The authors are grateful to Professor Grant Sutherland from Adelaide, Australia, for helpful discussions on cytogenetic methods during his academic visit to our Unit; to Professor Andries Retief from the University of Stellenbosch; Dr. Stander Jansen and

Ms. Analise Hillebrandt from the University of the Orange Free State for assistance with controlling the chromosomal fragility techniques performed at the Pretoria laboratory; and to Susann Steyn and Santie Swarts for expert secretarial assistance and typing the manuscript.

REFERENCES

Barabas G, Wardell B, Shapiro M, Mathews WS (1986) Coincident Down's and Tourette syndromes: Three case reports. J Child Neurol 1:358-360

Barletta C, Ragua RM, Garofalo G, Scilatto F, Ruggeri M (1991) segregation analysis of autosomal fragile sites in three families with fragile X chromosome. Ann Genet 34:111-114

Baron M, Shapiro ES, Shapiro AK, Rainer JD (1981) Genetic analysis of Tourette syndrome suggesting major gene defect. Am J Hum Genet 36:767-775

Comings DE, Comings BG, Dietz G, Muhleman D, Okada TA, Sarinana F, Simmer R, Stock D (1986) Evidence the Tourette syndrome gene is at 18q22.1. Seventh International Congress of Human Genetics, West Berlin, part II:620 (Abstr)

Comings DE, Comings BG, Devor EJ, Cloninger CR (1984) Detection of a major gene for Gilles de la Tourette syndrome. Am J Hum Genet 36:586-600

Comings DE (1990) "Tourette syndrome and human behavior." California: Hope Press, pp 205-214

Comings DE, Comings BG (1993) Comorbid behavioral disorders. In Kurlan R (ed) "Handbook of Tourette's syndrome and related tic and behavioral disorders." New York: Marcel Dekker Inc., pp 111-147

Crowe RR (1993) Candidate genes in psychiatry: An epidemiological perspective. Am J Med Genet 48:74-77

Devor EJ (1984) Complex segregation analysis of Gilles de la Tourette syndrome: Further evidence for a major locus mode of transmission. Am J Hum Genet 36: 704-709

Devor EJ (1993) Why there is no gene for alcoholism. Behav Genet 23:145-151

Donnai D (1987) gene localization in Tourette syndrome. Lancet 1:627

Freund LS, Reiss AL, Hagerman R, Vinogradov S (1992) Chromosome fragility and psychopathology in obligate female carriers of the fragile X-chromosome. Arch Gen Psychiatry 49:54-60

Garofalo G, Ragusa RM, Barletta C, Spina E (1992) Schizophrenia and chromosomal fragile sites. Am J Psychiatry 149:1116

Heutink P, Breedveld GJ, Niermeijer MF, van de Wetering BJM, Oostra BA (1993). Progress in gene localization. In: Kurlan R (ed) "Handbook of Tourette's syndrome and related tic and behavioral disorders." New York: Marcel Dekker Inc., pp 317-333

Jaworski MA, Severini A, Mansour G, Hennig K, Slater JD, Jeske R, Schlant J, Yoon JW, Maclaren NK, Nepom GT (1989) Inherited diseases in North American Mennonites: Focus on old colony (Chortitza) Mennonites. Am J Med Genet 32:158-168

Kerbeshian J, Burd L, Matsolf J, (1984) Fragile X-syndrome associated with Tourette symptomatology in a male with moderate mental retardation and autism. J Dev Behav Pediatr 5:201-203

Kidd KK, Pauls DL (1982) Genetic hypotheses for Tourette syndrome. In: Friedhoff AJ, Chase TN (eds) "Advances in neurology." New York: Raven Press, Vol 35, pp 243-249

Kidd KK (1993) Associations of disease with genetic markers: Déjà vu all over again. Am J Med Genet 48:71-73

Kimmel M, Axelrod DE, Sahe GM (1992) A branching process of gene amplification following chromosome breakage. Mutat Res 276:225-239

Marskey H (1974) A case of multiple tics with vocalization (partial syndrome of Gilles de la Tourette) and XYY karyotype. Br J Psychiatr 125:593-594

Pato CN, Pato MT, Kennedy JL(1993) Summary of the Third World Congress on Psychiatric Genetics. Am J Med Genet 48:234-237

Pauls DL, Leckman JF (1986) The inheritance of Gilles de la Tourette's syndrome and associated behaviors: Evidence for autosomal dominant transmission. New Engl J Med 315:993-997

Pulver AE, Karayiorgou M, Wolniec PS, Lasseter VK, Kasch L, Nestadt G, Antonorakis S, Houseman D, Kazazian HH, Meyers D, Ott J, Lamacz M, Liang KY, Hanfelt J, Ullrich G, Dimarchi N, Ramu E, McHugh PM, Adler L, Thomas M, Carpenter WT, Manschreck T, Gordon CT, Kimberland M, Babb R, Puck J, Childs B (1994) Sequential strategy to identify a susceptibility gene for schizophrenia: Report of potential linkage on chromosome 22q12-q13.1: Part 1. Am J Med Genet 54:36-43

Singer HS, Rosenberg LA (1989): The development of behavioral and emotional problems in Tourette syndrome. Pediatr Neurol 5:41-44

Singh DN, Howe GL, Jordan HW, Hara X (1982) Tourette's syndrome in a black woman with associated triple X and 9p mosaicism. J Natl Med Assoc 74:675-682

Sutherland GR, Ledbetter DH (1989) Report of the Committee on Cytogenetic Markers (HGM 10). Cytogenet Cell Genet 51:452-458

Sutherland GR (1991) The detection of fragile sites on human chromosomes. In Adolph KW (ed) "Edvanced Techniques in Human Chromosome Research." New York: Marcel Dekker, pp 203-222

Sverd J (1991) Tourette's syndrome and autistic disorder: a significant relationship. Am J Med Genet 39:173-179

Taylor LD, Krizman DB, Jankovic J, Hayani A, Steuber PC, Greenberg F, Fenwick RG, Caskey CT (1991) 9p monosomy in a patient with Gilles de la Tourette's syndrome. Neurology 41:1513-1515

Trimble M (1989) Psychopathology and movement disorders: a new perspective on the Gilles de la Tourette syndrome. J Neurol Neurosurg Psychiat (Suppl) 52:90-95

Windle B, Draper BW, Yin Y, O'Gorman S, Wahe GM (1991) A central role for chromosome breakage in gene amplification and amplicon integration. Genes Dev 5:160-174

CHAPTER 3

American Journal of Medical Genetics (Neuropsychiatric Genetics) 67:25-30 (1996)

Increased Expression of Aphidicolin-Induced Common Fragile Sites in Tourette Syndrome: The Key to Understand the Genetics of Comorbid Phenotypes?

George S. Gericke*, Ingrid Simonic*, Elma Cloete*, and Piet J. Becker**

* Department of Human Genetics and Developmetal Biology, University of Pretoria, and
** Center for Epidemiological Research in Southern Africa (CERSA), Medical Research Council, Pretoria, South Africa

In a comparison of 80 common aphidicolin-inducible fragile sites (FS) between 26 DSM-IV Tourette syndrome (TS) and 24 control individuals, the mean of the summed break frequencies following mild aphidicolin pretreatment was significantly higher in TS individuals than in controls (P<0.001). Other breakpoints encountered during this study, i.e., random breaks, breaks corresponding to rare FS, and breakpoints recorded by others but not listed as common FS according to the Chromosome Coordinating Meeting (1992) were listed as category II breakpoints. By using the most significantly different mean FS breakage figures between TS and control individuals, further stepwise discriminant analysis allowed identification of TS individuals from only a few sites in both the common FS and category II breakpoint groups. Future research needs to focus on confirmation of altered common fragile site expression in association with behavioral variation, whether expression of certain discriminatory sites concurs with specific comorbid disorder expression; the nature of the molecular alterations at these FS and the implications of a genomic instability phenotype for the mapping of a primary TS gene or genes.

Key words: Tourette syndrome, aphidicolin, common fragile sites, gene mapping, tandem repeat sequences

Introduction

Tourette syndrome (TS) is a common, heritable tic disorder associated with obsessive compulsive behavior and attention deficit-hyperactivity which are regarded as integral components of the phenotype (Kurlan et al. 1994), while the relationship to a wide range of additional reported comorbid problems (Comings and Comings 1993) remains controversial. Most research findings indicate that TS is due to the effect of a single major autosomal gene (Baron et al. 1981, Kidd and Pauls 1982, Comings et al. 1984, Devor 1984, Pauls and Leckman 1986), which is supported by genealogical research indicating a gene founder effect in South Africa (Torrington and Gericke, unpublished data). Alternatively, the complex spectrum of manifestations suggested to others the possibility of polygenic inheritance (Comings 1994).

The recent observation of increased chromosomal breakage in TS has been proposed as a basis for the consideration of variable multiple secondary gene involvement at chromosomal fragile sites in complex behavioral disorders (Gericke et al. 1995).

Fragile sites (FS) are nonrandom heritable sites on chromosomes that can be induced to form gaps, breaks, and rearrangements under specific conditions (Jordan et al. 1990). Both rare and common FS are expressed in culture under conditions which inhibit DNA synthesis. FS may represent "active genomic sites that are vulnerable to physiological and environmental disturbance" (Yunis and Hoffman 1989). Because the different classes of FS reveal cross induction (Yunis et al. 1987, Hecht et al. 1988), it was suggested that FS may be indicative of areas of shared molecular homology in the sequence composition of nonrandom chromosomal DNA (Stopera 1989), and that such sites can be general targets of mutagenic action.

Since aphidicolin-induced common FS have been postulated to be of pathologic importance (Hecht 1991), we decided to investigate whether increased fragility of these sites are present in TS individuals, and whether any specific aphidicolin-induced

breakpoints could be demonstrated to be characteristically associated with TS clinical diagnosis.

Material and methods

Five to ten milliliters of heparinized blood for chromosome culturing was collected from individuals after informed consent was obtained.

Patients

This group of young adult Caucasians included 15 males and 11 females. All of these 26 individuals had mild, though clearly recognizable motor and/or vocal tics, but were moderately to severely handicapped by associated obsessive-compulsive disorder, residual learning problems, or conduct disorders. Nineteen relatives from 6 families were included in this group. They were diagnosed as having TS by means of DSM-IV criteria, had positive family histories of a tic disorder in all instances, and were selected according to whether their clinic visits coincided with days that were convenient to the laboratory, whether they gave informed consent to participate in the study and whether matched controls were available. The choice of this age cohort allowed a clear retrospective evaluation of the longitudinal patient histories and permitted exclusion of other potentially confounding psychiatric or neurologic differential diagnoses.

Controls

Twenty-four randomly selected emotionally stable Caucasian medical students volunteers consisting of 14 males and 10 females. There were no significant differences in the mean age between patients and controls, and in the mean age of males and females within the groups.

Procedure

Coded blood specimens from both patients and controls arrived simultaneously to the laboratory in batches of varying sizes. Patient and control specimens were thus handled together, and culture preparation and analysis of slides were performed by a single investigator who had no knowledge of the clinical status of the individuals being tested.

Cell cultures

Phytohemagglutinin (PHA) stimulated blood lymphocytes were cultured in RPMI 1640 medium (Highveld) supplemented with 10% fetal calf serum (GIBCO) for 72 hours at 37°C. Twenty-four hours prior to harvesting, aphidicolin dissolved in 70% ethanol (MERCK) was added to each 10 ml culture to obtain a final concentration of 0.1μM. Colcemid (0.2 ml)(SIGMA) was added for the last 35 minutes. Cells were then treated with a hypotonic 0.075 M KCl solution for 12 minutes and fixed in methanol : acetic acid fixative. Slides were made, air dried and GTG banded.

Analysis

Three to four slides from each individual were analyzed, and for each patient one hundred complete metaphases were assessed at 350-400 band level. Aberrations (gaps and breaks) were counted as single event at the band(s) involved. In some metaphases, it was not possible to identify the individual chromosomes and chromosome bands because of the extent of breakage. Such pulverized cells were not included in the analysis.

All events on the band were recorded, i.e. chromatid gap or break, chromosome gap or break, or chromatid exchange. In case a gap or break occurred on both homologues in the same band, it was counted as two events.

Frequencies of 80 common aphidicolin-inducible FS were evaluated according to criteria formulated by the Chromosome Coordinating Meeting (CCM) (1992). (Table

1) All other breakpoints, including rare FS, breaks reported by others but not listed by the CCM 1992 as common fragile sites, and all breakpoints regarded as random, were listed separately. (Table 2) These sites included only those breakpoints from the group, which occurred in most of the individuals in low frequencies, or which occurred in only some of the samples, but in higher frequencies, comparable to those of common FS, in both TS and non-TS groups.

Results

The differences in the mean number of aberrations per cell between TS and controls for known aphidicolin-inducible FS, and for breakpoints, not listed as common FS by CCM 1992 are shown in Tables 1 and 2.

In order to develop a classification function for characterization of TS patients, a stepwise discriminant analysis was utilized. The sample size was not adequate to include all the sites into the analysis. Hence to decide which sites to include, the two groups were compared with respect to the mean percentage of breaks at each site, using the appropriate t-test after comparing the groups for equal variance using Levene's test. Those sites for which the two groups were significantly different ($P<0.05$) with regard to the mean percentage of breakages, and who additionally displayed mean differences >5% between patients and controls were subsequently included into a further stepwise discriminant analysis.

Standard Aphidicolin-Inducible Sites (Table1)

From the sites, 1p21, 1q44, 3p14, 6q26, 7p22, 7q31, 7q32, 11p14, 11q14, 16q23, 22q12, Xp22, and Xq22, utilization of stepwise discriminant analysis indicated optimum standardized coefficients for the sites 1q44, 16q23, and 22q12 which were used in the classification function:

$$\text{TS:} \quad y = -16.922 + 0.934(1q44) + 0.395(16q23) + 0.321(22q12)$$

Non-TS: $y = -5.085 + 0.295(1q44) + 0.245(16q23) + 0.133(22q12)$

This could be simplified to the single function: 0.639(1q44) + 0.150(16q23) + 0.188(22q12). With all controls scoring less than 11.837, this function indicated 100% specificity (all controls can be assigned to non-TS group by means of these three FS expression frequencies), and 88% sensitivity (88% of clinically diagnosed TS individuals can be assigned to TS group by means of above three FS expression frequencies/cell) under these study circumstances.

Category II Sites (Table 2) (Fig. 1)

It was decided to keep the analyses separate for the two categories of sites defined in this article. The overall difference in category II sites between TS and control individuals were significant (P=0.001). By including sites 3p26, 3q13, 10p11.2, 12q12-q13, 14q12-q13, and Xq13, utilization of stepwise discriminant analysis indicated optimum standardized coefficients for the sites 10p11.2, 14q13, and 3q13. Employing these sites in an analysis similar to the one performed for group I FS, 73.1% of clinically diagnosed TS patients were assigned to the affected group by means of FS expression (sensitivity), as was the case for 79.2% of controls (specificity).

The classification function for category II sites was:

TS: $y = -5.810 + 0.888(10p11.2) + 1.184(14q12\text{-}q13) + 0.766(3q13)$

Non-TS: $y = -1.812 + 0.346(10p11.2) + 0.623(14q12\text{-}q13) + 0.336(3q13)$

This could be reduced to the single classification function:

$$0.542(10p11.2) + 0.561(14q12\text{-}q13) + 0.430(3q13),$$

meaning that an individual could be classified as TS if the observed value of this function for above three sites was more than 4.

Table 1

Mean (SD) percentage of common aphidicolin-inducible breaks at different sites for TS and control individuals

Site	TS Mean	TS SD	Control Mean	Control SD	t-test	Site	TS Mean	TS SD	Control Mean	Control SD	t-test
1p36	2.69	1.57	1.42	1.25	0.003	**7q21**	4.31	2.15	2.42	1.72	0.001
1p32	5.00	2.95	2.83	1.61	0.002	**7q22**	2.85	1.89	1.33	1.24	0.002
1p31	4.65	3.15	2.00	1.22	0.001	**7q31**	12.62	4.93	6.83	2.97	0.000
1p22	0.77	1.07	0.75	1.03	0.949	**7q32**	17.38	5.26	9.29	3.22	0.000
1p21	14.42	5.12	7.83	3.20	0.000	**7q36**	1.35	1.13	0.46	0.59	0.001
1q21	1.42	1.33	0.67	0.92	0.025	**8q22**	5.38	2.97	3.29	1.99	0.005
1q25	6.19	3.15	2.63	1.79	0.000	**8q24.1**	3.08	2.28	1.88	1.73	0.042
1q31	1.50	1.24	0.75	0.85	0.016	**8q24.3**	2.15	1.54	1.13	0.99	0.007
1q42	0.73	1.28	0.67	0.96	0.844	**9p21**	1.31	1.38	0.54	0.78	0.019
1q44	9.85	0.67	3.96	2.03	0.000	**9q12**	0.85	0.88	0.54	1.18	0.303
2p24	9.96	4.27	6.04	2.97	0.001	**9q22**	3.27	1.89	2.08	1.47	0.017
2p16	6.77	2.57	3.92	1.95	0.000	**9q32**	12.73	4.88	8.04	3.26	0.000
2p13	6.96	2.86	3.50	1.79	0.000	**10q21**	0.73	0.96	0.46	0.78	0.279
2q21	6.58	2.91	2.71	1.78	0.000	**10q22**	3.00	1.96	1.58	1.32	0.005
2q31	5.96	4.15	2.04	1.70	0.000	**10q25**	3.08	2.33	1.25	1.07	0.001
2q32	7.77	3.66	5.79	2.20	0.025	**10q26**	6.50	2.94	4.29	2.89	0.010
2q33	5.69	3.21	2.21	1.47	0.000	**11p15**	3.62	1.94	2.42	1.18	0.011
2q37	8.12	3.70	4.42	2.70	0.000	**11p14**	10.38	4.01	5.08	2.93	0.000
3p24	7.19	4.07	4.67	2.58	0.012	**11p13**	9.27	3.64	5.92	3.60	0.002
3p14	58.50	13.64	38.42	9.10	0.000	**11q14**	10.19	4.38	5.75	2.66	0.000
3q25	4.00	2.48	2.71	2.81	0.091	**11q23**	0.85	0.83	0.54	0.72	0.175
3q27	3.73	2.22	2.25	1.65	0.011	**12q21**	4.00	2.51	2.13	1.60	0.003
4p16	2.85	1.85	2.21	1.35	0.173	**12q24**	2.65	1.92	1.92	1.35	0.121
4p15	2.92	1.90	2.46	1.64	0.361	**13q13**	8.50	3.71	4.29	2.37	0.000
4q21	3.19	2.47	1.96	1.71	0.047	**13q21**	1.35	1.44	1.17	1.37	0.655
4q31	12.19	3.90	8.17	3.13	0.000	**13q32**	2.65	1.72	1.21	1.25	0.002
5q31	2.04	1.40	0.96	0.86	0.002	**16q23**	49.62	10.18	27.92	9.49	0.000
6p25	6.62	3.58	3.42	1.74	0.000	**17q23**	2.88	1.99	0.63	0.65	0.000
6p22	1.08	1.23	0.67	0.92	0.191	**18q12**	5.85	2.65	3.08	1.77	0.000

Table 1 Continued

Site	TS Mean	TS SD	Control Mean	Control SD	t-test	Site	TS Mean	TS SD	Control Mean	Control SD	t-test
6q15	1.58	1.45	1.46	1.06	0.744	**18q21**	1.65	1.47	1.08	1.25	0.147
6q21	3.23	1.75	2.29	1.57	0.052	**19q13.1**	1.50	1.50	1.17	1.43	0.427
6q26	12.15	4.86	7.13	2.91	0.000	**20p12**	5.73	3.21	3.33	2.04	0.003
7p22	9.23	3.87	4.13	2.89	0.000	**22q12**	10.48	4.35	5.83	3.38	0.000
7p14	1.96	1.48	0.54	0.83	0.000	**Xp22.3**	23.00	8.60	14.96	6.84	0.001
7p13	11.85	4.22	6.04	3.64	0.000	**Xq22**	15.46	6.41	9.22	3.68	0.000
7q11	4.62	2.14	3.00	1.82	0.006	**Xq27**	2.27	1.97	1.21	1.28	0.028

Table 2

Mean (SD) percentage of other breaks for TS patients and controls

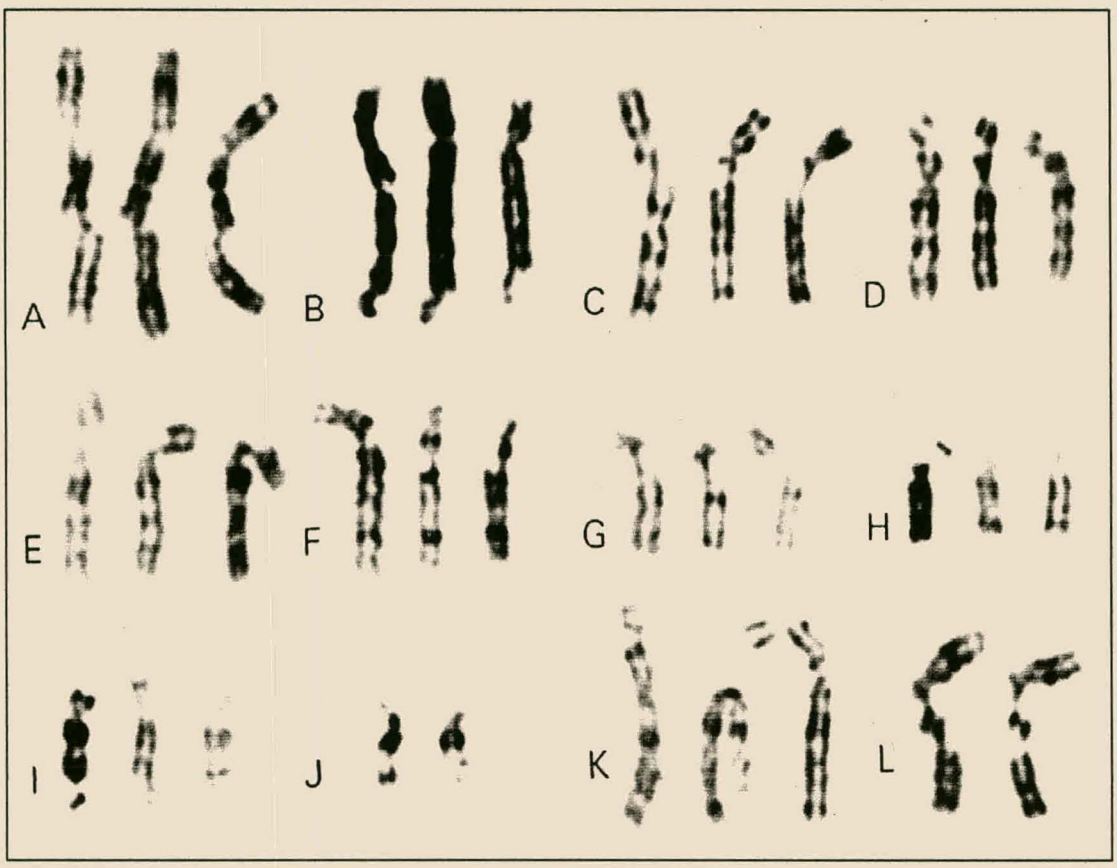
Site	TS Mean	TS SD	Control Mean	Control SD	t-test	Site	TS Mean	TS SD	Control Mean	Control SD	t-test
3p26[a]	2.42	2.32	1.21	1.32	0.027	**13q34**[a]	0.42	0.81	0.46	0.66	0.867
3q13[a]	3.73	2.44	1.63	1.17	0	**14q13**	3.92	1.94	2	1.38	0
4q23[a]	0.65	0.94	0.25	0.53	0.065	**15q13-q14**	0.5	0.65	0.67	0.82	0.426
4q27[a]	0.69	1.09	0.5	0.72	0.469	**18p11.2**	1.15	1.35	0.96	1.27	0.6
4q34-q35	1	1.26	0.63	1.01	0.256	**18q22**	0.96	1.08	0.46	0.88	0.079
5q13	1.38	1.27	0.67	0.82	0.022	**20q11.2**	0.73	0.78	0.5	0.72	0.283
5q33-q34	2.15	1.8	1.42	1.5	0.125	**21q22**	0.92	0.8	0.46	0.66	0.03
8p21	0.62	0.75	0.33	0.56	0.139	**22q13**[b]	0.42	0.75	0.08	0.28	0.041
8q11.2	0.35	1.02	0.46	0.98	0.693	**Xp22.1**	0.96	1.15	0.29	0.55	0.012
10p13	1.19	1.3	1.96	1.49	0.583	**Xq13**	2.08	1.38	0.96	1.04	0.002
10p11.2	3.08	1.72	1.29	1.27	0	**Xq24**	0.58	0.76	0.29	0.69	0.172
10q11.2	0.12	0.33	0.21	0.41	0.381	**Xq26**	0.62	0.9	0.5	0.72	0.621
12p12	0.77	0.82	0.75	0.79	0.933	**Xq27.3**[b]	0.04	0.2	0.13	0.34	0.28
12q13[b]	1.23	1.21	0.38	0.65	0.003						

[a] Breakpoints detected in our study and reported by others, but not listed in CCM92

[b] Breakpoints considered as rare folate-sensitive FS

Figure 1

Examples of some category II breaks in TS patients: A. 3q13; B. 4q33-q34; C. 5q13; D. 10p13; E. 10p11.2; F. 14q13; G. 15q13-q14; H. 18p11.2-p11.3; I. 18q22; J. 21q22; L. Xq13.



In six instances, similarly affected family members of index cases with TS formed part of the study. The sites recurring within these families include:

- family 1(a male and a female): 15q15 in 2/2 family members;
- family 2 (three females and a male): 3q21 in 4/4 cases and 6q23 in 3/4 cases;
- family 3 (three males and a female): 17p12 in 3/4 cases and Yq11-q12 in 2/3 males;
- family 4 (three males and a female): 17p12 in 4/4 cases and Yq11-q12 in all males;
- family 5(a male and two females): 8p23 in 2/3 cases.

In family 6, in two females, there was no concordance with regard to any FS expression. Yq11-q12 fragile site was expressed in 7/8 males from the above family groups. Some of these breaks are not reflected in Table I or II because they do not qualify for inclusion according to the criteria mentioned in the methods section of this paper.

Discussion

This paper reports increased aphidicolin induced common fragile sites (FS) expression associated with Tourette syndrome (TS), as well as the possible existence of discriminatory FS with regard to TS patients. The particular sites found to be indicative of TS during this early study may be altered during subsequent research.

Confirmation of increased aphidicolin-induced FS expression in TS needs to take into account the large number of conditions which may influence chromosome breakage (Craig-Holmes et al. 1987, Chudley et al. 1990, Smeets and Merkx 1990, Tedeschi et al. 1992), including tissue specificity of FS expression (Morgan et al. 1988, Murano et al. 1989), culture conditions such as ethanol concentration (Kuwano and Kajii 1987), menstrual stage cycle in females (Furuya et al. 1991), and knowledge of the population background with regard to common aphidicolin-induced FS (Rao et al. 1988). Although our study could be criticized for not adequately evaluating the role of such variables, the differences between TS and control individuals were consistently distinct to such a marked degree that it was considered worthwhile to open this avenue of research to the widest possible scrutiny.

A major question arises whether increased expression of aphidicolin-induced FS in individuals with TS could be potentially informative for the molecular analysis of genes involved in the expression of a TS spectrum disorder. The future isolation and cloning of the DNA sequences involved at some of these sites might show that they are polymorphic regions and targeted association studies could potentially be applied to examine their role in TS as well as other neuropsychiatric disorders.

A number of reports correlating behavioral alteration with chromosomal breakage have already been published and include schizophrenia (DeLisi et al. 1988, Garofalo et al. 1992), the psychopathology found in obligate fra-X female carriers (Freund et al. 1992), and Rett syndrome (Telvi et al. 1994).

To our knowledge, the only published finding relating chromosomal fragility to TS by another author can be found in an article on Huntington disease and childhood-onset Tourette syndrome (Kerbeshian et al. 1991), where the karyotype 46, XY / 46, XY, fra(16)(q22) was found in 12% of cells. The authors stated that the clinical significance of FS in phenotypically normal individuals is unknown.

In addition to being considered genetically active areas (Yunis et al. 1987, Hecht et al. 1988, Hecht and Hecht 1991, Austin et al. 1992), FS may represent regions of DNA repeat sequences (Sutherland et al. 1985). Organized repetitive DNA sequences in the genome are considered to bear a relationship to a highly conserved chromatin folding code (Vogt 1990) and may predispose these areas to selective forces such as environmentally induced breakage.

Individuals expressing the rare folate-sensitive sites FRAXA and FRAXE have unstable expanded CCG repeats and methylation of adjacent CpG islands. An explanation has been proposed according to which GCC and CGG repeat sequences lead to delayed replication because they form unusual DNA structures that present a block for the replication apparatus (Knight et al. 1993). Chromatin is subsequently rendered fragile through late replication following failure to erase an imprinted X-inactivation signal in these instances (Laird et al. 1987).

In addition, the molecular basis of FRA16A, another autosomally located, rare folate sensitive fragile site, was found to be expansion of a normally polymorphic p(CCG)n repeat (Richards et al. 1994). It is not clear whether one is allowed to extrapolate from these findings concerning rare fragile sites. Molecular analysis of a 4.5kB fragment containing 6 of 13 aphidicolin-induced breakpoints at constitutive fragile site 3p14.2 failed to identify traditional motifs, such as a trinucleotide repeat sequence to explain fragility (Paradee et al. 1994). However, in a cell line with a reciprocal translocation

between human chromosome 3 (with breakpoint at 3p14.2) and a hamster chromosome, the fragile site was expressed on both derivative chromosomes, suggesting in this instance that the fragile site represents a repeated sequence (Glover and Stein 1988).

If multiple genes, acting either independently (genetic heterogeneity), additively (polygenic inheritance), or epistatically are required for expression of a neurobehavioral spectrum phenotype, this will create obstacles for linkage detection of a major gene by the standard lod score method (Crowe 1993). Nonparametric linkage analyses, such as the affected sib-pair method, which is useful for the analysis of multifactorial disorders, and which was suggested as the next step in TS-gene linkage at the Tenth Genetic Workshop on Tourette syndrome held during August 1994 in Toronto (personal communication, J. Weber) is less powerful to detect linkage and cannot satisfactorily address a situation where multiple genes may be operative, which is what our fragility studies could suggest.

For TS, a dual situation may exist, according to which various associated phenotypes may arise through modified gene activity at fragile sites ("the component factors of multidimensional phenotypes")(Cloninger 1994), but the primary phenotype for which the TS gene is responsible may actually be chromosomal instability.

A heritability estimate of 0.88 for aphidicolin-inducible common FS expression in a twin study indicated that fragility could be considered to represent the secondary expression of a more fundamental mechanism operative within the genome (Austin et al. 1992). Similarly, a basic genetic mechanism was proposed for the concurrent expression of both autosomal FS and fragile-X sites in individuals from three families with the fra-X syndrome (Barletta et al. 1991). This concept was also discussed by other groups (Amarose et al. 1987, Smeets and Ares 1990).

Since individuals with TS are usually phenotypically normal, an association between FS and neurobehavioral characteristics (rather than with dysmorphic or disease-related phenomena) may easily have been overlooked during earlier population based surveys of common FS. The quotation by Roger Kurlan that "...there is at least a little

bit of TS in us all" (Kurlan 1993) seems entirely appropriate in view of the population frequency of expression of common FS and the proposed role of some change at these sites being associated with neurobehavioral alteration.

In conclusion, if differential expression of certain fragile sites can be confirmed to be associated with variant behavior, and these FS can also be demonstrated to show a consistent relationship with certain classes of repeat sequences, such sites may be important with regard to both normal evolutionary processes (King 1994), as well as having the potential to be involved in potentially deleterious dynamic mutations.

As indicated, these findings do not rule out the need to search for a major TS gene, which may however not be as strongly linked to any of the behavioral phenotypical features as to the chromosomal phenotype which is described in this paper.

Acknowledgments

The authors are grateful to Cecile Buckle for her efforts with regard to obtaining blood samples, Dawid Rood for technical laboratory assistance, and to Susann Steyn for expert secretarial assistance and typing.

References

Amarose AP, Huttenloker PR, Sprudzs RM, Laitsch TS, Pettenati MJ (1987) A heritable fragile site 12q14.13 segregating in a family with the fragile X chromosome. Hum Genet 15:4-6

Austin MJ, Collins JM, Carey LA, Nance WE, Neale MC, Schieken RM, Brown JA (1992) Aphidicolin inducible common fragile site expression results from a population survey of twins. Am J Hum Genet 50:76-83

Barletta C, Ragusa RM, Garafalo G, Scillato F, Ruggeri M (1991) Segregation analysis of autosomal fragile sites in three families with the fragile X chromosome. Ann Genet 34:111-114

Baron M, Shapiro ES, Shapiro AK, Rainer JD (1981): Genetic analysis of Tourette syndrome suggesting a major gene defect. Am J Hum Genet 36:767-775

Chromosome Coordinating Meeting (1992) Cuticchia AJ, Pearson PL, Klinger HP (eds) "Genome Priority Reports". Vol. 1. Basel:Krager, 1993

Chudley AE, Ray M, Evans JA, Cheang M (1990) Possible association of rare autosomal folate sensitive fragile sites and idiopathic mental retardation: A blind controlled population study. Clin Genet 38:241-256

Cloninger CR (1994) Turning point in the design of linkage studies of schizophrenia. Am J Med Genet 54:83-92

Comings DE, Comings BG, Devor EJ, Cloninger CR (1984) Detection of a major gene for Gilles de la Tourette syndrome. Am J Hum Genet 36:586-600

Comings DE, Comings BG (1993) Comorbid behavioral disorders. In Kurlan R (ed) "Handbook of Tourette's syndrome and related tic and behavioral disorders." New York: Marcel Dekker, pp 111-147

Comings DE (1994) Role of genetic factors in human sexual behavior based on studies of Tourette syndrome and ADHD probands and their relatives. Am J Med Genet 54:227-241

Craig-Holmes AP, Strong LC, Goodacre A, Pathak S (1987) Variation in the expression of aphidicolin-inducible fragile sites in human lymphocyte cultures. Hum Genet 76:134-137

Crowe RR (1993) Candidate genes in psychiatry: An epidemiological perspective. Am J Med Genet 48:74-77

DeLisi LE, Reiss AL, White BJ, Gershon ES (1988): Cytogenetic studies of males with schizophrenia: Screening of the fragile X chromosome and other chromosomal abnormalities. Schizophrenia Research 1:277-281

Devor EJ (1984) Complex segregation analysis of Gilles de la Tourette syndrome: Further evidence for a major locus mode of transmission. Am J Hum Genet 36:704-709

Freund LS, Reiss AL, Hagerman R, Vinogradov S (1992): Chromosome fragility and psychopathology in obligate female carriers of the fragile X-chromosome. Arch Gen Psychiatry 49:54-60

Furuya T, Hagiwara J, Ochi H, Tokuhiro H, Kikiawada R, Karube T, Watanabe S (1991) Changes of common fragile sites on chromosomes according to menstrual cycle. Hum Genet 86:471-474

Garofalo G, Ragusa RM, Barletta E, Spina E (1992) Schizophrenia and chromosomal fragile sites. Am J Psychiat 149:1116

Gericke GS, Simonic I, Cloete E, Buckle C, Becker PJ (1995) Increased chromosomal breakage in Tourette syndrome predicts the possibility of variable multiple

gene involvement in spectrum phenotypes: Preliminary findings and hypothesis. Am J Med Genet 60:444-447

Glover TW, Stein CK (1988) Chromosome breakage and recombination at fragile sites. Am J Hum Genet 43:265-273

Hecht F, Tajara EH, Lockwood D, Sandberg AA, Hecht BK (1988) New common fragile sites. Cancer Genet Cytogenet 33:1-9

Hecht F, Hecht BK (1991) Genetic activity and fragile sites. Cancer Genet Cytogenet 55:113 (Editorial)

Hecht F (1991) Biologic agents and induction of chromosome fragile sites. Cancer Genet Cytogenet 54:127 (Editorial)

Jordan DK, Burns TL, Devilbiss JE, Woolson RF, Patil SR (1990) Variability in expression of common fragile sites – in search of a new criterion. Hum Genet 85:462-466

Kerbeshian J, Burd L, Leech C, Rorabaugh A (1991) Huntington disease and childhood-onset Tourette syndrome. Am J Med Genet 39:1-3

Kidd KK, Pauls DL (1982) Genetic hypotheses for Tourette syndrome. In: Friedhoff AJ, Chase TN (eds) "Advances in neurology." New York: Raven Press, Vol 35, pp 243-249

King DG (1994) Triplet repeat DNA as a highly mutable regulatory mechanism. Science 263:595-596

Kinght SJL, Flannery AV, Hirst MC, Campbell L, Christodoulou Z, Phelps SR, Pointon J, Middleton-Price HR, Barnicoat A, Pembrey ME, Holland J, Oostra BA, Bobrow M, Davies KE (1993) Trinucleotide repeat amplification and hypermethylation of a CpG island in FRAXE mental retardation. Cell 74:127-134

Kurlan R, Whitmore D, Irvine C, McDermott MP, Como PG (1994) Torette's syndrome in a special education population: A pilot study involving a single school district. Neurology 44:699-702

Kuwano A, Kajii T (1987) Synergistic effect of aphidicolin and ethanol on the induction of common fragile sites. Hum Genet 75:75-78

Laird C, Jaffe E, Karpen G, Lamb M, Nelson R (1987) Fragile sites in human chromosomes as regions of late-replicating DNA. Trends Genet 3:274-281

Morgan R, Morgan SS, Hecht BK, Hecht F (1988) Fragile sites at 4q23 and 7q11.23 unique to bone marrow cells. Cancer Genet Cytogenet 31:47-53

Murano I, Kuwano A, Kajii T (1989) Cell type-dependent difference in the distribution and frequency of aphidicolin-induced fragile sites: T and B lymphocytes and bone marrow cells. Hum Genet 84:71-74

Paradee W, Jayasankar V, Mullins C, Wilke C, Glover TW, Smith DI (1994) Molecular characteriztion of the 3p14.2 constitutive fragile site. Am J Hum Genet 55:A114

Pauls DL, Leckman JF (1986) The inheritance of Gilles de la Tourette's syndrome and associated behaviors: Evidence for autosomal dominant transmission. New Engl J Med 315:993-997

Rao PN, Heerema NA, Palmer CG (1988) Fragile sites induced by FudR, caffeine and aphidicolin: their frequency, distribution and analysis. Hum Genet 78:21-26

Richards RI, Nancarrow JK, Mangelsdorf M, Kremer E, Holman K, Eyre H, Doggett NA, Le Paslier D, Hon T, Callen DF, Sutherland GR (1994) Chromosomal fragile site FRA16A: Implications for fragile site genesis. Am J Hum Genet 55: A12

Smeets DF, Ares A (1990) Genetic determination of fragile site expression. Am J Hum Genet 47:196-201

Smeets DF, Merkx G (1990) Neither age nor sex influence the expression of folate sensitive common fragile sites on human chromosomes. Hum Genet 86:76-78

Stopera SA, Ray M (1989) Expression and distribution of aphidicolin-induced fragile sites in chronic myeloid leukemia, acute lymphocytic leukaemia and acute myeloid leukaemia. Cytobios 60:103-109

Sutherland GR, Baker E, Fratini A (1985) Excess thymidine induces sensitive fragile sites. Am J Med Genet 22:433-444

Tedeschi B, Vernole P, Sanna ML, Nicoletti B (1992) Population cytogenetics of aphidicolin-inducible fragile sites. Hum Genet 89:543-547

Telvi L, Leboyer M, Chirou C, Feingold J, Pensot G (1994) Is Rett syndrome a chromosome breakage syndrome? Am J Med Genet 51:602-605

Vogt P (1990) Potential genetic functions of random repeated DNA sequence blocks in the human genome are based on a highly conserved "chromatin folding code". Hum Genet 84:301-336

Yunis JJ, Soreng AL, Bowe AE (1987) Fragile sites are targets of diverse mutagens and carcinogens. Oncogene 1:59-69

Yunis JJ, Hoffman WR (1989) Nuclear enzymes, fragile sites and cancer. J Gerontol 44:37-44

CHAPTER 4

Hum Genet (1996) 97:524-531

ORIGINAL INVESTIGATION

I. Simonic, G. S. Gericke

The enigma of common fragile sites

Received: 5 July 1995 / Revised: 25 September 1995

Abstract

One hundred and fifty breakpoint sites were recorded during an analysis of aphidicolin-ethanol inducible fragile sites (FS) in 56 blood samples and 13 amniocyte cultures and were classified according to the criteria formulated by the Chromosome Coordinating Meeting. The finding of previously unlisted FS in this sample, the altered expression of FS in conditions not usually associated with chromosomal abnormalities and the apparent lack of tissue specificity indicate the importance of one or more fundamental mechanisms operating to produce the diverse associated clinical phenotypes, with the chromosomal fragility representing an intermediate phenotype. Several lines of evidence converge towards the conclusion that FS are a manifestation of an altered state of genetic activity at areas associated with transcriptional regulation, because of their concordance with CpG islands, nuclease sensitive sites, replication origins, zinc finger protein domains and viral integration sites. An investigation is required whether this phenomenon could contribute both to evolutionary diversity through increased recombination, the formation of unstable repeat sequences and variable methylation, and to the expression of multigene disease processes resulting in the production of variable and complex phenotypes, even within families.

Introduction

The basic biology and evolutionary significance of chromosomal fragile sites (FS) remain enigmatic. FS on human chromosomes can be defined as vulnerable regions where lesions occur spontaneously or after induction with certain break-inducing agents (Smeets and Merkx 1990). These sites have been divided into two main groups, viz. "rare" and "common" FS (Hecht 1986). Those present in less than 2.5% of the population are considered to be rare (Berger et al. 1985), whereas common FS are considered to be present in more than 50% of the population (Hecht 1986) or in almost every individual (Samadder et al. 1993).

Aphidicolin-inducible common FS appear to be ubiquitous in humans and other mammals. The estimated high degree of genetic determination for pooled FS frequencies suggests that these sites result from a common process that is under stringent genetic control (Austin et al. 1992). Aphidicolin-induced common FS have been considered as candidates of pathological importance (Hecht 1991).

We have documented the increased expression of these common FS in individuals with Gilles de la Tourette syndrome (TS), a complex neurobehavioural spectrum disorder (Gericke et al. 1995, 1996). This finding has raised questions concerning possible relationship between FS and gene activity and suggests an underlying role for these genomic areas with regard to the pattern of expression of co-morbid disorders in TS. If FS expression reflects altered transcriptional activity at multiple sites, common FS may play a role in the expression of polygenic disorders in general.

Materials and methods

Specimens from 21 Caucasians (4 females and 17 males) diagnosed as having TS by means of DSM IV criteria, were analysed as one group. These criteria require the presence of both motor and one or more vocal tics, occurring many times a day for more than 1 year with onset before age 18, causing significant social and occupational impairment and are not due to the effects of a medical condition or substance use

(American Psychiatric Association 1994). A separate group consisted of randomly chosen Caucasian blood samples referred to our laboratory for chromosomal analysis, viz. Five for fragile X (fra-X) screening, one patient with velocardiofacial (Shprintzen) syndrome together with blood samples from non-affected sibling and both parents, two patients with Duchenne muscular dystrophy, two patients with Cornelia de Lange syndrome and a patient with Russel-Silver syndrome, two couples with a history of recurrent fetal loss and three healthy males referred for chromosomal damage screening after being exposed to ethylene dioxide together with matching control blood specimens. The remainder included blood samples of newborn children with dysmorphic features where a syndromic diagnosis had not been made. The number of patients included in the second group consisted of 13 females and 22 males. All individuals included in the study had normal karyotypes, according to standard banded chromosome analyses and no cytogenetic evidence was found for any of the suspected fra-X cases having increased breakage at Xp27.3 in folate-deficient culture media.

In total, 56 blood samples were aphidicolin treated. Thirteen amniocyte cultures, reported to be "well growing", 48 h before harvesting, were also included on a random basis.

Two phytohaemaglutinin (PHA)-stimulated lymphocyte cultures were established for each individual by using 0.5ml whole blood in RPMI-1650 medium supplemented with 10% fetal bovine serum (FBS). After 48 h, aphidicolin dissolved in 70% ethanol was added (24 h treatment) to each 10 ml culture to obtain a final concentration of 0.1μM. Standard cytogenetic techniques were used for harvesting and slide preparation.

Complete metaphase spreads (n=50) for each culture (100 per individual) were assessed at the 350-400 band level. All types of aberrations (gaps, breaks, chromatid exchanges) were scored as single events at the band(s) involved.

Fibroblasts derived from amniotic fluid were maintained and subcultured in Ham's F10 medium and 20% FBS. Aphidicolin dissolved in 70% ethanol was added to obtain a final concentration of 0.1μM, 25h prior to harvesting. Cultures were treated with colcemid for the last three hours. Cells were processed in the standard manner for the preparation of chromosomes. Cultured amniocytes were shown to be extremely sensitive to aphidicolin-ethanol treatment and the mitotic index in most samples was low. However it was decided to evaluate all available complete metaphases for the distribution of gaps and breakpoints.

Results

In total, 150 breakpoint sites were recorded. These site were grouped and evaluated separately according to their status as common FS, rare FS (Chromosome Coordinating Meeting 1992), breakpoints recorded by others, but not listed as common FS (Table 1), and so-called random breakpoints (Tables 2, 3). All sites reported as being tissue-specific for other tissues were also found in T-lympocyte cultures. It was therefore decided to include the analysis of 576 metaphases from 13 amniocyte cultures to observe differences in their expression when compared with lymphocytes. A breakpoint was counted as "one" when occurring on one of the chromosomes, and as "two" when occurring on both homologues.

There was no intention, from the beginning of the study, to make any statistical comparison of FS expression between samples from different groups of patients. Our only interest was to screen for the frequencies of random breaks. As a much higher expression at most sites assessed as random breakpoints was found in patients with TS, we decided to evaluate FS expression in this group separately.

The computed means per 100 cells of common FS in T-lymphocyte cultures of both investigated groups of patients are summarized in Table 1. Despite the observation that the rank orders of most expressed sites differed between the two groups (TS and non-TS) in agreement with previous findings (Gericke et al. 1995, 1996), we found

the sites 3p14, 16q23 and Xp22.3 as being the most prevalent in both groups. The hierarchy of the remaining sites in the TS group was: 1q44, 7q32, Xq22, 14q24, 1p21, 7p22, 7p13, 6q26, 2p24, 4q31, 11p13, 8q22, 2q32, 9q32, 22q12, 2q37, etc. The hierarchy of these sites in the non-TS group was: 7q32, 1q44, Xq22, 14q24, 1p21, 7p21, 6q26, 7q31, 2p24, 4q31, 7p13, 2q32, 11p13, 9q32, 8q22, 22q12, etc.

Some 22 breakpoints are listed according to their frequencies of occurrence in Table 2. All of these are expressed at least in one subject in 4% or more of the metaphase spreads, which is a cut-off frequency necessary to cytogenetically classify an individuals as fra-X positive (Jacobs et al. 1980). The rank orders of the means for the above described group of aphidicolin-inducible FS were found to be different between the patient groups, viz. for the TS group: 10p12-p13, 10p11.2, 14q13, 6p21.1, Xq13, 18p11.2, 6q23, Xp22.1, 17q25, 8q11.2, 4q34-q35, 15q12-q13, Xq12, 21q22, 9q34, 4q31, 3p13, 5q12-q13, 15q14-q15, 18p11.3, 5q34, 20q11.2; and for the non-TS group: 10p12-p13, 18p11.3, 10p11.2, Xp22.1, 5q12-q13, 6q23, Xq13, 14q13, 18p11.2, 4q34-q35, 6p21.1, 8q11.2, 15q14-q15, 15q12-q13, Xq12, 20q11.2, 9q34, 5q34, 14q31, 17q25, 21q22, 3p13.

The means for the listed group of breakpoints is in most cases approximately twice as high in the TS group. The insufficient scope of the study and the lack of complete clinical data for non-TS individuals, especially with regard to their neurobehavioral status, do not allow definite conclusions to be made from the observed differences in this study.

The localization of all other breakpoints repeatedly found in our study is given in Table 3. According to the most widely used criteria discussed in Jordan et al. (1990), they cannot be assigned as FS. These sites can only be recognized as being fragile based on Craig-Holmes et al. (1987) approach, where the site is scored as positive if it is expressed at least once in two or more subjects.

Table 1

Mean frequency (%) of common aphidicolin-inducible fragile sites for TS and non-TS patients.

Site	TS (mean)	TS (SD)	non-TS (mean)	non-TS (SD)	Site	TS (mean)	TS (SD)	non-TS (mean)	non-TS (SD)
1p36	2.43	1.37	1.54	1.25	**8q24.3**	4.19	2.7	2.29	1.28
1p32	3.29	1.58	2.86	1.99	**9p21**	1.24	1.34	0.8	0.91
1p31	3.19	2.04	2.94	3.08	**9q12**	0.86	0.89	0.49	0.79
1p22	0.62	0.72	0.77	0.88	**9q22**	0.9	1.11	0.71	0.8
1p21	15.86	5.37	10.8	4.45	**9q32**	9.33	5.15	7.03	3.91
1q21	0.52	0.66	0.63	0.88	**9q34**	1.24	1.44	0.66	1.12
1q25	6.57	2.95	3.94	1.82	**10p13**	5.81	3.25	3.31	2.55
1q31	0.9	0.75	0.97	0.96	**10p11.2**	3.43	1.99	1.2	1.24
1q42	1.48	1.18	0.69	0.87	**10q21**	0.67	1.21	0.89	0.87
1q44	22.9	5.66	15.86	7.72	**10q22**	1.81	1.1	1.77	1.56
2p24	11.81	3.02	9.4	4.87	**10q23**[b]	0.52	0.66	0.49	0.72
2p16	3.57	1.73	2.89	2.31	**10q25**	2.81	2.28	1.83	2.06
2p13	4.67	2.83	3.69	2.02	**10q26**	6.1	3.19	5.06	3.06
2q13[b]	1.05	0.95	0.83	1.21	**11p15**	3.14	1.39	2.77	1.9
2q21	6.1	4.87	3.71	2.71	**11p14**	4.05	2.95	2.91	1.69
2q31	2.67	1.83	2.4	2.06	**11p13**	10.95	4.25	7.54	3.02
2q32	9.81	3.2	7.54	2.86	**11q13**	0.62	0.65	0.2	0.4
2q33	5.1	2.86	3.91	2.27	**11q14**	6.57	2.84	5.03	3.3
2q37	8.29	5.03	5.43	3.06	**11q23**	0.81	1.1	0.63	0.92
3p36[a]	5.76	2.67	3.23	2.16	**12q13**[b]	1.19	1.53	0.8	1.02
3p24	3.76	1.9	2.74	2.4	**12q21**	3.81	2.11	2.54	1.6
3p21[a]	2.1	1.27	0.89	0.87	**12q24**	2.67	2.1	1.86	1.51
3p14	59.05	9.51	52.31	12.82	**13q13**	7.24	4.46	4.94	2.32
3p13	1.19	1.1	0.54	0.69	**13q21**	0.95	1.05	1.2	1.15
3q13[a]	4.71	2.6	2.74	1.85	**13q32**	1.86	1.64	1.06	0.78
3q21[a]	1.1	1.06	0.71	0.84	**13q34**[a]	1.57	1.22	0.89	1.15
3q25	2.19	1.82	1.4	0.95	**14q13**	3.48	1.99	1.2	1.24
3q27	5.1	2.16	3.29	2.13	**14q23**	2.86	1.96	1.86	1.55
4p16	3.52	2.77	2.49	2.24	**14q24**	16.14	4.33	11.83	4.37
4p15	3.67	2.51	2.29	1.89	**14q31**	1.24	1.19	0.63	0.96
4q21	1.24	1.41	1.26	1.01	**15q12-13**	1.57	1.26	0.74	0.81
4q23[a]	1.9	1.51	1.83	1.36	**15q14-15**	0.97	1.09	0.86	0.83
4q27[a]	1	1.02	0.66	0.66	**15q22**	1	1.02	0.57	0.86
4q31	11.24	3.83	8.29	3.3	**16p13**[b]	1.1	1.06	0.46	0.76
4q35	1.62	1.86	0.97	1.16	**16q22**	4.67	5.54	1.89	1.74
5p14	3.29	3.47	2.83	2.6	**16q23**	49.9	10.88	41.71	10.19
5p13	2.14	1.73	1.69	1.45	**17p12**[b]	1.14	0.99	0.17	0.37
5q13	1.14	1.28	1.34	1.37	**17q23**	1.62	1.81	1.23	1.06
5q15	5.43	1.84	4.54	2.7	**17q25**	2.14	1.93	0.63	1.07
5q21	2	1.88	1.14	1.18	**18p11.3**	0.48	1.01	2.2	3.62
5q31	0.81	0.85	0.77	0.75	**18p11.2**	2.52	1.79	1	1.33
6p25	12.29	4.24	7.31	3.84	**18q12**	5.24	3.04	4.49	2.91
6p22	2.29	1.61	1.26	1.66	**18q21**	1.38	1.17	0.71	0.93
6p21.1	3.05	2.17	0.91	0.94	**18q22**	1.62	1.33	0.74	0.76

Table 1

Continued

Site	TS (mean)	TS (SD)	non-TS (mean)	non-TS (SD)	Site	TS (mean)	TS (SD)	non-TS (mean)	non-TS (SD)
6q15	1.81	1.56	1.03	0.83	**19q13.1**	2.43	1.73	1.46	1.36
6q21	3.71	1.86	3.66	2.28	**20p12**	6.81	3.22	4.69	2.16
6q25	2.29	1.93	1.23	1.12	**20q11.2**	0.76	0.87	0.49	0.82
6q26	11.1	5.31	10.14	3.79	**20q13.2**	0.33	0.47	0.23	0.59
7p22	15.57	5.96	10.51	4.03	**21q21**	0.14	0.35	0.11	0.4
7p14	1.05	1.17	0.66	0.78	**21q22**	1.43	1	0.63	1.15
7p13	12.43	3.49	8.17	4.62	**22q11.2**	0.67	0.94	0.49	0.73
7q11	3.71	1.86	3.74	2.53	**22q12**	9.24	4.56	5.71	3.15
7q21	2.81	1.68	2.29	1.91	**22q13**[b]	0.67	0.84	0.49	0.72
7q22	0.71	0.7	0.91	0.92	**Xp22.3**	31.05	12.27	15.8	10.19
7q31	11.33	3.68	10.03	4.62	**Xp22.1**	2.24	1.41	1.69	1.41
7q32	21.33	5.19	15.94	5.62	**Xq13**	2.52	1.99	1.23	1.53
7q36	0.9	0.97	0.86	1	**Xq22**	17.05	5.35	13.89	6.27
8p21	1.05	0.79	0.6	0.73	**Xq26**	0.38	0.65	0.23	0.42
8q11.2	2.05	1.86	0.89	1.04	**Xq27**	1	1.2	0.71	0.9
8q21[b]	0.24	0.43	0.09	0.28	**Yq11**	0.71	1.39	0.34	0.75
8q22	10.19	2.91	5.83	2.93	**Yq12**	0.48	0.73	0.2	0.58
8q24.1	4.38	2.68	2.74	1.95					

[a] Breakpoints detected in our study and reported by others but not listed as common FS in Chromosome Coordinating Meeting (1992)

[b] Breakpoints considered as rare folate-sensitive FS

Table 2

Mean frequency (%) of non-specific aphidicolin-inducible breakpoints for TS and non-TS groups

Site	TS (mean)	TS (SD)	non-TS (mean)	non-TS (SD)	Site	TS (mean)	TS (SD)	non-TS (mean)	non-TS (SD)
3p13	1.19	1.1	0.54	0.69	**14q31**	1.24	1.19	0.63	0.96
4q35	1.62	1.86	0.97	1.16	**15q12-13**	1.57	1.26	0.74	0.81
5q13	1.14	1.28	1.34	1.37	**15q14-15**	0.97	1.09	0.86	0.83
6p21.1	3.05	2.17	0.91	0.94	**17q25**	2.14	1.93	0.63	1.07
6q25	2.29	1.93	1.23	1.12	**18p11.3**	0.48	1.01	2.2	3.62
8p21	1.05	0.79	0.6	0.73	**18p11.2**	2.52	1.79	1	1.33
8q11.2	2.05	1.86	0.89	1.04	**18q22**	1.62	1.33	0.74	0.76
9q34	1.24	1.44	0.66	1.12	**21q22**	1.43	1	0.63	1.15
10p13	5.81	3.25	3.31	2.55	**Xp22.1**	2.24	1.41	1.69	1.41
10p11.2	4.1	2.24	2.03	1.78	**Xq11-12**	1.52	1.74	0.74	1.02
14q13	3.48	1.99	1.2	1.24	**Xq13**	2.52	1.99	1.99	1.53

Table 3

Mean frequency (%) of sporadically occurring aphidicolin-inducible breakpoints for TS and non-TS groups

Site	TS (mean)	TS (SD)	non-TS (mean)	non-TS (SD)	Site	TS (mean)	TS (SD)	non-TS (mean)	non-TS (SD)
2p12	0.48	0.91	0.17	0.51	**17q21**	0.52	0.73	0.4	0.73
4q12	0.33	0.47	0.37	0.59	**19q13.3**	0.62	0.72	0.29	0.78
5q33	0.76	1.06	0.66	0.79	**20q11.2**	0.76	0.87	0.49	0.82
8p25	0.62	0.79	0.31	0.52	**20q13.2**	0.33	0.47	0.23	0.59
9p24	0.62	0.84	0.46	0.65	**21q21**	0.14	0.35	0.11	0.4
10p14-15	0.38	0.58	0.23	0.48	**22q11.2**	0.67	0.94	0.49	0.73
10q11.2	0.76	1.11	0.34	0.58	**Xp21**	0.1	0.29	0.03	0.17
12p12	0.67	0.71	0.74	0.73	**Xp11.4**	0.14	0.35	0.34	0.58
12q12	0.48	0.66	0.03	0.17	**Xp11.2**	0.29	0.55	0.11	0.32
12q24.3	0.29	0.63	0.09	0.28	**Xq21**	0.19	0.39	0.03	0.23
14q12	0.67	0.99	0.66	0.79	**Xq24**	0.24	0.53	0.31	0.62
14q32	0.52	1.14	0.2	0.47	**Xq26**	0.38	0.65	0.23	0.42
16p11.2	0.24	0.68	0.17	0.38	**Yq11**	0.71	1.39	0.34	0.75
16q12	0.24	0.43	0.11	0.32	**Yq12**	0.48	0.73	0.2	0.58
16q21	0.57	0.73	0.37	0.48					

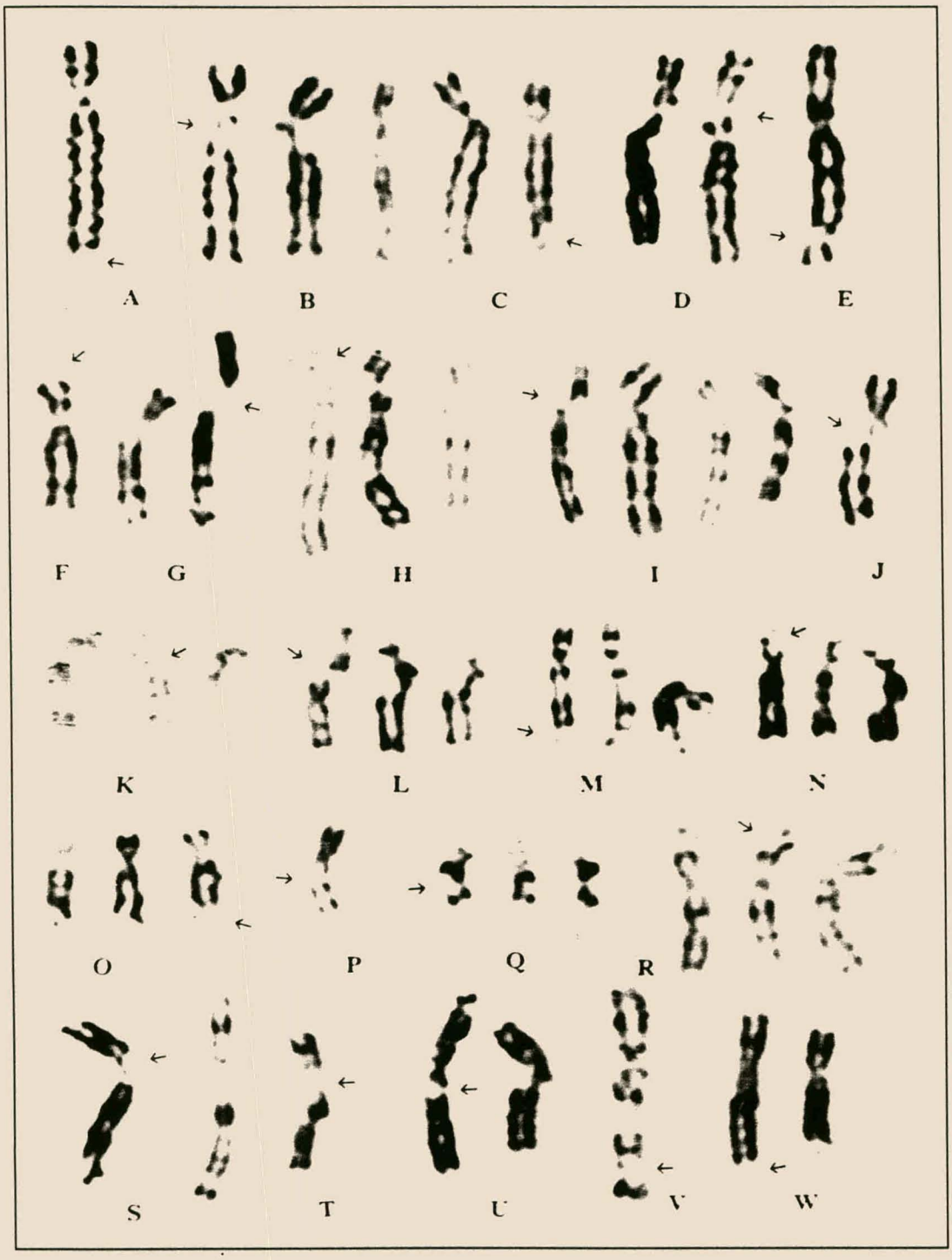
Table 4

Mean frequency (%) of aphidicolin-inducible breakpoints in cultured amniocytes

Site	Mean	Site	Mean	Site	Mean	Site	Mean
1p31	14.93	**3q13**	2.6	**7q21**	1.22	**15q22**	1.39
1q25	1.04	**3q27**	7.81	**7q31**	2.08	**16q23**	15.1
1q44	1.39	**4q23**	2.26	**7q32**	1.04	**18q12**	1.39
2q21	2.26	**4q31**	1.04	**10q11.2**	4.69	**18q21**	4.86
2q32	1.04	**4q35**	1.04	**10q26**	1.22	**19q13.1**	1.91
2q33	1.04	**5q13**	2.97	**11q14**	1.74	**20q12**	2.6
3p26	2.26	**6q26**	2.43	**12p13**	1.39	**22q12**	1.57
3p21	1.04	**7p22**	1.04	**13q32**	3.47	**Xq12**	1.04
3p14	14.1	**7q11**	9.9	**14q24**	1.39	**Xq22**	1.74

Figure 1

Examples of some breakpoints (*arrows*) included in Tables 2 and 3, and breakpoint Xq28 included in Table 1. A. 4q35, B. 5q13, C. 5q34, D. 6p21.1, E. 6q25, F. 8p25, G. 8q11.2, H. 10p13, I. 10p11.2, J. 10q11.2, K. 14q13, L. 15q14-q15, M. 17q25, N. 18p11.2, O. 18q22, P. 10q11.2, Q. 21.q22, R. Xp22.1, S. Xp11.2, T. Xq12, U. Xq13, V. Xq26, W. Xq28



Breakpoints expressed in cultured amniocytes (Table 4) usually corresponded to those frequently found in skin fibroblasts from normal individuals (Murano et al. 1989a). The average number of FS induced in amniocytes was low compared with PHA-stimulated T-lymphocytes, and because there was a strong inhibition of cell division, a comparison of induced frequencies between these cell types was meaningless. Whereas fragile sites at 16q23 and 3p14 were again most frequent, Xp22.3 expression was low. In addition, sites at 1p31, 7q11, 3q27 (probably described as 3q26.2 in the above mentioned study), 10q11.2, 18q21, 13q32, 5q13, 3p26, 20p12.1 or 2q21 were not the most frequently seen in cultured T-lymphocytes.

Sites at the 4q23 and 7q11.2 bands were previously regarded as bone-marrow-specific (Morgan et al. 1988) and but were subsequently found in B lymphocytes and skin fibroblasts by Murano et al. (1989a) together with the site at 10q11.2. These sites were also observed in cultured amniocytes and a low level of their expression compared with other common FS was found in T lymphocytes. This finding appears to support our conclusion that aphidicolin-induced FS is not as tissue-specific as suggested. It remains to be seen whether the differences in cell-type-dependent frequency of expression are attributable to either different replication rates or the sequence of DNA replication, differences in gene activity, or tissue sensitivity to aphidicolin treatment (Murano et al. 1989b). Most of the sites included in Table 2 were also sporadically found in cultured amniocytes.

It was not possible to localize all "new" breakpoints at the sub-band level because of differences in chromosomal resolution (Fig. 1). In some cases, involved areas coincided with cancer chromosome breakpoints (Hecht 1988): 5q13, 5q35, 9q34, 14q32, 18q11.2, 20q13, 21q22, 22q11.2, Xp11.2. The breakpoint in 14q13 was reported in patients with possible acquired chromosomal instability (Tedeschi et al. 1987).

Sites at 5q13, 17q21, 20q11.2, 20q13.3, 21q22, 22q11.2, Xq13 were recognized as common chromosome breaks (not FS) by Michels (1985) and Xq26 as a FS by the same author. The breakpoint at 6p21, discussed by Tedeschi et al. (1992) as being a common FS, is not listed as such by the Chromosome Coordinating Meeting (1992).

The 18p11.3 breakpoint reported to be associated with translocation and heritable FS at 2q13 and fetal demise at 10 weeks of gestation (Jacky et al. 1995) was expressed, respectively, in 14% and 16% of cells in patients with Cornelia de Lange and Russel-Silver syndromes.

Discussion

In this paper, we report FS expression phenomena in a wide range of genetically compromised individuals. The finding of previously unlisted FS in this sample, the altered expression of FS in conditions not usually associated with chromosomal fragility and the apparent lack of tissue specificity indicate the importance of one or more fundamental mechanisms operating to produce the diverse associated phenotypes recorded here, with the chromosomal fragility representing an intermediate phenotype. In this regard, the comment by Hecht (1991) that common FS can be classified by means of chemical inductive methods, with a large currently unclassified group representing a "wastepaper basket", could perhaps be reinvestigated by ascertainment of the type of individuals in which their expression is altered.

It will be of interesting to see whether groups of common FS exist that are over- or underexpressed in specified groups of patients. The high frequency of breakpoints in 18p11.3 in the patients with Cornelia de Lange syndrome and Russel Silver syndrome as included in our study, or the overall higher expression of breakpoints as summarized in Table 2 in TS patients may serve as examples of the sofar unknown importance of "fragile" chromosome areas or the specific regulatory mechanisms involved in their expression.

Aphidicolin, an inhibitor of DNA polymerase alpha (Glover et al. 1984), may alter DNA synthesis in specific regions that concur with FS expression (Ahuja 1990). The enhancing action of ethanol (Kuwano and Kajii 1987) is considered to act through the suppression of RNA synthesis, which in turn probably influences the repair capacity of the cells (Obe and Ristow 1979).

The application of low stringency criteria for FS assignment and the enhancement of the effect of aphidicolin via the inhibition of DNA repair would probably lead to increased FS expression. This may have been responsible for the divergent results arising between studies attempting to correlate specific structural chromosome defects recorded in neoplasia with FS, or between breakpoints involved in reciprocal translocations and FS (Daniel 1986).

An association of the localization of common FS with active gene regions has been postulated primarily according to their occurrence in the light G-bands (Hecht and Hecht 1991). The obvious example supporting this idea is the expression of the common FS Xp22.3 in both homologues in females, whereas the site Xq22.1 is expressed only on the active X chromosome (Austin 1991).

As shown in our study, additional breakpoints occurring on the X chromosome may represent active transcription areas, since some of them are present in bands, where genes escaping X-inactivation have been found, such as Xp22.1 (*ZFX*), Xp11.2 (*UBE 1*, *SMCX*), and Xq13 (*RPS4X*) (Rappold 1993, Disteche 1995). The clustering of FS within light G bands, which represents areas of apparent GC richness with a high level of gene expression, and their concordance with CpG islands, nuclease-sensitive sites, replication origins, zinc finger protein domains and viral integration sites have been demonstrated in many studies (Porfirio et al. 1989, Ahuya 1990, Tadeschi et al. 1991, Austin et al. 1991, 1992, Lichter et al. 1992, Nancarrow et al. 1994). Synergy seems to occur between adenoviruses and aphidicolin because adenoviruses type 5 and 12 act at FS sensitive to aphidicolin (Caporossi et al. 1991).

A histogram giving the distribution of sequences hybridizing with the isochores of the G + C-richest H3 family (Saccone et al. 1992) on the human G-banded chromosomes is in good agreement with the areas of highest expression of common FS. These hybridizing sequences can also be correlated with other breakpoints not currently classified as common FS (Tables 2, 3).

A strong association has been found between bands sensitive to restriction enzymes *Msp*I and *Hpa*II and those bands containing common FS (Porfirio et al. 1989, Tedeschi et al. 1991). Recognition sites of the isoschizomeric restriction

endonucleases *Hpa*II and *Msp*I are more frequent in CpG islands than in bulk DNA. The high expression of *Msp*I breakpoints at some bands, which also coincide with FS areas, indicates a role for CpG islands in their expression.

CpG sequences may represent one identifiable subset of *DNase-I*-hypersensitive sites (Elgin 1981). Moreover, agents known preferentially to attack chromatin *DNase-I*-hypersensitive sites induce non-randomly distributed chromosomal damage in areas showing striking concordance to locations of known common FS (Austin 1991).

The FS that have been cloned so far are the rare folate-sensitive sites FRAXA (Verkerk et al. 1991), FRAXE (Knight et al. 1993), FRAXF (Parrish et al. 1994), FRA16A (Nancarrow et al. 1994) and FRA11B (Jones et al. 1995) and all represent unstable poly $(CCG)_n$ repeats adjacent to CpG islands that become hypermethylated when the number of copies of the repeat exceeds certain limits. It has been suggested that FS originate at regions in the genome that are not normally associated with methylation and that the observed methylation associated with the above mentioned sites represents a consequence rather than a cause of FS mutation (Nancarrow et al. 1994).

Methylation is associated with imprinting; since both FS and imprinting appear to be evolutionary conserved responses and since more than 50% of breakpoints that have occurred during chromosome evolution in primates are reported at or close to FS (Miro et al. 1987), the possibility of a molecular basis for a link between the processes of imprinting and FS formation would clearly be of interest. Parent of origin effects have been recorded in relation to TS (Lichter et al. 1995) and as being associated with the processes of recombination, late replication and allelic expansion (J. G. Hall personal communication). These circumstances represent genetic conditions that can be considered as having an effect on FS expression. The possibility exists that FS represent a dynamic part of the process of imprinting.

The localization of FRA11B within a previously described gene indicates the possibility that other poly$(CCG)_n$ repeats contained within known genes are also associated with folate-sensitive FS expression. The association between FRA11B and

Jacobsen 11q-syndrome represents the first evidence and direct link between an autosomal FS and chromosome breakage (Jones et al. 1995).

Several authors have suggested that FS represent an amplification of specific DNA sequences (Sutherland 1985, Nussbaum 1986) including naturally occurring polypurine/polypyrimidine sequences (Sutherland and Hecht 1985). The $(CG)_n$ and $(TG)_n$ repeat families are not highly repetitive in the human genome and seem to be clustered. The potential of repeat unit involvement in areas other than rare FS has been discussed in connection with their preferential occurrence in promotor areas of active genes.

$(TG)_n$ sequence blocks have been shown to have an enhancer function in vitro (Hamada et al. 1984) and the potential to adopt a Z-DNA structure (Rich et al. 1984) that may attract specific nuclear proteins and therefore be able to arrest the DNA replication fork (Vogt 1990) or produce a nuclease-sensitive site. Arrest of the DNA replication fork coupled with incomplete or aberrant DNA repair may result in FS expression and increase the probability of double-strand breaks where daughter strands that have just replicated will join with the unreplicated parental strand. When replication is completed, the normal ligation process at the junction will generate the exchange, which can be visualized as chromatid exchange in metaphase chromosomes (Ikushima 1989).

Although no direct evidence exists for exchange between tandem arrays on sister chromatids or homologous chromosomes at minisatellite sequences, the results of several studies using simultaneous treatment for sister chromatid differentiation and common FS expression support the idea that sister chromatid exchanges on chromosomes may preferentially occur in FS regions (Hirsch 1991, Tsuji et al. 1991). It will be interesting to determine whether similar mechanisms are involved in rare but distinct phenomenon of partial endoreduplication which may originate in common FS, as documented in Figure 2.

The 5' flanking regions of transcriptionally active genes can often be structurally distinguished as nucleosome-free zones occupied by non-histone DNA-binding proteins (Mitchell and Tjian 1989). Stabilization of these domains is one of the

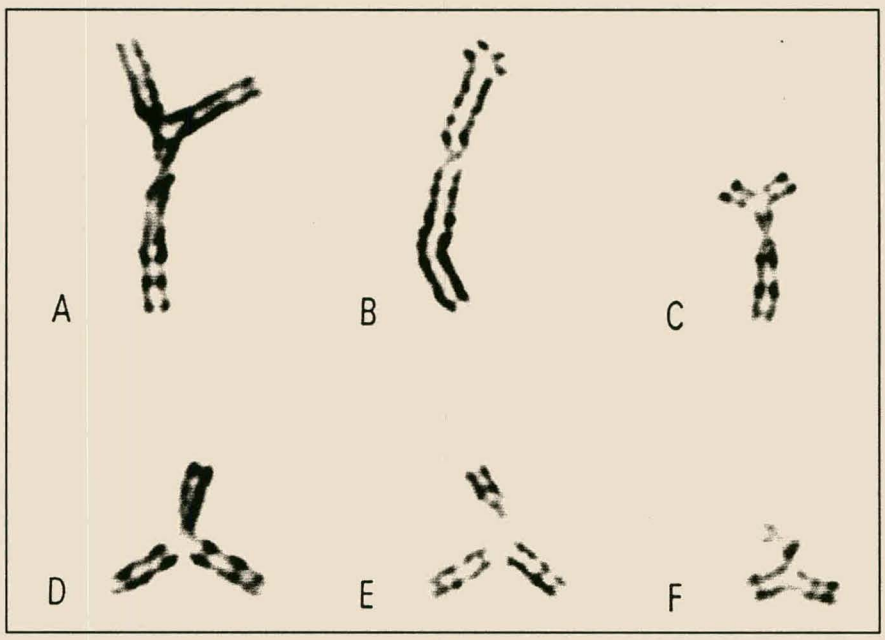
properties of locus control regions (LCR). The LCRs are saturated with DNA-binding sequences for transacting factors that enter into cooperative binding interactions to keep the promoter free of histones. They guarantee that transcription factor binding is stabilized at the replication fork in dividing cells and that it does not dissociate during chromatin assembly (Felsenfeld 1992).

Since DNA replication must be completed prior to initiation of mitosis (Navas et al. 1995) and since the addition of colcemid to the cell culture overcomes this important checkpoint, delayed or absent replication at the LCR may lead to incomplete packaging and collapse of the chromosome structure during metaphase, a process that results in FS expression (Sutherland 1988). LCR play an important role in the regulation of gene cluster transcription (Craddock et al. 1995). Whether FS localized in areas with multiple zinc finger protein domains (Lichter et al. 1992) are a cytogenetic expression of such regulatory regions, or whether increased frequencies of common FS in certain clinically distinguished population groups represent an induced response in active genes remains to be solved.

In conclusion, several lines of information indicate that FS are a manifestation of an altered state of genetic activity at areas associated with transcriptional regulation. Despite the current lack of molecular evidence, our findings and observations from routine cytogenetic screening, together with reports published to date, strongly support the idea that FS observed under specific culture conditions represent the cytogenetic expression of so called chromatin folding code in specific areas (see Vogt 1990, although the author himself denies any such correlation). If the processes described in this paper can be demonstrated to be an important interactive mechanism in the expression of genetic diseases, the potential for widespread changes caused by FS formation and/or repeat sequence formation could result in the production of variable and complex phenotypes, even within families. This could be responsible for disease as well as evolutionary advantageous genetic diversification. If found to be true, the mechanisms responsible for producing such divergent outcomes merit further study.

Figure 2

A-F Examples of tri-radials produced by chromatid breaks in a previous division, followed by nondisjunction of the distal fragment, where the origin of the partial duplication corresponds to the area of common FS expression. A. 1p21, B. 2p24, C. 7p13, D. 7q11, E. 9q12, F. 18q12



Acknowledgements

For helpful discussions, we are indebted to Elna Myburgh, Adrene Saayman, Edna Aucamp and Margaret Rasmussen at the Cytogenetic Laboratory, Department of Human Genetics, University of Pretoria.

References

Ahuja YR (1990) Fragile site analysis: its significance in environmental mutagenesis. In: Mendelsohn ML, Albertini RJ (eds) Mutation and the Environment, part B. Progress in clinical and biological research. Vol.340B. Wiley-Liss, New York, pp 371-384

American Psychiatric Association (1994) Diagnostic and Statistical Manual of Mental Disorders, 3rd edition, Washington DC

Austin MJF (1991) Expression of common fragile sites on the X chromosome corresponds with active gene regions. Cancer Genet Cytogenet 54:71-76

Austin MJF, Collins JM, Corey LA, Nance WE, Neale MC, Scieken RM, Brown JA (1992) Aphidicolin-inducible common fragile site expression: results from a population survey of twins. Am J Hum Genet 50:78-83

Berger R, Bloomfield CD, Sutherland GR (1985) Report of the committee on chromosome rearrangements in neoplasia and on fragile sites. (8th International Workshop on Human Gene Mapping) Cytogenet Cell Genet 40:490-535

Caporossi D, Bacchetti S, Nicoletti B (1991) Synergism between aphidicolin and adenoviruses in the induction of breaks at fragile sites on human chromosomes. Cancer Genet Cytogenet 54:39-53

Craddock CF, Vvas P, Sharpe JA, Ayyub H, Wood WG, Higgs DR (1995) Contrasting effects of alpha and beta globin regulatory elements on chromatin structure may be related to their different chromosomal environments. EMBO J 14:1718-1726

Chromosome Coordinating Meeting (1992) In: Cuticchia AJ, Pearson PL, Klinger HP (eds) Genome Priority reports, Vol. 1. Krager, Basel 1993

Craig-Holmes AP, Strong LC, Goodacre A, Pathak S (1987) Variation in the expression of aphidicolin-induced fragile sites in human lymphocyte cultures. Hum Genet 76:134-137

Daniel A (1986) Clinical implications and classification of the constitutive fragile sites. Am J Med Genet 23:419-427

Disteche CM (1995) Escape from X inactivation in human and mouse. Trends Genet 11:17-22

Elgin SCR (1981) Dnase-1-hypersensitive sites of chromatin. Cell 27:413-415

Felsenfeld G (1992) Chromatin as an essential part of the transcriptional mechanism. Nature 355:219-224

Gericke GS, Simonic I, Cloete E, Buckle C, Becker PJ (1995) Increased chromosomal breakage in Tourette syndrome may be responsible for variable secondary gene involvement in complex neurobehavioral disorders: preliminary findings and hypothesis. Am J Med Genet 60:444-447

Gericke GS, Simonic I, Cloete E, Becker PJ (1996) Increased expression of aphidicolin induced common fragile sites in Tourette syndrome: the key to understanding the genetics of comorbid phenotypes? Am J Med Genet (in press)

Glover TW, Berger C, Coyle J, Echo B (1984) DNA polymerase alpha inhibition by aphidicolin induces gaps and breaks at common fragile sites in human chromosomes. Hum Genet 67:136-142

Hamada H, Petrino MG, Kakunaga T, Seidman M, Stollar BD (1984) Characterization of genomic poly(dT-dG) poly(dC-dA) sequences. Structure, organisation and conformation. Mol Cell Biol 4:2610-2621

Hecht F (1986) Rare, polymorphic, and common fragile sites: a classification. Hum Genet 74:207-208

Hecht F (1988) Fragile sites, cancer chromosome breakpoints, and oncogenes all cluster in light G bands. Cancer Genet Cytogenet 31: 17-24

Hecht F, Hecht BK (1991) Genetic activity and fragile sites. Cancer Genet Cytogenet 55:113

Hirsch B (1991) Sister chromatid exchanges are preferentially induced at expressed and nonexpressed common fragile sites. Hum Genet 87:302-306

Ikushima T (1989) SCE enigma: methodology, mechanism and meaning of sister chromatid exchange. Annu Rep Res Reactor Inst Kyoto Univ 22:57-77

Jacky PB, Berry TL, Freesemann SM, Foxley RA, Lamb OA, Langlois MI, Meagher KL, Olson JS, Rubinstein JS, Reiss JA, Weeks FF, Yoshitomi MJ (1995) Association of heritable fragile site at 2q13 with fetal demise. Cytogenet Cell Genet 69:118

Jacobs PA, Glover TW, Mayer M, Fox P, Gerard JW, Dunn HG, Herbst DS (1980) X-linked mental retardation. A study of seven families. Am J Med Genet 7:471-489

Jones C, Penny L, Mattina T, Yu S, Baker E, Voullaire L, Langdon WY, Sutherland GR, Richards RI, Tunnacliffe A (1995) Association of a chromosome deletion syndrome with a fragile site within the proto-oncogene CBL-2. Nature 376:145-149

Jordan DK, Burns TL, Devilbiss JE, Woolson RF, Patil SR (1990) Variability in expression of common fragile sites: in search of new criterion. Hum Genet 85:462-466

Knight SJL, Flannery AV, Hirst MC, Campbell L, Christodoulou Z, Phelps SR, Pointon J, Middleton-Price HR, Barnicoat A, Pembrey ME, Holland J, Oostra BA, Bobrow M, Davies KE (1993) Trinusleotide repeat amplification and hypermethylation of a CpG island in FRAXE mental retardation. Cell 74:127-134

Kuwano A, Kajii T (1987) Synergistic effect of aphidicolin and ethanol on the induction of common fragile sites. Hum Genet 75:75-78

Lichter DG, Jackson A, Schachter M (1995) Clinical evidence of genomic imprinting in Tourette's syndrome. Neurology 45:924-928

Lichter P, Bray P, Ried T, Dawid IB, Ward DC (1992) Clustering of C2-H2 zinc finger motif sequences within telomeric and fragile sites regions on human chromosomes. Genomics 13:999-1007

Michels VV(1985) Fragile sites on human chromosomes: description and clinical significance. Mayo Clin Proc 60:690-696

Miro R, Clemente IC, Fuster C, Egozcue J (1987) Fragile sites, chromosome evolution and human neoplasia. Hum Genet 75:345-349

Mitchell PJ, Tijian R (1989) Transcriptional regulation in mammalian cells by sequence specific DNA binding proteins. Science 245:371-378

Morgan R, Morgan SS, Hecht BK, Hecht F (1988) Fragile sites at 4q23 and 7q11.23 unique to bone marrow cells. Cancer Genet Cytogenet 31:47-53

Murano I, Kuwano A, Kajii T (1989a) Fibroblast-specific common fragile sites induced by aphidicolin. Hum Genet 83:45-48

Murano I, Kuwano A, Kajii T(1989b) Cell type-dependent difference in the distribution of aphidicolin-induced fragile sites: T and B lymphocytes and bone marrow cells. Hum Genet 84:71-74

Nancarrow JK, Kremer E, Holman K, Eyre H, Doggett NA, LePaslier D, Callen DF, Sutherland GR, Richards RI (1994) Implications of FRA16A structure for the mechanism of chromosomal fragile site genesis. Science 264:1938-1941

Navas TA, Zhou Z, Elledge SJ (1995) DNA polymerase links the DNA replication machinery to the S phase checkpoint. Cell 80:29-39

Nussbaum RL, Airhart SD, Ledbetter DH (1986) Recombination and amplification of pyrimidine-rich sequences may be responsible for initiation and progression of the Xq27 fragile site: an hypothesis. Am J Med Genet 23:715-721

Obe G, Ristow H (1979) Mutagenic, carcinogenic and teratogenic effects of alcohol. Mutat Res 65:229-259

Parrish JE, Oostra BA, Verkerk AJMH, Richards CS, Reynolds J, Spikes AS, Shaffer LG, Nelson DL (1994) Isolation of a GCC repeat showing expansion in FRAXF, a fragile site distal to FRAXA and FRAXE. Nature Genet 8:229-235

Porfirio B, Tedeschi B, Vernole P, Caporossi D, Nicoletti B (1989) The distribution of MspI-induced breaks in human lymphocyte chromosomes and its relationship to common fragile sites. Mutat Res 213:117-124

Rappold GA (1993) The pseudoautosomal regions of the human sex chromosomes. Hum Genet 92:315-324

Rich A, Nordheim A, Wang AHJ (1984) The chemistry and biology of left-handed Z-DNA. Annu Rev Biochem 53:791-856

Saccone S, De Sario A, Della Valle G, Bernardi G (1992) The highest gene concentration in the human genome are in telomeric bands of metaphase chromosomes. Proc Natl Acad Sci USA 89:4913-4917

Samadder P, Evans JA, Chudley AE (1993) Segregation analysis of rare autosomal folate sensitive common fragile sites. Am J Med Genet 46:165-171

Smeets DFCM, Merkx G (1990) Neither age nor sex influence the expression of folate sensitive common fragile sites on human chromosomes. Hum Genet 86:76-78

Sutherland GR (1985) The enigma of the fragile X chromosome. Trend Genet 1:108-112

Sutherland GR, Hecht F (1985) Fragile sites on human chromosomes. Oxford University Press, New York

Sutherland GR (1988) The role of nucleotides in human fragile site expression. Mutat Res 200:207-213

Tedeschi B, Porfirio B, Vernole P, Caporossi D, Dallapicolla B, Nicoletti B (1987) Common fragile sites: their prevalence in subjects with constitutional and acquired chromosomal instability. Am J Med Genet 27:471-482

Tedeschi B, Vernole P, Lucia Sanna M, Nicoletti B (1992) Population cytogenetics of aphidicolin-induced fragile sites. Hum Genet 89:543-547

Tsuji H, Hitomi A, Takahashi E, Murata M, Ikeuchi T, Yamamoto K, Tsuji S, Hori T (1991) Induction of distamycin A-inducible rare fragile sites and increased sister chromatid exchanges at fragile sites. Hum Genet 87:254-260

Verkerk AJMH, Pieretti M, Sutcliff JS, Fu YH, Kuhl DPA, Pizzuti A, Reiner O, Richards S, Victoria MF, Zhang F, Eussen BE, Ommen GJB van, Blonden LAJ, Riggins GJ, Chastian JL, Kunst CB, Galjaard H, Caskey CT, Nelson DL, Oostra BA, Warren ST (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

Vogt P (1990) Potential genetic functions of tandem repeated DNA sequence blocks in the human genome are based on a highly conserved "chromatin folding code". Hum Genet 84:301-336

CHAPTER 5

Am. J. Hum. Genet. 63:839-846, 1998

Identification of Genetic Markers Associated with Gilles de la Tourette Syndrome in an Afrikaner Population

Ingrid Simonic,[1] George S. Gericke,[1] Jurg Ott,[2] James L. Weber[3]

[1]MRC Neurogenetics Research Laboratory, 1 Soutpansberg Rd., 0001 Arcadia, SA; [2]Rockefeller University, 1230 York Ave., New York, NY 10021-6399, USA; [3]Center for Medical Genetics, 1000 North Oak Ave., Marshfield, WI 54449, USA

Summary

Because gene-mapping efforts, using large kindreds and parametric methods of analysis, for the neurologic disorder Tourette syndrome have failed, efforts are being redirected toward association studies in young, genetically isolated populations. The availability of dense marker maps makes it feasible to search for association throughout the entire genome. We report the results of such a genome scan using DNA samples from Tourette patients and unaffected control subjects from the South African Afrikaner population. To optimize mapping efficiency, we chose a two step strategy. First we screened pools of DNA samples from both affected and control individuals, using a dense collection of 1,167 short tandem-repeat polymorphisms distributed throughout the genome. Second, we typed those markers displaying evidence of allele frequency-distribution shifts, along with additional tightly linked markers, using DNA from each affected and unaffected individual. To reduce false positives, we tested two independent groups of case and control subjects. Strongest evidence for association (P values 10^{-2} to 10^{-5}) were obtained for markers within chromosomal regions encompassing D2S1790 near the chromosome 2 centromere,

D6S477 on distal 6p, D8S257 on 8q, D11S933 on 11q, D14S1003 on proximal 14q, D20S1085 on distal 20q, and D21S1252 on 21q.

Introduction

Gilles de la Tourette syndrome (MIM 137580) is a childhood-onset neurological disorder characterized by chronic, involuntary motor and vocal tics. Both twin and family studies have indicated a high degree of heritability of Tourette syndrome, especially when individuals with chronic motor tics are considered affected (Patel 1996). Concordance rates, among MZ twins, of 53% and 56% for full-blown Tourette syndrome and 77% and 94% for Tourette syndrome plus chronic motor tics have been reported (Price et al. 1985). Earlier segregation analyses indicated an autosomal dominant mode of inheritance, with reduced penetrance (Comings et al. 1984, Pauls and Leckman 1986, Eapen et al. 1993). More-recent studies, however, have indicated a more complex mode of inheritance (Hasstedt et al. 1995, Walkup et al. 1996).

Tourette syndrome has been reported to cosegregate with several neuropsychiatric disorders. A connection between Tourette syndrome and obsessive-compulsive disorder is generally accepted (Pauls et al. 1991, 1995). Connection between Tourette syndrome and obsessive-compulsive disorder is generally accepted (Pauls et al. 1991, 1995). Conections between Tourette syndrome and attention deficit disorder and between other disorders such as alcoholism, panic attacks, and conduct disorders are controversial (Comings and Comings 1988, Pauls et al. 1988, Comings 1994).

A number of large Tourette kindreds have been identified (Kurlan et al. 1986, Robertson and Trimble 1991, McMahon et al. 1992, Hasstedt et al. 1995, Heutink et al. 1995). Several of these kindreds have undergone whole-genome polymorphism screening. Parametric linkage analysis of the resulting data, under the assumption of autosomal dominant inheritance, failed to locate genes (Pakstis et al. 1991, Wilkie et al. 1992, van de Wetering and Heutink 1993). Reasons for the failure are unknown but may be related to the wide range of disease severity within Tourette kindreds and/or to

misspecification of the mode of inheritance. Associations between three dopamine receptor genes and Tourette syndrome have been reported (Comings et al. 1993, Nöthen et al. 1994, Grice et al. 1996), but replication has not been achieved (Hebebrand et al. 1997).

The South African Afrikaner population arose largely from a small group of European, primarily Dutch, immigrants who began settling in Cape Town in 1652 (Torrington et al. 1984; Jenkins 1990; see also Government Communication and Information System [South Africa]). By 1701, there were 1,265 Europeans at the Cape. In the 18th century, European immigration to South Africa was small. In the first half of the 19th century, particularly in the 1830s, the Afrikaners began the "Great Treks," a series of migrations of Europeans away from the coastal areas, to establish farming communities in the interior of South Africa. The small number of European founders of the Afrikaner population at Cape Town, the lack of large waves of European immigration, the Great Treks, and the relatively rapid increase in the Afrikaner population, today at ~3 million, all have contributed to the genetic isolation and relative homogeneity of this population. Common mutations and/or marker haplotypes among Afrikaners have been established for a number of disorders, including variegate porphyria, keratolytic winter erythema, hypercholesterolemia, and progressive familial heart block (Jenkins 1990, 1996, Meissner et al. 1996, Warnich et al. 1996, Starfield et al. 1997, Groenewald et al. 1998).

We report here our efforts to map Tourette syndrome genes, using Afrikaner patients. Starting with a whole-genome screen of >1,000 polymorphisms, we have identified several markers that show significant differences in allele frequency distributions between affected and control individuals.

Subjects, Material and Methods

Patients

Probands were selected at random, on consecutive clinic days, from a group of existing patients at the Tourette Syndrome Clinic in Pretoria, South Africa. All subjects were previously identified as Tourette syndrome patients according to criteria of *the Diagnostic and Statistical Manual of Mental Disorders*. Motor and vocal tics were in the moderately severe to severe range in all affected individuals. Tics started at <18 years, occurred for ≥1 year, and were not absent for any periods >3 months. Prior to their recruitment, probands and available family members were evaluated personally by two educational psychologists for history of tics and presence and severity of comorbid problems. The following diagnostic tools were used: a self-report form designed by the Tourette Syndrome Association Genetic Consortium (January 1995 version), the Yale Global Tic Severity Scale (Leckman et al. 1989), observation/examination methods for the purpose of excluding other movement or neurological disorders and for confirming observable tics typical of those associated with Tourette syndrome, and the *Stony Brook Psychiatric Exclusion Checklist* (Gadow and Sprafkin 1998). English language versions of all forms were used, as most Afrikaners, including all patients, control subjects, and evaluators, were bilingual.

Forty Tourette syndrome patients were recruited in the initial phase of the study. For the follow up investigations an additional sixty unrelated patients agreed to participate (second group). All study subjects spoke Afrikaans and had Afrikaner family names but were not known to belong to any Afrikaner subgroup. Individuals of ascertainable English descent were excluded. The first and second groups of patients (and control subjects) were selected from the same population base. All blood samples were coded and made anonymous prior to DNA isolation. Only the information on age, sex, tics severity range, and comorbid behavioral problems remained attached to each sample. The study was approved by the South African Medical Research Council Review Board.

Controls

The Afrikaner individuals (n = 96) included in the control group came from two sources: clinic personnel and medical students. Like the patients, all control subjects spoke Afrikaans and had Afrikaner family names. Control subjects were required to complete the same selfreport form used by the patients and, in addition, were personally evaluated by one of us (G.S.G.) for the presence of chronic motor and vocal tics. Their blood samples were coded and made anonymous prior to DNA isolation.

Primers

Primer pairs for the detection of short tandem-repeat polymorphisms (STRPs) were obtained from Research Genetics. Approximately 40% of the markers were tri- and tetranucleotide-repeat polymorphisms developed within the Cooperative Human Linkage Center, and nearly all remaining 60% Généthon dinucleotide-repeat polymorphisms. We determined allele sizes, using known genotypes of three CEPH individuals: 133101, 133102 and 134702. Marker spacing was determined and close flanking markers selected by use of comprehensive sex-averaged genetic maps produced by the program CRI-MAP (Lander and Green 1987, Broman et al. [in press], see also Center for Medical Genetics).

Sampling and DNA Pooling

Genomic DNA was isolated from individual blood samples by standard methods (Sambrook et al. 1989). DNA concentrations were measured by spectrophotometric readings at OD_{260}. Pools of template DNA were prepared by combining 1μg of DNA from each of 20 individuals from the first group of subjects. Two non-overlapping pools were prepared from the affected individuals (A1 and A2) and from the controls (C1 and C2). The pooled DNA was diluted to a final concentration of 20ng/μl.

PCR

Amplification of STRP markers was performed in 96-well microtiter plates with 45 ng of either pooled DNA or individual DNA in a 10-µl volume containing 1.5 mM $MgCl_2$; 50 mM KCl; 10 mM Tris-HCl, pH 8.3; 0.01% (w/v) gelatin; 200 µM each dGTP, dATP, and dTTP; 2.5 µM dCTP; 0.35 µCi $\alpha[^{32}P]$-dCTP (NEN Du Pont; 800 Ci/mmole, 10 µCi/µl); 10ng (~3 pmol) each PCR primer; and 0.3 U *AmpliTaq* polymerase (Boehringer Mannheim). In most cases, two markers were amplified simultaneously. Samples were subjected to 27 cycles consisting of 30 sec at 94°, 75 sec at 55°, and 15 sec at 72°, with a final 6 min at 72° after the last cycle. PCR products were denatured by adding formamide and by heating for 10 min at 95° prior to loading (1µl) onto vertical 6.5% polyacrylamide, 7.7 M urea DNA sequencing gel and running at 70 W constant power for ~3 hours. Gels were dried on filter paper and exposed on X-ray film.

Statistical analysis

The following statistics were calculated to compare affected and control allele frequency distributions for each polymorhic marker:

"Heterogeneity" approach. - For the *m* alleles at a given marker, an *m* by *2* contingency table was formed, with columns corresponding to alleles in case and control individuals. Rows were ordered by increasing allele size, and rows containing cells with expected values <1 were pooled with neighboring rows. From a χ^2 analysis, empirical (two-sided) significance levels were obtained. This approach is expected to be powerful when several alleles occur in different frequencies in the case and control subjects.

"Single alleles" approach. - A given allele, *i*, was selected and all other alleles combined into a second category, *not i*. The resulting 2 by 2 table was analyzed by use of Fisher's exact test. The smallest (one-sided) P-value was selected (and the allele identified at which it occurred) and adjusted for multiple comparisons by means of a

Bonferroni correction. This approach is expected to be powerful when a single allele shows association. One-sided P values were used because, when the frequency of a given allele is increased in case versus control subjects, some other allele(s) must necessarily be decreased. When one allele is tested after another, each allele should thus be tested only for positive association.

"t-test" approach. - Mean allele sizes for case and control subjects were compared with use of the t-test. Empirical p-values (two-sided) were calculated numerically. This approach is expected to be powerful when allele sizes are shifted in a constant direction between case and control subjects.

Table 1 Chromosomal Breakdown of Markers Used in Whole-Genome Screen

Chromosome	No. of STRPs	Chromosome	No. of STRPs
1	96	13	41
2	91	14	48
3	60	15	31
4	65	16	35
5	78	17	41
6	69	18	47
7	66	19	24
8	67	20	32
9	39	21	14
10	43	22	17
11	44	X	52
12	67		

Results

A first group of 40 Afrikaners clearly affected with Tourette syndrome, along with 40 unaffected Afrikaner controls, were recruited. DNA was pooled from two nonoverlapping sets of 20 individuals for both affected (A1 and A2) and control (C1 and C2) subjects. Two sets of pools were chosen, to reduce the false-positive rate due to slight differences in DNA concentrations and pipetting inaccuracies. The pooled DNA, along with standard DNA from the parents of CEPH family 1331, was used as template in the PCR amplification of 1,167 STRPs distributed throughout the entire genome. Chromosomal breakdown of the STRPs is shown in table 1. Sex-averaged spacing between STRPs was 3.0 ± 3.2 cM (mean ± SD). Ninety-five percent of the intervals between markers were <9.0 cM. Fifteen intervals were >15 cM, with the very largest at 22 cM.

Visual examination of the autoradiographic images of pooled PCR products revealed 51 loci with consistent differences in allele distributions between the two affected and two control DNA pools. Results from six of these loci are displayed in figure 1. Criteria for selection of the positive loci were (1) consistency of observed differences between affected and control subjects as determined by repeat PCR and by use of two independent sets of control and affected pools, and (2) prevalence of one allele in pooled PCR products from the affected pools (indicated by arrows in fig. 1). Amplification of the vast majority of the markers resulted in indistinguishable radiographic images between affected and control pools. For a small proportion of STRPs, differences between affected and control pools were detected for A1/C1 or A2/C2, but not for both. This may be the consequence of pooling DNA from only 20 individuals.

To evaluate the statistical significance of the observed differences in pooled PCR products, we subjected the 51 putative positive markers to individual typing, using the same group of 40 affected and 40 control DNA samples used for the initial genomic screen. We than compared the resulting pairs of allele frequency distributions, using three different statistical approaches (see the Subjects, Material, and Methods section).

For 15 of the 51 loci, at least one of the calculated statistics showed significance at the .05 level (table 2). For the remaining 36 markers, apparent differences between affected and control subjects were not confirmed at the .05 level.

In an attempt to reproduce these initial results, a second group of unrelated Afrikaners (60 affected and 56 control individuals) were recruited, and their DNA was typed individually with the 15 positive markers from the first phase of the study. Results of the typings for the second group are also listed in table 2. For the majority of loci, evidence for association was not confirmed. However, in four instances (D2S1790, D6S477, D11S933, D20S1085), the P-values for the second group also reached significance and were even lower than those for the first group. For all 15 markers, P (heterogeneity) values from the first and second groups were then combined by summing χ^2 and df values (table 2). This approach of combining results was chosen as a safeguard, to eliminate potential effects of sizing alleles differently in the two samples. Combining results increased support for association in nearly every case relative to the results from either the first or second group of samples.

To extend these results, we typed a number of markers closely linked to positive markers from the whole-genome scan, using both groups of samples jointly. Markers at several of these loci gave significant results, including D2S440 adjacent to D2S1790, D8S257 and D8S1132 close to D8S1119, D11S1377 adjacent to D11S933, D20S468 and D20S469, both very close to D20S1085, and GATA45C03 near D21S1252 (table 3). Results for D11S1377 were particularly impressive. In addition, two tightly linked markers on chromosome 14 also produced significant results (table 3). The marker at locus D14S742 was identified in the first stage of the study but did not quite yield P values <0.05 for the first set of samples. No additional markers with significant allele distribution differences were identified from chromosomes 1, 4, 5, 6, 12, or 13. Allele frequency distributions for D11S1377 and GATA45C03 are displayed in table 4. Note, that for each marker, more than two common alleles were enriched in the affected subjects compared with control subjects.

Figure 1

Electrophoretic profiles for markers that demonstrated consistent allele frequency distribution between affected and control subjects in the first group. Segments of autoradiographs from polyacrylamide gels are displayed for six STRPs at the indicated loci. The DNA templates used to generate the amplified DNA fragments were in the same order for each marker: standard DNA from CEPH family parents 133101 (1) and 133102 (2), DNA pool from set 1 of unaffected control subjects (C1), DNA pool from set 1 of affected subjects (A1), independent amplification of pools C1 and A1, DNA pool from set 1 of affected subjects (A1), independent amplification of pools C1 and A1, DNA pool from set 2 of unaffected control subjects (C2), and DNA pool from set 2 of affected subjects (A2). Arrows mark alleles enriched in the affected subjects. Sizes (in nucleotides) for the enriched alleles are listed on the left.

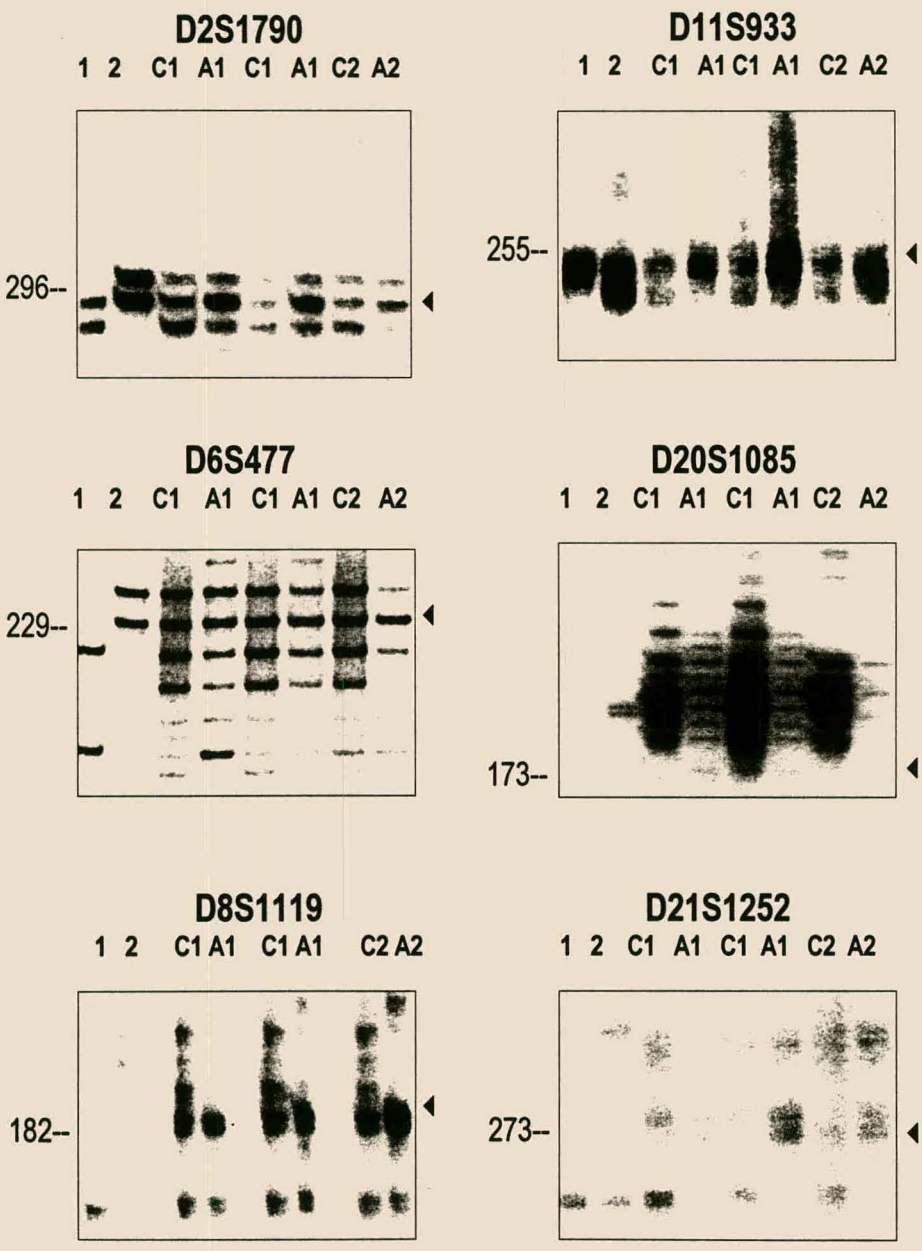


Table 2

Loci associated with Tourette syndrome in first and second groups of affecteds

		FIRST GROUP OF AFFECTEDS			SECOND GROUP OF AFFECTEDS			FIRST & SECOND GROUPS COMBINED
LOCUS	cM	P(het)	P(single alleles)	P(t-test)	P(het)	P(single alleles)	P(t-test)	P(het)
D01S0485	176.7	0.02	0.005	0.12	0.25	0.24	0.22	0.02
D01S1665	136.2	0.01	0.001	0.04	0.12	0.06	0.45	0.008
D02S1391	227.9	0.04	0.16	0.41	0.14	0.1	0.29	0.03
D02S1790	113.4	0.01	0.04	0.03	0.006	0.01	0.86	0.0006
D04S1551	40.7	0.01	0.02	0.12	0.42	0.64	0.34	0.03
D05S0666	176.4	0.03	0.02	0.07	0.19	0.62	0.92	0.03
D06S0477	9.2	0.005	0.06	0.08	0.02	0.04	0.41	0.0009
D06S0470	17.9	0.0006	0.02	0.001	0.43	0.21	0.13	0.004
D08S1119	133.4	0.03	0.06	0.66	0.09	0.05	0.82	0.01
D11S0933	158.1	0.05	0.006	0.75	0.002	0.009	0.38	0.0009
D12S0327	126	0.001	0.1	0.04	0.38	0.3	0.63	0.006
D13S0788	50.7	0.005	0.16	0.5	0.56	0.43	0.95	0.02
D20S1085	119.7	0.03	0.03	0.65	0.0005	0.0004	0.76	0.0001
D21S1252	60.2	0.00001	0.006	0.42	0.07	0.12	0.36	0.000008
D21S1435	41.9	0.04	0.05	0.03	0.46	0.17	0.24	0.09

NOTE. – het = heterogeneity. P values indicate the probability of obtaining a difference as large or larger than the one observed if there is no association between disease and marker. Values were calculated by means of three different approaches, as described in the Subjects, Material, and Methods section.

Table 3 Additional Loci in Selected Chromosomal Regions Associated with Tourette Syndrome

LOCUS	cM	*P* VALUE FOR APPROACH Heterogeneity	Single Alleles	t-test
D02S0440	113.4	0.002	0.001	0.54
D08S0257	145	0.01	0.004	0.74
D08S1132	155	0.05	0.1	0.33
D11S1377	119.2	<10-6	0.0005	0.01
D14S0742	12.5	0.06	0.02	0.03
D14S1003	12.5	0.002	0.003	0.00003
D20S0468	121.3	0.05	0.06	0.005
D20S0469	121.9	0.1	0.06	0.66
GATA45C03 chromosome 21	31.3	0.0004	0.004	0.04

NOTE. – Results were obtained with DNA from affected and control individuals in both first and second groups. See note for table 2.

Table 4 Allele Frequency Distributions for Two Markers with Especially Low *P* Values

Mfd 316 at D11S1377			GATA45C03 at chromosome 21		
Allele	Affected	Control	Allele	Affected	Control
124	1	0	266	0	3
128	4	0	268	62	66
130	4	7	286	1	0
132	2	6	288	10	5
134	45	20	289	6	5
136	62	42	291	32	47
138	22	52	302	1	3
140	6	19	305	4	14
142	16	5	308	15	16
144	4	9	310	28	19
146	8	6	312	14	1
148	0	6	314	3	0
			318	0	1

NOTE. – Results were obtained with DNA from affected and control individuals in both first and second groups.

Discussion

The primary goal of this study was to identify loci associated with Gilles de la Tourette syndrome in the Afrikaner population, using the straightforward comparison of polymorphic allele frequency distributions between severely affected probands and unaffected control subjects. We believe we have accomplished this goal for the markers at the loci listed in tables 2 and 3.

Several groups have identified linkage disequilibrium over broad chromosomal intervals in isolated populations. Disequilibrium has readily been detected over intervals ranging up to 15 cM in the Finish population (Peltonen and Uusitalo 1997), which is likely to be considerably older than the Afrikaners. Among populations thought to be similar in age to the Afrikaners (~12 generations), Houwen et al. (1994) identified disequilibrium over a 19-cM interval in patients with intrahepatic cholestasis from an isolated Netherlands fishing village, and Puffenberger et al. (1994) found disequilibrium over at least 10 cM among American Mennonites with Hirschsprung disease. Among Afrikaners, Starfield et al. (1997) detected disequilibrium over 10 cM for keratolytic winter erythema, and Groenewald et al. (1998) recently reported shared haplotypes extending over ~17 cM for variegate porphyria. Therefore our use of a highly informative marker density of 3.0 cM gave us a good chance of successfully detecting association among Afrikaner Tourette patients.

The use of DNA pools in the whole-genome screen dramatically reduced the amount of required laboratory work. Allele frequency distributions obtained from pooled DNA templates have been found to match, reasonably well, those determined by typing individuals (Pacek et al. 1993, Graff et al. 1997). Pooling approaches are rapidly becoming standard for the mapping of rare recessive disorders within isolated populations (Sheffield et al. 1994, Peltonen and Uusitalo 1997). Although careful scanning of the electrophoretic profiles of the amplified fragments might be ideal (Barcellos et al. 1997, Graff et al. 1997), visual selection of markers with differences in allele frequency distributions is efficient and has worked well for recessive

disorders. The fact that only 15 of 51 loci that appeared positive visually were confirmed by typing individuals indicated that we were generous in our selection of candidate loci and that, although false-negative loci cannot be completely ruled out, true-positive loci were not easily missed.

Analysis of the amplified fragments from the DNA pools gave no indication that a single predominant Tourette gene exists among the Afrikaners or, that Tourette genes were introduced into the Afrikaner population by a single founder. Although specific alleles at various loci were clearly enriched in affected versus control subjects, no single predominant alleles were found at any loci (see, for example, table 4). Therefore, even in the isolated Afrikaner population, several founding Tourette alleles at several loci likely exist.

Since no one yet knows how many Tourette genes exist, how they interact, how many Afrikaner founders introduced Tourette genes into this population, or when the genes might have been introduced, it is not possible to project how many patients would be required to detect association and what P values would be significant (Kruglyak 1997). It is likewise impossible to completely rule out subtle, undetected population differences between the patient and the control groups. Nevertheless, we feel that the P-values displayed in tables 2 and 3 are very promising. Given 1,167 markers tested, and on the basis of the conservative Bonferroni correction for multiple testing, we would by chance expect 0.1 results significant at the 10^{-4} level. In fact, we found four such markers on chromosomes 11, 14, 20, and 21. Our use of repeat PCR (testing pools A1 and C1 twice, as shown in the figure) and separate pools of affected and control individuals (pools A1, A2, C1, and C2), the confirmation of results using an entire separate group of Tourette patients, and the typing of additional STRPs tightly linked to the original positives were all designed to reduce the possibility of false positives.

We assert that the only practical route to isolation of Afrikaner Tourette genes is to pursue loci that give strong evidence for association and that eventually are confirmed by independent studies. In this regard, Leppert and McMahon, in their whole-genome-

polymorhism screen of unilineal branches of a very large Utah Tourette kindred (McMahon et al. 1992), obtained some of their strongest positive LOD scores for several markers (including D8S257) in the exact same region of chromosome 8q as that identified in this study (unpublished results). Also, Devor and Magee (in press) very recently reported a family in which individuals with Tourette syndrome or tics showed segregation with a balanced chromosome 1 - 8 translocation t[(1:8)(q21.1;q22.1)]. D8S257 is located at or very close to 8q22.1 (Bray-Ward et al. 1996). We plan to continue our efforts among the Afrikaners by collecting parents of affected individuals, so that shared haplotypes can be identified, and by collecting Tourette families from South African Gereformeerde Church members who are themselves a subgroup of Afrikaners (Torrington et al. 1984).

Acknowledgments

We thank all patients participating in the study and members of the Tourette Syndrome supportive group in Pretoria for their cooperation, fruitful discussions at meetings, and private financial support of the study. The work was also supported by the Human Genome Organization (HUGO Europe) travel grant (to I. S.), by National Human Genome Research Institute grant HG0008 (to J. O.), and by Marshfield Medical Research Foundation. Karl Broman provided statistical analysis of marker density, and Donna David assisted in the laboratory.

Electronic-Database Information

Accession numbers and *urls* for data in this article are as follows:

Center for Medical Genetics, http://www.marshmed.org/genetics (for polymorphism and genetic map information);

Cooperative Human Linkage Center, http://www.chlc.org (for polymorphism information);

Généthon, http://www.genethon.fr (for polymorphism information);

Government Communication and Information System [South Africa], (for Afrikaner history) http://www.gcis.gov.za/level12/history.htm ;

Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim (for Tourette syndrome) [MIM 137580]

References

Barcellos LF, Klitz W, Field LL, Tobias R, Bowcock AM, Wilson R, Nelson MP, et al. (1997) Association mapping of disease loci, by use of a pooled DNA genomic screen. Am J Hum Genet 61:734-747

Bray-Ward P, Menninger J, Lieman J, Desai T, Mokady N, Banks A, Ward DC (1996) Integration of the cytogenetic, and physical maps of the human genome by FISH mapping of CEPH YAC clines. Genomics 32:1-14

Broman KW, Murray JC, Sheffield VC, White RL, Weber JL Comprehensive human genetic maps: individual and sex-specific variation in recombination. Am J Hum Genet (in press)

Comings DE (1994) Tourette syndrome: a hereditary neuropsychiatric spectrum disorder. Ann Clin Psychiatry 6:235-247

Comings DE, Comings BG (1988) A controlled study of Tourette syndrome – revised: a reply to the letter of Pauls et al. Am J Hum Genet 43:209-217

Comings DE, Comings BG, Devor EJ, Cloninger CR (1984) Detection of major gene for Gilles de la Tourette syndrome. Am J Hum Genet 36:586-600

Devor EJ, Magee H. Multiple childhood behavioral disorders (Tourette syndrome, multiple tics, ADHD, and OCD) presenting in a family with a balanced chromosome translocation t(1;8)(q21.1; q22.1). Psychiatric Genet (in press)

Diagnostic and statistical manual of mental disorders (1994) 4th ed. American Psychiatric Association, Washington, DC

Eapen V, Pauls DL, Robertson MM (1993) Evidence for autosomal dominant transmission in Tourette's syndrome. United Kingdom Cohort Study. Br J Psychiatry 162:593-596

Gadow KD, Sprafkin J (1998) Child symptom inventory – 4, screening manual. Checkmate Plus, Stony Brook, NY

Graff C, Persson A, Ulfendahl PJ, Wahlberg J, Wadelius C (1997) Genetic mapping using fluorescent quantification of allele frequencies in pooled DNA loaded by solid support. Clin Genet 51:145-152

Grice DE, Leckman JF, Pauls DL, Kidd KK, Pakstis AJ, Chang FM, Buxbaum JD, et al (1996) Linkage disequilibrium between an allele at the dopamine D4 receptor locus and Tourette syndrome, by the transmission-disequilibrium test. Am J Hum Genet 59:644-652

Groenewald JZ, Liebenberg J, Groenewald IM, Warich L (1998) Linkage disequilibrium analysis in a recently founded population: evaluation of the variegate porhyria founder in South African Afrikaners. Am J Hum Genet 62:1254-1258

Hasstedt SJ, Leppert M, Filloux F, van de Wetering BJM, McMahon WM (1995) Intermediate inheritance of Tourette syndrome, assuming assortative mating. Am J Hum Genet 57:682-689

Hebebrand J, Nöthen MM, Ziegler A, Klug B, Neidt H, Eggerman K, Lehmkuhl G, et al. (1997) Nonreplication of linkage disequilibrium between the dopamine D4 receptor locus and Tourette syndrome. Am J Hum Genet 61:238-239

Heutink P, van de Wetering BJM, Pakstis AJ, Kurlan R, Sandor P, Oostra BA, Sandkuijl LA (1995) Linkage studies on Gilles de la Tourette syndrome: What is the strategy of choice? Am J Hum Genet 57:465-473

Houwen RHJ, Baharloo S, Blankenship K, Raeymaekers JJ, Sandkuijl LA, Freimer NB (1994) Genome screening by searching for shared segments: mapping gene for benign recurrent intrahepatic cholestasis. Nat Genet 8:380-386

Hyde TM, Aaronson BA, Randolph C, Rickler KC, Weinberger DR (1992) Relationship of birth weight to the phenotypic expression of Gilles de la Tourette syndrome in monozygotic twins. Neurology 42:652-658

Jenkins T (1990) Medical genetics in South Africa. J Med Genet 27:760-779

Jenkins T (1996) The South African malady. Nature Genetics 13:7-9

Kruglyak L (1997) What is significant in whole-genome linkage disequilibrium studies? Am J Hum Genet 61:810-812

Kurlan R, Behr J, Medved L, Shoulson I, Pauls D, Kidd JR, Kidd KK (1986) Familial Tourette's syndrome: report of a large pedigree and potential for linkage analysis. Neurology 36:772-776

Lander ES, Green P (1987) Construction of multilocus genetic linkage maps in humans. Proc Natl Acad Sci USA 84:2363-2367

Leckman JF, Riddle MA, Hardin MT, Ort SI, Swartz KL, Stevenson J, Cohen DJ (1989) The Yale Global Tic Severity Scale: initial testing of a clinician-rated scale of tic severity. J Am Acad Child Adolesc Psychiatry 28:566-573

McMahon WM, Leppert M, Filloux F, van de Wetering BJM, Hasstedt S (1992) Tourette symptoms in 161 related family members. In: Chase TN, Friedhoff AJ, Cohen DJ (eds) Advances in neurology, Vol 58. Raven Press, New York, pp 159-165

Meissner PN, Dailey TA, Hift RJ, Ziman M, Corrigall AV, Roberts AG, Meissner DM, Kirsch RE, Dailey HA (1996) A R59W mutation in human protoporphyrinogen oxidase results in decreased enzyme activity and is prevalent in South Africans with variegate porphyria. Nature Genetics 13:95-97

Nöthen MM, Hebebrand J, Knapp M, Hebebrand K, Camps A, von Gontard A, Wettke-Schäfer R, et al. (1994) Association analysis of the dopamine D2 receptor gene in Tourette's syndrome using the haplotype relative risk method. Am J Med Genet 54:249-252

Pacek P, Sajantila A, Syvänen AC (1993) Determination of allele frequencies at loci with length polymorphism by quantitative analysis of DNA amplified from pooled samples. PCR Methods and Applications 2:313-317

Pakstis AJ, Heutink P, Pauls DL, Kurlan R, van de Wetering BJM, Leckman JM, Sandkuyl LA (1991) Progress in the search for genetic linkage with Tourette syndrome: an exclusion map covering more than 50% of the autosomal genome. Am J Hum Genet 48:281-294

Patel PI (1996) Quest for the elusive genetic basis of Tourette Syndrome. Am J Hum Genet 59:980-982

Pauls DL, Alsobrook MP II, Goodman W, Rasmussen S, Leckman JF (1995) A family study of obsessive compulsive disorder. Am J Psychiatry 152:76-84

Pauls DL, Cohen DJ, Kidd KK, Leckman JF (1988) Tourette syndrome and neuropsychiatric disorders: is there a genetic relationship? Am J Hum Genet 43:206-209

Pauls DL, Raymond CL, Stevenson JM, Leckman JF (1991) A family study of Gilles de la Tourette syndrome. Arch Gen Psychiatry 42:815-820

Peltonen L, Uusitalo A (1997) Rare disease genes – lessons and chalanges. Genome Res 7:765-767

Price RA, Kidd KK, Cohen DJ, Pauls DL, Leckman JF (1985) A twin study of Tourette syndrome. Arch Gen Psychiatry 42:815-820

Puffenberger EG, Kauffman ER, Bolk S, Matise TC, Washington SS, Angrist M, Weissenbach J, et al. (1994) Identity-by-descent and association mapping of a recessive gene for Hirschsprung disease on human chromosome 13q22. Hum Mol Genet 3:1217-1225

Robertson MM, Trimble MR (1991) Gilles de la Tourette syndrome in the Middle East: report of a cohort and a multiply affected large pedigree. Br J Psychiatry 158:416-419

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual, vol 1. 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

Sheffield VC, Carmi R, Kwitek-Black A, Rokhlina T, Nishimura D, Duyk GM, Elbedour K, et al. (1994) Identification of a Bardet-Biedl syndrome locus on chromosome 3 and evaluation of an efficient approach to homozygosity mapping. Hum Mol Genet 3:1331-1335

Starfield M, Hennies HC, Jung M, Jenkins T, Wienker T, Hull P, Spurdle A, et al. (1997) Localization of the gene causing keratolytic winter erythema to chromosome 8p22-p23, and evidence for a founder effect in South African Afrikaans-speakers. Am J Hum Genet 61:370-378

Torrington M, Gericke GS (1993) Ancestral links between current South African families with Tourette Syndrome. Biannual Congress of the South African Society of Human Genetics, Bloemfontein, South Africa

van de Wetering BJM, Heutink P (1993) The genetics of the Gilles de la Tourette syndrome: a review. J Lab Clin Med 121:638-645

Walkup JT, LaBuda MC, Singer HS, Brown J, Riddle MA, Hurko O (1996) Family study and segregation analysis of Tourette syndrome: evidence for a mixed model of inheritance. Am J Hum Genet 59:684-693

Warnich L, Meissner PN, Hift RJ, Louw JH, van Heerden CJ, Retief AE (1996) Mapping of Variegate Porphyria (VP) gene: contradictory evidence for linkage between VP and microsatellite markers at chromosome 14q32. Hum Genet 97:690-692

Wilkie PJ, Ahmann PA, Hardacre J, LaPlant RJ, Hiner BC, Weber JL (1992) Application of microsatellite DNA polymorphisms to linkage mapping of Tourette syndrome gene(s). In: Chase TN, Friedhoff AJ, Cohen DJ (eds) Advances in neurology, Vol 58. Raven Press, New York, pp 173-180

CHAPTER 6

South African Medical Journal (submitted for publication, Jun 1999)

THE SEARCH FOR TOURETTE SYNDROME GENES: AN OVERVIEW

Simonic I., Gericke G. S.

Neurogenetics Research Initiative, 1 Soutpansberg Rd., MRC Pretoria

Clinical description of Tourette syndrome

By working together in Paris in the 1880s on the classification of movement disorders, Jean Martin Charcot, the leading French neurologist at the time and his intern Georges Gilles de la Tourette were able to draw attention to a distinguished form of movement disorder that is of great interest today (Lajonchere et al. 1996) The major features of the disorder as described by Tourette in his 1885 article were involuntary movements and sounds, markedly enhanced startle reactions, a tendency to repeat both vocalizations (echolalia) and movements (echopraxia), and uncontrolable verbal obscenities (coprolalia). Tourette also believed that this childhood onset syndrome, later named after him, did not affect the senses or intellect, and that the condition was hereditary, variable in severity across the patient's life span, and incurable.

According to the Diagnostic and Statistical Manual of Mental Disorders, 4th ed. (DSM-IV) (American Psychiatric Association 1994), the diagnostic criteria for Gilles de la Tourette syndrome (GTS) include:

1. the presence of motor and one or more vocal tics at some time during the illness, although not necessarily concurrently;
2. occurrence of tics throughout a period of >1 year, with no tic-free period of >3 consecutive months;

3. marked distress or significant impairment in social, occupational, or other important areas of functioning;
4. onset at <18 years of age;
5. lack of identifiable environmental causes or other contributing medical conditions.

The primary symptom of GTS – tics - are defined as sudden, rapid, recurrent, non-rhythmic, stereotype movements or vocalization (Robertson and Stern 1997) and can range in affected individuals from simple tics, often around the eyes, simple clearing of the voice, or sniffing to more complex utterances (Robertson 1989). Although of involuntary nature, tics can wax and wane and even be suppressed. Other symptoms including echolalia, arithmomania, grunting, and coprolalia may develop, as the disease progresses. Although coprolalia is often regarded as the most notorious symptom associated with GTS, it has been proven rather infrequent (~ 8%) in the population of clinic patients (Goldenberg et al. 1994).

It has been documented that biological relatives of GTS patients more frequently exhibit obsessive-compulsive symptoms and chronic tics, than relatives in control families (Pauls et al. 1991). The finding supports the hypothesis that obsessive-compulsive disorder (OCD) and chronic tic disorder (CT) together with GTS are representing an alternative manifestation of the same underlying disorder. The distribution of OCD symptoms recognized in individuals with GTS (Frankel et al. 1986, Pauls et al. 1986, Kurlan 1989, Leonard et al. 1992, Swedo and Leonard 1994) and those more prevalent in pure OCD patients are, however, different. Moreover, comorbid expression of OCD appears to be sex-related with OCD by itself occurring more frequently in female relatives of GTS probands (Eapen et al. 1997a).

GTS has been observed in all studied populations (Staley et al. 1997) with the vast majority of affected individuals being able to lead normal, productive lives, and remaining undiagnosed because their symptoms do not require medical attention (Leckman et al. 1997). However, with increasing severity of either tics or associated behavioral problems such as hyperactivity, temper control, aggressive behavior, and obsessive compulsive symptoms, GTS may become a socially and psychologically devastating condition.

A strong genetic component has been determined for GTS by twin, and family based studies (Price et al. 1986, LaBuda et al. 1993) which provoked attempts for the elucidation of its genetic basis. So far no biochemical or morphological impairment has been consistently associated with GTS. The suspected implication of head of the nucleus caudatus in GTS, based on one SPECT tomographical study is currently under further investigation. Therefore, the first step in the elucidation of the genetics of GTS is the identification of a genetic map location of GTS underlying gene(s). This can be done via genome-wide search for linked/associated genetic loci, or via candidate gene approach and association of the functional polymorphism in known candidate genes with GTS phenotype.

Genetic models and linkage studies

The most powerful method in genome wide searches for genetic loci linked to a trait is standard lod score (linkage) analysis. However, in order to be successfully applied, it requires precise specification of the phenotype (diagnostic criteria) and genetic model, including mode of inheritance, penetrance (a proportion of individuals with specific genotype that show the expected phenotype), and the population prevalence of the condition (Lander and Schork 1994).

The diagnosis of GTS is a subject of ongoing debate and controversy (Patel 1996). Many scientists, attempting to elucidate the genetic basis of GTS expressed reservations about the DSM-IV criteria and proposed an alternative classification scheme, recognized by the Tourette Syndrome Association (TSA) Genetic Linkage Consortium (Kurlan 1997). The scheme emphasizes GTS as tic-based dysfunction, includes detailed exclusion criteria and a list of clinical features for increasing confidence for the diagnosis of GTS.

In order to specify the mode of inheritance and penetrance, complex segregation analysis were carried out in both nuclear families and extended GTS pedigrees. Earlier studies supported a mixed model of inheritance for GTS including a major semidominant gene with low heritability of the multifactorial background variation

(Comings et al. 1984). In later family studies evidence was found for the presence of rare, semidominant, incompletely penetrant allele considered to lead to expression of the disorder (Devor 1984), autosomal dominant pattern of inheritance (Pauls and Leckman 1986) or intermediate pattern of inheritance (Hasstedt et al. 1995). All family studies consistently rejected purely polygenic and autosomal recessive models of inheritance, with the latest accepted model indicating that the susceptibility for GTS is conveyed by an additive major locus in combination with a multifactorial background (Walkup et al. 1996).

When first described, GTS was considered to be rare. Currently, the population prevalence of about 4-5 per 10,000 is the most acknowledged estimate (Eapen et al. 1997b) with the risks to first-degree relatives of GTS individuals being about 200 times the population prevalence (Santangelo et al. 1996). The most recent prevalence estimates of GTS in mainstream school population, however, suggest that the mild form of GTS is much more common than previously thought with a prevalence rate of almost 3/100 in the age group of 13-14 years (Mason et al. 1998). The reason for such differences is based on whether one just considers fully expressing cases versus all those that fulfil diagnostic criteria in the community i.e. outside clinic settings.

All attempts to link genetic markers with GTS in systematic genome searches to date were unsuccessful. The failure, as discussed by number of authors (Devor 1990, Patel 1996, Robertson and Boardman 1996, Barr and Sandor 1998), could be due to ambiguous phenotypic delineation of GTS and, perhaps most importantly, a genetic model complicated by equivocal gene frequency, incomplete (reduced) penetrance, and bilineal transmission in GTS families (Kurlan et al. 1994). Moreover, sex-specific expression of GTS behaviors and sex-associated differences in genetic transmission consistent with possible role of genomic imprinting in phenotypic expression of GTS have been documented (Lichter et al. 1995, Eapen et al. 1997c).

The lack of success using a family based linkage approach has recently led researchers to the adoption of model-free, non-parametric methods which ignore unaffected people, and look for chromosomal segments that are shared by affected individuals (e.g. affected sib-pair studies). These methods do not require precise

specification of genetic model and concentrate on clearly affected individuals only. Another approach currently supported in a search for GTS gene(s) involves linkage disequilibrium (LD) mapping in historically isolated populations with greater genetic homogeneity.

Candidate genes and association studies

Among the human genes that have recently been characterized, no obvious candidates for GTS have been identified. Nevertheless, dopamine-receptor genes, which are very popular candidates for most neuropsychiatric disorders, because of the earlier observations of the therapeutic effects of dopamine inhibitor drugs, have been investigated for the association with GTS in numerous studies. An initial positive association of the A1 allele of the dopamine D2 receptor gene (DRD2) with GTS was subsequently disproved in several studies (Devor et al. 1990, Gelender et al. 1994, Nothen et al. 1994). The finding of positive linkage disequilibrium (LD) reported between the dopamine D4 receptor allele (DRD4*7R) and GTS (Grice et al. 1996) was also not replicated (Barr et al. 1996). All other known dopamine receptor genes as well as the dopamine beta hydroxylase gene and the tyrosine hydroxylase gene were excluded by classical linkage analysis in large pedigrees from playing a role in etiology of GTS (Brett et al. 1995a, Barr et al. 1997, Devor et al. 1998).

Serotonergic pathway disturbances have also been implicated in number of neuropsychiatric disorders including GTS, substance abuse and depression. Significantly lower levels of hydroxyphenyl glycol and serotonin have been reported after examination of urinary amines and their metabolites in GTS patients (Bornstein and Baker 1992) and serotoninergic agents have been successfully used for treatment of GTS symptoms, particularly tics (Silvestri et al. 1994). Numerous searches were performed in order to define an association between GTS and some of the variants of serotonin receptor genes with no success (Brett et al. 1995b, Gelender et al. 1995).

Other candidate gene studies led to the exclusion of the whole chromosome regions from linkage with GTS (Devor et al. 1991, Brett et al. 1997), or when positive awaiting independent confirmation (Gate et al. 1998).

Since no sufficient support was found for association of the most obvious candidate genes with GTS, and virtually all the genes, expressed in brain, might be good candidates, further candidate gene approaches can only be effective if proceeded by the identification of suggestive chromosome location of GTS gene(s). Such areas of interest may be identified by linkage/association studies, or by identification of a chromosomal abnormality co-segregating with disease phenotype in proband and his affected family members.

Chromosome abnormalities reported in probands with GTS

There are no characteristic chromosome abnormalities associated with GTS. Most clinical cases are therefore not referred for cytogenetic analysis. Nevertheless, some investigators continue searching for chromosome abnormalities in their GTS probands, because such finding may represent an additional valuable information about possible marker areas for genetic loci involved in the susceptibility to the trait.

Numerous seemingly nonspecific chromosomal abnormalities have been reported in GTS families up to date. Among the most significant findings is an apparently balanced translocation, t(7;18)(q22-q31;q22.1), identified in a proband with GTS and several of his relatives presenting various clinical GTS symptoms (Heutink et al. 1990). Subsequently a proband with GTS and terminal deletion, del (18)(q22.2-qter) was identified (Donnai 1987), as a second reported case of chromosome 18q22 abnormality in conjunction with GTS phenotype. Both findings led to the tentative assignment of GTS gene locus at 18q21-q22, a region currently subjected to molecular cloning (Boghosian-Sell et al. 1996).

Among the chromosome areas detected more than once in probands with chromosome abnormality and GTS phenotype are chromosome 9p, deleted in two GTS probands (Sigh et al. 1982, Taylor et al. 1991), and sex chromosome abnormalities, namely 47,XYY and 47,XXX karyotypes (Marskey 1974, Sigh et al. 1982).

The GTS symptomatology has also been observed in several individuals with trisomy 21 (Barabas et al. 1986, Karlinsky et al. 1986, Collacott and Ismail 1988), in a male

with Fragile X syndrome, moderate mental retardation and autism (Kerbeshian et al. 1984), and in probands with Asperger's syndrome (Kerbeshian and Burd 1986, Marriage et al. 1993). In one instance a co-occurrence of childhood onset GTS and adult onset Huntington disease has been described (Kerbeshian et al. 1991).

Traditional linkage approaches have not found evidence for a linked marker at most of the above chromosomal regions so far. As such, the co-segregation of any chromosome translocation, micro-deletion, or -duplication of chromosome region with GTS could be coincidental and unrelated to the phenotype, since the population frequency of balanced translocation carriers is 1/1,000 and the prevalence of chromosome deletions / duplications is approximately 1/10,000 at birth. It is, however, possible that some of the above chromosome variations are causally related to the phenotype and their molecular identification will provide valuable insights into GTS.

Similarly, a relatively high incidence of 1/1,000 live births for sex chromosome trisomies and 1/800 for trisomy 21 argue against an interrelationship between sex-chromosome abnormalities and GTS as well as Down syndrome and GTS (Myers and Pueschel 1995).

More recently, a balanced chromosomal translocation, t(1;8)(q21;q22.1), co-segregating with the GTS affection status in the family was described (Devor and Magee, in press). This is a third report associating GTS with the chromosome 8q22.1 region. A suggestive linkage (=significant at low stringency level) was found between GTS and the 8q22 region by Leppert et al. (1996), as well as in a case-control association study in the Afrikaner population of South Africa (Simonic et al. 1998). The cloning of the chromosome 8 breakpoint in the above family is therefore justified and currently well under way (J.L. Weber - personal communication).

GTS and environmental factors

While the role of genetics in GTS etiology is now well documented and accepted, the environmental factors influencing penetrance as well as severity of phenotypic expression of GTS gene(s) are in the early stages of investigation.

Clinical observations and recent results of immunological studies indicate that infections with group A beta-hemolytic streptococci, among others, may trigger onset or worsening of tics in subset of GTS pediatric cases (Allen et al. 1995). These cases have been identified as pediatric autoimmune neuropsychiatric disorders (OCD and TD) associated with streptococcal infections (PANDAS) (Garvey et al. 1998).

It has been hypothesized that Sydenham's chorea (SC), a major manifestation of rheumatic fever (RF), may provide a medical model for OCD and GTS. SC is a sequel of an untreated Group A ß-hemolytic streptococcal infection (GABHS) with the pathophysiology probably being due to antibodies to GABHS cross-reacting with certain brain regions. The expression of D8/17, a B-lymphocyte cell surface antigen, which was assigned as trait marker for susceptibility to SC and RF, has been also found greater in subset of patients with childhood onset OCD and GTS without documented SC or RF (Murphy et al. 1997). Furthermore, significantly higher serum levels of antineuronal antibodies against putamen have been demonstrated in children and adolescents with GTS than in the controls (Singer et al. 1998).

Autoimmunity, T-cells and chromosome fragility

The cells responsible for immune specificity are lymphocytes. During the 1960s it was discovered that the two major classes of immune responses are mediated by different classes of lymphocytes: T- cells, which develop in the thymus, are responsible for cell-mediated immunity; B-cells, which in mammals develop in the adult bone marrow or the fetal liver, produce antibodies. The majority of T-lymphocytes play a regulatory role in immunity, acting either to enhance or suppress the responses of other white blood cells. These cells are called helper T-cells or Th-cells.

Differentiated Th-cells produce restricted set of lymphokines, allowing their subdivision into two major subsets, Th1- and Th2-cells, which led to a new paradigm for immunoregulation based on the Th1/Th2 dichotomy (Carter and Dutton 1996). In general, Th1 cells produce interleukin (IL)-2 and interferon (IFN)-gamma, while Th2 cells characteristically produce IL-4, IL-5, and IL-6.

According to the results obtained in different experimental models of autoimmune diseases, the Th1-cells contribute to the pathogenesis of several organ-specific autoimmune diseases, whereas Th2-cells may inhibit disease development (Charlton and Lafferty 1995). There is a general agreement that the different functional subsets of Th-cells arise post-thymically from a common pool of precursors and as a consequence of antigen activation. However, the factors affecting differentiation of Th precursors into Th1 or Th2 subsets are still unclear.

Fragile sites (FS) on human chromosomes which are defined as vulnerable regions where lesions occur spontaneously or after induction with certain break-inducing agents, have been divided into two main groups, viz. "rare" and "common" FS according to their expression frequencies and mode of induction. Most of the rare FS have been cloned by now, while the molecular and biological significance of common fragility remains enigmatic (Sutherland and Richards 1999).

The finding of increased breakage rates at numerous common FS in two subsequent studies (Gericke et al. 1995, 1996) indicate that for unknown reasons, the T-lymphocyte DNA at the chromosomal level of organization is more sensitive to the inhibition of DNA synthesis in GTS individuals than in controls without GTS. Such observation may well be a result of different initial cell type ratio (e.g. Th1/Th2) prior to mitogen stimulation in two groups of samples and therefore it may prove important to investigate the T-lymphocyte balance (e.g. Th1/Th2) in purified T-cell samples from GTS vs. non-GTS individuals. Any alterations in such balance would be suggestive of the type of candidate genes contributing to the GTS phenotype.

The relation between the markers for streptococcal infections, higher serum levels of specific antineuronal antibodies, increased common chromosome fragility and clinical

characteristics of GTS remain unknown. Nevertheless, if proven, autoimmune etiology of tics/OCD even in a subset of GTS/OCD patients will have large implications for the treatment, prevention (Swedo et al. 1997, 1998), and genetic studies of GTS.

Summary

The attempts to narrow phenotypic definition of GTS and other neurobehavioral and neuropsychiatric traits will undoubtedly continue until genetic and environmental factors playing a role in their etiology are elucidated.

The unsuccessful gene-mapping efforts and number of spurious candidate gene association findings over the years of searching for GTS loci are by most researchers in the field viewed as an argument for pursuing alternative strategies, rather than a failure. As a consequence, there is a move towards non-parametric methods of analysis, collection of smaller, less densely affected families as well as analysis of large data sets obtained by international collaboration. Candidate gene and candidate loci association studies as well have good chances of success, if the results are interpreted with caution and are followed by well-designed confirmatory studies.

It remains important to pursue the genetic studies to find a definitive diagnostic test and understand the pathophysiology of a common neurobehavioral disorder which has the potential of disrupting the development of children of all population groups. Several discoveries in GTS research in recent years deserve further elucidation. One is the observation of tic- aggravation following streptococcal infections, offering a whole new perspective in understanding of GTS. The other is a threefold independent association of one chromosome region (8q22.1) with GTS. Discoveries of similar kind are promise of even more challenging and hopefully successful future developments in the search for the GTS gene(s).

REFERENCES

Allen AJ, Leonard HL, Swedo SE. Case study: a new infection-triggered, autoimmune subtype of pediatric OCD and Tourette's syndrome. J Am Acad Child Adolesc Psychiatry 1995; 34(3):307-311

Barabas G, Wardell B, Shapiro M, Mathews WS. Coincident Down's and Tourette syndromes: Three case reports. J Child Neurol 1986; 1:358-360

Barr CL, Wigg KG, Zovko E, Sandor P, Tsui LC. No evidence for a major gene effect of the dopamine D4 receptor gene in the susceptibility to Gilles de la Tourette syndrome in five Canadian families. Am J Med Genet 1996; 67(3):301-305

Barr CL, Wigg KG, Zovko E, Sandor P, Tsui LC. Linkage study of the dopamine D5 receptor gene and Gilles de la Tourette syndrome. Am J Med Genet 1997; 74(1):58-61

Barr CL, Sandor P. Current status of genetic studies of Gilles de la Tourette syndrome. Can J Psychiatry 1998; 43(4):351-357

Boghosian-Sell L, Comings DE, Overhauser J. Tourette syndrome in a pedigree with a 7;18 translocation: identification of a YAC spanning the translocation breakpoint at 18q22.3. Am J Hum Genet 1996; 59(5):999-1005

Bornstein RA, Baker GB. Urinary amines in adults with Tourette's syndrome. Psychiatry Res 1992; 43(3):277-285

Brett PM, Curtis D, Robertson MM, Gurling HM. The genetic susceptibility to Gilles de la Tourette syndrome in a large multiple affected British kindred: linkage analysis excludes a role for the genes coding for dopamine D1, D2, D3, D4, D5 receptors, dopamine beta hydroxylase, tyrosinase, and tyrosine hydroxylase. Biol Psychiatry 1995a; 37(8):533-540

Brett PM, Curtis D, Robertson MM, Gurling HM. Exclusion of the 5-HT1A serotonin neuroreceptor and tryptophan oxygenase genes in a large British kindred multiply affected with Tourette's syndrome, chronic motor tics, and obsessive-compulsive behavior. Am J Psychiatry 1995b; 152(3):437-440

Brett PM, Curtis D, Robertson MM, Gurling HM. Neuroreceptor subunit genes and the genetic susceptibility to Gilles de la Tourette syndrome. Biol Psychiatry 1997; 42(10):941-947

Carter LL, Dutton RW. Type 1 and type 2: a fundamental dichotomy for all T-cell subsets. Curr Opin Immunol 1996; 8(3):336-42

Charlton B, Lafferty KJ. The Th1/Th2 balance in autoimmunity. Curr Opin Immunol 1995; 7(6):793-798

Collacott RA, Ismail IA. Tourettism in a patient with Down's syndrome. J Ment Defic Res 1988; 32:163-166

Comings DE, Comings BG, Devor EJ, Cloninger CR. Detection of major gene for Gilles de la Tourette syndrome. Am J Hum Genet 1984; 36:586-600

Devor EJ. Complex segregation analysis of Gilles de la Tourette syndrome: further evidence for major locus mode of transmission. Am J Hum Genet 1984; 36(3):704-709

Devor EJ. Untying the gordian knot: the genetics of Tourette syndrome. J Nerv Ment Dis 1990; 178(11):669-679

Devor EJ, Grandy DK, Civelli O, Litt M, Burgess AK, Isenberg KE, van de Wetering BJ, Oostra B. Genetic linkage is excluded for the D2-dopamine receptor lambda HD2G1 and flanking loci on chromosome 11q22-q23 in Tourette syndrome. Hum Hered 1990; 40(2):105-108

Devor EJ, Henderson V, Sparkes RS. Linkage to Tourette syndrome is excluded for red-cell acid phosphatase (ACP1) and flanking markers on chromosome 2pter-2p23. Hum Biol 1991; 63(2):221-226

Devor EJ, Dill-Devor RM, Magee HJ. The Bal I and Msp I polymorphisms in the dopamine D3 receptor gene display, linkage disequilibrium with each other but no association with Tourette syndrome. Psychiatr Genet 1998; 8(2):49-52

Devor EJ, Magee HJ. Multiple childhood behavioral disorders (Tourette syndrome, multiple tics, ADHD and OCD) presenting in family with balanced chromosome translocation t(1;8)(q21.1;q22.1) Psychiatric Genetics. *In press*.

Donnai D. Gene localization in Tourette syndrome. Lancet 1987; 1:627

Eapen V, O'Neill J, Gurling HM, Robertson MM. Sex of parent transmission effect in Tourette's syndrome: evidence for earlier age at onset in maternally transmitted cases suggests a genomic imprinting effect. Neurology 1997c; 48(4):934-937

Eapen V, Robertson MM, Alsobrook JP 2nd, Pauls DL. Obsessive compulsive symptoms in Gilles de la Tourette syndrome and obsessive compulsive disorder: differences in diagnosis and family history. Am J Med Genet 1997a; 74(4):432-438

Eapen V, Robertson MM, Zeitlin H, Kurlan R. Gilles de la Tourette syndrome in special education schools: a United Kingdom study. J Neurol 1997b; 244(6):378-382

Frankel M, Cummings JL, Robertson MM, Trimble MR, Hill MA, Benson DF. Obsessions and compulsions in the Gilles de la Tourette syndrome. Neurology 1986; 36:379-382

Garvey MA, Giedd J, Swedo SE. PANDAS: the search for environmental triggers of pediatric neuropsychiatric disorders. Lessons from rheumatic fever. J Child Neurol 1998; 13(9):413-423

Gate R, Muhleman D, Blake H, MacMurray J, Johnson P, Verde R, Saucier G, Comings DE. Correlation of length of VNTR alleles at the X-linked MAOA gene and phenotypic effect in Tourette syndrome and drug abuse. Mol Psychiatry 1998; 3(1):50-60

Gelender J, Pauls DL, Leckman J, Kidd KK, Kurlan R. D2 dopamine receptor alleles do not influence severity of Tourette's syndrome. Results from four large kindreds. Arch Neurol 1994; 51(4):397-400

Gelender J, Rao PA, Pauls DL, Hamblin MW, Sibley DR, Kidd KK. Assignment of the 5HT7 receptor gene (HTR7) to chromosome 10q qnd exclusion of genetic linkage with Tourette syndrome. Genomics 1995; 26(2):207-209

Gericke GS, Simonic I, Cloete E, Buckle C, Becker PJ. Increased chromosomal breakage in Tourette syndrome predicts the possibility of variable multiple gene involvement in spectrum phenotypes: Preliminary findings and hypothesis. Am J Med Genet (Neuropsych Genet) 1995; 60:444-447

Gericke GS, Simonic I, Cloete E, Becker PJ. Increased expression of aphidicolin-induced common fragile sites in Tourette syndrome: The key to understand the genetics of comorbid phenotypes? Am J Med Genet (Neuropsych Genet) 1996; 67:25-30

Goldenberg JN, Brown SB, Weiner WJ. Coprolalia in younger patients with Gilles de la Tourette syndrome. Movement Dis 1994; 9:622-625

Grice DE, Leckman JF, Pauls DL, Kurlan R, Kidd KK, Pakstis AJ, Chang FM, Buxbaum JD, Cohen DJ, Gelernter J. Linkage disequilibrium between an allele at the dopamine D4 receptor locus and Tourette syndrome, by the transmission-disequilibrium test. Am J Hum Genet 1996; 59(3):644-652

Hasstedt SJ, Leppert M, Filloux F, van de Wetering BJM, McMahon WM. Intermediate inheritance of Tourette syndrome, assuming assortative mating. Am J Hum Genet 1995; 57:682-689

Heutink P, van de Wetering BJ, Breedveld GJ, Weber J, Sandkuyl LA, Devor EJ, Heiberg GA, Niermeijer MF, Oostra BA. No evidence for genetic linkage of Gilles de la Tourette syndrome on chromosomes 7 and 18. J Med Genet 1990 27(7):433-436

Karlinsky H, Sandor P, Berg JM, Moldofsky H, Crawford E. Gilles de la Tourette's syndrome in Down's syndrome – a case report. Br J Psychiatry 1986; 148:601-604

Kerbeshian J, Burd L, Martsolf JT. Fragile X syndrome associated with Tourette symptomatology in a male with moderate mental retardation and autism. J Dev Behav Pediatr 1984; 5(4):201-203

Kerbeshian J, Burd L. Asperger's syndrome and Tourette syndrome: the case of the pinball wizard. Br J Psychiatry 1986; 148:731-736

Kerbeshian J, Burd L, Leech C, Rorabaugh A. Huntington Disease and childhood-onset Tourette syndrome. Am J Med Genet 1991; 39:1-3

Kurlan R. Tourette's syndrome: current concepts. Neurology 1989; 39:1625-1620

Kurlan R, Eapen V, Stern J, McDermott MP, Robertson MM. Bilineal transmission in Tourette's syndrome families. Neurology 1994; 44(12):2336-2342

Kurlan R. Diagnostic criteria for genetic studies of Tourette syndrome. Arch Neurol 1997; 54(5):517-518

LaBuda MC, Gottesman II, Pauls DL. Usefulness of twin studies for exploring the etiology of childhood and adolescent psychiatric disorders. Am J Med Genet 1993; 48(1):47-59

Lajonchere C, Nortz M, Finger S. Gilles de la Tourette and the discovery of Tourette syndrome. Includes a translation of his 1884 article. Arch Neurol 1996; 53(6):567-574

Lander ES, Schork NJ. Genetic dissection of complex traits. Science 1994; 265:2037-2048

Leckman JF, Peterson BS, Pauls DL, Cohen DJ. Tic disorders. Psychiatr Clin North Am 1997; 20(4):839-861

Leonard HL, Swedo SE, Rapoport JL, Rickler KC, Topol D, Lee S, Rettew D. Tourette syndrome and obsessive-compulsive disorder. Adv Neurol 1992; 58:83-93

Leppert M, Peiffer A, Snyder B, van de Wetering, Filloux F, Coon H, McMahon WM. Two loci of interest in a family with Tourette syndrome. Am J Hum Genet 1996; 59(Suppl):A225

Lichter DG, Jackson LA, Schachter M. Clinical evidence of genomic imprinting in Tourette's syndrome. Neurol 1995; 45:924-928

Marriage K, Miles T, Stokes D, Davey M. Clinical and research implications of the co-occurrence of Asperger's and Tourette syndromes. Aust N Z J Psychiatry 1993; 27(4):666-672

Marskey H. A case of multiple tics with vocalization (pertial syndrome of Gilles de la Tourette) and XYY karyotype. Br J Psychiatry 1974; 125:593-594

Mason A, Banerjee S, Eapen V, Zeitlin H, Robertson MM. The prevalence of Tourette syndrome in a mainstream school population. Dev Med Child Neurol 1998; 40(5):292-296

Murphy TK, Goodman WK, Fudge MW, Williams RC Jr, Ayoub EM, Datal M, Lewis MH, Zabriskie JB. B lymphocyte antigen D8/17: a peripheral marker for childhood-onset obsessive-compulsive disorder and Tourette's syndrome? Am J Psychiatry 1997; 154(3):402-407

Myers B, Pueschel SM. Tardive or atypical Tourette's disorder in a population with Down syndrome? Res Dev Disabil 1995; 16(1):1-9

Nothen MM, Hebebrand J, Knapp M, Hebebrand K, Camps A, von Gontard A, Wettke-Schafer R, Lisch S, Cichon S, Poustka F, et al. Association analysis of the dopamine D2 receptor gene in Tourette's syndrome using the haplotype relative risk method. Am J Med Genet 1994; 54(3):249-252

Patel PI. Quest for the elusive genetic basis of Tourette Syndrome. Am J Hum Genet 1996; 59:980-982

Pauls DL, Leckman JF. The inheritance of Gilles de la Tourette's syndrome and associated behaviors: evidence for autosomal dominant transmission. N Engl J Med 1986; 315:993-997

Pauls DL, Towbin KE, Leckman JF, Zahner GE, Cohen DJ. Gilles de la Tourette syndrome and obsessive compulsive disorder: Evidence supporting a genetic relationship. Arch Gen Psychiatry 1986; 43:1180-1182

Pauls DL, Raymond C, Stevenson J, Leckman JF. A family study of Gilles de la Tourette syndrome. Am J Hum Genet 1991; 48:154-163

Price RA, Leckman JF, Pauls DL, Cohen DJ, Kidd KK. Gilles de la Tourette's syndrome: tics and central nervous system stimulants in twins and nontwins. Neurology 1986; 36:232-237

Robertson MM. The Gilles de la Tourette syndrome: The current status. Br J Psychiatry 1989; 154:147-69

Robertson MM, Boardman J. Tourette's syndrome in the year 2000. Aust N Z J Psychiatry 1996; 30(6):749-759

Robertson MM, Stern JS. The Gilles de la Tourette syndrome. Crit Rev Neurobiol 1997; 11(1):1-19

Santangelo SL, Pauls DL, Lavori PW, Goldstein JM, Faraone SV, Tsuang MT. Assessing risk for the Tourette spectrum of disorders among forst-degree relatives of probands with Tourette syndrome. Am J Med Genet (Neuropsych Genet) 1996; 67:107-116

Sigh DN, Howe GL, Jordan HW, Hara X. Tourette's syndrome in a black woman with associated triple X and 9p mosaicism. J Natl Med Assoc 1982; 74:675-682

Silvestri R, Raffaele M, De Domenico P, Tisano A, Lagana A, Di Perry R. Serotoninergic agents in the treatment of Gilles de la Tourette's syndrome. Acta Neurol (Napoli) 1994; 16(1-2):58-63

Simonic I, Gericke GS, Ott J, Weber JL. Identification of genetic markers associated with Gilles de la Tourette Syndrome in an Afrikaner population. Am J Hum Genet 1998; 63:839-846

Singer HS, Giuliano JD, Hansen BH, Hallett JJ, Laurino JP, Benson M, Kiessling LS. Antibodies against human putamen in children with Tourette syndrome. Neurology 1998; 50(6):1618-1624

Staley D, Wand R, Shady G. Tourette disorder: a cross-cultural review. Compr Psychiatry 1997; 38(1):6-16

Sutherland GR, Richards RI. Fragile sites – cytogenetic similarity with molecular diversity. Am J Hum Genet 1999; 64:354-359

Swedo SE, Leonard HL. Childhood movement disorders and obsessive compulsive disorder. J Clin Psychiatry 1994; 55 (Suppl):32-37

Swedo SE, Leonard HL, Mittleman BB, Allen AJ, Rapoport JL, Dow SP, Kanter ME, Chapman F, Zabriskie J. Identification of children with pediatric autoimmune neuropsychiatric disorders associated with streptococcal infections by a marker associated with rheumatic fever. Am J Psychiatry 1997; 154(1):110-112

Swedo SE, Leonard HL, Garvey M, Mittleman B, Allen AJ, Perlmutter S, Dow S, Zamkoff J, Dubbert BK, Lougee L. Pediatric autoimmune neuropsychiatric disorders associated with streptococcal infections: clinical description of the first 50 cases. Am J Psychiatry 1998; 155(2):264-271

Taylor LD, Krizman DB, Jankovic J, Hayani A, Steuber PC, Greenberg F, Fenwick RG, Caskey CT. 9p monosomy in a patient with Gilles de la Tourette's syndrome. Neurology 1991; 41(9):1513-1515

Walkup JT, LaBuda MC, Singer HS, Brown J, Riddle MA, Hurko O. Family study and segregation analysis of Tourette syndrome: evidence for a mixed model of inheritance. Am J Hum Genet 1996; 59:684-693

CHAPTER 7

American Journal of Human Genetics (submitted for publication, July 1999)

Significant Evidence for Linkage Disequilibrium over a 5 cM region among Afrikaners

Derek Gordon[1], Ingrid Simonic[2], Jurg Ott[1]

[1] Department of Statistical Genetics, Rockefeller University, New York; [2]Neurogenetics Research Initiative, MRC Pretoria, South Africa

Abstract

The extent of disequilibrium on the population level, or background disequilibrium, can provide important information regarding complex-trait-mapping study design. Therefore, it is vitally important that the extent of background disequilibrium in differing populations be investigated. In this work, we explore the extent of deviations from Hardy-Weinberg equilibrium (HWE) at a marker locus and linkage disequilibrium (LD) between pairs of marker loci in the Afrikaner population of South Africa. DNA samples were used for genotyping of 24 loci at six chromosomes. The samples were collected from 91 healthy unrelated Afrikaner adults. Exact tests were used to determine evidence for deviations from HWE at a single marker locus or LD between pairs of marker loci. At the 0.05 level of significance, evidence was found for deviation from HWE at only one of the 24 loci. At the same level of significance, LD was found among 10 of the 38 intra-chromosomal pairs of loci. On chromosome 21, there was evidence for LD between a pair of loci with an estimated genetic distance of 5.51 cM ($p = 0.0186$). On chromosome 2, in which 5 loci were typed, one locus showed LD with three of the four other loci. The highest estimated genetic distance for pairs of loci with corrected $p \leq 0.05$ for LD was 5.28 cM (nominal p =

0.0019; p = 0.0268, corrected for multiple testing. Our findings indicate that Afrikaans-speaking Afrikaners represent one of those special populations deemed particularly suitable for disequilibrium mapping. This population may also be highly suitable for mapping genes underlying complex traits because strong LD is expected in the vicinity of disease loci, which can add power to genetic linkage analysis.

Introduction

Suitability of a specific population for the localization of disease genes by disequilibrium mapping may be assessed by investigating background LD, that is, LD among marker loci rather than between a disease and a marker locus (or loci). It has been pointed out that "quantifying the degree of such 'background' linkage disequilibrium is a crucial undertaking in paving the way for whole genome association studies" (Freimer et al. 1997). Studies of background LD have been performed in genetically isolated populations (Weir 1992; Peterson et al. 1995; Laan and Paabo 1997; Terwilliger et al. 1998); in larger outbred populations (Maiste 1993; Maiste and Weir 1995); among human and chimpanzee populations (Crouau-Roy et al. 1996); analytically (Terwilliger et al. 1998), and in computer simulations (Terwilliger et al. 1998; Kruglyak 1999). To date, we are unaware of any such studies in the Afrikaner population of South Africa, although, as described below, this population may be a strong candidate for the mapping of complex-trait genes through whole-genome association studies.

The Afrikaner population of South Africa was formed by immigrants from Europe, primarily Dutch, German and French, who began settling in Cape Town in 1652 (Torrington et al. 1984, Jenkins 1990). During the 18th century, the population was virtually cut from Europe and expanded from just over 1,000 to 17,000 in 85 years (Jenkins 1996). After Great Treks (mass movements of Afrikaners away from Cape in the 1830s), geographical isolation of new Afrikaner settlements was also accompanied by cultural isolation of the population, mainly due to the afrikaans language (derived from Dutch) and religion (Dutch reformed). The lack of larger immigration waves from Europe also contributed to the genetic isolation of the

Afrikaners, which was partially broken only around the turn of this century. Today after an estimated 12 generations since founding, there are about 3 million Afrikaners living in South Africa and abroad. Common mutations and/or marker haplotypes among Afrikaners have been established for several heritable disorders. Such disorders include variegate porphyria, keratolytic winter erythema, hypercholesterolemia, and progressive heart block (Jenkins 1990, Jenkins 1996, Meissner et al. 1996, Warnich et al. 1996, Starfield et al. 1997) suggesting that this recently founded population has maintained its relative genetic homogeneity. The extent of conserved haplotypes around disease genes according to the above studies is between 8-11cM, making the Afrikaner population of South Africa exceptionally valuable for genetic mapping studies (Groenewald et al. 1998).

Methods

Subjects

Coded and anonymized DNA samples collected for a previous case-control association study (Simonic et al. 1998) were used for genotyping of 24 loci (22 genetic loci and 2 cryptic duplicates) on six chromosomes. The samples were collected from paid healthy volunteers coming from two sources: medical students and personnel at the Afrikaans University. All study subjects were unrelated, older than 18 years, spoke Afrikaans and had Afrikaner family names. The male:female ratio in the group was ~ 1:1.

Primers

24 primer pairs for the detection of short tandem repeat polymorphisms (STRPs) were obtained from Research Genetics (Huntsville, AL). 16 of the genetic markers were tetra- and trinucleotide repeat polymorphisms developed within the Cooperative Human Linkage Center (CHLC) and 6 genetic markers were Genethon dinucleotide repeat polymorphisms. Allele sizes were determined utilizing the known genotypes of three CEPH individuals: 1331-01, 1331-02 and 134702. Marker spacing was determined using comprehensive sex-averaged genetic maps (Broman et al. 1998). A list of all markers used may be found in Table 1.

PCR

Amplification of STRP markers was performed with 45ng of pooled DNA in a 10µl PCR reaction mixture containing 1.5µl PCR buffer (1.5mM $MgCl_2$, 50mM KCL, 10mM Tris-HCl, 0.01% w/v gelatin); 200µM each dGTP, dATP, dTTP, 2.5µM dCTP, 0.35µCi α^{32}P-dCTP (NEN Du Pont, 800 Ci/mmole, 10µCi/µl); 10ng (about 3 pmoles) each PCR primer; and 0.3 U *AmpliTaq* polymerase (Boehringer Mannheim). Most markers were amplified in sets of two markers per amplification reaction in 96-well microtiter plate (*Techno*). Samples were subjected to 27 cycles consisting of 30 sec at 94°, 75 sec at 55°, 15 sec at 72° and a final 6 min at 72° after the last cycle. Products were denatured at 95° for 10 min, loaded (1µl) onto a vertical 6.5% polyacrylamide DNA sequencing gel and run at 70 watts constant power for about 3 hours. The amplified DNA fragments for each marker were evaluated visually on the autoradiographs.

Table 1

Markers used in this study and their estimated locations

Chromosome 1		**Chromosome 2**		**Chromosome 5**	
Marker	cM	Marker	cM	Marker	cM
GATA26C12(D1S1648)	101.48	AFM177xh4(D2S139)	101.56	GATA52A12(D5S1501)	85.25
GATA61A06(D1S1665)	102.02	GATA62B10	103.16	AFM284vd1(D5S641)	92.38
AFM294wg1(D1S481)	104.23	GATA88G05	103.16	GATA142H05	92.38
GAAT1D09	104.23	16AFM044xa1(D2S289)	103.16	GATA91E02(D5S1726)	94.80
		Mfd 337(D2S417)	106.84	GATA69H12	94.80
Chromosome 11		**Chromosome 20**		**Chromosome 21**	
Marker	cM	Marker	cM	Marker	cM
Mfd 316(D11S1377)	120.87	GATA45B10(D20S480)	79.91	GATA45C03	31.26
GATA64D03(D11S4464)	123.00	GATA46C01(D20S1085)	82.07	AFM276za5(D21S1254)	31.26
GATA140F03	123.00	GGAA11E12(D20S469)	84.78	ATA27F01(D21S1440)	36.77
AFM240ye1(D11S933)	124.07				

Statistical Tests

There are several statistical methods available for testing deviations from HWE at a single locus (Guo and Thompson 1992; Zaykin et al. 1995) or LD among several loci, when haplotype information is available (Lazzeroni and Lange 1995; Slatkin 1994; Schneider et al. 1997) or not available (Zaykin et al. 1995; Slatkin and Excoffier 1996; Schneider et al. 1997). Since haplotype information was not available for this data set, we used the tests implemented by Zaykin et al. (1995) in their MLD software (see Electronic Database Information).

As outlined below, exact tests were used in all cases. To test for departure from HWE at a marker locus, the probability of the set of genotypes in the sample, conditional on the allelic counts (i.e., the marginals) was calculated from multinomial theory under the null hypothesis of HWE (Guo and Thompson 1992; Zaykin et al. 1995). Alleles were permuted and, for each permutation, the conditional probability was calculated. The proportion of permutations no more probable than the original sample provided an estimate for the empirical significance level associated with the test result. Exact confidence intervals were computed based on the binomial distribution as implemented in the BINOM program (Ott 1999, Electronic Database Information). These tests of HWE are implemented in the MLD software program (Electronic Database Information), and all reported p-values are determined by use of that software. Here, 17000 permutations were selected for each run of MLD on a marker locus to narrow the 99 % confidence interval for p-value reported.

For a given pair of loci, the probability of the set of multi-locus genotypes in the sample, conditional on the allelic counts (i.e., the marginals) was calculated from multinomial theory under the null hypothesis that each two-locus genotype frequency was the product of the respective one-locus genotype frequencies (Zaykin et al. 1995). Permutations were implemented by keeping the one-locus genotypes intact and permuting these genotypes among individuals at one of the loci (Zaykin et al. 1995). An estimate for the empirical significance level was determined as in the one-locus case. These tests are also implemented in the MLD software program, and all reported

p-values were determined by use of that software. As above, 17000 permutations were chosen for each run of MLD (see Table 3).

The one-locus test was applied to all markers, while the test for LD among pairs of loci was applied to all pairs of markers. Here and elsewhere, the term *intra-chromosomal pair* refers to a pair of marker loci on the same chromosome, while *inter-chromosomal pair* refers to a pair of loci on two different chromosomes. Results of tests for intra-chromosomal markers are presented in Table 2.

Results and Discussion

One-Locus Tests

Only one marker, D5S1501, had a p-value less than 0.05 when testing for deviations from HWE ($p = 0.0299$). By chance alone we would expect approximately one of 22 markers (not counting the cryptic duplicates) to show a p-value less than 0.05, even if all the markers are in HWE. Thus, no markers showed a significant deviation from HWE.

Two-Locus Tests

Table 2 presents those intra-chromosomal marker pairs for which the test for LD showed evidence for LD, along with the estimated genetic map distance between the markers. The reported distances were computed using Table 1. The p-values (uncorrected for multiple testing) for all intra-chromosomal marker pairs may be found in Table 3. One can see from Table 2 that there is evidence for LD in this sample between markers D21S1254 and D21S1440 on chromosome 21. The estimated genetic distance between these two markers is 5.51cM. On chromosome 2, there is evidence for LD between D2S417 and all other markers (with the exception of D2S289). The estimated genetic distance between marker D2S417 and D2S139 is 5.28 cM. In fact, 6 of the 10 marker pairs in Table 2 have estimated genetic distances over 2 cM.

Table 2

Marker Pairs that show evidence for LD in sample of 91 Afrikaners.

Marker Pair		Estimated Genetic Distance (cM)	Chromo-some	p-value and 99% Confidence Interval (not corrected for multiple testing)	p_{corr} and 99% Confidence Interval (corrected for multiple testing)
D21S1254	D21S1440	5.51	21	.0186 (.0160,.0214)	.2733
D2S417	D2S139	5.28	2	.0016 (.0009,.0026)	.0269 (.0152,.0433)
D2S417	GATA62B10	3.68	2	.0004 (.0001,.0010)	.0068 (.0017,.0169)
D2S417	GATA88G05	3.68	2	.0000 (.0000,.0003)	.0000 (.0000, .0051)
D20S1085	D20S469	2.71	20	.0461 (.0421,.0504)	.5517
D20S480	D20S1085	2.16	20	.0233 (.0204,.0264)	.3302
D1S1648	D1S1665	0.54	1	.0021 (.0013, .0032)	.0351 (.0219,.0530)
GATA62B10	GATA88G05	0.0*	2	.0000 (.0000,.0003)	.0000 (.0000, .0051)
D5S1726	GATA69H12	0.0*	5	.0000 (.0000,.0003)	.0000 (.0000, .0051)
D11S4464	GATA140F03	0.0	11	.0104 (.0085, .0126)	.1628

* indicates that one marker in pair was identified by Marshfield Labs as being a "cryptic" duplicate of another marker

To correct for multiple testing, we applied a form of Bonferroni correction as follows. For N markers on a chromosome, we assume that $N - 1$ independent tests are carried out when all pairwise comparisons are considered. Over all chromosomes, this leads to 17 independent intra-chromosomal tests. Thus, a corrected empirical significance level is given by $p_{corr} = 1-(1-p)^{17}$. With this, we observe that all p-values for

Chromosome 2 are still significant at the 0.05 level (Table 2). Additional marker pairs with $p_{corr} < 0.05$ are found on chromosomes 1 and 5.

We also tested for LD among more than two marker loci jointly as implemented in the MLD program (Zaykin et al. 1995). Table 4 reports results for chromosome 2 in which the marker set tested includes D2S417 and excludes D2S289, since this marker showed no evidence for LD with any other marker. All p-values in Table 4 are significant at the 0.05 level, even when using the conservative assumption that 3 independent tests have been performed (corrected p-value $= 1-(1-p)^3$).

Table 3

- Exact Test p-values for all intra-chromosomal pairs
- 99% confidence intervals for significant p-values (< 0.05) determined by BINOM program (Ott 1999)
- Uncorrected for Multiple Testing

Chromosome 1

	D1S1665 102.02cM	GAAT1D09 104.23cM	D1S481 104.23cM
D1S1648 101.48cM	.0021 (.0013, .0032)	.4878	.0885
D1S1665 102.02cM	***	.6745	.5837
GAAT1D09 104.23cM		***	.6679

Chromosome 2

	GATA62B10 103.16cM	GATA88G05 103.16cM*	D2S289 103.16cM	D2S417 106.84cM
D2S139 101.56cM	.0656	.0856	.1474	.0016 (.0009,.0026)
GATA62B10 103.16cM	***	.0000 (.0000,.0003)	.6314	.0004 (.0001,.0010)
GATA88G05 103.16cM*		***	.5521	.0000 (.0000,.0003)
D2S289 103.16cM			***	.1033

Chromosome 5

	D5S641 92.38cM	GATA142H05 92.38cM	D5S1726 94.8cM	GATA69H12 94.8cM*
D5S1501 85.25cM	.7798	.8395	.9601	.7597
D5S641 92.38cM	***	.7393	.4508	.4648
GATA142H05 92.38cM		***	.2444	.4784
D5S1726 94.8cM			***	.0000 (.0000,.0003)

Chromosome 11

	D11S4464 123.00 cM	GATA140F03 123.00 cM	D11S933 124.07cM
D11S1377 120.87cM	.5867	.1154	.3382
D11S4464 123.00cM	***	.0104 (.0085, .0126)	.7036
GATA140F03 123.00cM		***	.7576

Chromosome 20

	D20S1085 82.07cM	D20S469 84.78cM
D20S480 79.91cM	.0233 (.0204,.0264)	.7953
D20S1085 82.07cM	***	.0461 (.0421,.0504)

Chromosome 21

	D21S1254 31.26cM	D21S1440 36.77cM
GATA45C03 31.26cM	.4415	.1180
D21S1254 31.26cM	***	.0186 (.0160,.0214)

NOTE. - All map positions determined by using the Marshfield Website sex-averaged maps (www.marshmed.org/genetics/)
*marker was identified as being a "cryptic" duplicate of another marker

Table 4

Multi-locus Exact-Test p-values for Chromosome 2 using sets of markers that include D2S419 and that exclude D2S289.

3-Locus Tests

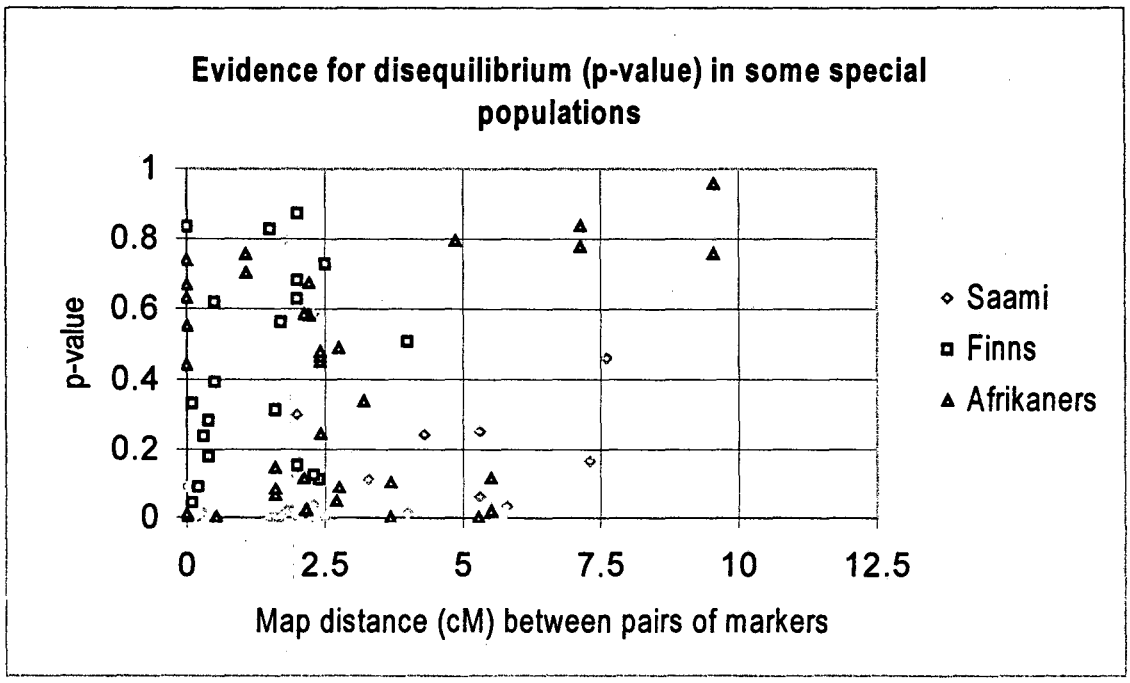
Marker Set	P-value	Corrected P-value
D2S139 – GATA62B10-D2S417	.0025	.0075
D2S139 – GATA88G05-D2S417	.0030	.0090
GATA62B10-GATA88G05-D2S417	.0000	.0000

4-Locus Test

Marker Set	P-value
D2S139 – GATA62B10- GATA88G05-D2S417	.0000

Our findings indicate that Afrikaans-speaking Afrikaners represent one of those special populations deemed particularly suitable for disequilibrium mapping (Figure 1). This population may also be highly suitable for mapping genes underlying complex traits because strong LD is expected in the vicinity of disease loci, which can add power to genetic linkage analysis.

Figure 1



◇ Laan and Pääbo (1997)
□ Peterson et al. (1995)
△ Simonic et al. (1998)

Acknowledgements

The authors gratefully acknowledge Dr. James Weber for providing resources for this study and enabling Ingrid Simonic to perform genotyping in his laboratory. Also, Harald Göring is gratefully acknowledged for software provided in early stages of this work. The authors thank Dimitri Zaykin for providing answers to questions regarding the theory and usage of the MLD program. Support for this research comes from grant HG00008 from the National Institutes of Health.

Electronic Database Information

History of the Afrikaners' migration to South Africa may be found at

http://www.sacs.org.za/level2/history/.

The Arlequin program is available via ftp from the URL
http://anthropologie.unige.ch/arlequin/.

The MLD program is freely available from the North Carolina State University Department of Statistics. The URL is http://statgen.ncsu.edu/#software/.

The BINOM program is available via ftp from the Rockefeller University Laboratory of Statistical Genetics. The URL is
ftp://linkage.rockefeller.edu/software/utilities/.
Documentation may be found at
http://linkage.rockefeller.edu/ott/linkutil.htm#BINOM.

Information for the Marshfield Genetic Maps may be found at the URL
http://www.marshmed.org/genetics/.

References

Broman KW, Murray JC, Sheffield VC, White RL, Weber JL (1998) Comprehensive human genetic maps: individual and sex-specific variation in recombination. Am J Hum Genet 63:861-869

Crouau-Roy B, Service S, Slatkin M, Freimer N (1996) A fine scale comparison of the human and chimpanzee genomes: linkage, linkage disequilibrium and sequence analysis. Hum Mol Genet 5: 1131-1137

Freimer NB, Service SK, Slatkin M (1997) Expanding on population studies. Nat Genet 17, 371-373

Groenewald JZ, Liebenberg J, Groenewald IM, Warnich L (1998) Linkage disequilibrium analysis in a recently founded population: evaluation of the variegate porphyria founder in South African Afrikaners. Am J Hum Genet 62:1254-1258

Guo SW, Thompson EA (1992) Performing the exact test of Hardy-Weinberg proportion for multiple alleles. Biometrics 48(2):361-72

Jenkins T (1990) Medical genetics in South Africa. J Med Genet 27:760-779

Jenkins T (1996) The South African malady. Nat Genet 13:7-9

Kruglyak L (1999) Prospects for whole-genome linkage disequilibrium mapping of common disease genes. Nat Genet 22:139-144

Laan M, Pääbo S (1997) Demographic history and linkage disequilibrium in human populations. Nat Genet 17: 435-438

Lazzeroni LC, Lange K (1995) Markov chains for Monte Carlo tests of genetic equilibrium in multidimensional contingency tables. Technical Report No. 178, Division of Biostatistics, Stanford University

Maiste PJ (1993) Comparison of statistical tests for independence at genetic loci with many alleles. Ph.D. Dissertation, North Carolina State University, Department of Statistics

Maiste PJ, Weir BS (1995) A comparison of tests for independence in the FBI RFLP data bases. Genetica 96:125-138

Meissner PN, Dailey TA, Hift RJ, Ziman M, Corrigall AV, Roberts AG, Meissner DM, Kirsch RE, Dailey HA (1996) A R59W mutation in human protoporphyrinogen oxidase results in decreased enzyme activity and is prevalent in South Africans with variegate porphyria. Nat Genet 13:95-97

Ott J (1999) Analysis of Human Genetic Linkage. Johns Hopkins University Press. Baltimore

Peterson AC, Di Rienzo A, Lehesjoki AE, de la Chapelle A, Slatkin M, Freimer NB (1995) The distribution of linkage disequilibrium over anonymous genome regions. Hum Mol Genet 4:887-94

Schneider S, Kueffer J-M, Roessli D, Excoffier L (1997) Arlequin ver. 1.1: A software for population genetic data analysis. Genetics and Biometry Laboratory, University of Geneva, Switzerland

Simonic I, Gericke GS, Ott J, Weber JL (1998) Identification of genetic markers associated with Gilles de la Tourette syndrome in an Afrikaner population. Am J Hum Genet 63:839-846

Slatkin M (1994) Linkage disequilibrium in growing and stable populations. Genetics 137: 331-336

Slatkin M, Excoffier L (1996) Testing for linkage disequilibrium in genotypic data using the EM algorithm. Heredity 76:377-383

Starfield M, Hennies HC, Jung M, Jenkins T, Wienker T, Hull P, Spurdle A, et al. (1997) Localization of the gene causing keratolytic winter erythema to chromosome 8p22-p23, and evidence for a founder effect in South African Afrikaans-speakers. Am J Hum Genet 61:370-378

Terwilliger JD, Ott J (1994) Handbook of Human Genetic Linkage. Johns Hopkins University Press. Baltimore

Terwilliger JD, Zollner S, Laan M, Pääbo S (1998) Mapping genes through the use of linkage disequilibrium generated by genetic drift: 'drift mapping' in small populations with no demographic expansion. Hum Hered 48:138-54

Torrington M, Botha JL, Pilcher GJ, Baker SG (1984) Association between familial hypercholesterolaemia and church affiliation. SA Med J 65:762-767

Warnich L, Meissner PN, Hift RJ, Louw JH, van Heerden CJ, Retief AE (1996) Mapping of Variegate Porphyria (VP) gene: contradictory evidence for linkage between VP and microsatellite markers at chromosome 14q32. Hum Genet 97:690-692

Weir BS (1992) Disequilibrium on chromosome 21 in some Utah families. Cytogenet Cell Genet 59: 128-129

CHAPTER 8

Human Molecular Genetics (submitted for publication, September 1999)

CONFIRMATION OF GILLES DE LA TOURETTE SYNDROME (GTS) SUSCEPTIBILITY LOCI ON CHROMOSOMES 2p11, 8q22 and 11q23-24 IN SOUTH AFRICAN AFRIKANERS

[1]I. Simonic, [1]G.S. Gericke, [2]D. Nyholt, [2]D. Gordon, [2] J. Ott, [3] J.L. Weber

[1] Neurogenetics Research Initiative, MRC Pretoria, South Africa

[2] Department of Statistical Genetics, Rockefeller University, New York, NY

[3] Center for Medical Genetics, Medical Research Foundation, Marshfield, WI

Summary

Five genomic regions on chromosomes 2, 8, 11, 20, and 21 that gave evidence for association with GTS in previous case-control association studies were investigated for linkage and association with GTS utilizing DNA samples from 91 Afrikaner nuclear families with one or more affected children. Highly polymorphic markers with mean heterozygosity of 0.77 were typed and resulting genotypes evaluated using single marker transmission disequilibrium (TDT), single marker haplotype relative risk (HRR), and multi-marker 'extended' TDT and HRR methods. Single marker TDT analysis showed evidence for linkage or association, with p-values near 0.05, for markers D2S139, GATA28F12 and D11S1377 on chromosomes 2p11, 8q22 and 11q23-24, respectively. Extended TDT and HRR analysis provided further evidence for linkage or association on chromosome 2 with p values of 0.007 and 0.025, and chromosome 8 with p values of 0.059 and 0.013. These results provide strong additional evidence for the location of GTS susceptibility loci.

Introduction

While the etiology of Gilles de la Tourette syndrome (GTS), a childhood-onset neurologic disorder characterized by chronic involuntary movements (either motor, vocal, or both at the time of the diagnosis) remains hidden, it has a strong genetic component (Price et al. 1985). According to recent complex segregation analyses, the susceptibility for GTS is more complicated than previously suggested, conveyed by an additive major locus in combination with a multifactorial background (Walkup et al. 1996). A failure to identify GTS-linked genetic loci in large, multiple affected kindreds motivated the redirection of gene-mapping efforts toward allele-sharing methods and linkage disequilibrium (LD) studies in genetically isolated populations.

Lander and Kruglyak (Lander and Kruglyak 1995) proposed stringent guidelines for the assignment of genome-wide significance levels in linkage studies to diminish the rates of false positive results. In case-control LD studies, because of the danger of spurious associations due to mismatching of controls and population admixture, consistent replication may be the best evidence for a true association (Kidd 1993).

Our previous whole-genome search for association with GTS among Afrikaners (Simonic et al. 1998) with 1,167 short tandem repeat polymorphisms, using a cross-sectional case-control strategy, DNA pooling, and follow-up individual typing of two independent sets of case-control samples resulted in identification of 15 markers in 11 chromosomal regions with significant allele distribution differences between the two groups of samples.

For the reasons outlined above, our current study was designed as an attempt to replicate our previous results. Therefore, in this study, we collected additional independent affected individuals and their parents from the Afrikaner population. We then applied transmission disequilibrium (TDT) and haplotype relative risk (HRR) tests to genotype data generated at 5 earlier determined regions of interest, to further investigate the significance of our initial findings. In three instances, our previous results were confirmed ($p < 0.05$) even after the application of conservative Bonferroni corrections for multiple testing.

Subjects, Material & Methods

Patients

85 randomly selected GTS individuals from 74 independent families (64 singletons, 9 pairs and 1 affected trio), not participating in our previous studies, were invited together with their parents for interviews and blood sample collection. They were identified as GTS sufferers at previous visits at the Tourette Syndrome Clinic in Pretoria and interviewed for the second time using the criteria of the Tourette Syndrome Association Genetic Consortium (Simonic et al. 1998). For 19 individuals (including two pairs of siblings), only one parent was available for the study. In addition, parental and affected individual blood samples were collected for 22 GTS individuals from 17 independent families (12 singletons and 5 affected pairs), whose DNA samples were used in our initial genome screen. Our combined material thus consisted of 91 families with 107 GTS individuals, 88 with both parental DNA samples available, and 19 with one parental DNA only. All families spoke Afrikaans and had Afrikaner family names on both maternal and paternal sides.

Polymorphisms

In total, 31 short tandem-repeat polymorphisms (STRPs) mapping to 5 genomic regions were selected for this replication study (Table 1). The main criteria for marker selection were map position (Center for Medical Genetics, Marshfield), spacing (~1.5 cM separation on sex-averaged genetic map), and heterozygozity (mean 0.77) (Broman et al. 1998). Primer pairs were mostly obtained from Research Genetics. Primers for CA dinucleotide repeat marker T7-27 were designed from the partial sequence of BAC clone 127D17, which contains the MTG8 gene. Allele sizes for the markers were determined using known genotypes of two CEPH individuals (133101 and 133102).

PCR

Amplification of STRPs was performed in 96 well microtiter plates with 50 ng of DNA in a 5 µl volume containing 3 mM $MgCl_2$; 50 mM KCl; 10 mM Tris-HCl, pH 8.3; 0.01% (w/v) gelatin; 200 µM each dGTP, dCTP, and dTTP; 2.5 µM dATP; 0.35

μCi α[^{33}P]-dATP (NEN Du Pont; 800 Ci/mmol, 10 μCi/μl); 10 ng (~3 pmol) each PCR primer; and 0.3 U Taq polymerase (Sigma). Samples were subjected to 29 cycles consisting of 30 sec at 94°C, 75 sec at 55°C, and 30 sec at 72°C, with a final 6 min at 72°C after the last cycle. All markers were amplified individually. All genotyping was performed in duplicate to minimize errors. PCR products were denatured by heating for 10 min at 95°C prior to loading (0.5μl) onto vertical 6.5% polyacrylamide, 7.7 M urea DNA sequencing gels and running at 70 W constant power for ~2 hours. Gels were dried on filter paper and exposed on X-ray film.

Table 1

31 markers in five genomic regions were selected for the replication study

Chr.	*Marker*	*Locus*	*cM*	*Het.*	*Chr.*	*Marker*	*Locus*	*cM*	*Het.*
2	AFM177xh4	D2S139	101.56	0.82	**11**	DRD2		105	0.63
2	GATA6F08	D2S440	103.16	0.68	**11**	Mfd316	D11S1377	120.87	0.78
2	AFMa126zb1	D2S2161	105	0.78	**11**	AFM331yc5	D11S1353	122.47	0.79
2	Mfd337	D2S417	106.84	0.75	**11**	GATA64D03	D11S4464	123	0.78
					11	AFM240ye1	D11S933	124.07	0.8
8	ATA19G07	D8S1119	101.01	0.8	**11**	Mfd251	D11S975	126.21	0.79
8	AFM147yb6	D8S1707	101.69	0.7					
8	AFM165yb10	D8S271	102.62	0.77	**20**	GATA45B10	D20S480	79.91	0.74
8	GATA8B01		103.69	0.74	**20**	GATA46C01	D20S1085	82.07	0.86
8	MTG8*		103.69	0.84	**20**	AFM276xh1	D20S120	83.51	0.85
8	T7-27**		103.69	0.43	**20**	GGAA11E12	D20S469	84.78	0.79
8	AFM165xh4	D8S270	103.69	0.79					
8	GATA28F12		104.33	0.63	**21**	GATA45C03		31.26	0.7
8	AFMa052wh1	D8S1822	107.97	0.77	**21**	AFMa086yf9	D21S1920	31.99	0.75
8	GATA23D12	D8S1129	110.2	0.29	**21**	AFMb280xd9	D21S1895	33.84	0.82
8	AFM077ya5	D8S257	111.68	0.71	**21**	AFM261zg1	D21S1252	35.45	0.8
8	AFM352td9	D8S559	112.42	0.72					
8	AFM333vb9	D8S1808	113.16	0.53					

* MTG8 3'UTR microsatetellite polymorphism (Wolford et al. 1998)

** T7-27 primers: AGCATAACATTGCTGCTAGAG and GTTACTTTGCAGATCTTTGAGC

Statistical Analysis

Several different statistics were used to evaluate marker genotypes (calculation of p-values is described in paragraph 4 below):

1. Single marker transmission disequilibrium test (TDT), first proposed by Spielman et al. (1996), based on a statistical comparison of the frequencies of transmissions and nontransmissions of a marker alleles from heterozygous parents to affected offspring. The distribution of transmitted versus nontransmitted alleles were compared using the T_m (χ^2) TDT statistic (Bickeböller and Clergot-Darpoux 1995).

2. Single marker haplotype relative risk (HRR) method, first proposed by Falk and Rubinstein (1987), uses two alleles of the parents (both homo- and heterozygous for particular marker alleles) which are not transmitted to their affected offspring (only one affected offspring for each pair of parents is used) to create a so called "pseudo-control" (Schaid 1998). All case alleles are from the first affected child in each pedigree. The control alleles are assigned in relation to the number of typed parents, using the following rules: (i) if both parents are typed then the two nontransmitted alleles in the parents of the affected children are included in the control sample; (ii) if only one parent is typed and there is an unambiguous solution (i.e., parent and child have different genotypes) then the one remaining (nontransmitted) allele is included in the control sample; (iii) if neither parent is typed, or if only one parent is typed and has the same genotype as the affected child, in which case the probabilities of the transmitted and nontransmitted alleles are ambiguous, then no allele(s) is included in the control sample. The association of disease with marker alleles was then assessed by a traditional case-control χ^2 statistic.

3. Multipoint TDT and HRR methods were devised to increase meiotic informativeness within families. Multi-marker haplotypes were built across two or three contiguous markers (i.e., 1-2, 2-3, 3-4 and 1-2-3, 2-3-4, 3-4-5, etc.). A

transmitted haplotype consisted of alleles at contiguous loci which could be unambiguously determined to be transmitted by a parent to an affected offspring, and non-transmitted haplotypes were defined analogously. The only case in which transmitted haplotypes are ambiguous occurs when both of the typed parents plus the affected child are all heterozygous for the same alleles at the same locus, in which case these individuals were excluded from the analysis. The frequencies of these transmitted and non-transmitted haplotypes were then statistically tested for linkage and association using the same methods used for the single marker tests.

4. Because the large number of allele and haplotype combinations produce large, sparse contingency tables, significance for all TDT and HRR tests was determined by exact methods. Associated p-values were approximated by Monte Carlo simulation as implemented in the StatXact3 program (CYTEL Software Corp., Cambridge, USA) with 10,000 iterations. Also because this is a replication analysis focussing on previously identified genomic regions, p-values are reported without correction for multiple testing.

Results

Genotypes for 4, 13, 6, 4 and 4 markers (Table 1), mapping to the regions on chromosomes 2, 8, 11, 20 and 21, respectively, previously found to be associated with GTS, were tested for linkage and association in 91 independent families. Results of single and multiple marker TDT and HRR analyses are summarized in Table 2. Single marker TDT analysis showed evidence for linkage or association, producing p-values of 0.039, 0.056 and 0.022, for markers D2S139, GATA28F12 and D11S1377 on chromosomes 2, 8 and 11, respectively. Extended TDT and HRR analysis provided further evidence for linkage or association on chromosomes 2 and 8, producing two-marker p-values of 0.007 and 0.025, and 0.059 and 0.013, respectively. Moreover, three-marker HRR analysis on chromosome 8 also showed significant association, producing a p-value of 0.011. No evidence, however, was found for linkage or association with markers on chromosomes 20 and 21.

Table 2

Results from TDT and HRR Analyses

	Original Case-Control Study (5)			Follow Up Familial TDT and HRR Study			
					Extended Haplotypes		
Chromo Some	Locus	cM from p-tel (sex-avg)	*P*	1-Locus TDT *P*	2-Locus TDT *P*	2-Locus HRR *P*	3-Locus HRR *P*
2	D2S139	101.56		**0.039**	**0.007**	**0.025**	
	D2S440	103.16	0.002	0.734			
	D2S417	106.84		0.160			
8	D8S1119	101.01	0.01	0.349			
	T7-27	103.69		0.835			
	D8S270	103.69		0.823	**0.059**	**0.013**	**0.011**
	GATA28F12	104.33		**0.056**			
	D8S257	111.68	0.01	0.638			
11	D11S1377	120.87	< 0.000001	**0.022**	0.109	0.108	
	D11S1353	122.47		0.135			
	D11S933	124.07	0.0009	0.535			
20	D20S1085	82.07	0.0001	0.240			
	D20S469	84.78	0.1	0.411			
21	GATA45C03	31.26	0.0004	0.221			
	D21S1252	35.45	0.000008	0.279			

Discussion

Tests for disease-marker association within nuclear families via TDT or HRR are generally considered more rigorous than case-control studies. We therefore recruited Afrikaner families with affected children in an attempt to confirm our original case-control GTS results. Although we used a relatively small sample, we were still able to confirm 3 out of 5 chromosomal regions identified previously. Regions centered at about 2p11, 8q22 and 11q23-24 are therefore strengthened as locations of genes which influence GTS in the Afrikaner population.

There are several potential reasons why we could not confirm the remaining two regions on chromosomes 20 and 21. Instead of true association with disease-predisposing alleles, the results for these chromosomes could be due to factors such as population admixture, slight differences in ancestry between cases and controls, and multiple founding alleles. Alternatively, the original findings could be true, but replication failed because we searched in the wrong locations or selected markers, which had particularly high mutation rates or had common alleles associated with the disease allele.

In general, TDT has greater power to detect linkage for a 'recessive'-type model of a genetic trait than for a 'dominant'-type. Its power is also higher when there is a greater difference in marker allele frequency between disease and normal chromosomes and the marker mutation rate is minimal (Xiong and Guo 1998). Given these limitations, it is exciting that we were able to confirm even a fraction of the original locations.

Our findings provide an important addition to the molecular genetics of GTS and will aid in further studies. For example, our evidence for linkage or association with GTS by TDT and HRR analysis, at marker locus GATA28F12, maps proximally to D8S257 which has previously been suggestively associated with GTS by us, and linked with GTS by Leppert et al. (1996). In addition, D8S257 is distal (< 1.5 cM) to a breakpoint (unpublished results) in a family with GTS co-segregating with a chromosome translocation t(1,8) (q21.1, q22.1) described by Devor and Magee (in press).

In summary, the results from this study represent significant replication and strong evidence for the presence of susceptibility loci in some of these genomic regions. In particular, the involvement of chromosomes 8q22, 2p11, and 11q23-24 as candidate regions for GTS susceptibility deserves further study. We plan to continue these investigations either by increasing our current sample size; by focusing on historical subgroups of Afrikaners and thus escalating the chances for detection of shared haplotypes; and ultimately by screening of candidate genes in these regions.

Acknowledgments

The authors wish to express their gratitude to all patients and their family members for active participation in the study, to Mrs Cicely van Straaten, the chair of the GTS Support Group in Pretoria and Dr Marie Torrington for their help in contacting and recruitment of GTS families, and also Mrs Donna David for her expert assistance in the laboratory. This work was supported by a grant from the Tourette Syndrome Association and grant HG00008 from the National Human Genome Research Institute.

Electronic-Database Information

Center for Medical Genetics, http://www.marshmed.org/genetics (for polymorphism and genetic map information)

References

Bickeböller H and Clergot-Darpoux (1995) Statistical properties of the allelic and genotypic transmission/disequilibrium test for multiallelic markers. *Genet Epidemiol* 12:865-870

Broman KW, Murray JC, Sheffield VC, White RL, Weber JL (1998) Comprehensive human genetic maps: individual and sex-specific variation in recombination. *Am J Hum Genet* 63:861-869

Devor EJ, Magee HJ (1999) Multiple childhood behavioral disorders (Tourette syndrome, multiple tics, ADHD and OCD) presenting in family with balanced chromosome translocation t(1;8)(q21.1;q22.1) *Psychiatric Genetics*. In press.

Falk CT, Rubinstein P (1987) Haplotype relative risks: an easy reliable way to construct a proper control sample for risk calculations. *Ann Hum Genet* 51:227-233

Kidd KK (1993) Associations of disease with genetic markers. *Déjà vu* all over again. *Am J Med Genet* 48:71-73

Lander ES, Kruglyak L (1995) Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat Genet* 11:241-247

Leppert M, Peiffer A, Snyder B, van de Wetering, Filloux F, Coon H, McMahon WM (1996) Two loci of interest in a family with Tourette syndrome. *Am J Hum Genet* 59(Suppl):A225

Price RA, Kidd KK, Cohen DJ, Pauls DL, Leckman JF (1985) A twin study of Tourette syndrome. *Arch Gen Psychiatry* 42:815-820

Schaid DJ (1998) Transmission disequilibrium, family controls, and great expectations. *Am J Hum Genet* 63:935-941

Simonic I, Gericke GS, Ott J, Weber JL (1998) Identification of genetic markers associated with Gilles de la Tourette s yndrome in an Afrikaner population. *Am J Hum Genet* 63:839-946

Spielman RS, Ewens WJ (1996) The TDT and other family-based tests for linkage disequilibrium and association. *Am J Hum Genet* 59:983-989

Walkup JT, LaBuda MC, Singer HS, Brown J, Riddle MA, Hurko O (1996) Family study and segregation analysis of Tourette syndrome: evidence for a mixed model of inheritance. *Am J Hum Genet* 59:684-693

Wolford JK, Bogardus C, Prochazka M (1998) Polymorphism in the 3' untranslated region of MTG8 is associated with obesity in Pima Indian males. *Biochemical and Biophysical Research Communications* 246:624-626

Xiong M, Guo SW (1998) The power of linkage detection by the transmission/disequilibrium tests. *Hum Hered* 48:295-312

CHAPTER 9

Discussion

9.1. Contributions made against the background of the current state of knowledge concerning the genetics of neuropsychiatric disorders.

Despite intense international efforts, utilization of classical genetic linkage studies and affected sib pair studies have not yielded any strong and consistent leads for any major psychiatric disorder such as schizophrenia, bipolar affective disorder, or behavioral traits such as the alcoholism.

Positive findings from numerous whole genome scans are scattered throughout the genome with no apparent consistency among the results. Some chromosomes have not been putatively linked to any psychiatric disease. Others contain large genomic regions, with a spread of positive findings well over 40 cM, which may contain a susceptibility locus to several distinct neuropsychiatric traits. At this stage, however, none of the findings appear strong enough to be followed by physical cloning and sequencing of surrounding regions (DeLisi and Crow 1999).

Views vary widely from the notion that there is one major gene for all psychoses, to a single major gene within some families but several different genes within separate populations, to the concept that a number of interacting genes are necessary for the expression of the disorder within one individual. If the multitude of positive regions reported for neuropsychiatric traits by various research teams represent true psychosis vulnerability loci, this would support the latter concept (DeLisi and Crow 1999) as well as the view that some of the susceptibility loci will have different prevalence rates in different populations. Under such circumstances coordinated international collaboration in genetic mapping of neuropsychiatric traits is essential, including sample collection in various populations and integration of all existing and newly generated data.

A good example of such an international effort is the work of the International Tourette Syndrome Genetic Linkage Consortium, created as a result of the Tourette Syndrome Association (TSA) initiative. The Consortium has supported international collaborative studies for over 10 years, with both administrative and direct research funding provided by the TSA. The Consortium includes investigators from 11 centers in United States, Canada, United Kingdom, Netherlands, as well as the Neurogenetics Research Initiative at the Medical Research Council in Pretoria, South Africa.

The molecular genetic investigations presented in this study were initiated and partially supported by the Tourette Syndrome Genetic Linkage Consortium and signify an important step forward in Tourette syndrome gene mapping efforts as well as mapping genes for neuropsychiatric illness in general.

The design and the results achieved throughout the subsequent studies provide evidence that initial case-control association findings in special populations have the potential to be confirmed by linkage in the nuclear family design. These studies also suggest that several genes at different chromosomal regions shape the susceptibility to Tourette syndrome. Even though the initial hypothesis of single founder effect for Tourette syndrome in the Afrikaner population was probably incorrect, this type of population may still prove to be an extremely valuable source of genetic material for mapping complex traits, as documented in this study.

9.11. The possible significance of cytogenetic findings

It is sometimes assumed that there is no longer much scope for chromosome studies in genetic research because everything is known about phenotypic consequences of chromosome abnormalities and also because traditional cytogenetic techniques are slowly overtaken by molecular methods (Rutter 1994). Despite these developments, further exploration of cytogenetic investigations is needed:

1. in the area of submicroscopic chromosomal deletion/duplication(s) investigations in relation to complex phenotypes which are often characterized by comorbidity and phenotypic overlaps as exemplified by the contiguous gene syndromes;
2. in the identification of cytogenetically detectable anomalies in proband(s) which may help to accelerate the search for disease gene location, as it was in case of neurofibromatosis type 1 or Duchenne muscular dystrophy;
3. and finally, in the search for disease-specific chromosome fragile sites, which may be expressed in the affected individuals as a consequence of sequence variations, such as repeat expansions, at the disease locus.

Chromosome fragile sites (FS) are regarded as primary targets for environmental insults and as such may facilitate subtle DNA changes during prenatal and postnatal development resulting in susceptibility/resistance to a large variety of complex genetic traits and/or somatic cell mutations involved in malignancy. The role of FS in amplification, loss of heterozygosity (LOH), or gain of function at or near cancer related genes is currently well recognized and there are several vague, but intriguing indications that a subgroup of chromosome FS may give rise to genetic variations involved in the expression of some neurological disorders or may appear consequently to such genetic variations. The best known example of such a relationship is the rare FS on chromosome X (FRAXA).

Our findings of increased breakage rates at large number of common fragile sites in two subsequent case-control studies involving Tourette syndrome probands indicate that for unknown reasons their lymphocyte DNA at the chromosomal level of organization is more sensitive to the inhibition of DNA synthesis (due to folic acid deficiency or aphidicolin treatment).

The number of FS exhibiting increased expression rates between Tourette syndrome and non-Tourette syndrome samples is high and the fragility itself may be influenced by numerous sample and culture related conditions, therefore the FS expression cannot serve as discriminatory cytogenetic marker for Tourette syndrome and it is improbable that the individual FS represent candidate chromosome regions for Tourette syndrome loci.

Overall, there are several possible explanations for the findings of increased fragility rates in several groups of Tourette syndrome probands, two of which are:

1. As discussed in chapter 6 one possible cause of increased fragility rates may be the lack or the presence of distinct T-lymphocyte type(s) in the Tourette-positive group of samples as opposed to controls. This may be a consequence of an autoimmune reaction to certain types of bacterial infection, which is considered by some researchers associated with phenotypic expression of Tourette syndrome in at least a subgroup of cases.

2. Another alternative, not discussed in Chapter 2, 3 and 6 could relate to a prolonged cell cycle in mitogen-stimulated T-lymphocytes of Tourette syndrome probands. Such a prolonged cell cycle could result in the observation of earlier metaphase stages, consequently leading to an increased number of gaps and breaks in the chromosome structure due to DNA under-replication caused by aphidicolin treatment.

Similar findings were documented for the group of genetic conditions known as chromosome breakage syndromes, which are characterized by increased breakage rates observed in cultured lymphocytes as a result of faulty DNA repair and packaging. Completion of DNA replication/repair/packaging is an important check-point for cytokine induced initiation of mitosis, and if delayed may directly influence the duration of the whole cell cycle.

If an increased FS expression in Tourette syndrome T-lymphocyte cultures is a function of a prolonged cell cycle, then genes directly or indirectly influencing the duration of certain cell cycle stages would be good candidates for Tourette syndrome susceptibility. Such candidates must, however, be different from the group of genes involved in the expression of chromosome breakage syndromes, which are characterized by a high prevalence of childhood cancers.

Increased rates of fragile site expression were also found in association with other disease states accompanied by neurological and psychiatric disturbances (Petronis and Kennedy 1995), e.g. schizophrenia (Garofalo et al. 1993, Chen et al. 1998), type I bipolar disorder (Turecki et al. 1995), autism (Arrieta et al. 1996), and Rett syndrome (Telvi et al. 1994, Simonic et al. 1997). Increased chromosome breakage was also observed in autoimmune disorders, e.g. scleroderma (Emerit et al. 1976, 1980), rheumatoid polyarthritis (Vincent et al. 1986), and celiac disease (Fundia et al. 1996), suggesting that some common underlying mechanisms (certain type of immune responses?) play a role in phenotypic expression of all the above conditions including Tourette syndrome. In the case of Tourette syndrome, the latter conclusion is supported by a growing body of evidence that a subset of Tourette syndrome cases could be related to the presence of D8/17, an immune marker for rheumatic fever.

9.12. The significance of association/linkage findings

For most neurological, psychiatric, or neurobehavioral genetic disorders, the patterns of inheritance has proven to be more complicated than previously thought. Examples include schizophrenia, bipolar affective disorder, alcohol dependence, autism, obsessive compulsive disorder (OCD), attention deficit hyperactivity disorder (ADHD), Tourette syndrome, conduct disorder, etc., where disease-associated psychopathologies presumably develop due to multiplicative and/or additive interactions of several genes rather than a single major gene. The interplay of several genes as well as unknown environmental factors could well explain other common features of this group of complex traits, such as clinical heterogeneity and comorbidity. These factors together with high phenocopy rates in the populations and frequent bilineal transmissions of susceptibility factors in families contribute to overall resistance of the traits to gene-mapping efforts.

Heutink et al. (1995) performed an extensive computer simulation study in order to investigate the best strategy for mapping Tourette syndrome gene(s) with respect to data collection and diagnostic classification of the disease. Their results suggested, that the scenario of multiple tests that included various Tourette syndrome diagnostic

models (narrow, mild, and broad) gives much higher probability for detecting linkage than does the use of a single test including narrowly defined phenotype. However, classical linkage studies using large pedigrees with multiple affected relatives have not produced unambiguous results under the assumption of a single dominant model of inheritance, despite incorporating different diagnostic models into data analysis. On the contrary, the implementation of classical linkage methods resulted in exclusion of >90% of the genome for harboring Tourette syndrome genes.

Complexities arising with more complicated models of inheritance cannot be assumed when parametric methods, such as classical linkage methods, are used for data analysis. Therefore, a great deal of hope in current Tourette syndrome gene-mapping efforts is focused on the notion that non-parametric methods, such as the affected sib-pair method (ASP), which do not require assumptions with regard to a specific genetic model will provide a more powerful tool for genetic mapping. As a result, the International Tourette Syndrome Genetic Linkage Consortium has initiated a preliminary ASP genome-wide linkage study using hundred affected siblings collected at several sites in Canada, United States, United Kingdom and Netherlands. The results of single and multipoint maximum likelihood score (MLS) analyses were recently submitted for a publication in Am J Hum Genet (Aug 1999).

The ASP approach is expected to have sufficient power for detection of linkage with complex traits only when extensive samples are subjected to statistical analysis, requiring about 200 to probably a thousand or more affected siblings. A similar situation with respect to robustness accounts for other analytical approaches based on nuclear family or case-control data collection. To overcome this problem, sequential approaches to data collection/generation/analysis have been proposed, based on subsequent follow-up studies using smaller sets of samples. Moreover, to avoid shortcomings of each different statistical method, combined semiparametric methods for mapping complex traits were established via linkage/linkage disequilibrium (LD) analysis.

Two critical tasks regarding sequential and combined approaches in study designs are the selection of genetic markers with respect to their average genetic map spacing and

their mean heterozygosity, as well as the interpretation of the results and statistically generated significance levels. With this regard, a number of general rules were established:

1. The average spacing of the markers in the whole genome searches has been recommended as < 0.5 cM for LD mapping efforts in heterogeneous populations and between 5-10 cM in young (< 20 generations) genetically isolated populations. An efficient strategy in scanning the chromosomes for linkage/LD with disease genes also involves typing markers along sparse map, throughout the genome, and then following up on promising areas from the initial scan with a dense array of markers, to extract the full inheritance information (Durham and Feingold 1997).

2. Historically, most methods for detecting LD (in particular) were designed for use with diallelic marker loci, for which the analysis is straightforward. These analytical methods have now been extended for use of polymorphic markers with many alleles, leading to an increase of the number of observations and corresponding loss of power. Alternatively, multiallelic systems were reduced to diallelic systems by selecting one allele believed to be associated with the trait and collapsing the other alleles, leading to the introduction of substantial bias into the analysis (Terwilliger 1995). Ott and Rabinowitz (1997) examined the relationship between marker heterozygosity and the power to detect LD and concluded that, despite the penalties for multiple testing incurred with multiple alleles, in general greater heterozygosity of markers results in greater power.

3. In interpreting linkage/LD findings, consistent significant results are usually considered suggestive for the identification of a candidate gene locus. However, inconsistent findings with positive linkage results and negative LD results could have several different interpretations: a/ markers significantly linked with disease locus may still be located far away from the actual disease locus, resulting in linkage-equilibrium; b/ linked markers identify the disease locus, but may have multiple disease-causing mutations, diminishing the allelic association with closely spaced markers; 3/ the linkage signals could be false-positive. On the other

hand, an inconsistent finding with positive LD and negative linkage is likely to suggest that LD resulted from population-based forces other than co-segregation of disease and marker alleles. Finally, a consistent negative result between LD and linkage analyses provides a strong support that the corresponding region is less likely to include any major disease locus (Zhao et al. 1999).

4. The interpretation of significance levels resulting from multi-locus analysis (particularly genome-wide) is matter of current debate and persistent controversy. Therefore, repeated replication of statistically significant or suggestive linkage/LD findings from several independent experiments is still regarded as the best evidence for candidate disease gene location.

The study design, presented in this thesis, attempted to identify candidate gene location(s) for Tourette syndrome based on a modified sequential and semi-parametric model of analysis. Several independent studies were performed including an initial whole-genome search for association (case-control design), a confirmatory LD study with denser marker maps in regions of interest and two independent non-overlapping groups of case-control samples, and finally the confirmatory linkage/association study of case-control data based on nuclear family design using TDT and HRR statistics. Throughout the study an attempt was made to comply with all the recommendations discussed earlier with respect to marker density and heterozygosity, as well as the interpretation of the results.

To date, three genetic regions were identified on chromosomes 8q22.1-q22.3, 11q23-24, and 2p11 during this study, as being suggestive for Tourette syndrome candidate loci.

The chromosome 8q22 region was found to manifest a translocation breakpoint in a family with a cytogenetically balanced chromosome translocation co-segregating with Tourette syndrome. The breakpoint at 8q22 was found in close proximity to the marker D8S257, associated with Tourette syndrome in our study, and the *MTG8* gene (Weber, unpublished), also known as the *CDR/ETO* gene (MIM 133435). The 'myeloid translocation gene on 8q22 (*MTG8*), is juxtaposed to the acute myeloid

leukemia (*AML1*) gene, which is the most frequent target for the AML associated chromosome translocation t(8;21). The *MTG8* gene encodes a protein with two putative zinc finger motifs and several proline-rich regions, and is presumed to function as a transcription factor (Kitabayashi et al. 1998). The *MTG8* mRNA is abundantly expressed in human adipose tissue, skeletal muscle and neural cells (Sacchi et al. 1998) and as such might be a good candidate for Tourette syndrome gene.

Several other genes have been mapped at or near the 8q22 region, e.g. the *tsBN51* gene (MIM 187280), encoding for a temperature sensitivity complementation protein with cell cycle-specific expression. The mutations in the *tsBN51* gene lead to a block in progression through the G1 phase of the cell cycle at non-permissive temperatures. Another interesting gene mapped to the region is *GEM* (MIM 600164), a GTP-binding mitogen-induced T-cell protein, which is transiently overexpressed in human peripheral blood cells after the mitogen induction. The product of the gene has been found to be overexpressed in skeletal muscle and in individuals with type II diabetes mellitus. Both the above genes might be good candidates for Tourette syndrome, particularly from the standpoint of chromosomal fragility findings viewed earlier in this thesis.

It is of interest, that the search for shared haplotypes at the 8q22 region among Tourette syndrome probands based on Afrikaner nuclear family data revealed only two individuals who shared an haplotype identical by state (IBS) throughout the investigated region (Appendix, Table 3). By narrowing the region towards marker GATA28F12, which yielded the highest P-value with TDT statistics, the number of individuals sharing an IBS haplotype increased [Note, that two alleles 12 and 13 were pooled for the MTG8 polymorphism (Table 3). The reason is a high mutation rate of the marker, in which case clustering of neighboring alleles is common even in statistical analyses for linkage/LD. The Afrikaner haplotyping data have not been statistically evaluated and Table 3 is included in appendices for illustration purposes only.] However, significant values for haplotype sharing among affected individuals were not reached as one would expect from assuming a single founder effect for Tourette syndrome among Afrikaners. Without detailed genealogical study of the

ancestry of families sharing IBS haplotypes at 8q22, it is hard to predict whether identified regions are also identical by descent (IBD), and if so, whether they are linked to Tourette syndrome or simply occur as a result of the background kinship in our study population.

The region on chromosome 11q23, suggestively linked to Tourette syndrome throughout this study (Chapter 5 and 8), has previously been extensively investigated by others particularly because it harbors the dopamine D2 receptor (*DRD2*) gene. The initial positive association finding between Tourette syndrome and the A1 allele of the *DRD2* (MIM126450) gene has been repeatedly disproved by linkage studies in large pedigrees. Due to the genetic location of the markers D11S933 and D11S1377 (Chapter 5) mapped distal to the *DRD2* locus (~15.1 cM), transmission of one out of three allelic polymorhisms identified within *DRD2* gene was investigated in Afrikaner nuclear families. The investigation did not yield any significant distortion of allelic transmission patterns in the group of samples used for the study reported in this thesis (Appendix, Table 5) which leads to the conclusion, that if there is a gene playing a significant role in susceptibility to Tourette syndrome at the 11q23 region, it is probably distinct from the DRD2 gene and most likely located distal to the DRD2 locus.

Another obvious candidate for the Tourette syndrome gene at the 11q23.1-q23.2 region is the 5-hydroxytryptamine (serotonin) receptor-3 gene (*HTR3*, MIM182139), which has not been investigated for association with Tourette syndrome before.

It is of interest, that the ASP study of the Tourette Syndrome Genetic Linkage Consortium identified marker D11S912 as one of 18 markers, which yielded positive single- and multipoint maximum likelihood scores (MLS) in their whole genome screen. Marker D11S912 is located distal from marker D11S1377 (~11 cM) and marker D11S933 (~7 cM away).

The search for a common haplotype across the 11q23 region in South African patients did not reveal any IBS region present in > 4 individuals. It is, however, of interest, that a rare allele (236) of marker D11S975 present only four times across the entire

sample of >200 individuals (Appendix, Table 5), has been transmitted to affected probands three times, each time on the same background (IBS) haplotype (Appendix, Table 4). Only in one case was the allele not transmitted to the affected proband, in which case it resided on a different background haplotype. A denser map of markers has to be typed at chromosome 11q23 to see whether this particular IBS haplotype also demonstrates IBD haplotype.

There have been no previous reports of Tourette syndrome linkage with the chromosome 2p11 region. As illustrated in Fig.1 (Appendix) several markers along the whole length of the chromosome 2 were identified in the initial Afrikaner genome screen (Chapter 5) with consistent differences in marker allele distribution between affected and control pools. However, only two loci have subsequently been confirmed for association by individual typings (Appendix, Table 1 and 2). Increased haplotype sharing identified by extended HRR analysis among the markers at the 2p11 region may either reflect the association with a putative Tourette syndrome gene locus, or may alternatively represent high levels of background LD in the region, which has been documented for the same region in the Afrikaner background LD study (Appendix, Figure 2). Such high levels of LD may be explained by centromeric chromosomal location of the region, a site, which is usually characterized by lower recombination rates.

There have been no obvious candidates for Tourette syndrome mapped to the 2q11 region, however, genes important with respect to immune responses, such as CD8 antigen, alpha polypeptide (*CD8A*, MIM186910), and immunoglobulin kappa light chain gene cluster (*IGK@*) been mapped to the region.

A gene, such as transcriptional factor-9 (*TCF9*, MIM189901), which binds GC-rich sequences (common fragile sites cluster in GC-rich chromosome regions), and maps to 2p11.2-p11.1, may also serve as a candidate gene in future studies with attempt to elucidate the genetics of Tourette syndrome.

It is intriguing, that the 2q12-q11 region has also been associated with alcohol dependence in genome-wide linkage investigation of the Collaborative Study on the

Genetics of Alcoholism (COGA), with one of two largest lod scores in the region being achieved with marker D2S1790 (Reich et al. 1998), positively associated with Tourette syndrome in the Afrikaner study.

The regions on chromosome 20q and 21q21, associated with Tourette syndrome throughout the case-control study (Chapter 5), have not been confirmed for linkage by subsequent nuclear family study design (Chapter 8). The high original P-values for the markers in the region most likely resulted from spurious associations detected in a rather small initial group of samples (Appendix, Table 1 and 2).

The remaining 10 regions associated with Tourette syndrome in earlier case-control study (Chapter 5) have yet to be subjected to confirmatory studies. Particularly the region on chromosome 1p is of interest, where two markers (D1S485 and D1S1665) span a ~34 cM region harboring marker D1S1728, which also yielded a suggestive single point MLS value in the genome-wide ASP study of the Tourette Syndrome Genetic Linkage Consortium. The region has also been suggestively linked to schizophrenia in several recent studies (Shaw et al. 1998), bipolar disorder (Rice et al. 1997), and alcohol dependence (Reich et al. 1998).

The chromosome 6p22-p24 region has been subjected to extensive linkage and ASP analyses in connection with common psychoses, since the report by Straub et al. (1995) of linkage to schizophrenia. While markers which yielded suggestive association results throughout the Afrikaner study, D6S477 and D6S470 (Appendix, Table 1 and 2), map to the 6p24-p25 region, a marker identified by the Tourette Syndrome Genetic Linkage Consortium in the ASP study with a positive (>1) single point MLS value (D6S1053) maps proximal to the region (55.17 cM) in the HLA and dyslexia gene region (Grigorenko et al. 1997, Fischer et al. 1999, Gayán et al. 1999).

Wide-spread positive linkage findings throughout the 6p region (~40 cM) are documented in the search for schizophrenia genes and may also become common in the search for Tourette syndrome genes as more linkage/association studies using non-parametric methods of analysis are performed. The HLA antigens (HLA-A, B, C,

and DR) have been excluded from linkage with Tourette syndrome in earlier studies (Caine et al. 1985).

Chromosomes 12, and 14 markers associated with Tourette syndrome in the initial genome scan (Appendix, Table 1 and 2), have not been repeatedly associated with either Tourette syndrome or other major neurological or psychiatric traits (Detera-Wadleigh 1999, Craddock and Lendon 1999).

Chromosome 5q and 13q regions including markers D5S666 and D13S788 (Appendix, Table 1 and 2) have been suggestively linked to schizophrenia and bipolar disorder respectively (Crowe and Vieland 1999, Barden and Morissette 1999).

One of the potentially important areas also includes the Xp11.4-p21 region, where differences were observed in marker allele distributions for two markers in the initial genome scan with DNA pools from affected and control males only (Appendix, Figure 1). These initial findings could not be confirmed with subsequent evaluation of individual genotyping data (Appendix, Table 1 and 2). However, because of the fact that the male- to female-ratio in the prevalence of Tourette syndrome as well as other childhood onset neurological disorders (e.g. autism, ADHD) is distorted, our initial finding may still be of importance because of the possibility of imprinted susceptibility loci at the Xp chromosome region.

Further investigations are necessary to support this claim, e.g. by separate investigations of marker allele transmissions from fathers and mothers to affected probands in nuclear families.

9.2. Future research strategies

The localization and characterization of genes important for the expression of the Tourette syndrome phenotype would represent a major advance in the understanding of the pathogenesis of this disorder and would also provide a model for the study of other developmental disorders. Furthermore, once genes conferring susceptibility to Tourette syndrome and associated behaviors have been characterized, further research

will allow the identification of additional non-genetic factors important for the manifestation as well as amelioration of the symptoms of the disorders (Pauls, 1990). It is generally accepted that co-segregation of a chromosomal abnormality with a disease phenotype, even in the case of common complex trait, may represent additional valuable information about the genetic locus or character of the gene product responsible for the susceptibility to the trait. Therefore, it is important, that at least until the genes responsible for Tourette syndrome are found, researchers will continue to search for chromosomal abnormalities in Tourette syndrome cases.

Further research is needed to finalize and confirm the molecular data generated throughout this study:

1. Confirmatory linkage/LD studies in the remaining 10 loci identified during the initial whole genome screen (Chapter 5) are of immediate importance.
2. A second complete genome scan on the Afrikaner Tourette syndrome nuclear families, which would allow replication and extension of the initial case-control association results.
3. Further examination of current chromosomal areas of interest, based on the identification of regions that are shared identical by descent (IBD) among Tourette syndrome patients. These must preferably be drawn from genealogically well-characterized and presumably more genetically isolated sub-population of Afrikaners.
4. The follow-up on the initial Xp association findings by linkage/LD analysis using nuclear families. Particularly differences in the parental marker allele transmissions (paternal vs. maternal transmissions) would be of interest in light of distorted male:female ratio among affected individuals and observed imprinting effects as a part of the inheritance pattern of certain Tourette syndrome-associated pathologies (e.g. tics vs. OCD).

In agreement with the current goals of the Tourette Syndrome Genetic Mapping Consortium, the recommended gene-mapping designs for further exploration of the existing and newly generated data will incorporate extensive clinical information about the study participants for future quantitative trait loci analysis (QTL). E.g. tics

plus associated psychopathologies (either OCD or ADHD, etc.) vs. tics only may serve as one discriminatory factor for QTL analysis, the presence/expression vs. non-presence of D8/17 antibody in Tourette syndrome probands may serve as another discriminatory factor.

The co-localization of some of the currently identified genetic regions of interest for Tourette syndrome with regions linked to other psychoses (schizophrenia, bipolar affective disorder) or neurobehavioral traits (alcohol dependence) may also serve as an alternative basis for further investigations aiming to identify susceptibility genes for this intriguing condition.

REFERENCES

Allen AJ, Leonard HL, Swedo SE (1995) Case study: a new infection-triggered, autoimmune subtype of pediatric OCD and Tourette's syndrome. J Am Acad Child Adolesc Psychiatry 34(3):307-311

Alsobrook JP II, Pauls DL (1997) The genetics of the Tourette syndrome. Neurol Clinics 15:381-393

Arnheim N, Strange C, Erlich H (1985) Use of pooled DNA samples to detect linkage disequilibrium of polymorphic restriction fragments and human disease: Studies of HLA class II loci. Proc Natl Acad Sci 82:6970-6974

Arrieta I, Nunez T, Gil A, Flores P, Usobiaga E, Martinez B (1996) Autosomal folate sensitive fragile sites in an autistic Basque sample. Ann Genet 39(2):69-74

Association American Psychiatric, Diagnostic and statistical manual of mental disorders, 4th ed. 1994, Washington D.C.: American Psychiatric Association

Barcellos LF, Klitz W, Field LL, Tobias R, Bowcock AM, Wilson R, Nelson M, Nagatomi J, Thomson G (1997) Association mapping of disease loci, by use of pooled DNA genomic screen. Am J Hum Genet 61:734-747

Barden N, Morissette J (1999) Chromosome 13 workshop report. (VIth World Congress on Psychiatric Genetics 1998). Am J Med Genet (Neuropsych Genet) 88:260-262

Barr CL, Sandor P (1998) Current status of genetic studies of Gilles de la Tourette syndrome. Can J Psychiatry 43:351-357

Caine ED, Weitkamp LR, Chiverton P, Guttormsen S, Yagnow R, Hempfling S, Kennelly D (1985) Tourette syndrome and HLA. J Neurol Sciences 69:201-206

Chen CH, Shih HH, Wang-Wuu S, Tai JJ, Wuu KD (1998) Chromosomal fragile site expression in lymphocytes from patients with schizophrenia. Hum Genet 103(6):702-706

Comings DE, Comings BG, Devor EJ, Cloninger CR (1984) Detection of a major gene for Gilles de la Tourette syndrome. Am J Hum Genet 36:586-600

Comings DE, Comings BG (1987) Tourette syndrome and attention deficit disorder with hyperactivity. Arch Gen Psychiatry 44:1023-1026

Comings DE (1989) Presidential address: the genetics of human behavior – lessons for two societies. Am J Hum Genet 44:452-460

Craddock N, Owen MJ (1996) Modern molecular genetic approaches to psychiatric disease. British Medical Bulletin 52(2):434-452

Craddock N, Lendon C (1999) Chromosome workshop: Chromosomes 11, 14, and 15. (VIth World Congress on Psychiatric Genetics 1998). Am J Med Genet (Neuropsych Genet) 88:244-254

Crowe RR, Vieland V (1999) Report of the chromosome 5 workshop (VIth World Congress on Psychiatric Genetics 1998). Am J Med Genet (Neuropsych Genet) 88:229-232

Curtis D (1996) Genetic dissection of complex traits (Correspondence). Nat Genet 12:356-357

Daniels J, Holmans P, Williams N, Turic D, McGuffin P, Plomin R, Owen MJ (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

Davies JL, Kawaguchi Y, Bennett ST, Copeman JB, Cordell HJ, Pritchard LE, Reed PW, et al. (1994) A genome-wide search for human type 1 diabetes susceptibility genes. Nature 371:130-136

De la Chapelle A, Wright FA (1998) Linkage disequilibrium mapping in isolated populations: the example of Finland revisited. Proc Natl Acad Sci USA 95(21):12416-12423

De la Tourette G (1885) Etude sur une affection nerveus, characterisee par l'incoordination motorice accompagnee de l'echolalie et de coprolalie. Arch Neurol 9:158-200

DeLisi LE, Crow TJ (1999) Current State of Psychiatric Linkage. Am J Med Genet (Neuropsych Genet) 88:215-218

Detera-Wadleigh SD (1999) Chromosome 12 and 16 workshop. (VIth World Congress on Psychiatric Genetics 1998). Am J Med Genet (Neuropsych Genet) 88:255-259

Devlin B, Risch N (1995) A comparison of linkage disequilibrium measures for fine-scale mapping. Genomics 29:311-322

Devor EJ (1984) Complex segregation analysis of Gilles de la Tourette syndrome: Further evidence for a major locus mode of transmission. Am J Hum Genet 36:704-709

Durham LK, Feingold E (1997) Genome scanning for segments shared identical by descent among distant relatives in isolated populations. Am J Hum Genet 61:830-842

Eapen V, Robertson MM, Alsbrook JP II, Pauls DL (1997) Obsessive compulsive symptoms in Gilles de la Tourette's syndrome and obsessive compulsive disorder: Differences by diagnosis and family history. Am J Med Genet (Neuropsych Genet) 74:432-438

Eisenberg R, Ascher E, Kanner L (1959) A clinical study of Gilles de la Tourette's disease (maladie des tics) in children. Am J Psychiatry 115:715-723

Emerit I, Housset E, Feingold J (1976) Chromosomal breakage and scleroderma: studies in family members. J Lab Clin Med 88(1):81-86

Emerit I, Michelson AM (1980) Chromosome instability in human and murine autoimmune disease: anticlastogenic effect of superoxid dismutase. Acte Physiol Scand (Suppl) 492:59-65

Falk CT, Rubinstein P (1987) Haplotype relative risks: An easy reliable way to construct a proper control sample for risk calculations. Ann Hum Genet 51:227-233

Farmer AE, Williams J, Jones I (1994) Phenotypic definitions of psychotic illness for molecular genetic research. Am J Med Genet (Neuropsych Genet) 54:365-371

Freimer NB, Service SK, Slatkin M (1997) Expanding on population studies. Nat Genet 17:371-373

Friedman TB, Liang Y, Weber JL, Hinnant JT, Barber TD, Winata S, Arhya IN, et al. (1995) A gene for congenital recessive deafness DFN3 maps to the pericentromeric region of chromosome 17. Nat Genet 9:86-91

Fisher SE, Marlow AJ, Lamb J, Maestrini E, Williams DF, Richardson AJ, Weeks DR, Stein JF, Monaco A (1999) A quantitative-trait locus on chromosome 6p influences different aspects of developmental dyslexia. Am J Hum Genet 64:146-156

Fulker DW, Cardon LR (1994) A sib-pair approach to interval mapping of quantitative trait loci. Am J Hum Genet 54:1092-1103

Fundia A, Gomez JC, Maurino E, Boerr L, Bai JC, Larripa I, Slavutsky I (1996) Chromosome instability in untreated adult celiac disease. Acta Paediatr Suppl. 412:82-84

Garofalo G, Ragusa RM, Argiolas A, Scavuzzo C, Spina E, Barletta C (1993) Evidence of chromosomal fragile sites in schizophrenic patients. Ann Genet 36(2):132-135

Gayán J, Smith SD, Cherny SS, Cardon LR, Fulker DW, Brower AM, Olson RK, Pennington BF, DeFries JC (1999) Quantitative-trait locus for specific language and reading deficits on chromosome 6p. Am J Hum Genet 64:157-164

Gecz J, Gedeon AK, Sutherland GR, Mulley JC (1996) Identification of the gene FMR2, associated with FRAXE mental retardation. Nat Genet 13(1):105-108

Glover TW (1998) Instability at chromosomal fragile sites. Recent Results Cancer Res 154:185-199

Goldenberg JN, Brown SB, Weiner WJ (1994) Coprolalia in younger patients with Gilles de la Tourette syndrome. Movement Dis 9:622-625

Göring HHH, Ott J (1997) Relationship estimation in affected sib pair analysis of late-onset diseases. Eur J Hum Genet 5:69-77

Graff C, Persson A, Ulfendahl PJ, Wahlberg J, Wadelius C (1997) Genetic mapping using fluorescent quantification of allele frequencies in pooled DNA loaded by solid support. Clin Genet 51:145-152

Greenberg DA (1993) Linkage analysis of "necessary" disease loci versus "susceptibility" loci. Am J Hum Genet 52:135-143

Grigorenko EL, Wood FB, Meyer MS, Hart LA, Speed WC, Schuster A, Pauls DL (1997) Susceptibility loci for distinct components of developmental dyslexia on chromosomes 6 and 15. Am J Hum Genet 60:27-39

Gunter C, Paradee W, Crawford DC, Meadows KA, Newman J, Kunst CB, Nelson DL, et al. (1998) Re-examination of factors associated with expansion of CGG repeats using a single nucleotide polymorphism in FMR1. Hum Mol Genet 7:1935-1946

Hallett JJ, Harling-Berg CZ, Agrawal JR, Kiessling LS (1996) Tic-like phonation and dyskinesia after intracaudate microfusion of sera from children with Tourette syndrome. FASEBJ 10:A1357

Hallett JJ, Kiessling LS (1997) Neuroimmunology of tics and other childhood hyperkinesias. Neurol Clin 15(2):333-344

Hansen RS, Canfield TK, Fjeld AD, Mumm S, Laird CD, Gartler SM (1997) A variable domain of delayed replication in FRAXA fragile X chromosomes: X inactivation – like spread of late replication. Proc Natl Acad Sci USA 94:4587-4592

Hasstedt SJ, Leppert M, Filloux M, van de Wetering BJM, McMahon WM (1995) Intermediate inheritance of Tourette syndrome, assuming assortative mating. Am J Hum Genet 57:682-689

Heubner K, Hadaczek P, Siprashvili Z, Druck T, Croce CM (1997) The FHIT gene, a multiple tumor suppressor gene encompassing the carcinogen sensitive chromosome fragile site FRA3B. Biochim Biophys Acta 1332(3):M65-70

Heubner K, Garrison PN, Barnes LD, Croce CM (1998) The role of the FHIT/FRA3B locus in cancer. Annu Rev Genet 32:7-31

Heutink P, van de Wetering BJM, Pakstis AJ, Kurlan R, Sandor P, Oostra BA, Sandkuijl LA (1995) Linkage studies on Gilles de la Tourette syndrome: What is the strategy of choice? Am J Hum Genet 57:465-473

Hewett DR, Handt O, Hobson L, Mangelsdorf M, Eyre H, Baker E, Sutherland GR, et al. (1998) FRA10B structure reveals common elements in repeat expansion and chromosomal fragile site genesis. Mol Cell 1:773-781

Hirst MC, Grewal PK, Davies KE (1994) Precursor arrays for triplet repeat expansion at the fragile X locus. Hum Mol Genet 3:1553-1560

Hodge SE (1981) Some epistatic two-locus models of disease. I. Relative risks and identity-by-descent distributions in affected sib pairs. Am J Hum Genet 33:381-395

Hodge SE (1993) Linkage analysis versus association analysis: Distinguishing between two models that explain disease-marker associations. Am J Hum Genet 53:367-384

Hodge SE (1994) What association analysis can and cannot tell us about the genetics of complex disease. Am J Med Genet (Neuropsych Genet) 54:318-323

Holmas P (1993) Asymptotic properties of affected-sib-pair linkage analysis. Am J Hum Genet 52:362-374

Holmas P, Craddock N (1995) Strategies for efficient genome screening using maximum likelihood sib-pair analysis. Psychiatr Genet 5 Suppl:S106

Houwen RHJ, Baharloo S, Blankenship K, Raeymaekers JJ, Sandkuijl LA, Freimer NB (1994) Genome screening by searching for shared segments: mapping gene for benign recurrent intrahepatic cholestasis. Nat Genet 8:380-386

Huang H, Qian J, Proffit J, Wilber K, Jenkins R, Smith DI (1998) FRA7G extends over a broad region: coincidence of human endogenous retroviral sequences (HERV-H) and small polydispersed circular DNAs (spcDNA) and fragile sites. Oncogene 16(18):2311-9

Huang H, Reed CP, Mordi A, Lomberk G, Wang L, Shridhar V, Hartmann L, Jenkins R, Smith DI (1999) Frequent deletions within FRA7G at 7q31.2 in invasive epithelial ovarian cancer. Genes Chromosomes Cancer 24(1):48-55

Hyde TM, Aaronson BA, Randolph C, Rickler KC, Weinberger DR (1992) Relationship of birth weight to the phenotypic expression of Gilles de la Tourette's syndrome in monozygotic twins. Neurology 42:652-658

Jones C, Slijepcevic P, Marsh S, Baker E, Langdon WY, Richards RI, Tunnacliffe A (1994) Physical linkage of the fragile site FRA11B and a Jacobsen syndrome chromosome deletion breakpoint in 11q23.3. Hum Mol Genet 3(12):2123-2130

Jorde LB (1995) Linkage disequilibrium as a gene-mapping tool. Am J Hum Genet 56:11-14

Kennedy JL, Billett EA, Macciardi FM, Verga M, Parsons TJ, Meltzer HY, Lieberman J, Buchanan JA (1995) Association study of dopamine D3 receptor gene and schizophrenia. Am J Med Genet (Neuropsych Genet) 60:558-562

Kidd KK (1993) Associations of disease with genetic markers: Déjà vu all over again. Am J Med Genet (Neuropsych Genet) 48:71-73

Kiessling LS, Marcotte AC, Culpepper L (1993) Antineuronal antibodies in movement disorders. Pediatrics 92(1):39-43

Kitabayashi I, Ida K, Morohoshi F, Yokoyama A, Mitsuhashi N, Shimizu K, Nomura N, Hayashi Y, Ohki M (1998) The AML1-MTG8 leukemic fusion protein forms a complex with a novel member of the MTG8(ETO/CDR) family, MTGR1. Molecular and Cellular Biology 18(2):846-858

Kruglyak L, Lander ES (1995a) High-resolution genetic mapping of complex traits. Am J Hum Genet 56:1212-1223

Kruglyak L, Lander ES (1995b) Complete multipoint sib-pair analysis of qualitative and quantitative traits. Am J Hum Genet 57:439-454

Kruglyak L (1997) What is significant in whole-genome linkage disequilibrium studies? Am J Hum Genet 61:810-812

Kruglyak L (1999) Prospects for whole-genome linkage disequilibrium mapping of common disease genes. Nat Genet 22:139-144.

Kuo MT, Sen S, Hittelman WN, Hsu TC (1998) Chromosomal fragile sites and DNA amplification in drug resistant cells. Biochem Pharmacol 56(1):7-13

Laan M, Pääbo S (1997) Demographic history and linkage disequilibrium in human populations. Nat Genet 17:435-438

Lajonchere C, Nortz M, Finger S (1996) Gilles de la Tourette and the discovery of Tourette syndrome. Includes a translation of his 1884 article. Arch Neurol 53:567-574

Lander ES, Schork NJ (1994) Genetic dissection of complex traits. Science 265:2037-2048

Lander ES, Kruglyak L (1995) Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. Nat Genet 11:241-247

Leckman JF, Grice DE, Boardman J, Zhang H, Vitale A, Bondi C, Alsbrook J, Peterson BS, Cohen DJ, Rasmussen SA, Goodman WK, McDougle CJ, Pauls DL (1997) Symptoms of obsessive compulsive disorder. Am J Psychiatry 154:911-917

Le Beau MM, Rassool FV, Neilly ME, Espinosa R, III, Glover TW, Smith DI, McKeithan TW (1998) Replication of a common fragile site, FRA3B, occurs late in S phase and is delayed further upon induction: implications for the mechanism of fragile site induction. Hum Mol Genet 7:755-761

Leckman JF, Zhang H, Vitale A, Lahnin F, Lynch K, Bondi C, Kim YS, Peterson BS (1998) Course of tic severity in Tourette syndrome: The first two decades. Pediatrics 102:14-19

Leonard HL, Lenane MC, Swedo SE, Rettew DC, Gershon ES, Rapoport JL (1992) Tics and Tourette's disorder: A 2- to 7-year follow-up of 54 obsessive-compulsive children. Am J Psychiatry 149:1244-1251

Li Z, Bailey AD, Buchowski J, Weiner AM (1998) A tandem array of minimal U1 small nuclear RNA genes is sufficient to generate a new adenovirus type 12-inducible chromosome fragile site. J Virol 72(5):4205-4211

Luo ZW (1998) Detecting linkage disequilibrium between a polymorphic marker locus and a trat locus in natural populations. Heredity 80:198-208

Michailis RC, Velagaleti GV, Jones C, Pivnick EK, Phelan MC, Boyd E, Tarleton J, Wilroy RS, Tunnacliffe A, Tharapel AT (1998) Most Jacobsen syndrome deletion breakpoits occur distal to FRA11B. Am J Med Genet 76(3):222-228

Morton ME (1998) Significance levels in complex inheritance. Am J Hum Genet 62:690-697

Murphy TK, Goodman WK, Fudge MW, Williams RC, Jr, Ayoub EM, Dalal M, Lewis MH, Zabriskie JB (1997) B lymphocyte antigen D8/17: A peripheral marker for childhood-onset obsessive-compulsive disorder and Tourette's syndrome? Am J Psychiatry 154(3):402-407

Nancarrow JK, Kremer E, Holman K, Eyre H, Doggett NA, Le Paslier D, Callen DF, Sutherland GR, Richards RI (1994) Implications of FRA16 structure for the mechanism of chromosomal fragile site genesis. Science 264(5167):1938-1941

Nancarrow JK, Holman K, Mangelsdorf M, Hori T, Denton M, Sutherland GR, Richards RI (1995) Molecular basis of p(CCG)n repeat instability at the FRA16A fragile site locus. Hum Mol Genet 4(3):367-372

Newport MJ, Huxley CM, Huston S, Hawrylowicz CM, Oostra BA, Williamson R, Levin M (1996) A mutation in the interferon γ-receptor gene and susceptibility to mycobacterial infection. N Engl J Med 335:1941-1949

Nomura Y, Segawa M (1982) Tourette syndrome in oriental children: clinical and pathophysiological considerations. In: Friedhoff AJ, Chase TN (eds) Gilles de la Tourette Syndrome. Raven Press, New York:105-114

Nystuen A, Benke PJ, Merren J, Stone E, Sheffield VC (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

Ott J (1989) Statistical properties of the haplotype relative risk. Genet Epidemiol 6(1):127-130

Ott J (1991) Analysis of Human Genetic Linkage, revised edition. John Hopkins University Press, Baltimore

Ott J, Rabinowitz D (1997) The effect of marker heterozygosity on the power to detect linkage disequilibrium. Genetics 147(2):927-930

Owen MJ, McGuffin P (1993) Association and linkage: complementary strategies for complex disorders. J Med Genet 30:638-639

Pacek P, Sajantila A, Syvänen AC (1993) Determination of allele frequencies at loci with length polymorphism by quantitative analysis of DNA amplified from pooled samples. PCR Methods and Applications 2:313-317

Pauls DL, Hurst CR, Kruger SD, Leckman JF, Kidd KK, Cohen DJ (1986a) Gilles de la Tourette's syndrome and attention deficit disorder: Evidence against a genetic relationship. Arch Gen Psychiatry 43:1177-1179

Pauls DL, Leckman JF (1986b) The inheritance of Gilles de la Tourette's syndrome and associated behaviors: Evidence for autosomal dominant transmission. N Eng J Med 315:993-997

Pauls DL, Leckman JF, Raymond CL, Hurst CR, Stevenson JM (1988) A family study of Tourette's syndrome: evidence against the hypothesis of association with a wide range of psychiatric phenotypes. (Abstract) Am J Hum Genet 43:A64

Pauls DL (1990) Emerging genetic markers and their role in potential preventive intervention strategies. In: Muehrer P (ed) Conceptual Research Models for Preventing Mental Disorders. NIMH: Rockville, MD.

Pauls DL, Raymond CL, Leckman JF, Stevenson JM (1991) A family study of Tourette's syndrome. Am J Hum Genet 48:154-163

Pauls DL, Leckman JF, Cohen DJ (1993) Familial relationship between Gilles de la Tourette syndrome, attention deficit disorder, learning disabilities, speech disorders, and stuttering. J Am Acad Child & Adol Psychiatry 32:1044-1050

Pauls DL, Alsbrook JP II, Goodman W, Rasmussen S, Leckman JF (1995) A family study of obsessive compulsive disorder. Am J Psychiatry 152:76-84

Paz-y-Mino C, Ocampo L, Narvaez R, Narvaez L (1992) Chromosome fragility in lymphocytes of women with cervical uterine lesions produced by human papillomavirus. Cancer Genet Cytogenet 59(2):173-176

Peltonen L (1996) Identification of disease genes in genetic isolates. Methods 9(1):129-135

Peterson AC, Di Rienzo A, Lehesjoki AE, de la Chapelle A, Slatkin M, Freimer NB (1995) The distribution of linkage disequilibrium over anonymous genome regions. Hum Mol Genet 4(5):887-894

Petronis A, Kennedy JL (1995) Unstable genes—unstable mind? Am J Psychiatry 152(2):164-172 0

Popescu NC, Zimonjic D, DiPaolo JA (1990) Viral integration, fragile sites, and proto-oncogenes in human neoplasia. Hum Genet 84(5):383-386

Popescu NC (1994) Chromosome fragility and instability in human cancer. Crit Rev Oncog 5(2-3):121-140

Price RA, Kidd KK, Cohen DJ, Pauls DL, Leckman JF (1985) A twin study of Tourette syndrome. Arch Gen Psychiatry 42:815-820

Price RA, Pauls DL, Kruger SD, Caine ED (1988) Family data support a dominant major gene for Tourette syndrome. Psych Res 24:251-261

Reich T, Edenberg HJ, Goate A, Williams JT, Rice JP, Van Eerdewegh P, Foroud T, Hesselbrock V, et al. (1998) Genome-wide search for genes affecting the risk for alcohol dependence. Am J Med Genet (Neuropsych Genet) 81:207-215

Rice JP, Goate A, Williams JT, Bierut L, Dorr D, Wu W, Shears S, Gopalakrishnan G, et al. (1997) Initial genome scan of the NIMH genetics initiative bipolar pedigrees: chromosome 1, 6, 8, 10, and 12. Am J Med Genet (Neuropsych Genet) 74:247-253

Riordan JR, Rommens JM, Kerem B, Alon N, Rozmamel R, Grzelczak Z, Zielenski J, Lok S, Plavsic N, Chou JL et al. (1989) Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. Science 245(4925):1437

Risch N (1990) Linkage strategies for genetically complex traits. I. Multilocus models. Am J Hum Genet 46:222-228

Risch N, Merikangas K (1996) The future of genetic studies of complex human diseases. Science 273:1516-1517

Robertson MM, Stern JS (1997) Gilles de la Tourette syndrome. Br J Hosp Medicine 58(6):253-256

Robertson MM, Baron-Cohen S (1998) Tourette syndrome: The Facts. Oxford University Press, Oxford, England

Rubinstein P, Ginsberg-Fellner F, Falk C (1981) Genetics of Type I diabetes mellitus: A single, recessive predisposition gene mapping between HLA-B and GLO. With an appendix on the estimation of selection bias. Am J Hum Genet 33(6):865-882

Rutter M (1994) Psychiatric genetics: Research Challenges and pathways forward. Am J Med Genet (Neuropsych Genet) 53(3):185-198

Sacchi N, Tamanini F, Willemsen R, Denis-Donini S, Campiglio S, Hoogeveen AT (1998) Subcellular localization of the oncoprotein MTG8 (CDR/ETO) in neural cells. Oncogene 16:2609-2615

Sbrana I, Zavattari P, Barale R, Musio A (1998) Common fragile sites on human chromosomes represent transcriptionally active regions: evidence from camptothecin. Hum Genet 102(4):409-414

Schaid DJ, Sommer SS (1993) Genotype relative risks: Methods for design and analysis of candidate-gene association studies. Am J Hum Genet 53(5):1114-1126

Schaid DJ, Sommer SS (1994) Need to confirm promising case-control association studies: Reply to Sham. Am J Med Genet (Neuropsych Genet) 54:156-157

Schaid DJ (1998) Transmission disequilibrium, family contrals, and great expectations. Am J Hum Genet 63(4):935-941

Scott DA, Carmi R, Elbedour K, Yosefsberg S, Stone EM, Sheffield VC (1996) An autosomal recessive nonsyndromic-hearing loss locus identified by DNA pooling using two inbred Bedouin kindreds. Am J Hum Genet 59:385-391

Service SK, Temple Lang DW, Freimer NB, Sandkuijl LA (1999) Linkage-disequilibrium mapping of disease genes by reconstruction of ancestral haplotypes in founder populations. Am J Hum Genet 64:1728-1738

Sham P (1994) Sequential analysis and case-control candidate gene association studies: Reply to Sobell et al. Am J Med Genet (Neuropsych Genet) 54:154-155

Shapiro AK, Shapiro E (1982) Clinical efficacy of haloperidol, pimozine, penfluridol, and clonidine in the treatment of Tourette syndrome. Adv Neurol 35:383-386

Shapiro E, Shapiro AK, Fulop G, Hubbard M, Mandeli J, Nordlie J, Phillips RA (1989) Controlled study of haloperidol, pimozide and placebo for the treatment of Gilles de la Tourette's syndrome. Arch Gen Psychiatry 46(8):722-730

Shaw SH, Carrasquillo MM, Kashuk C, Puffenberger EG, Chakravarti A (1998) Allele frequency distributions in pooled DNA samples: Applications to mapping complex disease genes. Genome Research 8:111-123

Shaw SH, Kelly M, Smith AB, Shields G, Hopkins PJ, Loftus J, Laval SH, Vita A, et al. (1998) A genome-wide search for schizophrenia susceptibility genes. Am J Med Genet (Neuropsych Genet) 81:364-376

Sheffield VC, Carmi R, Kwitek-Black A, Rokhlina T, Nishimura D, Duyk GM, Elbedour K, Sunden SL, Stone EM (1994) Identification of a Bardet-Biedl syndrome locus on chromosome 3 and evaluation of an efficient approach to homozygosity mapping. Hum Mol Genet 3:1331-1335

Simonic I, Gericke GS, Lippert M, Schoeman JF (1997) Additional clinical and cytogenetic findings associated with Rett syndrome. Am J Med Genet (Neuropsych Genet) 74:331-337

Slatkin M (1994) Linkage disequilibrium in growing and stable populations. Genetics 137(1):331-336

Smith DI, Huang H, Wang L (1998) Common fragile sites and cancer. Int J Oncol 12(1):187-196

Sobell JL, Heston LL, Sommer SS (1993) Novel association approach for determining the genetic predisposition to schizophrenia: Case-control resource and testing of a candidate gene. Am J Med Genet (Neuropsych Genet) 48:28-35

Spielman RS, McGinnis RE, Ewens WJ (1993) Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). Am J Hum Genet 52:506-516

Spielman RS, McGinnis RE, Ewens WJ (1994) The transmission/disequilibrium test detects cosegregation and linkage. Am J Hum Genet 54(3):559-560

Spielman RS, Ewens WJ (1996) The TDT and other family-based tests for linkage disequilibrium and association. Am J Hum Genet 59:983-989

Straub RE, MacLean CJ, O'Neill FA, et al. (1995) A potential vulnerability locus for schizophrenia on chromosome 6p24-22: evidence for genetic heterogeneity. Nature Genet 11:235-236

Sutherland GR, Baker E, Richards RI (1998) Fragile sites still breaking. Trends Genet 14:501-506

Sutherland GR, Richards RI (1999) Fragile sites – cytogenetic similarity with molecular diversity. Am J Hum Genet 64:354-359

Swedo SE, Leonard H, Mittleman BB, Allen AJ, Rapoport JL, Dow SP, Kanter ME, Chapman F, Zabriskie J (1997) Identification of children with Pediatric autoimmune neuropsychiatric disorders associated with sterptococcal infections by a marker associated with rheumatic fever. Am J Psychiatry 154(1):110-112

Tedeschi B, Vernole P, Sanna ML, Nicoletti B (1992) Population cytogenetics of aphidicolin-induced fragile sites. Hum Genet 89(5):543-547

Telvi L, Leboyer M, Chiron C, Feingold J, Ponsot G (1994) Is Rett syndrome a chromosome breakage syndrome? Am J Med Genet (Neuropsych Genet) 51:602-605

Terwilliger JD, Ott J (1992) A haplotype-based "haplotype relative risk" approach to detecting allelic association. Hum Hered 42:337-346

Terwilliger JD, Ott J (1994) Handbook of human genetic linkage. Johns Hopkins, Baltimore

Terwilliger JD (1995) A powerful likelihood method for the analysis of linkage diseaquilibrium between trait loci and one or more polymorphic marker loci. Am J Hum Genet 56(3):777-787

Terwilliger JD, Zöllner S, Laan M, Pääbo S (1998) Mapping genes through the use of linkage disequilibrium generated by genetic drift: 'Drift mapping' in small populations with no demographic expansion. Hum Hered 48:138-154

The VIth World Congress of Psychiatric Genetics (1999) Chromosome Workshops

Tienari PJ, Wikstrom J, Sajantila A et al. (1992) Genetic susceptibility to multiple sclerosis linked to myelin basic protein gene. Lancet 340:987-991

Turecki G, Smith M, Mari JJ (1995) Type I bipolar disorder associated with a fragile site on chromosome 1. Am J Med Genet 60(3):179-182

Van Woert MH, Yip LC, Balis ME (1977) Purine phosphoribosyltransferase in Gilles de la Tourette syndrome. New Engl J Med 296:210-212

Vincent G, Croquette MF, Houvenagel E, Leloire O (1986) Chromosome anomalies in rheumatoid polyarthritis. Breakage level and study of the breaking factor. Rev Rhum Mal Osteoartic 53(11):625-630

Walkup J, LaBuda MC, Singer HS, Brown J, Riddle MA, Hurko O (1996) Family study and segregation analysis of Tourette syndrome: evidence for a mixed model of inheritance. Am J Hum Genet 59:684-693

Watkins WS, Zenger R, O'Brien E, Nyman D, Eriksson AW, Renlund M, Jorde LB (1994) Linkage disequilibrium patterns vary with chromosomal location: a case study from the von Willebrand factor. Am J Hum Genet 55(2):348-355

Wilke CM, Hall BK, Hoge A, Paradee W, Smith DI, Glover TW (1996) FRA3B extends over a broad region and contains a spontaneous HPV16 integration site: direct evidence for the coincidence of viral integration sites and fragile sites. Hum Mol Genet 5(2):187-195

Witte JS, Elston RC, Schork NJ (1996) Genetic dissection of complex traits (Correspondence). Nat Genet 12:355-356

Xiong M, Guo SW (1998) The power of linkage detection by the transmission/disequilibrium tests. Hum Hered 48(6):295-312

Yu S, Mangelsdorf M, Hewett D, Hobson L, Eyre H, Lapsys N, et al. (1997) Human chromosomal fragile site FRA16B is an amplified AT-rich minisatellite repeat. Cell 88:367-374

Zhao LP, Quiaoit F, Aragaki C, Hsu L (1999) An efficient, robust, and unified method for mapping complex traits (III): Combined linkage/linkage disequilibrium analysis. Am J Hum Genet 84:433-453

Zohar AH, Pauls DL, Ranzoni G, Apter A, Dycian A, Binder M, King R, Leckman JF, Kron S, Cohen DJ (1997) OCD with and without tics in an epidemiological sample of Israeli adolescents: Are they distinct subtypes? Am J Psychiatry 154:274-276

ELECTRONIC-DATABASE INFORMATION

Advanced PubMed search: http://www.ncbi.nlm.nih.gov/PubMed/medline.html

Center for Medical Genetics, Marshfield, WI, for genetic marker and genetic map information: http://www.marshmed.org/genetics

Centre d'Etuds du Polymorphisme Humain (CEPH) database, for marker allele frequencies in Caucasians: http://www.cephb.fr/

Cooperative Human Linkage Center, for marker information: http://www.chlc.rg/ChlcMarkers.html; http://www.cephb.fr/quickmap.html

Généthon, database, for genetic marker information: http://www.genethon.fr

Genemap database, for the physical location of over 30,000 genes: http://www.ncbi.nih.gov/genemap/

Government Communication and Information System [South Africa], for Afrikaner history: http://www.sacs.org.za/level2/history.htm

On Line Mendelian Inheritance in Man (OMIM), for information regarding Tourette syndrome (MIM 137580) and other genetic diseases: http://www3.ncbi.nlm.nih.gov/Omim/searchomim.html

The NCBI Database, for the human gene map location of candidate genes: http://www.ncbi.nlm.gov/cgi-bin/SCIENCE96

The Tourette Syndrome Association: http://tsa.mgh.harvard.edu

APPENDICES

Figure 1

Electrophoretic profiles for additional markers (see Chapter 5) that demonstrated consistent allele frequency distribution between affected and control subjects in the first group. Segments of autoradiographs from polyacrylamide gels are displayed for six STRPs at the indicated loci. The DNA templates used to generate the amplified DNA fragments were in the same order for each marker: standard DNA from CEPH family parents 133101 (1) and 133102 (2), DNA pool from set 1 of unaffected control subjects (C1), DNA pool from set 1 of affected subjects (A1), independent amplification of pools C1 and A1, DNA pool from set 1 of affected subjects (A1), independent amplification of pools C1 and A1, DNA pool from set 2 of unaffected control subjects (C2), and DNA pool from set 2 of affected subjects (A2). Arrows mark alleles enriched in the affected subjects.

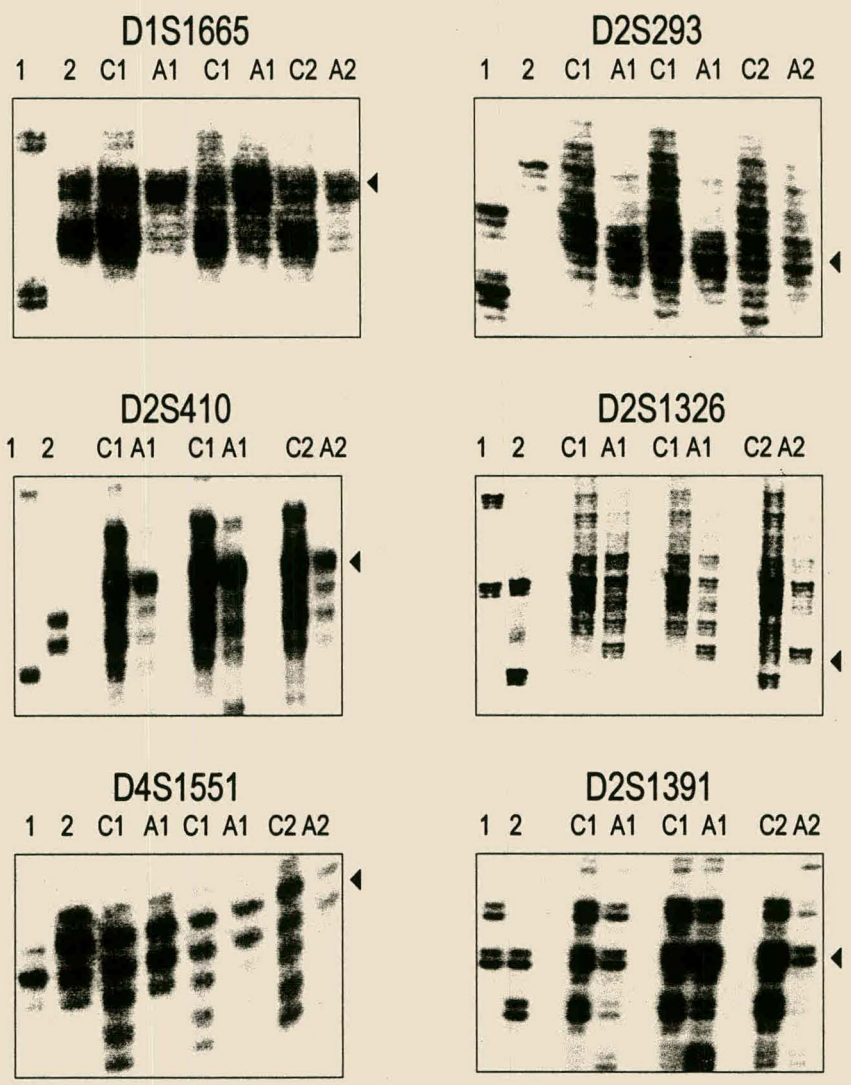


Figure 1 (Continued)

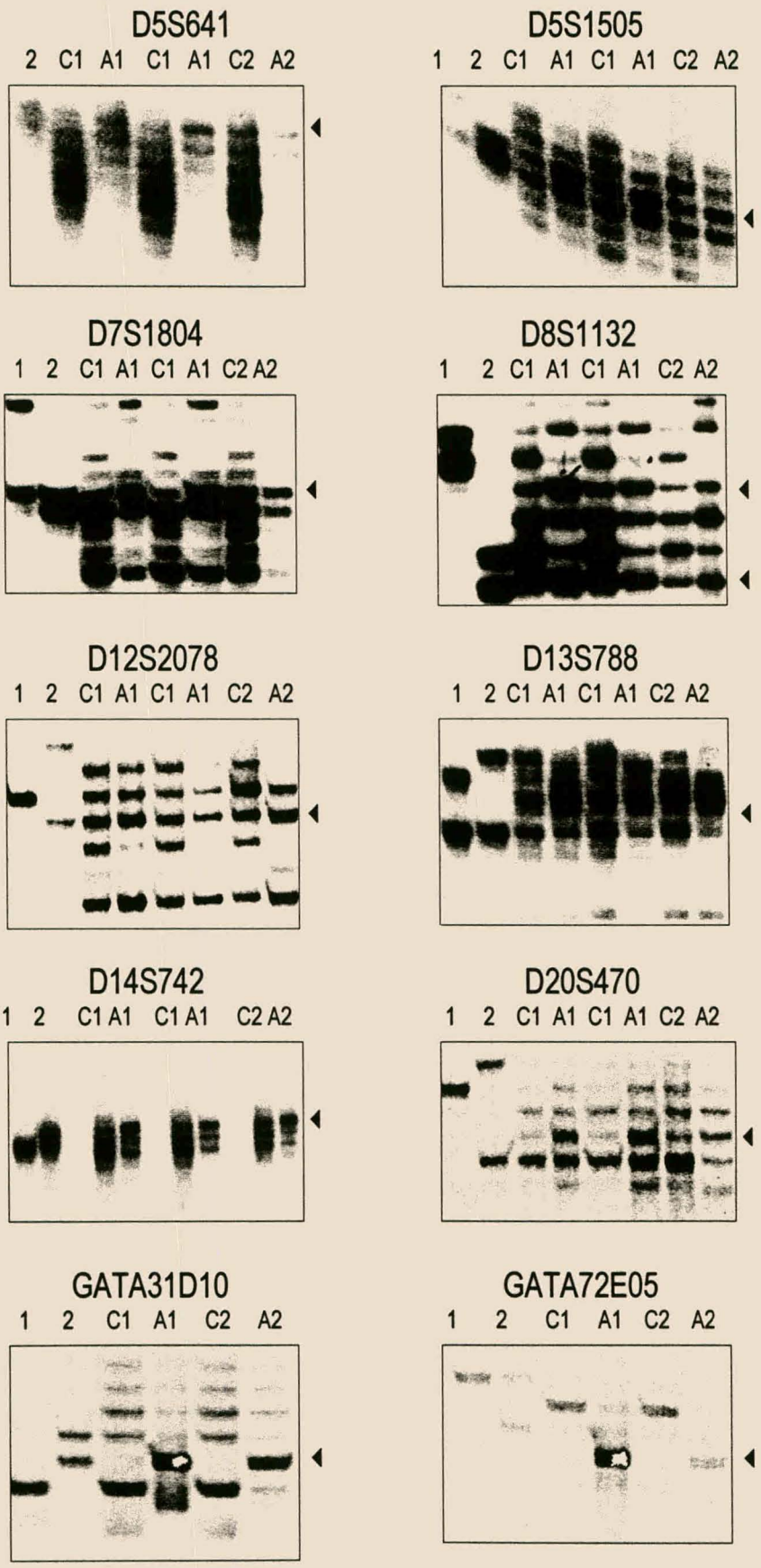


Table 1

Statistical analysis of the marker allele frequencies at the loci investigated in the first set of samples (~40 cases,~40 controls). "*2p(chi-sq)*" refers to the (two-sided) *P*-value of the chi-square test when all alleles are distinguished in an 2-by-n contingency table (2 affection classes, *n* alleles); alleles with expected values smaller than 1 were suitably pooled. "*Allele*" is the marker allele most strongly (in terms of *P*-value) associated with disease (or control). "*P(assoc)*" refers to the (one-sided) *P*-value for Fisher's test in a 2-by-2 table, given allele versus all other alleles, cases versus controls. One such table was calculated for each allele. The value reported is the smallest *P*-value, adjusted for multiple testing using the Bonferroni correction. "*t*" and "*df*" refers to the t-test for a difference in mean allele sizes between cases and controls. Positive t-values mean larger allele sizes in the group of cases (affected individuals). The t-test is expected to be powerful when marker alleles generally are shifted in the length in cases versus controls (Chapter 5).

Chr.	Marker	2P(chi-sq)	Allele	P(assoc)	t	df
1	D1S485	0.017664	177	0.005112*	-1.5556	178
1	D1S1665	0.009487*	231	0.001256*	2.0439	180
2	D2S410	0.237736	152	0.222738	-2.5914	180
2	D2S1326	0.106139	236	0.075417	-1.3403	180
2	D2S1360	0.066787	152	0.29132	1.3521	182
2	D2S1391	0.04229	109	0.155708	-0.8236	178
2	D2S1790	0.012772	324	0.034551	-2.1946	170
3	D3S1261	0.217071	213	0.338027	-1.5857	176
3	D3S1286	0.05662	141	0.47967	0.7068	180
4	D4S1551	0.009769*	178	0.025258	1.5437	182
4	D4S1613	0.234327	261	0.358959	-0.4513	172
4	D4S1647	0.037058	152	0.012706	1.5706	180
5	D5S666	0.025713	249	0.022373	-1.8311	182
6	D6S274	0.124564	182	0.086487	-2.3028	180
6	D6S277	0.433172	116	0.262028	1.3867	172
6	D6S344	0.855934	141	0.692204	-1.8502	172
6	D6S429	0.216727	230	0.505481	0.9265	176
6	D6S439	0.535595	284	0.63383	-0.499	174
6	D6S470	0.000621**	134	0.022729	3.3466**	180
6	D6S477	0.004796*	229	0.06495	1.6926	178
8	D8S1119	0.025608	182	0.057083	-0.4463	180
8	D8S1128	0.453483	244	0.312992	-0.4471	178
8	D8S1130	0.019856	148	0.076935	3.6133**	178

Table 1

Continued

Chr.	Marker	2P(chi-sq)	Allele	P(assoc)	t	df
8	D8S1145	0.111139	261	0.257379	-2.2015	180
8	D8S1477	0.136501	159	0.088622	-0.3253	182
9	D9S910	0.496543	105	0.504005	-0.2093	178
9	D9S922	0.262896	255	0.218491	-1.95	150
9	D9S925	0.369299	171	0.225581	1.2767	170
10	D10S186	0.001720*	155	0.000181**	0.3809	176
10	D10S198	0.8642	187	0.583652	0.5003	172
10	D10S1432	0.254791	181	0.149307	0.097	180
11	D11S933	0.049813	255	0.005541*	-0.3139	180
12	D12S327	0.001229*	196	0.101758	-2.0964	182
12	D12S356	0.360483	217	0.355306	0.749	182
12	D12S358	0.082245	256	0.024513	-0.0065	180
13	D13S788	0.005115*	258	0.163327	-0.6732	170
13	D13S1493	0.227108	135	0.041064	-0.7115	174
16	D16S771	0.532954	253	0.694941	-0.2014	168
16	D16S3253	0.289044	187	0.221844	-1.1191	170
20	D20S478	0.34899	263	0.31646	1.6931	180
20	D20S604	0.44301	127	0.331024	0.6893	174
20	GATA46C01	0.027842	189	0.028089	0.4601	172
20	GGAA7E02	0.088129	266	0.111662	1.3189	176
21	D21S1252	0.000005**	243	0.006422*	0.8027	152
21	D21S1435	0.04453	171	0.046633	-2.2074	180
X(m)	DXS1218	0.312308	263	0.172211	1.9379	62
X(m)	DXS1221	0.575005	155	0.304773	-0.6525	63
X(m)	DXS6799	0.866395	257	0.692356	-0.3485	61
X(m)	DXS6800	0.076792	197	0.100052	1.9878	60
X(m)	DXS6810	0.159977	223	0.09872	-1.6041	61
X(m)	DXS7132	0.304675	287	0.184667	-1.4221	61

m – investigated in the male samples only

* results with P = 0.01 or smaller

** results with P = 0.001 or smaller

Table 2

Statistical analysis of the marker allele frequencies at the loci investigated with the second set of samples (~60 cases, ~60 controls).

* results with P = 0.01 or smaller

** results with P = 0.001 or smaller

Chr.	Marker	2P(chi-sq)	Allele	P(assoc)	t	df
1	D1S1665	0.120449	239	0.057488	-0.7509	200
1	D1S485	0.245221	173	0.242315	-1.2204	194
1	GATA124B02	0.722764	198	0.463694	0.6729	396
1	D1S495	0.415392	146	0.454405	0.6122	386
2	D2S428	0.244404	158	0.138064	-1.3824	420
2	GATA6E12	0.512984	264	0.631871	0.7613	412
2	GATA5B07	0.930603	192	0.758404	-0.5789	410
2	D2S1396	0.158605	130	0.286638	-0.6296	376
2	D2S435	0.222442	215	0.079522	-0.9384	424
2	D2S440	0.001505*	207	0.001086**	-0.6062	426
2	D2S1790	0.006018*	324	0.012406	-0.1754	198
2	GATA62B10	0.392579	161	0.740326	-0.7276	410
2	D2S1391	0.141234	133	0.103959	1.0698	179
3	D3S1286	0.777862	145	0.211394	1.326	200
3	GATA123C09	0.156317	169	0.138956	-1.5572	386
3	GATA8C11	0.909997	124	0.691808	-0.5488	406
4	D4S1551	0.416725	176	0.639875	-0.9522	189
5	D5S666	0.189512	233	0.624466	-0.1008	186
6	D6S477	0.02357	237	0.040963	0.8224	422
6	ATA109H09	0.178322	203	0.204661	-0.2115	416
6	D6S470	0.434915	130	0.206133	-1.5249	242
8	D8S1130	0.42089	136	0.422439	-0.8546	210
8	D8S1138	0.238189	260	0.105402	1.3155	420
8	D8S273	0.130718	137	0.164897	-1.3345	420
8	D8S257	0.013695	118	0.003529*	-0.3329	393
8	Mfd45	0.164226	84	0.31805	1.5887	422
8	D8S0271	0.752514	271	0.750539	-0.4636	414
8	D8S1119	0.088352	182	0.052879	0.2282	409
10	D10S1222	0.27078	118	0.200439	-1.0616	212
13	D13S788	0.562724	270	0.429572	0.0673	196
13	ATA17C06	0.92774	213	0.794166	-0.7704	374
13	GATA148B01	0.453457	188	0.533182	0.3577	428

Table 2

Continued

Chr.	Marker	2P(chi-sq)	Allele	P(assoc)	t	df
14	D14S742	0.060972	399	0.018396	-2.1493	410
14	D14S275	0.79853	201	0.663896	-0.0732	392
14	D14S283	0.691032	133	0.836835	-0.3339	382
14	Mfd130	0.842671	180	0.524172	-0.6704	396
14	D14S1003	0.001486*	163	0.002918*	3.9910*	402
20	D20S913	0.613341	252	0.46292	-0.8893	396
20	D20S1085	0.000462**	181	0.000401**	-0.3111	196
20	ACT1A04	0.053424	194	0.060856	2.8233	394
20	ATA7B01	0.543936	110	0.510919	-0.101	412
20	D20S469	0.097633	215	0.06269	-0.4355	420
21	D21S260	0.150657	269	0.108757	-0.3152	374
21	ATA22G04	0.747804	320	0.464855	0.3443	192
21	D21S1252	0.072452	249	0.121975	-0.9262	189
21	D21S1435	0.461993	179	0.174308	-1.1872	196
21	GATA116E08	0.360294	230	0.203998	0.7338	402
21	D21S1255	0.081902	118	0.223159	-1.126	194
X(m)	GATA186D06	0.861824	204	0.867496	-0.474	135
X(m)	ATA70F04	0.371436	140	0.566302	1.7733	92
X(m)	DXS1068	0.379466	251	0.197329	-0.6172	117
X(f)	GATA186D06	0.040756	212	0.019096	1.5016	132
X(f)	ATA70F07	0.114496	140	0.11832	2.025	106
X(f)	DXS1068	0.663063	249	0.593089	-0.0499	136

m – investigated in the male samples only
f – investigated in the female samples only

Thresholds for "t" values (absolute values)

df	P = 0.01	P = 0.001	P = 0.000
120	2.617	3.373	4.025
∞	2.576	3.291	3.891

Figure 2

Chromosomal regions investigated for the extent of the background LD in the Afrikaner population (Chapter 7).

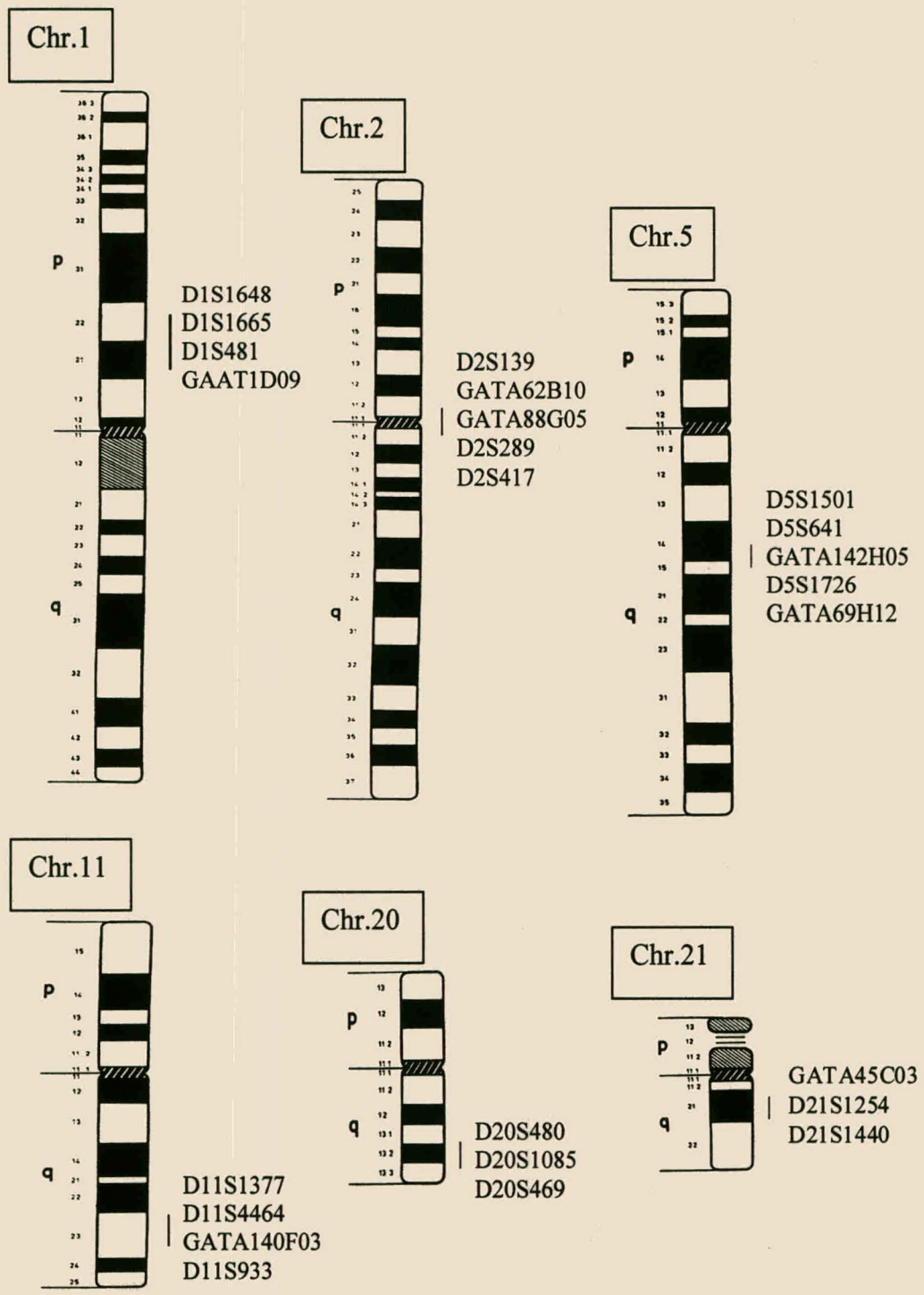


Table 3

IBS haplotype sharing among unrelated Tourette syndrome probands on chromosome 8q22 (see Chapter 8, 9)

		Tourette syndrome probands																								
Marker	cM	2-3	7-3	16-3	18-3	23-3	24-3	33-3	38-3	41-3	44-3	46-3	49-3	50-3	57-3	61-3	69-3	79-3	81-3	05-3	08-3	013-3	016-4	019-3	020-3	023-3	025-3
D8S1119	101.01										9								9			9					
D8S1707	101.69	107									107								107			107					
D8S271	102.62	267							267		267								267			267					
GATA8B01	103.69	234	234		234	234		234	234		234	234	234	234	234		234		234		234	234	234	234	234	234	234
T7-27	103.69	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
D8S270	103.69	191	191	191	191	191	191	191	191	191	191	191	191	191	191	191	191	191	191	191	191	191	191	191	191	191	191
MTG8	103.69	13	13	13	13	13	12	13	13	13	13	12	12	13	13	12	12	13	13	12	13	13	12	13	12	12	13
GATA28F12	104.33		286	286	286	286		286	286	286	286		286	286	286	286		286	286	286	286	286		286	286	286	
D8S1822	107.97		14		14						14			14	14	14		14	14	14	14			14		14	
D8S1129	110.2				144						144			144		144			144		144			144		144	
D8S257	111.68				110						110								110								
D8S559	112.42						218				218								218								
D8S1808	113.16						174				174								174								

Table 4

IBS haplotype sharing among unrelated Tourette syndrome probands on chromosome 11q23

		Tourette syndrome probands		
Marker	cM	8-3	13-3	49-3
D11S1377	120.87	136	136	
D11S1353	122.47	202	202	202
D11S4464	123	4	4	4
D11S933	124.07	255	255	255
D11S975	126.21	236	236	236

Nontransmitted haplotype associated with the allele 236 (D11S975)

4-2
136/136
200/200
2/5
247/259
224/236

Table 5

Actual numbers of transmitted (T) versus non-transmitted (NT) alleles for investigated markers (Chapter 8) subjected to TDT and HRR statistics. Only informative nuclear families for particular marker-allele are included.

D8S1119			D8S1707			D8S271			GATA8B01			MTG8		
Alelle	T	NT	Alelle	T	NT	Alelle	T	NT	Alelle	T	NT	Alelle	T	NT
1	0	1	107	40	39	272	1	1	258	0	1	19	2	1
2	18	13	105	27	29	269	6	7	254	3	4	17	7	10
3	14	23	103	7	11	267	40	36	250	7	5	16	17	17
4	6	6	99	47	42	265	6	10	246	20	12	15	11	13
5	12	17				263	21	22	242	7	16	14	10	4
6	53	48	GATA28F12			261	8	2	238	29	18	13	28	32
7	5	1	Alelle	T	NT	259	29	30	236	0	1	12	28	27
9	26	25	298	2	1	257	35	37	234	38	46	11	19	27
			294	2	2	253	0	1	230	5	8	10	20	30
			290	22	41				226	4	2	9	10	9
			286	44	28							8	3	6
			282	9	9							7	2	0
			274	4	2							5	1	0
												2	1	0

T7-27			D8S270			D8S1822			D8S1129			D8S257		
Alelle	T	NT	Alelle	T	NT	Alelle	T	NT	Alelle	T	NT	Alelle	T	NT
11	26	26	197	3	4	2	0	1	148	3	4	122	0	1
10	4	4	195	7	5	3	3	1	144	19	17	118	1	1
6	5	1	193	28	24	4	1	1	140	6	8	116	14	19
5	1	2	191	49	43	5	4	7	136	4	3	114	44	54
4	35	35	189	12	18	6	10	13	130	7	7	112	37	31
3	2	5	187	2	4	7	4	0				110	32	23
2	1	1	185	1	2	8	12	13				108	1	1
			183	5	10	9	2	7				106	1	0
			181	31	27	10	42	34						
			179	1	1	12	4	2						
			175	0	1	13	4	5						
						14	41	43						
						15	14	14						

D8S559			D8S1808			DRD2			D11S1377			D11S1353		
Alelle	T	NT	Alelle	T	NT	Alelle	T	NT	Alelle	T	NT	Alelle	T	NT
238	1	0	176	1	0	1	1	3	146	8	6	214	0	1
232	1	0	174	9	6	2	31	31	144	3	5	208	6	3
230	4	6	172	6	11	3	49	43	142	19	8	206	34	38
228	7	7	170	38	49	4	19	32	141	0	1	204	28	29
226	30	29	168	51	39	5	24	14	140	10	19	203	5	0
224	16	20				6	1	2	138	17	27	202	15	24
222	5	9							136	26	37	200	33	28
220	42	40							134	37	23	196	3	1
218	24	19							132	8	2			
216	0	1							130	3	1			
212	1	0							128	5	7			

Table 5

Continued

D11S4464			D11S933			D11S975			D2S139			D2S440		
Alelle	T	NT	Alelle	T	NT	Alelle	T	NT	Alelle	T	NT	Alelle	T	NT
1	10	4	267	1	3	240	1	8	201	1	2	211	4	2
2	13	20	265	3	1	238	0	1	199	1	2	207	8	8
3	38	33	263	8	3	236	3	1	197	7	11	203	18	21
4	38	43	261	3	4	234	11	14	195	30	27	199	44	39
5	15	12	259	11	7	232	43	51	193	16	29	195	43	42
6	8	10	257	31	37	230	6	1	191	13	9	191	12	16
			255	37	41	228	6	9	189	0	1	187	1	0
			253	6	7	226	3	1	187	0	3	183	0	2
			251	34	26	224	22	11	183	0	2			
			249	3	3	222	36	33	181	29	32			
			247	4	9	216	6	7	177	4	1			
									175	43	25			
D2S2161			D2S417			D20S480			D201085			D20S469		
Alelle	T	NT	Alelle	T	NT	Alelle	T	NT	Alelle	T	NT	Alelle	T	NT
204	3	4	213	0	2	308	2	3	199	2	0	239	2	0
202	4	7	211	6	5	304	4	10	195	1	2	235	1	0
200	8	7	209	13	13	300	33	27	193	0	2	231	8	4
198	9	16	207	32	23	296	36	31	191	8	8	229	0	1
196	13	7	205	36	51	292	32	37	189	4	6	227	3	2
194	33	42	203	47	34	288	13	12	188	2	1	223	8	3
192	29	21	201	1	4	284	2	2	187	20	22	219	15	17
190	11	4	199	0	1				185	20	19	215	6	10
188	3	5	197	7	9				184	6	2	211	32	41
184	3	2							183	16	33	207	17	14
180	1	0							181	20	25	203	1	1
176	0	3							180	21	14			
172	1	0							179	5	4			
									177	15	9			
									176	14	9			
									173	3	3			
									171	2	0			
D20S120			GATA45C03			D21S1920			D21S1895			D21S1252		
Alelle	T	NT	Alelle	T	NT	Alelle	T	NT	Alelle	T	NT	Alelle	T	NT
249	0	1	1	1	2	234	2	2	280	1	1	253	1	4
247	0	1	2	0	2	232	23	22	278	10	8	251	2	3
245	1	2	3	13	13	230	25	27	276	8	15	249	11	10
243	8	7	4	19	18	228	9	8	274	18	17	247	21	22
241	8	8	5	16	9	226	29	27	272	12	16	245	9	17
239	27	19	6	2	2	224	33	39	270	27	27	243	5	0
237	43	28	7	1	2	222	4	1	268	17	11	241	3	2
235	12	15	8	32	33	220	11	10	266	11	22	239	25	23
233	19	28	9	10	2				264	46	31	237	42	32
231	20	23	10	2	2				262	3	4	233	2	2
229	0	2	12	0	1				260	0	1	231	20	27
223	0	1	16	1	1							227	1	0
221	1	1												
219	6	11												
217	1	1												
213	8	6												