A LINKAGE STUDY OF CHROMOSOME 5q11.2-q13.3 IN SCHIZOPHRENIA

by

ROBIN PAUL SHERRINGTON

A thesis submitted for the degree of PhD of the University of London

The work reported in this thesis was carried out whilst registered at the University College and Middlesex School of Medicine, University of London

May 1992

ABSTRACT

Early studies investigating the inheritance of schizophrenia estimated that between 66% to 93% of the variance in the aetiology is genetic. However, segregation analysis studying the patterns of affected individuals within families has not been able to establish a common mode of transmission. Nevertheless linkage analysis of marker data obtained from families is considered to be the most efficient and robust approach to studying the genetic aetiology of schizophrenia. Linkage studies in schizophrenia can be directed towards regions of the genome where there is a priori grounds for the presence of a susceptibility locus. This study investigated chromosome 5 after a report of an association (though weak) between a partial trisomy of chromosome 5q 11.2-13.3 and schizophrenia.

In the original investigation seven families fulfilled stringent selection criteria, five from Iceland and two from Britain. Two RFLPs mapping to 5q11.2-q13.3 were studied for co-segregation with schizophrenia (D5S76 and D5S39). Weak evidence of linkage (lod = 2.45) was found when the analysis was restricted to cases with schizophrenia alone with the "unaffecteds" diagnosed as phenotype unknown. Using the same model but classifying the unaffecteds as phenotypically normal gave a maximum lod of 3.22 at 99% penetrance. However the degree of co-segregation increased when the spectrum cases were included in the schizophrenia phenotype (lod = 4.33 at 99% penetrance). This is not unexpected since these disorders are known to be genetically related to schizophrenia. More surprising is the observation that the evidence for linkage further increased when the fringe cases were included (lod = 6.49 at 86% penetrance), but only five of the fringe cases were informative for linkage so this could represent nothing more than a chance finding. These results were interpreted as demonstrating the existence of a dominant susceptibility gene on chromosome 5 for schizophrenia. Furthermore this defect seemed to predispose to schizophrenia spectrum disorders and possibly also to a variety of other psychiatric conditions.

Numerous investigators promptly examined these and proximate markers in other family samples, but all obtained negative results. This discrepancy may represent true genetic heterogeneity, or a type I statistical error. In any event, initial linkage may be difficult to replicate due to complex mode of transmission. If the significant lod from the Icelandic/British study represented genetic heterogeneity, this could be tested by increasing the informativeness of the markers used, consequently increasing the

proportion of informative meioses for linkage. Fortunately 50% of the meioses were uninformative for linkage with the original markers and these meioses could be studied with new markers. Another approach would be to collect a new cohort of families from Iceland using the same stringent selection criteria.

The aim was to screen genomic cosmid and lambda libraries with markers mapping to 5q11.2-q13.3 to expand these loci and detect new polymorphisms to increase their PIC values. Cosmid clones were obtained for the 5HT1a receptor gene and lambda clones were obtained for D5S76, D5S39, D5S6 and HEXB. Originally potential new RFLPs were to be investigated, but at this time the PCR had allowed the development of several new types of multi-allelic markers collectively called microsatellites. Highly PIC value microsatellites were found for D5S39 and D5S76 and a low PIC value microsatellite for HEXB. The two microsatellites for D5S76 showed strong linkage disequilibrium to each other. Unlike highly polymorphic VNTRs these markers are stable during evolution and can be used in tests of association. Additionally two polymorphic microsatellites were identified in cosmids for D5S127 and D5S125 known to be linked to the previous markers. This collection of highly polymorphic microsatellites would greatly increase the number of meioses informative for linkage. A new Taq I RFLP was found for the 5HT1a receptor which showed tight linkage to the microsatellites for D5S76 and the RFLP for D5S6. The receptor is a potential candidate gene for the putative chromosome 5 susceptibility locus for schizophrenia.

Linkage analysis with these new markers excluded the region of the trisomy in the original cohort of families with only one family showing an interesting lod of 2.2 (F41) with the 5HT1a receptor Taq I RFLP. However this family had a clear recombinant individual with a diagnosis of schizophrenia so it is unlikely that this gene plays a role in the disease even for this family. A non-significant positive lod was resulted with markers proximal to D5S76, but a new cohort of Icelandic families (N = 6) excluded this region for a dominant susceptibility allele. This new cohort using the new microsatellites also excluded the region of the trisomy and the region covered by the original significant positive lod. Taking all this evidence there remains only one possible conclusion. The original significant lod was a type 1 statistical error caused to some degree by multiple testing of different clinical definitions of affection. However, this negative finding does not invalidate the pursuit of genetic marker strategies in studying the genetic basis of schizophrenia. Despite the variety of methodological problems a systematic search throughout the human genome can now be justified as a means of detecting major gene effects in schizophrenia.

ACKNOWLEDGEMENTS

I wish to thank Dr Hugh Gurling for his helpful advice and guidance throughout this project. In addition, I am grateful to Dr Baljinder Mankoo and Dr Georg Melmer for assistance with screening genomic libraries, Dr Mike Dixon for help with microsatellite analysis, and Dr David Curtis for the analysis of the linkage data with the LINKAGE package. Thanks must be extended to the members of the Medical Molecular Biology Unit and Molecular Psychiatry for their many helpful suggestions and invaluable technical assistance. In particular I would like to mention Dr Debbie Holmes, Dr Chris Cook, Ms Gursharon Kalsi, Mr John Vincent, Dr Aviva Symes, Dr Colleen Higgins, Dr Mukesh Patel, Dr Annie Rowe, Mr Stephen Hodgkinson for their encouragement during the project and Mr Peter Brett for managing the finances of the laboratory and falling asleep at the weekly lunch time meetings. I am also indebted to all the patients and their families for their co-operation in this study and the following clinicians for the family collection, Dr Jon Brynjolfsson, Dr Hannes Petursson, Dr Pordur Sigmundsson, Dr Keith Duddleston, Dr Brian Barraclough, Dr Timothy Read, Dr Mark Potter, Dr Patrice Murphy, Dr Eamonn Maloney, Dr Larry Rifkin, and Dr Tomoy Sharma. I should also thank Iceland Air for my safe arrival in Iceland and free transport of blood samples.

I would also like to thank my parents and friends for their endless encouragement. In particular Ms Elaine Muldrew for her patience and support, Mr Richard Fiddes for a roof over my head, Mr Speedy Fathers for asking me, if I had finished my PhD yet, every time I saw him and for advising me on the purchase of my car, Mr Steve Ball for fixing the said car, Dr Jonny Poynter for games of tennis, philosophical discussions and taking me swimming in the Atlantic last new year resulting in glue ear and having them syringed. I would also like to mention in no particular order, Louise, Mark, Val, Tim, Ken, Si, Dan, Della, Helen, Vincent, Span, Chris, Judith, Dylon, Anne, Shelagh, Boon, and Anabel so they do not feel left out.

Lastly, I wish to mention Dr Amador del la Concha who died in such tragic circumstances last year.

ABBREVIATIONS

amp: ampicillin

ATP: adenosine triphosphate

bp: base pair

cDNA: complementary DNA

dATP: deoxyribosyladenine 5' triphosphate

dCTP: deoxyribosylcytosine 5' triphosphate

dGTP: deoxyribosylguanine 5' triphosphate

dTTP: deoxyribosylthymine 5' triphoosphate

DNA: deoxyribonucleic acid

DSM III: diagnostic statistical manual III

DTT: dithiothreitol

DOMS: dominant inheritance with schizophrenia

DOMSS: dominant inheritance with schizoophreia and spectrum disorders

DOMSSF: dominant inheritance with schizophrenia and all other psychiatric disorders

EDTA: ethylenediamine tetra acetic acid

hrs: hours

IPTG: isopropyl-beta-D-thiogalactopyranoside

kb: kilobase pair

LB: Luria Bertani medium

mA: milliamps

ug: micrograms

mins: minutes

ul: microlitres

ml: millilitres

mRNA: messenger RNA

PBS: phosphate buffered saline

PCR: polymerase chain reaction

RDC: research diagnostic criteria

RNA: ribonucleic acid

RNase: ribonuclease

SADS-L: schizophrenia and affective disorders questionnaire (lifetime version)

SDS: sodium dodecyl sulphate

SSC: standard sodium citrate

Tris: tris(hydroxymethyl)aminomethane

V: volts

XGAL: 5-Bromo-4-chloro-3-indolyl-beta-D-galactoside

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

Table 1. Two point lods for each family, for each model and marker

Figure 4. Multipoint maps for each model for the original and new cohorts of families, separately and combined for the new marker genotypes

Figure 5. PCR amplification of parraffin-embedded tissue with p599$(CA)_n$

Table 2. Microsatellite genotypes for the parraffin-embedded tissue samples

CHAPTER 1: GENERAL INTRODUCTION

1.1 SCHIZOPHRENIA: CLINICAL FEATURES AND SYMPTOMS

Mental illness should by no means be considered a disorder of modern society. The ancient civilizations of Egypt and the Near East attributed mental illness to magical influences of malevolent deities. The Greek philosopher Hippocrates (450-355 B.C.) was the first to produce a rational classification of the mental state including epilepsy, mania, depression, paranioa and phobias. Galen (131-200 A.D.) linked mania and depression together, each being due to an excess of one of the four humours out of which the body was composed.

At the turn of this century the major advances in clinical psychiatry resulted from work by Germans physicians. They were convinced that mental illnesses resulted from disorders of the brain and were the first to attempt to classify the psychoses. They initiated modern phenomenology of mental illness and recognised that the natural history of mental illness had to be adequately delineated. For example they recognised catatonia as psychosis with a motor disorder consisting of strange attitudes, odd movements, stupor and mental deterioration. Later they introduced the terms cyclothymia for circular insanity and hebephrenia for a rapid mental deterioration occuring during puberty.

One of the German school's most outstanding psychiatrists was probably Emil Kraeplin (1855-1926). He developed a classification for the mental disorders using clinical observation and an awareness of natural history. He brought together the disorders dementia praecox, hebephrenia, catatonia and dementia paranoides as a process of psychological degeneration. He grouped these disorders as subtypes of a single disease entity ubiquitiously called dementia praecox. In 1893 he recognised a separate psychosis which had a periodic nature consisting of manic illness and depressions. He grouped recurrent mania, depressions and circular insanity as manic-depressive insanity. This was the first recorded separation of the two major psychoses namely schizophrenia and affective disorders which has been of primary importance for the development of modern classification systems used in clinical psychiatry.

Eugen Bleuler in 1911 coined the term schizophrenia because he believed that the functions of the mind were split off from each other in the disease. He wanted to make obsolete the term dementia preacox which implied intellectual impairment. He was the

first clinical psychiatrist to apply Freud's theories of psychotherapy to the study of psychotic symptoms.

Modern clinical psychiatry has concerned itself with improving diagnostic systems, psychopharmacology, and research to elucidate the true aetiological factors for the major psychosis. It is crucial to understand that in the absence of any diagnostic marker concepts of shizophrenia may eventually prove to be misleading. Nevertheless comparison of present day concepts with descriptions of madness in the early nineteenth century indicate that the major clinical features have remained constant. However owing to the development of antipsychotic drugs and modern medical practices the symptoms seem to have become less severe and bizarre in recent years with some investigators claiming that the disease is abating (Murray et al 1990).

Schizophrenia has a life time risk of approximately 0.85% (Eaton 1985) and an onset in early adult life. The main symptoms of schizophrenia consist of a disorder of thought, of perception, emotion, and abnormal ideas. These occur in symptom complexes, but for the sake of lucidity they will be considered individually even though this is somewhat artificial. For further details see Hamilton (1984) and Sadock et al (1989).

1.1.1 DISORDERS OF THOUGHT CONTENT AND THOUGHT PROCESS

Abnormalities in mental disorders can be divided into those of content which reflect ideas, beliefs, and interpretations of stimuli, and those of process and form observed in the spoken language and in how ideas and language are formulated.

Delusions are examples of a disorder of form and content. Often bizarre they may be persecutory, grandiose, religious or somatic. The schizophrenic can sometimes believe that some outside force is controlling their thoughts and behaviours or that they are controlling outside events in some extraordinary fashion. Paranoid delusions commonly observed in patients are that they are being spied upon, talked about or at risk from being harmed. Patients may also experience thought broadcasting, thought insertion, thought withhdrawal, and thought control. The common central theme of schizophrenic delusions is the direct, immediate and total certainty with which a patient holds these beliefs.

Schizophrenic speech has a tendency to be filled with bizzare or symbolic images the meaning of which is difficult to comprehend. There is sometimes a preoccupation with abstract, psychological and philosophical ideas.

Patients may lack a clear sense of where their own body, mind and influence ends and where these characteristics in other animate and inanimate objects begin. For example the patient may believe that other people or the television are talking about them or that they have fused with another object or disintegrated completely.

Disorders of the form of thought are observable in the spoken and written language of the patient. Disorders of the process of thought refer to how ideas and language are formulated in the brain, consequently the two are interrelated. Forms of thought disorders found in schizophrenia include loosening of associations, incoherence, tangentiality, flight of ideas, neologisms, echolalia, verbigeration, mutism, thought blocking, impaired attention, overinclusion, illogical ideas, vagueness and poverty of content. The patients however do not seem to be aware that their comunication is abnormal.

Thought blocking is the experience that some patients have of the sensation of having their thoughts physically removed from their heads, stopping suddenly in the middle of a sentence. Poverty of speech content refers to a lack of information an examiner may receive even though the patient may have been quite loquacious. Mutism or inhibition of speech before the use of modern neuroleptics often lasted for years in the chronic schizophrenic, even with treatment the patients can sometimes be monosyllabic.

1.1.2 PERCEPTUAL DISTURBANCES

Hallucinations and delusions are very common in schizophrenia. On the other hand illusions are common but also experienced by many normal people though to a lesser extent. An illusion is a misinterpretation of a sensory stimulus, while a hallucination is a perception in the absence of an external stimulus. Hallucinations can occur in any of the five sensory modalities, however upto 75% are auditory. The auditory hallucinations may consist of voices of God or the devil; sometimes they are of neighbours, deceased relatives or unrecognised individuals. The schizophrenic may experence two voices discussing himself in the third person; voices may make obscene comments about the patient, and the patient may hold audible conversations with the voices. Visual hallucinations are rare in schizophrenia and can be confused with strongly held delusions. Hallucinations of smell are not uncommon; the patient complains of gas, odours of decomposition, chemical smells and so on and these are normally intertwined with their paranoid delusions. A third frequently found hallucination is of the bodily or somatic kind. Patients experience induced sensations

of heat, cold, pain, or electric shock. In the acute or chronic schizophrenics hallucinations can be bizarre with feelings that their flesh is being torn away, their bowels are torn out, animals or machines are inserted into their bodies. Illusions are difficult to differentiate clinically from hallucinations but perceptual hypersensitivity to light, sound, touch, smell, taste can occur. For example small changes in the lighting on another person's face might be perceived as a dramatic change by a patient and interpreted in a delusional way.

1.1.3 MOOD, FEELINGS AND AFFECT

In schizophrenia these can be grossly reduced, extremely exaggerated, or patently bizarre. Many patients show reduced emotional responses and seem to be indifferent with emotional shallowness. An extreme form of this results in a profound emotional barrenness in which the patient is incapable of experiencing any pleasure. Often the emotional responses of a schizophrenic are inappropriate to the situation. For example they may smile whilst talking about a morbid subject or show unaccountable anger. Schizophrenics also suffer bizzarre emotions with states of exaltation, feelings of omnipotence, and religious ecstasy. The patients may also show marked sensitivity to emotional trauma, being easily hurt by very mildly aggressive or rejecting behaviour by others.

1.1.4 ABNORMAL BEHAVIOURS

First impressions of a schizophrenic can sometimes consist of extreme bizarreness, agitation or withdrawal, exhibited as a set facial expression, lack of sustained eye contact and staring at inanimate objects. Their personal appearance tends to deteriorate and they may exhibit idiosyncratic manners or offensive behaviour. Lack of motivation and will is demonstrated by an inability to continue an occupation or showing complete disinterest in future plans. In chronic patients stereotypic behaviour may sometimes present as repetitive patterns of moving or walking, strange gestures, or endless repetitions of the same phrase or question. Social withdrawal is very common symptom in schizophrenia; contacts feel unable to establish rapport with the patient which often prevents others from feeling empathy or sympathy towards them. Until the mid-1950s, mental hospitals contained many severely affected patients, who had catatonic symptoms, stereotypies and grossly disorganised behaviour. Catatonia occurs today but is thought to be the severest form of schizophrenia with paranoid being the least severe. The antipsychotic chlorpromazine has a dramatic effect in reducing many disabling symptoms of schizophrenia especially catatonia and bizarre behaviours.

1.1.5 COURSE OF THE ILLNESS

Onset is usually in adolesence. Onset before the age of 10 is rarely reported, whilst onset after 45 is not uncommon paticularly in women. Early features of schizophrenia may involve patients feeling overwhelmed by external and internal pressures resulting in anxiety, irritability, distractibility, and impaired performance at work. Onset may also be characterised by boredom, apathy, hopelessness, loneliness, and unexplained aggressive behaviour can arise in many individuals. The stage of actual psychosis may occur after many obvious changes in general behaviour. Even then delusions can be hidden by denial and paranoid ideation. The classical course is one of remissions with a lack of return to the patients previous norm. The signs of a relapse are hallucinations, bizarre behaviour, sleeping problems, and problems in thinking clearly and lack of self care. On average the deterioration in the quality of life continues for many years by which stage a plateau is reached. The positive symptoms tend to become less severe with time and the patient is left with the more socially and functionally debilitating negative symptoms. Many schizophrenics remain in a state of stable chronicity with clearly visible signs and symptoms of severe mental illness. 20-30% recover to lead relatively normal lives, a similar number continue to experience moderate symptoms whilst 40-60% remain significantly impaired for life (Sadock et al 1989).

1.2 CLASSIFICATION OF SCHIZOPHRENIA

The past thirty years has witnessed a renaissance of clinical research on the diagnosis and classification of mental disorders. The classification of mental disorders into discrete categories has been a prerequisite for the scientific study of mental disorders. In addition, the introduction of antipsychotic drugs was also a spur to improve differential diagnosis between schizophrenia and mood disorders because evaluation of their relative treatment responses was necessary. The emergence of laboratory and familial genetic strategies for exploring potential aetiological factors has benefitted from the use of consistent diagnosis criteria.

Improved diagnosis has resulted from determining whether symptoms cluster into characteristic patterns or syndromes. As a consequence operationalised criteria have been developed in order to increase the reliability of diagnosis as well as to improve diagnostic validity.

Modern diagnostic schemes were developed out of the earlier nosological theories of Kraepelin and Bleuler. The major diagnostic systems for schizophrenia have been Schneider's; Langfeldt's (Schneider et al 1959); the various revisions of the

International Classification of Disease (ICD-10, WHO 1989); the St Louis Criteria (also called Feighner's criteria, Feighner et al 1972); the Reseach Diagnostic Criteria (Spitzer et al 1978a); the the Present State Examination (PSE/CATEGO, Wing et al 1974); and those of the American Psychiatric association eg DSM II and DSM III-R (American Psychiatric Association 1987). All of these systems include the symptoms of psychosis in their criteria. The Langfeldt, St Louis and DSM III-R criteria require a reduced level of functioning for the diagnosis. The RDC, St Louis and DSM III-R systems stipulate minimum duration of symptoms, and the St Louis system requires onset of symptoms before the age of 45. The criteria for these diagnostic systems are listed in Table 1.

All the diagnostic systems for schizophrenia are based on the presence of particular symptoms; course of illness; various exclusion criteria; and information drawn from the patient, family, past psychiaric status, and present mental status. The ideal situation would be to diagnose all patients into non-overlapping groups. Not only have diagnostic systems evolved to characterise schizophrenia from the other major psychosis but they have also subtyped schizophrenia based on differences in clinical presentations. The following subtypes are those from the DSM III-R (Table 2).

Paranoid subtype

Characterized by delusions of persecution or grandeur, it has a relatively later age of onset and less deterioration of thought and social behaviour and emotional response.

Catatonic subtype

In catatonia the patient may be in a state of complete stupor with waxy flexibility, stereotypies and a pronounced decrease in spontaneous movements and activity may also occur. The reverse is excited catatonia in which patients are in a state of extreme psychomotor agitation and talk and shout almost constantly.

Hebephrenic subtype

The major characteristics of this subtype are primitive, disinhibited and unorganised behaviour. Pronounced thought disorder and inappropriate emotional responses expressed as explosive laughter without apparent reason, and incongruouse grinning and grimacing.

Undifferentiated and Residual subtypes

If a patient exhibits psychotic symptoms that do not fulfill any one of the previous subtypes they are classified as undifferentiated. The residual subtype describes those

patients that have had previous symtoms that have met the criteria for schizophrenia but at the present evaluation show no prominent psychotic symptoms.

Positive and Negative symptoms

Positive symptoms or florid, productive or type 1 symptoms refer to the delusions, hallucinations, and bizarre or agitated behaviours which are associated with acute onset. The negative or defect, deficit or type II symptoms are used to describe the characteristics of affective blunting, poverty of speech and thought content, apathy, anhedonia and poor social functioning. These symptoms are sometimes associated with an insidious onset, and a chronic course but can also occur after positive symptoms have faded.

Schizoaffective disorder

This diagnosis represents a syndrome when manic or depressive symptoms occur concurrently with the major symptoms of schizophrenia. The psychotic episode has to have persisted for at least two weeks without the presence of affective symptoms. It may represent an overlap between the two major psychoses or a group of individuals exhibiting specific features of a schizoaffective psychosis.

Schizophreniform disorder and brief reactive psychosis

Schizophreniform disorder is diagnosed when all the criteria for schizophrenia have been met except that the symptoms have been present for less than six months. Brief reactive psychosis is diagnosed when schizophrenia like symptoms have been present for less than one month and when there was a clear precipitating stressor.

Schizotypal personality disorders

A variety of personality disorders may present with some features of schizophrenia, but without meeting the entire list of diagnostic criteria for schizophrenia. These individuals seem aloof, with often eccentric and bizarre behaviours.

KURT SCHNEIDER

1. First-rank symptoms
   a. Audible thoughts
   b. Voices arguing, discussing, or both
   c. Voices commenting
   d. Somatic passivity experiences
   e. Thought withdrawal and other experiences of influenced thought
   f. Thought broadcasting
   g. Delusional perceptions
   h. All other experiences involving volition, made affects, and made impulses
2. Second rank symptoms
   a. Other disorders of perception
   b. Sudden delusional ideas
   c. Perplexity
   d. Depressive and euphoric mood changes
   e. Feelings of emotional impoverishment
   f. ". . . and several others as well"

GABRIEL LANGFELDT

1. Symptom criteria
   Significant clues to a diagnosis of schizophrenia are (if no sign of organic mental disorder, infection, or intoxication can be demonstrated):
   a. Changes in personality, which manifest themselves as a special type of emotional blunting followed by lack of initiative, and altered, frequently peculiar behavior. (In hebephrenia, especially, these changes are quite characteristic and are a principal clue to the diagnosis.)
   b. In catatonic types, the history as well as the typical signs in periods of restlessness and stupor (with negativism, oily facies, catalepsy, special vegetative symptoms, etc.)
   c. In paranoid psychoses, essential symptoms of split personality (or depersonalization symptoms) and a loss of reality feeling (derealization symptoms) or primary delusions
   d. Chronic hallucinations
2. Course criterion
   A final decision about diagnosis cannot be made before a follow-up period of at least 5 years has shown a chronic course of disease.

Table 1. Essential features of various diagnostic criteria for schizophrenia.

RESEARCH DIAGNOSTIC CRITERIA

Criteria 1 through 3 required for diagnosis.

1. At least two of the following for definite illness, and one for probable (not counting those occurring during period of drug or alcohol abuse or withdrawal):
   a. Thought broadcasting, insertion, or withdrawal
   b. Delusions of being controlled or influenced, other bizarre delusions, or multiple delusions
   c. Delusions other than persecution or jealousy lasting at least 1 month
   d. Delusions of any type if accompanied by hallucinations of any type for at least 1 week
   e. Auditory hallucinations in which either a voice keeps up a running commentary on subject's behaviors or thoughts as they occur or two or more voices converse with each other
   f. Nonaffective verbal hallucinations spoken to subject
   g. Hallucinations of any type throughout day for several days or intermittently for at least 1 month
   h. Definite instances of marked formal thought disorders accompanied by blunted or inappropriate affect, delusions or hallucinations of any type, or grossly disorganized behavior
2. One of the following:
   a. Current period of illness lasted at least 2 weeks from onset of noticeable change in subject's usual condition
   b. Subject has had previous period of illness lasting at least 2 weeks during which he or she met criteria, and residual signs of illness have remained (e.g., extreme social withdrawal, blunted or inappropriate affect, formal thought disorder, or unusual thoughts or perceptual experiences)
3. At no time during active period of illness being considered did subject meet criteria for probable or definite manic or depressive syndrome to the degree that it was a prominent part of illness

ST. LOUIS CRITERIA

1. Both necessary:
   a. Chronic illness with at least 6 months of symptoms before index evaluation, without return to premorbid level of psychosocial adjustment
   b. Absence of period of depressive or manic symptoms sufficient to qualify for mood (affective) disorder or probable mood (affective) disorder
2. At least one of the following:
   a. Delusions or hallucinations without significant perplexity or disorientation
   b. Verbal production that makes communication difficult owing to lack of logical or understandable organization (in presence of muteness, diagnostic decision must be deferred)
3. At least three for definite, two for probable, illness:
   a. Never married
   b. Poor premorbid social adjustment or work history
   c. Family history of schizophrenia
   d. Absence of alcoholism or drug abuse within 1 year of onset
   e. Onset before age 40

Table 1 (continued). Essential features of various diagnostic criteria for schizophrenia.

PRESENT STATE EXAMINATION

The following 12 items from the Present State Examination correspond to a 12-point diagnostic system for schizophrenia, with varying levels of certainty of diagnosis based on the cut off score determined by the examiner. Nine of the symptoms are scored 1 point each when present (+), and three are scored 1 point each when absent (–).

1. Restricted affect (+)
2. Poor insight (+)
3. Thoughts aloud (+)
4. Awaking early (–)
5. Poor rapport (+)
6. Depressed facies (–)
7. Elation (–)
8. Widespread delusions (+)
9. Incoherent speech (+)
10. Unreliable information (+)
11. Bizarre delusions (+)
12. Nihilistic delusions (+)

Table 1 (continued). Essential features of various diagnostic criteria for schizophrenia.

Paranoid Type
A type of schizophrenia in which there are:
A. Preoccupation with one or more systematized delusions or with frequent auditory hallucinations related to a single theme
B. None of the following: incoherence, marked loosening of associations, flat or grossly inappropriate affect, catatonic behavior, grossly disorganized behavior

Specify stable type if criteria A and B have been met during all past and present active phases of the illness.

Catatonic Type
A type of schizophrenia in which the clinical picture is dominated by any of the following:
(1) catatonic stupor (marked decrease in reactivity to the environment and/or reduction in spontaneous movements and activity) or mutism
(2) catatonic negativism (an apparently motiveless resistance to all instructions or attempts to be moved)
(3) catatonic rigidity (maintenance of a rigid posture against efforts to be moved)
(4) catatonic excitement (excited motor activity, apparently purposeless and not influenced by external stimuli)
(5) catatonic posturing (voluntary assumption of inappropriate or bizarre postures)

Disorganized Type
A type of schizophrenia in which the following criteria are met:
A. Incoherence, marked loosening of associations, or grossly disorganized behavior
B. Flat or grossly inappropriate affect
C. Does not meet the criteria for catatonic type

Undifferentiated Type
A type of schizophrenia in which there are:
A. Prominent delusions, hallucinations, incoherence, or grossly disorganized behavior
B. Does not meet the criteria for paranoid, catatonic, or disorganized type

Residual type
A type of schizophrenia in which there are:
A. Absence of prominent delusions, hallucinations, incoherence, or grossly disorganized behavior
B. Continuing evidence of the disturbance, as indicated by two or more of the residual symptoms listed in criterion D of schizophrenia

Table 2. Diagnostic criteria for subtypes of schizophrenia.

1.3 FAMILY RELATIVE RISK STUDIES

Family studies of schizophrenia constitute one of the largest bodies of literature in psychiatry. In his earliest descriptions of dementia praecox, Kraepelin stressed the aetiologic importance of familial factors in this disorder (Kraepelin 1904). Generally the studies have observed the lifetime risk of the illness in the relatives of schizophrenics and compared this with the risk in control families and the general population. This method does not distinguish between genetic and environmental factors. However if a disease is non-familial then it is less likely that heritable factors play a substantial role in its aetiology.

The first such study performed by Rudin in Munich (Rudin 1916) showed an increased incidence of schizophrenia in their relatives, this has been subsequently confirmed by many other workers (reviewed Gottesman and Shields 1982). A valid criticism of these earlier studies was that they did not use standardised diagnostic criteria or controls and the probands as well as the relatives were not diagnosed blindly.

For a meaningful interpretation of family data the following methodological requirements should be used.

1. Structured interview schedules with explicit inclusion and exclusion criteria for psychiatric illness and good interrater reliability.
2. Blind evaluation of family members with respect to diagnosis and kinship status.
3. The family study method (ie., direct interviews).
4. Reliable operational criteria for the schizophrenia spectrum disorders.

Recent studies that fulfill most of these requirements will be reviewed here. They have often produced conflicting results and possible reasons for this will be critically examined.

One of the first such studies by Tsuang et al (1980) reported on the morbidity risks of schizophrenia and affective disorders among the first degree relatives of probands with schizophrenia, mania, depression and surgical conditions as a control group. Relatives of 200 schizophrenics, 100 manics, 225 depressives and 160 controls matched for sex, pay status, and age were personally interviewed some 30 to 40 years after the probands admission (so that most relatives had passed through the risk period for schizophrenia 15 to 45 years) using the Iowa structured psychiatric interview. Blind diagnosis of first degree relatives of the probands gave a morbidity risk for

schizophrenia of 4.3%. This may be lower than expected since many of the relatives of the probands were parents and since schizophrenics have a reduced fertility the sample may contain many non-penetrant carriers. However the result was significantly higher than the control group.

Baron et al 1983 studied the familial relatedness of schizophrenia and schizotypal states using the method of parental mating types. Matings between two affected individuals should have a greater risk in the offspring than matings with one affected individual or two unaffected individuals. RDC and DSM III criteria gave a morbidity risk for schizophrenia of 7.3% with one affected parent, and a risk of 14.5% for definite schizotypal personality disorder. This study indicates that schizotypal disorder may be genetically related to schizophrenia or at least it occurs in the same families. However Baron noted that the siblings in the sample were not all through the age of risk for schizophrenia and this raises the possibility that some of the cases of schizotypal disorder represent preschizophrenic states and not a separate illness category.

Using modern criteria two studies failed to replicate this finding (Pope et al 1982 and Abrams et al 1983). Pope and coworkers studied the first degree relatives (N = 199) of 39 probands using DSM III categories. They found no cases of even "possible" schizophrenia and argued against a large genetic component for schizophrenia. They suggested that in earlier studies some schizophrenics would now be classified as DSM III personality disorder and that schizoaffective and acute schizophrenia represent primarily cases of affective disorder. Since both of these display a significant hereditary component a family study of schizophrenia containing cases of these other disorders might exaggerate the actual genetic component. Abrams studied the relatives (N = 128) of 30 schizophrenic probands satisfying their own and RDC criteria. They found a morbidity risk of 1.6% which was not significantly different from the general population.

The inconsistencies in these studies arise from methodologic inadequacies. The inability of these two studies to find a significant risk for schizophrenia is likely to be due to their small sample sizes, neither were large enough to significantly reduce the possibility of a type II error. If you remove the parents of the probands from Pope et al study (since they have a reduced risk) out of the children who are left only 50% have reached the age of risk. Abrams had no control group in their study, and used telephone interviews which are know to be less sensitive in diagnosing schizophrenia. Pope also used the chart review family history method (obtaining information from other family

members). A study by Winokur et al 1972 showed that this method found a risk for schizophrenia 40% less than that found by individual field follow-up. These inadequacies could lead to incomplete ascertainment in both the Pope and Abrams studies.

Two larger studies (Guze et al 1983; Baron et al 1985a) did show significantly elevated risks for schizophrenia in relatives of probands. Guze et al carried out a 6-12 year follow up study of 500 schizophrenic probands and 1249 relatives using Feighner's classification. The overall risk in first degree relatives for definite and probable schizophrenia was 8.1%, significantly higher than the control group. Interestingly the sensitivity of these criteria at the start of the study was lower than at follow up because approximately one third of the cases only became evident at follow up. Many of these later cases received diagnosis of unipolar depression or schizophreniform disorder. The risks reported by other studies in other populations could be lower than the true value when longitudinal follow up information is not collected.

In the study by Baron et al (1985) 750 first degree relatives of 90 chronic schizophrenics were personally interviewed. The control group was taken from randomly chosen acquaintances of the well siblings of the schizophrenic probands and matched for age, socioeconomic status and ethnicity. Schizophrenia was diagnosed according to the RDC since it has a more restrictive definition of chronic schizophrenia. All other diagnosis were according to DSM III. The average risk to siblings and parents was 8.5 and 4.4% respectively. The difference was attributed to the reduced fertility of schizophrenics. The risks were significantly different to the risk in the control group. The study nullifies the argument of Pope et al, that other types of psychiatric disorders may exaggerate the familiarity of schizophrenia since in this study the sample was based on chronic schizophrenia. The study of Guze et al (1983) showed no evidence of major depression in schizophrenic relatives and this also argues against the hypothesis of Pope et al.

A study by Kendler et al (1985) further resolved the controversy over the familiarity of DSM III schizophrenia. Personal interviews of 723 first degree relatives of 253 patients and 1056 first degree relatives of 261 surgical control patients were performed. In all of the analyses the risk for schizophrenia was significantly greater (18 times) in the relatives of schizophrenics than those of the controls. The risk of other non-affective psychotic disorders was also significantly higher in the schizophrenic group than in the controls.

Gershon et al (1988) studied a relatively small sample of 24 schizophrenic probands and 108 relatives, diagnosed according to RDC criteria. A significant increase in risk for schizophrenia compared to the control group was found even though the less sensitive telephone interview system was chosen for the data collection.

A second study with a similar sample size failed to find an increased risk for either schizophrenia or schizophrenia spectrum disorders in first degree relatives of DSM III schizophrenic probands compared to controls (Coryell et al 1988). In Kendler's (1988) critique of this work he found four methodologic limitations. First, the sample size was relatively small (185 control and 80 schizophrenic relatives); Kendler calculated that this would give a 40% chance of a type II error. Second, a different interview instrument for the schizophrenia spectrum was used (the structured interview for DSM III personality disorders) compared to other studies. Thirdly many of the relatives were interviewed by telephone making diagnosis of the spectrum disorders without visual clues and lastly the control group was unmatched.

Family studies can also help clarify the familial relationship between schizophrenia and other psychiatric disorders. In the early history of psychiatric genetics the view was held that a single inherited predisposition to mental disease existed (Griesinger 1861). In constrast Kraepelin(1904) articulated that the causes of different psychiatric disorders should be distinct. The data from Tsuang et al (1980) supported the distinction between schizophrenia and affective disorders as separate entities. The distinction between schizophrenia and mania alone was less clear cut. Similar results were demonstrated in Kendler's 1985 study in which schizophrenia and unipolar disorder were found to be independent and unrelated. The risk for bipolar disorder and affective disorders with psychotic symptoms was elevated in the relatives of schizophrenics but not significantly so from the controls. One interpretation is that if affective disorder develops, a familial predisposition to schizophrenia may lower the threshold for the development of psychotic symptoms. Another interpretation is that the presence of atypical features could complicate diagnosis. Thirdly there may be assortative mating for schizophrenia and bipolar disorder.

The situation with schizoaffective disorder is less straight forward. Gershon et al (1988) demonstrated that there was no tendency for schizoaffective disorder to aggregate separately from schizophrenia. An increased incidence of bipolar disorder was found in the relatives of patients with schizoaffective disorder but not in the

relatives of patients with schizophrenia. This leaves the conundrum as to whether schizoaffective disorder is a bridge in a continuum of the major psychoses or a separate entity. Baron et al (1982) attempted to solve the issue. He subdivided a group of schizoaffective disorder probands according to the predominance of affective or schizophrenic symptomatology according to the RDC and carried out a family study. This subdivision enhanced the familial resemblance between the schizoaffective affective type and affective disorder on the one hand and between the schizoaffective schizophrenic type and schizophrenia on the other hand.

In summary the results of these family studies using modern diagnostic criteria verified the original findings that schizophrenia is a familial disorder. They also indicate that the spectrum disorders are also familial and related to schizophrenia. Affective disorders tended to aggregate independently of schizophrenia, but the case of bipolar illness is less clear cut. This could be due to atypical features of psychosis and might become apparent at follow up. Schizoaffective disorder clearly segregates in both schizophrenic and bipolar affective disorder families. Whether it provides a link between the major psychosis is unclear. Separation of schizoaffective disorder into mainly affective or schizophrenic types (Baron et al 1982a) provides evidence for a dichotomy of the psychoses.

These results have important implications for future familial genetic studies of segregation and linkage. The selection of disorders which are possibly genetically related to schizophrenia on the basis of family studies will reduce the effects of false positive and false negative cases in the analyses. In addition a careful choice of diagnostic criteria and interview methods must be made to clearly distinguish between disorders related and unrelated to schizophrenia.

1.4 ADOPTION STUDIES OF SCHIZOPHRENIA

Adoption and cross fostering studies seek to further clarify the respective contributions of the genetic endowment and the environment because this method separates the effect of genes and the environment more clearly than in family risk and twin studies.

The most plausible non-genetic form of familial transmission for a behavioural syndrome like schizophrenia is "vertical cultural transmission" in which a particular trait is learned by offspring from their parents. Probable examples are religious belief and political affiliation. For schizophrenia pure cultural transmission would predict that the high risk

for schizophrenia in the children of schizophrenic parents results from "learning" to be schizophrenic from their parents. Consequently children of schizophrenic biologic parents who have been reared by non-schizophrenic adoptive parents should have a low risk for schizophrenia, and children of a non-schizophrenic biologic parent reared by an adoptive schizophrenic parent should have a high risk for schizophrenia.

Several studies have examined the role of this simple form of vertical cultural transmission in schizophrenia. Heston (1966) found a significant excess of schizophrenia in foster home reared children of schizophrenic mothers. Probably the best designed studies addressing this point are those of a large Danish-American collaboration which had the advantage of the excellent Danish National Adoption Register. In the original study (Rosenthal et al 1971) approximately 5500 non-familial adoptees were identified who had entered the age of risk for schizophrenia. Their biological parents were identified through a central psychiatric register. Those with mental illness had their adoptees matched with adoptees from psychiatrically normal parents for sex, age, age at adoption and socioeconomic status. Diagnosis of schizophrenia and schizophrenic spectrum disorders in the adoptees demonstrated rates of 31.6% for the index group and 17.8% for the control group. The finding was a significant excess of schizophrenia and spectrum disorders in adopted away offspring of schizophrenic biologic parents compared with controls. Nearly all of the parents became ill some time after the birth of the adopted child (average of 11 years) consequently the child's early environment was psychiatrically normal. Lowing et al (1983) confirmed this original result by re-analysing the data but using the modern operational criteria of DSM III. The use of stricter criteria still yielded three times as many schizophrenic spectrum disorders in the index group as in the control group and this difference was statistically significant.

Kety et al (1975) expanded the Danish adoption study further by attempting to interview personally all available biologic and adoptive relatives of the index and control adoptees. In addition the control adoptees were also personally interviewed. However the index adoptees were not interviewed personally and diagnoses were made with psychiatric hospital records. The preliminary results of a blind review of the psychiatric interviews were reported using global diagnostic categories (based in part on DSM II). The results replicated and extended those of the previous report, chronic, borderline and uncertain schizophrenia again were significantly concentrated in the biologic relatives of the index adoptees.

Since these original studies the approach to psychiatric diagnosis has undergone considerable change. The previous emphasis on global diagnostic descriptions has shifted to operationalized diagnostic criteria. These changes have prompted re-examinations of the interviews collected. Spitzer (1979) applied RDC diagnosis to all those relatives and adoptees that contained cases of chronic, borderline, acute or uncertain schizophrenia by the original study. They found a high rate of agreement with the diagnosis of Kety et al (1975). Another study examined the interviews of the "schizotype" (Gunderson et al 1983) in particular. This study made both clinical and DSM III based diagnoses of borderline and schizotypal personality disorder. They identified a personality syndrome in the biologic relatives of schizophrenic adoptees, which was particularly characterized by social dysfunction, eccentricity, poor interpersonal rapport and suspiciousness.

A complete follow-up applying modified DSM III criteria to all the available interviews with adoptees and relatives was carried out by Kendler and Gruenberg (1984). The DSM III diagnosis had reasonable agreement with the study of Kety et al (1975). For chronic schizophrenia there was 64.7% agreement. For the other cases 17.6% met schizoaffective (mainly schizophrenic subtype) criteria. 71.4% of the acute schizophrenic adoptees met DSM III criteria for schizophreniform disorder. However only 10% of those receiving borderline schizophrenia were considered DSM III schizotypal personality disorder. Nevertheless the conclusions were in favour of a significant excess of both schizophrenia and schizophrenia spectrum disorders in the biologic relatives of the schizophrenic (25.7%) compared to the control adoptees (2.2%). A low frequency of major depressive disorder was found in the biologic relatives of the three adoptees with a diagnosis of schizoaffective disorder (mainly schizophrenia subtype). This suggests that this type of schizoaffective disorder is genetically part of the schizophrenia spectrum and not genetically related to affective illness a finding confirmed by Baron et al (1982). A high frequency of schizotypal personality disorder in the biologic relatives of adoptees with the same diagnosis was demonstrated which suggests that genetic liability to this disorder can be directly transmitted.

Further expansion of the Danish adoption study analysed adoptees not only from Copenhagen but the whole of Denmark (Kety 1987). 9300 adoptions were identified and this larger sample confirmed the original findings that chronic schizophrenia and borderline or uncertain schizophrenia were found to concentrate significantly in the biological relatives of schizophrenic adoptees as compared to controls, but not in their

adoptive relatives. Criticisms by Lidz and Blatt (1983) of the Danish adoption study have been refuted in the responce by Kety et al (1983). Additional evidence that genes are probably more "schizophrenogenic" than parenting was illustrated by Wender et al (1974) who found a significantly higher rate of schizophrenia spectrum disorders in the offspring of schizophrenics raised by normal adopting parents than in the offspring of psychiatrically normal adults raised by parents who subsequently became schizophrenic.

The adoption study results taken together suggest that simple cultural transmission is not likely to play a major aetiologic role in schizophrenia. However there is still the possibility that schizophrenia is transmitted by a more "subtle" form of learning from parents who have some features of schizophrenia. If true this hypothesis would predict that the schizophrenia spectrum disorders would be common in the adoptive parents of schizophrenic patients. The re-analysis of the Danish adoption study (Kendler and Gruenberg 1984) demonstrated no evidence for this hypothesis. Another prediction of this "subtler" form of vertical cultural transmission is that stepsiblings of schizophrenic patients should be at a high risk for the disorder because they are exposed to the same rearing environment that produced a schizophrenic patient. Kety et al (1975) found no difference in the frequency of the schizophrenia spectrum disorders in the stepsiblings of schizophrenic adoptees compared to those of the control adoptees. Recently a Finnish study (Tienari et al 1989) suggested a closer link between genes and the environment for the development of schizophrenia. They found an increased risk of schizophrenia and spectrum disorders in adoptees of schizophrenic mothers, if the adoptive family also suffered from a severe mental illness compared to "healthy" adoptive families.

The available evidence from the adoption studies indicate that schizophrenia is not learned from parents but genetic factors are the most influential cause of the familial transmission. These genetic factors are also important in the liability of several schizophrenic like syndromes including schizoaffective disorder (RDC schizophrenia subtype) schizotypal disorder and paranoid personality disorder. Conversely the studies demonstrate that these factors do not contribute to the liability of Affective disorders including the Bipolar disorders. Since the familial clustering of schizophrenia is an expression of shared genetic factors, the application of molecular genetic techniques to informative high-density pedigrees can be justified.

1.5 TWIN STUDIES OF SCHIZOPHRENIA

For more than 50 years a major focus of psychiatric genetics has been to determine to what extent the familial clustering of schizophrenia is due to genetic versus non genetic factors. To address this question the use of the twin method has been an important source of information about the aetiological role of genetic factors in schizophrenia.

The twin method has been used to evaluate the role of genetic factors in a trait by the comparison of concordance rates for the disorder in monozygotic or identical twins and in same sex dizygotic or fraternal twins. Monozygotic twins have the same genotype while dizygotic twins, like full siblings, share on average 50% of their genes. The assumption often adopted in twin studies is that monozygotic and same sex dizygotic twins share environmental factors to approximately the same extent. However models where there are unequal effects in twins have been used for behavioural traits. It seems likely that for such a severe disorder as schizophrenia then the equal environments assumption is valid. Therefore in the case of schizophrenia differences in concordance between the two twin types are likely to be due to the influence of genetic factors.

Despite this there have been several objections to the use of twin studies for schizophrenia. One was that monozygotic twins may be more susceptible to schizophrenia because of their physical similarity and psychological relationship. However, the frequency of schizophrenia in monozygotic twins compared to the general population shows no significant differance (Rosenthal 1960). Another criticism of the twin approach is based on the observation that monozygotic twins share more of their social environment than do dizygotic twins, and the chance that this similarity in social environment rather than genetic similarity makes monozygotic twins more similar than dizygotic twins. Kendler (1983) in his thorough review argues against this conclusion. There are two possible hypotheses. The first is that similar phenotypes in monozygotic twins are caused by their similar social environments. The second is that the similar phenotypes are caused by their identical genotypes, if true this would predict that monozygotic twins would develop similar phenotypes regardless of their social environment, and the similarity in social environment of monozygotic twins is the product of the phenotypic similarity.

Twin studies evaluating personality and cognitive, perceptual and language skills have investigated these two hypotheses. One natural experiment is where the true blood typed zygosity does not agree with the family perceived zygosity. Studies investigating these situations showed that twins who were really monozygotic but thought they were dizygotic behaved in a more similar manner than twins who were really dizygotic

but thought they were monozygotic (Matheny 1979). The trend was towards twins behaving on the bases of their genetic similarity. A second approach was to take advantage of the fact that monozygotic twins though genetically identical differ considerably in their degree of physical similarity. Plomin et al (1976) showed no significant correlation in monozygotic twin pairs between degree of physical and personality similarity for 228 pairs. Again the social environment seemed to have little effect on twin behaviour. If such an effect was truely operating the prediction would be that the physically similar monozygotic twins would be more similar in their behaviour. The available evidence suggests that the similarity of the social environment of monozygotic twins is a result of the behavioural similarity of the twins.

In an attempt to rule out the possibility of environmental bias in twin studies in schizophrenia Kendler (1983) showed that monozygotic twins who were very physically similar were no more likely to be concordant for schizophrenia than twins who were relatively dissimilar. His results suggest that differential treatment by the social environment on the basis of physical similarity between twins is not likely to be a significant bias in twin studies of schizophrenia. Further clarification on this point came from a study by Gottesman and Shields (1982) who reported on 12 monozygotic twins who had been separated early in life and reared apart. Seven of these (58%) were concordant for schizophrenia and this has been interpreted as being highly suggestive of a genetic influence.

The evidence to date seems to validate the twin method as an approach to estimate the genetic contribution to schizophrenia. This evidence, as pointed out by Kendler (1983) is susceptible to a number of "reverse biases", which may decrease the difference in concordance in monozygotic and same sexed dizygotic twins. There are two main factors for this possible bias; assortative mating and the twin transfusion syndrome.

Assortive mating has been shown to occur for a number of mental illnesses (Slater 1971a). This would result in dizygotic twins sharing on average more than 50% of their genes. This would increase the dizygotic twin concordance whilst the monozygotic twin concordance would remain the same, thereby decreasing the difference between the two. As a consequence an underestimate of the genetic contribution would result.

Monozygotic twins are either dichorionic or monochorionic, and in monochorionic twins sometimes differential transfusion of blood occurs between twins. This is called the

twin transfusion syndrome. The syndrome can result in large differences in the weight and size of monozygotic twins at birth. Obstetric complications may be of aetiologic importance in schizophrenia (Jacobsen et al 1980, Owen et al 1988). These are hypothesised to act as precipitating factors on a genetic susceptibility since no increase in the frequency of schizophrenia is recorded amongst monozygotic twins. It might be argued that twin transfusion could decrease the concordance for schizophrenia in monozygotic twins leading to an underestimation of the importance of genetic factors. In conclusion the evidence points to the fact that twin studies for schizophrenia are valid and other complicating factors probably reduce the estimate of the importance of genetic factors.

Concordance ratios for monozygotic versus dizygotic schizophrenic twins are consistently in excess of 3:1 (Gottesman et al 1982). It is true to say, however, that some of the earlier studies suffered from methodological deficiencies the most important being selection bias, poor zygosity determination, lack of operational criteria and age correction. Gottesman and Shields (1972) used systematic ascertainment by selecting their probands from the Maudsley twin register where all patients who were twins were consecutively listed. They found a concordance ratio of 58:12 and taking a weighted average from their own and four other methodologically similar studies (Gottesman and Shields 1976) they found a concordance ratio of 46:14. The discordance rate of over 50% in monozygotic twins tends to support the importance of non genetic biological or environmental factors.

The raw concordance rates for different twin studies are quite heterogeneous (Table 1). The probandwise concordance ranges from 8% to 28% in the same sexed dizygotic twins and from 33% to 78% in monozygotic twins. This variation could occur for several methodologic reasons. For example, differences in severity of illness in index twins, differences in diagnostic criteria, differences in completeness of ascertainment and chronogenetic effects such as age related penetrance can all substantially alter observed concordance rates. However in any single study these factors should affect equally the concordance for monozygotic and dizygotic twins. Therefore methods combining data from both monozygotic and same sexed dizygotic twins to provide a single estimate of the role of genetic factors should allow a comparison between studies.

One method (Kendler 1983) is the H'c statistic.

H'c = 1 - Cdz/Cmz

Cdz and Cmz equal the probandwise concordance rate for the disorder in dizygotic and monozygotic twins respectively. This value can vary from 0 to 1 ; higher values reflect the greater aetiologic importance of genetic factors. Table 1 give the H'c values for each of the studies. The method assumes that monozygotic and same sexed dizygotic twins share aetiologically relevant environmental variables to a similar extent. The H'c values for the studies in Table 1 are all similar and do not vary significantly from one another. The mean genetic contribution is 0.71. So when the possible methodological differences between studies are corrected for, the results of all twin studies of schizophrenia are quite consistent. They all suggest that genetic factors are of major aetiologic importance in schizophrenia and the proportion of genetic variance is as high as for diabetes (0.71) and higher than those for hypertention (0.58. Harvald et al 1965) and coronary heart disease (0.49. Berg 1981).

Table 1. Probandwise concordance for schizophrenia in monozygotic and same-sex dizygotic twins and their H'c statistic.

		Monozygotic concordance			Dizygotic concordance			Author
Year	Nb	Nc	%	Nb	Nc	%	H'c	
Essen-Moller	1941	11	7	64	27	4	15	.77
Slater	1953	41	32	78	61	14	23	.71
Inouye	1961	55	33	60	11	2	18	.70
Kringlen	1967	69	31	45	96	14	15	.68
Fisher et al	1969	23	14	61	43	12	28	.54
Gottesman et al	1972	26	15	58	34	4	12	.79
Allen et al	1972	111	42	38	130	11	8	.78
Tienari	1975	21	7	33	42	6	14	.57
Leonhard*	1982	44	30	68	34	7	21	.70

(*see Kendler 1983)

Nb = Total number of co-twins of proband twin.
Nc = Number of co-twins of proband with schizophrenia.

Estimates made by Fulker (1973) by combining all the available twin study data at that time suggests that the variance contributed by the family environment in the susceptibility to develope schizophrenia is less than 1%. This implies that the role of mothering and fathering in causing schizophrenia is virtually nil. This does not preclude a role for the family in causing readmission and affecting overall prognosis as described in theories concerning expressed emotion in schizophrenic families. Fulker also estimated the effect of unique environmental variance that affects one member of a twin pair and not the other is about 20%. This indicates that the non familial environment has some role to play prehaps triggering episodes of illness by such factors as head injury, birth trauma, viruses, and child birth in women (Eaton 1985).

A number of studies have investigated the effects of varying definitions of schizophrenia in twin studies. Gottesman and Shields (1976) showed that definitions of schizophrenia which were either very broad or very narrow reduced the ratio of monozygotic to dizygotic twin concordance. They suggested that a moderately loose definition seemed to correlate with the strongest genetic tendency. Further studies extended this work by applying several operational diagnostic criteria for schizophrenia on the same series of twins (McGuffin et al 1984; Farmer et al 1987). The probable level of caseness for schizophrenia by the RDC (Spitzer and Endicott 1975), the criteria of Feighner et al (1972) and the DSM III criteria (American psychiatric association 1980) all identified a highly genetically loaded form of the disorder. However this does not mean that the resultant diagnoses are valid. Despite the lack of external validation these diagnostic systems are found to be the most effective at predicting outcome which is an indirect method of validation (Stephens et al 1982). Other criteria were found not to be so sucessful at identifying a genetic loading. The Schneider orientated interview, the present state examination (Reich et al 1972) had a heritability which was effectively zero and was relatively poor at predicting outcome. The criteria of Taylor et al (1975) were found to be extremely restrictive and had poor inter-rater agreement, they reached the conclusion that schizophrenia was non familial or at best only weakly so.

In conclusion the criteria of the RDC, Feighner and DSM III all identified a highly heritable disorder as well as providing good inter-rater agreement. These two factors are important in the choice of opperational criteria to be used in future family linkage studies. As a final point of interest these twin studies demonstrated that genetically identical persons may be discordant for all subtypes of schizophrenia but the hebephrenic subtype had a higher heritability than the paranoid subtype. This fact might

be used to influence adjustments of phenocopy rates in linkage studies.

An extension of the twin method has been to investigate the frequency of schizophrenia in the offspring of schizophrenic monozygotic twins and their normal co-twins (Fischer 1971; Gottesman et al 1989; Kringlen et al 1989). This would allow the comparison of the frequency of shizophrenia in a group where one of the parents has schizophrenia with another group where one of the parents have the same genes as a schizophrenic person but no psychotic illness. One goal of this strategy is to try to differentiate between genetic factors that are necessary for the development of schizophrenia and other factors that may increase risk. Such factors could be triggers, exacerbators, or contributors to severity or poor outcome. Fischer's sample contained eight monozygotic pairs discordant for schizophrenia, with a total of 11 offspring from the probands and 28 offspring from the unaffected cotwins. The conclusion of this Danish study was that the frequency of schizophrenia and schizophrenia like psychosis in the children of schizophrenic and normal monozygotic twins was equal. The result being interpreted as demonstrating that a genetic factor present equally in the schizophrenic and normal monozygotic twins was responsible for causing schizophrenia and that environmental factors associated with being reared by a schizophrenic parent had little influence. However because of the limited number of individuals investigated the results were not statistically significant and could have occurred by chance. At best they demonstrate a trend.

Kringlen et al (1989) analysed offspring of discordant monozygotic twins from Norway. They found no significant difference in the rate of DSM III schizophrenia and schizotypal disorder in the offspring of the proband compared with the normal cotwin. However they did report a non-significant trend that schizophrenia spectrum disorders were less frequent in the offspring of non psychotic cotwins compared to that of the index twin. They argued that this might be evidence for the role of environmental factors modifying schizophrenia by the effect of a schizophrenic parent on a child. However this study also had a limited number of individuals and the findings could have arisen by chance. Another criticism of the Norway study was that only 40% of the offspring had gone through the risk period for schizophrenia (15-45 years).

Gottesman et al (1989) carried out a follow up study on Fischer's twin group some 18 years after the original study in order to clarify the situation. All individuals were classified using DSM III-R criteria. They found a morbid risk for schizophrenia and related disorders in the offspring of schizophrenic monozygotic twin and the normal

cotwin to be 16.8% and 17.4% respectively. They also analysed the risks in the offspring of discordant dizygotic twins. The risks were 17.4% for the offspring of the schizophrenic twin and 2.1% for the normal cotwin. The risk for the offspring of the normal monozygotic cotwin compared to the normal dizygotic cotwins were found to be significantly different. They concluded that there was no support for the hypothesis that rearing by a schizophrenic parent is necessary or sufficient for causing schizophrenia in the offspring.

Hypotheses as to the cause of the discordance between monozygotic twins have included, stochastic processes, a high incidence of new mutations or to a high incidence of non genetic phenocopies. The evidence of equal risks in the offspring of the proband monozygotic and normal cotwin does not support a high incidence of phenocopies. The results indicate that there is a degree of incomplete penetrance or expression of a schizophrenia genotype. If these results are valid then they are encouraging to molecular genetic strategies as they suggest that phenocopies are infrequent within any one pedigree.

1.6 SEGREGATION ANALYSES OF SCHIZOPHRENIA

To test a particular genetic hypothesis in human populations the geneticist has to fit probability models to family data: that is by comparing the observed proportion of affected siblings and offspring with the proportion expected according to a particular genetic hypothesis. This is referred to as segregation analysis. The main problems of such studies arise through the different methods of ascertaining families and affected individuals, pooling data from different families, incomplete ascertainment, inaccurate diagnosis and genetic heterogeneity.

Statistical support for a specific genetic model can be considered as evidence in favour of an underlying genetic aetiology. This information would be invaluable in genetic linkage studies, because demonstration of a major gene effect would provide impetus for the linkage strategy and greater precision in the estimation of the genetic parameters for that analysis. Identification of the mode of inheritance could help to elucidate the evolutionary and genetic mechanisms which have maintained schizophrenia in the population despite its apparent selective disadvantage (reduced fertility). Lastly a precise knowledge of genetic transmission would be useful in genetic counselling.

Genetic modelling of the transmission for complex disorders can take into account reduced penetrance, phenocopies and multifactorial-polygenic (MFP) background effects. Reduced penetrance is interpreted as the incomplete or absent manifestation of a disorder in individuals who have the disease genotype. For schizophrenia evidence of this is demonstrated from family studies and the finding of discordance in monozygotic twins. Phenocopies are those individuals who manifest the disorder even though they do not carry the disease genotype, possibly due to environmental or other genetic effects. MFP is interpreted as more than one gene of additive effect (equal weight) contributing to an individuals liability to the disorder acting jointly with environmental effects.

A further two factors that are important for the correct parameterisation of transmission models in schizophrenia are the classification of who is affected and genetic heterogeneity. A too restrictive definition of schizophrenia would result in a large proportion of false negative cases, whilst a too inclusive definition would generate many false positive cases. As described above studies of twins have indicated that RDC and DSM III criteria for schizophrenia provide an empirically useful approach for genetic studies. However the genetic models available do not easily allow for heterogeneity of linkage. In this case statistical methods to detect the fact that single major gene effects in schizophrenia at several loci can operate independently to produce schizophrenia must be employed.

Segregation analysis of family data employs likelihood methods which permit comparisons of genetic models under certain assumptions. In practise conclusive evidence for one model cannot be found because the extent of underlying heterogeneity of gene effects is unknown. Likelihood refers to the probability of observing a given data set under a particular hypothesis. The values of the genetic parameters that maximize the likelihood are the maximum likelihood estimates of these parameters. The ratio of the maximum likelihood under the null hypothesis (the observed pattern occurring by chance) to the maximum value of the likelihood under a particular genetic model is termed the ratio criterion. Therefore the smaller the likelihood ratio the less likely the null hypothesis is true. The statistical analysis consists of rejecting the null hypothesis by the ratio being smaller than that of a certain test statistic (for example a Chi-square value). If the underlying assumptions are varied, analyses may favour a number of models none of which can be conclusively proven.

Three types of genetic analyses have been used for schizophrenia with family pedigree data. These are the single major locus (SML), multifactorial-polygenic (MFP) and mixed models. A brief description of each follows.

1.6.1 THE SML MODEL

This model is based on the assumption that the inheritance of a disorder is a consequence of a single locus with two alleles (A,a) (Morton et al 1971). Three genotypes are distributed throughout the population AA and aa the homozygous states, and Aa the heterozygotes (FIG 1). These genotypes are each given a mean on a liability scale and are represented by three normal distributions whose variances are determined by the amount of other non-allelic (genetic or environmental) variance. The peaks of each curve is the mean value for the population as a whole for each genotype. The penetrance of the genotypes is the measure of the area of the normal curves above a certain threshold value on the liability scale (shaded area). ie dd has 100% penetrance, and DD 0% where D represents the normal allele and d represents the mutant allele.

The values of the population prevalence can be compared to the risk to relatives of a proband. Several hypotheses can be tested. Mendelian autosomal dominant and autosomal recessive models with contributions from the environment can be compared to the observed risk for relatives from family studies or to recurrence risk in multigenerational pedigrees. The effects of two or more related phenotypes can be tested by incorporating different thresholds, that is relatives of a more severe form of the disorder are at a greater risk than relatives of a proband with a less severe form. Paranoid schizophrenia could be considered less severe because the evidence from family studies points to a lower recurrence risk to relatives than for other subtypes of schizophrenia (Elston 1980).

1.6.2 THE MFP MODEL.

The MFP inheritance (Falconer 1965) involves more than one gene and random environmental factors each contributing a small but additive effect to the expression of the phenotype. The liability to the disorder corresponds to an underlying quantitative effect defined by the sum of all genetic and environmental factors. This liability is assumed to have a normal distribution in the population (FIG 2). The threshold is the point on the liability scale where all individuals above this point are affected and those

Figure 1. The SML model (adapted from Reich et al (1971). The horizontal axis represents the liability scale. The three genotypes are AA, Aa and aa (a denotes disease allele). M_0, M_1 and M_2 are the genotypic means of the three genotypes. The penetrance is represented by the shaded area above the threshold T. Below the penetrance of aa is 100% and that of AA is 0%.

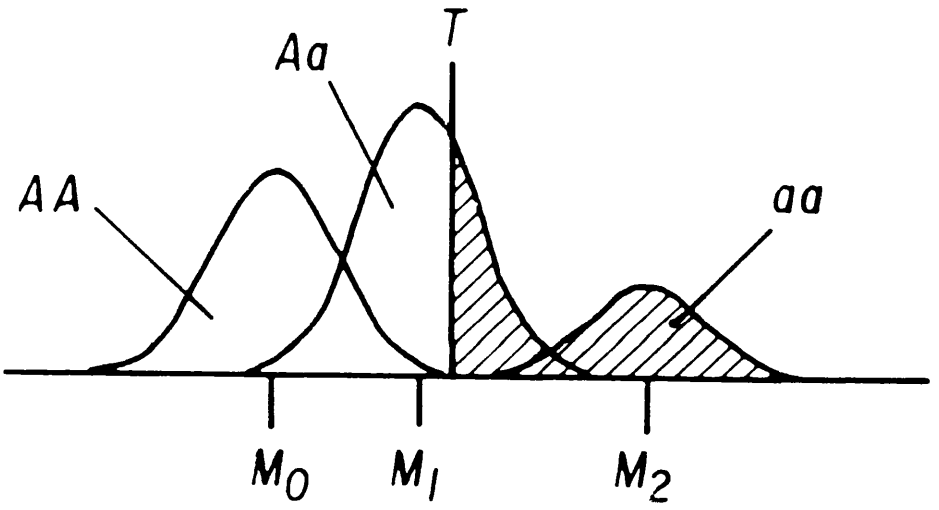
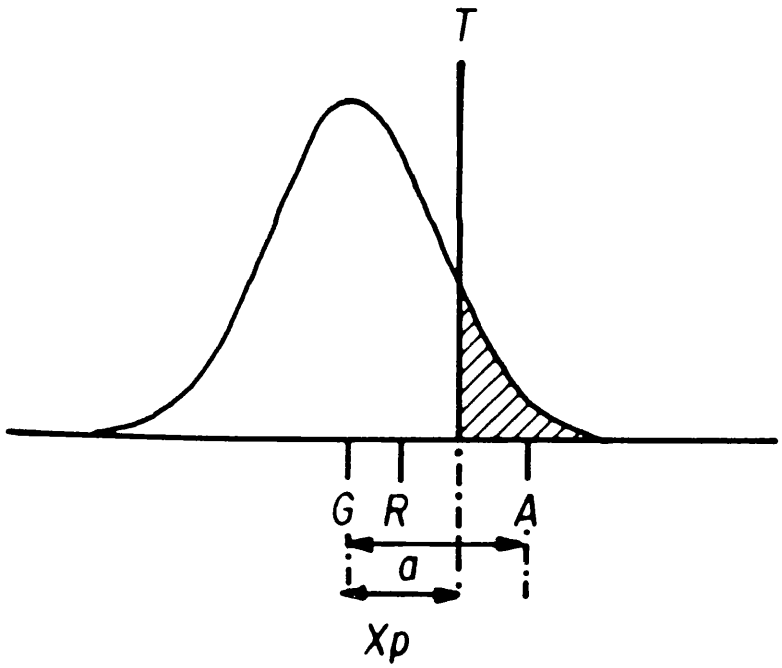


Figure 2. The MFP model (adapted from Falconer 1965). The underlying liability to the disorder is represented on the horizontal axis. The normal curve depicts the phenotypic distribution in the general population; the mean of the distribution is denoted by G. T is the threshold of the liability above which all individuals whose liability to the disorder exceeds T are affected with the disorder. Thus the shaded area represents the prevalence of the disorder in the general population. The liability at point T exceeds G by Xp (the distance between the two points). A is the mean liability of affected individuals, and a is the distance between A and the mean of the phenotypic distribution, G. R is the mean liability of the relatives of affected persons.



below are normal. The shaded area denotes the population prevalence for the disorder. G represents the mean liability for the general population whilst A stands for the mean of affected individuals. Relatives have a greater risk for a genetic disorder than that of the general population and their mean liability is higher as denoted by R. The heritability (the proportion of the variance that is genetically inherited), can be calculated from knowing the incidence in relatives of a proband and by knowing the population prevalence (Falconer 1965). According to the model the consistency of observed and expected heritability estimates across different classes of relatives can be interpreted as favouring or weakening the MFP hypothesis.

1.6.3 OTHER MODELS.

The SML and MFP are rather simple models for analysing complex behavioral traits, consequently less restrictive models have been devised. The mixed model (Lalouel et al 1983) combines that of the SML with a polygene background and environmental effects split into common environment within sibships and random environment. The advantage of this model is that the relative contributions of a SML and MFP effects to the overal liability can be evaluated.

Other models include the two major locus theory which examines the inheritance of two separate loci that interact in the development of the disorder (Kidd and Gladstein 1978) and a polygenic model with graduated gene effects to allow variable contributions to the liability from several loci (Matthysse et al 1979).

1.6.4 RESULTS FROM SEGREGATION ANALYSES APPLIED TO SCHIZOPHRENIA.

The general findings of a number of studies are given in the Table 1 which shows that there is considerable variation between studies. In summary seven out of a total of twelve SML analysis found evidence to support this model (see Table 1), whilst five out of six MFP analyses were in favour of this model(see Table 1). Two studies rejected both of these hypotheses as opposed to one study which found an acceptable fit to both (Matthysse et al 1976, Baron 1982b, Kidd et al 1973). Two studies using more complex analyses provided evidence for the two locus hypothesis however one of these investigations also could not reject the SML and four locus models (Debray et al 1978 and 1980, Book et al 1978). Only two studies have investigated the mixed model and both found the family data compatible with this hypothesis (Risch et al 1984). These conflicting results may be a consequence of population differences, but it is more likely that diagnostic uncertainties and over simplifications of the true genetic model are to

blame.

Nearly all of the studies were based on average risk figures for different classes of relatives of the probands and did not use the information available by including the particular segregation pattern of the disorder within families. Five studies which did take into account pedigree structure (see Table 1) nevertheless produced conflicting results. One study (Risch et al 1984) provided statistical evidence for a major locus with a recessive mode of transmission, however the ascertainment of schizophrenic pedigrees might be said to have an inherent bias favouring this model because of the presence of reduced fertility resulting in the selection of families where the parents of schizophrenics are unaffected. Another potential pitfall is that only eight of the studies in the table tested different models simultaneously and support for one mode of inheritance can only be considered strong if all other types of inheritance have been excluded for the same data.

In general the analyses are inadequate to account for all the variables affecting the true mode of transmission. Important factors such as genetic heterogeneity are ignored, and interactions between genes or epistasis are not accounted for in the models. Furthermore assortive mating and reduced fertility which are both relevant to schizophrenia are not dealt with adequately, and the populations analysed may have not been large enough to produce a statistically unequivocal result. The limitations of the analytical methodologies are further compounded by diagnostic uncertainties. Therefore the following requirements seem necessary for family studies and genetic modelling in schizophrenia. 1.Studies should be based on a well defined series of probands and families from within a single population in order to alleviate the extent of heterogeneity in the sample studied. 2.Sampling bias should be reduced by using prospective proband identification. 3. Operationalized diagnostic criteria should be used to increase the reliability of psychiatric diagnosis and allow comparison between studies. 4. Personal interviews rather than the family history method should be used to reduce ascertainment biases particularly for the schizophrenia spectrum. 5.Blind diagnoses by the clinicians should be carried out. In Table 1 it is disappointing that only two studies have taken these obvious rules into account (Tsuang et al 1982 and 1983; Baron et al 1984). It is also noteworthy that these two studies included the spectrum disorders. On the assumption that such disorders are truly manifestations of the same genetic aetiology as full blown schizophrenia, which the family and adoption studies seem to indicate, studies which do not include these cases as affected will have a number of false negatives.

At this juncture, it is argued that there is no firm conclusion about the mode of inheritance for schizophrenia, the inconsistency in the results must be partly attributable to the variations in the methods of data collection and analysis and the apparent heterogeneous nature of the disorder. Classical segregation analysis to sort out these factors appears hopeless, because there are seldom enough family data to estimate the large number of unknown parameters required to accurately model the complex trait. Nevertheless linkage analysis is considered robust enough to detect a major or minor gene effects despite incorrect assumptions being incorporated in the linkage analysis (Clerget Darpoux et al 1986). The risk ratio, another method of genetic modelling, decreases with the degree of relationship between a proband and relatives, and the rate of decrease depends on the mode of inheritance underlying the trait. Risch (1990b) applied the risk ratio to prevalence figures for schizophrenia and demonstrated that they best fitted an epistatic model of two or three major loci. This supports the use of linkage analysis to find these genes.

TABLE 1. GENETIC MODELS OF SCHIZOPHRENIA: APPLICATION TO FAMILY DATA

Study	Sample	Origin	Finding
SML analyses			
Book et al 1953	Sibs/parents	North-Sweden	SML
Slater 1958	As above	Germany	SML
Garrone 1962	Sibs	Switzerland	SML
Heston 1970	MZ twins sibs/offspring parents	Europe/USA	SML
Elston et al 1971	MZ/DZ twins relatives	Germany/USA	SML
Slater et al 1971b	Sibs/offsring relatives	Europe	SML
Karlsson 1974	Pedigrees	Iceland	SML
Kay et al 1975	Sibs/parents	England	SML rejected
Elston et al 1978	Pedigrees	USA	SML rejected
Tsuang et al 1982	Sibs/parents	USA	SML rejected
O'Rourke et al	MZ/DZ twins	Europe	SML

	1982	sibs/parents offspring		rejected
Risch et al	1984	sibs/parents	USA	SML rejected
MFP analyses				
Gottesman et al	1967	MZ/DZ twins relatives	Europe	MFP
Rao et al	1981	MZ/DZ twins relatives	Europe	MFP
Tsuang et al	1983	Sibs/parents	USA	MFP rejected
Ungvari	1983	As above	USSR	MFP
McGue et al	1983	MZ/DZ twins relatives	Europe	MFP
Studies of both models				
Kidd et al	1973	As above	Europe	SML and MFP
Matthysse et al	1976	MZ twins Sibs/offsring	Europe	SML/MFP rejected
Baron	1982b	Sibs/parents offspring	Germany	SML/MFP rejected
Two-locus model				
Elston et al	1977	Pedigrees	USA	rejected
Studies of multiple loci				
Debray et al and	1978 1979	Pedigrees	France	SML/2-locus/ 4-locus
Book et al	1978	Pedigrees	Sweden	2-locus
Mixed SML-polygenic model				
Carter et al	1980	Sibs/ offspring	USA	SML and MFP
Risch et al	1984	Sibs/parents	USA	SML with polygenic effects. MFP rejected.

1.7 VULNERABILITY TRAITS OF SCHIZOPHRENIA

Vulnerability traits are a potential sorce of aetiological information underlying a disorder. Such traits are biological characteristics that are correlated with the genetic susceptibility to the disorder and are assumed to be a part of the pathway from the genotype to phenotype. They may include biochemical (enzymes, metabolites), neurophysiological (attentional, electroencephalographic), and morphological (brain anatomy) factors that can be measured. For these traits to be useful as genetic risk factors they should demonstrate the following features (Rieder and Gershon 1978):

1. that they have a genetic endowment.
2. that they are associated with the disorder at the population level.
3. that they are state-independent (stability of the trait over time independent of clinical state).
4. that they show familial segregation with the illness. Baron (1986) has reviewed the uses and limitations of vulnerability traits in relation to schizophrenia. Current findings for a number of susceptibility traits will be summarised.

1.7.1 NEUROTRANSMITTER ENZYMES

Neurotransmitter enzymes have a role in the metabolism of biogenic amines and some are thought to be involved in the biology of schizophrenia. These include platelet monoamine oxidase (MAO), plasma amine oxidase (PAO), erythrocyte catechol-o-methyltransferase (COMT) and plasma dopamine-beta-hydroxylase (DBH). A proportion of the activity level of these enzymes is under the control of the major autosomal genes and consequently they are heritable and stable. Chronic schizophrenia has been associated with reduced activities of both MAO and PAO (Baron et al 1985a,b), whereas the data for COMT and DBH is contradictory (Baron et al 1984a; 1980). Unfortunately there is little systematic family segregation data to make the apparent association of the enzyme activity with schizophrenia more obvious. Two studies of MAO (Wetterberg et al 1979; Baron et al 1984b) did find a significant association of reduced activity with ill relatives compared with well relatives of schizophrenic probands. However there are certain limitations to the usefulness of MAO as a trait marker. For example there is a considerable overlap of the activity levels between affecteds and unaffecteds, MAO activity may be influenced by neuroleptics and there is a high frequency of the genotype for low level activity of MAO in the general population. Unfortunately these considerations limit the scope for MAO as a major trait marker for schizophrenia. Similar findings and caveats also apply to PAO (Baron et al 1985b). Studies with COMT showed no familial relation with schizophrenia (Baron et

al 1984a) but for DBH there is some inconclusive evidence for reduced levels of activity in schizophrenics with a family history (Wetterberg et al 1979; Baron et al 1980). Despite considerable efforts in the past these enzymes have yet to be implicated as useful vulnerability traits.

1.7.2 THE IMMUNE CONCEPT OF SCHIZOPHRENIA

This has been widely explored with inconsistent and conflicting findings (DeLisi 1984) and has included proposed pathological mechanisms such as serum abnormal immunoglobulins, altered states of the lymphoid system, and production of brain autoantibodies. The binding of serum globulin substance by human brain (a putative measure of brain autoantibodies) was found to be associated with schizophrenics. However high brain serum binding level differences between ill and well relatives of probands failed to reach statistcal significance (Baron et al 1977).

1.7.3 NEUROPHYSIOLOGICAL AND PSYCHOLOGICAL TRAITS

A set of psychological tests known as the information overload tasks (IOT) have been shown to be deficient in schizophrenic patients and in populations at genetic risk for schizophrenia (Cornblatt et al 1985). These deficiencies do show some state independence but they cannot be considered as definite vulnerability traits because evidence for genetic transmission has not been reported.

Eye tracking is a heritable and stable characteristic, because of this it has been investigated in schizophrenia. It is a potentially useful susceptibility trait since smooth pursuit eye movement (SPEM) dysfunctions are found in schizophrenics. The disruptions consist of a larger than expected number of saccadic intrusions and of saccadic smooth pursuit tracking that include "square wave jerks" during pursuit movements and during fixation eye movements. Investigations have estimated that 50% to 85% of schizophrenic patients and 40% of manic depressives have these dysfunctions, whilst the population prevalence is only 8% (Holzman et al 1984). Holzman et al also demonstrated a significant increase of eye tracking dysfunctions in the parents of schizophrenic probands (34% compared with only 10% for manic depressive parents). The eye tracking dysfunction in manic depressives however may be a consequence of lithium treatment to some extent. The first degree relatives of schizophrenic probands tend to have abnormal smooth pursuit eye tracking even if the proband's smooth pursuit is normal (Matthysse et al 1986). To account for this finding Matthysse et al proposed that schizophrenia and disturbed eye tracking are independent expressions of an underlying latent trait which is genetically transmitted. Holzman et

al (1988) later studied the prevalence of both schizophrenia and abnormal eye tracking in the first degree relatives of monozygotic and dizygotic twins discordant for schizophrenia and showed that the transmission of schizophrenia can be accounted for by a single dominant gene when the illness is considered together with abnormal eye tracking . However there are several important caveats of this study. In particular the data was only compared with the one model of transmission, and that model is probably oversimplified. The actual mode of inheritance of the eye tracking dysfunctions is still not clearly understood. In another study there was no co-segregation of SPEM dysfunctions in multi-affected schizophrenic pedigrees (Bryley et al 1992). Despite these limitations the latent trait model has considerable heuristic value and deserves further investigation before it can be considered as a model for genetic linkage studies.

Another approach to the study of cognitive deficit in schizophrenia has been the analysis of long-latency auditory event-related potentials (ERPs). In particular thr P300 (P3), the long-latency "cognitive" potential which occurs 300 milliseconds or so post stimulus when the subject is attending to an event that is unexpected. Abnormalities of P300 latency (increased) and amplitudes (reduced) have been found in schizophrenics (Blackwood et al 1987; Romani et al 1987). However the same abnormalities are known to occur in Alzheimer's disease, borderline/ schizotypal personality disorder and in other psychiatric and neurological disorders (St Clair et al 1988; Blackwood et al 1990,1986). It is of interest that in multi- affected pedigrees a number of unaffected family members also have the abnormalities (Blackwood et al 1990), and in some cases they have become affected at follow-up (Blackwood et al 1992). Though these results are promising, in that a subtype of schizophrenia may be identified by ERPs abnormalities, further investigation is required and the mode of inheritance for this trait needs to be addressed.

1.7.4 BRAIN MORPHOLOGY

The introduction of computed tomography (CT) offered the opportunity to study brain structure in living subjects. Over 90 investigations of brain structure in schizophrenics have been published (Shelton and Weinberger 1986). The general findings were enlargement of the lateral ventricles, enlargement of the third ventricle, atrophy of the frontal lobe, atrophy of the cerebellar vermis, abnormalities of brain density and reverse cerebral asymmetry in schizophrenics compared with normal controls. Further studies using the more sensitive magnetic resonance imaging (MRI) demonstrated that the area of the third ventricle in its most anterior coronal slice was increased by 73% in

schizophrenic subjects and lateral ventricular volume was increased by 62% compared to normal controls (Kelsoe et al 1988) and a reduction in the overall volume of cortical grey matter and of limbic structures such as the hippocampus (Zipursky et al 1992; Harvey et al 1992). These results have given rise to the hypotheses that schizophrenia is a neurodevelopmental disorder (Weinberger 1987). It has been demonstrated that cerebral ventricular volume is familial and probably due to a genetic effect (Reveley et al 1984; DeLisi et al 1986). Weinberger (1984) examined pairs of siblings and found that the schizophrenic sibling had a greater ventricular size than the normal sibling whilst DeLisi et al (1986) demonstrated that change in ventricular volume was associated with the diagnosis of schizophrenia within families. Nevertheless further study is required because in a group of schizophrenics of twin birth no evidence of ventricular enlargement was found where there was a family history of schizophrenia (Reveley et al 1984). However cerebral ventricular size was significantly increased in sporadic cases of schizophrenia associated with birth complications (Owen et al 1988). Although an attractive hypothesis has been proposed (Lewis et al 1987) that a distinction can be made between familial and sporadic forms of schizophrenia, and that this is reflected in brain appearance (with scan abnormalities occurring more frequently in the sporadic cases), it has been subjected to theorectical criticism and fails to completely account for the reports of brain scan abnormalities in apparently familial schizophrenia (Pearlson et al 1985; Nasrallah et al 1983). The original hypothesis that ventricular enlargement was largely confined to those patients with no evident family history is now probably untrue (Lewis 1990). In a study of two British families (Gurling et al 1992) one family showed a significant co-segregation between schizphrenia and multiple focal white matter high signal lesions.

The results show a potential association between schizophrenia and brain morphology changes but there is a considerable overlap with the normal population and controls have to be carefully matched for age, sex, height, weight and head size. Unfortunately the mode of inheritance for brain morphology is unknown though some abnormalities may represent an endophenotype for subtypes of familial schizophrenia.

In summary vulnerability traits provide a useful avenue of investigation but their use in determining underlying genetic aetiology is restricted because there has never been any strong a priori evidence for their involvement and because, in most cases, the genetics of the traits themselves are unclear. The most promising avenues of search for an endophenotype appear to be brain morphology, SPEM dysfunctions and abnormalities of ERPs. The use of genetic linkage markers still provide the most robust investigation

into the genetic aetiology of schizophrenia. The true relationship between the genetics of schizophrenia and these endophenotypes will probably become known when a genetic subtype of schizophrenia has been identified through linkage studies.

1.8 THE FEASIBILITY OF LINKAGE STUDIES IN SCHIZOPHRENIA

1.8.1 GENETIC LINKAGE ANALYSIS

Linkage analysis consists of identifying polymorphic genetic markers that are sufficiently close on a chromosome such that they are inherited together with the disease mutation from one generation to the next. In such cases the marker and disease are said to be linked. The distance between the disease gene and the marker locus can be calculated by observing the number of recombinations that occur (recombination is the rearrangement or crossing over of alleles following exchange of material between pairs of homologous chromosomes during meiosis). The closer the disease locus is to a marker, the less likely recombination is to occur. Recombination is measured by the recombination fraction theta ($0.0 < \text{theta} < 0.5$). This is the proportion of instances within a pedigree that the disease and the marker are not inherited together. A value of 0.5 indicates random segregation of the disease and marker alleles, a value less than 0.5 indicates that linkage may be present.

The method adopted in determining linkage is the maximum likelihood estimate of the recombination fraction based upon the relative probability (Pr) of having obtained the family. The latter is determined by calculating the probability of having obtained the various combinations of the particular traits under consideration on the assumption of there being no measurable linkage (theta $=0.5$) and comparing this with the probabilities based on a range of recombination fractions from 0.0 to 0.5.

$$Pr = \frac{P(\text{family, given theta} = 0\text{-}0.5)}{P(\text{family, given theat} = 0.5)}$$

For convenience Pr is expressed as its logarithm. The Log10 of the relative probability is called the lod score (Morton 1955). The maximum likelihood estimate of theta may be obtained by summing the lods for all the families studied against various values of theta, the recombination fraction value which corresponds to the maximum lod is taken as the best estimate. At any specific value of theta, a lod exceeding 3.0 is said to confirm linkage and a value less than -2.0 rejects linkage. A comprehensive account of linkage analysis has been given by Ott (1985). By considering the relative lengths of

all 22 autosomes it has been calculated that the prior odds of linkage for any two genes (ie. that two genes are syntenic) is 1:17.5 (Renwick, 1971). As a consequence a lod of 3 represents odds in favour of linkage of approximately 20:1 and not 1000:1.

There are now a large number of DNA markers and the effect of using many markers has been investigated in relation to the statistical significance of the lod. Thompson (1984) demonstrated, for the use of multiple markers used in independent tests of segregation that the interpretation of lod scores for a particular disease locus, requires the same level of significance as a single test.

Multipoint mapping (Lathrop et al 1984; Ott 1985) is a method of using linked markers to resolve gene order and linkage with disease loci. However Morton (1988) has recently criticised its use in relation to the mapping of disease genes because of the effects that genotyping errors may have on the multipoint lod with a disease gene. Nevertheless such likelihood analyses are widely used in the search for disease mutations.

1.8.2 THE HUMAN GENOMIC MAP

Since the first description of "inborn errors of metabolism" by Sir Archibald Garrod (Childs 1970) a large number of human genetic diseases have been found to be associated with major genetic components. There are now approximately 4000 recognised genetic loci or disorders (McKusick 1988), including those responsible for prevalent diseases - Cancer, neuropsychiatric, degenerative and cardio vascular diseases and others. Recent advances in molecular genetics have revolutionized linkage studies in man. Prior to the development of these methods, human linkage studies were severely hampered by the paucity of informative markers in any given family. Linkage studies in schizophrenia applying blood group markers have produced inconclusive results (Elston et al 1973; Andrew et al 1987), partly as a consequence of the small number of pedigrees, their small size, and limited area of the genome studied using blood group markers.

The use of restriction endonucleases, and the ability to clone human DNA has uncovered a wealth of hitherto inaccessible genetic variation. This variation, detected by restriction fragment length polymorphisms (RFLP), representing a point mutation in the genomic sequence which either creates or destroys a restriction endonuclease recognition site, has dramatically increased the number of potential informative markers in any given pedigree. A second type of polymorphic marker (VNTRs) whose length

variation is dependent on the number of tandem repeats of a short sequence (Jefferys et al 1986; Nakamura et al 1987) provided a set of highly polymorphic markers with multiple alleles for linkage studies. These methods will eventually define the organisation and even the nucleotide sequence of the entire human genome. At present many genetic disorders have become susceptible to efficient diagnosis at the genetic level, either through characterisation of mutant genes identified and isolated through aberrant metabolic pathways (forward genetics) or through linkage with other mapped loci (reverse genetics). There is obviously a great need to be able to identify and characterise genes associated with the vast majority of human genetic disorders with no identified biochemical defect. Reverse genetics uses as its starting point nucleotide sequences known to map near the disease locus followed by isolation of the entire relevant nucleotide sequence by gene mapping and chromosome walking or jumping procedures.

Just over 10 years ago the goal was to construct a human genetic map using RFLP's (Botstein et al 1980). Intensified mapping activity over the last decade has generated linkage maps for many human chromosomes (Drayna et al 1985, Donis-Keller et al 1987, Dracopoli et al 1988, Lathrop et al 1988, Nakamura et al 1988 a,b,c, O'Connel et al 1988, Tanzi et al 1988 and Rouleau et al 1989). With the development of these polymorphic markers the rate of acquisition of new mapping data has increased dramatically. A recent compilation has identified approximately 1800 expressed sequences and more than 2500 RFLPs and other markers (Human Gene Mapping 11, 1991). Based on these markers a complete linkage map of the human genome at the 3-5 cM level is now very close to reality, in principal permitting linkage analysis of most genetic diseases. Linkage maps use the centimorgan unit of measure (cM), based on the frequency of recombination. Approximately one cM is equal to 1% recombination and the predicted length of a sex averaged human map is 3300 cM (Morton et al 1982). Using RFLP's many human genetic diseases have been localized by linkage analysis. These include Huntington's chorea (Gusella et al 1983), Duchenne muscular dystrophy (Davies et al 1985) and cystic fibrosis (White et al 1985). These localizations have brought about the sequencing of some of the relevant genetic mutations (Monaco et al 1986; Royer-Pokora et al 1986; Friend et al 1987). The success with chronic granulomatous disease, retinoblastoma, and Duchenne muscular dystrophy and others have come largely through the prior identification of a translocation breakpoint, a deletion or some other cytological landmark. However the recent isolation and characterisation of the gene for cystic fibrosis (Riordan et al 1989) has been accomplished entirely by reverse genetics, thereby providing important support for this

approach to the identification and understanding of human genetic disease loci. Though these methods are impressive there are potential problems. For example the gene for Huntingtons disease has yet to be isolated despite large collaborative efforts (Pritchard et al 1991). Of course the availability of a more thoroughly saturated linkage map will speed progress in the future.

A detailed genetic linkage map of the human genome will define positions of all known genetic loci, and allow the identification of RFLPs linked to most and perhaps to all human genetic diseases. These will provide a starting position for the isolation of the relevant disease loci themselves by a variety of rapidly developing and improving physical methods, including genomic walking and jumping and physical mapping with pulse-field gel electrophoresis (Rommens et al 1989). However a linkage map will not provide all relevant information on intragenic and regulatory regions of the genome that may affect the development of human diseases, nor will it remove the difficulties of chromosome walking between linked markers. These problems can be over come in part by the physical characterisation of the human genome and the development of a fine structure physical map. The ultimate goal of the development of a detailed physical map of the human genome is to obtain the ultimate map possible - that is the complete sequence of the genome. For the human genome roughly $3X10^9$ bp of DNA will have to be sequenced which would be aided by improvements in technology (Prober et al 1987). The initial phase of the genome mapping effort is now to focus on the order and spacial relationship of genetic polymorphisms, genes and DNA sequences. Primary goals are to produce a dense linkage map, and a variety of physical maps. Physical maps are either cytogenetic or molecular in nature. Cytogenetic maps order loci with respect to the visible banding pattern or relative position along the chromosomes by the means of somatic cell hybrids, and insitu hybridisation (Eubanks et al 1992). Ultimately the construction of a contig map, a collection of overlapping and contiguous clones is required. The recent development (Burke et al 1987) of yeast cloning vectors of higher cloning capacity than the previously used cosmid vectors and the use of pulse field electrophoretic mapping methods makes this approach even more attractive. Olson et al (1989) have designed a scheme to augment this combination of approaches. This plan uses the universally available sequence information from the termini of all large cloned human fragments, and cloned genes that constitute "sequence-tagged sites" (STSs). These would also include primer sequences from which the intervening fragment sequences could be generated by all interested investigators by the polymerase chain reaction (PCR; Saiki et al 1988). Eventually a complete STS map of the human genome would be available, providing a set of overlapping, cloned and

ordered sequences spanning an entire chromosome or portion of a chromosome.

One of the major impacts of this intensive mapping activity will be to increase markedly the number of human diseases that are recognised as having major genetic components. Even multigenic human diseases could be resolved into discrete Mendelian factors by the use of RFLP mapping (Paterson et al 1988). The identification of such genes allow the development of powerful new approaches diagnosis, detection, screening and even therapy of these disorders. Therefore it seems that human molecular mapping offers the best approach to identifying the genes involved in the aetiology of schizophrenia.

1.8.3 THEORETICAL CONSIDERATION OF THE APPLICATION OF LINKAGE ANALYSIS TO SCHIZOPHRENIA

A workshop organised by the ECE has published (Merikangas et al 1989) recommendations aimed at resolving difficulties applying linkage analyses to the psychiatric disorders. The problems are often subsumed under the term "non-Mendelian inheritance", but there are a number of different concepts which can be considered separately with advantage. These consist of difficulty in defining caseness, incomplete penetrance, presence of phenocopies, and genetic heterogeneity. In addition the mode of transmission is unknown and so the quantitative values attaching to each of the above values is unknown as is the gene frequency of the susceptibility allele. When a disease is inherited in a straightforward Mendelian fashion, this implies that it is possible to reliably diagnose the disease and that everybody who inherits a copy (or, in the case of a recessive gene, two copies) of the abnormal gene will go on to develop the disease at some age during their life span and that everyone else will not. Because under these conditions, the mode of inheritance is fairly obvious it is also possible to make reasonably accurate estimates of the frequency of the abnormal gene in the population based on the prevalence of the disease. However a Mendelian mode of inheritance is not essential for modern methods of linkage analysis, and in fact only represents a special case of a more general model as implemented in the computer program LINKAGE (Lathrop et al, 1985). The general genetic model for detecting linkage between a disease and genetic markers depends only on the concept of there being some major gene effect, which means that the presence or absence of one or two copies of a certain allele at the susceptibility locus has a substantial effect in terms of altering the individual's liability to develop the disease. The genetic mode of transmission is then characterised by the probability of affection, or penetrance, conditional on each of the three possible genotypes (pAA, pAa and paa) and by the

gene frequency of the abnormal allele conferring increased susceptibility. A fully penetrant dominant gene for a disease with no phenocopies is characterised by setting pAA, pAa and paa equal to 0, 1 an 1 respectively, while for a recessive gene the values would be 0, 0 and 1. Where a disease is only partially penetrant the probability of an individual inheriting the abnormal alleles developing the illness will fall to less than 1 and if phenocopies exist, that is to say if some cases have a non-genetic aetiology, then the probability of developing the disease despite having a normal genotype will rise to above 0. Since schizophrenia is clearly not inherited exclusively as either a Mendelian dominant or a recessive disease, and because no cast iron agreement exists as to how they should be most reliably diagnosed, the first problem is to find a way of defining cases which will reflect as far as possible any major gene effect which may be present. The greater the differences between the penetrances conditional on genotype, the more powerful linkage analysis will be to detect linkage. If the definition of caseness is too narrow then many people with the abnormal gene will be classed as unaffected and the penetrance must be set low, but if the definition is too wide, for example including all psychiatric illness, then many people classed as affected will not in fact possess the abnormal gene and so the penetrance conditional on normal genotype (or, more loosely, phenocopy rate) must be set high.

A useful approach to establishing a diagnostic scheme which detects an underlying genetic abnormality is to choose a measure which maximises the ratio of concordance between monozygotic (MZ) and dizygotic (DZ) twins. McGuffin et al (1984) tested 6 sets of operational criteria and found that the Research Diagnostic Criteria (RDC) (Spitzer et al, 1978) and Feighner definitions (Feighner, 1972) gave the highest MZ:DZ concordance ratio, whereas Schneiderian and Present State Examination (PSE) gave only low scores. Farmer et al (1987) found that the maximum MZ:DZ ratio was found when schizophrenia was diagnosed as a psychosis with mood incongruent delusions, together with either schizotypal personality and atypical psychosis. The inclusion of paranoid disorder and affective disorder reduced this ratio, suggesting that they were not genetically related to schizophrenia.

It should be pointed out that there may be a difference between criteria used to produce an optimal strategy for defining cases within a family and those criteria which should be used in order to decide which families to select for study in the first instance. This is because the probability of an individual who has a minor abnormality being a "genetic case" is higher if he is known to have a close relative with the full-blown disease. To maximise the power of linkage analysis one wishes first to

select families in which the disease gene is probably segregating, and then to select as cases those individuals within the family who are most likely carriers of the gene. This means that one may need to use conservative criteria to select families and more liberal criteria to define cases within them (Gurling et al 1991). It is important to note the sensitivity of the lod to misclassification (Skolnick et al 1984). The sensitivity is very dependent on the genetic parameters specified during the linkage analysis. For example, one individual misclassified into affected/unaffected will influence the lod under high and low penetrances to different degrees.

An issue that cannot be completely distinguished from the definition of caseness is that of heterogeneity of linkage in which some cases of an illness may be due to the actions of a particular disease gene whilst other cases may be associated with abnormalities at different genetic loci. Linkage analysis will generally fail if such non-allelic heterogeneity of linkage goes unrecognised. It is possible to analyse the data under the assumption that in only a proportion of families is the disease due to an abnormality at a particular locus whilst other families may be unlinked (Ott, 1986). However such elaborations do considerably reduce the power of linkage analysis. In general the presence of heterogeneity increases the recombination value at which the lod scores peak, by an amount that increases with the amount of heterogeneity (Cavalli-Sfortza et al 1986). There is a corresponding increase in the number of families necessary to establish linkage (Gershon et al 1987). Highly polymorphic markers however will greatly reduce the number of families required to establish linkage (Nakamura et al 1987). It is also true that using two genetic markers at a known genetic distance will increase the power of detecting heterogeneity (Lander et al 1986; Martinez et al 1989). A situation that is catastrophic for the linkage approach is when non-allelic heterogeneity occurs within a family, that is when the multiple cases seen within a family occur as the result of actions of two different genes at separate loci.

With the above considerations to be taken into account one particularly pressing question is whether or not to include cases of manic depression and schizoaffective illness in linkage studies of schizophrenia, and vice versa. Although for much of the modern history of psychiatry manic depression and schizophrenia have been thought of as essentially separate, some advocates have been found for the proposal that the two diseases represent different ends of a continuous spectrum (Kendell and Brockington 1980, Crow 1986). In the context of a genetic aetiology this could be interpreted as meaning that a single genetic defect might predispose to the development of psychosis, and that other genetic or environmental factors might

determine whether the schizophrenic or affective disease developed. However this view has not found universal acceptance, and most authors find that although there may be some diagnostic overlap the different disorders do tend to "breed true". This view has been supported by monozygotic twin and dual-mating studies (Bertlesen et al, 1972, Gottesman and Bertlesen, 1988). If there is a tendency for schizophrenia and manic depression to occur at increased rates in the same families then this may be due to a true shared genetic aetiology (perhaps occurring in only a subset of cases), or it may be due to assortative mating (Merikangas et al, 1983) or difficulty in distinguishing the diagnoses.

The application of linkage analysis to "non-Mendelian" illnesses may be seen formally as performing the analysis with an increased number of degrees of freedom. The lod score method depends on comparing the likelihoods of obtaining the observed distribution of cases and genetic markers under the conditions of linkage and non-linkage. The lod score is the log to the base 10 of the ratio of these likelihoods. With classical linkage analysis there is only one degree of freedom which is the marker-disease gene recombination fraction or, equivalently, the map position of the disease gene. For a non-Mendelian disease with unknown mode of transmission each genotype-specific penetrance represents an added degree of freedom, as does the gene frequency, disease definition and, if heterogeneity is suspected, the proportion of families studied in which the disease gene is segregating. This would seem to add 6 degrees of freedom to the model and accordingly to achieve equivalent statistical significance to a lod score of 3 with the classical method one would have to obtain around double the lod score and use a threshold of around 6 (Kelsoe et al 1989). In fact this is rather a conservative approach, because varying the penetrances and phenotype definitions does not usually make all that much difference to the lod score. The accepted practice of presenting the lod score under several representative conditions is perfectly valid. It is however essential to report that a range of models has been studied and to give some idea of the degree to which the lod score varies between the plausible models. The reader can then form an idea of whether or not the lod score is critically dependent on unjustified assumptions in the model - frequently it seems that this is not the case and that a positive lod score is quite robust to such parameters. An alternative or complementary approach to deal with the problem of unknown penetrance is to obtain a prior estimate of penetrance by a segregation analysis independent of any genetic marker data, for example using the computer programmes ILINK, from the LINKAGE package, or POINTER (Lalouel et al 1981). If this value is then held fixed throughout subsequent linkage analysis then penetrance no longer

contributes to the degrees of freedom of the model. However there are problems with this approach, since families selected for linkage analysis are generally not suitable for segregation analysis, and frequently one does want to alter the penetrance to find a maximum likelihood estimate that does take account of linkage data.

Simulation studies have shown (Cox et al 1988) that linkage is likely to be detected even in the absence of knowledge of the mode of transmission, if a range of models are examined. An important effect of erroneous modelling is that it decreases the maximum lod (or at least the expected one), that can be obtained and furthermore the wrong genetic model for the disease may incorrectly reject linkage for a given recombination fraction. In no instances does erroneous specification of parameters artificially increase evidence for linkage (Clerget-Darpoux 1991). The greatest effect when the genetic model is badly specified is in estimating the recombination fraction which is highly sensitive to the misspecification of the genetic parameters at the disease locus (Clerget-Darpoux 1991). The robustness of lod score estimates can be improved, however by using the information from several genetic markers simultaneously (multipoint linkage analysis) assuming all genotypes are error free.

A second method of linkage which does not require any assumptions about the mode of inheritance is the sib pair method (Green and Woodrow 1977; Weeks and Lange 1988). This method examines the distribution in members of a sibship of the alleles at the marker locus and disease phenotype. However the power of this method is a magnitude less than that of the lod score pedigree method. Furthermore, its power is closely related to the ability to determine whether the alleles of a sib pair are identical by descent, that is they represent the same parental chromosomal locus. This can be achieved only by studying highly polymorphic markers or including the parents. Linkage using this approach can be detected in the presence of heterogeneity (Goldin and Gershon 1988), however, the power of the sib pair method for the detection and estimation of the extent of heterogeneity is much lower than the lod method. Consequently the role that sib pair analysis should play in studying schizophrenia will remain unclear until such time as the underlying modes of transmission and the extent of heterogeneity are known.

CHAPTER 2: LOCALIZATION OF A SUSCEPTIBILITY LOCUS FOR SCHIZOPHRENIA ON CHROMOSOME 5

2.1 INTRODUCTION

2.1.1 PEDIGREE SELECTION STRATEGIES

The application of linkage analysis in schizophrenia is not straight forward necessitating specific strategies in order to maximise success. The following strategies were employed for the selection of suitable pedigrees.

Families sampled were medium and large pedigrees in which schizophrenia occurred in at least three generations with multiply affected generations. These pedigrees might enrich for a particular genetic form of the illness with an unusually high penetrance in which dominant rather than recessive transmission is most likely. Although these pedigrees are highly selected for linkage analysis they still would not increase the chance of finding a false positive linkage result (Cox et al 1988). However they may not represent the genetic aetiology for the majority of schizophrenia cases. Interestingly such pedigrees do fit the most likely mode of transmission of schizophrenia in Iceland (Karlsson 1982; 1988). The use of large families is particularly important if genetic heterogeneity is suspected because the power to detect heterogeneity and the power to detect linkage in the presence of heterogeneity is greater if a few large families rather than many small families are studied (Ott 1986; Risch 1988). Pedigrees with many generations are more useful than nuclear families providing considerably more genetic information, though the amount of genetic information varies considerably with the precise structure of the pedigree (Thompson et al 1978). Consequently fewer individuals are required to establish linkage at higher recombination fractions (Botstein et al 1980). The presence of a large number of affected cases in the families also enables a special approach to linkage analysis to be used because linkage information can be extracted only from affected cases ignoring all unaffected individuals. The results of the analysis would not then be dependent on whether some unaffected individuals were in fact non-penetrant carriers of the disease gene. This circumvents the necessity to vary the penetrance in the linkage analyses and is termed "penetrance free analysis".

The evidence from twin and family studies (see general introduction) suggest that manic depression and schizophrenia have predominantly different genetic

aetiologies. The aim was to obtain a sample of families as homogeneous as possible, therefore all families containing cases of bipolar illness were excluded from the study. The reason for not excluding just those individuals who were bipolar was the possibility that there might be two genetically distinct diseases segregating within the pedigree and as a result there might be some cases (eg schizoaffective disorder or atypical cases) which could not be reliably distinguished as belonging to the schizophrenia or the bipolar genotypes. Interestingly, in every family in which bipolar illness and schizophrenia was encountered it was possible to establish that there were two independent genetic sources for these psychoses within the extended kindreds, suggesting that assortative mating was the explanation for the coincidence of the two diseases. Equal care was taken to ensure that there was only one possible source for a schizophrenia susceptibility allele segregating within the pedigree. Such an approach clearly presupposes a dominant mode of transmission, which did not seem unreasonable in the families recruited and requires that the spouses families are free from major psychiatric conditions. Although it is theoretically possible to perform linkage analysis on a pedigree with two dominant alleles segregating within it, such an approach is inherently problematic and certainly has reduced power. Additionally, if the two alleles were in fact at different genetic locations linkage analysis would certainly be unsuccessful. Using these stringent selection criteria seven families from a larger sample were deemed suitable for linkage analysis. 5 Icelandic, and 2 British kindreds with multiple cases of schizophrenia in at least three generations. The five Icelandic families were ascertained after identification of schizophrenic probands who were amongst currently hospitalised patients during 1986 and 1987. Pedigree-tracing in the Icelandic families was aided by the extensive genealogical records available there, which frequently also contain some psychiatric data. A beneficial side-effect was that the families came from a country which is relatively genetically isolated (Kidd et al 1974), increasing the probability that there would be genetic homogeneity amongst them. The two British families were selected on the basis of having four or more affected individuals. Extensive genealogical tracing was carried out to ensure, as best we could, that there was a single unilateral source for any possible dominant disease allele entering into each kindred.

Genetic linkage research is very sensitive to diagnostic error, especially to false positive diagnosis of caseness. Failure to identify a true case has less severe consequences because such individuals will join the pool of individuals who are either showing recombination or who are non-penetrant carriers. This group of

individuals can receive special attention in the linkage analysis, unlike false positive cases which will continue to have a hidden negative effect on all types of analysis. With this in mind the RDC and DSM III criteria were chosen, since twin studies have shown them to identify psychopathology related to schizophrenia (McGuffin et al 1984; Farmer et al 1986), and would be the most useful for linkage research.

In the seven pedigrees there were 104 individuals who were still alive and who had descendants or ancestors who had been diagnosed as schizophrenic. All 104 people were personally interviewed by a psychiatrist for the study. 98 of these received a full interview using the Schizophrenia and Affective Disorders Schedule - lifetime version (SADS-L) (Spitzer et al 1978) and the remaining 6 unaffected individuals were screened more briefly for the absence of psychiatric disorder. The SADS-L information obtained at interview and extensive medical information made available to us from other sources were used to classify all cases according to both the Research Diagnostic Criteria (RDC) and some of the clinical categories of the American Psychiatric Association's Diagnostic and Statistical Manual of Mental Disorders (DSM-III). DSM-III and RDC diagnoses are the two most widely used systems of psychiatric diagnosis for research purposes and have been shown to be reliable and valid. Independent validation of diagnosis from all sources of information was carried out by two further psychiatrists. This consisted of reviewing all the available information on individuals who received any psychiatric diagnosis and in the event of uncertainty more information was obtained until a consensus was reached. Psychiatric interviewing and diagnosis was carried out blind to the results of linkage work in the laboratory. The results were as below.

(1) Amongst the 104 individuals there were 39 cases of schizophrenia as defined by the RDC. Under the slightly more stringent DSMIII criteria 31 of these qualified as schizophrenic. The remaining eight (8) were classified as having schizophreniform disorder (DSMIII) or unspecified functional psychosis (RDC) and were included as cases of schizophrenia. The patients exhibited a wide variety of the clinical subtypes of schizophrenia (paranoid, hebephrenic etc.) and all had histories of psychiatric hospitalisation and often prolonged treatment with antipsychotic drugs.

(2) A further 5 individuals were diagnosed as having schizoid personality disorders (DSMIII) with (N = 2) or without (N = 3) schizotypal features as defined by the RDC. We refer to this group as those with schizophrenic spectrum disorders.

(3) A further ten individuals had psychiatric illnesses which were not part of the schizophrenic or schizophrenic spectrum disorders. Some of these individuals had received continuous periods of treatment with anti-psychotic drugs even though they had not been diagnosed as schizophrenic. We refer to these cases as "fringe" phenotypes (see Figure 1 for detailed pedigree structure and diagnoses).

These three levels of diagnosis were subsequently referred to as schizophrenia, schizophrenia spectrum and schizophrenia fringe cases respectively. The inclusion of all other psychiatric illness as being possibly indicative of carrying a susceptibility allele for schizophrenia is not quite so anomalous as it might at first appear. In the first instance the high density of cases of schizophrenia in the relatives of such people made it far more likely that their illness represented a "forme fruste" of schizophrenia than would be the case in someone without affected relatives. Secondly, the RDC categories frequently did not fully convey the clinical impression gained from studying the case histories of these individuals - for example some had chronic, atypical "depression" for many years and had been treated continually with neuroleptic medication, yet fell into the RDC category of major depressive illness. For further details of the diagnoses see Brynjolfsson et al (1990). DNA from collected individuals was extracted from peripheral blood lymphocytes and the pedigree structures were tested for paternity as best they could using DNA fingerprints (Figure 2). No cases were found in these pedigrees.

2.1.2 FAVOURED LOCI FOR MOLECULAR GENETIC RESEARCH INTO SCHIZOPHRENIA.

Numerous reviews have publicized the role of molecular genetics in unravelling the genetic aetiologies of schizophrenia (Gurling 1986a,b; Sturt and McGuffin 1985; McGuffin and Sturt 1986; Gershon and Goldin 1987a; Gershon et al 1987b). Recent advances in recombinant DNA technology and the potential availability of many more polymorphisms make genetic marker studies an increasingly attractive prospect in schizophrenia. Linkage analysis is considered to be the most powerful and robust approach although important limitations are referred to, subsumed under "unclear modes of inheritance". Three related strategies are considered appropriate, the random genome search, candidate genes and favoured loci.

Before the availability of DNA markers, linkage studies in schizophrenia were severely hampered by the paucity of traditional protein markers, and produced inconclusive results (McGuffin et al 1983). In the random search method DNA

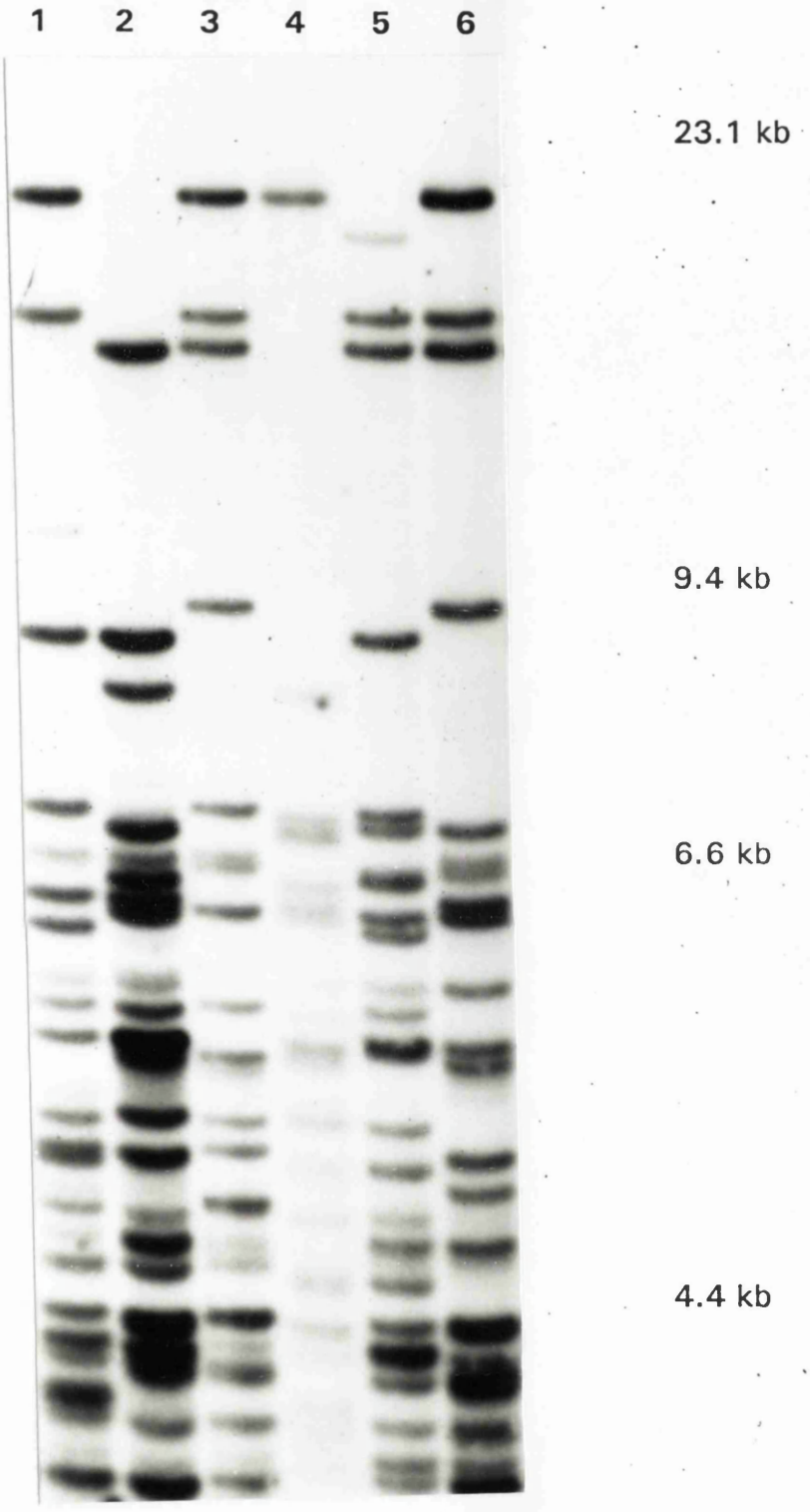


Figure 2. DNA fingerprints of a mother (1) and five children (2-6) for family F20. The fingerprints were generated from Hinf I digested DNA using Lambda 33.6 as a DNA probe (Jefferys et al 1985).

probes are used to analyse the adjacent areas of the chromosome in an attempt to detect a locus of major effect. Probes are chosen either randomly or because of their known chromosomal location. The aim is to exclude regions either as a series of independent two point lods (Edwards 1987) or possibly more efficiently using groups of linked markers in a multipoint analysis (Lathrop et al 1984).

The human genome is approximately 3300 centimorgans (cM) in length and Ott (1985) has calculated that in order to insure with 95% accuracy that one or more marker loci would be within 10cM of a given disease locus 498 markers would need to be typed. However, using markers that are ideally spaced along the genome could considerably reduce this figure to as low as 200 to 300 markers. Ott (1985) demonstrated that in order to reach 95% power in the detection of linkage at recombination distances of 10, 20 and 30cM one needs samples approximately 2,4 and 8 times larger than those required to detect linkage at 5cM. It seems likely that with realistic sample sizes linkage will be rarely detected at distances greater than 10 or 20cM. Both incomplete penetrance and heterogeneity increase the sample sizes required to detect linkage (McGuffin et al 1986; Cavalli-Sforza 1986). For incomplete penetrance at close linkage (theta = 0.05) the detection of linkage for a dominant trait with a penetrance of 80% requires a sample size four times larger than would be needed to detect linkage if the trait was fully penetrant. Though statistically problematic, a random genome search would eventually identify any major gene effects in schizophrenia.

Recombinant DNA technology has generated different types of DNA marker for mapping the human genome. Restriction fragment length polymorphisms (RFLPs) refers to the differences in the sizes of fragments resulting from digestion of DNA with restriction endonucleases. These enzymes recognise specific sequences in DNA and catalyse endonucleolytic cleavage, yielding fragments of defined length which can then be separated according to molecular size by gel electrophoresis. Thus fragments encoding specific sequences and any variations thereof that effect their size can be detected by southern blot hybridisation with a radiolabelled probe (Southern 1975). The different sized fragments are called RFLPs and these can arise as a result of either the creation or loss of a restriction enzymes recognition site or by the insertion or deletion of DNA. Generally RFLPs have a maximum heterozygosity of 50%, but most are considerably less than this figure and as such many meioses are uninformative for linkage analysis. The genetic measure of informativeness is the polymorphic information content (PIC) value of a probe. The

PIC value or informativeness represents the probability that a given offspring of a parent carrying the disease allele at the index locus will allow deduction of the parental genotype at the marker locus (Botstein et al 1980).

Another type of DNA variation is the variable number tandem repeat (VNTR; Nakamura et al 1987). These consist of relatively short tandem repeats of DNA whose length is determined by the number of repeat units. VNTRs are detected by cleaving genomic DNA with a restriction enzyme which cuts close to but not within the tandem repeat, hybridization with a specific probe visualises this variation on a southern blot. These markers can generate more linkage information by having many alleles and consequently higher PIC values. Certain probes which have a common core sequence for VNTRs can recognise many variable loci simultaneously generating the human "DNA fingerprint" (Jeffreys et al 1985). Though highly individual they are difficult to analyse for linkage and the location of the alleles is unknown, their major use is for paternity and forensic testing.

The candidate gene approach uses polymorphic markers for genes suspected of having a pathological role in schizophrenia. For example the hyperactivity of dopaminergic neurotransmission in schizophrenia (Reynolds 1989) indicates that mutations in genes encoding dopamine pathway enzymes and receptors may cause schizophrenia. Examples are the recently cloned dopamine D2 receptor (Grandy et al 1989), the rate limiting enzyme tyrosine hydroxylase (Craig et al 1985) and dopamine beta hydroxylase (Kobayashi et al 1989). Linkage studies using candidate genes usually requires a recombinant between an affected individual and the gene for it to be excluded as having a role. However for schizophrenia there are problems with this approach, a recombinant could arise by miss specifying the mode of inheritance in the linkage analyses, and by the presence of incomplete penetrance, phenocopies and heterogeneity. Candidate genes can only be excluded from having a major role if negative lods (less than -2) are obtained. Association tests in unrelated affected individuals can also be used to investigate a genetic role of a candidate gene, and with large enough samples even minor genetic effects should be detectable.

The third approach that of investigating favoured loci, involves carrying out linkage studies in regions of the genome where there is a prior reason for suspecting an involvement in the genetic aetiology of schizophrenia. An example was the apparent association of albinism and schizophreniform psychosis in a single

pedigree (Baron 1976), with a lod of 1.55 at 10% recombination. The tyrosinase gene for albinism has been localised to chromosome 11 (Barton et al 1988). The cosegregation of Marfans syndrome and schizophrenia in a number of pedigrees has been reported (Romano et al 1987; Sirota et al 1992). Other favoured loci include associations between cytogenetic abnormalities and schizophrenia. Delisi and Lovett (1990) summarized a number of X chromosome abnormalities in schizophrenics, including XXY, and XXX individuals. Crow (1988) postulated that pairs of psychotic first degree relatives are more often than expected of the same sex and this points to a major locus for schizophrenia on the pseudoautosomal region of the X chromosome. The reduced rate of recombination between the Y and X chromosome at the autosomal region implies that affected sibs would often be the same sex if the disease gene was paternally inherited. However the validity of the statistical analysis of this study has been questioned (Curtis et al 1989). Many autosomal abberations have been reported to be associated with the psychoses (Bassett 1992). Some are more likely to have a pathological role and constitute favoured loci for linkage studies. These including chromosome 18q translocations and deletions, the chromosome 19 fragile site (Bassett 1992), a chromosome 11q translocation (St Clair et al 1990) and a chromosome 5 trisomy.

2.1.3 PARTIAL TRISOMY CHROMOSOME 5 COSEGREGATING WITH SCHIZOPHRENIA.

Two individuals of an Asian family had schizophrenia - a 20 year old male college student and his 52 year old maternal uncle. The 20 year old proband was noted on admission to have slightly different facial features from his parents, but looked similar to his uncle. Further investigation showed that the two patients had a constellation of slightly dysmorphic features, including folded protuberant ears, short stature, frontal bossing and renal abnormalities. Cytogenetic investigation of these two individuals and the normal transmitting mother revealed a partial trisomy of chromosome 5 q11.2-13.3 in the cases of schizophrenia, whilst the mother had a balanced translocation to chromosome 1 q32.3 (Bassett et al 1988). The fact that the mother psychiatrically normal suggests that it is not the disruption to chromosome 1 and 5 but rather the partial trisomy that is causing the physical abnormalities and schizophrenia.

2.1.4 AIM

The aim was to investigate the trisomic region of chromosome 5 (5q11.2-13.3) as a candidate region for a shizophrenia susceptibility locus by using DNA markers

mapping to this region for linkage analysis in the seven selected pedigrees. Because the available chromosome 5 DNA markers had not been mapped in relation to the trisomy or localised by in situ hybridisation, two markers were chosen for an initial investigation of this region. The glucocorticoid receptor gene thought to map to 5q11.2 (GRL; Murray et al 1987) and L372-CRI (D5S49), both produced negative two point lods -3.134,-3.533 (theta = .01) respectively for linkage to schizophrenia (Table 1). However, GRL was later shown to map to 5q32 by linkage and in situ hybridisation (Francke et al 1989) and L372-CRI mapped distal to the trisomy (Donis-Keller et al 1989). Linkage analysis was then carried out using two markers which mapped proximal to GRL and L372-CRI, that is p105-599Ha (D5S76) and p105-153Ra (D5S36).

2.2 METHODS AND RESULTS

The inheritance of a two restriction fragment length polymorphisms (RFLPs) known to be on chromosome 5 probably mapping to 5q11.2-13.3 (Leppert et al 1987) were studied in relation to schizophrenia in the seven families. The RFLP at locus D5S76 is defined by TaqI polymorphisms detected at Southern hybridisation (see materials and methods) by the probe p105-599Ha (Figure 3). This probe contained human repetitive elements and required competing with total human DNA for its analysis (see materials and methods). Locus D5S39 is defined by MspI and XbaI restriction enzyme polymorphisms detected by the probe p105-153Ra (Figure 3; Leppert et al 1987). Digested DNA was separated by size in 0.8% agarose gels and transferred by southern blotting to Hybond-N (Amersham Int plc) as described in materials and methods. P105-599Ha reveals a three allele system giving allelic fragments 17 kilobase (kb), 14kb and 10kb in size shown as W, X and Z (allele frequencies 0.32, 0.16 and 0.52 respectively). Probe 105-153Ra hybridizes to Msp I allelic fragments that are 8.7kb and 5.8kb (shown as S and T) with allele frequencies of 0.33 and 0.67. The Xba I RFLP was found to be in linkage disequilibrium with the Msp I RFLP with fragments of 6.3kb and 5.2kb with frequencies .71 and .29 respectively. These loci are linked together at the male genetic distance of about 13.7cM, with a constant male:female recombination ratio of 1.5 (Leppert et al 1987). The genotypes and pedigree structure of all the families are shown in Figure 2.

Three different models for the linkage analysis were employed because of the uncertainty of defining as cases those individuals with spectrum disorders as well

A.

Theta	0.00	0.01	0.05	0.1	0.2	0.3	0.4
F20	-0.81	-0.52	-0.10	0.09	0.22	0.22	0.14
F27	-0.42	-0.30	-0.05	0.07	0.12	0.08	0.03
F35	-0.81	-0.54	-0.15	0.00	0.06	0.03	0.00
F36	0.00	0.00	0.00	0.00	0.00	0.00	0.00
F40	-3.94	-1.38	-0.34	0.04	0.23	0.17	0.05
F41	0.00	0.00	0.00	0.00	0.00	0.00	0.00
F74	-2.41	-0.40	0.18	0.33	0.33	0.20	0.06
Total	-8.39	-3.14	-0.46	0.53	0.96	0.70	0.28

B.

F20	-2.55	-2.17	-1.12	-0.63	-0.21	-0.06	-0.01
F27	-0.30	-0.24	-0.09	0.00	0.06	0.05	0.01
F35	-2.72	-0.75	-0.18	-0.03	0.01	0.00	0.00
F36	-0.38	-0.37	-0.34	-0.30	-0.20	-0.09	-0.02
F40	0.00	0.00	0.00	0.00	0.00	0.00	0.00
F41	0.00	0.00	0.00	0.00	0.00	0.00	0.00
F74	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Total	-5.95	-3.53	-1.73	-0.96	-0.34	-0.10	-0.02

Table 1. Two point MLINK lods for GRL (A) and D5S49 (B) and schizophrenia classified as affected at 99% penetrance.

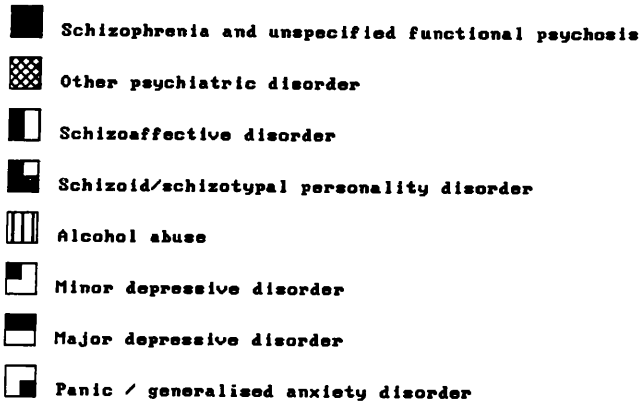
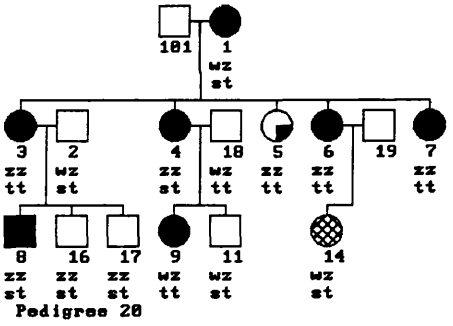
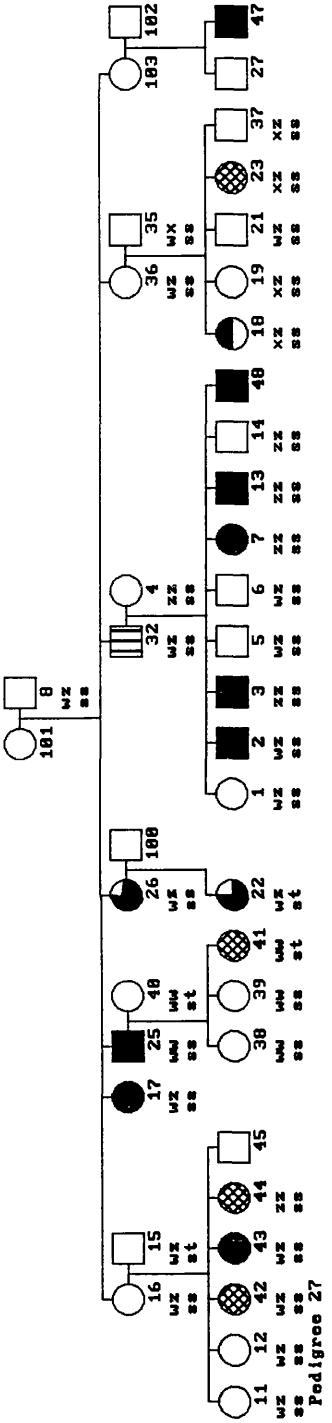
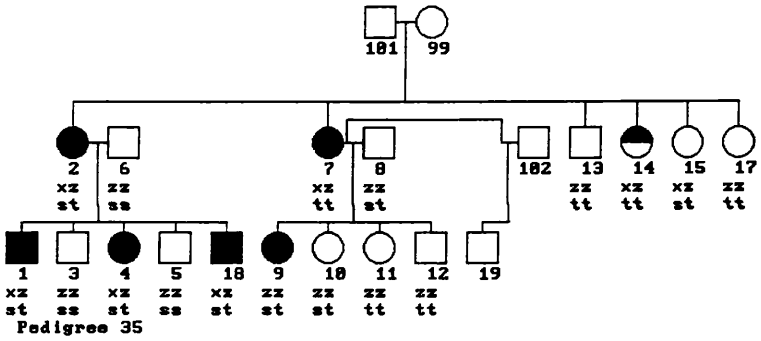
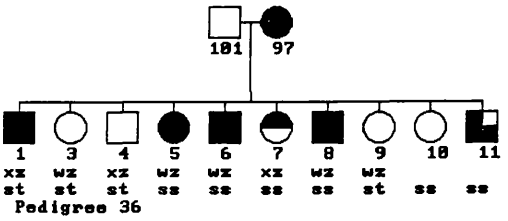


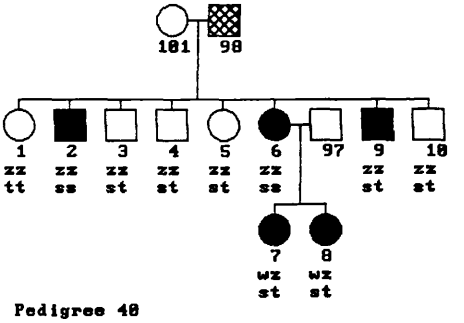
Figure 1. Pedigree structures, diagnoses for the seven analysed families and genotypes for D5S76 and D5S39. F27, F35, F36, F40, F41 Icelandic families and F20, F74 British families. W,X,Z D5S76 alleles and S,T D5S39 alleles.


101
1
wz
st
3
zz
tt
2
wz
st
4
zz
st
18
wz
tt
5
zz
tt
6
zz
tt
19
7
zz
tt
8
zz
st
16
zz
st
17
zz
st
9
wz
tt
11
wz
st
14
wz
st
Pedigree 20

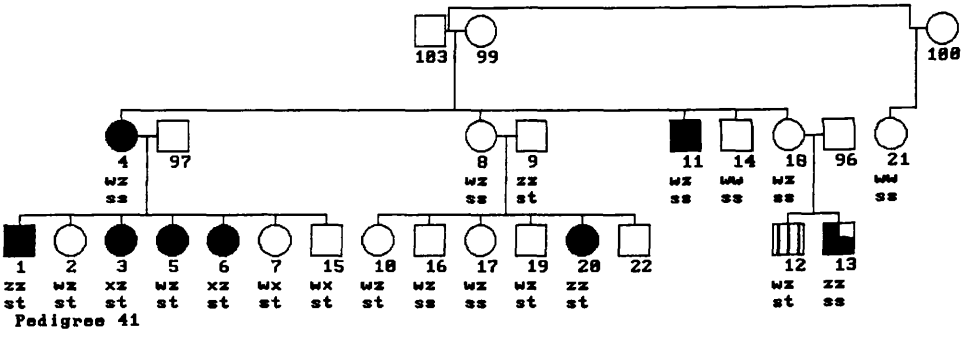

Pedigree 27

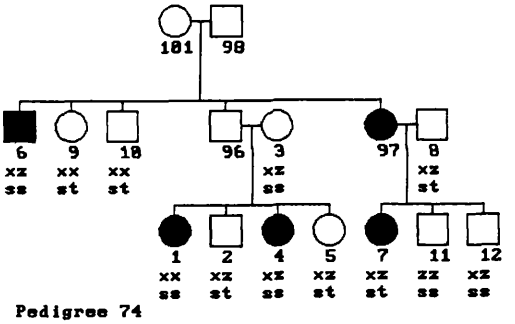

101
99
2
xz
st
6
zz
ss
7
xz
tt
8
zz
st
102
13
zz
tt
14
xz
tt
15
xz
st
17
zz
tt
1
xz
st
3
zz
ss
4
xz
st
5
zz
ss
18
xz
st
9
zz
st
10
zz
st
11
zz
tt
12
zz
tt
19
Pedigree 35


101
97
1
3
4
5
6
7
8
9
10
11
xz wz xz wz wz xz wz wz
st st st ss ss ss ss st ss ss
Pedigree 36


101
98
1
zz
tt
2
zz
ss
3
zz
st
4
zz
st
5
zz
st
6
zz
ss
97
9
zz
st
10
zz
st
7
wz
st
8
wz
st

Pedigree 40


103
99
100
4
wz
ss
97
8
wz
ss
9
zz
st
11
wz
ss
14
ww
ss
18
wz
ss
96
21
ww
ss
1
zz
st
2
wz
st
3
xz
st
5
wz
st
6
xz
st
7
wx
st
15
wx
st
10
wz
st
16
wz
ss
17
wz
ss
19
wz
st
20
zz
st
22
12
wz
st
13
zz
ss
Pedigree 41


101
98
6
xz
ss
9
xx
st
10
xx
st
96
3
xz
ss
97
8
xz
st
1
xx
ss
2
xz
st
4
xz
ss
5
xz
st
7
xz
st
11
zz
ss
12
xz
ss
Pedigree 74

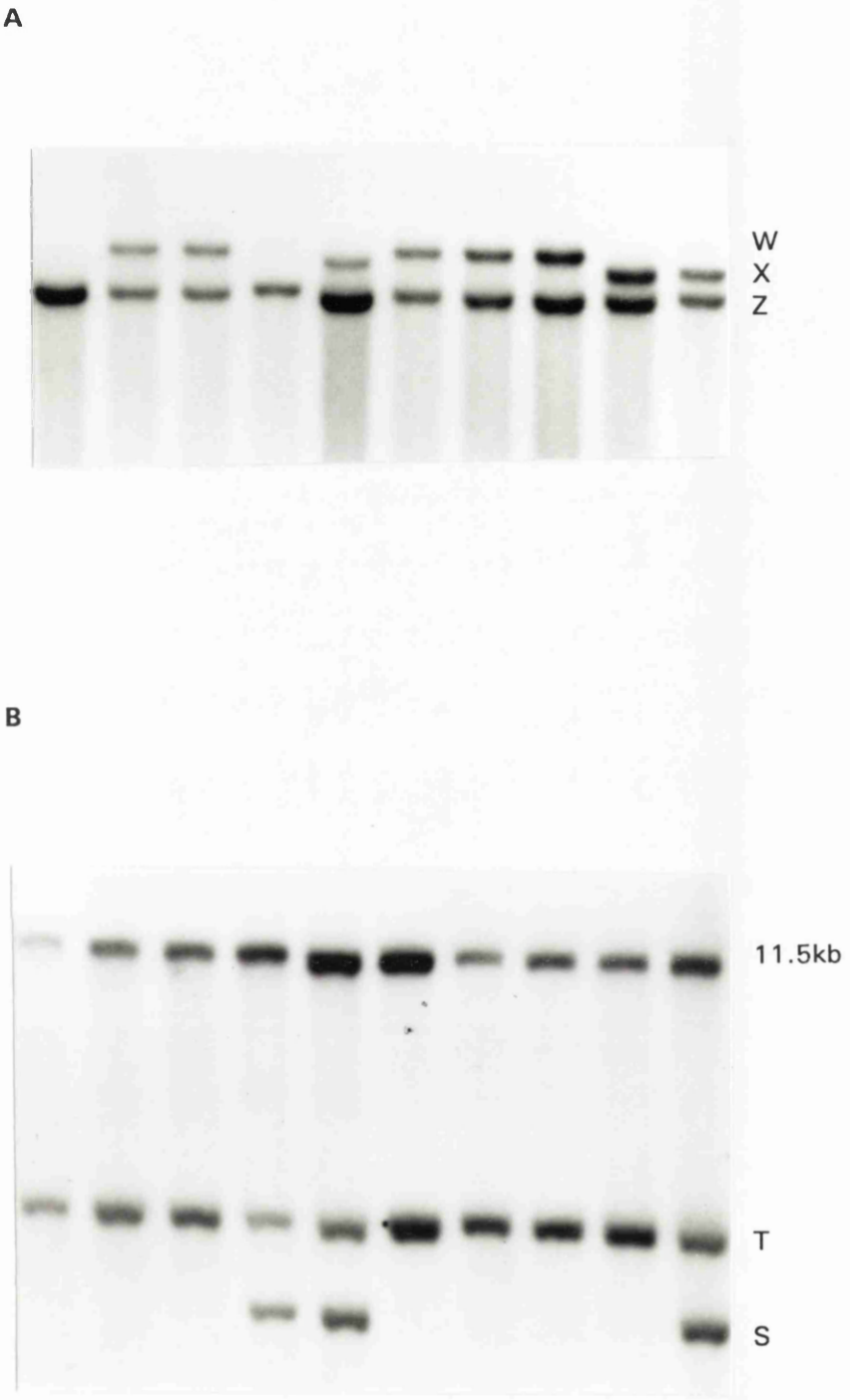


Figure 3. A. TaqI RFLP for D5S76, alleles W,X and Z are 17, 14 and 10kb respectively. B. XbaI RFLP for D5S39, has alleles S and T of 5.2 and 6.3kb respectively and a constant band of 11.5kb.

C

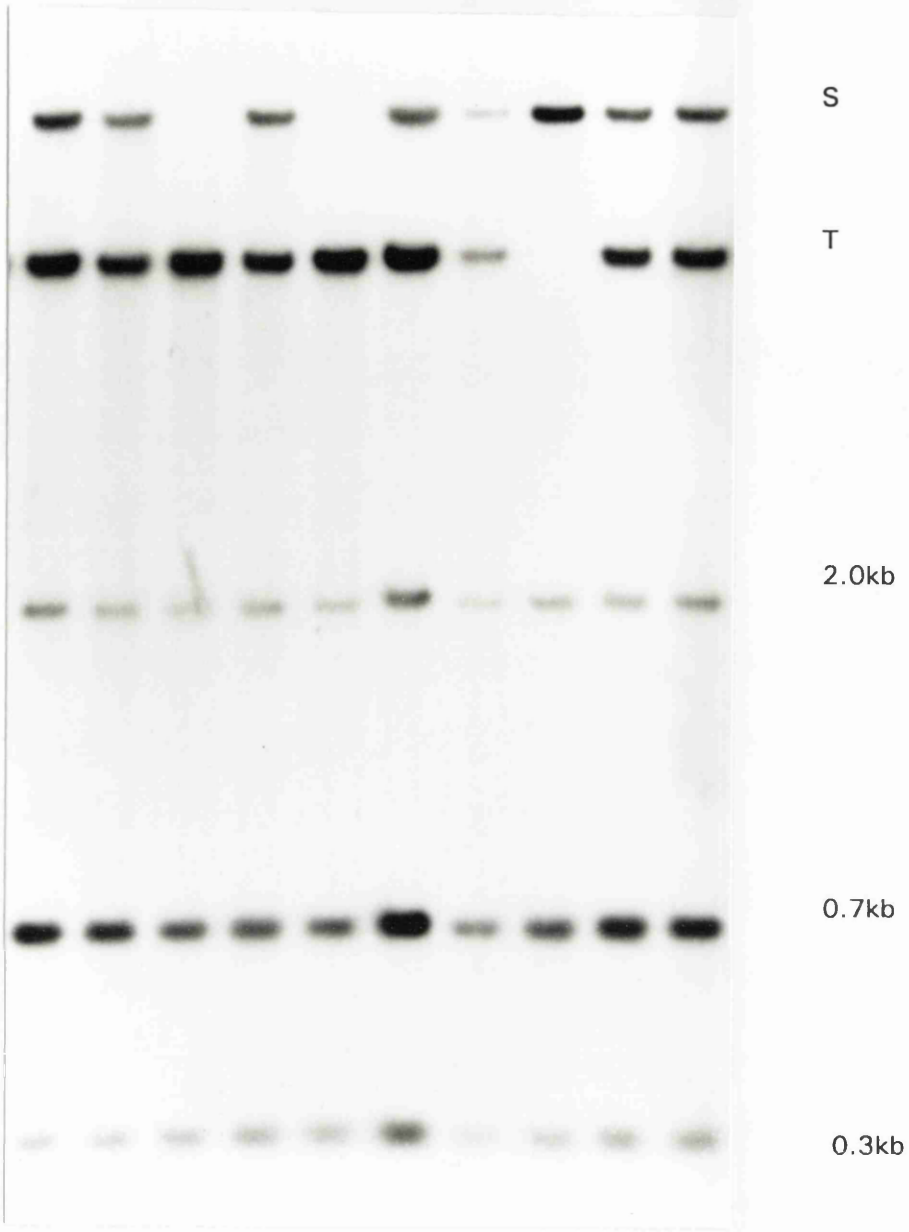


Figure 3 (continued). C. MspI RFLP for D5S39, showing alleles S and T of 8.7 and 5.8kb respectively and constant bands of 2.0, 0.7 and 0.3kb.

as other psychiatric disorders. These models were (1) DOMS, in which only individuals diagnosed with schizophrenia were scored as affected. (2) DOMSS, in which individuals with schizophrenia and schizophrenia spectrum disorders were regarded as cases. (3) DOMSSF, in which schizophrenia, schizophrenia spectrum disorders and all other fringe phenotypes were included. For each of the models described above we varied the degree of penetrance for a dominant gene from 50% to 100% in increments of 10% and computed lod scores for all seven families for a putative schizophrenia susceptibility locus. Equal penetrances were used for schizophrenia homozygotes and heterozygotes. Though the evidence from segregation analysis for a mode of inheritance has been inconsistent, two studies in Iceland did provide evidence for a modified dominant gene effect (Karlsson 1982;1988). In addition to analysing the data by varying penetrance we also employed a "penetrance free" (PF) model in which all "unaffected" individuals were analysed as "phenotype unknown". Although this discards a large fraction of the information it is robust to incomplete penetrance. In all the analyses the gene frequency for schizophrenia was set at 0.0085 (Gottesman et al 1982) to account for a morbid risk of approximately 1%, and the frequency of phenocopies was set with an arbitrary value of 0.1%. Two point lods were calculated using MLINK programme from the LINKAGE package (Ott 1985), and the results given in Table 2 are the lods obtained for each model at 99% penetrance. One family F41 gives a maximum lod of 3.25 at the D5S76 locus. No evidence for non-allelic heterogeneity of linkage was found using HOMOG, and the B-TEST (Ott 1986; Risch 1988). Maximum likelihood estimates of the lod score were calculated from location scores from the LINKMAP programme of LINKAGE (Figure 4), for each of the proposed models using the same parameters as for the two point lods, but maximising the lod with respect to penetrance. The schizophrenia locus was moved from left to right, as shown, starting at an arbitrary distance of -0.4 Morgans (M) from the p105-599Ha locus to a position of 0.8M to the right of p105-153Ra. P105-599Ha and p105-153Ra are at map positions 0.00 and 0.137M respectively. The maximum lods for schizophrenia and the two markers were found at 99% penetrance in the DOMS and DOMSS models and 86% in the DOMSSF analysis as shown. In the PF analysis complete penetrance was assumed. The highest lod obtained was 6.49 with schizophrenia in the middle of the two markers at approximately 0.08 M from each marker under the DOMSSF model. The maximum lods for the other models were 4.33 (DOMSS), 3.22 (DOMS) and under the stringent PF model 2.45. The data was not adequate to make a reliable determination of the gene order of schizophrenia in relation to the two marker loci.

DOMS.

Theta	0.00	0.01	0.05	0.10	0.20	0.30	0.40
F20	-1.19	-0.92	-0.56	-0.43	-0.38	-0.32	-0.16
F27	-0.01	-0.01	-0.01	0.00	0.00	0.00	0.00
F35	-0.86	-0.58	-0.17	0.01	0.10	0.08	0.03
F36	-0.30	-0.30	-0.30	-0.29	-0.21	-0.10	-0.03
F40	0.60	0.59	0.53	0.46	0.32	0.17	0.05
F41	-0.01	-0.01	-0.01	0.00	0.00	0.00	0.00
F74	0.92	0.90	0.81	0.71	0.49	0.27	0.08
Total	-0.85	-0.32	0.29	0.46	0.32	0.10	-0.03

DOMSS.

F20	-1.19	-0.92	-0.56	-0.43	-0.38	-0.32	-0.16
F27	0.00	0.00	0.00	0.00	0.00	0.00	0.00
F35	-0.86	-0.58	-0.17	0.01	0.10	0.08	0.03
F36	-0.41	-0.40	-0.36	-0.17	0.09	0.12	0.05
F40	0.60	0.59	0.53	0.47	0.32	0.17	0.05
F41	-0.01	-0.01	-0.01	0.00	0.00	0.00	0.00
F74	0.91	0.89	0.81	0.71	0.49	0.27	0.08
Total	-0.96	-0.43	0.24	0.59	0.62	0.32	0.05

DOMSSF.

F20	0.75	0.72	0.60	0.44	0.15	-0.05	-0.09
F27	0.00	0.00	0.00	0.00	0.00	0.00	0.00
F35	1.14	1.12	1.04	0.92	0.68	0.42	0.14
F36	-0.40	-0.27	0.25	0.52	0.60	0.43	0.16
F40	0.60	0.59	0.53	0.47	0.32	0.17	0.05
F41	-0.03	-0.02	-0.02	-0.01	0.00	0.00	0.00
F74	-1.00	-0.75	-0.37	-0.20	-0.06	-0.02	0.00
Total	1.06	1.39	2.03	2.14	1.69	0.95	0.26

Table 2. D5S39 MLINK two point lods for the DOMS, DOMSS and DOMSSF models at 99% penetrance.

DOMS

Theta	0.00	0.01	0.05	0.10	0.20	0.30	0.40
F20	-0.80	-0.52	-0.12	0.06	0.14	0.10	0.03
F27	-9.92	-3.59	-2.00	-0.97	-0.16	0.05	0.04
F35	-3.13	-0.21	0.40	0.63	0.74	0.56	0.24
F36	-0.31	-0.31	-0.30	-0.27	-0.17	-0.07	-0.02
F40	0.00	0.00	0.00	0.00	0.00	0.00	0.00
F41	-2.66	1.38	1.95	2.01	1.69	1.09	0.36
F74	-0.23	-0.22	-0.15	-0.08	0.01	0.03	0.02
Total	-17.05	-3.47	-0.22	1.38	2.25	1.76	0.67

DOMSS

F20	-0.80	-0.52	-0.12	0.06	0.14	0.10	0.03
F27	-13.85	-3.35	-0.95	-0.08	0.42	0.39	0.16
F35	-3.13	-0.21	0.40	0.63	0.74	0.56	0.24
F36	-0.21	-0.21	-0.18	-0.14	-0.07	-0.03	-0.01
F40	0.00	0.00	0.00	0.00	0.00	0.00	0.00
F41	-1.10	2.88	3.25	3.11	2.47	1.58	0.55
F74	-0.23	-0.21	-0.15	-0.07	0.01	0.04	0.02
Total	-19.32	-1.62	2.25	3.51	3.71	2.64	0.99

DOMSSF

F20	1.15	1.13	1.04	0.93	0.68	0.40	0.13
F27	-9.37	-4.52	-1.78	-0.59	0.30	0.47	0.30
F35	-2.82	0.26	1.10	1.33	1.24	0.85	0.32
F36	-0.21	-0.21	-0.18	-0.15	-0.09	-0.04	-0.01
F40	0.00	0.00	0.00	0.00	0.00	0.00	0.00
F41	-1.97	2.04	2.50	2.47	2.02	1.33	0.51
F74	-0.23	-0.22	-0.15	-0.08	0.01	0.03	0.02
Total	-13.45	-1.52	2.53	3.91	4.16	3.04	1.27

Table 2 (continued). D5S76 MLINK two point lods for DOMS, DOMSS and DOMSSF models at 99% penetrance.

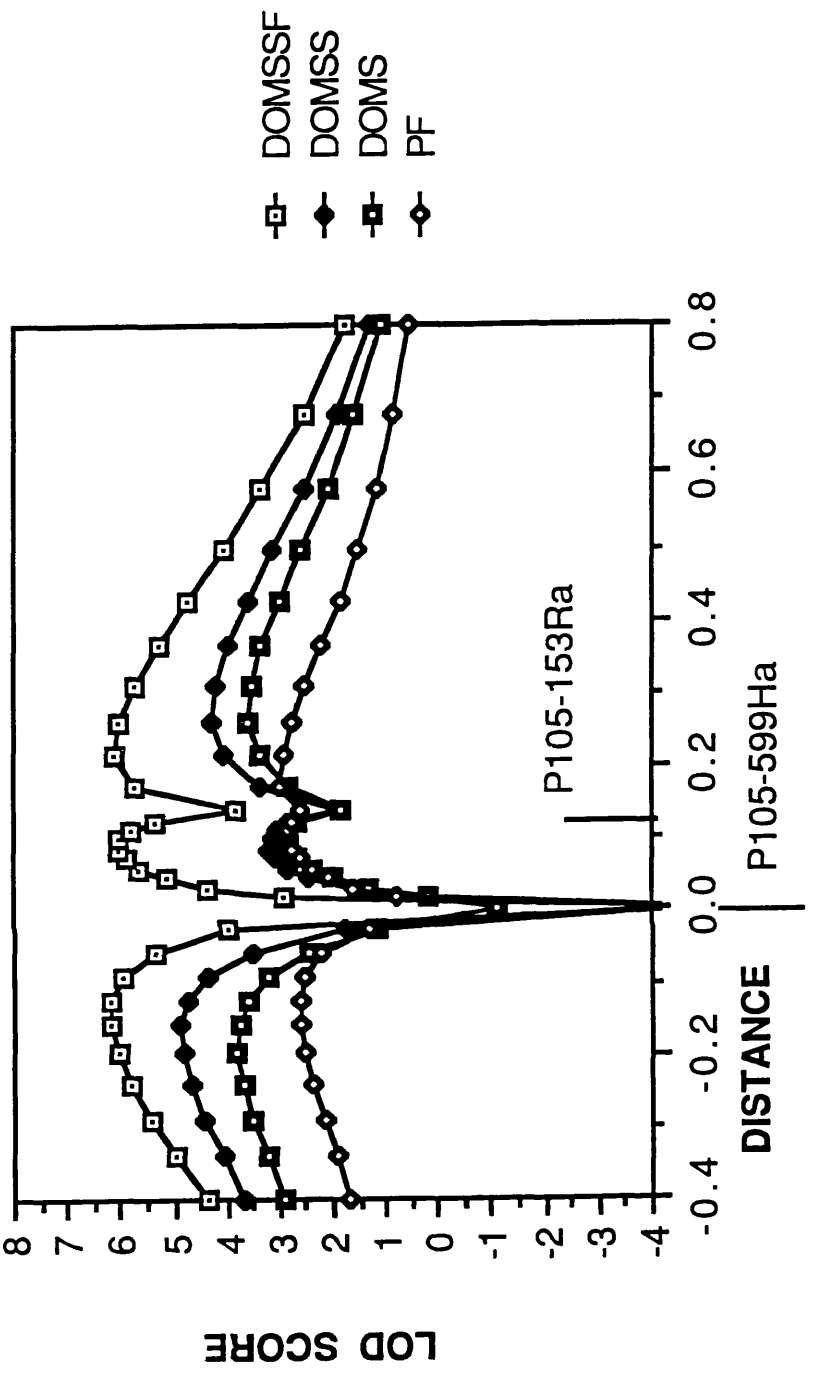


Figure 4. Multipiont linkage map of each model of affection with D5S76 and D5S39.

Maximum likelihood estimates of penetrance for schizophrenia were calculated with ILINK and were 86%, 76% and 73% for the DOMSSF, DOMSS and DOMS models respectively.

A lod of 3.00 is good evidence for linkage for a disease with complete penetrance but when penetrance is used as a parameter the higher lod score of 3.85 is required (E.S. Lander, Whitehead institute USA, personal communication).

2.3 DISCUSSION

The data provided strong evidence for the segregation of a dominant schizophrenia susceptibility allele on chromosome 5 in some or all of the families. Family and twin studies of schizophrenia have shown that the spectrum disorders are genetically related and as expected there is an increase in the lod for the DOMSS compared to the DOMS model (3.22 to 4.33). Most surprisingly the highest lod score of 6.49 was found in the DOMSSF analysis at 86% penetrance, which had not been expected from family relative risk studies. This suggested that the schizophrenia susceptibility allele may have alternative phenotypic manifestations and appear to represent a "schizophrenia fringe" group of cases. However caution must be expressed in interpreting this result because only 5 out of the 10 potential fringe cases were informative for linkage analysis and because it is possible that this result may have occurred by chance. The informative fringe phenotypes included two with minor depression, a major depressive, a phobic disorder and one drug abuse disorder. The range of phenotypic expression merits further attention because it may throw light on the genetic or environmental factors which modify or limit the expression of the schizophrenic susceptibility allele, which may have important therapeutic applications.

The relation between the trisomy and these markers was determined from a human hamster cell line containing the deleted chromosome 5 from the transmitting mother (Bassett et al 1988), as its only human complement. P105-153Ra mapped within the trisomy, whilst p105-599Ha mapped proximal to the 5q11.2 breakpoint (Gilliam et al 1989). Therefore a marker showing significant linkage to schizophrenia also mapped within the cytogenetic abnormality causing schizophrenia. Other markers were also mapped within the trisomy, D5S6 (M4), D5S78 (p105-798Rb), DHFR and HEX B (Gilliam et al 1989).

The interpretation of these results in relation to other families must also be made with caution. If valid the locus on chromosome 5 might represent only one cause of schizophrenia. Additional genes on other chromosomes may exist and there may also be non-genetic forms. By the procedure we adapted of selecting families with a large proportion of affected cases (primarily in Iceland) we may have enriched for a particular genetic subtype of schizophrenia, and if Iceland is a genetic isolate then genetic drift may increase the relative frequency of the chromosome 5 schizophrenia susceptibility allele compared to other populations. In addition the use of multiple and high penetrance values may have falsely elevated the lod scores. In other populations the penetrance of any putative chromosome 5 schizophrenia susceptibility alleles may be lower than the value which we have observed in our families. Despite the limitations on the validity of these findings for schizophrenia as a whole the demonstration of a schizophrenia susceptibility locus on chromosome 5 appeared to have provided the first concrete evidence for a genetic basis and possible mode of inheritance for schizophrenia. Moreover the presence of all the traditionally defined subtypes of schizophrenia in these families appeared to show that clinical heterogeneity in schizophrenia can still have a common genetic basis. The general applicability of these findings needs to be studied in other families, since all preliminary linkages need to be confirmed. Synthetic gene frequency maps established by principal components analysis has shown that the Icelandic population is genetically similar to those populations in northern England, north Germany and the south east of Sweden (Menozzi et al 1978), therefore chromosome 5 linkage might be found in schizophrenia families from these populations.

The identification of linkage in schizophrenia allows the study of the relationship between various endophenotypes or vulnerability traits and the underlying genotype. For example brain ventricular size may be a marker for a less genetic form of the illness (Reveley 1984), especially when associated with perinatal complications (Owen et al 1988), the two aetiological factors may operate in a cumulative way (Murray et al 1988) thus providing an example of how penetrance might be affected by environmental factors. The power of studies involving other traits such as eye tracking movements (Holzman et al 1988) and the p300 component of the auditory evoked potential (Blackwood et al 1987) can be increased significantly by testing hypotheses with respect to a genetically defined population, for example a chromosome 5 linked subtype of schizophrenia. If an association is found then this will lead to an enhanced ability for psychiatric

researchers to pursue separate subtypes of schizophrenia in their own right, and might provide an avenue to isolate families which are more likely to show chromosome 5 linkage. Preliminary work showed that the schizophrenics who were partially trisomic for chromosome 5 had eye tracking abnormalities.

Eventually linkage might show that there are truly non-genetic types of schizophrenia. Comparison between the genetic and non-genetic types might lead to a new aetiologic classification and this avenue of research could lead to management and treatment advances even in the absence of further recombinant DNA research to clone and sequence the underlying mutations. Related to this are the factors affecting penetrance. The families studied were selected for high penetrance and the factors affecting penetrance which could be environmental, biological or genetic are likely to be a productive research area in the future with implications in preventative psychiatry. Polygenetic background effects are likely to contribute as are specific environmental factors such as stressful psychological experiences. If clear evidence of environmental precipitants for schizophrenia could be identified acting jointly with the genetic susceptibility this could lead to dramatic and simple methods of prevention.

Finally if this result can be confirmed and much more closely linked markers can be found then it may eventually be possible for psychiatrists to consider genetic counselling in families where chromosome 5 linkage can be reliably established. It should also be possible to identify the 5q locus precisely by applying chromosome specific cloning techniques to isolate and characterize the mutation, allowing new treatments to be developed with a precise knowledge of the disease pathway.

CHAPTER 3: MICROSATELLITE POLYMORPHISMS PROVIDE A HIGHLY POLYMORPHIC MAP FOR CHROMOSOME 5 BANDS q11.2-13.3.

3.1 INTRODUCTION

3.1.1 OTHER LINKAGE STUDIES OF CHROMOSOME 5q11.2-13.3 AND SCHIZOPHRENIA

A study carried out simultaneously by a collaboration between American and Swedish groups failed to find linkage to this area of chromosome 5 (Kennedy et al 1988). They analysed families of a single large pedigree from an isolated northern Swedish population which has a 3% morbid risk for schizophrenia (Book et al 1956). They only analysed those individuals with "core schizophrenia" all other psychiatric disorders such as atypical psychoses or affective disorders with psychotic features were excluded. The pedigree also contained cases of mental retardation. The narrowest analysis DOMS approximates to Kennedy et al's approach and because of this difference it was argued that the results should be compared with caution (Kennedy et al 1989). Nevertheless positive evidence for genetic heterogeneity was demonstrated between the two studies (Kennedy et al 1989). A further four studies have also failed to replicate the positive finding (St Clair et al 1989; Wadleigh et al 1989; McGuffin et al 1990; Kaufmann et al 1989). Two of these studies (St Clair et al 1989 and Wadleigh et al 1989) contained families with mixed psychoses (schizophrenia and bipolar illness). As argued before the addition of other potential psychoses genes in combination with atypical diagnostic features may be sufficient to produce a type II statistical error (false negative). However the same criticisms cannot be made of the other studies. When all the available published data was combined it provided statistical evidence for genetic heterogeneity. However the removal of this study with a positive finding provided no evidence for either heterogeneity or linkage in the remaining families (McGuffin et al 1990). Elston (Cold spring harbor meeting on mental illness 1991) demonstrated that if six major loci for schizophrenia existed then it would be easier to find a new locus than replicate a positive finding.

3.1.2 HYPOTHESES FOR THE LINKAGE HETEROGENEITY

There appear to be two possible hypotheses:

1. Heterogeneity exists and the gene on chromosome 5 accounts for only a small

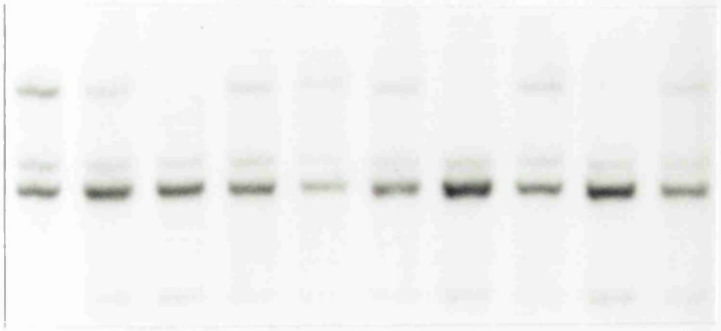
proportion of schizophrenia. This frequency may be elevated in the positive study by genetic drift in the isolated population of Iceland.

2. The positive finding was a type I statistical error (false positive).

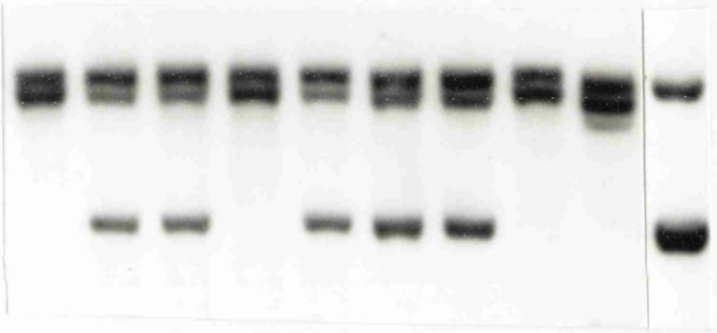
There are two obvious solutions. One is to collect a new cohort of schizophrenic families from Iceland in the hope that genetic drift and family selection strategies have produced the apparent heterogeneity in the linkage results. The second approach would be to increase the genetic information available for linkage analysis in the original cohort and new cohort of families. It is important to note that in the study of the seven schizophrenic families and the two chromosome 5 markers, approximately 50% of the available meioses for linkage analysis were informative. Increasing this proportion and establishing phase in certain pedigrees should allow the true situation to become more lucid. Using the original pedigrees alone in this way would rule out the addition of possible "other genes" for schizophrenia.

Three new RFLP's were analysed for these original seven pedigrees, M4 (D5S6), HEX B and CRI-L1155 (D5S51) to increase the number of informative meioses (Dietzsch et al 1986; Bikker et al 1988; Donis-Keller et al 1987). The Hex B probe pHEXB/0.49X identifies a XmnI RFLP with a PIC value of 0.36 (Figure 1), it produced a two point lod of -2.06 at 10% recombination for the DOMSSF model at 99% penetrance with a phenocopy rate of 0.01 (Table 1). This marker had not been mapped definitively relative to D5S76 and D5S39, nevertheless it did map within the trisomic region associated with schizophrenia (Weiffenbach et al 1991; Gilliam et al 1989). A distance of 30cM could potentially occur between D5S39 and HEX B with Hex B mapping distal to D5S39, however the result indicates that the probable schizophrenia susceptibility allele is unlikely to be distal to D5S39. A second marker CRI-L1155 recognises a Hinc II RFLP at the locus D5S51, this marker with a PIC value of 0.36 provided little genetic information for linkage analysis with a maximum two point lod of 0.84 at 20% recombination (Table 1). This marker position was also not known relative to the D5S76 and D5S39. The third marker M4 identifying a BamHI RFLP (Figure 1) with a PIC value of 0.52 produced a maximum two point lod of 1.33 at 20% recombination with the same model and parameters as for HEX B (Table 1). This marker had been tentatively mapped between D5S39 and D5S76 at a recombination distance of 8cM from D5S39 with odds of 20:1 in favour of this order (Giuffra et al 1989). A multipoint map of these three markers and the DOMSSF model (Figure 2) produced a

A



B



C

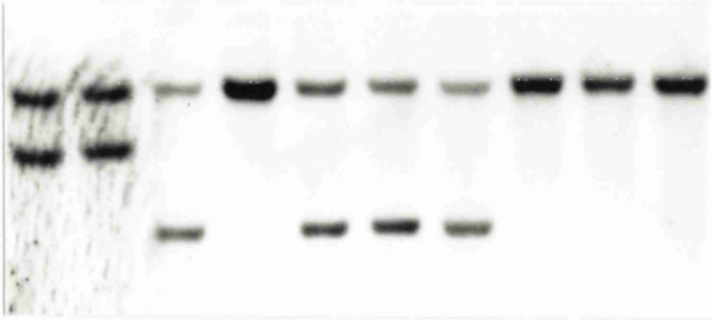


Figure 1. A. pHEXB/0.49X (HEX B) XmnI RFLP with alleles of 7.0 and 2.7 kb with frequencies of .38 and .62 respectively. B. CRI-L1155 (D5S51) HincII RFLP with alleles of 8.8 and 6.2 kb with frequencies of .67 and .33 respectively. C. M4 (D5S6) BamHI RFLP with alleles of 11.0, 9.6 and 7.6 kb with frequencies of .33, .52 and .15.

Theta	HEX B	D5S51	D5S6
0.001	-6.28	-3.34	-8.65
0.01	-4.96	-1.68	-5.03
0.05	-3.17	0.08	-1.39
0.1	-2.06	0.69	0.27
0.2	-0.82	0.84	1.33
0.3	-0.24	0.53	1.18
0.4	-0.03	0.16	0.54

Table 1. Two point lods for HEX B, D5S51 and D5S6 calculated at 99% penetrance for the DOMSSF model. The recombination fraction (Theta) was varied between 0.001 and 0.4 and the lods shown are the summed totals from all the families.

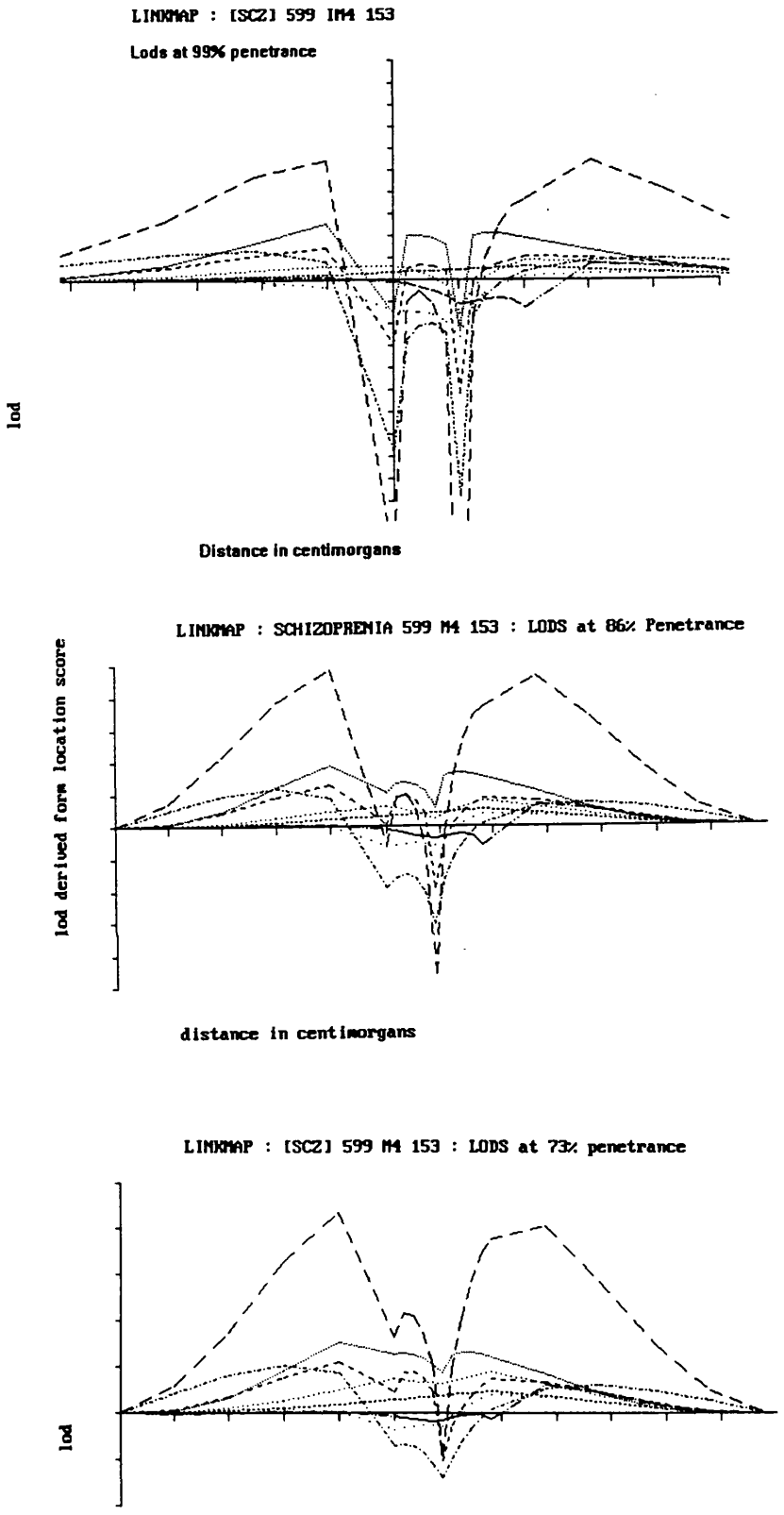


Figure 2. Multipoint LINKMAP of the DOMSSF model at three penetrance values. D5S76 and D5S6 are placed at the troughs whilst D5S39 is 13cM from D5S76.

maximum lod of 5.44 close to D5S76 and a lod of 5.412 close to D5S39 at 99% penetrance. Negative lods were registered either side of D5S6, which became slightly positive at lower penetrances (Figure 2). Inspection of the pedigrees showed that only three meioses which were uninformative for D5S76 and D5S39 were informative for M4 and identified two unaffected recombinants as the only new recombinants in this data. The informative meioses were for F20/16, 17 and 8 with F20/16 and 17 as the recombinants. Clearly these unaffecteds could be regarded as non-penetrant cases. Because of the uncertainty of the map position of the available markers at this time and the inconclusive results which were obtained a different approach was necessary to address the two hypotheses.

3.1.3 MICROSATELLITE POLYMORPHISMS

A major problem with RFLP's for genetic linkage analysis are their limited heterozygosity (Botstein et al 1980). Dimorphic RFLP's have a maximum heterozygosity of 50% but most fall well short of this and consequently in linkage studies a great deal of information is lost. On the other hand VNTR's often have very high PIC values (Nakamura et al 1987; Wong et al 1987) but they tend to be biased in their distribution, located in telomeric bands (Royle et al 1987) or centromeres (Willard et al 1986).

Recently the polymerase chain reaction (PCR; Saiki et al 1988) has permitted the development of new types of DNA markers. PCR allows the specific amplification of DNA sequences from genomic DNA by using short oligonucleotide primers limiting the 5'ends of the two complementary strands. The amplification process is based on repetitive cycling of three simple reactions, it is self contained and fully automated. The first step involves heat denaturation of native double stranded DNA. In the second step, performed at a reduced, primer specific temperature, two short DNA primers are annealed to complementary sequences on the opposite strands of the target DNA, thus defining the ends of the amplified stretch of DNA. The specificity of the PCR derives from the precision of the annealing reaction. The third step of the reaction is the extension of each annealed primer by Taq polymerase to synthesise a complementary second strand of DNA. After extension of the primers the cycle is repeated. All previously synthesised products act as templates for new primer extension reactions in each cycle. The result is an exponential increase in the quantity of DNA synthesised.

These new markers are collectively termed microsatellites or variable simple

sequence motifs, consisting of very simple short tandem repeats. One type of microsatellite consist of polymorphic interspersed repeats such as the dinucleotide (dC-dA).(dG-dT)$_n$ hereafter (CA)$_n$ (Weber and May 1989; Litt and Luty 1989; Smeets et al 1989). Another was short repeated sequences found commonly adjacent to Alu elements such as the trinucleotide (TTA)$_n$ or (TAA)$_n$ (Zaliani et al 1990). A third type was called Alu variable 3'poly deoxyadenylate tracts (Economou et al 1990) consisting of either short tandem repeats or simply 3' poly-adenine tails of Alu elements. Approximately 90% of Alu elements have either a patterned repeat or adenine tail and they are likely to be a rich source of polymorphisms.

Microsatellites like VNTR's vary in length due to the number of repeat units. Consequently alleles can differ by as little as two base pairs for (CA)$_n$ and cannot be separated by conventional southern blot analysis, however combining PCR amplification of a microsatellite and denaturing polyacylamide gel electrophoresis enables its alleles to be resolved. In contrast to VNTR's microsatellites are distributed throughout the human genome, there are approximately 50,000 copies of (CA)n alone where n>10 (Miesfeld et al 1981; Hamada et al 1982). It has been estimated that over 7000 (CA)n will have PIC values in excess of 0.7 (Weber 1990) and should yield a human genetic map of 0.3 to 0.5 resolution.

3.1.4 AIM

To increase the informativeness of D5S76 and D5S39 and other DNA markers mapping to 5q11.2-13.3 (D5S6, HEX B, D5S125, D5S127 and D5S78), by identifying microsatellites in cosmid or lambda genomic clones for these markers. To map these new markers in available reference pedigrees. This would enhance the linkage analysis of the original cohort as well as to maximise information from a second cohort of families.

3.2 METHODS

3.2.1 Cloning microsatellites

A human genomic cosmid (pcosEMBL) and phage libraries (lambda EMBL3, Clontech Laboratories, Inc.) were screened according to standard procedures with the following chromosome 5 probes: L599Ha (D5S76); p105-153Ra (D5S39); M4 (D5S6); pHEX-X. No positive cosmid clones were obtained probably as a consequence of using an over amplified library. However positive phage clones

were plaque purified through three rounds of rescreening. Purified phage DNA was restricted and southern blotted onto Hybond-N, and then the clones were back-hybridised with the original probe. As a final check that they were genuine the Lambda clones were tested for the original RFLP. In addition to the positive phage clones, three cosmid clones from chromosome 5, YN5.132 (D5S127), EF5.15 (D5S125), MC5.60, and the plasmid clone p105-798Rb known to map to 5q11.2-13.3 were screened for the presence of dinucleotide repeats using random primer radiolabelled poly(dC-dA).(dG-dT) and poly(dC-dT).(dG-dA) (Pharmacia). Hybridizations were performed overnight at 65°C in 0.9M NaCl, 1% SDS, 40ug/ml tRNA. Post hybridization washes were at 65°C in 1X SSC. The genomic clones were cut with AluI, HaeIII and Sau3A and southern blotted onto Hybond-N. Positive clones gave a strong signal after 2 hours exposure at room temperature. DNA from each positive clone was cut with either AluI, HaeIII or Sau3A and shotgun cloned into the vector M13 mp18 (Boehringer). Which enzyme that was used for each genomic clone depended on having a insert size between 100-400bp. Plaque lifts (on Hybond-N, Amersham) of these subclones were screened as above and positive subclones were sequenced by the dideoxy method using sequenase TM (US Biochemical). From the sequence data PCR primers were synthesised .

3.2.2 PCR amplification of microsatellites

Human DNA was isolated from nucleated blood cells according to standard procedures (Maniatis et al 1982). 20 to 50ng of this genomic DNA was amplified in a total volume of 25ul containing 25pM of each primer, 200uM each of dGTP, dCTP, dTTP and 25uM dATP (Pharmacia), 1.5mM MgCl2, 20mM Tris-HCl pH8.3, 50mM KCl, 0.1% w/v gelatin and 1ul 35S-dATP (NEN 034-S) and 0.75 units AmpliTaq (Cetus). 35 cycles of amplification were performed using a Perkin Elmer DNA thermocycler. Each cycle consisted of 30s at 94°C, 30s at the primer annealing temperature (Table 2) and 30s at 72°C. The final elongation step was extended to 10 min. The amplified product was extracted once with chloroform and 2ul was mixed with an equal volume of sequencing loading buffer and boiled for one minute. The alleles were separated by electrophoresis in 6% denaturing polyacrylamide DNA sequencing gels. The gels were fixed, dried and exposed to high sensitivity X-ray films (GRI Ltd). The PCR alleles for each block were sized by comparison to a M13 dideoxy sequencing ladder. The concentration of the magnesium ions were varied in the PCR to optimise the results.

3.2.3 Linkage analysis

Genotypings were recorded in 18 reference pedigrees. Two-point and multipoint linkage analysis was carried out using the LINKAGE package (version 5.03; Lathrop et al 1984), in order to localize and map the new microsatellite polymorphisms.

3.3 RESULTS

A total of ten microsatellites were sequenced and unique primers produced for their amplification (Figure 3). The primers were chosen such that they were close to the microsatellite and each pair had similar melting temperatures (TM). The TM's were calculated from the following equation:

$$TM = 4(G+C) + 2(A+T)$$

Though not accurated this equation gives a good estimate of the actual TM.

Two microsatellites were sequenced from the lambda clones for L599Ha (D5S76) one of which was present in the original plasmid clone. These were called p599$(CA)_n$ and Lambda599$(CA)_n$ and they consisted of uninterrupted runs of $(CA)_{11}$ and $(CA)_{18}$ respectively. Two microsatellites were also sequenced from the lambda clones for p105-153Ra (D5S39) and were identified as 6741$(CA)_n$ and 6741$(GT)_n$ consisting of $(CA)_7$ and $(GT)_{20}$, respectively. Sequence close to 6741(GT)n was not obtained because of a highly compressed region at one end of the repeat. One primer was chosen outside this region for the PCR resulting in a large PCR product. The lambda clones for pHEX-X contained only one complex microsatellite with sequence repeat $(GA)_9(GT)_{21}$ and called HEXB$(GA/GT)_n$. A further four microsatellites were sequenced from two chromosome 5q11-q13 cosmids. YN5.132 (D5S127) revealed a dinucleotide repeat of $(CT)_{19}$ called YN$(CT)_n$. EF5.15 (D5S125) contained three dinucleotide repeats, $(CA)_{30}$, $(CA)_{28}$ and $(TG)_7TA(TG)_4(AG)_6AC(AG)_4AC(AG)_2AC(AG)_4$. These were identified as EF1$(CA)_n$, EF2$(CA)_n$ and EF$(TG/AG)_n$ respectively. The plasmid p105-798Rb was also found to have a dinucleotide repeat, labelled p798$(GT)_n$ consisting of $(GT)_3AT(GT)_2(AT)_3(GT)_9AT$. The sequence and the position of the primers for all of the microsatellites are given in Figure 3. No dinucleotide repeats were found for genomic lambda clones for M4 or for the cosmid clone MC5.60.

PCR amplification and denaturing polyacryamide gel electrophoresis revealed that seven of the ten microsatellites had length polymorphisms, p599$(CA)_n$, Lambda599$(CA)_n$, 6741$(CA)_n$, 6741$(GT)_n$, HEXB$(GA/GT)_n$, YN$(CT)_n$ and EF$(TG/AG)_n$ (Figure 4 and Table 2). The annealing temperatures (Table 2) were chosen arbitrarily 5°C below the calculated TM. The polymorphic microsatellites (excluding HEXB$(GA/GT)_n$) were analysed in 18 Icelandic and British pedigrees. Linkage

p599(CA)$_n$

GATCTAAAACAGTAAAGTACATTTAGGAATTGCAACT**5'AATTCAGTACTGC**
TGAAG3'GATTGTGTGTGTGTGTGTGTGTGCGTGCCAGTGCGCACCTACG
TGTATGTGTGTTGGTGAAGGCAAAGTAGGAATGGCAATGAGTAGAGAGGGTA
3'TTCCGTTTCATCCTTACC5'
TAACAAGTAATGACGATGGAGATGTAAACAAGGGTAAGATC

Lambda 599(CA)$_n$

GGGTTGATAGGTACAGCAAACCACCATGGCGCATGTATACCTATGTAACGAA
CCCGCATGTTCTGCACATGTATCCCAGAACTTAAATTAAAAAAAATTTGTAA
ATAACAAAAAAAAAAAAAAAGAAATTC**5'CAGTCCTCGTGGAATCATGC3'**A
TTCTAGTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTTGGAAAGG
3'CC
AGCAGTAAATAAGTGCAAATA
TCGTCATTTATTCACGTTTAT5'

Figure 3. DNA sequence for the ten cloned microsatellites. PCR primers are marked in bold, the foreward primers within the sequence and the reverse primers below their complimentary sequence. (**********) indicates a region of unresolvable DNA sequence, probably as a consequence of DNA secondary structure.

Figure 3 continued.

6741(GT)$_n$ and 6741(CA)$_n$

CATGTCATGCAGGTCTA**5'GTCTTGAACTCCTGAGC3'**TCA(**********
**********)AGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG
TGTGTGTAAGGTCTCCTATTGATTCTGTTTCT**5'CTGGAGAACCCTAATACA**
3'GACCTCTTGGGATTATGT
ATGG3'ACATCCTGTCTCTTCTAAGAAAACGGTGAGCCACATGCACCCCACA
TACC5'
ACACACACACACACCTTCCTTTGGTTTCCAGTTGTCTGGCATTTACCTTGGA
3'CAGACCGTAAATGGAACCT

GCAGAG
CG5'

HEXB(GA/GT)$_n$

GCCCACTTTTGTACTGGATTATTTTTATTCTATTTAT**5'TGTTTTGTGGGAA**
TTCCATATGTC3'TATGTATGTATTGTGAGAGAGAGAGAGAGAGAGTGTGTG
TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTTGGGTTTATTCTCCTGG
3'CCCAAATAAGAGGACC
TTCATAGCTTCTTTCCCGTTGGTGGGATGCCTACCGCAAGCAGGACTCTAGA
AAGTATC5'
GGATCCCC

p798(GT)$_n$

ATAAGAGATGTGTGC**5'TGAGAAAACGCCATTCGAGG3'**CAGGTCTCTCGTG
CCACAAGCCTGGTGTGTATTGTGTATATATGTGTGTGTGTGTGTGTATGA
3'CT
TTGTGGGGAGGAAAGTACAAATCTATGCGTTAGAAAGAGTAGCGTAAGAGATC
AACACCCCTCCTTTCATG5'

Figure 3 continued.

YN(CT)$_n$

AATTAGGTTTAGATGCAGTTATGAGGGACTA5'GTACCCTTACAGGAAGAGC
CAAG3'ATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTGCCATG
TGAGGAACAACAAGAAGGTAGCCATCTGCAACCCAGGAATAAGG
3'CTTCCATCGGTAGACGTTGGGTC5'

EF1(CA)$_n$

AG5'CTCTCAAACACGGGCATTATTT3'TCCACACACACACACACACACA
CACACACACACACACACACACACACACACACACACACACACCCTCTGTATTTTA
GGACTCTAGAGGATCCCCGGGTACCGAG
CCTGAGATCTCCTAGGGGCCCAT5'

EF2(CA)$_n$

TTTTATTCAGTCATAAAAGGTGAAATTAT5'GCCAGTTGCAAGACAATGAAT
G3'AAGTTGGACCTTTTTTTTACAAAAAAGCCAGAATCAGATAAGAAACAGT
TGGTGTTATATTCGAACATACAACACACACACACACACACACACACACAC
ACACACACACACACACACACACGACATGAAAGTAGAAAGGAGCCATCCTGAA
3'CTTTCATCTTTCCTCGGTAGGA5'
TGGGTCT

EF(TG/AG)$_n$

GGCTTTTGTACACGTTCTGAGTGACTCTTATGAGCCAGAAATTT5'TAGAAA
ATACTACAGTCTTCTTG3'GGAGTGGTGTGTGTGTTTGTGTGTGTGTGTGTA
TGTGTGTGAGAGAGAGAGAGACAGAGAGAGACAGAGACAGAGAGAGAAATAG
GTGAGAGAGTTACAACAAGCATGGGAGTCCTCTTTTGAAAAGTGTCTGCTCA
3'GTTGTTCGTACCCTCAGGAGAA5'
TGTCCTTTGCCTACTTTTTAATGGTTTTTTTCCTTGTAAA

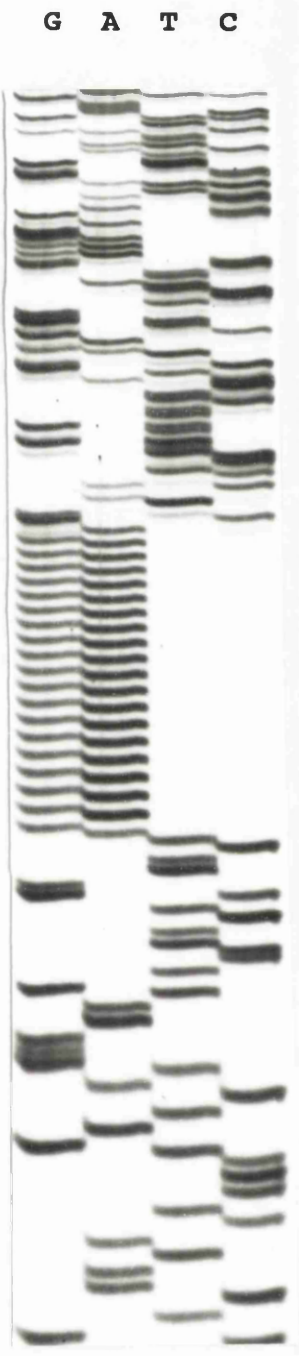


Figure 3 continued. DNA sequencing gel for the perfect repeat of $YN(CT)_n$. The sequencing reactions were electrophoresed on a 6% denaturing polyacrylamide gel.

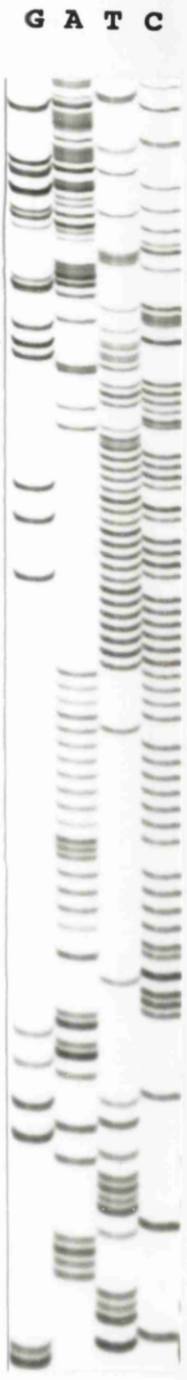


Figure 3 continued. DNA sequencing gel results for the complex repeatitive element of EF(TG/AG)$_n$.

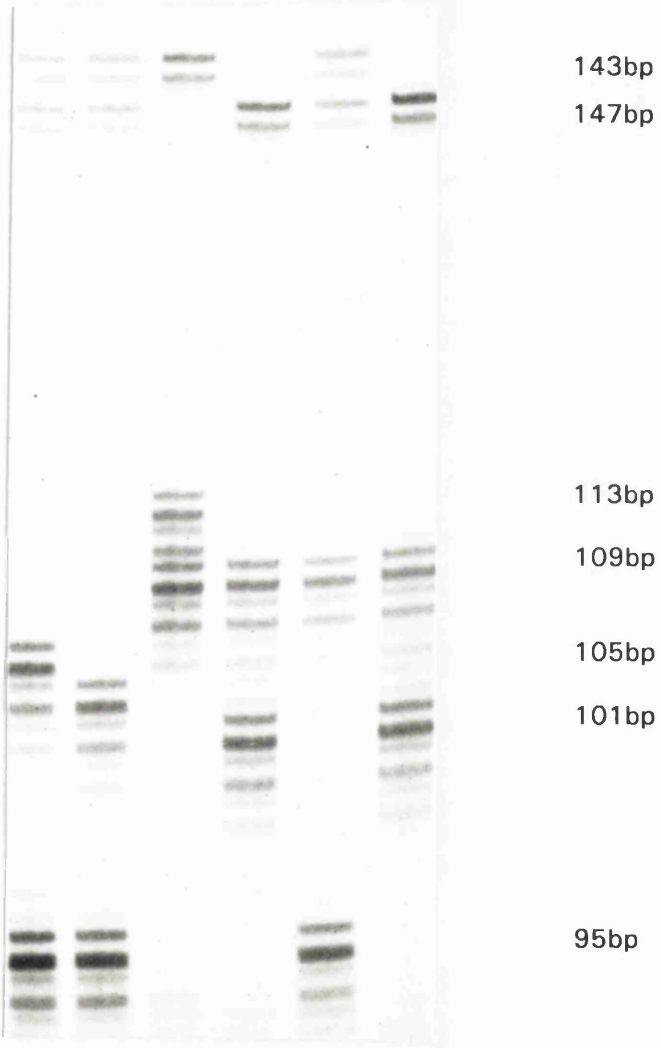
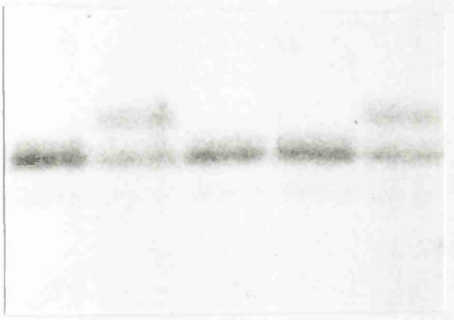


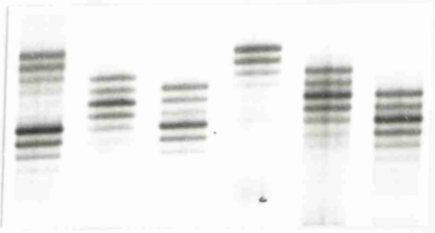
Figure 4. Microsatellite alleles for EF(TG/AG)$_n$ and YN(CT)$_n$. Each lane represents one individual with EF(TG/AG)$_n$ alleles of 143 and 147bp, and YN(CT)$_n$ alleles between 95 to 113bp. The two microsatellites were amplified simutaneously using an annealing temperature of 62°C and the alleles were resolved on a 6% polyacrylamide sequencing gel.

6741(CA)$_n$



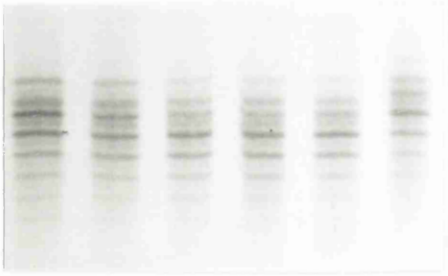
122bp
124bp

6741(GT)$_n$



224bp
218bp
212bp

HEX(GA/GT)$_n$



122bp
124bp

Figure 4 continued. Microsatellite alleles for 6741(CA)$_n$, 6741(GT)$_n$ and HEX(GA/GT)$_n$. Alleles sizes are given in base pairs (bp) of DNA. Figures of the microsatellite alleles for D5S76 are shown in chapter 4.

	Annealing temp	Allele sizes	Het
p599(CA)$_n$	47°C	97-109bp	0.70
Lambda599(CA)$_n$	57°C	94-110bp	0.70
6741(GT)$_n$	60°C	212-220bp	0.88
6741(CA)$_n$	60°C	122-124bp	0.47
HEX(GA/GT)$_n$	60°C	122-124bp	0.40
p798(GT)$_n$	55°C	102bp	0.00
YN(CT)$_n$	62°C	96-114bp	0.96
EF(GA/GT)$_n$	62°C	143-147bp	0.50
EF1(CA)$_n$	63°C	121bp	0.00
EF2(CA)$_n$	59°C	132bp	0.00

Table 2. Annealing temperatures for PCR of the ten cloned microsatellites and their resulting allele sizes and heterozygosity values.

	Allele size (bp)	Allele frequency	PIC
6741(GT)$_n$			
1	220	0.29	0.74
2	218	0.25	
3	216	0.22	
4	214	0.12	
5	212	0.12	
6741(CA)$_n$			
1	124	0.57	0.37
2	122	0.43	
YN(CT)$_n$			
1	113	0.148	0.84
2	109	0.130	
3	107	0.130	
4	105	0.204	
5	103	0.037	
6	101	0.037	
7	99	0.037	
8	97	0.111	
9	95	0.166	
EF(TG/AG)$_n$			
1	143	0.500	0.38
2	147	0.500	
HEX(GA/GT)$_n$			
1	122	0.167	0.24
2	124	0.833	

Table 2 (continued). Allele sizes and frequencies for 6741(GT)$_n$, 6741(CA)$_n$, YN(CT)$_n$, EF(TG/AG)$_n$ and HEX(GA/GT)$_n$. Details of the D5S76 microsatellites are given in chapter 4.(analysis of 100 chromosomes).

analysis was performed with the LINKAGE programs (Lathrop et al 1985). This confirmed that there were no recombinants between 6741(CA)$_n$ and 6741(GT)$_n$ at D5S39 (maximum lods 14.8 at theta = 0.0). Yn(CT)$_n$ and EF(TG/AG)$_n$ were linked to each other (maximum lod 12.2 at theta = 0.05) and were also linked to the other microsatellites and to the BamHI RFLP detected by M4 at D5S6 (see Table 3 for two point lods between the markers). In order to localise these microsatellites more precisely a multipoint analysis was preformed using LINKMAP (overlapping 4 points) against p599(CA)$_n$, M4, 6741(GT)n, and p105-798Rb positioned according to Weissenbach et al (1991). For YN(CT)$_n$ this produced a maximum lod of 34.3 at a distance of 1 centimorgan from 6741(GT)$_n$ and for EF(TG/AG)$_n$ a maximum lod of 17.7 at zero recombination with M4, however in neither case was it possible to establish an unambiguous order (Figure 5). Instead, published RFLP data from the CEPH 4 database was used to map the MspI RFLP detected by YN5.132 at D5S127 and the TaqI RFLP detected by EF5.15 at D5S125 against L599Ha (D5S76), M4 (D5S6), p105-153Ra (D5S39) and p105-798Rb (D5S78). This analysis produced the order: D5S76-D5S6-D5S125-D5S39-D5S127-D5S78 (Figure 6). The most likely position for D5S125 was 2cM distal to D5S6 and the position for D5S127 was 9cM distal to D5S39. Odds in favour of this order over the next most likely order are 10000:1 for D5S125 and 1000:1 for D5S127.

3.4 DISCUSSION

Just over ten years ago the goal of human genetics was to construct a linkage map using RFLP's (Botstein et al, 1980). Since then preliminary linkage maps for many human chromosomes have been published (Donis-Keller et al, 1987). These maps have localized many classical human genetic disorders by linkage analysis, such as Duchenne muscular dystrophy, Huntingtons chorea and cystic fibrosis, which has brought about the isolation of some of the relevant disease mutations (Rommens et al, 1989).

These encouraging results initiated linkage searches for disease genes for more genetically complex disorders. Psychiatric disorders have caused a great deal of controversy. A gene on chromosome 11 for manic depression in the old order Amish seemed to show tight linkage to the gene for insulin (Egeland et al 1988). On expansion of the pedigree the strong evidence disappeared, though the possibility of genetic heterogeneity could not be ruled out altogether (Kelsoe et al 1989). Another locus for manic depression seems to be on the X chromosome but only accounting for a proportion of the disease (Baron et al 1988). A gene for

EF(TG/AG)$_n$

Theta	0.00	0.01	0.05	0.10	0.20	0.30	0.40
6741(GT)	-99.00	14.31	14.70	13.41	9.78	5.82	2.14
M4	6.54	6.39	5.79	5.05	3.55	2.00	0.65
p599(CA)	10.95	10.92	10.37	9.29	6.68	3.90	1.38
YN(CT)	-99.00	10.21	12.19	11.61	8.67	5.16	1.87

YN(CT)$_n$

Theta	0.00	0.01	0.05	0.10	0.20	0.30	0.40
p599(CA)	-99.00	16.17	21.28	21.27	17.36	11.30	4.44
M4	-99.00	7.27	11.76	12.12	9.80	6.08	2.21
6741(GT)	-99.00	33.28	33.88	31.53	24.59	16.00	6.52
p105-798Rb	-99.00	-0.72	4.92	6.13	5.34	3.21	0.99

Table 3. MLINK two point lods for EF(TG/AG)$_n$ and YN(CT)$_n$ and other chromosome 5 markers.(For details of p105-798Rb see chapter 6).

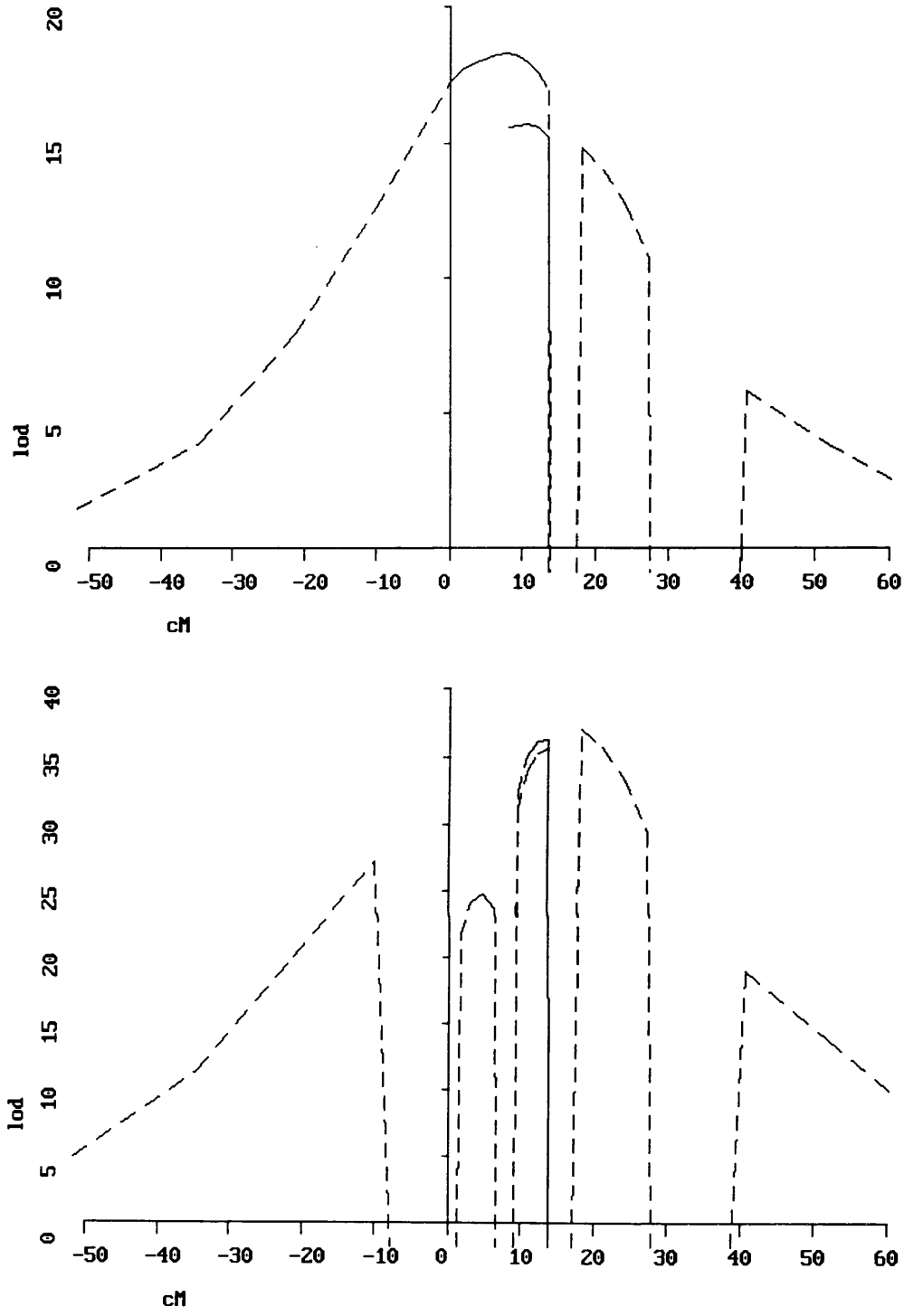


Figure 5. LINKMAP analyses of EF(TG/AG)$_n$ (above) and YN(CT)n (below) in the schizophrenic pedigrees. D5S76 (p599(CA)n) is at zero, D5S6 (M4) is at 8cM, D5S39 (6741(GT)$_n$ is at 15cM and D5S78 (p105-798Rb) is at 31cM.

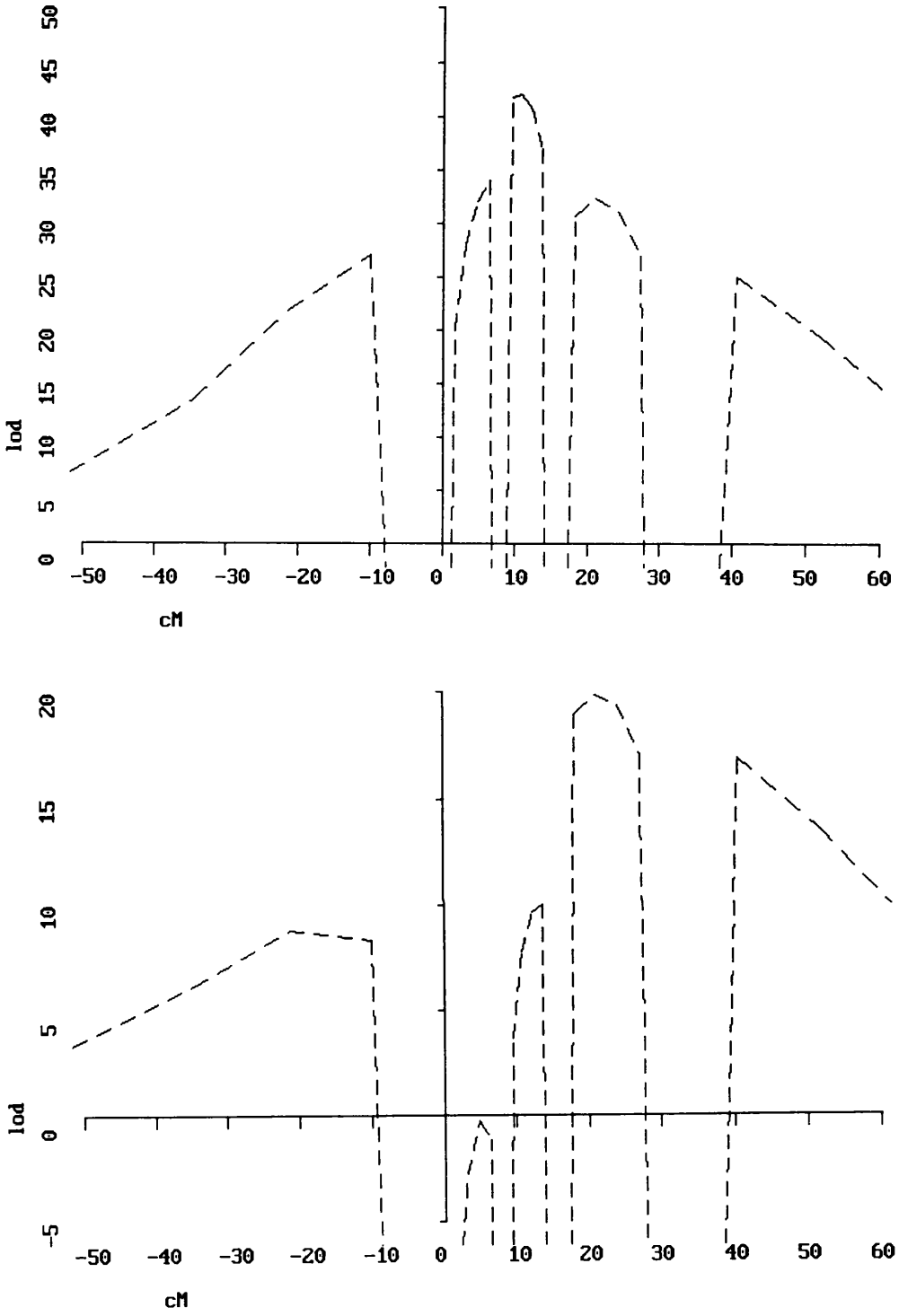


Figure 6. LINKMAP analyses for EF5.15 TaqI RFLP (above) and YN5.132 MspI RFLP (below) in published CEPH data (vs 4). The position of D5S76, D5S6, D5S39, and D5S78 was the same as Figure 5.

Alzheimers disease has been localized to chromosome 21 (St George-Hyslop et al 1987; Goate et al 1989). Although there is good evidence for linkage in early onset families, there is also a suggestion of non allelic heterogeneity and apparently no evidence for linkage in late age onset families (St George-Hyslop et al 1990). Useful pedigree for linkage in Alzheimers are difficult to find because of the late age of onset.

It is clear that there are certain practical considerations that limit the application of linkage mapping in human disorders. These seem to be especially true for common diseases with a genetic component. Highly polymorphic markers would ameliorate some of these potential difficulties. Highly polymorphic microsatellites would increase the number of informative matings, and help establish phase which would have the knock on effect of increasing the power of heterogeneity tests (Ott 1986).

All of the lambda, cosmid and plasmid clones studied from chromosome 5 contained a dinucleotide repeat except the cosmid clone MC5.60 and the lambda clones for M4. This is in keeping with the estimate by Litt and Luty (1989) that one in three cosmids would contain a dinucleotide repeat of the $(CA)_n$ type. Obviously this figure will increase by the searching for $(CT)_n$ type as found in the cosmid YN5.132. Most of the dinucleotide repeats were of the perfect type (uninterrupted runs of the dinucleotide; Weber 1990) only three had a more complex pattern of repeats, this is also in line with the estimates of Weber (1990). Weber also suggested a relationship between the length of a perfect dinucleotide repeat and the informativeness of the marker. The longer the repeat the greater the PIC value (Botstein et al 1980). Though this relationship is probably true, the informativeness of p599$(CA)_n$, Lambda599$(CA)_n$, 6741$(CA)_n$, 6741$(GT)_n$ and YN$(CT)_n$ seem to follow this rule of thumb, there do however appear to be important exceptions. In our study two perfect dinucleotide repeats were not informative, EF$(CA)_{30}$ and EF$(CA)_{28}$, even though they were the two longest sequenced. In general the informativeness of the complex repeats is less than that of the perfect repeats (Weber 1990) even when length is taken into account. Our results seem to follow this rule, EF$(TG.AG)_n$ and HEX$(GA.GT)_n$ were both polymorphic but only exhibited diallelic systems whilst p798$(GT)_n$ was monomorphic even though its repeated pattern was of considerable length.

The faint bands associated with a microsatellite allele have been suggested to be

caused by strand slippage by Taq polymerase during the amplification (Luty et al 1990) or in combination with somatic mutations (Smeets et al 1989). The latter observation is based on a greater intensity of these bands when comparing amplified genomic DNA to amplifying a clone of the same DNA. It is of interest to note that the non polymorphic microsatellites had no faint bands associated with the major band. This suggests that the flanking sequences of the microsatellite may be important in causing strand slippage by DNA polymerase during replication. This may be important in producing the polymorphisms in vivo with the same phenomenon occuring in vitro. When more microsatellites have been cloned this phenomenon can be studied in greater detail. It can also be seen that the intensity of the associated bands varies between microsatellites, and is not dependant on the batch of enzyme. In our study it was difficult to distinguish the alleles of the HEX(GA.GT)$_n$ so we excluded this marker from the linkage map. Similar results were also found for a microsatellite reported by Dean et al (1990). The alleles for the HEX B microsatellite may be clearer if one of the primers is end radiolabelled rather than radioactivity being incorporated during the amplification.

It is clear that increasing the informativeness of markers in an area of interest by locus expansion and searching for polymorphic microsatellites is a much more powerful method than searching for new RFLP's. It is likely too that very many RFLP's would have been required to reach the high PIC values obtained with our microsatellites. The wide distribution (Luty et al 1990) and the large number available throughout the human genome make these type of marker the preferred choice over both RFLP's and VNTR's, in increasing the informativeness of a locus. The polymerase chain reaction also has benefits over the Southern blot. Tiny amounts of starting material are required (ng compared to ug). Methods to multiplex the markers (Luty et al 1990) will reduce the amount of work. Deceased individuals from pedigrees can be typed if pathological specimens are available (Mankoo et al 1991). This has potential advantages where fatalities could limit the usefulness of certain families, for example late age of onset disorders such as Alzheimers, and childhood death disorders such as the acute type of childhood SMA.

In particular for the chromosome 5 "linked" schizophrenia the three markers Lambda599(CA)$_n$, 6741(GT)$_n$ and YN(CT)$_n$ should make nearly all the available meioses in the original pedigrees informative for at least two markers. They will also establish phase for families where either one or both parents are missing. This

alone should provide enough information to confirm or refute the original localization, without the use of new pedigrees where arguments of heterogeneity cannot be ruled out altogether (Kelsoe et al 1989). In addition the use of the microsatellite EF(TG/AG)$_n$ which map between D5S76 and D5S39 will provide linkage information between the markers used in the original study. The use of all these markers will increase the power of the heterogeneity test. Mapping of the microsatellite markers for D5S39 and D5S125 with the RFLP for D5S6 in 40 CEPH reference pedigrees produced the same order as reported here, with D5S125 mapping distal to D5S6 with a recombination value of 0.02 and D5S125 mapping proximal to D5S39 with a recombination value of 0.08 (Morrison et al 1992).

Another disorder has recently been mapped to this region of chromosome 5, namely Spinal Muscular Atrophy (SMA; Brzustowicz et al 1990; Gilliam et al 1990).The most likely map position for the chronic childhood SMA gene is between D5S39 and D5S6 whilst the acute form either maps to the same locus, ie shows probable allelic heterogeneity , or lies telomeric to D5S39. The microsatellite markers for D5S39 and YN(CT)$_n$ should provide the linkage information to define this position accurately, and give greater evidence for genetic homogeneity between the two types of SMA (Daniels et al 1992a; Morrison et al 1992). Furthermore these new microsatellite markers will be useful in perinatal diagnosis of SMA (Daniels et al 1992b).

In general the novel microsatellite type of markers will provide the genetic information needed to map disorders with complex patterns of inheritance or genetic heterogeneity, either by linkage or using the sib-pair approach where no mode of inheritance is assumed (Risch 1990). A question remains as to whether microsatellites can be used to find allelic association (linkage disequilibrium) with disease mutations as was necessary for the successful cloning of the cystic fibrosis gene. Fortunately the data collected offered a unique chance to answer this question because two highly polymorphic microsatellites were found in a single clone.

CHAPTER 4: LINKAGE DISEQUILIBRIUM BETWEEN TWO HIGHLY POLYMORPHIC MICROSATELLITES.

4.1 INTRODUCTION

In common with variable number of tandem repeat polymorphisms (VNTR's, Nakamura et al. 1987) microsatellites vary in length due to the number of repeat units. There has been some speculation concerning the mutation rate for the generation of new microsatellite lengths, if the rate were as high as for VNTR's it might be possible to identify new mutations developing within families (Jeffreys et al 1988). It has also been demonstrated that increased informativeness or heterozygosity as expressed in the polymorphism information content (PIC) of microsatellite sequence polymorphism is positively correlated with length (Weber 1990). Consequently the mutation rate might be higher as the microsatellite length increases. The relationship between the number of alleles, length of microsatellite and mutation rate is compounded by the fact that a new mutation is as likely to produce a new allele of identical length to a pre-exisiting allele as to a new length variation. The mutation rate therefore may be higher than expected by simply counting the number of alleles present. This effect would tend to weaken any chance of observing linkage disequilibrium with microsatellite polymorphisms.

Because of these considerations the general application of these markers in tests of association and linkage disequilibrum in mapping genetic diseases ought to be tested. If the microsatellite sequences are hypothesised to be stable during evolution then we should expect to find linkage disequilibrium between two adjacent highly polymorphic microsatellites in the same genomic clone.

4.1.1 AIM

The degree of linkage disequilibrium was to be investigated between the two highly polymorphic microsatellites cloned for D5S76.

4.2 METHODS AND RESULTS

The cloning and amplification of p599$(CA)_n$ and Lambda599$(CA)_n$ is described in chapter 3. A third microsatellite of the polydeoxyadenylate tract type consisting of 15 repeated adenine residues was found to be non polymorphic in ten unrelated individuals.

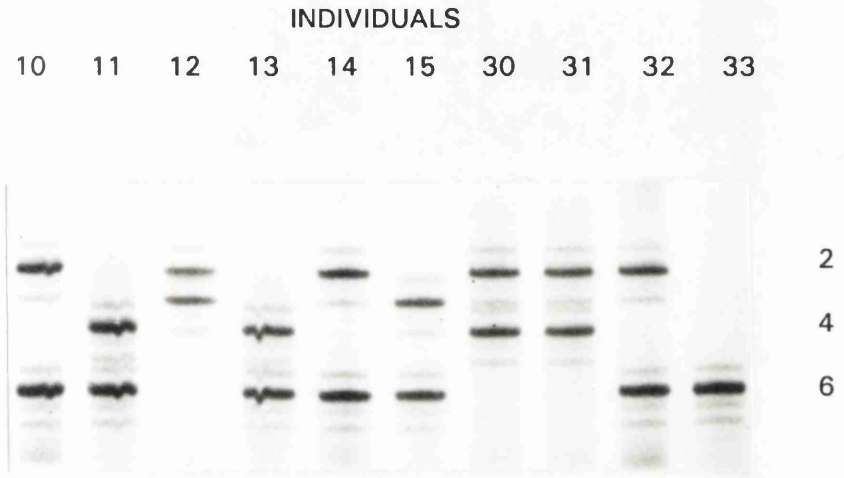
Genomic DNA from 46 unrelated individuals was amplified and typed for both of

the microsatellite repeats. Six alleles (Figure 1) were observed for p599$(CA)_n$ ranging upwards in size from $(CA)_{11}$ to $(CA)_{17}$. The frequency and actual sizes of the alleles in base pairs are given in Table 1. The PIC value (Botstein et al 1980) for this microsatellite was 0.71. A total of eight alleles were observed for lambda 599$(CA)_n$ with a PIC value of 0.74 (Figure 1). The sizes and frequencies of these alleles are also shown in Table 1. Restriction maps for the lambda clones at p105-599Ha were determined and the relative positions of the two CA dinucleotide repeats were identified by hybridization with poly (dC-dA).(dG-dT) and one of the PCR primers. The two microsatellites were within 7.7kb (+/- 1.5kb) of each other (Figure 2).

The resulting genotypes (Table 2) were analysed with the computer program ASSOC (Ott 1985b) which demonstrated strong linkage disequilibrium between the two microsatellites. There was no evidence for any of the observed linkage disequilibrium to have been caused other than by allelic association (chi-squared = 31, 665 df, NS). The evidence for allelic association (chi-squared = 118, 35 df, $p<0.001$), was investigated further in order to overcome the problem of small expected values in some of the cells of the contingency table. Unrelated individuals from a number of multigeneration families were typed so that phase could be determined. In almost all cases it was possible to determine both haplotypes from the parental genotypes, although in a few cases only one haplotype could be determined unambiguously. A total of 86 haplotypes were assigned in this way, and they included the same alleles as had been detected in the original sample. Haplotype frequencies are presented in Table 3. This table was collapsed to a smaller 4x3 contingency table by amalgamating rows and columns with small totals until all expected cell frequencies were greater than 5. A standard chi-squared test was then performed which yielded highly significant evidence in favour of linkage disequilibrium (chi-squared = 67, 9 df, $p<0.001$). It is interesting to note that there is also a positive correlation between the sizes of the two alleles in each haplotype (Kendall's correlation coefficient: $r=0.42$, $p<0.001$).

The nucleotide sequences in the immediate vicinity of the two microsatellites (approximately 380bp each) were analysed using the University of Wisconsin genetic group computer package. Direct or inverted repeats and sequences suspected of being involved in gene conversion events (Kenter and Birshtein.1981) were not found for either microsatellite. A proportion of the surrounding sequence of the Lambda599$(CA)_n$ was 80% homologous to the human KpnI repeat.

p599(CA)$_n$



Lambda599(CA)$_n$

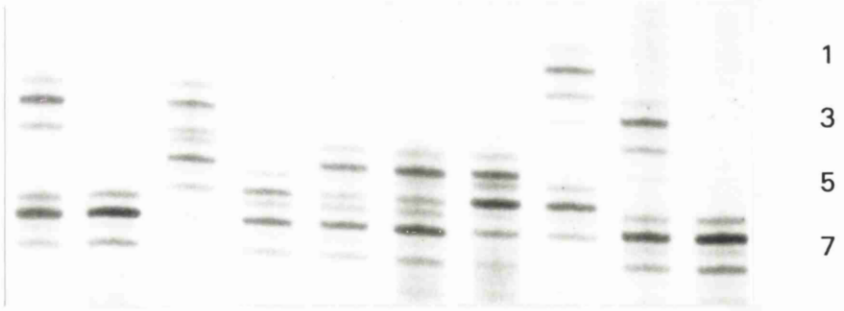


Figure 1. Microsatellite polymorphisms for p599(CA)$_n$ and Lambda599(CA)$_n$ for ten unrelated individuals (10-33). The corresponding alleles are given for each microsatellite. The alleles were resolved in a 6% polyacryamide sequencing gel.

p599(CA)$_n$

Primers 5'ATTCAGTACTGCTGAAGG 3'

5'CCATTCCTACTTTGCCTT 3'

Annealing temp. 47^{0}C

No. Alleles	Allele frequency	PCR product size(bp)
1 CA_{17}	0.021	109
2 CA_{15}	0.213	105
3 CA_{14}	0.213	103
4 CA_{13}	0.127	101
5 CA_{12}	0.064	99
6 CA_{11}	0.362	97

Lambda599(CA)$_n$

Primers 5'CAGTCCTCGTGGAATCATGC 3'

5'TATTTGCACTTATTTACTGCTCC 3'

Annealing temp. 57^{0}C

No. Alleles	Allele frequency	PCR product size(bp)
1 CA_{26}	0.011	110
2 CA_{25}	0.043	108
3 CA_{24}	0.098	106
4 CA_{23}	0.065	104
5 CA_{22}	0.228	102
6 CA_{21}	0.174	100
7 CA_{20}	0.358	98
8 CA_{18}	0.022	94

Table 1. Nucleotide sequences of oligonucleotide primers, number of repeats for each allele, PCR product size and allele frequencies.

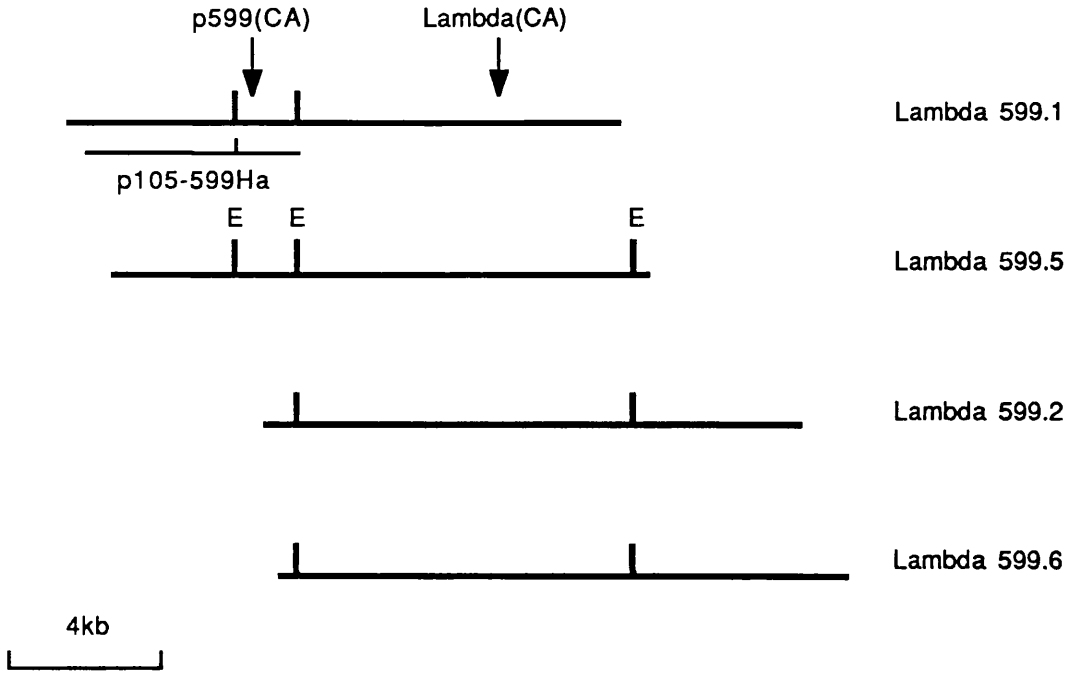


Figure 2. An EcoRI (E) restriction map of the genomic lambda clones for p105-599Ha. The position of the original plasmid clone and the polymorphic microsatellites are shown.

Raw data	Alleles	
Individuals	p599(CA)$_n$	Lambda 599(CA)$_n$
1	3 3	5 4
2	6 6	7 7
3	2 2	2 2
4	2 6	3 7
5	2 6	3 5
6	2 5	3 5
7	4 6	7 7
8	3 6	5 7
9	2 6	7 7
10	2 6	3 7
11	4 6	7 7
12	2 3	3 5
13	4 6	6 7
14	2 6	5 7
15	3 6	5 7
16	1 3	5 8
17	3 6	5 7
18	3 6	6 7
19	2 3	5 6
20	5 6	7 7
21	3 6	5 5
22	4 4	6 6
23	3 6	5 7
24	2 4	3 6
25	3 6	6 7
26	5 6	5 7
27	3 6	7 7
28	2 4	2 6
29	3 6	5 5

Table 2. Genotypes for 46 individuals of the two microsatellites for D5S76.

30	2 4	5 6
31	2 4	1 6
32	2 6	3 7
33	6 6	7 7
34	6 6	5 7
35	2 2	2 4
36	5 5	4 4
37	3 6	6 7
38	6 6	7 7
39	6 6	5 7
40	2 2	4 4
41	3 3	6 6
42	3 3	5 5
43	5 6	3 7
44	3 4	6 7
45	1 2	3 8
46	4 4	6 6

Table 2 continued.

		p599(CA)n						
# $(CA)_n$		17	16	15	14	13	12	11
	26	0	0	0	0	0	0	0
	25	0	0	0	0	0	0	0
	24	0	0	5	0	0	0	0
	23	0	0	15	0	0	7	0
Lambda599(CA)n	22	0	0	1	8	0	0	7
	21	0	0	0	9	14	0	0
	20	0	0	0	1	1	0	17
	19	0	0	0	0	0	0	0
	18	1	0	0	0	0	0	0

Table 3. Haplotype frequencies for the two microsatellites at D5S76.

4.3 DISCUSSION

The conservation of the dinucleotide repeat (CA)n in the genomes of organisms as diverse as yeast to man (Tautz and Renz 1984) has generated a number of ideas on their possible biological role. They may be involved in gene regulation (Hamada et al 1984; Norheim and Rich 1983), DNA replication (Roo et al 1988), chromation structure (Pardue 1987) and recombination where they may be "hot spots" (Kmiec and Holtman 1986; Slightom et al 1980). They are suspected of forming conformational changes such as Z-DNA, recombination tetrads (Treco and Arnheim 1986) and triple stranded structures (Johnston 1988). Several mutation mechanisms are possible in order to create the length variation, unequal crossing over events, gene conversion and strand slippage during DNA replication and repair (Tautz 1986). In M13 bacteriophage the rate of strand slippage increases as the length of the (CA)n increases (Levinson and Gutman 1987). This is consistent with the observation that marker informativeness correlates with the number of repeats (Weber 1990). Zuliani and Hobbs (1990) suggested an association between their repeated polymorphisms and the physical proximity to Alu sequences. Similar results were also demonstrated by Economou et al (1990) with Alu 3' poly deoxyadenylate tracts. It is interesting to note that Alu elements have also been suspected of being involved in recombination events, possible being transcriptionally active and hot spots for gene rearrangements.

For minisatellites with heterozygosity values >97% it was possible to detect new mutations within pedigrees (Jeffreys et al., 1988) with a rate estimated at 0.004 per DNA fragment per gamete. Similar findings occured with DNA fingerprints detected with a (CAC)5 oligonucleotide (Nurnberg et al., 1989), however most microsatellites analysed have PIC values ranging from 0.3 to 0.8 (Weber, 1990) so it is unlikely that new mutations within pedigrees will be a problem. The evidence so far shows that the observed degree of heterozygosity at $(CA)_n$ microsatellite loci can vary even when length is taken into account (Weber 1990). Thus mutation rates must differ between microsatellite loci, or genetic admixture has been more important for some microsatellites than others. This assumes selection has had no role. It is therefore important to establish more extensive data on their stability before their use as diagnostic markers can be validated. One explanation for the finding of less intense bands associated with microsatellite alleles at PCR (see Fig 1) is that they may have occurred by strand slippage during the amplification (Luty et al. 1990) or because of somatic mutations (Smeets et al. 1989). Nevertheless

the ability to demonstrate disequilibrium indicates that they are stable in the germ line and have been stable during evolution.

Non-uniform recombination with the human B-globin gene cluster established a "hot spot" for recombination within 9.1 kb segment of DNA (Chakravarti et al., 1984). Increasing the recombination rate 3-30 fold would have a profound effect on the establishment of association. The disequilibrium established between the two microsatellites in this study has not been overcome by recombination ,it is unlikely that (CA)n repeats can be "hot spots" for recombination as suggested by Kmiec and Holloman (1986) and Slightom et al (1980). In contrast, linkage disequilibrium for the alpha-globin VNTR (Higgs et al 1986) was not demonstrated with other physically linked RFLPs even though this region was a "cold spot" for recombination (Chakravarti et al 1984; Michele and Jenkins 1988; Ramsay and Jenkins 1988). The explanation for this may be an elevated mutation rate for this VNTR, so that disequilibrium cannot be maintained, a situation which is not apparent for some highly polymorphic microsatellites.

A linear relationship between disequilibrium and physical distance may occur in some chromosome regions but not in others. This is because for small regions of DNA, the effects of mutation, genetic drift and population admixture may outweigh those of recombination (Litt and Jorde 1986). As markers are separated by greater distance recombination should eventually become the major effect and the power to establish disequilibrium will be reduced. Thompson et al (1988) demonstrated that very large sample sizes are required to establish negative disequilibrium (i.e. repulsion phase) for diallelic systems, therefore a failure to demonstrate does not imply it's absence. The analysis of multiallelic systems is less straightforward, grouping of alleles can either increase or decrease the power of the test (Weir and Cockerham 1978), when allowing for multiple testing . Generally though, the power will be greater. Potentially multiallelic microsatellites should increase the usefulness of mapping by the supratype method (Morton and Lew 1985) when recombination has the greatest influence, because relatively few individuals are required to establish disequilibrium. This method calculates the maximum likelihood for alternative gene orders for trios of loci. Thus information from microsatellites could prove useful in organising procedures like chromosome walking, jumping and YAC cloning for the identification of disease genes (Rommens et al 1989) before physical maps or maps based on linkage data are available.

Linkage disequilibrium between a marker and disease gene encourages the search for candidate genes in the near vicinity of the marker (Estivill et al 1987; Dean et al 1990; Hanover et al 1990; Ikonen et al 1990). For example a highly polymorphic microsatellite has increased the power to establish disequilibrium between the locus D9S15 and the Friedreich's Ataxia gene (Fujita et al 1990). If these findings are generally applicable then microsatellites will prove to be powerful new tools in mapping human genetic diseases by linkage analysis and help isolate the genes by increasing the power of disequilibrium tests. In another study linkage disequilibrium was demonstrated between a microsatellite in the CFTR gene and RFLP's at that locus (Morral et al 1991). Haplotypes in this study provided evidence that the microsatellite polymorphisms are produced by strand slippage during replication rather than unequal cross-over. Microsatellites will also increase the power of association tests for genetic diseases where family data may be difficult to obtain or where a complex mode of inheritance is hypothesised. Population geneticists now have a new set of tools for the study of inter and intra population variations, migration patterns and gene flow between populations (Ramsay and Jenkins 1988).

The finding of a positive length correlation between the sizes of the two microsatellite sequences is not easy to explain. It could be postulated that the length relationship is important for chromatid folding, or gene expression possibly by a stem-loop structure (Hamada et al 1984; Vogt 1990). Consequently selection may be preserving the apparent disequilibrium. However no evidence for this was seen with ASSOC, and the correlation may be coincidental.

CHAPTER 5: A COSMID CLONE FOR THE 5HT1a RECEPTOR REVEALS A TaqI RFLP WHICH SHOWS TIGHT LINKAGE TO DNA LOCI D5S6, D5S39, AND D5S76.

5.1 INTRODUCTION

5.1.1 THE 5HT RECEPTORS

A gene encoding the 5HT1a neuroreceptor on chromosome 5 q11.2-13.3 was a potential candidate gene for schizophrenia. It is one of a series of 5HT related receptors that are potentially involved in either the cause or treatment of schizophrenia and other psychotic disorders. The effects of 5-Hydroxytryptamine are mediated by the activation of distinct receptor subtypes. Based on pharmacological properties six 5HT receptor subtypes have been defined, 5HT1a, 5HT1b, 5HT1c, 5HT1d, 5HT2 and 5HT3. (Leysen et al 1989). These receptors enable 5HT to exert its physiological effects as a neurotransmitter, a hormone and a mitogen. In particular 5HT present in the cortex of the brain contributes to affective and perceptual states through the excitation of tryptaminergic synapses. Lysergic acid diethylamine (LSD) is thought to act by antagonism of 5HT receptors an action which it effects in very small quantities. Hallucinations and delusions as a result of LSD use are well known and may bear some relation to the phenomena of schizophrenia. These synapses also represent one of the sites of action for psychotropic drugs (Jacobs 1984). The cloning and sequencing of a neuroreceptor gene allows the determination of its structure and enables new understanding about ligand binding and coupling of the receptors to messenger systems within the cell.

5.1.2 CLONING OF 5HT AND RELATED RECEPTORS

Three of the six 5HT receptor genes have been cloned so far. These are the human 5HT1a (Kobilka et al 1987a; Fargin et al 1988) and the rat 5HT1c and 5HT2 (Julius et al 1988; Pritchett et al 1988) genes. The 5HT1c receptor was isolated by cloning with RNA expression vectors in combination with patch clamp characterisation of the receptor in Xenopus oocyte membranes. Homologous hybridization with this receptor led to the isolation of the rat 5HT2 gene. The 5HT1a receptor gene was isolated by making use of sequence homology between the human beta 2 adrenergic receptor gene and the 5HT1a receptor (43% across the transmembrane regions at the amino acid level). The clone that was isolated was called G21 and was originally thought to encode an adrenergic receptor but was later identified as the 5HTla receptor gene through ligand binding assays (Fargin et al 1988). Recently the 5HT1d receptor has been cloned (Branchek et al

1990).

The cloning of the beta 2 adrenergic receptor gene (Kobilka et al 1987c) has in fact brought about the identification of a super family of receptor genes. All these genes including the 5HTIa, 5HTIc and 5HT2 genes have several features in common. They contain seven hydrophobic and alpha helical transmembrane regions which form ligand binding pockets. They all exert their functions through guanine nucleotide regulatory proteins (G-proteins). The specificity for coupling to the G-protein lies within a region extending from the amino terminus of the V transmembrane domain to the carboxy-terminus of the VI. For some receptors the seventh domain is important in agonist and antagonist ligand binding specificity (Kobilka et al 1988). They have sites for N-linked glycosylation near the amino termini and consensus sequences for phosphorylation by regulatory kinases on the cytoplasmic domains. There is also striking amino acid homology within the transmembrane regions which has enabled the cloning of these related receptor genes by homologous hybridization. Hybridization using the hamster beta 2 adrenergic receptor identified the rat dopamine D2 receptor (Bunzow et al 1988). The human dopamine D1 receptor was cloned using oligonucleotides homologous to the D2 receptor transmembrane region by use as either as DNA probes (Dearry et al 1990) or degenerate primers for PCR (Zhou et al 1990). Similar methods using the D2 receptor gene eventually led to the isolation of clones for the dopamine D3 receptor (Sokoloff et al 1990), and the dopamine D4 receptor (Van Tol et al 1991), and using the D1 receptor as a probe isolated the D5 receptor (Sunahara et al 1991). The human beta 1 adrenergic receptor was cloned by cross hybridization with the 5HT1a receptor (Frielle et al 1987). Another four probable members of the family of receptors have also been isolated from thyroid cDNA using degenerate transmembrane primers for PCR, but their exact functions have not yet been defined (Libert et al 1989). Other family members cloned include the alpha 1 and alpha 2 adrenergic receptors (Cotecchia et al 1988; Kobilka et al 1987b).

The 5HTIa, beta 2 adrenergic and D1 receptors represent a group of intronless genes all localised to chromosome 5. The beta 2, and alpha 1 adrenergic receptors map to chromosome 5 at the band 5q32-34 (Yang-Feng et al 1990; Sunahara et al 1990). The beta 2 and alpha 1 adrenergic receptor genes are within 300 kilo bases of each other. The glucocorticoid and dopamine D1 receptor genes have also been localised to this region (Francke et al 1989; Grandy et al 1990). The beta I and alpha 2 adrenergic receptors have been localised to chromosome 10 at 10q24-

26 and are within 250 kilobases of each other (Yang-Feng et al 1990). Both the receptor gene clusters and the other G-protein coupled receptors are thought to have arisen by gene and chromosomal duplication from an ancestral receptor gene. The gene family subsequently diverged with respect to function.

5.1.3 AIM

The 5HTIa receptor was localized by in situ hybridization to chromosome 5q11.2-13 (Kobilka et al 1987). As such it could be considered a candidate gene for the chromosome 5 type schizophrenia, however it was not found to lie within the deleted region of a chromosome 5 cell line deleted for 5q11.2-13.3 (Gilliam et al 1989). To confirm the original localization and to refine its map position in relation to other DNA markers a DNA polymorphism is required for linkage studies. The original G21 probe reveals a Sac1 RFLP (Khan et al 1990) which unfortunately has a low heterozygosity value and was uninformative in the pedigrees with multiple cases of schizophrenia. The aim was to increase the informativeness of the 5HTIa locus by isolating large genomic clones and searching for new polymorphisms with the extra DNA material.

5.2 METHODS AND RESULTS

5.2.1 IDENTIFICATION OF COSMID CLONES FOR THE 5HT1a RECEPTOR

Random primer labelled G21 was used to screen a human cosmid library (cloned into pcos2EMBL, Ehrich et al 1987; see materials and methods). After selection and purification of the colonies showing positive signals two cosmids containing the 5HTIa gene were identified (G21pcos6 and G21pcos2). The cosmids were cut with a variety of restriction endonucleases and the resulting fragments were size separated by agarose (0.8%) gel electrophoresis. After southern blotting to Hybond-N the filter was hybridized with radiolabelled G21 as a control, both cosmids gave a strong resulting signal on autoradiography (30 minutes at room temperature). For further definitive evidence that these cosmids contained the 5HT1a receptor an EcoRI/PstI 1.2kb restriction endonuclease fragment, with the Pst I site within the coding region of 5HT1a receptor, was subcloned into PUC19 and double stranded sequencing was carried out to confirm that the cosmids contained the 5HTIa gene. Approximately 200bp of sequence was obtained which was identical to published sequence (Kobilka et al 1987).

5.2.2 IDENTIFICATION OF A NEW POLYMORPHISM FOR THE 5HT1a RECEPTOR

Southern hybridization of the 5HTla cosmids with poly (dCA-dGT) and poly (dCT-dGA) (Pharmacia) did not identify any dinucleotide repeats so this avenue of potential highly polymorphic microsatellite markers was not available. However attempts were made to identify Alu elements within the cosmids which may have 3' repetitive sequences (Economou et al 1990). To acheive this the cosmids were cut with Sau3A1 and shotgun cloned into the M13 mp18 vector (BamHI cloning site; see materials and methods). Human Alu elements have two highly conserved regions (Kariya et al 1987) at the 5' and 3' ends of the element consisting 25bp and 16bp respectively. The 5' conserved region is located between nucleotide positions 23 and 47 whilst the other conserved region is found between nucleotides 245 and 260. Alu elements are approximately 280 nucleotides in length and the conserved regions are separated by a more variable region containing a adenine rich region. A 20mer oligonucleotide was designed for the 3' conserved region of Alu elements such that the last base (3' end of the oligonucleotide) was known to be identical for the majority of 50 randomly chosen and sequenced Alu elements (Kariya et al 1987). This oligonucleotide had the following sequence 5'GATCGCGCCACTGCACTCCA 3'. The Alu element 20mer was 5' end labelled with ^{32}P (using $^{32}PATP$ and T4 polynucleotide kinase see materials and methods) to a specific activity of at least 5×10^8 counts per minute per ug of oligonucleotide. This was used as a hybridization probe to identify M13 plaques containing human Alu elements. Hybridizations were carried out in 0.9M NaCl and 1% SDS with 40ug/ml yeast tRNA at 55°C for at least 16 hours. Post hybridization washes were at a stringency of 2 x SSC at 55°C. Positive clones were identified by autoradiography for 2-16 hours at -70°C these were picked and template DNA was prepared for direct sequencing. Dideoxy method of sequencing (Sanger et al 1977) using Sequenase (US Biochemical), and the Alu element 20mer identified one Alu element present in both of the cosmids. This Alu element had a 22 base pair 3'poly deoxyadenylate tract (Figure 1). A second oligonucleotide homologous to the 5' conserved region of Alu elements was generated to provide sequence data 5' of the poly deoxyadenylate tract (5'GTAATCCCAGCACTTTGGGAG 3'). The complete sequence for this Alu element is shown in Figure 1. One PCR primer was taken from the sequence after the poly deoxyadenylate tract (non Alu element primer), whilst the other was taken from the variable region of Alu elements (Figure 1). This was to minimize the background Alu contamination of the amplified product. PCR amplification was carried out according to Economou et al (1990). The PCR were carried out as for other microsatellites except that 50pM of ^{32}P end labelled non Alu

GG**5'GAAGTGAAAGTTGCAGTCATCTG3'**AGATTGCGCCACTGCACTCCAGC
CTGGGCGACAGAGCAAGACTCGGCTCTCAAAAAAAAAAAAAAAAAAAAAATT
AGCTGGGGCCTGGTGCGACATGCCTGTGGTCCCCAGCTACTTAGGAGGCTGA
3'GGTCGATGAATCCTCCGACT
AGTGGGAGGATTGCTTGAGCCTGGCAGGTGGAGGCTGCAGTGAGCTGTGAGT
TC5'
CCAGTA

Figure 1. DNA sequence for the human Alu element identified in the two cosmids for the 5HT1a receptor. The 22bp poly deoxyadenylate tract was amplified by PCR using the primers marked in bold. The conserved region of Alu elements is underlined.

element primer and 5pM of the Alu element primer were used in the amplification. PCR amplification of this poly deoxyadenylate tract in ten unrelated individuals revealed no polymorphism.

To identify potential RFLP's whole cosmid DNA was radiolabelled using the random primer method (Feinberg and Vogelstein, 1983). The human repeat elements present in the cosmid were competed out by hybridization with 2.5mg of unlabelled human genomic DNA (see materials and methods). The cosmid probe was then southern hybridized against an RFLP panel consisting of DNA from eight individuals cut with twenty separate restriction enzymes (ApaI, BamHI, BclI, BglI, BglII, EcoRI, EcoRV, HaeIII, HincII, HindIII, HinfI, MspI, PstI, PvuII, RsaI, SacI, StuI, TaqI, XbaI, and XmnI) and southern blotted onto Hybond-N. The cosmid (G21pcos6) identified a new Taq I polymorphism with alleles of 6.2 and 4.2kb. This cosmid was cut with TaqI and the fragments were separated by agarose gel electrophoresis, and southern hybridization with total human genomic DNA identified those fragments containing repetitive elements. A repeat free fragement corresponding to the smaller allele was present in the cosmid and was subcloned into a plasmid vector (pUC18 AccI site) to give cG21.T.4.2. The allele sizes and frequencies are shown in Figure 2 for the probe cG21.T.4.2.

5.2.3 LINKAGE MAPPING OF THE 5HT1a RECEPTOR

Linkage analysis between this RFLP and new microsatellite polymorphisms at the D5S76 and D5S39 loci, the RFLP at D5S6 and the RFLP at D5S21 (Dietzsch et al 1987) was performed in 18 Icelandic and British pedigrees. Two point lod scores are shown in Table 1 and multipoint analysis in Figure 3. The results demonstrate tight linkage between the 5HTIa receptor gene and chromosome 5 loci D5S76, D5S39 and D5S6. However a definitive position of the 5HT1a receptor in relation to D5S6 and D5S76 was not demonstrated. These results confirm the original localisation by in situ hybridisation for the 5HT1a receptor. D5S6 and D5S39 were both found to lie within a chromosome 5 cell line deleted for q11.2-13.3 whereas D5S76 and the 5HTIa receptor gene were not within this region (Gilliam et al 1989). This finding in combination with these results localizes the 5HTIa receptor to the proximal portion of the band 5q 11.2 either lying between D5S76 and D5S6 or proximal to D5S76.

Figure 2. Demonstrates the codominant segregation of 5HTIa/TaqI RFLP in a reference pedigree. 5ug of TaqI digested genomic DNA was size separated in agarose gels and blotted onto a nylon membrane. ^{32}P labelled probe cG21.T.4.2 was hybridized at 65°C in 0.9M NaCl,1% SDS and 10% dextran sulphate. The final wash was at a stringency of 0.5 x SSC at 65°C. Autoradiography overnight. The probe reveals a diallelic system with allele sizes 6.2kb and 4.2kb with frequencies 0.22 and 0.78 respectively (analysis of 70 chromosomes).

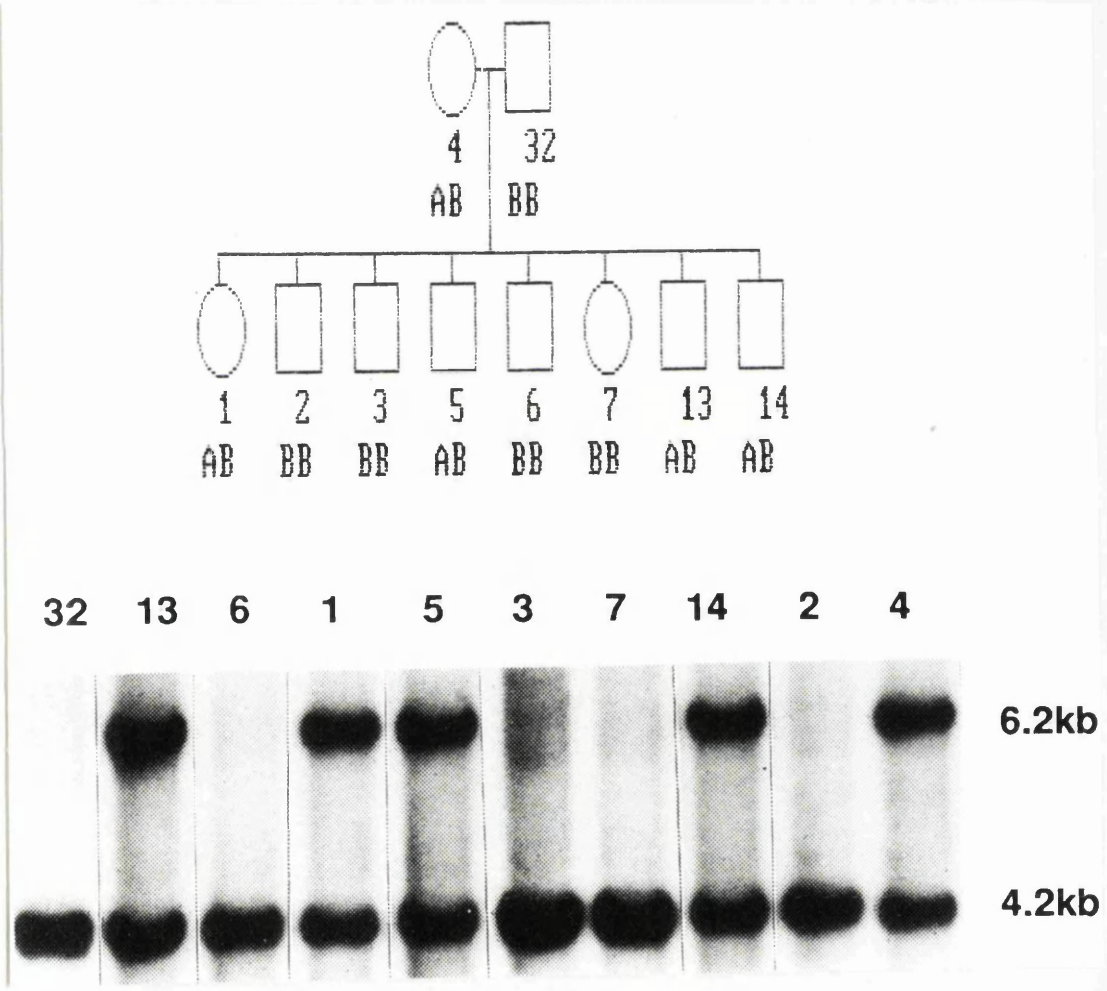
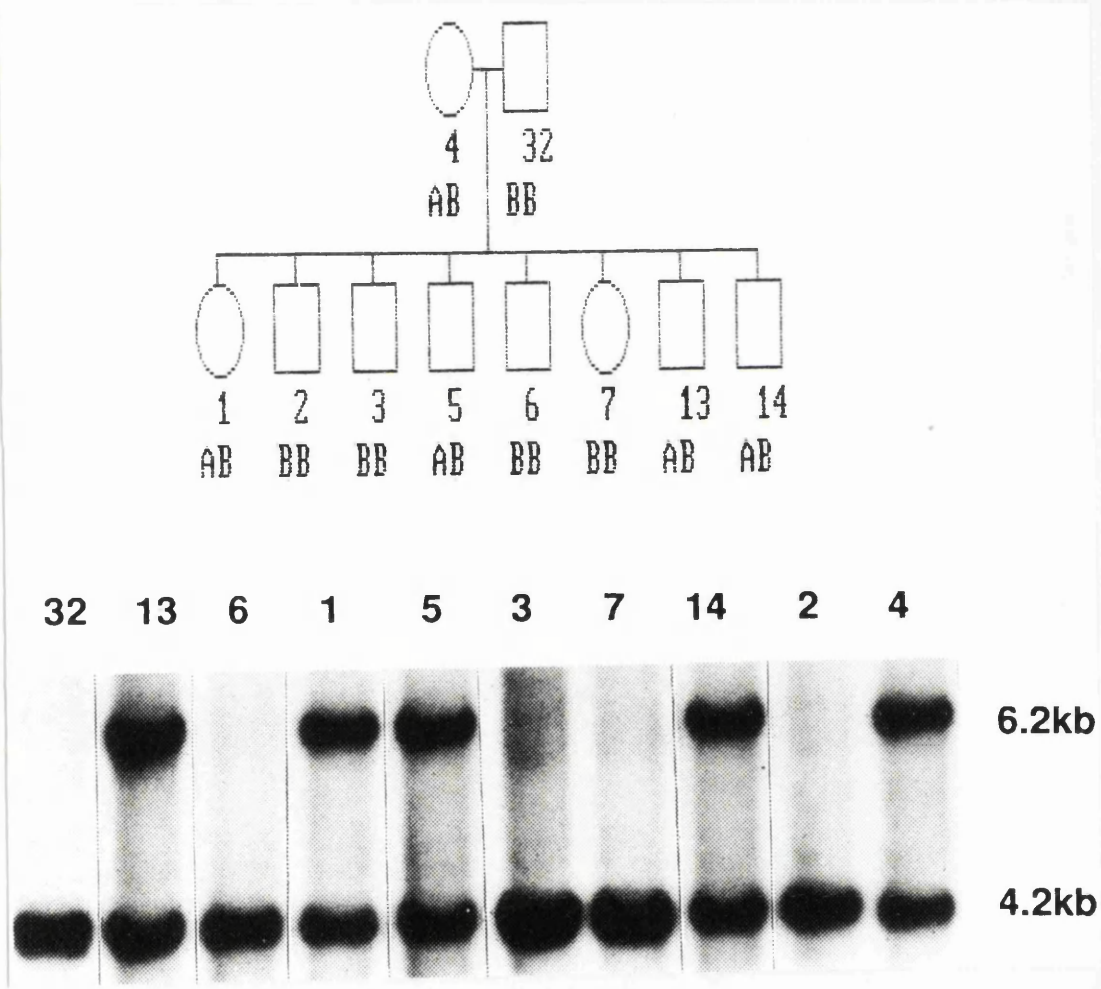


Figure 2. Demonstrates the codominant segregation of 5HTIa/TaqI RFLP in a reference pedigree. 5ug of TaqI digested genomic DNA was size separated in agarose gels and blotted onto a nylon membrane. ^{32}P labelled probe cG21.T.4.2 was hybridized at 65°C in 0.9M NaCl,1% SDS and 10% dextran sulphate. The final wash was at a stringency of 0.5 x SSC at 65°C. Autoradiography overnight. The probe reveals a diallelic system with allele sizes 6.2kb and 4.2kb with frequencies 0.22 and 0.78 respectively (analysis of 70 chromosomes).



Theta	0.00	0.01	0.05	0.10	0.20	0.30	0.40
6741(GT)	-99.00	5.07	6.64	6.69	5.50	3.54	1.28
p599(CA)	12.27	12.02	11.03	9.78	7.19	4.45	1.68
M4	4.05	3.97	3.64	3.29	2.57	1.66	0.66
YN(CT)	-99.00	2.86	5.12	5.53	4.77	3.12	1.13
EF(AG/TG)	-99.00	5.29	5.89	5.69	4.56	3.00	1.22

Table 1. MLINK two point lods for the cG21.T.4.2 TaqI RFLP with chromosome 5q11.2-13.3 markers.

5.3 DISCUSSION

It is likely that the G-protein coupled receptors arose by gene and chromosomal duplication from the same ancestral receptor gene. The 5HTIa receptor locus may represent another receptor gene cluster such as those on 5q32-34 and 10q24-26. On the assumption that the linkage reported on 5q11-13 for a subtype of schizophrenia is valid this cluster may contain important candidate genes as well as the 5HT1a receptor for schizophrenia. Alternatively the 5HTIa receptor gene at 5q11 may have arisen as a result of an isolated chromosomal rearrangement or duplication in which no further receptor genes have duplicated and diverged at this locus.

A number of reviews have been published on the possible role of serotonin in schizophrenia (Saddock et al 1989; Wyatt etal 1971; DeLisi et al 1985). The general findings are summarised below. Serotonin-depleting agents, such as reserpine , appear to alleviate some symptoms of schizophrenia. Whole blood serotonin was found to be significantly higher in schizophrenic patients with cerebral atrophy detected by CT scan than those with noormal CT scans or controls. Lower CSF 5-hydroxyindoleacetic acid (5-HIAA) the principal metabolite of serotonin was correlated with the presence of cerebral atrophy by CT scanning. Elevated CSF 5-HIAA was also reported to be associated with a positive family history. Postmortem brain studies have produced conflicting results. Regional elevations in brain serotonin in areas such as the putamen, hypothalamus, medial olfactory area, nucleus accumbens and globus pallidus still remain unconfirmed. In addition reduced serotonin in the hypothalamus, medulla oblongata and hippocampas of chronic schizophrenics complicate the above findings. These inconsistencies may simply represent difficulties inherent in obtaining reliable data from human postmortem studies. Cause of death, interval between death and autopsy, medication status, and reliability of postmortem diagnosis are only a few of the problems and these are exacerbated by the heterogeneous nature of schizophrenia. 5HT receptors are candidate genes for schizophrenia as well as other psychiatric disorders. They are one of the sites of blockade by a number of neuroleptics used to treat patients (Douglas, 1980; Fuller 1991). The neuroleptic Clozapine has a strong antagonist effect for both 5HT2 and 5HT1c receptors and this has been postulated to be relevant to its usefulness in the treatment of schizophrenia. Risperidone has a mixed serotonin 5HT2 and D2 dopamine antagonist properties being ten times more potent for 5HT2 receptors, and is known to exhibit a strong antipsychotic effect to improve both negative and and

affective symptoms of schizophrenia (Niemegeers et al 1991). Recent findings have shown an increase in 5HT1a receptors in prefrontal and temporal cortices of patients with chronic schizophrenia and 5HT1a receptor agonists antagonize neuroleptic catalepsy in rats (Fuller 1991). The relationship of serotonin to the illness, its aetiology and other biochemical abnormalities all warrant further investigation. To this end the localization of the 5HT1a gene and the identification of the TaqI RFLP will enable further linkage studies of schizophrenia.

CHAPTER 6: RE-ANALYSIS OF THE SCHIZOPHRENIA SUSCEPTIBILITY LOCUS ON CHROMOSOME 5

6.1 INTRODUCTION

6.1.1 RESUME

A partial trisomy of chromosome 5 q11.2-13.3 was found to co-segregate with schizophrenia in a single family (Bassett et al 1988). Following this report significant evidence for linkage using two RFLP's mapping to this area, (D5S39, and D5S76) was found in five Icelandic and two British families. This result was interpreted as providing evidence for a dominant schizophrenia susceptibility allele on chromosome 5. The maximum lod was recorded when all psychiatric disorders present in the families were classified as affected (Lod 6.42), therefore this susceptibility allele might predispose to a number of other psychiatric disorders. Five additional studies have failed to confirm this initial finding (Kennedy et al 1988; Detera-Wadleigh et al 1989; St Clair 1989; Kaufmann et al 1989; McGuffin et al 1990).

There appeared to be two possible hypotheses:

1. Heterogeneity exists and the gene on chromosome 5 accounts for only a small proportion of schizophrenia. This frequency may be elevated in the positive study by genetic drift in the isolated population of Iceland.

2. The positive finding was a type I statistical error (false positive).

To address these two hypothesis a highly polymorphic microsatellite map of chromosome 5q11.2-13.3 was developed to dramatically increase the number of informative meioses for linkage in the original families from the positive study. Approximately 50% of meioses were informative with the original RFLPs, increasing this proportion and establishing phase in certain pedigrees should confirm the linkage. Using the original pedigrees alone would rule out the addition of possible genes for oother subtypes of schizophrenia. In this way it should be possible to conclusively detect a type 1 statistical error or true linkage.

The first hypothesis of heterogeneity was to be studied with a second cohort of families from Iceland (N = 6). These were collected using the same stringent criteria

as for the first cohort and SADS-L interviews and concensus RDC dianoses were obtained from three clinicians for the individuals. These six families contained 78 individuals, with 18 diagnosed as having schizophrenia, 9 as schizophrenia spectrum disorders and a further 11 individuals had a diagnosis of some other psychiatric disorder including two with schizoaffective disorder (Figure 1). Pedigree structure was confirmed using the highly polymorphic microsatellites. If the chromosome 5 type schizophrenia does occur at elevated population levels in Iceland compared to other populations then this second group of families should be expected to replicate and confirm the original result. As a consequence true heterogeneity and a susceptibility allele for schizophrenia on chromosome 5 will have been shown. Using the highly polymorphic microsatellites not only will increase the power to detect linkage in these families, but will also increase the power of heterogeneity tests.

The new TaqI RFLP will elucidate to the potential involvement of the 5HT1a receptor as a candidate gene for the suceptibility locus for schizophrenia on chromosome 5. This RFLP provided genetic linkage information for the 5HT1a locus in the cohort of pedigrees whilst the published RFLP (Khan et al 1989) was uniformative. This information also provided a preliminary map position for the 5HT1a receptor, being tightly linked to D5S76 and D5S6.

6.1.2 MICROSATELLITE ANALYSIS OF DECEASED INDIVIDUALS

A third benefit of the microsatellite polymorphisms generated for this study has been the possibility to genetically type deceased individuals from the families if paraffin tissue blocks were available. Schizophrenics often have reduced family size compared to normal families (Kendler et al 1985). The availability of formalin-fixed paraffin-embedded tissue of deceased affected members families being studied by linkage offers the opportunity to genotype additional individuals and hence obtain valuable genetic information, in particular marker information in the older generations which may help to establish phase. The use of the polymerase chain reaction to examine paraffin-embedded tissue provides a relatively simple and extremely sensitive method for analysing archival material which may be unsuitable for techniques that require relatively large and pure quantities of high molecular weight DNA. PCR can produce a detectable signal with only a few nucleated cells (Shibata et al 1988), in contrast Southern blotting requires about 5ug DNA, equivalent to 5 $x10^6$ cells, to produce a detectable signal. Furthermore, the amplification product of microsatellites can be designed to be within 100-300

nucleotides in length, and so does not necessitate the use of high molecular weight DNA as substrate.

6.1.3 AIM

To analyse the original and new cohorts of multiply affected pedigrees with the new highly polymorphic microsatellites and the 5HT1a receptor RFLP for linkage to schizophrenia and related disorders. The area covered by the linkage study was to be expanded to cover all of the trisomic region associated with schizophrenia (Gilliam et al 1989) including markers proximal (D5S21 and D5S20) and distal (DHFR) to the markers producing evidence for linkage.

6.2 METHODS

6.2.1 POLYMORPHIC MARKERS ANALYSED

New DNA samples for individuals of the original cohort and the second cohort of families were extracted from frozen blood samples. The following markers were typed for the original seven families: p599(CA)$_n$ (Lambda599(CA)$_n$ was typed but provided identical linkage informatiom as p599(CA)$_n$ due to the strong linkage disequilibrium between these two markers) at the D5S76 locus, EF(AG/TG)$_n$ at the D5S125 locus, 6741(GT)$_n$ (6741(CA)$_n$ was typed but provided identical linkage information as p105-153Ra MspI RFLP) at D5S39, YN(CT)$_n$ at D5S127, Jo110Hc MspI RFLP at D5S21 (Figure 2, Leppert et al 1987), pJo71Ha Xba RFLP at D5S20 (Overhauser et al 1987), p105-798Rb MspI RFLP at D5S78 (Figure 2, Leppert et al 1987) and DHFR RsaI RFLP (Figure 2, Detera-Wadleigh et al 1989). The second cohort of Icelandic families were typed for the following markers: p599(CA)$_n$, EF(AG/TG)$_n$, 6741(GT)$_n$, YN(CT)$_n$, Jo110Hc, and M4 BamHI RFLP at D5S6 (Dietzsch et al 1986). All of the families were genotyped for the new 5HT1a TaqI RFLP.

6.2.2 ANALYSIS OF PARAFFIN-EMBEDDED TISSUE SAMPLES

Formalin-fixed paraffin-embedded samples of liver and brain biopsies of five individuals were obtained (F27/48, F41/29, F125/8, F125/9, and F141/8). All five fulfilled RDC criteria for schizophrenia. The blocks ranged in age from 7 to 18 years and had been stored at room temperature. Excess paraffin was trimmed from each block and several thin sections were sliced off with an autoclaved scaple blade. To minimise cross-contamination of samples, a fresh scaple blade was used for each block and all operations performed in an acid-washed biological safty cabinet. Sectioning was performed on Nescofilm and gloves were replaced between

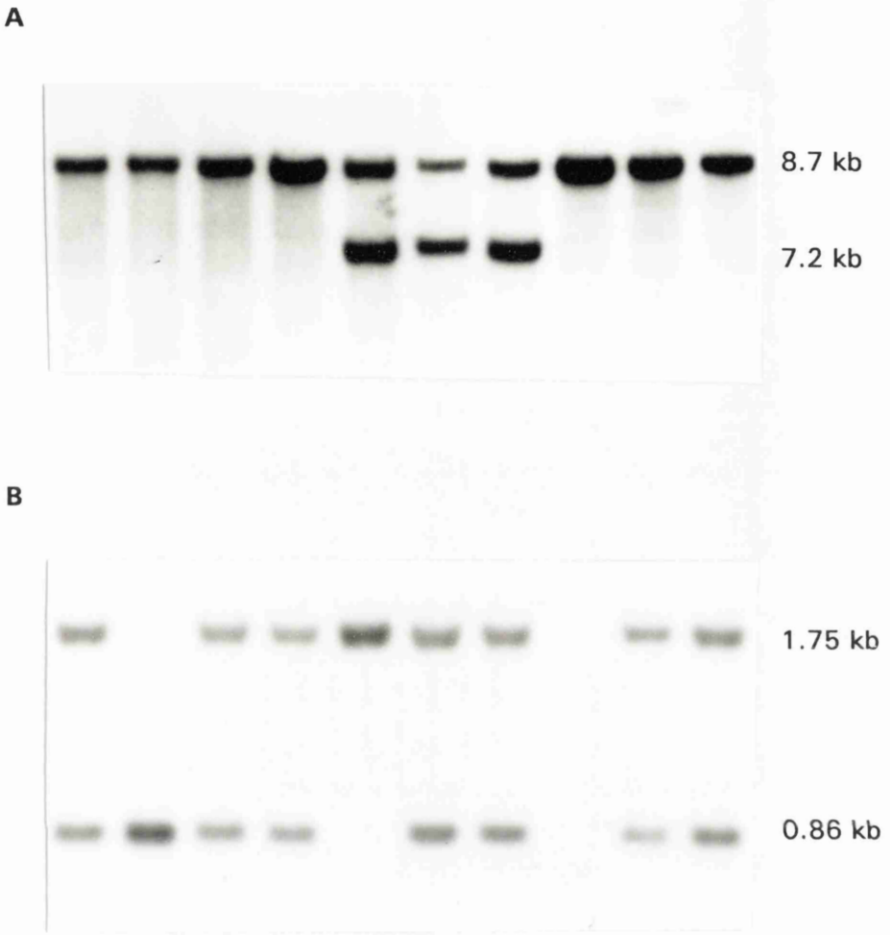


Figure 2. A. Jo110Hc MspI RFLP for the locus D5S21 with alleles of 8.7, 7.2 and 6.9 kb with frequencies of .54, .42 and .04 respectively. B. DHFR RsaI RFLP with alleles of 1.75 and 0.86 kb.

C

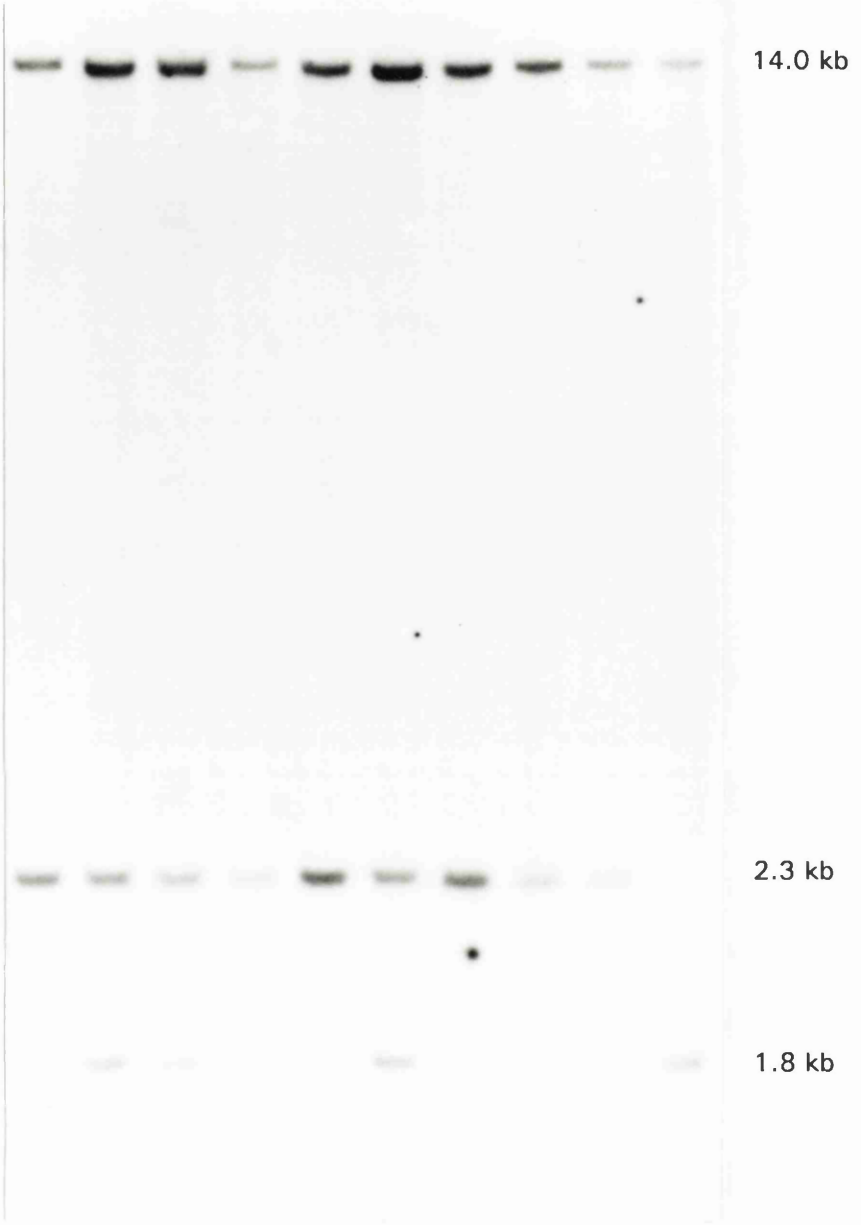


Figure 2 continued. C. p105-798Rb MspI RFLP at D5S78 with alleles of 2.3 and 1.8 kb, with frequencies of .57 and .43 respectively.

samples. Deparaffinisation was performed according to the protocol described by Wright and Manos (1990), each section was extracted twice with octane to remove paraffin, followed by two washes with 100% ethanol to remove the solvent. The ethanol was removed by evaporation at room temperature. The extracted tissue was treated with 200ug/ml proteinase K in 50 mM Tris-HCl pH8.5, 1mM EDTA, 0.5% Tween 20, overnight at 37oC. Proteinase K was inactivated by heating to 95oC for 10 minutes. All solids were pelleted by centrifugation for 30 seconds at 10,000g. Aliquots of 1 to 10 ul of the supernatant were used for the polymerase chain reaction. Prepared samples were stored at -20°C. Amplification was performed as previously described. Negative controls containing reagents only or reagents plus paraffin, but no tissue, were used.

6.2.3 LINKAGE ANALYSIS

All calculations were carried out for the same parameters, penetrances and inheritance models as in chapter 2.

6.3 RESULTS

To check for genotyping errors between the original RFLP's at D5S76 and D5S39 two point Lods were calculated with their respective microsatellites. One typing error for D5S76 microsatellite and the Taq I RFLP was identified as a recombinant. F35/7 and F35/17 were swapped for each other in the analysis of the TaqI RFLP. New blood samples confirmed this with a lod of ~~?~~ 24.44 at zero recombination. A second recombinant was found between p105-153Ra and 6741$(GT)_n$ at D5S39. F35/9 was a recombinant for his fathers alleles, no new blood sample was obtained to confirm this recombinant. Additional changes which may influence the linkage analysis includes the diagnosis. The variable age of onset for psychiatric disorders means that the diagnosis may change slightly within families with respect to time. Further analysis of the seven original families has found ~~three~~ two changes. An unaffected individual in F74 (F74/5) now has a RDC diagnosis of major depression and is recorded as affected in the DOMSSF model, whilst F27/26 previously classified as schizophrenia spectrum disorder is now considered to be unaffected. The genotyes recorded for each of the markers are given in Figure 1 and 3. Several new individuals have been collected F35/19, F27/45, F27/47, F27/48 and F41/29. F41/29 is an affected schizophrenic individual who is a first cousin to F41/99. Two point lods for all the families for all the markers, calculated for each model are shown in Table 1. Multipoint linkage analyses for each of the models was performed for both the original cohort and the new cohort separately and finally

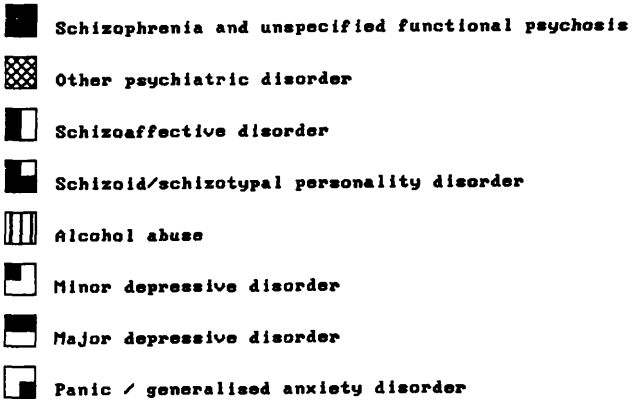
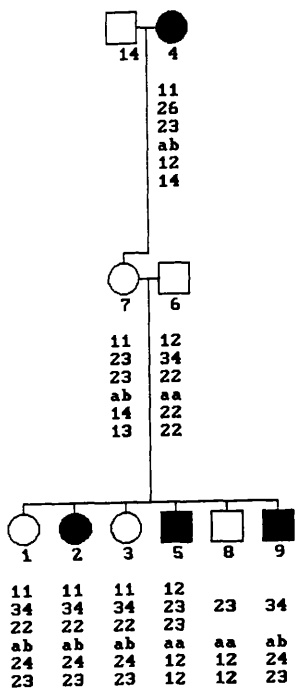
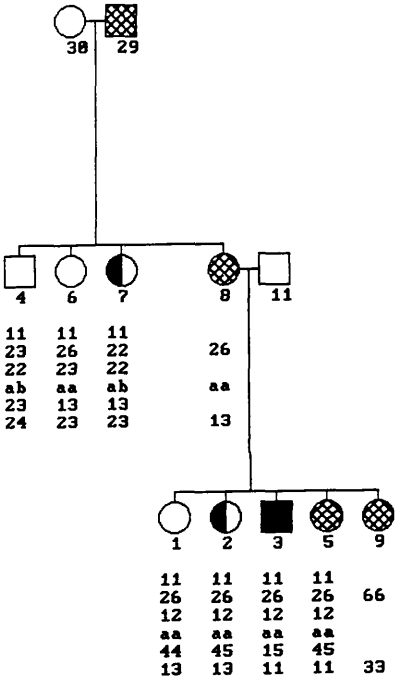


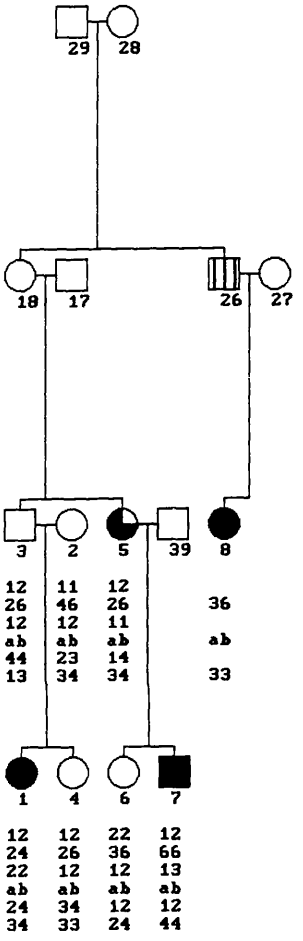
Figure 1. Pedigree structures, diagnoses and genotypes for the new cohort of Icelandic pedigrees. Families F121, F125, F141, F143, F184 and F224. The genotypes listed from top to bottom are for Jo110Hc, p599$(CA)_n$, M4, EF$(TG/AG)_n$, 6741$(GT)_n$ and YN$(CT)_n$.



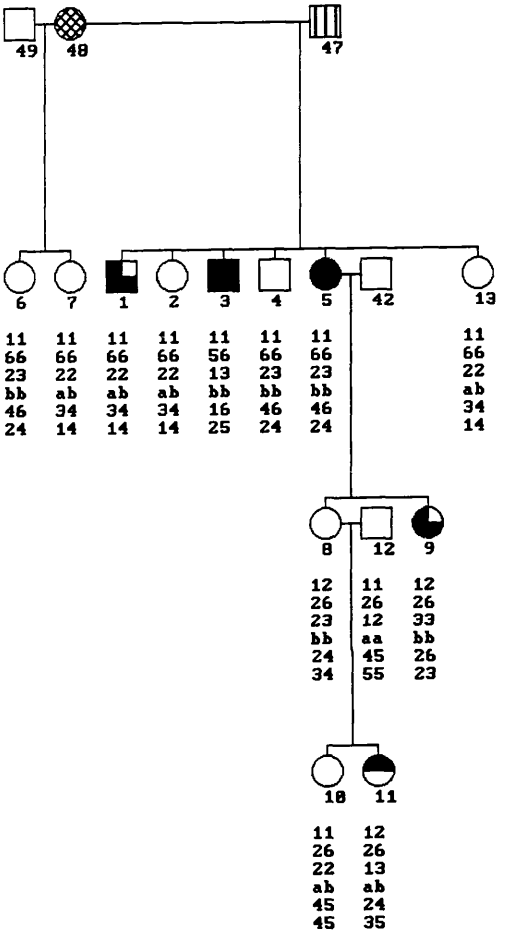
Pedigree 121



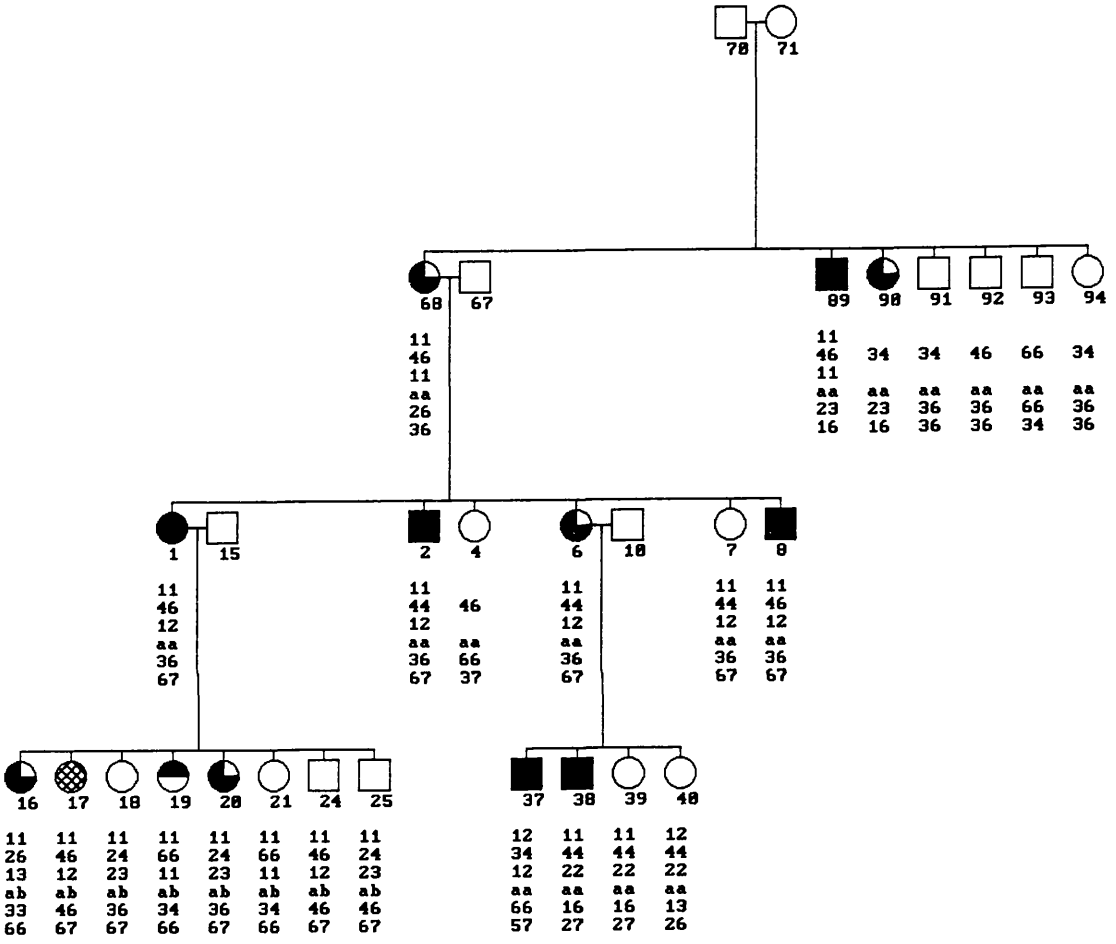
Pedigree 125



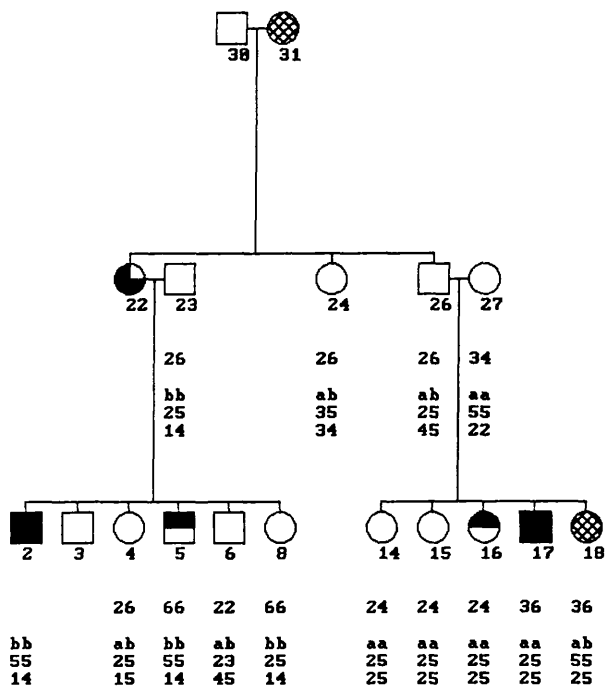
Pedigree 141



Pedigree 143



Pedigree 184



Pedigree 224

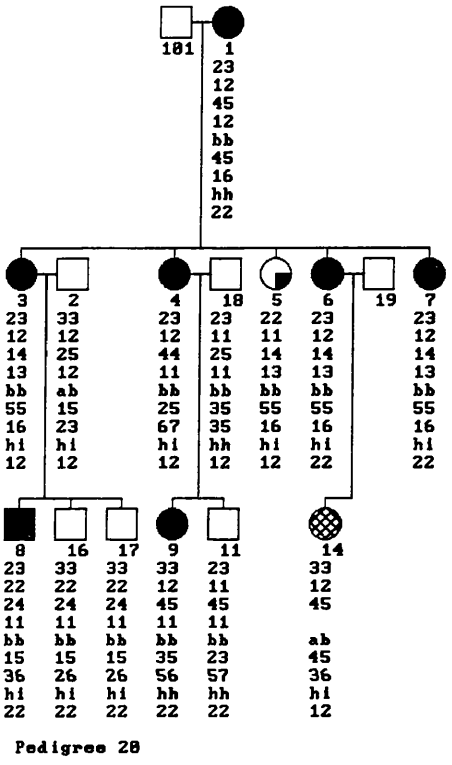
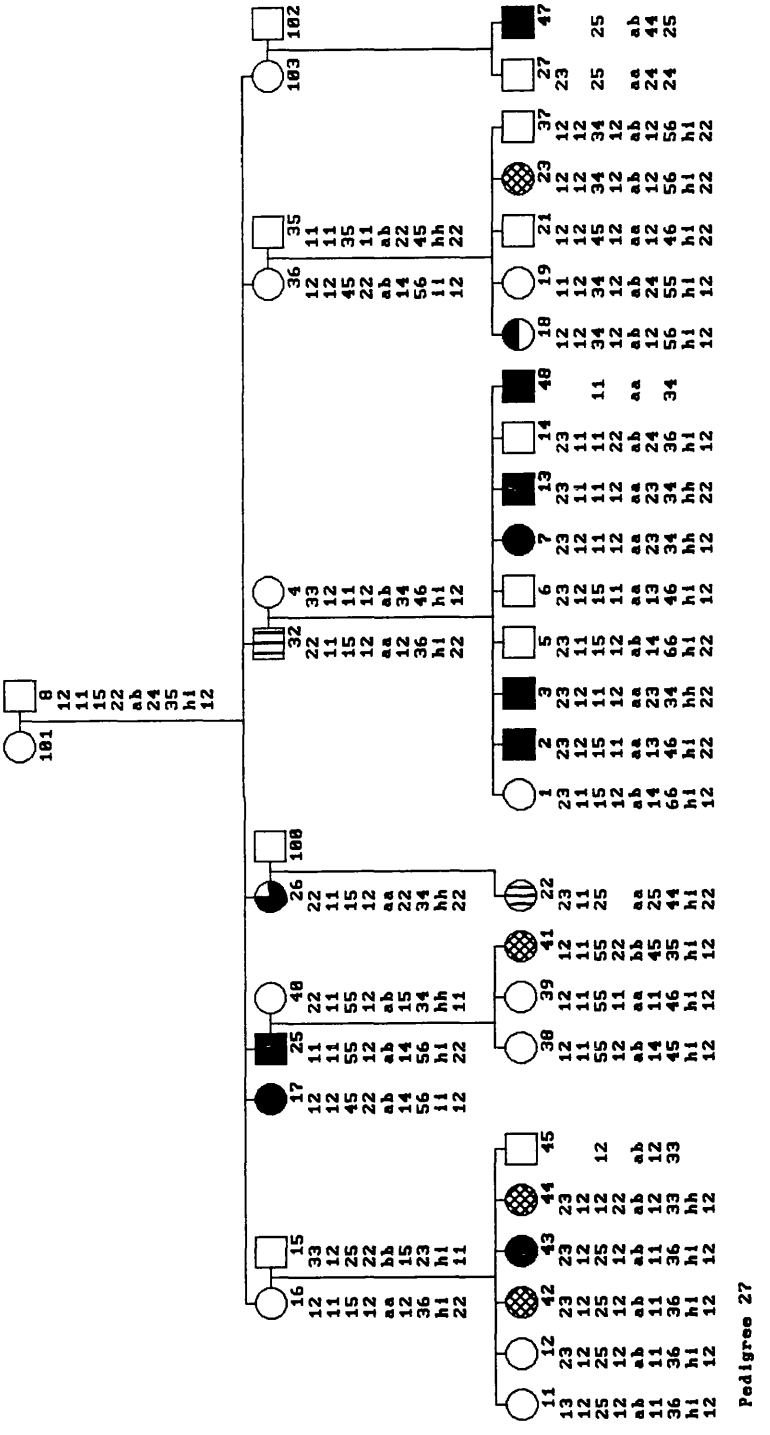
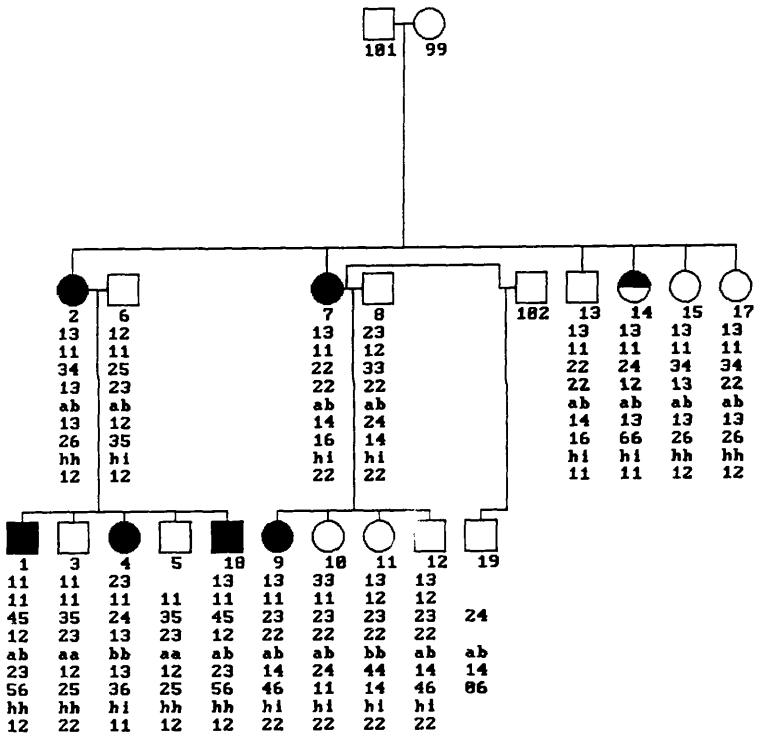


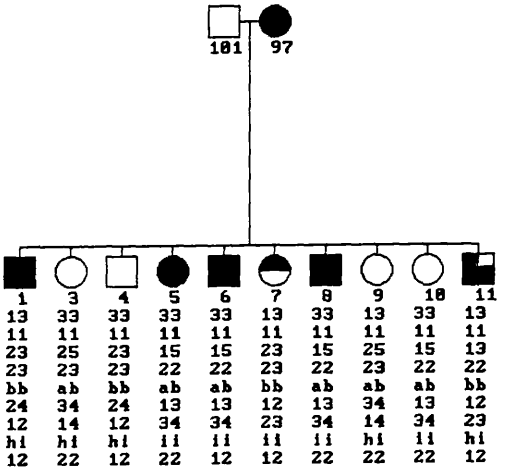
Figure 3. Pedigree structures, new diagnoses and genotypes for the original seven families analysed. F20, F27, F35, F36, F40, F41 and F74. The genotypes listed from top to bottom are for the following markers, J071Ha, Jo110Hc, p599$(CA)_n$, M4, EF(TG/AG)n, 6741$(GT)_n$, YN$(CT)_n$, p105-798Rb and DHFR.



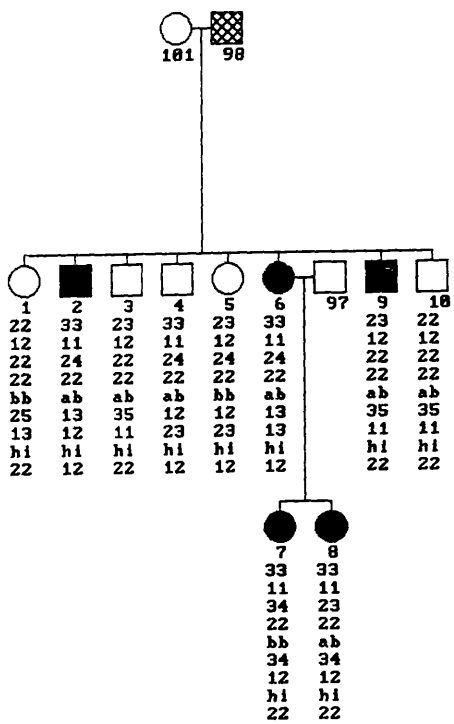
Pedigree 27


101 99
2 6 7 8 102 13 14 15 17
13 12 13 23 13 13 13 13 13
11 11 11 12 11 11 11 11 11
34 25 11 33 22 24 34 34 34
13 23 22 22 22 12 13 22 22
ab ab ab ab ab ab ab ab ab
13 12 14 24 14 13 13 13 13
26 35 16 14 16 66 26 26 26
hh hi hi hi hi hi hh hh hh
12 12 22 22 11 11 12 12 12
1 3 4 5 18 9 10 11 12 19
11 11 23 13 13 33 13 13
11 11 11 11 11 11 12 12
45 35 24 35 45 23 23 23 23 24
12 23 13 23 12 22 22 22 22
ab aa bb aa ab ab ab bb ab ab
23 12 13 12 23 14 24 44 14 14
56 25 36 25 56 46 11 14 46 66
hh hh hi hh hh hi hi hi hi
12 22 11 12 12 22 22 22 22

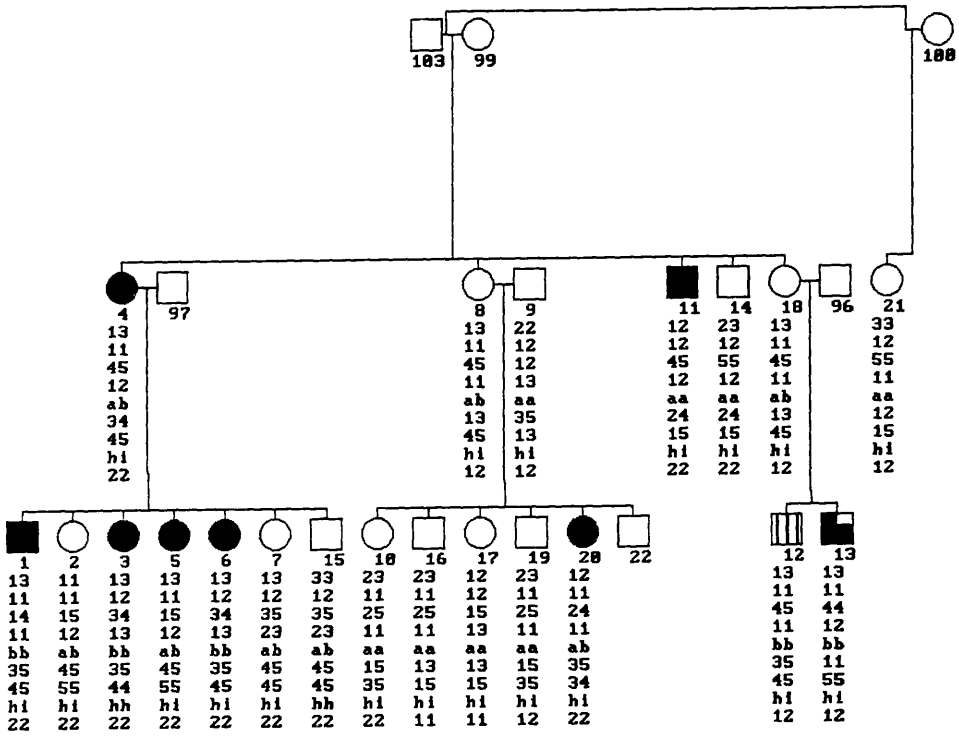
Pedigree 35


101
97
1 3 4 5 6 7 8 9 10 11
13 33 33 33 33 13 33 13 33 13
11 11 11 11 11 11 11 11 11 11
23 25 23 15 15 23 15 25 15 13
23 23 23 22 22 23 22 23 22 22
bb ab bb ab ab bb ab ab ab bb
24 34 24 13 13 12 13 34 13 12
12 14 12 34 34 23 34 14 34 23
hi hi hi ii ii ii ii hi ii hi
12 22 12 22 12 12 22 22 22 12

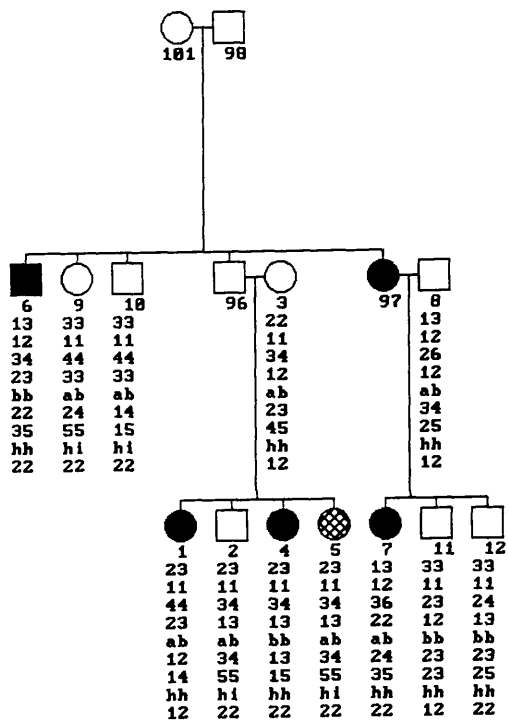
Pedigree 36


101
98
1
22
12
22
22
bb
25
13
hi
22
2
33
11
24
22
ab
13
12
hi
12
3
23
12
22
22
ab
35
11
hi
22
4
33
11
24
22
ab
12
23
hi
12
5
23
12
24
22
bb
12
23
hi
12
6
33
11
24
22
ab
13
13
hi
12
97
9
23
12
22
22
ab
35
11
hi
22
10
22
12
22
22
ab
35
11
hi
22
7
33
11
34
22
bb
34
12
hi
22
8
33
11
23
22
ab
34
12
hi
22

Pedigree 48


103
99
100
4
13
11
45
12
ab
34
45
hi
22
97
8
13
11
45
11
ab
13
45
hi
12
9
22
12
12
13
aa
35
13
hi
12
11
12
12
45
12
aa
24
15
hi
22
14
23
12
55
12
aa
24
15
hi
22
18
13
11
45
11
ab
13
45
hi
12
96
21
33
12
55
11
aa
12
15
hi
12
1
13
11
14
11
bb
35
45
hi
22
2
11
11
15
12
ab
45
55
hi
22
3
13
12
34
13
bb
35
44
hh
22
5
13
11
15
12
ab
45
55
hi
22
6
13
12
34
13
bb
35
45
hi
22
7
13
12
35
23
ab
45
45
hi
22
15
33
12
35
23
ab
45
45
hh
22
10
23
11
25
11
aa
15
35
hi
22
16
23
11
25
11
aa
13
15
hi
11
17
12
12
15
13
aa
13
15
hi
11
19
23
11
25
11
aa
15
35
hi
12
20
12
11
24
11
ab
35
34
hi
22
22
12
13
11
45
11
bb
35
45
hi
12
13
13
11
44
12
bb
11
55
hi
12

Pedigree 41


101
98
6
13
12
34
23
bb
22
35
hh
22
9
33
11
44
33
ab
24
55
hi
22
10
33
11
44
33
ab
14
15
hi
22
96
3
22
11
34
12
ab
23
45
hh
12
97
8
13
12
26
12
ab
34
25
hh
12
1
23
11
44
23
ab
12
14
hh
12
2
23
11
34
13
ab
34
55
hi
22
4
23
11
34
13
bb
13
15
hh
22
5
23
11
34
13
ab
34
55
hi
22
7
13
12
36
22
ab
24
35
hh
22
11
33
11
23
12
bb
23
23
hh
12
12
33
11
24
13
bb
23
25
hh
22

Pedigree 74

MLINK : DOMS J071Ha

theta	0.00	0.01	0.05	0.10	0.20	0.30	0.40
F20	-1.31	-0.93	-0.44	-0.20	-0.02	0.01	-0.02
F27	-2.16	-1.84	-1.21	-0.81	-0.39	-0.16	-0.04
F35	-3.76	-2.69	-1.46	-0.91	-0.40	-0.16	-0.04
F36	-0.30	-0.29	-0.25	-0.19	-0.10	-0.04	-0.01
F40	0.32	0.32	0.30	0.27	0.19	0.10	0.03
F41	0.13	0.15	0.20	0.22	0.19	0.11	0.03
F74	-0.36	-0.34	-0.26	-0.19	-0.10	-0.04	-0.01
total	-7.44	-5.62	-3.11	-1.81	-0.62	-0.18	-0.05

MLINK : DOMSS J071Ha

theta	0.00	0.01	0.05	0.10	0.20	0.30	0.40
F20	-1.35	-0.97	-0.46	-0.22	-0.03	0.00	-0.02
F27	-2.07	-1.77	-1.05	-0.64	-0.27	-0.11	-0.03
F35	-3.97	-2.73	-1.48	-0.92	-0.40	-0.16	-0.04
F36	-0.31	-0.31	-0.30	-0.28	-0.21	-0.11	-0.03
F40	0.30	0.30	0.29	0.27	0.19	0.11	0.03
F41	0.82	0.82	0.80	0.74	0.56	0.32	0.09
F74	-0.37	-0.34	-0.26	-0.19	-0.09	-0.04	-0.01
total	-6.94	-5.00	-2.47	-1.24	-0.25	0.02	0.00

MLINK : DOMSSF J071Ha

theta	0.00	0.01	0.05	0.10	0.20	0.30	0.40
F20	-0.74	-0.74	-0.56	-0.33	-0.09	-0.01	-0.01
F27	-1.87	-1.69	-1.14	-0.71	-0.25	-0.06	-0.01
F35	-4.57	-2.87	-1.55	-0.96	-0.42	-0.16	-0.04
F36	-0.31	-0.31	-0.30	-0.29	-0.20	-0.10	-0.03
F40	0.18	0.19	0.21	0.22	0.18	0.11	0.03
F41	0.97	0.98	0.97	0.91	0.69	0.39	0.11
F74	-0.35	-0.32	-0.22	-0.15	-0.07	-0.03	-0.01
total	-6.68	-4.75	-2.59	-1.31	-0.17	0.12	0.06

Table 1. Two point lods for Jo71Ha (old families).

MLINK : DOMS Jo110Hc

theta	0.00	0.01	0.05	0.10	0.20	0.30	0.40
F20	0.88	0.87	0.83	0.76	0.58	0.36	0.13
F27	-0.21	-0.20	-0.16	-0.11	-0.04	-0.01	0.00
F35	-0.01	-0.01	-0.01	-0.01	-0.01	0.00	0.00
F36	0.00	0.00	0.00	0.00	0.00	0.00	0.00
F40	-0.27	-0.23	-0.12	-0.04	0.02	0.03	0.01
F41	-0.31	-0.30	-0.27	-0.22	-0.13	-0.06	-0.01
F74	-0.35	-0.33	-0.25	-0.18	-0.09	-0.04	-0.01
total	-0.28	-0.21	0.02	0.20	0.34	0.27	0.12

MLINK : DOMSS Jo110Hc

theta	0.00	0.01	0.05	0.10	0.20	0.30	0.40
F20	0.82	0.81	0.79	0.73	0.57	0.35	0.13
F27	0.72	0.71	0.64	0.55	0.37	0.19	0.05
F35	-0.01	-0.01	-0.01	-0.01	-0.01	0.00	0.00
F36	0.00	0.00	0.00	0.00	0.00	0.00	0.00
F40	-0.27	-0.24	-0.13	-0.04	0.03	0.03	0.01
F41	-0.10	-0.08	-0.02	0.02	0.04	0.03	0.01
F74	-0.36	-0.33	-0.25	-0.18	-0.09	-0.04	-0.01
total	0.80	0.86	1.02	1.07	0.92	0.56	0.19

MLINK : DOMSSF Jo110HC

theta	0.00	0.01	0.05	0.10	0.20	0.30	0.40
F20	-1.34	-1.03	-0.49	-0.19	0.06	0.10	0.07
F27	-2.61	-2.40	-1.54	-1.02	-0.51	-0.22	-0.05
F35	-0.01	-0.01	-0.01	-0.01	-0.01	0.00	0.00
F36	0.00	0.00	0.00	0.00	0.00	0.00	0.00
F40	-0.30	-0.26	-0.14	-0.04	0.04	0.05	0.02
F41	-0.12	-0.10	-0.03	0.01	0.04	0.03	0.01
F74	-0.33	-0.30	-0.21	-0.14	-0.07	-0.03	-0.01
total	-4.70	-4.09	-2.42	-1.39	-0.44	-0.07	0.03

Table 1. Two point lods for Jo110Hc (old families)

MLINK : DOMS p599(CA)

theta	0.00	0.01	0.05	0.10	0.20	0.30	0.40
F20	-0.29	-0.27	-0.18	-0.11	-0.02	-0.02	-0.02
F27	-2.08	-1.96	-1.25	-0.80	-0.31	-0.06	0.01
F35	-2.15	-0.47	0.14	0.33	0.37	0.25	0.09
F36	-2.00	-1.18	-0.54	-0.27	-0.05	0.01	0.01
F40	-2.08	-1.55	-0.94	-0.63	-0.29	-0.12	-0.03
F41	-0.07	0.68	1.21	1.31	1.12	0.73	0.26
F74	-1.60	-0.95	-0.38	-0.16	-0.01	0.02	0.01
total	-10.27	-5.70	-1.95	-0.32	0.80	0.82	0.32

MLINK : DOMSS p599(CA)

theta	0.00	0.01	0.05	0.10	0.20	0.30	0.40
F20	-0.45	-0.42	-0.31	-0.21	-0.08	-0.04	-0.04
F27	-3.32	-1.71	-0.28	0.42	0.86	0.74	0.34
F35	-2.30	-0.46	0.15	0.34	0.38	0.26	0.09
F36	-1.48	-0.64	-0.06	0.13	0.20	0.13	0.04
F40	-2.12	-1.57	-0.95	-0.63	-0.30	-0.12	-0.03
F41	1.39	2.13	2.49	2.41	1.91	1.21	0.45
F74	-1.66	-0.97	-0.39	-0.16	-0.01	0.02	0.01
total	-9.94	-3.64	0.64	2.30	2.96	2.20	0.86

MLINK : DOMSSF p599(CA)

theta	0.00	0.01	0.05	0.10	0.20	0.30	0.40
F20	0.53	0.55	0.61	0.63	0.55	0.38	0.15
F27	-7.55	-6.47	-3.14	-1.51	-0.18	0.20	0.14
F35	-2.14	-1.19	-0.41	-0.05	0.21	0.21	0.09
F36	-2.09	-1.96	-1.20	-0.69	-0.24	-0.07	-0.01
F40	-2.30	-1.62	-0.98	-0.66	-0.32	-0.13	-0.03
F41	0.39	1.31	1.86	1.94	1.64	1.07	0.39
F74	-1.91	-1.09	-0.45	-0.19	-0.01	0.02	0.01
total	-15.08	-10.47	-3.72	-0.54	1.65	1.68	0.73

Table 1. Two point lods for p599$(CA)_n$ (old families)

MLINK : DOMS M4

theta	0.00	0.01	0.05	0.10	0.20	0.30	0.40
F20	0.08	0.09	0.13	0.16	0.16	0.11	0.03
F27	-2.84	-2.15	-1.36	-0.90	-0.41	-0.16	-0.04
F35	-2.14	-0.74	-0.10	0.14	0.26	0.21	0.08
F36	-0.29	-0.26	-0.14	-0.05	0.03	0.04	0.02
F40	0.00	0.00	0.00	0.00	0.00	0.00	0.00
F41	-1.32	-0.56	-0.01	0.18	0.24	0.16	0.05
F74	-2.03	-1.15	-0.53	-0.28	-0.09	-0.03	-0.01
total	-8.54	-4.77	-2.00	-0.76	0.20	0.33	0.13

MLINK : DOMSS M4

theta	0.00	0.01	0.05	0.10	0.20	0.30	0.40
F20	-0.04	-0.02	0.04	0.09	0.12	0.08	0.02
F27	-3.78	-2.94	-1.72	-1.13	-0.55	-0.24	-0.06
F35	-2.29	-0.78	-0.12	0.13	0.26	0.21	0.08
F36	-0.25	-0.12	0.14	0.26	0.29	0.19	0.06
F40	0.00	0.00	0.00	0.00	0.00	0.00	0.00
F41	-1.11	-0.33	0.20	0.35	0.35	0.21	0.06
F74	-2.12	-1.17	-0.54	-0.28	-0.09	-0.03	-0.01
total	-9.59	-5.37	-2.00	-0.58	0.38	0.43	0.15

MLINK : DOMSSF M4

theta	0.00	0.01	0.05	0.10	0.20	0.30	0.40
F20	0.23	0.26	0.33	0.37	0.35	0.23	0.07
F27	-4.27	-3.10	-1.52	-0.80	-0.23	-0.05	-0.01
F35	-2.09	-0.43	0.20	0.40	0.45	0.31	0.11
F36	-0.29	-0.29	-0.25	-0.16	-0.02	0.03	0.02
F40	0.00	0.00	0.00	0.00	0.00	0.00	0.00
F41	-1.18	-0.28	0.27	0.42	0.41	0.25	0.07
F74	-2.52	-1.30	-0.62	-0.33	-0.10	-0.03	-0.01
total	-10.12	-5.14	-1.59	-0.10	0.85	0.74	0.25

Table 1. Two point lods for M4 (old families)

MLINK : DOMS EF(TG/AG)

theta	0.00	0.01	0.05	0.10	0.20	0.30	0.40
F20	0.00	0.00	0.00	0.00	0.00	0.00	0.00
F27	-0.45	-0.43	-0.36	-0.29	-0.17	-0.08	-0.02
F35	0.01	0.02	0.02	0.03	0.02	0.01	0.00
F36	-0.32	-0.31	-0.26	-0.20	-0.10	-0.04	-0.01
F40	-0.37	-0.35	-0.27	-0.19	-0.08	-0.03	0.00
F41	-1.78	-0.92	0.02	0.44	0.65	0.52	0.23
F74	0.49	0.48	0.42	0.35	0.22	0.11	0.03
total	-2.42	-1.51	-0.42	0.15	0.54	0.50	0.23

MLINK : DOMSS EF(TG/AG)

theta	0.00	0.01	0.05	0.10	0.20	0.30	0.40
F20	0.00	0.00	0.00	0.00	0.00	0.00	0.00
F27	-0.87	-0.84	-0.61	-0.37	-0.13	-0.04	-0.01
F35	-0.01	0.00	0.01	0.02	0.02	0.01	0.00
F36	-0.31	-0.31	-0.30	-0.28	-0.21	-0.11	-0.03
F40	-0.38	-0.35	-0.27	-0.20	-0.09	-0.03	-0.01
F41	-0.48	0.22	0.92	1.15	1.09	0.76	0.30
F74	0.50	0.49	0.44	0.37	0.23	0.12	0.03
total	-1.54	-0.80	0.18	0.69	0.93	0.71	0.30

MLINK : DOMSSF EF(TG/AG)

theta	0.00	0.01	0.05	0.10	0.20	0.30	0.40
F20	0.00	0.00	0.00	0.00	0.00	0.00	0.00
F27	-0.70	-0.67	-0.31	-0.02	0.15	0.12	0.04
F35	-0.13	-0.12	-0.07	-0.03	0.00	0.01	0.00
F36	-0.28	-0.28	-0.27	-0.25	-0.18	-0.09	-0.02
F40	-0.39	-0.37	-0.30	-0.23	-0.12	-0.05	-0.01
F41	0.21	1.05	1.74	1.89	1.65	1.12	0.47
F74	0.56	0.54	0.48	0.41	0.26	0.13	0.04
total	-0.73	0.16	1.28	1.77	1.77	1.25	0.51

Table 1. Two point lods for $EF(TG/AG)_n$ (old families)

MLINK : DOMS 6741(GT)

theta	0.00	0.01	0.05	0.10	0.20	0.30	0.40
F20	1.01	1.00	0.94	0.86	0.66	0.42	0.18
F27	-1.58	-1.54	-1.34	-1.03	-0.48	-0.14	-0.01
F35	-2.41	-1.60	-0.88	-0.52	-0.17	-0.04	0.00
F36	-2.18	-1.51	-0.85	-0.53	-0.22	-0.08	-0.02
F40	0.75	0.74	0.70	0.63	0.44	0.23	0.06
F41	-2.20	-1.23	-0.02	0.45	0.68	0.54	0.23
F74	-1.29	-0.80	-0.26	-0.04	0.11	0.10	0.04
total	-7.90	-4.93	-1.71	-0.18	1.01	1.03	0.48

MLINK : DOMSS 6741(GT)

theta	0.00	0.01	0.05	0.10	0.20	0.30	0.40
F20	0.99	0.97	0.92	0.84	0.65	0.42	0.18
F27	-4.60	-2.88	-1.22	-0.32	0.37	0.45	0.22
F35	-2.59	-1.64	-0.90	-0.54	-0.18	-0.04	0.00
F36	-1.90	-1.21	-0.59	-0.32	-0.10	-0.02	0.00
F40	0.70	0.69	0.66	0.61	0.44	0.23	0.06
F41	-3.73	-2.03	-0.38	0.23	0.55	0.44	0.16
F74	-1.35	-0.82	-0.27	-0.04	0.11	0.11	0.04
total	-12.49	-6.90	-1.77	0.47	1.84	1.57	0.66

MLINK : DOMSSF 6741(GT)

theta	0.00	0.01	0.05	0.10	0.20	0.30	0.40
F20	1.69	1.66	1.54	1.38	1.04	0.66	0.27
F27	-7.23	-6.48	-3.66	-1.87	-0.36	0.12	0.13
F35	-2.56	-2.27	-1.56	-1.04	-0.43	-0.14	-0.02
F36	-1.49	-0.62	-0.02	0.20	0.28	0.19	0.06
F40	0.41	0.43	0.48	0.48	0.39	0.22	0.06
F41	-3.51	-1.21	0.44	0.97	1.10	0.80	0.32
F74	-3.03	-2.48	-1.45	-0.89	-0.36	-0.12	-0.02
total	-15.71	-10.96	-4.22	-0.77	1.65	1.73	0.73

Table 1. Two point lods for 6741(GT)$_n$ (old families)

MLINK : DOMS YN(CT)

theta	0.00	0.01	0.05	0.10	0.20	0.30	0.40
F20	-0.10	-0.08	0.02	0.12	0.22	0.21	0.13
F27	-1.19	-1.16	-0.98	-0.70	-0.22	0.03	0.05
F35	-2.85	-2.28	-1.10	-0.57	-0.13	0.00	0.01
F36	-2.18	-1.51	-0.85	-0.53	-0.22	-0.08	-0.02
F40	-2.17	-1.65	-0.98	-0.62	-0.26	-0.09	-0.02
F41	-2.78	-2.02	-0.86	-0.33	0.06	0.13	0.06
F74	-1.38	-1.32	-1.00	-0.66	-0.26	-0.08	-0.01
total	-12.66	-10.01	-5.76	-3.29	-0.81	0.12	0.21

MLINK : DOMSS YN(CT)

theta	0.00	0.01	0.05	0.10	0.20	0.30	0.40
F20	-0.25	-0.22	-0.09	0.05	0.20	0.21	0.13
F27	-4.07	-2.32	-0.77	0.05	0.66	0.65	0.32
F35	-3.10	-2.38	-1.14	-0.59	-0.14	0.00	0.01
F36	-1.90	-1.20	-0.59	-0.32	-0.10	-0.02	0.00
F40	-2.33	-1.75	-1.05	-0.67	-0.28	-0.10	-0.02
F41	-2.18	-2.01	-1.11	-0.50	-0.03	0.06	0.03
F74	-1.47	-1.39	-1.04	-0.68	-0.27	-0.08	-0.01
total	-15.30	-11.28	-5.78	-2.65	0.04	0.72	0.46

MLINK : DOMSSF YN(CT)

theta	0.00	0.01	0.05	0.10	0.20	0.30	0.40
F20	0.77	0.79	0.81	0.80	0.67	0.45	0.20
F27	-6.89	-6.13	-3.22	-1.50	-0.07	0.33	0.22
F35	-3.34	-3.08	-1.90	-1.13	-0.38	-0.08	0.00
F36	-1.49	-0.62	-0.02	0.20	0.28	0.19	0.06
F40	-2.87	-2.24	-1.39	-0.89	-0.38	-0.13	-0.03
F41	-2.23	-1.91	-0.59	0.02	0.35	0.29	0.11
F74	-3.06	-2.70	-1.95	-1.35	-0.61	-0.23	-0.05
total	-19.10	-15.89	-8.26	-3.86	-0.15	0.81	0.51

Table 1. Two point lods for YN(CT)^n (old families)

MLINK : DOMS p105-798Rb

theta	0.00	0.01	0.05	0.10	0.20	0.30	0.40
F20	-0.04	-0.03	-0.01	0.01	0.04	0.05	0.03
F27	-1.57	-0.94	-0.32	-0.05	0.12	0.11	0.04
F35	-0.33	-0.32	-0.27	-0.22	-0.13	-0.06	-0.01
F36	-0.29	-0.25	-0.14	-0.05	0.03	0.04	0.01
F40	0.17	0.17	0.15	0.12	0.07	0.03	0.01
F41	-0.90	-0.83	-0.62	-0.45	-0.23	-0.10	-0.02
F74	0.82	0.80	0.72	0.62	0.41	0.22	0.06
total	-2.12	-1.40	-0.51	-0.02	0.33	0.29	0.12

MLINK : DOMSS p105-798Rb

theta	0.00	0.01	0.05	0.10	0.20	0.30	0.40
F20	-0.08	-0.07	-0.04	-0.01	0.03	0.04	0.03
F27	-0.26	-0.24	-0.05	0.13	0.24	0.12	0.07
F35	-0.33	-0.32	-0.28	-0.23	-0.13	-0.06	-0.02
F36	-0.30	-0.30	-0.28	-0.25	-0.16	-0.07	-0.02
F40	0.17	0.17	0.15	0.12	0.07	0.03	0.01
F41	-1.09	-1.01	-0.75	-0.53	-0.26	-0.11	-0.03
F74	0.84	0.82	0.74	0.63	0.43	0.23	0.07
total	-1.05	-0.94	-0.51	-0.14	0.21	0.24	0.12

MLINK : DOMSSF p105-798Rb

theta	0.00	0.01	0.05	0.10	0.20	0.30	0.40
F20	-0.37	-0.35	-0.27	-0.20	-0.10	-0.04	0.00
F27	-2.01	-1.78	-1.12	-0.62	-0.12	0.04	0.04
F35	-0.34	-0.32	-0.26	-0.19	-0.09	-0.04	-0.01
F36	-0.25	-0.25	-0.21	-0.14	-0.01	0.03	0.02
F40	0.17	0.17	0.15	0.12	0.07	0.03	0.01
F41	-1.04	-0.94	-0.67	-0.45	-0.20	-0.08	-0.02
F74	-0.98	-0.73	-0.34	-0.17	-0.04	-0.01	0.00
total	-4.82	-4.20	-2.72	-1.64	-0.49	-0.05	0.03

Table 1. Two point lods for p105-798Rb (old families)

MLINK : DOMS DHFR

theta	0.00	0.01	0.05	0.10	0.20	0.30	0.40
F20	-0.51	-0.49	-0.39	-0.29	-0.14	-0.05	0.00
F27	-0.28	-0.26	-0.20	-0.14	-0.06	-0.02	0.00
F35	-2.03	-0.96	-0.34	-0.12	0.03	0.05	0.02
F36	-0.31	-0.31	-0.30	-0.28	-0.19	-0.10	-0.03
F40	0.28	0.27	0.24	0.20	0.11	0.05	0.01
F41	0.10	0.10	0.10	0.08	0.05	0.02	0.00
F74	-0.02	-0.02	-0.02	-0.02	-0.01	-0.01	0.00
total	-2.77	-1.66	-0.92	-0.56	-0.21	-0.05	0.01

MLINK : DOMSS DHFR

theta	0.00	0.01	0.05	0.10	0.20	0.30	0.40
F20	-0.64	-0.60	-0.49	-0.37	-0.19	-0.08	-0.01
F27	-0.97	-0.94	-0.81	-0.64	-0.35	-0.15	-0.04
F35	-2.11	-0.95	-0.33	-0.10	0.04	0.06	0.03
F36	-0.31	-0.31	-0.30	-0.26	-0.16	-0.08	-0.02
F40	0.28	0.27	0.24	0.19	0.11	0.05	0.01
F41	-0.18	-0.17	-0.13	-0.09	-0.06	-0.03	-0.01
F74	-0.03	-0.03	-0.02	-0.02	-0.01	-0.01	0.00
total	-3.95	-2.72	-1.83	-1.30	-0.62	-0.23	-0.04

MLINK : DOMSSF DHFR

theta	0.00	0.01	0.05	0.10	0.20	0.30	0.40
F20	-1.23	-1.16	-0.92	-0.70	-0.38	-0.18	-0.06
F27	-3.24	-2.44	-1.65	-1.17	-0.57	-0.23	-0.05
F35	-2.91	-1.69	-0.91	-0.54	-0.20	-0.06	-0.01
F36	-0.31	-0.31	-0.27	-0.18	-0.03	0.02	0.01
F40	0.28	0.27	0.24	0.19	0.11	0.05	0.01
F41	-0.06	-0.05	-0.02	-0.01	0.00	-0.01	0.00
F74	-0.03	-0.03	-0.03	-0.03	-0.02	-0.01	0.00
total	-7.50	-5.41	-3.55	-2.41	-1.09	-0.41	-0.10

Table 1. Two point lods for DHFR (old families)

MLINK : DOMS 5HT1a

theta	0.00	0.01	0.05	0.10	0.20	0.30	0.40
F20	-0.81	-0.78	-0.67	-0.54	-0.34	-0.19	-0.08
F27	-0.02	-0.02	-0.02	-0.01	0.00	0.00	0.00
F35	-0.04	-0.04	-0.04	-0.04	-0.02	-0.01	0.00
F36	-0.31	-0.30	-0.25	-0.20	-0.10	-0.04	-0.01
F40	-0.18	-0.18	-0.18	-0.16	-0.10	-0.05	-0.01
F41	-0.16	0.60	1.09	1.17	0.97	0.59	0.20
F74	-0.05	-0.04	-0.04	-0.03	-0.02	-0.01	0.00
total	-1.57	-0.77	-0.10	0.20	0.38	0.29	0.09

MLINK : DOMSS 5HT1a

theta	0.00	0.01	0.05	0.10	0.20	0.30	0.40
F20	-0.93	-0.90	-0.76	-0.62	-0.39	-0.22	-0.09
F27	0.05	0.05	0.05	0.05	0.03	0.02	0.00
F35	-0.04	-0.04	-0.04	-0.04	-0.02	-0.01	0.00
F36	-0.31	-0.31	-0.31	-0.29	-0.21	-0.11	-0.03
F40	-0.19	-0.19	-0.18	-0.16	-0.11	-0.05	-0.01
F41	1.08	1.84	2.20	2.13	1.66	1.00	0.33
F74	-0.06	-0.06	-0.05	-0.04	-0.02	-0.01	0.00
total	-0.40	0.40	0.91	1.03	0.93	0.62	0.20

MLINK : DOMSSF 5HT1a

theta	0.00	0.01	0.05	0.10	0.20	0.30	0.40
F20	-0.62	-0.58	-0.43	-0.29	-0.12	-0.02	0.02
F27	0.03	0.03	0.04	0.04	0.03	0.02	0.00
F35	-0.04	-0.04	-0.04	-0.04	-0.03	-0.01	0.00
F36	-0.32	-0.32	-0.32	-0.30	-0.21	-0.10	-0.03
F40	-0.19	-0.19	-0.19	-0.18	-0.12	-0.06	-0.02
F41	0.42	1.29	1.77	1.79	1.46	0.91	0.30
F74	-0.08	-0.08	-0.07	-0.06	-0.04	-0.02	0.00
total	-0.82	0.11	0.75	0.97	0.99	0.71	0.27

Table 1. Two point lods for 5HT1a (old families)

MLINK : DOMS Jo110Hc

theta	0.00	0.01	0.05	0.10	0.20	0.30	0.40
F121	-0.01	-0.01	-0.01	-0.01	0.00	0.00	-0.00
F125	-0.01	-0.01	-0.01	-0.01	0.00	0.00	0.00
F141	-0.55	-0.49	-0.33	-0.21	-0.09	-0.03	-0.01
F143	-0.02	-0.02	-0.02	-0.02	-0.01	0.00	0.00
F184	0.10	0.09	0.08	0.07	0.04	0.02	0.00
F224	0.00	0.00	0.00	0.00	0.00	0.00	0.00
total	-0.49	-0.43	-0.28	-0.17	-0.06	-0.02	0.00

MLINK : DOMSS Jo110Hc

theta	0.00	0.01	0.05	0.10	0.20	0.30	0.40
F121	-0.01	-0.01	-0.01	0.00	0.00	0.00	0.00
F125	-0.01	-0.01	-0.01	0.00	0.00	0.00	0.00
F141	-0.82	-0.70	-0.42	-0.26	-0.10	-0.04	-0.01
F143	-0.02	-0.02	-0.02	-0.02	-0.01	0.00	0.00
F184	0.11	0.11	0.09	0.08	0.04	0.02	0.01
F224	0.00	0.00	0.00	0.00	0.00	0.00	0.00
total	-0.75	-0.62	-0.36	-0.21	-0.07	-0.02	0.00

MLINK : DOMSSF Jo110Hc

theta	0.00	0.01	0.05	0.10	0.20	0.30	0.40
F121	-0.01	-0.01	-0.01	0.00	0.00	0.00	0.00
F125	-0.04	-0.04	-0.03	-0.03	-0.02	-0.01	0.00
F141	-1.00	-0.81	-0.46	-0.27	-0.10	-0.04	-0.01
F143	-1.61	-1.53	-1.26	-1.00	-0.63	-0.37	-0.16
F184	0.11	0.11	0.09	0.08	0.04	0.02	0.01
F224	0.00	0.00	0.00	0.00	0.00	0.00	0.00
total	-2.55	-2.28	-1.66	-1.22	-0.71	-0.39	-0.17

Table 1. Two point lods for Jo110Hc (new families)

MLINK : DOMS p599(CA)

theta	0.00	0.01	0.05	0.10	0.20	0.30	0.40
F121	-2.54	-2.37	-1.57	-1.04	-0.51	-0.23	-0.07
F125	-0.01	-0.01	-0.01	-0.01	0.00	0.00	0.00
F141	-0.61	-0.60	-0.44	-0.32	-0.15	-0.06	-0.01
F143	-0.38	-0.31	-0.12	-0.02	0.05	0.04	0.01
F184	-1.91	-1.83	-1.41	-0.95	-0.40	-0.11	0.00
F224	0.00	0.00	0.01	0.01	0.01	0.01	0.00
total	-5.45	-5.08	-3.55	-2.32	-1.00	-0.36	-0.07

MLINK : DOMSS p599(CA)

theta	0.00	0.01	0.05	0.10	0.20	0.30	0.40
F121	-2.71	-2.49	-1.61	-1.06	-0.51	-0.23	-0.07
F125	-0.01	-0.01	-0.01	0.00	0.00	0.00	0.00
F141	-1.61	-1.27	-0.76	-0.48	-0.21	-0.08	-0.02
F143	-0.44	-0.44	-0.40	-0.32	-0.18	-0.07	-0.02
F184	-4.03	-3.31	-2.25	-1.52	-0.66	-0.22	-0.02
F224	-0.42	-0.40	-0.30	-0.22	-0.10	-0.04	-0.01
total	-9.21	-7.91	-5.32	-3.60	-1.66	-0.63	-0.13

MLINK : DOMSSF p599(CA)

theta	0.00	0.01	0.05	0.10	0.20	0.30	0.40
F121	-3.51	-2.92	-1.73	-1.12	-0.52	-0.22	-0.06
F125	-2.10	-1.54	-0.89	-0.56	-0.25	-0.10	-0.03
F141	-1.77	-1.44	-0.90	-0.58	-0.26	-0.10	-0.02
F143	-0.08	-0.08	-0.05	-0.03	0.00	0.00	0.00
F184	-5.84	-5.35	-3.50	-2.27	-0.97	-0.32	-0.03
F224	-1.97	-0.99	-0.33	-0.07	0.11	0.11	0.04
total	-15.26	-12.32	-7.40	-4.63	-1.89	-0.63	-0.10

Table 1. Two point lods for p599$(CA)_n$ (new families)

MLINK : DOMS M4

theta	0.00	0.01	0.05	0.10	0.20	0.30	0.40
F121	-1.41	-0.95	-0.43	-0.20	-0.02	0.04	0.04
F125	-0.02	-0.02	-0.02	-0.01	-0.01	0.00	0.00
F141	-0.54	-0.49	-0.33	-0.21	-0.08	-0.03	0.00
F143	-0.26	-0.20	-0.03	0.07	0.12	0.09	0.03
F184	-1.43	-0.89	-0.30	-0.04	0.16	0.18	0.12
F224	0.00	0.00	0.00	0.00	0.00	0.00	0.00
total	-3.67	-2.54	-1.10	-0.39	0.17	0.28	0.19

MLINK : DOMSS M4

theta	0.00	0.01	0.05	0.10	0.20	0.30	0.40
F121	-1.43	-0.94	-0.41	-0.18	-0.01	0.05	0.04
F125	-0.02	-0.02	-0.01	-0.01	-0.01	0.00	0.00
F141	-1.23	-0.96	-0.53	-0.31	-0.12	-0.04	-0.01
F143	-2.38	-1.50	-0.86	-0.58	-0.31	-0.16	-0.06
F184	-3.71	-2.10	-0.85	-0.33	0.08	0.19	0.15
F224	0.00	0.00	0.00	0.00	0.00	0.00	0.00
total	-8.77	-5.52	-2.67	-1.42	-0.36	0.04	0.12

MLINK : DOMSSF M4

theta	0.00	0.01	0.05	0.10	0.20	0.30	0.40
F121	-1.48	-0.91	-0.36	-0.13	0.04	0.08	0.06
F125	-0.06	-0.06	-0.03	-0.01	0.01	0.01	0.00
F141	-1.19	-0.94	-0.52	-0.31	-0.12	-0.04	-0.01
F143	-1.34	-0.93	-0.43	-0.23	-0.12	-0.10	-0.06
F184	-4.86	-3.27	-1.38	-0.60	0.02	0.21	0.18
F224	0.00	0.00	0.00	0.00	0.00	0.00	0.00
total	-8.93	-6.10	-2.72	-1.27	-0.17	0.16	0.17

Table 1. Two point lods for M4 (new families)

MLINK : DOMS EF(TG/AG)

theta	0.00	0.01	0.05	0.10	0.20	0.30	0.40
F121	-1.59	-1.42	-1.01	-0.70	-0.34	-0.14	-0.03
F125	-0.01	-0.01	-0.01	-0.01	0.00	0.00	0.00
F141	-0.06	-0.06	-0.05	-0.03	-0.02	-0.01	0.00
F143	0.25	0.25	0.23	0.21	0.14	0.08	0.02
F184	-0.01	-0.01	-0.01	-0.01	0.00	0.00	0.00
F224	-0.07	-0.07	-0.07	-0.06	-0.04	-0.02	0.00
total	-1.49	-1.32	-0.90	-0.59	-0.25	-0.09	-0.02

MLINK : DOMSS EF(TG/AG)

theta	0.00	0.01	0.05	0.10	0.20	0.30	0.40
F121	-1.62	-1.46	-1.05	-0.73	-0.35	-0.14	-0.03
F125	-0.01	-0.01	-0.01	-0.01	0.00	0.00	0.00
F141	-0.08	-0.08	-0.06	-0.05	-0.02	-0.01	0.00
F143	-0.28	-0.27	-0.22	-0.16	-0.08	-0.04	-0.01
F184	-0.01	-0.01	-0.01	-0.01	0.00	0.00	0.00
F224	-1.49	-1.33	-0.94	-0.65	-0.31	-0.13	-0.03
total	-3.49	-3.16	-2.28	-1.60	-0.78	-0.32	-0.08

MLINK : DOMSSF EF(TG/AG)

theta	0.00	0.01	0.05	0.10	0.20	0.30	0.40
F121	-1.69	-1.60	-1.22	-0.86	-0.42	-0.17	-0.04
F125	-0.35	-0.33	-0.28	-0.23	-0.13	-0.06	-0.01
F141	-0.13	-0.13	-0.10	-0.07	-0.04	-0.02	0.00
F143	-0.30	-0.28	-0.23	-0.17	-0.09	-0.04	-0.01
F184	-0.01	-0.01	-0.01	-0.01	-0.01	0.00	0.00
F224	-1.54	-1.47	-0.94	-0.50	-0.10	0.01	0.02
total	-4.02	-3.82	-2.78	-1.83	-0.77	-0.27	-0.05

Table 1. Two point lods for $EF(TG/AG)_n$ (new families)

MLINK : DOMS 6741(GT)

theta	0.00	0.01	0.05	0.10	0.20	0.30	0.40
F121	-1.61	-1.59	-1.33	-0.97	-0.49	-0.22	-0.07
F125	-0.07	-0.07	-0.06	-0.05	-0.03	-0.01	0.00
F141	-0.70	-0.63	-0.44	-0.29	-0.13	-0.05	-0.01
F143	-0.14	-0.13	-0.06	0.00	0.06	0.05	0.02
F184	-2.73	-1.88	-1.06	-0.60	-0.13	0.05	0.06
F224	-0.03	-0.03	-0.03	-0.03	-0.03	-0.02	0.00
total	-5.28	-4.33	-2.98	-1.93	-0.75	-0.20	-0.01

MLINK : DOMSS 6741(GT)

theta	0.00	0.01	0.05	0.10	0.20	0.30	0.40
F121	-1.64	-1.62	-1.35	-0.98	-0.49	-0.22	-0.07
F125	-0.07	-0.06	-0.06	-0.04	-0.03	-0.01	0.00
F141	-1.35	-1.07	-0.63	-0.39	-0.17	-0.06	-0.01
F143	-2.23	-1.37	-0.74	-0.48	-0.25	-0.14	-0.06
F184	-4.48	-3.57	-2.29	-1.48	-0.53	-0.07	0.06
F224	-2.11	-1.92	-1.43	-1.03	-0.52	-0.22	-0.05
total	-11.87	-9.61	-6.49	-4.40	-1.98	-0.72	-0.14

MLINK : DOMSSF 6741(GT)

theta	0.00	0.01	0.05	0.10	0.20	0.30	0.40
F121	-1.70	-1.68	-1.42	-1.03	-0.51	-0.22	-0.06
F125	-1.83	-1.79	-1.32	-0.86	-0.37	-0.14	-0.03
F141	-1.47	-1.14	-0.67	-0.42	-0.18	-0.07	-0.02
F143	-3.57	-2.40	-1.72	-1.37	-0.84	-0.45	-0.18
F184	-7.20	-5.70	-3.62	-2.37	-0.92	-0.23	0.01
F224	-2.59	-2.47	-1.78	-1.19	-0.55	-0.22	-0.05
total	-18.35	-15.17	-10.52	-7.23	-3.37	-1.33	-0.33

Table 1. Two point lods for 6741(GT)$_n$ (new families)

MLINK : DOMS YN(CT)

theta	0.00	0.01	0.05	0.10	0.20	0.30	0.40
F121	-1.61	-1.59	-1.33	-0.97	-0.49	-0.22	-0.07
F125	-0.02	-0.02	-0.02	-0.01	-0.01	0.00	0.00
F141	-0.57	-0.49	-0.30	-0.18	-0.07	-0.03	-0.01
F143	-0.16	-0.14	-0.08	-0.01	0.05	0.05	0.01
F184	-0.84	-0.83	-0.75	-0.59	-0.24	-0.05	0.00
F224	-0.14	-0.13	-0.10	-0.07	-0.03	-0.01	0.00
total	-3.34	-3.21	-2.58	-1.83	-0.79	-0.26	-0.07

MLINK : DOMSS YN(CT)

theta	0.00	0.01	0.05	0.10	0.20	0.30	0.40
F121	-1.64	-1.62	-1.35	-0.98	-0.49	-0.22	-0.07
F125	-0.02	-0.02	-0.02	-0.02	-0.01	0.00	0.00
F141	-1.47	-1.03	-0.52	-0.30	-0.12	-0.05	-0.01
F143	-2.26	-1.39	-0.76	-0.49	-0.26	-0.15	-0.07
F184	-3.50	-2.52	-1.39	-0.78	-0.19	0.02	0.04
F224	-1.51	-1.09	-0.54	-0.28	-0.05	0.02	0.01
total	-10.39	-7.67	-4.58	-2.84	-1.13	-0.39	-0.10

MLINK : DOMSSF YN(CT)

theta	0.00	0.01	0.05	0.10	0.20	0.30	0.40
F121	-1.70	-1.68	-1.42	-1.03	-0.51	-0.22	-0.06
F125	-2.09	-1.92	-1.15	-0.68	-0.27	-0.09	-0.02
F141	-1.87	-1.19	-0.58	-0.32	-0.12	-0.04	-0.01
F143	-3.68	-2.51	-1.81	-1.43	-0.86	-0.45	-0.18
F184	-4.07	-3.06	-1.99	-1.26	-0.42	-0.04	0.05
F224	-1.70	-1.68	-1.52	-1.15	-0.53	-0.19	-0.04
total	-15.09	-12.04	-8.47	-5.87	-2.69	-1.04	-0.26

Table 1. Two point lods for $YN(CT)_n$ (new families)

MLINK : DOMS 5HT1a

theta	0.00	0.01	0.05	0.10	0.20	0.30	0.40
F121	0.00	0.00	0.00	0.00	0.00	0.00	0.00
F125	-0.03	-0.03	-0.02	-0.02	-0.01	0.00	0.00
F141	0.16	0.16	0.14	0.11	0.07	0.03	0.01
F143	-0.03	-0.03	-0.03	-0.03	-0.02	-0.01	0.00
F184	0.20	0.19	0.17	0.14	0.09	0.04	0.01
F224	0.00	0.00	0.00	0.00	0.00	0.00	0.00
total	0.30	0.29	0.25	0.20	0.12	0.06	0.02

MLINK : DOMSS 5HT1a

theta	0.00	0.01	0.05	0.10	0.20	0.30	0.40
F121	0.00	0.00	0.00	0.00	0.00	0.00	0.00
F125	-0.03	-0.03	-0.02	-0.02	-0.01	0.00	0.00
F141	0.20	0.19	0.17	0.14	0.08	0.04	0.01
F143	-0.03	-0.03	-0.03	-0.02	-0.02	-0.01	0.00
F184	0.21	0.21	0.18	0.15	0.09	0.04	0.01
F224	0.00	0.00	0.00	0.00	0.00	0.00	0.00
total	0.36	0.35	0.30	0.25	0.15	0.07	0.02

MLINK : DOMSSF 5HT1a

theta	0.00	0.01	0.05	0.10	0.20	0.30	0.40
F121	0.00	0.00	0.00	0.00	0.00	0.00	0.00
F125	-0.10	-0.09	-0.06	-0.04	-0.01	0.00	0.00
F141	0.20	0.19	0.17	0.14	0.08	0.04	0.01
F143	-0.03	-0.03	-0.02	-0.02	-0.01	-0.01	0.00
F184	0.21	0.21	0.18	0.15	0.09	0.04	0.01
F224	0.00	0.00	0.00	0.00	0.00	0.00	0.00
total	0.28	0.28	0.26	0.23	0.15	0.07	0.02

Table 1. Two point lods for 5HT1a (new families)

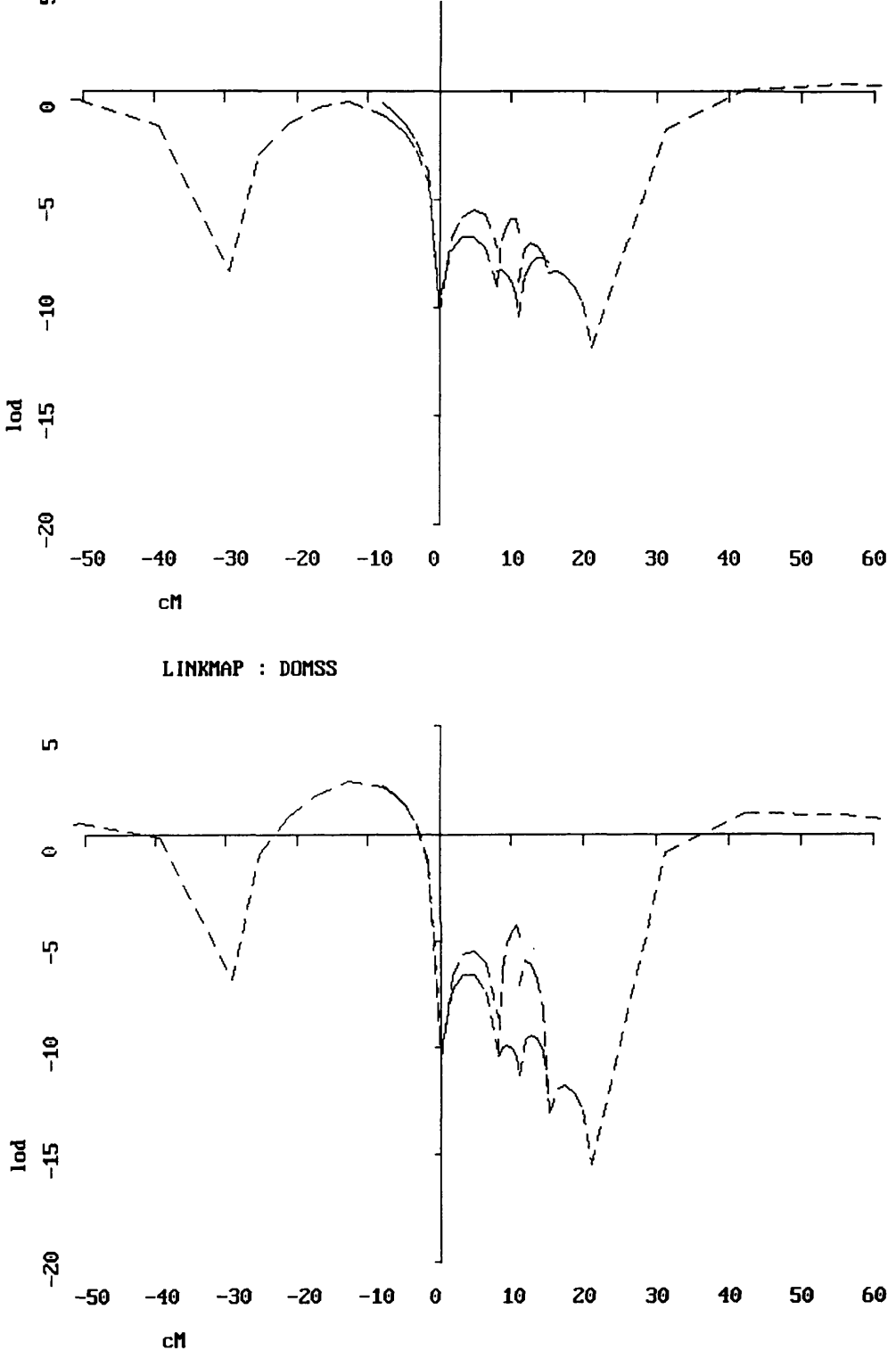


Figure 4. LINKMAP analyses of the original family cohort. Overlapping 4 point graphs for DOMS and DOMSS. D5S20 is positioned at -29cM, D5S21 at -8cM, D5S76 at 0cM, D5S6 at 8cM, D5S125 at 11cM, D5S39 at 15cM and D5S127 at 21cM.

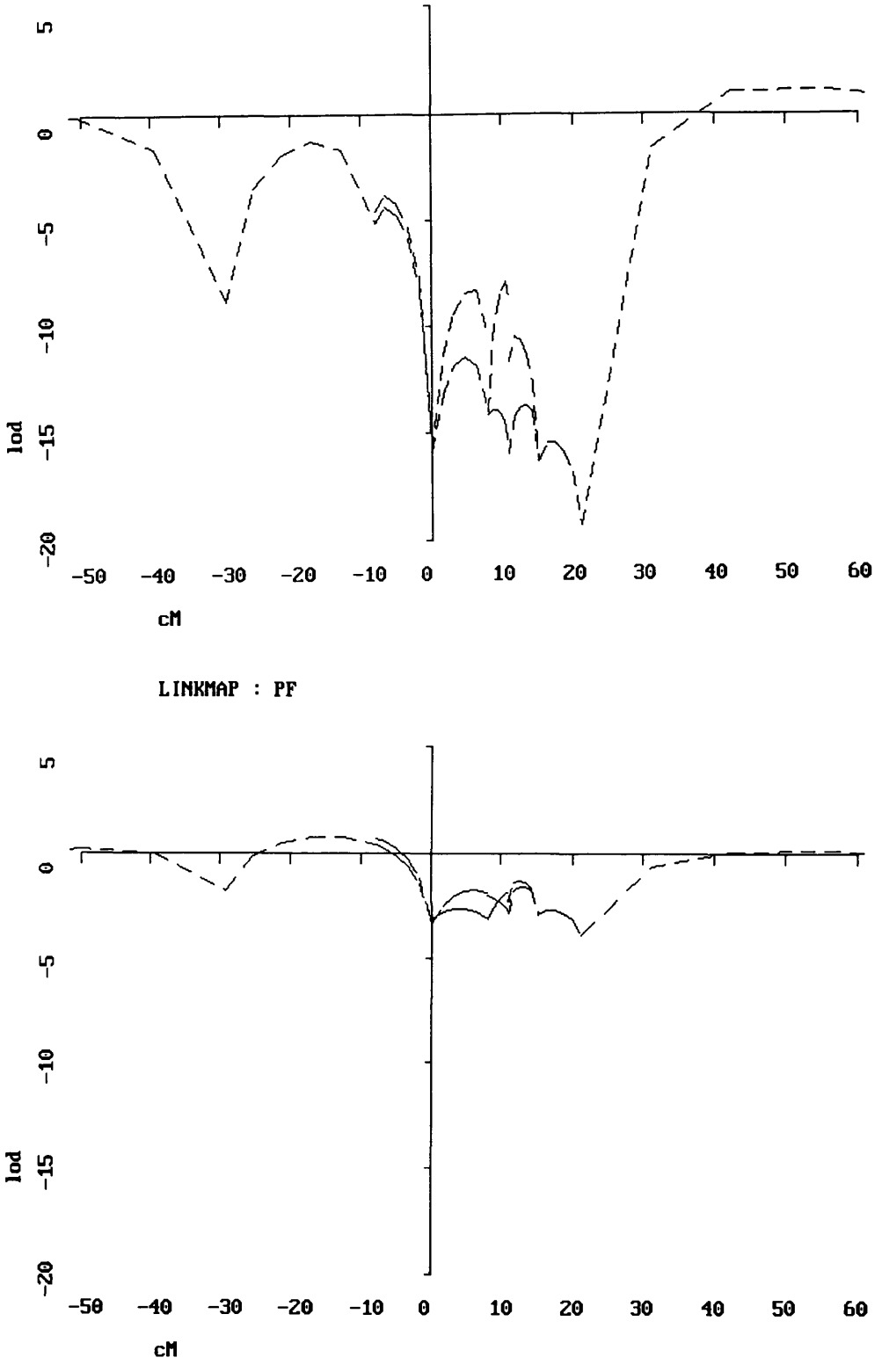
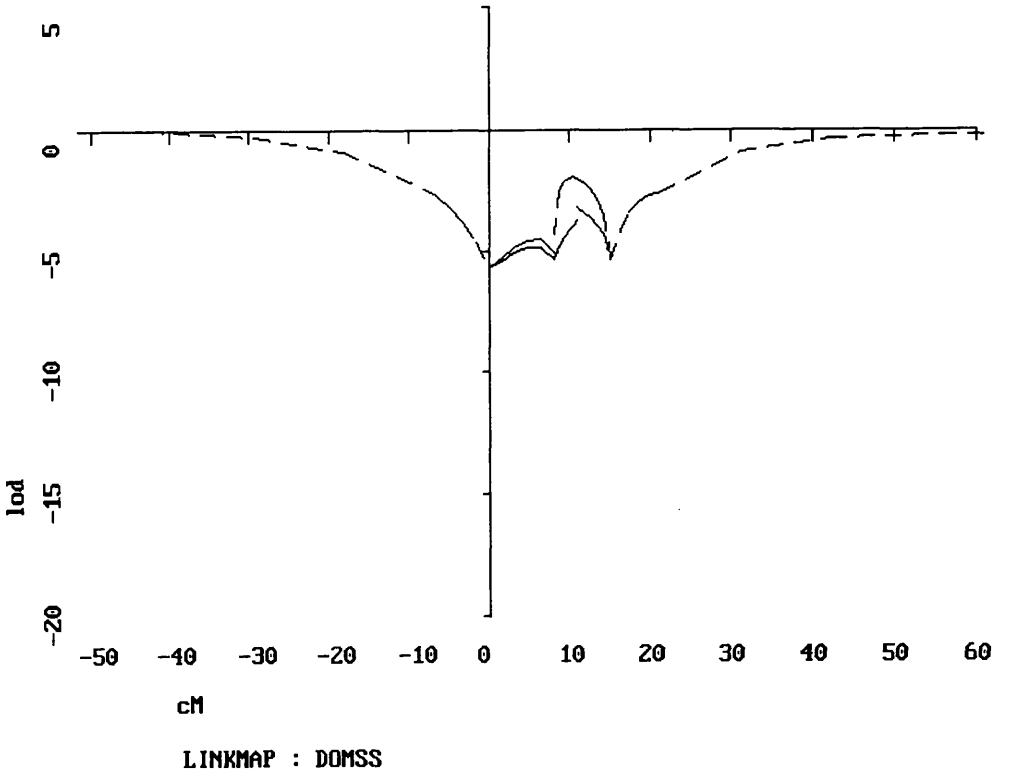


Figure 4 (continued). LINKMAP analyses of the original cohort of families with the DOMSSF and PF models.



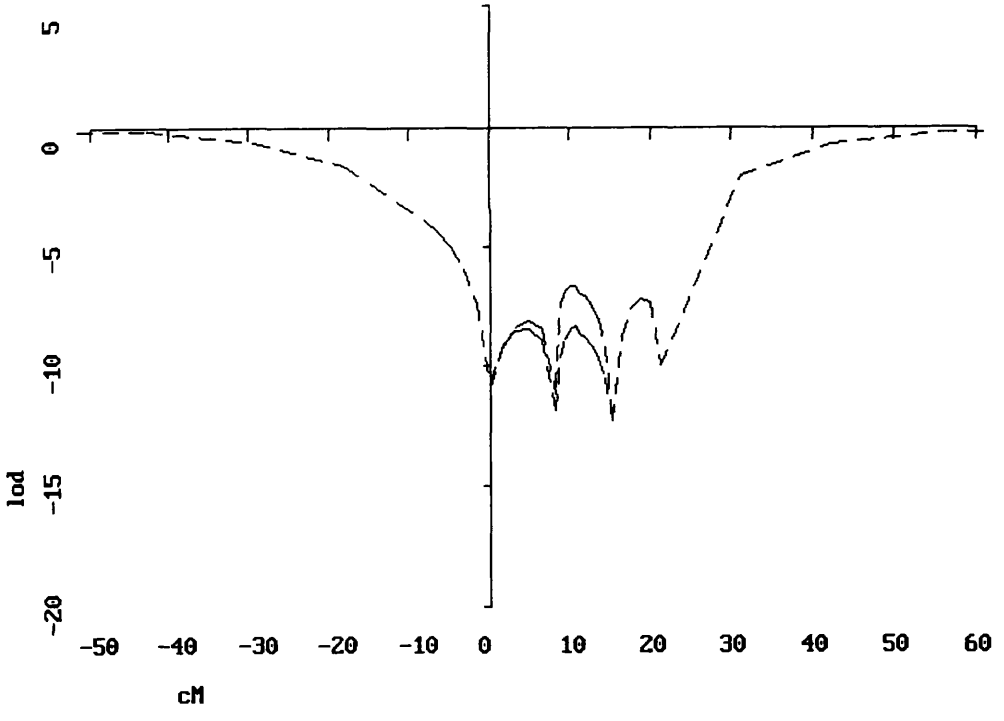


Figure 4 (continued). LINKMAP analyses of the new Icelandic cohort of families for the DOMS and DOMSS models. No information at D5S20 is shown.

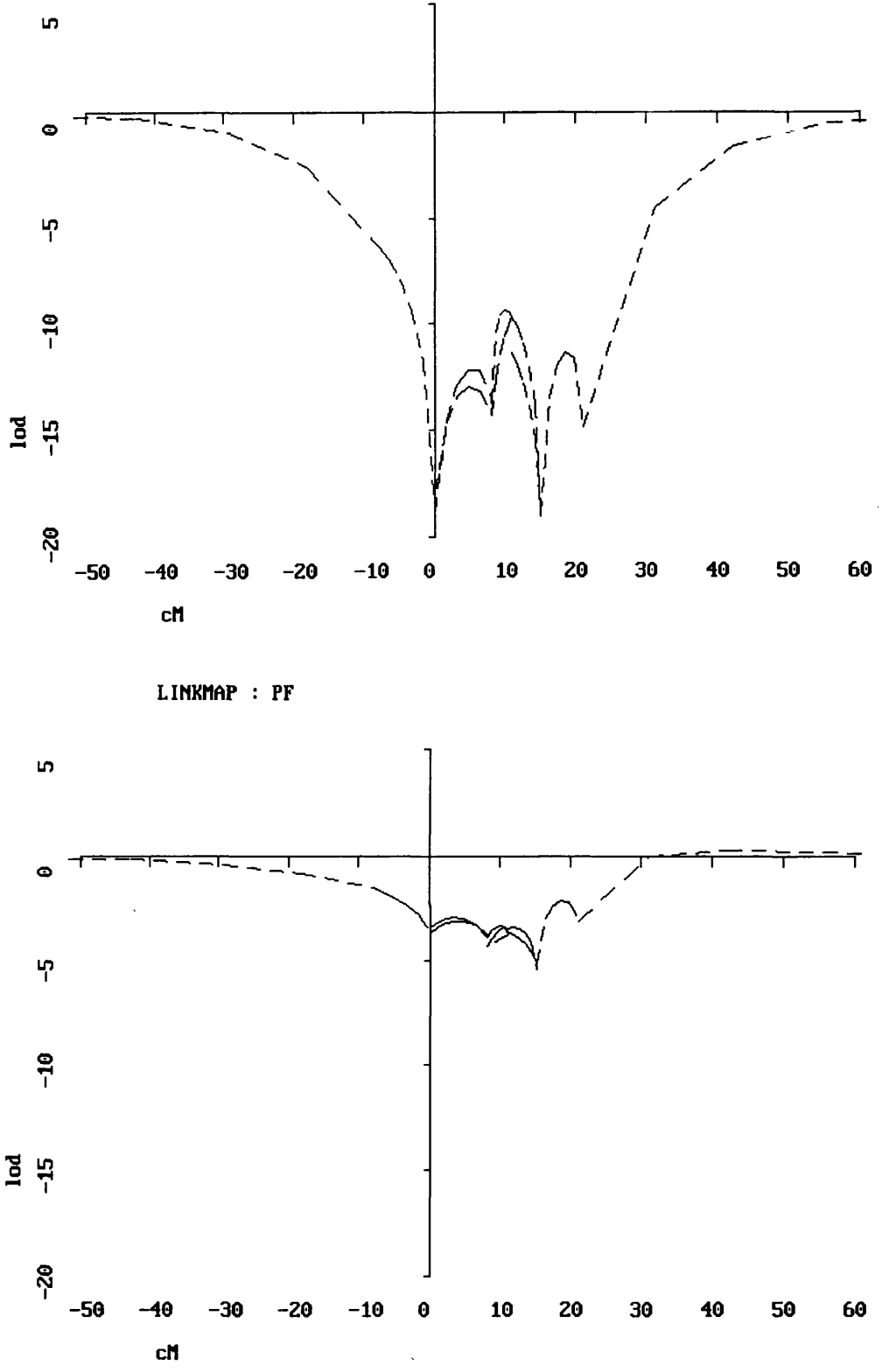


Figure 4 (continued). LINKMAP analyses of the new Icelandic cohort of families for the DOMSSF and PF models.

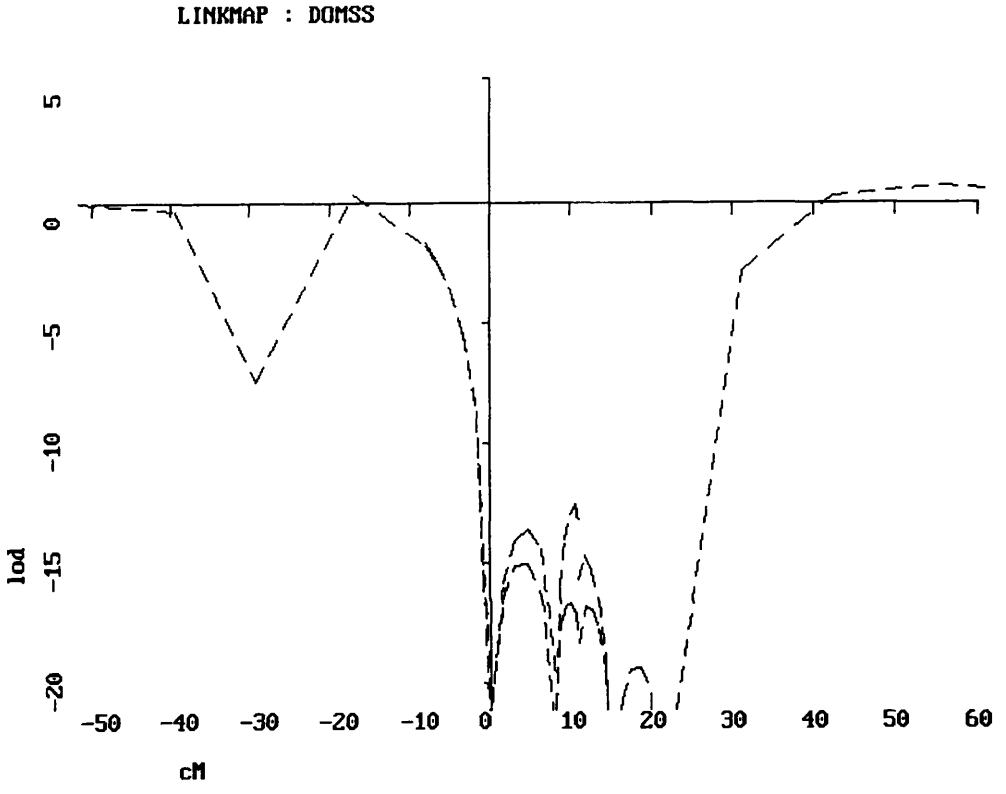


Figure 4 (continued). LINKMAP analysis for all the families combined for the DOMSS model.

combined for the DOMSS model. The map positions for the markers and their relative distances were based on the maps of Weissenbach et al (1991), and for the microsatellites Morrison et al (1992). The results are shown in Figure 4. DNA extracted from formalin-fixed paraffin-embedded tissue samples is an excellent substrate for PCR amplification of polymorphic microsatellites. All tissues, ranging from 7 to 18 years of age showed detectable amplification indicating that old samples are suitable for this technique. It was observed that some preparations gave no signal after one round of amplification, these were subjected to a further round of amplification before a signal was observed. There was no general correlation between the quality of signal that a sample produced when different primer sets for different loci were used. Typical results are presented in Figure 5; which shows the results obtained from PCR amplification of the p599$(CA)_n$ microsatellite. Each individual was genotyped by testing both liver and brain sections. In all instances both tissues produced identical genotypes confirming that the signal represents the true genotype of each sample and not a product of contamination. Furthermore, all negative controls gave no detectable signal. The alleles for each paraffin-embedded tissue for each microsatellite are shown in Table 2. Of the four markers tested, one set of primers, 6741(GT)n produced poor amplification. Only three individuals produced a signal on autoradiography, and it was not possible to ascertain the alleles due to a high background of non-specific bands. Also, one individual (F125/9) gave no signal for the EF(TG/AG)n marker.

6.4 DISCUSSION

PCR has been used extensively in forensic pathology to identify the source of blood, hair and sperm. There has been rapid progress in the application of PCR to the diagnosis of infectious diseases, sex determination of human embryos after in vitro fertilisation, prenatal diagnosis oof genetic disorders, and HLA analysis and tissue typing for organ transplantation (reviewed by Eisenstein, 1990). PCR has also been usefully applied to detect the presence of human papilloma virus DNA in contemporary pparaffin bloockks (Imparin et al 1987) and 40 year old paraffin-embedded sections (Shibata et al 1988). Wright and Manos (1990) have described the successful amplification of human papillomavirus L1 products of the size range 450 to 650 base pairs. In all these studies the target DNA for amplificatin was present in multiple copies. Our analysis shows that, despite probable degradation, amplification of single copy polymorphic DNA markers is also possible from old paraffin blocks. The increase in the number of cycles used to generate visible bands produced a corresponding increase in the background of non-specific bands. It may

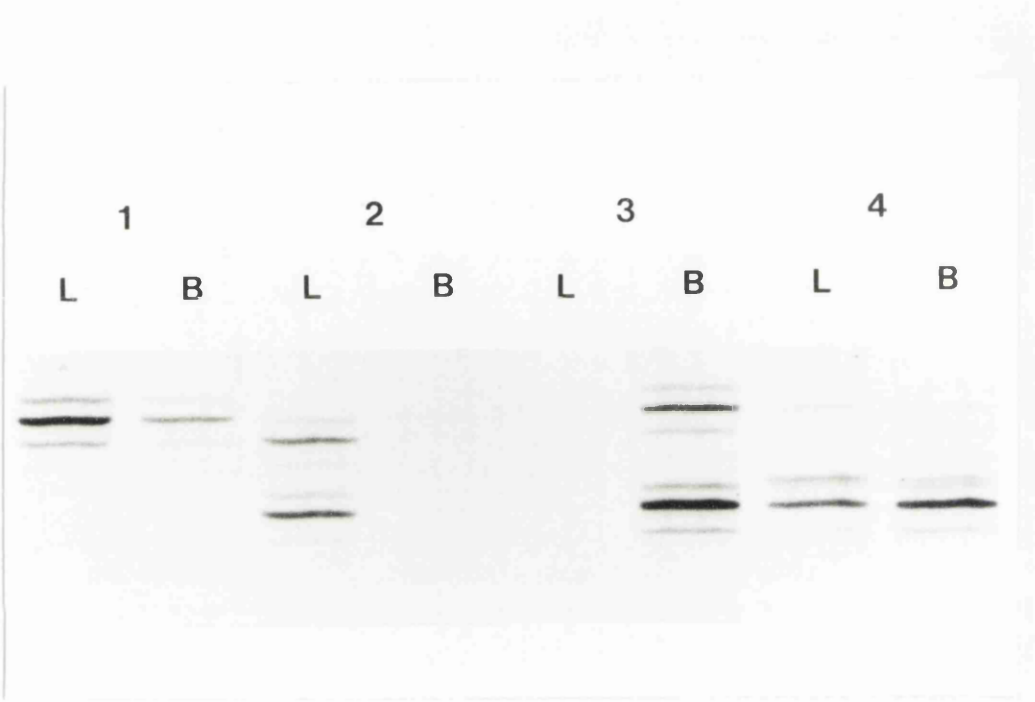


Figure 5. Autoradiograph of the product of PCR amplification of parraffin embedded tissue with the p599$(CA)_n$ microsatellite.

Individual	p599(CA)$_n$	EF(TG/AG)$_n$	YN(CT)$_n$
F27/48	1 1	a a	3 4
F41/29	3 5	a a	3 5
F125/8	2 6	a a	1 3
F125/9	6 6	- -	3 3
F141/8	3 6	a b	3 3

Table 2. Genotypes for the paraffin-embedded tissue samples.

be poossible to reduce this background by varying annealing temperature and annealing/polymerisation times, salt concentration, or by designing primers that produce more precise amplification. For these microsatellites it is noticeable that reducing the magnesium concetration to 1mM helps to reduce background.

One of the microsatellite markers, 6741$(GT)_n$, gave poor results and the genotypes of individuals could not be ascertained. It may be significant that the predicted allelles for 6741$(GT)_n$ are much larger (greated than 200 nucleotiddes) than for the other loci (less than 150 nucleotides). If the average size of DNA present in these old paraffin blocks was less than 200 base pairs this could explain the lack of successful amplification in samples tested with 6741$(GT)_n$. This could limit the testing of formalin-fixed tissue to only those microsatellites with short products. The reason why the formalin-fixed tissue should undergo degradation is unknown. It has been suggested that this could be due to acid depurination by unbuffered formalin (Shibata et al 1988). Alternatively it may reflect the nuclease activity in the original tissue. Preservation techniques that maintain the integrity of DNA will significantly enhance the potential of archival material for research. The microsatellite YN$(CT)_n$ maps distally to that of 6741$(GT)_n$. Using the genetic information obtained from the formalin-fixed tissue with this microsatelite in a multipoint analysis with the other microsatellites potentially supersedes the loss of information with the 6741$(GT)_n$ microsatellite. The availability of archival material will be of use in the study of the genetics of psychiatric disorders. Families are usually of small size and extant members of a pedigree may not provide sufficient lod scores. This paucity of large informative pedigrees, whilst not unique to psychiatric genetics, does generate considerable problems. The analysis of tissue of deceases individuals with highly polymorphic microsatellites is a quick, inexpensive method to obtain valuable genetic inforamtion.

The inclusion of the new microsatellite data, which gave appproximately 90% informativity at the two loci used in the original study in these families, dramatically reduced the evidence for a dominant schizophrenia susceptibility locus. No statistical evidence for non-allelic heterogeneity of linkage was found (data not shown, personal communication Dr D Curtis, HOMOG, B-Test, Ott 1986; Reish 1988) in the pedigree set, and by summing the lods across the pedigrees no evidence of linkage was found for any of the models (Table 1).

The markers between and including DHFR and D5S6 are known to be within the

candidate trisomic region (Gilliam et al 1989). Multipoint analysis using the microsatellites and other markers covering this region (Figure 4) in the original families not only demonsyratedno evidence of linkage but excluded the trisomic region for a dominant susceptibility allele for schizophrenia. However, the locus D5S21 on the p arm of chromosome 5 but linked to D5S76 produced a lod approaching 3 for the DOMSS model, as a consequence a dominant susceptibility allele for schizophrenia and spectrum disorders could not be excluded at this locus. Linkage analysis with the new cohort of Icelandic families with the microsatellites and RFLP's also demonstrated no evidence of linkage for this region of chromosome 5 for all the classification models (Table 1 and Figure 4). The exclusion for these families also included the marker D5S21, combining all the data provided no evidence of a dominant susceptibility allele for schizophrenia.
Interestingly one pedigree F41 on its own has a Lod of 2.49 with D5S76 and 2.2 with the Taq1 RFLP for the candidate 5HT1a receptor gene. Nevertheless there is a clear affected recombinant (F41/5) for both of these markers. So it is unlikely that this gene is of aetiologic importance in this family alone, but it cannot be ruled out (Table 1). Also archival material for one affected individual from an extention of this pedigree (F41/29) demonstrated that a second recombination event between the microsatellite marker for D5S76 and schizophrenia has occurred, since allele 4 co-segregates with schizophrenia in F41 and this individual has the genotype 3 5 for the microsatellite at D5S76.

It is reasonable to conclude that the original positive study was a type I statistical error (false positive). Quite why this should occur is difficult to discern, but in part it must have been caused by the assumptions and extra variables in the linkage analysis. Multiple testing utilising a number of different affection classes and penetrance values will increase the probability of creating a false positive linkage. To quantify this bias so that it can be generally applied in the linkage analysis will be difficult to achieve. It seems practical that a modest increase of the level of significance for linkage is required and that any such result should be considered as possible linkage or a candidate region until confirmed by an independent group would be most appropriate for linkage analysis in psychiatric disorders. Indiscriminately raising the level of significance for linkage for complex analyses will increase the chance of a type II statistical error.

At this juncture the type I error does not imply that the method of linkage analysis

will not be applied successfully in psychiatric genetics. It is still probably the best approach to understanding the true aetiology, as demonstrated by the recent successes for Alzheimers disease (Goat et al 1991).

Early studies investigating the inheritance of schizophrenia estimated that between 66% to 93% of the variance in the aetiology is genetic (Fulker 1973, Kendler 1983, McGue et al 1985). Segregation analysis by studying the patterns of affected individuals within families has not been able to establish a common mode of transmission (Baron 1986). This probably represents a failure of the segregation analysis to adequately model all the variables necessary for a disease which is highly likely to be aetiologically heterogeneous (Tsang et al 1990, Dalem et al 1990) and does not preclude the presence of major gene effects with near Mendelian modes of inheritance causing subtypes of schizophrenia. The advent of DNA markers through recombinant technology represents a renaissance for schizophrenia genetics allowing the search for major susceptibility loci contributing to the liability to develop the disease. These techniques also enable candidate genes such as those of the dopaminergic system to be studied for a genetic contribution to the disease. Linkage analysis (Ott 1985) or the less powerful identity by descent analysis (sib-pair) of marker data obtained from families is considered to be the most efficient and robust approach (Bodmer 1984).

Linkage studies in schizophrenia can be directed towards regions of the genome where there are a priori grounds for the presence of a susceptibility locus. The study under discussion here was encouraged to look at chromosome 5 after a report of an association (though weak) between a partial trisomy of chromosome 5q 11.2-13.3 and schizophrenia (Bassett et al 1988).

Families containing cases of bipolar affective disorder and assortive matings were excluded from the study to reduce the likelihood of a second gene causing a psychotic illness in a family, as detailed in chapter 2. In the original study seven families fulfilled these stringent criteria five from Iceland and two from Britain. In particular Iceland offered the potential to study a genetic isolate which might provide a more genetically homogeneous sample of families than else where. Of 104 individuals interviewed, 39 fulfilled RDC diagnosis for schizophrenia, five were diagnosed as having schizophrenia spectrum personality disorders, while ten additional family members received diagnosis including major and minor affective disorder, alcoholism and phobic disorder. These latter diagnoses are not generally believed to be genetically related to schizophrenia ("fringe" cases). Two RFLPs mapping to this area of chromosome 5 were studied for co-segregation with

schizophrenia (D5S76 and D5S39). Prior to this study the problem of having a range of diagnoses and penetrance levels in the linkage analysis had not been considereda and after discussions with E.S. Lander (Whitehead Institute Boston) and K.K. Kidd (Yale University New Haven) it was decided that three levels of affection states should be used and the lods maximised over a range of penetrance levels. Weak evidence of linkage (lod = 2.45) was found when the analysis was restricted to cases with shizophrenia alone with the "unaffecteds" diagnosed as phenotype unknown. Although this discarded a large fraction of the information it was robust to the influences of penetrance. Using the same model but classifying the unaffecteds as phenotypically normal gave a maximum lod of 3.22 at 99% penetrance. However the degree of co-segregation increased when the spectrum cases were included in the schizophrenia phenotype (lod = 4.33 at 99% penetrance). This is not unexpected since these disorders are known to be genetically related to schizophrenia. More surprising was the observation that the evidence for linkage further increased when the fringe cases were included (lod = 6.49 at 86% penetrance), but only five of the fringe cases were informative for linkage so this could represent nothing more than a chance finding. These results were interpreted as demonstrating the existence of a dominant susceptibility gene on chromosome 5 for schizophrenia. Furthermore this defect seemed to predispose to schizophrenia spectrum disorders and possibly also to a variety of other psychiatric conditions.

Although no evidence for heterogeneity was demonstrated with the HOMOG program (Ott 1986) between these seven families in the initial cohort certain caveats need to be stated on the validity of these results for schizophrenia as a whole, since the majority of the lod was obtained from three families (F41, F35, F20). The putative disease locus on chromosome 5 may represent only one cause of schizophrenia. Additional genes on other chromosomes may exist and non-genetic types of the illness cannot be ruled out. The selection of families with a large proportion of affected cases and primarily from Iceland may have enriched for a particular genetic subtype of schizophrenia. One might, at least in hindsight, question the use of relatively high penetrance values for a psychiatric disorder. This may have given undue weight to the information provided by unaffected individuals, falsely elevating lods. The analysis with varying affection classes and penetrance levels also add extra degrees of freedom. Multiple analyses increased the probability of a false positive finding and it would be wise to increase the lod score taken to indicate linkage. To account for this with these families Weeks et

al (1990) ran 2000 simulations with a 4 allele system assuming no linkage, their conclusions were that an increase to 3.66 for the schizophrenia and spectrum cases model would be applicable. The lod obtained still seemed significant in the study. Finally all linkages have to remain as only preliminary findings until replicated by another independent sample.

Numerous investigators promptly examined chromosome 5q markers in other schizophrenia family samples, but all obtained negative results (Kennedy et al 1988; Detera-Wadleigh et al 1989; St.Clair et al 1989; Kaufmann et al 1989; Aschauer et al 1990; Crowe et al 1990; McGuffin et al 1990). This discrepancy may represent true genetic heterogeneity, or a type I statistical error. In any event, initial linkage findings may be weak and difficult to replicate. For example, a study might identify a locus that ultimately proved to exert a major locus effect but only in certain samples or in a small proportion of cases, so that the linkage result would be difficult to replicate in independent samples. Alternatively an investigator might find a unique pedigree or geographical region where the occurrence of a polygenic disease depended on one or two loci simply because there is considerable homozygosity at the other relevant loci, a result that could be difficult to confirm. Alternatively differences in clinical methods could account for the discrepant findings. It is of interest that two of the studies contained families with Bipolar affective disorder (Detera-Wadleigh et al 1989; St.Clair et al 1989) even though twin, adoption and family studies have shown no evidence that this disorder is related to shizophrenia. However exclusion of these families from the analysis still provided no evidence for linkage. These two studies and those of Crowe et al (1990) and Aschauer et al (1990) contained cases of schizoaffective disorder which was not divided into schizophrenic or affective subtypes, consequently it is difficult to comment on these clinical differences. It is plausible that schizophrenia linkage studies could be confounded by families with severe variants of affective psychosis, or that assortative mating has resulted in diverse, and sometimes genetically mixed disorders within the pedigree. At this juncture it is important to comment that schizophrenia classification is still in evolution (Farmer et al 1991) and that the boundaries set for the diagnosis of schizophrenia remain arbitary until they can be validated either genetically or through other aetiological approaches. The problem of deciding exactly which definition for the phenotype of schizophrenia should be used should eventually be solved with successful linkage analysis. However for the time being analysing those individuals with schizophrenia and spectrum disorders should be robust enough to detect a linkage. Incomplete

penetrance, uncertain levels of heterogeneity, the presence of phenocopies and an unsure mode of inheritance incorporated into linkage analysis are likely to produce false negative results rather than false positive linkages. Two studies provided statistical evidence for heterogeneity. Kennedy et al (1989) refuted homogeneity for the North Swedish pedigree and the positive Icelandic and British study, and McGuffin et al (1990) between all the published data, heterogeneity ultimately requires the demonstration of another locus to be proven. However as stated earlier removal of the positive lod families from the McGuffin study provided no evidence of heterogeneity. This result was interpreted as evidence that the positive study was due to multiple testing. The possibility that the quasi-significant lod from the Icelandic/British study may still represent genetic heterogeneity could be tested by increasing the informativeness of the markers used, secondly a new cohort of families from Iceland could be studied.

The aim was to screen genomic cosmid and lambda libraries with the known mapped markers to expand these loci and detected new polymorphisms to increase their PIC values. Cosmid clones were only obtained for G21 (probably a consequence of using an over amplified library), however lambda clones were obtained for D5S76, D5S39, D5S6 and HEXB. Potential new RFLPs were also investigated, but at this time the PCR had allowed the development of several new types of highly polymorphic marker collectivelly called microsatellites (Weber 1990). Highly polymorphic microsatellites were found for D5S39 and D5S76 and a low PIC value microsatellite for HEXB. Additionally two polymorphic microsatellites were extracted from cosmids for D5S127 and D5S125 known to be linked to the previous markers.

Linkage analysis with these new markers excluded the region of the trisomy in the original cohort of families with only one family showing an interesting lod of 2.2 (F41) with the 5HT1a receptor Taq I RFLP. However this family had a clear recombinant individual with a diagnosis of schizophrenia so it is unlikely that this gene plays a role in the disease even for this family. A non-significant lod was seen with markers proximal to D5S76, but a new cohort of Icelandic families excluded this region for a dominant susceptibility allele. This new cohort using the new microsatellites also excluded the region of the trisomy and the region covered by the original positive lod. In conclusion the original significant lod was probably a type 1 statistical error caused to some degree by multiple testing. However, this negative finding does not invalidate the pursuit of genetic marker strategies in

studying the genetic basis of schizohrenia. Despite the variety of methological problems a systematic search throughout the human genome can now be justified as a means of detecting major gene effects in schizophrenia (Sturt and McGuffin 1985).

With the complete genetic map now available for the human genome (HGM 11 1991) a systematic search for linkage in schizophrenia is feasible. The use of the new type of micosatellite markers will help ameliorate the problems of heterogeneity by increasing the amount of information gained from the sample of pedigrees as a whole, allowing greater comparisons between pedigrees for heterogeneity. Guidelines for the numbers of pedigrees required to detect linkage assuming heterogeneity have been simulated (Boehnke 1990). In fact it could be postulated that heterogeneity could be beneficial, increasing the prior probability that any random marker may be linked to one of the many potential loci thereby allowing a theorectical reduction in the level of significance of the lod. However in practise this would be unwise given the caveats listed in previous chapters. The problems of the extra degrees of freedom introduced by varying penetrance, affection status, and phenocopy rate (Clerget-Darpoux 1991) in the linkage analysis can be accounted for by increasing the level of significance. However this may increase the likelihood of type II statistical errors to an unacceptable level. A number of improvements have been proposed which should increase the power of linkage analysis. The methods of interval mapping and simultaneous searches can reduce the numbers of families needed to detect linkage and heterogeneity (Lander and Botstein 1986). Alternatively a dummy quantitative variable can be used to deal with multiple affection categories in genetic linkage analysis (Curtis and Gurling 1991). The adverse affects of varying penetrance can be overcome by limiting the analysis to affected individuals only, however this will reduce the power of the analysis. Genetic subtypes of schizophrenia may also be identified by studying a variety of vulnerability traits and their co-segregation in multiplex schizophrenic pedigrees (Baron 1987). At present this approach is limited. To identify genes affecting the risk of developing schizophrenia or minor genes say for a polygenic mode or oligogenic of inheritance the improvements to the affected-sib-pair method of linkage analysis could be used (Weeks and Lange 1988; Amos et al 1990; Bishop and Williamson 1990). The disadvantage of this method is a loss of power compared to classical linkage analysis, and the problem that heterogeneity between families cannot be demonstrated very easily.

For a successful linkage strategy it is probably best to conceptualise schizophrenia as not one disorder but many with similar phenotypes. "The schizophrenias" will have more than one genetic and environmental cause. The genetic subtypes will be further subdivided by different modes of inheritance, different epistatic genetic effects and environmental factors, such as child birth in peuperal psychosis. It now seems feasible that with a large enough sample of families susceptibility alleles will be identified for the schizophrenias. An example of such a study is the European Science Foundation network in the molecular neurobiology of mental illness.

Other human diseases with some if not all of these complicating factors have been mapped using linkage. A clinically defined entity amoung the progressive myoclonus epilesies has been mapped by linkage to chromosome 21 q22 (Lehesjoki et al 1991). Two loci have been identified for the autosomal dominant disorder Tuberous Sclerosis on chromosomes 9 q34 (Fryer et al 1987) and 11q14-23 (Smith et al 1990). Linkage analysis in mice has identified two potential loci for non-obese diabetes (Todd et al 1991), unfortunately a mouse model for schizophrenia is unlikely. A gene in the insulin-IGF2 region on chromosome 11p has been identified for HLA-DR4-dependent diabetes susceptibility (Julier et al 1991). Diabetes has genetic and environmental subtypes with a number of loci contributing to the risk, however potential suceptbility loci are being identified, for example close linkage between the candidate gene glucokinase on chromosome 7 to early onset non-insulin dependent diabetes mellitus (Froguel et al 1992). Linkage has also been demonstrated between the common inherited disorder Familial combined hyperlipidaemia (prevalence of 0.5%-2%) and the apolipoprotein gene cluster on chromosome 11q23-24 (Wojciechowski et al 1991) following a report of an association with this cluster. Although the genetics of hypertension generally seem to favour polygenic inheritance, a mutation consisting of a chimaeric 11 beta hydroxylase and aldosterone synthase gene causes a mendelian subtype of the disorder (Lifton et al 1992). The autosomal dominant form of Retinitis Pigmentosa is both clinically and genetically heterogeneous. Subdivded into two groups based primarily on age of onset, there is considerable within-family variation in expression, variable penetrance, and late age of onset cases. Nevertheless using single large multiplex families in three separate studies three loci have been identified. Chromosome 3 (Farrar et al 1990), chromosome 8 (Blanton et al 1991) and chromosome 6 (Farrar et al 1991) this demonstrates the advantage of using large enough pedigrees to show a significant lod on their own in mapping genetically heterogeneous diseases. Alzheimer's disease is also genetically heterogeneous (St

George-Hyslop 1990) with some arbitarily chosen early age of onset families (< 65) showing linkage to chromosome 21. The amyloid precursor protein (APP), a strong candidate gene for Alzheimer's disease, was mapped to the same region of chromosome 21. APP is the precursor of amyloid beta peptide the major component of amyloid plaques. This candidate was originally excluded by the demonstration of recombinants in a few families (Van Broeckhoven et al 1987) but subsequent families which were compatible for linkage with APP identified rare mutations causing the disease (Goate et al 1991; Chartier-Harlin et al 1991). Over zealous exclusion of strong candidate genes in this case was not wise for a genetically heterogeneous diseases. It seems plausible that susceptibility loci for schizophrenia will eventually be identified. However the cloning of such genes present another set of obstacles. Below, these have been considered in relation to other disease genes already cloned by reverse genetics. The genetic diseases chronic granulomatous disease (CGD), Duchenne muscular dystrophy (DMD) and retinoblastoma were all found to be associated with cytogenetically identifiable chromosomal deletions. The use of hybrid cell lines containing these deletions enabled the selection of genomic or cDNA clones of candidate genes. The first clues to the identification of the affected locus in DMD came from the observation that females suffering from DMD have balanced X:autosome translocations with breakpoints within Xp21. These females are affected because the normal X chromosome is preferentially inactivated. Subtractive hybridisation experiments using a patient with a large deletion at Xp21 (Kunkel et al, 1985) were used to isolate DNA sequences form normal chromosomes which overlapped this deleted region. Sequences conserved across many mammalian species were selected and used to identify a large messenger RNA in foetal muscle cells (Monaco et al, 1985). The rationale for this was that a gene conserved throughout evolution and expressed in muscle would be an obvious candidate for the DMD gene. This RNA species has subsequently been identified as the mRNA (Koenig et al, 1987) for dystrophin which is the defective protein in DMD (Hoffman et al, 1987). Clones from the Xp21 region allow the detection of deletions in approximately 50% of DMD patients, ranging in size from six thousand to one million base pairs (den Dunnen et al, 1987). For CGD a B cell line was created from a patient who exhibited CGD and a cytogenetic deletion at Xp21 to which the disorder was linked. Subtractive screening using messenger RNA from this cell line with complementary DNA from a normal B cell line identified a phagocyte specific mRNA that was absent from some patients and which exhibited deletions in others (Royer-Pokora et al, 1986). These workers demonstrated the protein to be a member of

the cytochrome family. A similar methodology has been successful in identifying the molecular basis of retinoblastoma. Dryja et at (1986) demonstrated homozygous deletions greater than 25 kilobases (kb), at 13q14 in two out of 37 retinoblastoma tumours. cDNA clones of the gene encoding esterase D, which is linked to the retinoblastoma susceptibility locus in band 13q14.11, were used to isolate genomic clones of the esterase D locus. Distal DNA fragments of these clones were in turn used to isolate overlapping genomic clones. This chromosome walking technique allowed the isolation of DNA sequences over 500 kb. These were used to isolated cDNA clones from foetal retinal and placental libraries (Lee et al, 1987). Friend et al (1986) described a cDNA clone with properties attributed to the retinoblastoma gene. This clone detected a transcript absent specifically from retinoblastoma cell lines. Fung et al (1987) have shown that 60% of retinoblastomas tested contained no deletion in the retinoblastoma gene, but have aberrant RNA expression.

The contribution of cytological deletions that pinpointed the chromosomal location of each gene cannot be overstated. These deletions served not only to localize the genes, but were essential reagents for rapid cloning of the genes. Deletions have also contributed to the cloning of some genes for inheritable cancers. Inherited cancers normally account for a small proportion of the cases of that cancer, and are complicated by incomplete penetrance, and variable clinical phenotypes. Deletion mapping in individuals using alleles of a marker between normal lymphocyte DNA and the tumor cells DNA has proved invaluable in localising potential candidate regions. Pulse-field gel analysis and chromosome jumping isolated potential genes for the Wilm tumor gene (Gessler et al 1990; Rose et al 1990). One of these was found to be included in two partially overlapping homozygous deletions found in Wilms tumour DNA samples. Familial Adenomatous Polyposis Coli gene (FAP) was linked to chromosome 5 markers after chromosome 5 was implicated because of somatic cell mutations of chromosome 5q were observed. Two individuals narrowed the search for the gene to 100kb through deletion mapping, three candidate genes were examined for mutations in 61 unrelated patients identifying one as the gene for FAP and sporadic colon cancers (Groden et al 1991). Neurofibromatosis type 1 was assigned to chromosome 17 by linkage in affected pedigrees (Barker et al 1987), but the use of two patients with translocations of 17q11.2 focused the search for the gene to 60kb region and led to its eventual discovery (Viskochil et al 1990). A number of cytogenetic abnormalities have been reported in schizophrenic patients and are obvious

candidate regions for susceptibility genes (Bassett 1991). They also would facilitate the cloning of candidate genes from these regions.

The application of recombinant DNA techniques has been very successful in the pursuit of the cystic fibrosis (CF) gene, although there was no observed cytogenetic abnormality. Genetic linkage studies of CF and DNA markers localised the CF mutation to chromosome 7q22-31 (White et al, 1985). A large collaborative study using over 200 affected families was needed to establish the genetic order 917-PON-met-CF-D7S8. Thus genetic linkage allowed the CF to be localised to the region between met and D7S8, comprising 1-5 centimorgans corresponding to about 1-5 million base pairs. In order to proceed further, Williamson and co-workers made special use of the cell-transforming properties of the met oncogene to establish cell lines containing met and D7S8 but as little of other human genomic DNA as possible. Cosmid libraries were made from these cell lines and clones rich in G and C residues were selected (Estivill et al 1987). Such regions, termed CpG islands, are found near most (Bird 1986) and therefore by selecting clones in this way increased the chance of detecting a probe for a gene rather than a non-coding region of DNA. Initially one probe containing such a CpG island was thought to detect the gene for CF, after showing strong linkage disequilibrium with the disease (Estivill et al, 1987) but further analysis showed this not to be the case. Two new markers were found which were 0.5 centimorgans from met towards CF (Rommens et al, 1987). These probes were then used to approach closer to the CF gene by chromosome walking and jumping, with progress being monitored by measuring linkage disequilibrium between CF and polymorphisms in the cloned segments. Unfortunately these two probes used as starting points turned out to be only 10 kb away from each other so that progress towards CF was slower than had been hoped. Nevertheless, after cloning nearly 300 kb of DNA the CF gene was finally isolated (Rommens et al, 1989, Riordan et al, 1989) and the mutation identified (Riordan et al, 1989, Kerem et al, 1989).

With good fortune it may be possible to localise a schizophrenia susceptibility allele by genetic linkage in affected families to a map resolution of five to one million base pairs or less. This will entail the cloning of markers that "step over" recombinants by virtue of being genetically closer than previously used markers. This will need to be carried out with the use of samples biased heavily towards affected cases and not relying on unaffected cases for linkage information. This approach will only can provide robust estimates of the recombination fraction if

non-allelic heterogeneity is not too great. The inclusion of unaffected cases would bring about uncertainty because apparent recombinants may in fact be unaffected carrier individuals. Other factors which may affect the accurate mapping of a susceptibility locus are heterogeneity, phenocopies and uncertain modes of inheritance. Even in the absence of a cytogenetic abnormality to help clone susceptibility genes for schizophrenia, with the rapid development of a high density polymorphic map, and a physical map of the human genome it may be possible to establish linkage disequilibrium with a genetic subtype of schizophrenia. The power to establish linkage disequilibrium with genetically heterogeneous disorders has been discussed by Gershon et al (1989) but it will be improved by the use of highly polymorphic microsatellites. These are not only evolutionarily stable (see chapter 4) but using alleles with relatively low population frequencies may decrease the numbers of individuals needed to establish disequalibrium even in the presence of heterogeneity. However the problems of multiple testing when using multiallelic systems has to be accounted for. This will narrow the area to search within for candidate genes. The problems of heterogeneity for the establishment of linkage disequilibrium may be overcome in several other ways. For example, by collecting a large number of affected individuals each from a particular genetic subtype. Alternatively the search could be facilitated by the identification of a particular clinical marker associated with a genetic subtype. Other methods which could be beneficial in the identification of disease genes include Alu PCR differential hybridization (Bernard et al 1991) and laser microdissection and single unique primer PCR of a single human chromosome (Hadano et al 1991) which isolates many clones from an area of interest. These small clones can be used to identify larger clones and polymorphisms for genetic analysis. As already stated it may not be possible to genetically identify a schizophrenia susceptibility locus to less than 5-1 Mb of DNA. Yeast artificial chromosome vectors (YACs; Shero et al 1991) can produce large contigs of DNA rapidly, by use of the inverse PCR technique to rescue the genomic ends of YAC inserts (Silverman et al 1991). Candidate genes could be identified by using smaller subclones on zoo blots (Rommens et al 1989), or through HTF islands identified with rare restriction enzymes or oligonucleotides recognising these sites (Melmer et al 1991). Genes may also be identified by exon trapping (Buckler et al 1991) or using degenerate oligonucleotides for splice sites (Melmer et al 1991).

A number of techniques have recently been improved with the use of PCR which allow the rapid identification of mutations in a candidate genes. Single strand

conformation analysis (SSCP; Orita et al 1989; Poduslo et al 1991) detects changes in non-denaturing polyacryamide gels because the electrophoretic mobility of single stranded DNA depends not only on its size but also on its sequence. Denaturing gradient gel electrophoresis has been improved by the addition of GC clamps which alters the melting properties of the fragment. This greatly increases the fraction of possible mutations that are detected (Abrams et al 1990; Traystman et al 1990). Heteroduplex DNA with mismatches denatures at different levels of denaturant due to changes in the helix,and as a consequence are electrophoretically retarded. Another method uses RNase A to cleave at mutations between cRNA and genomic DNA (Kaufman et al 1990),alternatively chemical cleavage of mismatch heteroduplexes can also identify point mutations (Forrest et al 1991; Dianzani et al 1991). Identified mutations are then checked by direct genomic sequencing using PCR.

Unfortunately, positive linkage results do not lead inevitably to the rapid identification of the disease gene. For the autosomal dominant disorder myotonic dystrophy (DM), linkage was first demonstrated to the Lewis blood group and secretor loci in 1954, but despite chromosome 19 being small and containing many polymorphic markers the disease gene has only recently been identified (Brook et al 1992). Huntington's disease has long been known to be linked to G8 which is mapped to the short arm of chromosome 4 at 4p16.3, but because of its extremely telomeric location it has proved difficult to identify flanking markers and the gene involved has not been identified. Recently 2.5 Mb of DNA has been identified as containing the gene (Bates et al 1991).

The development of this range of efficient techniques at both the molecular and statistical levels will eventually help identify susceptibility mutations for schizophrenia. At this juncture there are two strong candidate regions for schizophrenia susceptibility loci. On chromosome 11q21 there is a report of an association of a balanced translocation with schizophrenia in a large pedigree with a maximum lod of 4.34 (St.Clair et al 1990). This area is further highlighted by the presence of the tyrosinase gene and the report of a large family with schizophreniform psychosis co-segregating with tyrosinase negative oculocutaneous albinism (Baron 1976; Barton et al 1988), and also by the mapping of the candidate human dopamine (D2) receptor gene to this area of chromosome 11 (Grandy et al 1989). A number of other families show co-segregation between Marfans syndrome and schizophrenia (Romano et al 1987; Sirota et al 1991).

Recently Marfans has been localized to chromosome 15q15-21.3 (Dietz et al 1991a) and mutations identified in the fibrillin gene (Dietz et al 1991b).

The last three years many genes for brain receptors have been cloned, some of these are strong candidate genes for schizophrenia. Neuroleptic drugs appear to act via an inhibition of dopamine receptors and the measurement of dopamine metabolites in post mortem brain tissue provide some evidence of a hyperactivity of dopaminergic neurotransmission in the disease (Reynolds 1989). Five dopamine receptors have been cloned by homology to each other or to other G-protein coupled receptors (Bunzow et al 1988; Sunahara et al 1990; Sokoloff et al 1990; Van Tol et al 1991; Sunahara et al 1991), some of these have produced negative results in preliminary linkage studies with schizophrenia (Kennedy et al 1991). In view of the finding of rare mutations in the amyloid protein for Alzheimers, a thorough search in these candidate genes for a role in schizophrenia may include association studies using microsatellites to improve their power, or searching for rare mutations in schizophrenic individuals using the methods mentioned previously. One such study for the D2 receptor proved negative for 14 schizophrenics (Sarkar et al 1991). Other candidate genes include the glutamate receptors, Freed (1988) has suggested that a down-regulation of glutaminergic activity is the means by which neuroletics exert their antipsychotic effect, and lowered concentration of mRNA that encode a non-NMDA glutamate receptor (KA/AMPA-R) has beeen reported in the post mortem brain hippocampal tissue of schizophrenics (Harrison et al 1991). Several receptors have been cloned (Hollmann et al 1989; Werner et al 1991; Egebjerg et al 1991) as has an NMDA recetor (Moriyoshi et al 1991; Kumar et al 1991). Phencyclidine which is an NMDA antagonist produces schizophreniform psychosis which can consist of both negative and positive symptoms. These as yet have not been tested for a genetic role in schizophrenia.

Recently an unstable DNA sequence (comprised of a trinucleotide repeat) shows expansion in successive generations (Harley et al 1992) providing a molecular basis for differences in severity and age of onset amoung successive generations with myotonic dystrophy and fragile X syndrome (Yu et al 1991) and is known as "anticipation". Enticingly, such unstable DNA sequences could be used to explain the complex familial segregation patterns for schizophrenia.

The successful localization and identification of schizophrenia susceptibility loci will improve the classification of the various subtypes of schizophrenia into distinct

groups with specific prognoses, drug treatment, and susceptibility for carriers. Genetic counselling may also eventually become possible if mutations can be detected directly, which has important ethical considerations. The early identification of incompletely penetrant carriers may give insights into simple and obvious cures, and a complete knowledge of the disease metabolic pathway will allow highly specific new drug treatments to be developed.

Schizophrenia does present problems for the molecular genetic research, but they do not really present "special" problems which are qualitatively different from those encountered in the genetic investigation of other diseases. Such research always demands a heavy investment of time and resources, and it is reasonable to suppose that if adequate effort is applied then at least some of the genes involved will eventually be identified. It could be argued that the potential benefits that would accrue are relatively greater than for non-psychiatric illness, because so little is now known about the aetiology of schizophrenia. Even a small increase in our understanding will represent a large step forward. The localisation, sequencing and cloning of a susceptibility locus with the subsequent identification of its product would mean that the whole disease processes culminating in a complex this syndrome could be fully understood.

8.1 SOURCES OF CHEMICALS

Ampicillin: Beecham Research labs (Brentford, UK)

Calf intestinal phosphatase: Boehringer Corporation Ltd

Chemicals: Sigma Chemicals Ltd

Cetyl trimethyl ammonium bromide (CTAB): Sigma Chemmicals Ltd

Ultra pure dATP, dCTP, dGTP, dTTP: Pharmacia Chemicals Ltd

DNA polymerase Klenow fragment: Boehringer Corporation Ltd

Fugi-RX film: Genetic Research Instruments

Hybond-N: Amersham International plc

Kanamycin: Sigma Chemicals Ltd

SeaKem GTG agarose: Flowgen Intruments Ltd

Ultra pure LMP agarose: Gibco BRL

pUC18/19 DNA: Boehringer Corporation Ltd

mp18/19 M13 RF DNA: Boehringer Corporation Ltd

Restrition enzymes: Boehringer Corporation Ltd

T4 DNA ligase: Boehringer Corporation Ltd

ATP (special grade): Boehringer Corporation Ltd

T4 DNA polynucleotide kinase: Boehringer Corporation Ltd

Ampli-Taq (Taq pol): Cetus Perkin Elmer

Sequenase kit: United States Biochemicals Corporation

Oligonucleotides synthesis: Oswell DNA services

Random hexanucleotides (pD6): Pharmacia Chemicals Ltd

Ribonuclease A: Sigma Chemicals Ltd

Sequagel kit: Flowgen Instruments Ltd

Agar: Difco laboratories

Yeast extract: Difco laboratories

Tryptone: Difco laboratories

Poly d(C-A).d(G-T): Pharmacia Chemicals Ltd

Poly d(C-T).d(G-A): Pharmacia Chemicals Ltd

DNA thermal cycler: Cetus Perkin Elmer

[alpha ^{32}P] dCTP 3000 ci/mMol: Amersham International plc

[gama ^{32}P] ATP 3000 ci/mMol: Amersham International plc

[gama ^{35}S] dATP NE0034S: Dupont NEN

8.2 SOURCES OF PROBES

L599Ha, p105-153Ra, p105-798Rb, Jo110Hc, Jo71Hc purchased from ATCC USA. M4 was a gift from Dr AE Retief, Stellenbosch University, Tygerberg, S Africa. DHFR was a gift from Dr S Detera-Wadleigh, NIMH Bethesda, Maryland USA. HEXB (pHEX-X) was a gift from BF O'Dowd, Hospital for sick children, University of Toronto, Ontario, Canada. G21 from Dr BK Kobilka, Howard Huughs Medical Institute, Durham, North Carolina USA. EF5.15, YN5.132, and MC5.60 were gifts from Dr Y Nakamura, Howard Hughes Medical Institute, Salt Lake City, USA. The human DNA minisatellite fingerprint clones were recieved from Professor A Jefferys of Liecester University..

8.3 BUFFERS AND SOLUTIONS

Ampicillin: prepared as a 100 mg/ml stock solution in water and stored at -20 °C. Used at a working concentration of 45ug/ml.

Blood lysis solution: 10 x, 100mM NaCl, 100mM EDTA pH approx 7.

Denaturization buffer: 1.5M NaCl, 0.5M NaOH

100x Denharts: 2% (w/v) bovine serum albumin, 2% (w/v) polyvinylpyrrolidone, 2% (w/v) Ficoll 400 made up in water.

DNA loading buffer: 10 x, 0.25% bromophenol blue, 0.25% xylene cyanol, 25% Ficoll (type 400) in water.

Chloroform: chloroform and isoamyl alcohol mixed at a ratio of 24:1.

Ethidium bromide: solution of 10mg/ml in water. Stored 4°C.

IPTG: 100mM solution of isopropyl-B-D-thiogalactopyranoside (23.8 mg/ml of water). Store -20 °C.

Ligation buffer:10 x, 660mM Tris.Cl (pH7.6), 50mM $MgCl_2$, 50mM DTT, 6mM ATP. Store -20 °C.

Lysozyme buffer: 50mM glucose, 10mM EDTA, 25mM Tris.Cl (pH8.0). Lysozyme added to 4mg/ml.

Neutralization buffer: 1M Tris.Cl (pH7.6), 1.5M NaCl.

Proteinase-K buffer: 100mM NaCl, 10mM Tris.Cl (pH8.5), 25mM EDTA.

Phage lysis buffer: 2.5% SDS, 0.5M Tris.Cl (pH9.0), 0.25 M EDTA.

Phenol: dissolved at 65°C, saturated with water and then equilibriated with TE pH7.6.

Spermdine: 100mM solution in water. Store -20 °C.

SM buffer: 100mM NaCl, 10mM $MgSO_4$, 50mM Tris.Cl (pH7.5), 0.01% gelatin.

SSC: 20 x, 3M NaCl, 0.3M tri-sodium citrate.

T4 polynucleotide kinase buffer: 800mM Tris.Cl (pH7.6), 120mM $MgCl_2$, 60mM dithiothreitol (DTT), made up in water and stored at -20 °C.

TAE: 40mM Tris.Cl, 1mM EDTA in water, pH adjusted to 8.0 with glacial acetic acid.

Taq polymerase buffer: 10 x, 100mM Tris.Cl (pH8.3), 500mM KCl, 20mM $MgCl_2$, 0.1% gelatin. $MgCl_2$ varied in concentration from 1.0mM-2.5mM for different amplifications. Stored -20 °C

TBE: 90mM Tris, 90mM Boric acid, 1.25mM EDTA, pH8.3.

TE: 10mM Tris.Cl, 1mM EDTA in water, pH 7.5-9.0 as stated.

X-Gal: 2% solution made up by dissolving 20mg of 5-Bromo-4-chloro-3-indoly-B-D-galactoside in 1ml of dimethyl formamide. Store -20 °C.

8.4 MEDIA

Glucose/minimal medium plates: 1.5% (w/v) minimal agar, 1 x M9 salts, 1mM $MgSO_4$, 0.1mM $CaCl_2$, 1mM thiamine HCl, 0.2% glucose.

Luria Bertani medium (LB): 1% (w/v) bacto tryptone, 0.5% (w/v) bacto yeast extract, 1% (w/v) sodium chloride made up in water and adjusted to pH7.5.

L-agar: 1.5% (w/v) Difco technical agar in LB.

H-agar: 1.2% (w/v) Difco technical agar in 1% bacto tryptone, 0.8% sodium chloride.

H-top agar: 0.5% agarose in 1% (w/v) bacto tryptone, 0.8% sodium chloride.

10 x M9 salts: per litre, 60g Na_2HPO_4, 30g KH_2PO_4, 10g NH_4Cl, 5g NaCl, stored at 4°C.

2 x TY: 1.6% (w/v) bacto tryptone, 1% (w/v) bacto yeast extract, 0.5% (w/v) sodium chloride made up in water.

All media were sterilized by autoclaving at 120°C for 20 minutes.

8.5 SOUTHERN BLOTTING AND HYBRIDIZATIONS ANALYSIS OF GENOMIC DNA

a) Isolation of genomic DNA from peripheral lymphocytes

The peripheral lymphocytes from 10mls of blood were lysed in 40mls of 1 x blood lysis solution. The nuclei were pelleted by centrifugation at 2000g for 10 minutes, resuspended in 40mls of 1 x blood lysis solution and pelleted. The cleaned nuclei were resuspended in 10mls of proteinase-K buffer and the nuclear membranes lysed with 1ml of 10% SDS (sodium dodecyl sulphate). Protein digestion was carried out with proteinase-K at 55°C for 16-20hrs. The DNA was extracted with an equal volume of of phenol and chloroform, and precipitated with one tenth the volume

of 3M sodium acetate and 2.5 x the volume of absolute ethanol. The DNA was resuspended in 0.5mls of TE and dissolved at 4°C. The DNA concentration was measured by its absorbance at 260nm (OD of 1 equals 50ug/ml),and the samples equilibriated to 5ug/20ul.

b) Digestion of genomic DNA with restriction enzymes

5ug of genomic DNA was digested with 20 units of a restriction enzyme, at the temperature and using the buffer recomended by the manufacturer with the addition of 1mM spermidine. The DNA was digested to completion overnight.

c) Southern blotting of genomic DNA and hybridization

Southern blotting (Southern 1975) was carried out using a modification of the protocol described by Amersham International plc for use with Hybond-N filter membranes. Digested DNA was separated by electrophoresis through 0.6-1% TAE agarose gels at 50V until the appropriate alleles had resolved (usually 18-48hrs). The gel was stained with ethidium bromide, photographed on a UV transilluminator, and denatured with denaturization buffer for 1hr. Transfer to Hybond-N membrane was carried out overnight in 20 x SSC after which the DNA was covalently linked to the membrane by UV irradiation for 2 minutes. Membranes were prehybridized at 65°C in 20mls of 0.9M NaCl, 1% SDS and 50ug/ml autoclaved denatured salmon testes DNA for a minimum of four hours. Hybridization was carried out overnight at 65°C in 20mls of fresh solution including 10% (w/v) dextran sulphate as well as 100ng denatured oligolabelled probe at 2 x 10^6 cpm/ml. The membranes were then washed to a stringency of 1 to 0.1 x SSC, 0.1% SDS at 65°C. The filters were exposed against autoradiographic film at -70°C using intensifying screens. In order to reprobe the filters were treated as recomended by the manufacturer.

8.6 DNA FINGERPRINTS USING THE LAMBDA 33.6 PROBE

a) Probe preparation

The M13 probe DNA was prepared as for single stranded sequencing. The probes were labelled according to Hu and Messing (1982), 100ng of M13 single strand DNA was hybridized with 2ng of M13 hybridization primer (Amersham International plc) at 65°C for 30 minutes in 10mM dithiothreitol,10mM Tris.HCl pH7.9, 60mM

NaCl, and 6.6mM $MgCl_2$. The primer hybridizes after the polylinker. This was then cooled to RT before dATP, dGTP, and dTTP were added to a final concentration of 50uM each along with 3ul of [alpha32P] dCTP (3000Ci/mM). This was incubated in the buffer above with 1U of DNA polymerase Klenow fragment for 90 minutes at 15°C. The reaction being inhibited by the addition EDTA pH8.3 to 25mM, and the probe stored on ice until used.

b) Hybridizations

Genomic DNA was digested with HinfI overnight to completion. 5ug was size separated by electrophoresis in 1x TBE 24cm agarose gels run at 1 volt/cm with several changes of the buffer. The gels were run until the 4kb marker was within 2cm of the end of the gel. The gels were denatured for 1hr then neutralized for 1hr before the DNA was southern transfered in 20 x SSC overnight onto Hybond-N. The DNA was UV fixed for 2 minutes. Hybridizations were carried out according to Jeffreys et al (1985). The Hybond-N filters were prehybridized at 65°C overnight in 1 x SSC, 5 x Denharts, 0.5% SDS and 1mg denatured herring sperm DNA (Sigma Chemical company). Hybridizations were in the same conditions but with 5% PEG(6000) overnight. Post hybridization washes were to a stringency of 1 x SSC, and autoradiography overnight at -70°C. Paternity of the families was checked by eye.

8.7 PLASMID AND COSMID VECTOR PREPARATIONS

solution 1: 50mM glucose, 25mM Tris-Cl and 10mM EDTA.
solution 2: 0.2M NaOH, 1% SDS.
solution 3: 147.2g Potassium acetate, 57.5 mls Glacial acetic acid. Resulting solution 3M Potassium ions and 5M acetate ions.

5ml overnight cultures prepared from single bacterial colonies were used to inoculate 500mls of sterile LB containing the appropriate antibiotic and grown overnight at 37°C with vigorous shaking. The bacteria were pelleted at 4200 rpm for 30 minutes at 4°C in a Beckman J6-B centrifuge and then resuspended in 50mls of solution 1. The cells were lysed by adding 100mls of solution 2 and the bacterial debri and chromosomal DNA was precipitated with the addition of 50mls of solution 3. After mixing the precipitate was pelleted by centrifugation at 4200 rpm for 15 minutes at 4°C. The supernatant was filtered through nylon gauze and the plasmid DNA and bacterial RNA precipitated by adding 120mls of propan-2-ol. After

centrifugation at 6000 rpm in a sorval GS3 rotor, the pellet was washed with 70% ethanol and resuspended in 5mls of TE pH8.0 before being transferred to a pre-weighed universal and made to 9g with TE pH8.0. To this was added 10g caesium chloride and 1ml of 5mg/ml ethidium bromide. The plasmid DNA was separated from the RNA according to density by centrifugation for 24-48hrs in a Beckman 70Ti rotor at 55000rpm at 20°C. The lower plasmid band was extracted from the gradient and made to 10mls with TE pH8.0. Plasmid DNA was precipitated with 20mls of absolute ethanol at room temperature and pelleted by centrifugation at 3000rpm for 20 minutes in a Beckman J6-B centrifuge. After washing with 70% ethanol and drying the DNA was resuspended in 500ul TE and treated with 40U/ml RNase A for 15 minutes at 37°C. The protein was then extracted with phenol/chloroform and the cleaned purified plasmid DNA precipitated with ethanol.

8.8 LARGE SCALE PHAGE DNA PREPARATION

Solutions:10% (w/v) maltose filter sterilized, 1M $MgSO_4$ (autoclaved),3M Sodium Acetate pH5.2, PEG 6000 50% (w/v) mix overnight to dissolve fully.

The host bacteria (LE392) was grown overnight in 10mls of LB media with 10mM $MgSO_4$ and 0.2% maltose. The cells were pelleted by centrifugation and resuspennded in 10mls of 10mM $MgSO_4$ and stored at 4°C for a maximum of three weeks. The optical density at 600nm was measured (OD 1 = 8 x 10^8 bacteria/ml). The titre of the phage lysate was also determined. In a total volume of 200ul containing 10mM $MgSO_4$ 10^{10} bacteria were incubated with 10^7 phage at 37°C for 20 minutes. After the absorbtion of the phage, the mixture was added to 500mls of LB, 10mM $MgSO_4$ and 0.2% maltose (in a smooth sided flask) and shaken vigorously overnight at 37°C. The flask was shaken for a further 30 minutes with 1ml of chloroform and the cell debris settled for 30 minutes. The lysate was removed leaving the chloroform behind and centrifuged at 8K and 4°C for 10 minutes (sorval GS3 rota). The cleared lysate was decanted into a clean flask and bacterial DNA and RNA were digested with 20ul of pancreatic DNase (20mg/ml) and 10ul pancreatic RNase (20mg/ml) by incubation at 37°C for 30 minutes. Sodium chloride was added to a final concentration of 1M dissolved and left on ice for 30 minutes after which 100mls of 50% PEG 6000 solution was added, mixed thoroughly and left on ice overnight to precipitate the phage. The phage and other debris were then pelleted at 8K and 4°C for 10 minutes (sorval GS3 rota). The

phage were then extracted with 3mls of SM buffer and spun at 8K for 2 minutes at RT (sorval HB4). The supernatant was collected and the remaining pellet was extracted once more with 1.5mls of SM buffer. To 0.5mls of the combined supernatants 5ul of 10% SDS and 10ul 0.25 M EDTA were added and the phage coats ruptured by incubation at 68°C for 15 minutes. This was extrated with an equal volume of phenol, then phenol chloroform and finally chloroform. The phage DNA was ethanol precipitated in the usual way.

8.9 PLASMID AND COSMID MINI-PREP WITH CTAB

The method described is a modification of Del Sal et al 1988. The bacterial cells from 1.5mls of an overnight culture were pelleted in an Eppendorf microfuge, and resuspended in 200ul of STET buffer (8% sucrose w/v, 0.1% TritonX-100 v/v, 50mM EDTA, 50mM Tris-HCl pH8.0). After the addition of lysozyme to break down the bacterial cell wall (4ul of 50mg/ml RT for five minutes) the samples were boiled for 45 seconds to precipitate proteins and cell debris. These were pelleted by centrifugation for 10 minutes and the supernatant removed to a fresh microfuge tube. The plasmid DNA was precipitated with 8ul CTAB (5% w/v) and centrifuged for 5 minutes at RT. The pellet was then resuspended in 300ul of 1.2M NaCl and the DNA reprecipitated with the addition of 750ul of absolute ethanol. The purified plasmid DNA was collected by centrifugation for 10 minutes, washed in 70% ethanol air dried and dissolved in 50ul of TE. For double stranded plasmid sequencing this DNA was further purified by several rounds of phenol/chloroform extractions.

8.10 PLASMID LIGATIONS

After restriction endonuclease digestion of vectors, an aliquot was run on a 0.8% TBE agarose gel to confirm that restriction of the DNA had gone to completion. The restriction endonuclease was inactivated at 65°C for 10 minutes. If the restricted ends were incompatible the plasmid vector was precipitated with propanol and ammonium acetate (Maniatis et al 1982), otherwise the vector was treated with calf intestinal phosphatase (15-20U) for 60 minutes at 37°C. Phenol and chloroform extracted and precipitated with absolute ethanol. Inserts were separated and purified in low melting point agarose in TAE, the correct band being visualised on a UV transluminator and excised. 100ng of insert was ligated to 20ng of vector as a general rule in 1 x ligation buffer with 1U of T4 DNA ligase and the agarose diluted to 0.1%. Ligations were carried out overnight at RT.

8.11 PREPARATION OF COMPETENT E.COLI CELLS

Bacteria were grown in 50 mls LB, 10mM $MgSO_4$.7H2O and 0.2% glucose to mid logarithmic phase (OD 550nm of 0.4 for JM83 Rec-, and 0.3 for TG2). The cells were then left on ice for 10 minutes and then gently centrifuged at 1500g for 10 minutes at 4°C. The pelleted cells were resuspended in ice cold 0.5mls of the above solution, to this 2.5mls of 36% glycerin, 12% PEG (MW7500), 12 mM $MgSO_4$.7H2O added to LB and sterilized by filtration, was added and mixed well without vortexing. The cells were aliquoted and stored at -80°C for up to 3 months.

8.12 TRANSFORMATION OF DNA INTO COMPETENT E.COLI

5ng of ligated vector was added to 100ul of competent cells and left on ice for at least 1hr. The bacteria were then heat shocked at 42°C for 3 minutes and left on ice. For plasmids an equal volume of 2 x LB was added and the cells incubated for 1hr to allow expression of the antibiotic resistance. The bacteria were briefly pelleted (15 seconds) and 1/10 and 1/100 dilutions were prepared in LB and spread onto an LB agar plate with the appropriate antibiotic and incubated overnight at 37°C. For PUC plasmid vectors X-Gal and IPTG were added to the agar to allow blue (non recombinants) white (recombinants) selection.

8.13 RANDOM PRIMER LABELLING OF PROBES

The method of Feinberg et al 1983 was used to label purified fragment DNA in low melting point agarose. 100ng was denatured by boiling for 7 minutes and quenched on ice before 10ul of labelling buffer, 30uCi of [alpha^{32}P] dCTP and 2 units of Klenow fragment were added to a final volume of 50ul. The reaction was incubated at 37°C for 1-2hrs. 50ul of TE buffer was added and the unicorporated radionucleotide removed by cetrifugation (Beckman GRI) 2500rpm for 5minutes through a Sephadex G50 medium column. The incorporation was measured by Cherenkoff counting.

8.14 COMPETITION OF HUMAN REPEAT CONTAINING PROBES

Probes that contain human repeat elements need competion with total human DNA before hybridization to genomic southerns. After the spun column probes were made up to a total volume 200ul with TE pH7.6 and to this 80ul of 1.2M sodium phosphate and 500ul of 2.5mg/ml sonicated human placental DNA were added. This mixture was boiled for 10 minutes and incubated for 5-8hrs at 65°C before being added directly to the hybridization solutions.

8.15 5' END LABELLING OLIGONUCLEOTIDES

Oligonucleotides (25ng) were end labelled by T4 polynucleotide kinase in the presence of 120uCi of [gama^{32}P] ATP and 1 x kinase buffer in a total volume of 30ul. After 1hr at 37°C, the labelled probe was separated from the unicorporated isotope by Sephadex G50 (medium) spun column.

8.16 HYBRIDIZATION OF GENOMIC LAMBDA LIBRARIES

The library was titred in a suitable host (LE392 for Clontech EMBL3 Sau3A1 partial HL1067j genomic library). 30,000 phage were plated out per 15cm petri dish, and a total of 20 plates were prepared such that 4-5 genomes were screened (Clarke and Carbon 1976). The phage were made upto a total volume of 300ul with SM buffer and to this 300ul of $MgSO_4$ stock host cells were added and the mixture incubated at 37°C for 15-20 minutes for absorbtion to occur. This was mixed with 7mls of moulton (45°C) top agarose 10mM $MgSO_4$ and poured directly onto LB agarose 15cm plate. The phage plaques were incubated overnight at 37°C (or until a near confluent phage plate had grown). Plaque lifts were taken directly onto Hybond-N, and the plates stored at 4°C. Hybond-N nylon filters were laid directly onto the plates and left for 20-30 seconds, and then carefully removed and placed onto filter paper soaked in denaturization solution for 5 minutes to denature the DNA. After neutralizing for a further five minutes the filters were washed gently in 2 x SSC and air dried. The DNA was fixed by baking at 80°C for 2hrs. Hybridizations were carried out in the same way as used for southern blots, with roughly seven filters per hybridization bag containing 100ng of labelled probe. After autoradiography (-70°C) usually overnight the positive clones were picked with the end of a sterile Pasteur pipette and stored in 500ul of SM buffer at 4°C. The filters were incubated at 75°C for 2hrs to remove the probe and were reused upto six times. Secondary screens were carried out by plating 10^{-5}, 10^{-6}, and 10^{-7} dilutions of the SM stock using 200ul of plating cells per 9cm petri dish. Plaque lifts and hybridizations were as above. Further screens were carried out at 10^{-3}, 10^{-4} and 10^{-5} dilutions until a plate contained none but the purified clone. Plaques from these plates were stored in SM buffer at 4°C. Large scale phage DNA preparations of these clones were carried out.

8.17 HYBRIDIZATIONS OF COSMID GENOMIC LIBRARIES

The cosmid glycerol stock was titred. The volume equalling 30,000 colonies was added to 500ul of LB and spread onto a dried LB agarose 15cm plate containing the appropriate antibiotic (kanamycin for pcos2EMBL). The plates were incubated

overnight at 37°C, and 15 plates were prepared representing approximately six genomes. Hybond-N filters were placed onto the colonies and carefully lifted so as not to spread the bacterial colony. The plates were stored at 4°C. The filters were denatured for 5 minutes and then neutralized for a further 5 minutes before being gently washed in 2 x SSC. The filters were air dried and baked at 80°C for 2hrs. The filters were prehybridized at 65°C with 1mg of proteinase K overnight to remove the cellular debris (five filters per bag), rinsed and prehybridized again before hybridization as described above. Positive clones were picked with a sterile Pasteur pipette and stored in 500ul of LB at 4°C after a brief incubation at 37°C (5-10 minutes). Secondary screens were carried out on serial dilutions of the stock by plating 10^{-2} to 10^{-5} dilutions out on 9cm plates. Further screens were performed until a positive clone well isolated from other colonies could be picked. Mini and large scale preparations of cosmid DNA were carried out.

8.18 SHOTGUN CLONING AND SEQUENCING OF MICROSATELLITES

a) Shotgun ligations

200ng of phage or cosmid DNA were restriction digested with Sau3A, AluI and HaeIII and combinations of these three. The digested clones were size separated in 1% agarose gels and southern blotted onto Hybond-N. The clones were screened for the presence of dinucleotide repeats using oligolabelled poly (dC-dT).(dG-dA) and poly (dC-dA).(dG-dT). Prehybridization of the filters was in 0.9M NaCl, 1% SDS and 40ug/ml tRNA at 65°C for at least 1hr. Hybridizations were carried out in the same buffer with the labelled probe overnight at 65°C. The filters were washed to a stringency of 1 x SSC at 65°C. Positive clones gave an autoradiographic signal after 1hr RT exposure to X-ray film. The genomic clones were digested such that the microsatellite containing fragment was approximately 200-300bp. 100ng of this DNA was ligated to 100ng of prepared M13 mp18 vector DNA (prepared as plasmid vectors), BamHI cut mp18 for Sau3A ligations and HincII cut vector for AluI or HaeIII ligations. Ligations were carried out overnight at RT, in a total volume of 30ul.

b) Identification of microsatellite containing clones

The ligations were diluted to a volume of 100ul with TE pH7.6, and 10ul of this was added to 100ul of competent TG2 cells (grown originally on glucose minimal medium plates to express the F-pilus) and left on ice for a minimum of 1hr. After heat shocking for 3 minutes at 42°C, this was added to 100ul of logarithmic phase

TG2 cells, 40ul X-gal, and 40ul IPTG mixed with 3mls of H-top agar poured onto warmed LB agar 9cm plates. The plates were incubated at 37°C overnight, clear plaques indicating recombinants (normally 100-200 recombinants/plate). Five plates for each genomic clone were routinely plated, and plaque lifts taken onto Hybond-N. The filters were denatured, neutralized and washed in 2 x SSC, air dried and finally UV fixed for two minutes. Recombinants containing microsatellites were identified by hybridization as described above. After 2hrs exposure to X-ray film positive M13 plaques could be identified. Alu elements with potential 3'poly deoxyadenylate tracts were identified by hybridizing Sau3A ligations with end labelled oligonucleotides representing the conserved regions of these elements. Hybridizations were carried out at 55°C in the same buffer as for dinucleotide repeats, with the filters washed to a stringency of 2 x SSC at RT. Positive clones were identified on X-ray film after overnight exposure.

c) Sequencing the microsatellites

A modification of the protocol described by Amersham in their booklet "M13 cloning and sequencing" was used to prepare single stranded DNA. An overnight culture of E.Coli TG2 was diluted 1:100 with 2 x TY medium and 1.5ml aliquots inoculated with a positive M13 plaque. After incubating with shaking for 6hrs at 37°C, the phage released into the medium was separated from the bacteria by centrifugation for 5 minutes in a microfuge. 1ml of the supernatant was removed and transferred to a new eppendorf tube. Phage particles were precipitated by addition of 200ul of ice cold 2.5M NaCl, 20% (w/v) polyethylene glycol 6000 (PEG) on ice for 15 minutes and pelleted by spinning in a microfuge for 5 minutes. All traces of PEG were removed from the phage pellet before resuspending in 100ul of TE pH7.6. The viral particles were lysed by extracting with phenol, then phenol/chloroform and finally chloroform, and the DNA template precipitated with ethanol overnight at -20°C before being finally washed, air dried and resuspended in 30ul of TE pH7.6. A check gel was run with about 1-2ul of phage DNA to ensure the presence of DNA for sequencing. DNA sequencing was performed by the chain termination method originally devised by Sanger et al (1977) and detailed by United States Biochemical Corporation for use with their Sequenase kit. Sequencing reactions were labelled with [^{35}S] dATP and run on 6% polyacryamide gels in 1 x TBE for 2-18hrs. Gels were then fixed in 10% acetic acid and 10% methanol before drying at 80°C for 30 minutes and exposing to X-ray film at RT without intensifying screens. 3'Alu element ends were sequenced using the same protocol but replacing the primer with oligonucleotides representing the conserved regions

of the elements.

8.19 DOUBLE STRANDED PLASMID SEQUENCING

Plasmid DNA to be sequenced directly was prepared by the large scale preparation or more convienient CTAB mini preparation. 5ug of plasmid DNA was added to an equal volume of freshly prepared 0.4M NaOH and left at room temperature for 10 minutes. Denatured DNA was precipitated by addition of 0.1 volume 3M sodium acetate and 4 volumes ice-cold absolute ethanol and left at -70°C for 20 minutes. The DNA was recovered by centrifugation in a microfuge for 10 minutes, washed in 70% ethanol and air dried. The pellet was resuspended in 35ul of water and 7ul used in the sequencing reaction.

REFERENCES

Abrams ES et al (1990) Comprehensive detection of single base changes in human genomic DNA using denaturing gradient gel electrophoresis and a GC clamp. Genomics 7:463-475.

Abrams R et al (1983). The genetics of schizophrenia: A reassessment using modern criteria. Am J Psychiatry 140:171-175.

Allen MG et al (1972). Schizophrenia in veteran twins: a diagnostic review. Am J Psychiatry 128:939-945.

American Psychiatric Association (1987). DSM-IIIR Diagnostic and Statistical Manual of Mental Disorders. Washington DC:APA.

Amos CI et al (1990) The probabilistic determination of identity by descent sharing for pairs of relatives from pedigrees. 47:842-853.

Andrew B et al (1987). A study of genetic linkage in schizophrenia. Psychological Medicine 17:363-370.

Aschauer HN et al (1990) No evidence for linkage between chromosome 5 markers and schizophrenia. Hum Hered 40:109-115.

Barker D et al (1987) Gene for von Recklinghausen neurofibromatosis is in the pericentromic region of chromosome 17. Science 236:1100-1102.

Baron M (1976) Albinism and schizophreniform psychosis: A pedigree study. Am J Psychiatry 133:1070-1073.

Baron M et al (1977). Tissue binding factor in schizophrenic sera: A clinical and genetic study. Biol Psychiatry 12:199-219.

Baron M et al (1980). Plasma DBH activity: Relation to genetic factors in schizophrenia. Commun Psychopharmacol 4:197-202.

Baron M et al (1982a). Schizaffective illness, schizophrenia and affective disorders: morbidity risk and genetic transmission. Acta Psychiat Scand 65:253-262.

Baron M (1982b). Genetic models of schizophrenia. Acta Psychiatr Scand 65:263-275.

Baron M et al (1983). Familial relatedness of schizophrenia and schizotypal states. Am J Psychiatry 140:1437-1442.

Baron M et al (1984a). Erythrocyte catechol-o-methyltransferase activity in schizophrenia: Analysis of family data. Am J Psychiatry 141:29-32.

Baron M et al (1984b). Platelet monoamine oxidase activity and genetic vulnerability to schizophrenia. Am

J Psychiatry 141:836-842.

Baron M et al (1985a). A family study of schizophrenic and normal control probands: Implications for the spectrum concept of schizophrenia. Am J Psychiatry 142:447-455.

Baron M et al (1985b). Genetic analysis of platelet monoamine oxidase activity in families of schizophrenic patients. J Psychiatr Res 19:9-21.

Baron M et al (1985c). Genetic analysis of plasma amine oxidase activity in schizophrenia. Psychiatry Res 15:121-132.

Barton DE et al (1988) Human tyrosinase gene, mapped to chromosome 11 (q14-q21), defines second region of homology with mouse chromosome 7. Genomics 3:17-24.

Bassett AS et al (1988) Partial trisomy chromosome 5 cosegregating with schizophrenia. Lancet 8589:799-801.

Bassett AS (1992) Chromosomal aberrations and schizophrenia: Autosomes. In Press Br J Psychiatry.

Bates GP et al (1991) Defined physical limits of the Huntington disease gene candidate region. Am J Hum Genet 49:7-16.

Berg K (1981). Twin research in coronary heart disease, in twin research 3 : epidemiological and clinical studies. New York, Alan R Liss.

Bernard LE et al (1991) Isolation of DNA fragments from a human chromosomal subregion by Alu PCR differential hybridization. Genomics 9:241-246.

Bikker H et al (1988) XmnI RFLP at 5q13 detedted by a 0.49kb XmnI fragment of human hexosaminidase (HEX B) cDNA. Nucleic Acids Res 16:8198.

Bishop DT et al (1990) The power of identity by state methods for linkage analysis. Am J HumGenet 46:254-265.

Bird AP (1986) CpG-rich islands and the function of DNA methylation. Nature 321:209-213.

Blackwood DHR et al (1986). P300 event-related potential abnormalities in borderline personality disorder. Biol Psychiatr 21:577-560.

Blackwood DHR et al (1987). Changes in auditory P3 event-related potential in schizophrenia and depression. Br J Psychiatry 150:154-160.

Blackwood DHR et al (1990). Cognitive brain potentials and their applications.Br J Psychiatry 157(suppl):96-99.

Blackwood DHR (1992). Trait markers are they worth effort? ESF MNMI 1992 workshop. Aussois France.

Bodmer WF (1984) DNA polymorphisms and genetic markers in population and family studies of genetic predisposition. Banbury report 16:287-296.

Boehnke M (1990) Sample size guidelines for linkage analysis of a dominant locus for a quantitative trait by the method of lod scores. Am J Hum Genet 47:218-227.

Book JA (1953). A genetic and neuropsychiatric investigation of a north Swedish population. Acta Genet Stat Med 4:1-100.

Book JA et al (1978). Schizophrenia in a north Swedish geographical isolate, 1900-1977. Epidemiology, genetics and biochemistry. Clin Genet 14:373-394.

Botstein D et al (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Genet 32:314-331.

Brook JD et al (1992) Molecular basis of myotonic dystrophy: expansion of a trinucleotide (CTG) repeat at the 3' end of a transcript encoding a protein kinase family member. Cell 68:799-808.

Brynjolfsson J et al (1990) Diagnostic issues in the study of schizophrenia genetics. Nord Psykiatr Tidsskr 44:499-505.

Brzustowicz LM et al (1990) Genetic mapping of chronic childhood-onset spinal muscular atrophy to chromosome 5q11.2-13.3. Nature 344:540-541.

Buckler AJ et al (1991) Exon amplification: A strategy to isolate mammalian genes based on RNA splicing. PNAS USA 88:4005-4009.

Bunzow JR et al (1988) Cloning and expressioon of a rat D2 dopamine receptor cDNA. Nature 336:783-785.

Burke DT et al (1987). Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. Science 236:806-812.

Carter CH et al (1980). Segregation analysis of schizophrenia under a mixed genetic model. Hum Hered 30:250-256.

Cavalli-Sforza et al (1986) Detecting linkage for genetically heterogeneous diseases and detecting heterogeneity with linkage data. Am J Hum Genet 38:599-616.

Chakravarti A et al (1984) Non uniform recombination within the human beta-globin gene cluster. Am J Hum Genet 36:1239-1258.

Chartier-Harlin MC et al (1991) Early onset Alzheimer's disease caused by mutations at codon 717 of the beta

amyloid precursor protein gene. Nature 353:844-846.

Childs B (1970). Sir Archibald Garrod's conception of chemical individuality: a modern appreciation. N Engl J Med 282:71-77.

Clerget-Darpoux F (1991). The uses and abuses of linkage analysis in neuropsychiatric disorder. In The new genetics of mental illness. Butterworth-Heinemann Ltd, Oxford.

Cornblatt BA et al (1985). Positive and negative schizophrenic symptoms, attention and information processing. Schizophren Bull 11:397-408.

Coryell W et al (1988). The heritability of schizophrenia and schizoaffective disorder. Arch Gen Psychiatry 45:323-327.

Cotecchia S et al (1988) Molecular cloning and expression oof the cDNA for the hamster alpha 1 adrenergic receptor. Proc Natl Acad Sci USA 85:7159-7163.

Cox NJ et al (1988). Some effects of selection strategies on linkage analysis. Genetic Epidemiology 5:289-297.

Craig SP et al (1985) Localization of the human tyrosine hydroxylase gene to chromosome 11p15. Cytogenet Cell Genet 40:610.

Crow TJ et al (1986). The genetics of the unitary psychosis. Proceedings of the 4th world congress of biological psychiatry. Amsterdam:Elsevier.

Crow TJ et al (1988) Sex chromosomes and psychosis. The case for a pseudoautosomal locus. Br J Psychiatry 153:675-683.

Crowe RR et al (1990) The Iowa multiplex family study of schizophrenia: Linkage analyses on chromosome 5. Europ Arch Psychoneuro Sci 239:290-292.

Curtis D et al (1990) Unsound methodology in investigating a pseudoautosomal locus in schizophrenia. Br J Psychiatry 156:415-416.

Curtis D et al (1991) Using a dummy quantitative variable to deal with multiple affection categories in genetic linkage analysis. Ann Hum Genet 55:321-327.

Dalen P et al (1990) Aetiological heterogeneity of schizophrenia: The problem and the evidence. Br J Psychiatry 157:119-122.

Daniels RJ et al (1992) Linkage analysis of Spinal Muscular Atrophy. Genomics 12:335-339.

Daniels RJ et al (1992) J Med Genet

Darpoux C et al (1986). Effects of mispecifying genetic parameters in Lod score analysis. Biometrics 42:393-399.

Davies KE et al (1985). Human X chromosome markers and Duchenne muscular dystophy. Nuc Acids Res 13:3419-3426.

Dean M et al (1990) Approches to localizing disease genes as applied to cystic fibrosis. Nucleic Acids Res 18:345-350.

Dearry A et al (1990) Molecular cloning and expression of the gene for a human D1 dopamine receptor. Nature 347:72-76.

Debray Q et al (1978). Schizophrenia: A study of genetic models and some of their implications. Neuropsychobiology 4:257-269.

Debray Q et al (1979). Schizophrenia: A study of genetic models. Hum Hered 29:27-36.

DeLisi LE (1984). Is immune dysfunction associated with schizophrenia? A review of the data. Psychopharmacol Bull 20:509-513.

DeLisi LE et al (1985) Neurochemical aspects of schizophrenia. In Handbook of Neurochemistry. Plenum New York.

DeLisi LE et al (1986). A family study of the association of increased ventricular size with schizophrenia. Arch Gen Psychiatry 43:148-153.

Delisi LE et al (1990) The role of molecular genetics in psychiatry: Unraveling the aetilogy for schizoophrenia. In recent advaces in schizophrenia New York:Springer-Verlag.

Del Sal G et al (1988) A one tube plasmid DNA mini preparation suitable for sequencing.

den Dunnen JT et al (1987) Direct Detection of more than 50% of Duchenne Muscular Dystrophy Mutations by field inversion gels. Nature 329:640-642.

Detera-Wadleigh SD et al (1989). Exclusion of linkage to 5q11-13 in families with schizophrenia and other psychiatric disorders. Nature 339:391-393.

Dianzani I et al (1991) Simultaneous screening for beta thalassemia mutations by chemical cleavage of mismatch. Genomics 11:48-53.

Dietz HC et al (1991a) The Marfan syndrome locus: confirmation of assignment to chromosome 15 and identification of tightly linked markers at 15q15-q21.3. Genomics 9:355-361.

Dietz HC et al (1991b) Marfan syndrome caused by a recurrent de novo missense mutation in the fibrillin gene. Nature 352:337-339.

Dietzsch E et al (1986) An anonymous human single copy genomic clone, D5S6 (M4) on chromosome 5 identifies a three allele RFLP. Nucleic Acids Res 14:1923.

Donis-Keller H et al (1987). A genetic map of the human genome. Cell 51:319-37.

Douglas WW (1980) Histamine and 5-Hydroxytryptamine (serotonin) and their antagonists. *In* "Pharmacological bases of therapeutics" (A.G.Gilman, L.S.Goodman, A.Gilman, Eds.),pp 609-646. MacMillan, New York.

Dracopoli NC et al (1988). A genetic linkage map of 27 loci from PDN to FY on the short arm of chromosome 1. Am J Hum Genet 43:462-70.

Drayna D et al (1984) Genetic mapping of the human X chromosome by using restriction fragment length polymorphisms. PNAS 81:2836-2839.

Dryja TP et al (1986) Molecular detection of deletion involving band q14 of chromosome 13 in retinoblastomas. PNAS USA 83:7391-7394.

Eaton WW (1985). Epidemeology of schizophrenia. Epidemeological Reviews 7:105-126.

Edwards JH (1987) Exclusion mapping. J Med Genet 24:539-543.

Economou EP et al (1990) The polydeoxyadenylate tract of Alu repetitive elements is polymorphic in the human genome. Proc Natl Acad Sci USA 87:2951-2954.

Egebjerg J et al (1991) Cloning of a cDNA for a glutamate receptor subunit activated by kainate but not AMPA. Nature 351:745-748.

Ehrich E et al (1987) A family of cosmid vectors with multi-copy R6K replication origin. Gene 57:229-237.

Eisenstein BI (1990) The polymerase chain reaction. New England J Med 322:178-183.

Elston RC et al (1971). Schizophrenia: Evidence for the major gene hypothesis. Behav Genet 1:3-10.

Elston RC et al (1973). Possible linkage relationships between certain blood groups and schizophrenia. Behavioral Genetics 3:101-106.

Elston RC et al (1977). Family studies of schizophrenia. Bull Int Stat Inst 47:683-697.

Elston RC et al (1978). A genetic study of schizophrenic pedigrees. II. One locus hypothesis. Neuropsychobiology 4:193-206.

Elston RC (1980). Segregation analysis. In Current developments in anthropological genetics. Eds Mielke JH, Crawford MH. New York:Plenum

Essen-Moller E (1941). Psychiatrische untersuchungen an einer serie von zwillingen. Acta Psychiatr Scand 23:.

Estivill X et al (1987) A candidate for the cystic fibrosis locus isolated by selection for methylation-free islands. Nature 326:840-845.

Eubanks JH et al (1992). Localization of D5 dopamine receptor gene to human chromosome 4p15.1-p15.3, centromeric to the Huntington's disease locus. Genomics 12:510-516.

Falconer DS (1965). The inheritance of liability to certain diseases estimated from incidence among relatives. Ann Hum Genet 29:51-76.

Fargin A et al (1988) The genomic clone G-21 which resembles a Beta adrenergic receptor sequence encodes the 5-HT1A receptor. Nature 355:358-360.

Farmer AE et al (1987). Twin concordance for DSM III schizophrenia. Arch Gen Psychiatry 44:634-641.

Farmer AE et al (1991) Schizophrenia: How far can we go in defining the phenotype? In The new genetics of mental illness. Butterworth Heinemann, Oxford.

Farrar JG et al (1990) Autosomal dominant retinitis pigmentosa: linkage to Rhodopsin and evidence for genetic heterogeneity. Genomics 8:35-40.

Feighner JP et al (1972). Diagnostic criteria for use in psychiatric research. Archives of General Psychiatry 26:57-62.

Feinberg AP et al (1984) A technique for radiolabelling DNA restriction endonuclase fragments to high specific

activity. Addendum Anal Biochem 137:266-267.

Fischer M et al (1969). A Danish twin study of schizophrenia. Br J Psychiatry 115:981-990.

Fischer M (1971). Psychoses in the offspring of schizophrenic monozygotic twins and their noormal co-twins. Br J Psychiatry 118:43-52.

Forrest SM et al (1991) Mutation detection in phenylketonuria by using chemical cleavage of mismatch: importance of using probes from both normal and patient samples. Am J Hum Genet 49:175-183.

Francke U et al (1989) The glococorticoid receptor gene is in 5q-q32. Genoomics 4:610-612.

Freed WJ (1988) The therapeutic latency of neuroleptic drugs and non specific post junctional supersensitivity. Schizphrenia Bull 14:269-277.

Frielle T et al (1987) Cloning of the cDNA for the human beta 1 adrenergic receptor. Proc Natl Acad Sci USA 84:7920-7924.

Friend SH et al (1986). A human DNA segment with properties of the gene that predisposes to retinoblastoma and osteosarcoma. Nature 323:643-646.

Froguel Ph et al (1992) Close linkage of glucokinase locus on chromosome 7p to early onset non insulin dependent diabetes mellitus. Nature 356:162-164.

Fryer AE et al (1987) Evidence that the gene for Tuberous Sclerosis is on chromosome 9. Lancet 1:659-661.

Fujita R et al (1990) Additional polymorphisms at marker loci D9S5 and D9S15 generated extended haplotypes in linkage disequilibrium with Friedreich ataxia. Proc Natl Acad Sci USA 87:1796-1800.

Fulker DW (1973). A biometrical genetic approach to intelligence and schizophrenia. Social Biology 20, 266-275.

Fuller RW (1991) Serotonergic receptor subsystems and their functional significance. Biological Psychiatry 1:529-531.

Fung Y-H et al (1987) Structural evidence for the authenticity of the human retinoblastoma gene. Science 236:1657-1661.

Garrone G (1962). Etude statistique et genetique de la schizophrenia a Geneve de 1901 a 1950. J Genet Hum 11:89.

Gershon ES et al (1987a) The outlook foor linkage research in psychiatric disorders. J Psychiat Res 21:541-550.

Gershon ES et al (1987b) The role of molecular genetics in psychiatry. Biol Psychiatry 22:1388-1405.

Gershon ES et al (1988). A controlled family study of chronic psychoses. Arch Gen Psychiatry 45:328-336.

Gershon ES et al (1989) Detection of marker associations with a dominant disease gene in genetically complex and heterogeneous diseases. Am J Hum Genet 45:578-585.

Gessler M et al (1990) Homozygous deletion in Wilm's tumours of a zinc finger gene identified by chromosome jumping. Nature 343:774-778.

Gilliam CT et al (1989) Deletion mapping of DNA markers to a region of chromosome 5 that cosegregates with schizophrenia. Genomics 5:940-944.

Giuffra LA et al (1988) Glucocorticoid receptor maps to the distal long arm of chromosome 5. Cytogent Cell Genet 49:313-314.

Giuffra LA et al (1989) A linkage group of 14 markers spanning >100cM on chromosome 5 based on non-CEPH families. Cytogenet Cell Genet 51:1004.

Goate AM et al (1989) Predisposing locus for Alzheimer's disease on chromosome 21. Lancet i:352-355.

Goate A et al (1991) Segregation of a missense mutation in the amyloid precursor protein gene with familial Alzheimer's disease. Nature 349:704-706.

Goldin LR et al (1988). Power of the affective sib pair method for heterogeneous disorders. Genetic Epidemiology 5:35-42.

Gottesman II et al (1967). A polygenic theory of schizophrenia. Proc Natl Acad Sci USA 58:199-205.

Gottesman II et al (1972). Schizophrenia and genetics: A twin study vantage point. New York, Academic Press.

Gottesman II et al (1976). A critical review of recent adoption, twin and family studies of schizophrenia: behavioral genetics perspectives. Schizophr Bull 2:360-398.

Gottesman II et al (1982). Schizophrenia the epigenetic puzzle. Cambridge University press, Cambridge.

Gottesman II et al (1989). Confirming unexpressed genotyes for schizophrenia. Arch Gen Psychiatry 46:867-872.

Grandy DK et al (1989) The human dopamine D2 receptor gene is located on chromosome 11 at q22-q23 and identifies a TaqI RFLP. Am J Hum Genet 45:778-785.

Grandy DK et al (1990) A human D1 dopamine receptor gene is located on chromosome 5 at q35.1 and identifies an EcoRI RFLP. Am J Hum Genet 47:828-834.

Green JR et al (1977). Sibling method for detecting HLA-linked genes in a disease. Tissue Antigens 9:31-35.

Griesinger W (1861). Die pathologie und therapie der psychischen krankheiten. Stuttgart:Krabbe.

Groden J et al (1991) Identification and characterization of the familial adenomatous polyposis coli gene. Cell 66:589-600.

Gunderson JG et al (1983). The search for a schizotype: crossing the border again. Arch Gen Psychiatry 40:15-22.

Gurling HMD (1986a) Application of molecular biology to mental illness. Analysis of genomic DNA and brain mRNA. Psychiatric Dev 3:257-273.

Gurling HMD (1986b) Candidate genes and favoured loci: Strategies for molecular genetic research int schizophrenia, manic depression, autism, alcoholism and Alzheimer's disease. Psychiatric Dev 4:289-309.

Gurling HMD et al (1991). Genetic linkage studies of schizophrenia. In The new genetic of mental illness. Butterworth-Heinemann Ltd, Oxford.

Gurling HMD et al (1992). Genetic linkage analysis and magnetic resonance imaging in British and Icelandic pedigrees. In New genetic research in psychiatry. Am Psychopathological assoc series. New York:Raven Press.

Gusella JF et al (1983). A polymorphic DNA marker genetically linked to Huntingtons's disease. Nature 306:234-238.

Guze SB et al (1983). A follow up and family study of schizophrenia. Arch Gen Psychiatry 40:1273-1276.

Hadano S et al (1991) Laser microdissection and single unique primer PCR allow generation of regional chromosome DNA clones from a single human chromosome. Genomics 11:364-373.

Hamada H et al (1982) A novel repeated element with Z-DNA forming potential is widely found in evolutionarily diverse eukaryotic genomes. Proc Natl Acad Sci USA 79:6465-6469.

Hamada H et al (1984) Enhanced gene expression by the poly(dT-dG).(dC-dA) sequence. Mol Cell Biol 4:2622-2630.

Hamilton M (1984). Fish's Schizophrenia. John Wright PSG Inc.

Hanauer A et al (1990) The Friedreich ataxia gene is assigned to chromosome 9q13-q21 by mapping of tightly linked markers and shows linkage disequilibrium with D9S15. Am J Hum Genet 46:133-137.

Harley HG et al (1992) Expansion of an unstable DNA region and phenotypic variation in myotonic dystrophy. Nature 355:545-546.

Harrison PJ et al (1991) Decreased hippocampal expression of a glutamate receptor gene in schizophrenia. Lancet 337:450-452.

Harvald B et al (1965). Heredity factors elucidated by twin studies, in genetics and the epidemiology of chronic diseases: Public health service publication 1163. Washington DC, US department of health, education and welfare.

Harvey I et al (1992). Reduction of cortical volume in schizophrenia on MRI. Arch Gen Psychiatry (in press).

Herrling PL (1989) Clinical implications of NMDA receptors. In The NMDA receptor Eds Watkins JC and Collingridge GL. IRL Press pp177-185.

Heston LL (1966). Psychiatric disorders in foster home reared children of schizophrenic mothers. Br J Psychiatry 112:819-825.

Heston LL (1970). The genetics of schizophrenia and schizoid disease. Science 167:249-256.

Higgs DR et al (1986) Analysis of the human alpha-globin gene cluster reveals a highly informative genetic locus. Proc Natl Acad Sci USA 83: 5165-5169.

Hoffman EP et al (1987) Dystrophin: The protein product of the Duchenne Muscular Dystrophy locus. Cell 51:919-928.

Holloman M et al (1989) Cloning by functional expression oof a member of the glutamate receptor family. Nature 342:643-648.

Holzman PS et al (1984). Pursuit eye movement dysfunctions in schizophrenia. Arch Gen Psychiatry 41:136-139.

Holzman PS et al (1988). A single doominant gene can accunt for eye tracking dysfunctions and schizophrenia in offspring of discordant twins. 45:641-647.

Hu Nt et al (1982) The making of strand specific M13 probes. Gene 17:271-277.

Human gene mapping 11 (1991) 11th international workshop on human gene mapping, London. Cyto Cell Genet 58 (1-4).

Ikonen E et al (1990) Huntington disease in Finland:Linkage disequilibrium of chromosome 4 RFLP haplotypes and exclution of a tight linkage between the disease and D4S43 locus. Am J Hum Genet 46:5-11.

Imparin CC et al (1987) Analysis of DNA extracts from formalin-fixed, parraffin embedded tissues by enzymic amplification and hybridization with sequence specific oligonucleotides. Biochem Biophys Res Comm 142:710-716.

Inouye E (1963). Similarity and dissimilarity of schizophrenia in twins, in proceedings, third world congress of psychiatry. Vol 1 Montreal, University of Toronto Press.

Jacobs BL (1984) *in*: Hallucinogens: Neurochemical, Behavioral and Clinical Perspectives, pp 183-202. Raven, New York

Jacobsen B et al (1980). Perinatal complications in adopted and non-adopted schizophrenics and their controls: preliminary results. Acta Psychiatr Scand 285:337-346.

Jeffreys AJ et al (1988) Spontaneous mutation rates to new length alleles at tandem repetitive hypervariable loci in humane DNA. Nature 332:278-281.

Johnston BH (1988) The S1-sensitive form of $d(C\text{-}T)_n.d(A\text{-}G)_n$: chemical evidence for a three - stranded structure in plasmids. Science 241:1800-1804.

Julier C et al (1991) Insulin IGF2 region oon chromosome 11p encodes a gene implicated in HLA DR4 dependent diabetes susceptibility. Nature 354:155-159.

Julius D et al (1988) Molecular characterization of a functional cDNA encoding the serotonin 1c receptor. *Science* 241:558-556.

Kariya Y et al (1987) Revision of consensus sequence of human Alu repeats - a review. *Gene* 53:1-10.

Karlsson JL (1974). Inheritance of schizophrenia. Acta Psychiatr Scand Suppl 247.

Karlsson JL (1982) Family transmission of schizophrenia: A review and synthesis. Brit J Psychiatry 140:600-606.

Karlsson JL (1988) Partly dominant transmission of schizophrenia in Iceland. Brit J Psychiatry 152:324-329.

Kaufman DL et al (1990) Detection of point mutations associated with genetic diseases by an exon scanning technique. Genomics 8:656-663.

Kaufmann CA et al (1989) Physical mapping, linkage analysis of a putative schizophrenia locus on chromosome 5q. Schizophrenia Bulletin 15:441-452.

Kay DWK et al (1975). Genetic hypotheses and environmental factors in light of psychiatric moorbidity in the families of schizophrenics. Br J Psychiatry 127:109-118.

Kelso JR et al (1988). Quantitative neuroanatomy in schizophrenia. Arch Gen Psychiatry 45:533-541.

Kelsoe JR et al (1989). Re-evaluation of the linkage of relationship between chromosome 11p loci and the gene for bipolar affective disorder in the Old Order Amish. Nature 342:238-243.

Kendell RE et al (1980). The identification of disease entities and relationship between schizophrenia and affective psychoses. Br J Psychiatry 137:324-331.

Kendler KS (1983). Overview: A current perspective on twin studies of schizophrenia. Am J Psychiatry 140:1413-1425.

Kendler KS et al (1984). An independent analysis of the Danish adoption study of schizophrenia. Arch Gen Psychiatry 41:555-564.

Kendler KS et al (1985). Psychiatric illness in first degree relatives of schizophrenic and surgical control patients. Arch Gen Psychiatry 42:770-779.

Kendler KS (1988). Familial aggregation of schizophrenia and schizophrenia spectrum disorders. Arch Gen Psychiatry 45:377-383.

Kennedy JL et al (1988) Evidence against linkage of schizophrenia to markers on chromosome 5 in a northern Swedish pedigree. Nature 336:167-170.

Kennedy JL et al (1989) Molecular genetic studies in schizophrenia. Schizophrenia Bulletin 15:383-391.

Kennedy JL et al (1991) Linkage studies of the dopamine D4 receptor gene and schizophrenia in Eastern Canadian families. Psychiatric genet 2:81.

Kerem BS et al (1989) Identification of the cystic fibrosis gene: Genetic analysis. Science 245:1073-1079.

Kety SS et al (1975). Mental illness in biologic and adoptive families of adopted individuals who became schizophrenic: a preliminary report based on psychiatric interviews, in Genetic Research in Psychiatry. Edited by Fieve RR, John Hopkins University Press.

Kety SS (1983). Mental illness in the biological and adoptive relatives of schizophrenic adoptees: findings relevant to genetic and environmental factors in etiology. Am J Psychiatry 140:720-727.

Kety SS (1987). The significance of genetic factors in the etiology of schizophrenia: results from the national study of adoptees in Denmark. J Psychiat Res 21:423-429.

Khan MAR et al (1990) A Sac1 RFLP is detected with the 5HTla serotonin receptor probe G-21. Nucleic Acids Res 18:691.

Kidd KK et al (1973). An analysis of the genetics of schizophrenia. Soc Biol 20:254-265.

Kidd KK et al (1974) The role of genetic drift in the differentiation of Icelandic and Norwegian cattle. Evolution 28.

Kidd KK et al (1978). Alternative genetic models for the analysis of complex traits. In The biology of cleft lip and cleft palate. New York:Alan R Liss.

Kmiec EB et al (1986) Homologous pairing of DNA molecules by Ustilago Rec 1 protein is promoted by sequences of Z-DNA. Cell 44:545-554.

Kobayashi K et al (1989) Human dopamine beta hydroxylase gene: Two mRNA types having different 3'terminal regions are produced through alternative polyadenylation. Nuc Acid Res 17:1089-1102.

Koenig M et al (1987) Complete cloning of the Duchenne Muscular Dystrophy (DMD) DNA and preliminary genomic organisation of the DMD gene in normal and affected individuals. Cell 50:509-517.

Kobilka BK et al (1987a) An intronless gene encoding a potential member of the family of receptors coupled to guanine nucleotide regulatory proteins. Nature 329:75-79.

Kobilka BK et al (1987b) Cloning, sequencing and expressionof the gene coding for the human platelet alpha 2 adrenergic receptor. Science 238:650-656.

Kobilka BK et al (1987c) cDNA for the human beta 2 adrenergic receptor: A protein with multiple membrane spanning domains and encoded by a gene whose chromosomal location is shared with that of the receptor for platelet derived growth factor. Proc Natl Acad Sci USA 84:46-50.

Kobilka BK et al (1988) Chimeric alpha 2, beta 2 adrenergic receptors: Delineation of domains involved in effector coupling and ligand binding specificity. Science 240:1310-1316.

Kraepelin E (1904). Clinical psychiatry: A textbook for students and physicians, Diefendorf AR (trans), New York Macmillan Publishing Co Inc.

Kringlen E (1967). Heredity and environment in the functional psychoses: an epidemiological clinical twin study. Oslo, Universitetsforlaget.

Kringlen E et al (1989). Offspring of monozygotic twins discordant for schizophrenia 46:873-877.

Kumar KN et al (1991) Cloning of cDNA for the glutamate-binding subunit of an NMDA receptor complex. Nature 354:70-73.

Kunkel LM et al (1985) Specific cloning of DNA fragments absent from the DNA of a male patient with an X chromosome deletion. PNAS USA 82:4778-4782.

Lalouel (1983) see Hugh

Lander ES et al (1986) Strategies for studying heterogeneous genetic traits in humans by using a linkage map of restriction fragment length polymorphisms. PNAS USA 83:7353-7357.

Lathrop GM et al (1985). Multilocus linkage analysis in humans: detection of linkage and estimation of recombination. Am J Hum Genet 37:482-498.

Lathrop M et al (1988). A mapped set of genetic markers for human chromosome 9. Genomics 3:361-66.

Lee W-H et al (1987) Human retinoblastoma susceptibility gene: cloning, identification and sequence. Science 235:1394-1399.

Leppert M et al (1987) The gene for familial polyposis coli maps to the long arm of chromosome 5 . Science 238:1411-1413.

Levinson G et al (1987) Slipped strand mispairing: a major mechanism for DNA sequence evolution. Mol Biol Evol 4:203-221.

Lewis SW et al (1987). The familial sporadic distinction as a strategy in schizophrenia research. Br J Psychiatry 151:306-313.

Lewis SW (1990). Computerized tomography in schizophrenia. 15 years on. Br J Psychiatry 157(suppl):16-24.

Leysen JE (1988) Drugs as Tools in Neurotransmitter Research. *In* "Neuromethods, Neuropharmacology II"(A.A.Boulten, G.B.Baker, and A.V.Juorio, Eds.). *Humana Press, Clifton, USA.*

Libert F et al (1989) Selective amplification and cloning of four new members of the G protein coupled receptor family. Science 244:569-572.

Lidz T et al (1983). Critique of the Danish-American studies of the biological and adoptive relatives of adoptees who became schizophrenic. Am J psychiatry 140:426-435.

Lifton RP et al (1992) A chimaeric 11 beta-hydroxylase aldosterone synthase gene causes glucocorticooid remediable aldersteronism and human hypertension. Nature 355:262-265.

Litt M et al (1986) Linkage disequilibrium between pairs of loci within a highly polymorphic region of chromosome 2q. Am J Hum Genet 39:166-178.

Litt M et al (1989) A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. Am J Hum Genet 44:397-401.

Lowing PA et al (1983). The inheritance of schizophrenia spectrum disorders: A re-analysis of the Danish adoptee study data. Am J psychiat 140:1167-1171.

Luty JA et al (1990) Five polymorphic microsatellite VNTRs on the human X chromosome. Am J Hum Genet 46:776-783.

Maniatis T et al (1982) Molecular cloning (a laboratory manual) Cold Spring Harbor Laboratory.

Mankoo BS et al (1992) Analysis of polymorphic microsatellite DNA sequences by PCR amplification of paraffin-embedded tissue: A method of genotyping deceased members of pedigrees in genetic linkage studies. Psych Genet 2:161-166.

Martinez M et al (1989). The detection of linkage and heterogeneity in nuclear families for complex disorders: one versus two linked marker loci. Am J Hum Genet 44:552-559.

Matheny AP (1979). Appraisal of parental bias in twin studies: ascribed zygosity and IQ differences in twins.

Acta Genet Med Gemellol 28:155-160.

Matthysse SW et al (1976). Estimating the genetic contribution to schizophrenia. Am J Psychiatry 133:185-191.

Matthysse S et al (1979). Continuous variation caused by genes with graduated effects. Proc Natl Acad Sci USA 76:2862-2865.

Matthysse S et al (1986). The genetic transmission of schizophrenia: Application of mendelian latent structure analysis to eye tracking dysfunctions in schizophrenia and affective disorder. J Psychiat Res 20:57-76.

McGue M et al (1983). The transmission of schizophrenia under multifactorial threshold model. Am J Hum Genet 35:1161-1178.

McGuffin P et al (1983) A family study of HLA antigens and other genetic markers in schizophrenia. Psychol med 13:31-43.

McGuffin P et al (1984). Twin concordance for operationally defined schizophrenia. Arch Gen Psychiatry 41:541-545.

McGuffin P et al (1986) Genetic markers in schizophrenia. Hum Hered 36:65-88.

McGuffin P et al (1990) Exclusion of a schizophrenia susceptibility gene from the chromosome 5q11-13 region: New data and a reanalysis of previous reports. Am J Hum Genet 47:524-535.

McKusick VA (1988). Mendelian inheritance in Man. 8th ed. John Hopkins University Press, Baltimore.

Miesfeld R et al (1981) A member of a new repeated sequence family which is conserved throughout eukaryotic evolution is found between the human delta and beta globin genes. Nucleic Acids Res 9:5931-5947

Melmer G et al (1990) Isolation of clones on chromosome 7 that contain recognition sites for rare-cutting enzymes by oligonucleootide hybridization. Genomics 7:173-181.

Melmer G et al (1991) Gene detection by oligonucleotide hybridization corresponding to splice site concensus sequences. HGM11 abstracts. Cell Cyto Genet 58.

Menozzi P et al (1978) Synthetic maps of human gene frequencies in Europeans. Science 201:786-792.

Merikangas KR et al (1989) Linkage studies of bipolar disorder: methodologic and analytic issues. Arch Gen

Psychiatry 46:1137-1141.

Monaco AP et al (1985) Detection of deletions spanning Duchenne Muscular Dystrophy locus using a tightly linked DNA segment. Nature 316:842-845.

Monaco AP et al (1986). Isolation of candidate cDNAs for portions of Duchenne muscular dystrophy gene. Nature 323:646-650.

Moriyoshi K et al (1991) Molecular cloning and characterization of the rat NMDA receptor. Nature 354:31-37.

Morrison et al (1992) Am J Hum Genet

Morton NE (1955). Sequential tests for the detection of linkage. Am J Hum Genet 7:277-318.

Morton NE et al (1971). Complex segregation analysis. Am J Hum Genet 23:602.

Morton NE et al (1982). Data and theory for a revised chiasma map of man. Hum Genet 62:266-70.

Morton NE et al (1985) Mapping genetic systems by the supratype method. Hum Genet 70:231-235.

Murray JC et al (1987) RFLP for the glucocorticoid receptor (GRL) located at 5q11-5q13. Nucleic Acids Res 15:6765.

Murray R et al (1990). Why is schizophrenia disappearing? Schizophrenia Res 3:5.

Nakamura Y et al (1987) Variable number of tandem repeat (VNTR) markers for human gene mapping. Science 235:1616-1622.

Nakamura Y et al (1988a). An extended genetic linkage map of markers for human chromosome 10. Genomics 3:389-92.

Nakamura Y et al (1988b). A mapped set of DNA markers for human chromosome 17. Genomics 2:302-09.

Nakamura Y et al (1988c). A mapped set of markers for human chromosome 15. Genomics 3:342-46.

Nasrallah HA et al (1983). Clinical differences between schizophrenic patients with and without large cerebral ventricals. J Clin Psych 44:407-409.

Niemegeers CJE et al (1991) 5HT2 receptor antagonists in schizophrenia: Preclinical and clinical

considerations. Biological Psychiatry 1:535-537.

Nordheim A et al (1983) Negatively supercoiled simian virus 40 DNA contains Z - DNA segments within transcriptional enhancer sequences. Nature 303:674-679.

Nurnberg P et al (1989) DNA fingerprints with the oligonucleotide probe $(CAC)_5/(GTG)_5$: somatic stability and germline mutations. Hum Genet 84:75-78.

O"Connell P et al (1988). Twelve loci form a continuous linkage map for human chromosome 18. Genomics 3:367-72.

Olson M et al (1989). A common language for physical mapping of the human genome. Science 245:1434-1435.

Orita M et al (1989) Rapid and sensitive detection of point mutations and DNA polymorphisms using the polymerase chain reaction. Genomics 5:874-879.

O'Rourke DH et al (1982). Refutation of the general loocus model for the aetiology of schizophrenia. Am J Hum Genet 34:630-649.

Ott J (1985a). Analysis of human genetic linkage. John Hopkins University Press, Baltimore.

Ott J (1985b) A chi square test to distinguish allelic association from othercauses of phenotypic association between two loci. Genet Epidemiol 2:79-84.

Ott J (1986). The number of families required to detect oor exclude linkage heterogeneity. Am J Hum Genet 39:159-165.

Overhauser J et al (1987) Identification of 28 DNA fragments that detect RFLPs in 13 distinct physical regiooons of the short arm of chromosome 5. Nucl Acids Res 15:4617-4627.

Owen MJ et al (1988). Obstetric complications of schizophrenia: A CT study. Psychological Medicine 18: 331 340.

Pardue ML et al (1987) $(dC\text{-}dA)_n(dG\text{-}dT)_n$ sequences have evolutionary conserved chromosomal locations in Drosophila with implications for roles in chromosome structure and function. EMBO J 6:1781-1789.

Paterson AH et al (1988). Resolution of quantitative traits into Mendelian factors by using a complete linkage map of restriction fragment lenght polymorphisms. Nature 335:721-726.

Pearlson GD et al (1985). Symptomatic, familial, perinatal and social correlates of computerized axial tomography (CAT) changes in schizophrenics and bipolars. J Nervous and Mental diseases.

Plomin R et al (1976). Resemblance in appearance and the equal environments assumption in twin studies of personality traits. Behav Genet 6:43-52.

Poduslo SE et al (1991) Detecting high resolution polymorphisms in human coding loci by combining PCR and single strand conformation polymorphisms (SSCP) analysis. Am J Hum Genet 49:106-111.

Pope HG et al (1982). Failure to find evidence of schizophrenia in first degree relatives of schizophrenic probands. Am J Psychiatry 139:826-828.

Pritchard C et al (1991). The end in sight for Huntington disease? Am J Hum Genet 49:1-6.

Pritchett DB et al (1989) Structure and functional expression of cloned rat serotonin 5HT-2 receptor. *EMBO J* 7:4135-4140.

Prober JM et al (1987). A system for rapid DNA sequencing with fluorescent chain-terminating dideoxynucleotides. Science 238:336-341.

Ramsay M et al (1988) Alpha - globin gene cluster haplotypes in the Kalahari San and Southern African bantu speaking blacks. Am J Hum Genet 43:527-533.

Rao BS et al (1988) Pausing in simian virus 40 DNA replication by a sequence containing $(dG\text{-}dA)_{27}.(dT\text{-}dC)_{27}$. Nucleic Acids Res 16:8077-8094.

Rao DC et al (1981). Path analysis of qualitative data on pairs of relatives: Application to schizophrenia. Hum Hered 31:325-333.

Reich T et al (1972). The use of multiple thresholds in determining the mode of transmission of semicontinuous traits. Ann Hum Genet 36:163-184.

Reider RO et al (1978). Genetic strategies in biological psychiatry. Arch Gen Psychiatry 35:866-873.

Renwick JH (1971). The mapping of human chromosomes. Ann Rev Genet 5:81-120.

Reveley AM et al (1984). Cerebral ventricular enlargement in nongenetic schizophrenia: A controlled twin study. Br J Psychiatry 144:89-93.

Reynolds GP (1989) Beyond the dopamine hypothesis; The neurochemical pathology of schizophrenia.

Brit.J.Psychiatry 155:305-316.

Riordan JR (1989) Identification of the cystic fibrosis gene: Cloning and characterization of complementary DNA. Science 245:1066-1073.

Risch N et al (1984). Segregation analysis of schizophrenia and related disorders. Am J Hum Genet 36:1039-1059.

Risch N (1988) A new statistical test for linkage heterogeneity. Am J Hum Genet 42:353-364.

Risch N (1990a) Linkage strategies for genetically complex traits II. The power of the affective relative pairs. Am J Hum Genet 46:229-241.

Risch N (1990b). Linkage strategies for genetically complex traits. 1 multilocus models. Am J Hum Genet 46:222-228.

Romani A et al (1987). P300 and CT scan in patients with chronic schizophrenia. Br J Psychiatry 151:506-513.

Romano J et al (1987) Marfan's syndrome and schizophrenia: a case report. Arch Gen Psychiatry 32:466-476.

Rommens JM et al (1989). Identification of the cystic fibrosis gene: chromosome walking and jumping. Science 245:1059-1065.

Rose EA et al (1990) Complete physical map of the WAGR region of 11p13 localizes a candidate Wilm's tumor gene. Cell 60:495-508.

Rosenthal D (1960) Confusioon of identity and frequencyof schizophrenia in twins. Arch Gen Psychiatry 3:297-304.

Rosenthal et al (1971). The adopted-away offspring of schizophrenics. Am J Psychiat 128:87-91.

Rouleau GA et al (1989). A genetic linkage map of the long arm of human chromosome 22. Genomics 4:1-6.

Royer-Pokora B et al (I986) Cloning the gene for an inherited human disorder-chronic granulomatous disease on the basis of its chromosomal location. Nature 322:32-38.

Royle NJ et al (1987) Preferential localisation of hypervariable minisatellites near human telomeres. Cytogenet Cell Genet 46:685

Rudin E (1916). Studien uber Verebung und Entstehung geistiger Storungen I. Zur Vererbung und Neuenstehung der Dementia Praecox. Berlin Springer.

Sadock J et al (1989). Comprehensive textbook of psychiatry (IV). Baltimore, Williams and Wilkins.

Saiki RK et al (1988) Primer directed amplification of DNA with a thermostable DNA polymerase. Science 239:487-491.

Sanger F et al (1977) DNA sequencing with chain terminating inhibitors. *Proc Natl Acad Sci USA* 74:5463-5467.

Sarkar G et al (1991) Direct sequencing of the dopamine D2 receptor (DRD2) in schizophrenics reveals three polymorphisms but no structural change in the receptor. Genomics 11:8-14.

Scneider K (1959). Clinical Psychopathology (Hamilton MW, trans). London, New York: Grune and Stratton.

Sealey PG et al (1985) Removal of repeated sequences from hybridization probes. *Nucleic Acids Res* 13:1905-1922.

Shelton RC et al (1986). X-ray computerized tomography studies in schziophrenia: a review and synthesis. In: Handbook of schizophrenia, Vol 1 The neurology of schizophrenia, Elsevier Science Pub., Amsterdam.

Shero JH et al (1991) Yeast artificial chromosome vectors for efficient clone manipulation and mapping. Genomics 10:505-508.

Shibata D et al (1988) Analysis of DNA sequences in forty-year-old parraffin embedded thin tissue sections: a bridge between molecular biology and classical histology. Cancer Research 48:4564-4566

Silverman GA et al (1991) Yeast artificial chromosome cloning of two-megabase-size contig within chromosomal band 18q21 establishes physical linkage between BCL2 and plasminogen activator inhibitor type-2. Genomics 9:219-228.

Skolnick MH et al (1984). Possible linkage of a breast cancer-susceptibility locus to the ABO locus: sensitivity of lod scoresto a single new recombinant observvation. Genetic Epidemiology 1:363-373.

Slater E (1953). Psychotic and neurotic illnesses in twins. London, Her Majesty's Stationary Office.

Slater E (1958). The monogenic theory of schizophrenia. Acta Genet Stat Med 8:50-56.

Slater E (1971a). Assortative mating, in man, mind, and heredity. Baltimore, John Hopkins University Press.

Slater E et al (1971b). The genetics of mental disorders. London: Oxford university press.

Slightom JL et al (1980) Human fetal G-gama and A-gama globin genes: complete nucleotide sequences suggest that DNA can be exchanged between these duplicated genes. Cell 21:627-638.

Smeets HJM et al (1989) Use of variable simple sequence motifs as genetic markers: application to study of myotonic dystrophy. Hum Genet 83:245-251.

Smith M et al (1990) Mapping of a gene determining Tuberous Sclerosis to human chromosome 11q14-23. Genomics 6:105-114.

Sokoloff P et al (1990) Molecular cloning and characterization of a novel dopamine receptor (D3) as a target for neuroleptics. Nature 347:146-151.

Southern EM (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. J Mol Biol 98:503-517.

Spitzer RL et al (1978a). Research diagnostic criteria rationale and reliability. Archives of General Psychiatry 35:773-782.

Spitzer RL et al (1978b) Research diagnostic criteria for a selected group of functional disorders 3rd edn (New York state psychiatric institute).

Spitzer RL et al (1979). Crossing the border into borderline personality and borderline schizophrenia: The development of criteria. Arch Gen Psychiatry 36:17-24.

St Clair D et al (1988). Measuring the course of Alzheimer's disease. A longitudinal study of neuropsychological function and changes in P3 event-related potential. Br J Psychiatry 152:48-54.

St Clair D et al (1989) No linkage of 5q11-13 markers to schizophrenia in scottish families. Nature 339:305-309.

St Clair D et al (1990) Association within a family of a balanced autosomal translocation with major mental illness. Lancet 336:13-16.

St George-Hyslop PH et al (1987) The genetic defect causing familial Altzheimer's disease maps on chromosome 21. Science 235:885-890.

St George-Hyslop et al (1990) Genetic linkage studies suggest that Alzheimer's disease is not a single homogeneous disorder. Nature 347:194-197.

Stephens JH et al (1982). A comparison of nine systems to diagnose schizoophrenia. Psychiatry Res 6:57-67.

Sturt E et al (1985) Can linkage and markerassociation resolve the genetic aetiology of psychiatric disorders? Review and argument. Psychological Medicine 15:455-462.

Sunahara RK et al (1990) Human dopamine D1 receptor encoded by an intronless gene on chromosome 5. Nature 347:80-83.

Sunahara RK et al (1990) Human dopamine D1 receptor encoded by an intronless gene on chromosome 5. *Nature* 347:80-83.

Sunahara RK et al (1991) Cloning of the gene for a human dopamine D5 receptor with higher affinity for dopamine than D1. Nature 350:614-619.

Tanzi RE et al (1988). Genetic linkage map of human chromosome 21. Genomics 3:129-136.

Tautz D et al (1984) Simple sequences are ubiquitous repetitive components of eukaryotic genomes. Nucleic Acids Res 12:4127-4138.

Tautz D et al (1986) Cryptic simplicity in DNA is a major source of genetic variation. Nature 322:652-656.

Taylor MA et al (1975). Manic depressive illness and schizophrenia: A partial validation of research diagnostic criteria utilizing neuropsychological testing. Compr Psychiatry 16:91-96.

Thompson EA et al (1978) Linkage and the power of a pedigree structure, in Genetic Epidemiology Academic Press pp247-253.

Thompson EA (1984). Interpretation of lod scores with a set of marker loci. Genetic Epidemiology 1:357-362.

Thompson EA et al (1988) The detection of linkage disequilibrium between closely linked markers:RFLPs at the AI-CIII apolipoprotein genes. Am J Hum Genet 42:113-124.

Tienari P (1975). Schizophrenia in Finnish male twins, in studies of schizoophrenia. Ed by Lader MH Ashford, Kent, Headley Brothers.

Tienari P et al (1989). The Finnish adoptive family study of schizophrenia. Possible jiont effects of genetic vulnerability and family environment. Br J Psychiatry 155:29-32.

Traystman MD et al (1990) Use of denaturing gradient gel electrophoresis to detect point mutations in the factor VIII gene. Genomics 6:293-301.

Treco D et al (1986) The evolutionarily conserved repetitive sequence d(TG.AC)$_n$ promotes reciprocal exchange and generates unusual recombinant tetrads during yeast meiosis. Mol Biol 6:3934-3947.

Tsuang MT et al (1980). Morbidity risks of schizophrenia and affective disorders among first degree relatives of patients with schizophrenia, mania, depression and surgical conditions. Brit J Psychiat 137:497-504.

Tsuang MT et al (1982). Testing the monogenic theory of schizophrenia: An application of segregation analysis to blind family study data. Br J Psychiatry 140:595-599.

Tsuang MT et al (1983). A search for "schizophrenia spectrum disorders": An application of multiple threshold model to blind family study data. Br J Psychiatry 143:572-577.

Tsuang MT et al (1990) Heterogeneity of schizophrenia conceptual models and analytic strategies. Br J Psychiatry 156:17-26.

Ungvari G (1983). Validity of the ICD-9 schizophrenia classification: A blind family history study. Acta Psychiatr Scand 68:287-296.

Van Broeckhoven C et al (1987) Failure of familial Alzheimer's disease to segregate with the A4 amyloid gene in several European families. Nature 329:153-155.

Van Tol HHM et al (1991) Cloning of the gene for a human dopamine D4 receptor with high affinity for the antipsychotic clozapine. Nature 350:610-614.

Viskochil D et al (1990) Deletions and translocation interupt a cloned gene at the neurofibromatosis type 1 locus. Cell 62:187-192.

Weber JL et al (1989) Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. Am J Hum Genet 44:388-396.

Weber JL (1990) Informativeness of human (dC-dA)$_n$.(dG-dT)$_n$ polymorphism. Genomics 7:524-530.

Weeks DE et al (1988) The affected pedigree member method of linkage analysis. Am J Hum Genet 42:315-326.

Weeks DE et al (1990) Measuring the inflation of the lod score due to its maximization over model parameter values in human linkage analysis. Genet Epidemiol 7:237-243.

Weiffenbach B et al (1991) A genetic linkage map of human chromosome 5 with 60 RFLP loci. Genomics 10:173-185.

Weinberger DR (1984). Computed tomography (CT) findings in schizophrenia: speculation on the meaning of it all. J Psychiatr Res 18:477-490.

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

Weir BS et al (1978) Testing hypotheses about linkage disequilibrium with multiple alleles. Genetics 88:633-642.

Wender PH et al (1974). Crossfostering: a research strategy for clarifying the role of genetic and experimental factors in the etiology of schizophrenia. Arch Gen Psychiatry 30:121-128.

Werner P et al (1991) Cloning of a putative high-affinity kainate receptor expressed predominantly in hippocampal CA3 cells. Nature 351:742-744.

Wetterberg L et al (1979). Genetics and biochemistry of schizophrenia in a defined population. In Catecholamines: Basic and Clinical Frontiers. New York: Pergamon Press, pp 1857-1859.

White RS et al (1985) A closely linked genetic marker for cystic fibrosis. Nature 318:382-384.

Willard HF et al (1986) Detection of restriction fragment length polymorphisms at the centromeres of human chromosomes by using chromosome specific alpha satellite DNA probes: implications for development of centomere based genetic linkage maps. Proc Natl Acad Sci USA 83:5611-5615.

Wing JK et al (1974). The measurement and classification of psychiatric symptoms. Cambridge: Cambridge University press.

Winokur G et al (1972). The Iowa 500, II: a blind family history comparison of mania, depression and schizophrenia. Arch Gen Psychiatry 27:462-464.

Wojciechowski AP etal (1991) Familial combined hyperlipidaemia linked to the apoolipoprotein AI CIII AIV gene cluster on chromosome 11q23-q24. Nature

Wong Z et al (1987) Characterisation of a panel of highly variable minisatellites cloned from human DNA. Ann Hum Genet 51:269-288.

World Health Organization (1989). International Classification of Disease 10th edn. Geneva:WHO.349:161-164.

Wright DK et al (1990) Sample preparation from parraffin embedded tissues. In PCR protocols: A guide to

methods and applications. pp 153-158 Academic press, San Diago.

Wyatt RJ et al (1971) Biochemical and sleep studies of schizophrenia: A review of the literature. Schizophrenia Bull 4:10

Yang-Feng TL et al (1990) Chromosomal organization of adrenergic receptor genes. *Proc Natl Acad Sci USA* 87:1516-1520.

Yu S et al (1991) Fragile X genotype characterized by an unstable region of DNA. Science 252:1179-1181.

Zhou QY et al (1990) Cloning and expression of human and rat D1 dopamine receptors. Nature 347:76-80.

Ziprusky RB et al (1992). Widespread cerebral gray matter volume deficits in schizophrenia. Arch Gen Psychiatry (in press).

Zuliani G et al (1990) A high frequency of length polymorphisms in repeated sequences adjacent to Alu sequences. Am J Hum Genet 46:963-969.

RELEVANT PUBLICATIONS RESULTING FROM THIS THESIS

Sherrington R, Brynjolfsson J, Petursson H, Potter M, Duddlestone K, Barraclough B, Wasmuth J, Dobbs M, Gurling HMD (1988) Localisation of a susceptibility Locus for schizophrenia on chromosome 5. Nature 336:164-167.

Gurling HMD, **Sherrington R**, Brynjolfsson J, Read T, Curtis D, Mankoo BS, Potter M, Petursson H (1989) Recent and future molecular genetic research into schizophrenia. Schizophrenia Bulletin 15:373-382.

Mankoo BS, **Sherrington R**, De La Concha A, Kalsi G, Curtis D, Melmer G Gurling HMD (1991) Two microsatellite polymorphisms at the D5S39 locus. Nucleic Acids Research 19:1963.

Melmer G, **Sherrington R**, Mankoo B, Kalsi G, Gurling HMD (1991) A cosmid clone for the 5HT1A receptor (5HTR1A) reveals a Taq1 RFLP that shows tight linkage to DNA loci D5S6, D5S39 and D5S76. Genomics 11:767-769.

Sherrington RP, Melmer G, Dixon M, Mankoo BJ, Kalsi G, Gurling HMD (1991) Linkage disequilibrium between two highly polymorphic microsatellite sequences. American Journal of Human Genetics 49:966-971.

Mankoo BS, **Sherrington R**, Kalsi G, Melmer G, Brynjolfsson J, Petursson H, Gurling HMD (1992) Analysis of Polymorphic Microsatellite sequences by PCR amplification of paraffin-embedded tissue: A method of genotyping deceased members of pedigrees in genetic linkage studies. Psychiatric Genetics 2:161-166.

Sherrington R, Mankoo B, Dixon M, Curtis D, Kalsi G, Melmer G, Gurling H (1992) Microsatellite polymorphisms for chromosome 5 bands q11.2-13.3. In press Human Heredity.

Sherrington R, Kalsi G, Curtis D, Mankoo B, Melmer G, Brynjolfsson J, Sharma T, Sigmundsson T, Read T, Murphy P, Moloney E, Barraclough B, Petursson H, Gurling H (1992) Increased informativeness of DNA markers fails to confirm the chromosome 5 schizophrenia susceptibility locus (SCZD1). Submitted Nature.