A workshop on genetic linkage studies in schizophrenia was held at Columbia University's Arden House Conference Center in October 1989. This report summarizes the contents of invited talks by Drs. Arno Motulsky and T. Conrad Gilliam and the discussions at the five workshop sessions. Topics of the workshop sessions were (1) diagnostic boundaries and hierarchies in schizophrenia, (2) genetic models and linkage parameters, (3) selection and ascertainment of pedigrees, (4) future extensions of molecular genetics strategies, and (5) possibilities for future collaboration.

A workshop on genetic linkage studies in schizophrenia was held at Columbia University's Arden House Conference Center in Harriman, New York, on October 4-6, 1989, under the sponsorship of the university's Department of Psychiatry, the New York State Psychiatric Institute, and the Franz and Helly Kallmann Memorial Fund. Dr. Herbert Pardes and Dr. Nikki Erlenmeyer-Kimling hosted the meeting, which opened with welcoming remarks from Dr. Pardes, who stressed the timeliness and importance of the conference's theme. Dr. John Rainer pointed out how appropriate the triple sponsorship of the meeting was since some of the early genetic studies of schizophrenia in North America were conducted at Columbia and the New York State Psychiatric Institute by Franz Kallmann.

While most of the meeting followed a workshop format, two exceptions were the formal invited talks presented by Dr. Arno Motulsky on considerations in studying the genetics of common, complex diseases and by Dr. T. Conrad Gilliam on strategies to be pursued once linkage has been established. The workshop topics included (1) diagnostic boundaries and hierarchies in schizophrenia, (2) genetic models and linkage parameters, (3) selection and ascertainment of pedigrees, (4) future extensions of molecular genetic strategies, and (5) exploration of possibilities for future collaboration. For three of the five topics, the participants divided into two groups (workgroups A and B) to discuss the issues and then reconvened to share each group's deliberations; the other two topics were discussed in plenary workshops.

Invited Talks

Dr. Arno Motulsky: "Considerations in Studying the Genetics of Common, Complex Diseases." Dr. Motulsky's talk set the tone for much of the discussion that followed by reminding the participants that schizophrenia is a common, complex disease, with all of the concomitant problems that the investigation of such disorders presents. Dr. Motulsky noted that medical genetics has been most successful with the study of single-locus traits such as sickle-cell anemia. The introduction of recombinant deoxyribonucleic acid (DNA) technology has allowed the enormous variability of DNA between individuals to be analyzed systematically, leading to the development of numerous restriction fragment length polymorphisms and other DNA variants as markers.
These, in turn, have led to remarkable successes in identifying the chromosomal region, if not the exact gene, involved in disorders such as Duchenne's muscular dystrophy, cystic fibrosis (CF), and Huntington's disease (HD). The rapidly increasing density of the human gene map augurs well for the mapping of additional single-locus diseases. Schizophrenia, however, is one of a large group of diseases—including coronary heart disease, hypertension, allergies, diabetes mellitus, bipolar disorder, and many cancers—that shows strong evidence of genetic influence but often displays non-Mendelian segregation. Linkage studies of these disorders are often difficult since such complex disorders require special considerations.

Dr. Motulsky pointed out that one starting point in studying complex diseases is the use of segregation analysis to fit models of inheritance to family and population data. One common model for schizophrenia is the polygenic model, involving many genes of small effect with a threshold for disease. However, although the polygenic model fits morbidity risk data on relatives of schizophrenic probands, these same data can also be explained by models with one, two, or three major genes in a polygenic background. Since a model with a few major genes is potentially resolvable, in contrast to a purely polygenic model, rational prevention and treatment guided by an understanding of the genetics will be most successful if schizophrenia is controlled by a few major genes. While segregation analyses may indicate the best genetic model for schizophrenia, the final "proof" of a putative gene's involvement would have to come through linkage analysis.

Using examples from the study of coronary heart disease and diabetes,

Dr. Motulsky discussed several problems for linkage analysis of complex diseases. One problem is that most pedigrees do not show Mendelian patterns of transmission. In addition, it is difficult to find large pedigrees with many affected members, and those that are identified may be unusual and not generalizable to the disease under study as a whole. Other problems include late or variable age of onset, common occurrence of sex-related differences in expression, variable phenotypic expression, reduced penetrance, significant susceptibility to environmental influences, and diagnostic misassignments. Allelic heterogeneity is not a serious problem, but nonallelic heterogeneity is troublesome. The presence of phenocopies may make matters still more difficult.

Dr. Motulsky considered several approaches for handling these problems. First, it is desirable to understand the physiological pathway from genotype to phenotype and to use intermediate phenotypes instead of uncertain diagnostic endpoints. For instance, in coronary heart disease there are many independent risk factors, such as the genetic hyperlipidemias, that can be studied genetically and pathologically. In schizophrenia, however, better intermediate markers are required than those now available because traits such as eye movement disorder or attention dysfunctions are presumably far removed from a putative genotype. Second, it may be useful to examine candidate genes—that is, genes that are likely to be related biochemically or pathophysio logically to the disease. In coronary heart disease, for example, there are many candidate genes involved in lipid metabolism and blood clotting. Third, it is important to be aware of potential clues from clinical genetics to gene locations—such as additional symptoms associated with specific chromosomal deletions or translocations. Dr. Motulsky was not very optimistic about finding homologues to human genes in animals because the same phenotype (e.g., obesity) often does not have the same genetic cause in humans and animals. Few diseases are exactly the same in mouse and man; dystrophin defects, for example, cause very different phenotypes in the two species. A different approach to detecting putative disease genes would involve looking for association (linkage disequilibrium) between a marker and the disease. However, association studies rarely give clear evidence with respect to complex phenotypes. For instance, while a given apolipoprotein mutation may influence lipid levels, it may not be associated with coronary heart disease since additional genetic influences may determine this more general phenotype.

In summary, Dr. Motulsky discussed some difficulties involved in linkage analysis of common complex diseases such as schizophrenia, and he noted that it can be dangerous to carry out a linkage analysis without a prior segregation analysis demonstrating a major gene. Nonallelic heterogeneity is assumed to be common in these disorders and must always be taken into account. The use of intermediate phenotypes, possible minor chromosomal abnormalities, putative candidate susceptibility genes, and association studies may be helpful.

Dr. T. Conrad Gilliam: "Strategies to Be Pursued Once Linkage Has Been Established." In a talk on postlinkage strategies, Dr. Gilliam summarized strategies for localizing and isolating genes for simple Mendelian disorders, and he discussed projections for
similar studies of schizophrenia. A "simple disorder" was defined as a monogenic disorder with a known mode of inheritance and without complicating factors such as wide genetic heterogeneity, diagnostic uncertainties, or restrictive size or numbers of families. Regarding the recent cloning and sequencing of the CF gene and the simultaneous dissolving of evidence for linkage between chromosome 11 markers and manic depressive illness, Dr. Gilliam commented that "it was the best of times and it was the worst of times."

Dr. Gilliam pointed out that the practical limit to genetic resolution using human pedigree analysis is roughly 1 percent recombination or 1 centimorgan (cM) of genetic distance, which equals (on average) 1 million base pairs in physical distance. For example, in HD, CF, Friedreich's ataxia, polycystic kidney disease, myotonic dystrophy, neurofibromatosis I, and Duchenne's muscular dystrophy, each disease locus has been linked to a marker not more than several cM's away. Characterization of a small region containing the disease locus depends on the unambiguous identification of crossovers between DNA markers and the disease locus. Such recombination events will be difficult or impossible to pinpoint in studies in which disease classification is uncertain. While the use of hierarchical diagnostic schemes may improve the odds of detecting linkage to a disease with ambiguous phenotypes, there will still be relatively large confidence intervals surrounding the disease locus, even if such linkage is detected. Dr. Lodewijk Sandkuijl calculated the 95 percent confidence intervals for markers surrounding familial Alzheimer's disease (AD), bipolar disorder on chromosome 11, and X-linked bipolar disorder as 45 cM, 43 cM, and 25 or 40 cM, respectively, depending on the study.

Dr. Gilliam pointed out that, under current molecular cloning strategies, one would first need to map a putative gene to a region of about one megabase (1 million base pairs) before attempting to clone the gene, which can be a mammoth undertaking. For example, it took several large groups of investigators about 2 years to characterize an approximately 600-kilobase (Kb) (600,000 base pair) region containing the CF gene. Currently, an effective search for a disease gene can be carried out only within a region spanning less than 1-2 megabases. The task of identifying schizophrenia susceptibility genes will require high-resolution genetic and physical maps, together with new cloning strategies and fortified gene data banks.

The strategies and time scales for establishing linkage, defining a minimal genetic region encompassing the disease locus, and identifying the disease gene were discussed for several monogenic disorders and then compared with familial AD and schizophrenia. Five years were required to identify and clone the huge dystrophin gene causing Duchenne- and Becker-type muscular dystrophy after the first report of linkage to a DNA marker in 1981. Four years after the report of linkage, the CF gene was found. On the other hand, the locus for HD was mapped in 1983, and approximately 6 years later we are still not certain exactly where, within a 2-3 million base pair region, the locus resides.

Several factors were discussed that determine the pace and success of a study. Foremost is the availability of informative families for genetic linkage analysis. For example, the genemapping effort to study HD was greatly facilitated by a unique Vennzelan pedigree of about 10,000 individuals with an elevated risk for HD. Linkage to the recessively inherited CF gene was established in 1985, and, within 2 years, analyses of over 200 families with two or more affected children had confined the locus to a 1-cM interval. Linkage to the frequently occurring type I neurofibromatosis was discovered only after excluding roughly 70 percent of the genome, with an international consortium involving 15 centers and 142 families narrowing the interval containing the gene to 1-2 cM 1 year after finding linkage. The identification of cytological abnormalities can greatly accelerate a study; the discovery of two type I neurofibromatosis-associated chromosomal translocations appears to have narrowed the disease locus to a 10-240 Kb chromosomal region, with the prospect of isolating the neurofibromatosis gene in the near future. The chromosomal location of a disease gene may also affect isolation attempts. For example, the extremely telomeric location of the HD gene has hindered the discovery of a distal flanking marker because of the difficulty of cloning and mapping telomeres.

Linkage to familial AD was reported on chromosome 21 in 1987. However, in contrast to the rapid progress in mapping the 'simple' monogenic disorders, little progress has been made in mapping familial AD, and the 95 percent confidence interval for the putative familial AD locus remains approximately 45 cM. While it is too soon to draw conclusions, a similar picture seems to be emerging with manic-depressive illness and schizophrenia. Dr. Gilliam reminded the workgroups that initial reports of linkage between markers and disease status set the stage for one type of "replication" that is not
contingent on the assumption of genetic homogeneity. This replication is the testing of more closely linked markers, which should support linkage in the same families if the initial results are correct.

Dr. Gilliam discussed the need for strategies that complement linkage analysis, such as methods for identifying most genes within a long stretch of the human genome. Since the search for schizophrenia genes will probably be lengthy, in time we can expect the mapping efforts to benefit from the international initiative to map and sequence the entire human genome. Strategies that use genes isolated from a "disease linkage region" as molecular probes might then be used to search for diseasespecific alterations or mutations. Dr. Gilliam also discussed the utility of searching large hospital data bases for schizophrenic patients with other symptoms that may be associated with cytological abnormalities.

The Five Workshop Sessions

Diagnostic Boundaries and Hierarchies in Schizophrenia. The workshops, which were coordinated by Drs. L. Erlenmeyer-Kimling and Anne Bassett (workgroups A and B, respectively), discussed the topic of diagnostic boundaries and hierarchies in schizophrenia. Drs. Jean Endicott and Irving Gottesman were discussion leaders for workgroup A, and Drs. Elliot Gershon and Peter McGuffin were leaders for workgroup B.

Workgroup A. Workgroup A focused chiefly on defining the schizophrenia spectrum and its use in linkage studies. The participants agreed that errors in the proband's diagnosis lead to uninterpretable results, and that errors in the relatives' diagnoses may reduce power and seriously bias the results of linkage analyses. It is essential, then, to choose probands with "core" schizophrenia and to characterize everyone in the pedigree as thoroughly as possible.

The possible use of intervening pathophysiological or clinical variables called endophenotypes (see Gottesman and Shields 1976) to identify "affected" relatives was discussed as premature but potentially powerful. Variables that have been suggested as possible endophenotypes include eye tracking, electrophysiological variables (e.g., P100, P300, and contingent negative variation), attention and information processing (e.g., Continuous Performance Tests [Rosvold et al. 1956], backward masking), electrodermal variables (skin conductance level and response), and structured interviews (e.g., the Minnesota Multiphasic Personality Inventory [MMPI; Golden and Meehl 1979]). Several of these variables, however, have not yet been established as having a genetic basis or even as being trait rather than state indicators.

There was agreement that spectrum or "fringe" phenotypes are more likely to be heterogeneous than "core" diagnoses and, thus, that inclusion of fringe phenotypes in a linkage study may decrease both power and the possibility of replicating results. Therefore, any attempt to replicate a linkage study should begin with the core diagnoses. A question was raised as to why the lod score (log to the base ten of the odds ratio) increased when fringe phenotypes were added in the pedigrees studied by Gurling and colleagues (Sherrington et al. 1988). It was concluded that the discrepancy between the findings of Gurling's group, with 7 positive pedigrees, and the results of other investigators who, cumulatively, have 16 negative pedigrees cannot be explained solely on the basis of disparate spectrum definition. It was suggested that the findings of Gurling's group could have resulted from sampling, unintentionally, several times from one extended pedigree with an isolated disease. Finally, regarding fringe phenotypes, participants in workgroup A emphasized that schizotypal personality disorder is an especially troublesome category, given that it is clearly found in the relatives of probands with disorders other than schizophrenia.

Workgroup B. Workgroup B considered which diagnostic system to use. Dr. McGuffin reported that the European Science Foundation Network has chosen as diagnostic instruments the Schedule For Affective Disorders and Schizophrenia—Lifetime Version (SADS-L; Endicott and Spitzer 1978), the Present State Examination (PSE 10; Wing et al. 1974), and a checklist of operationalized criteria for psychosis using all sources of data (hospital charts, family history, direct interviews). A program called OPCRIT (Operational Criteria) can generate diagnoses according to any of several systems. Data were presented on 144 patients who had been diagnosed according to the several systems that the OPCRIT computer program provides, and it was noted that few patients fulfilled criteria for all of the diagnostic systems. Thus, there is a need for a consensual way to use multiple criteria or a polydiagnostic approach in the various studies. Dr. McGuffin noted that one possible approach for determining which diagnostic system maximizes heritability is to compare the ratios of concordance rates in monozygotic and dizygotic (MZ/DZ) twins generated by several systems.
and select the system that yields the highest ratio; this method was also proposed by Dr. Gottesman in workgroup A. Dr. Kenneth Kendler offered a model to guide the choice of diagnostic thresholds that explicitly recognized the tradeoffs involving sensitivity (the penetrance of the vulnerable genotypes) and specificity (one minus the penetrance of the normal genotype).

Workgroup B then discussed how to order a hierarchy of certainty for acceptable diagnoses. It was agreed that a gradient or continuum of diagnoses consistent with genetic inheritance is needed. This led to a discussion of which specific diagnoses belong in such a hierarchy and whether a continuum of severity is appropriate. However, it was noted that a continuum of severity does not necessarily relate to the underlying genetics; for example, bipolar I and II may be different entities so severity in this situation does not correlate entirely with genetics.

There was consensus that the choice of diagnoses for the continuum should be derived from family studies of schizophrenia that are independent of the linkage studies. Dr. Neil Risch observed that the power added by using spectrum diagnoses in linkage analyses is proportional to the relative risks (i.e., the ratio of their recurrence in the families of schizophrenic probands to their population base rates).

The workgroups reconvened. The final “inclusion hierarchy,” a list of uncertain conditions, and a list of exclusion diagnoses (table 1) developed by workgroup B formed the initial basis for discussion when the two groups reconvened. Members of workgroup A pointed out that the lower categories of the inclusion hierarchy were undoubtedly heterogeneous etiologically. There was considerable discussion about the problems associated with category 2, schizoaffective disorder, and it was suggested that the definition be refined to be chronic schizoaffective disorder, primarily depressed, with chronic thought disorder. Most participants agreed that the proband’s diagnosis needs to be restricted to inclusion category 1, chronic schizophrenia, although some argued for allowing for category 2 diagnoses as well.

In discussing how the inclusion hierarchy would be used in linkage analyses, Dr. Jurg Ott commented on two possibilities. The first is to do several analyses with several different cutoff points. However, this requires methods to correct for the inflation of the lod score, since it is maximized over several models. The second possibility is to treat the hierarchical categories as continuous variables by assigning numerical values to each. Dr. Theodore Reich mentioned another approach that generates a continuous liability scale by quantifying the probability of being a true genetic case, as has been done in the work of Rice et al. (1987). The probabilities can be estimated on the basis of long-term stability of the diagnosis and treatment response. Both of the latter require longitudinal followup, and this needs to be built into the initial study design. Although the participants’ exact rankings on the hierarchy or continuous liability scale did not necessarily agree, the disagreements were not major.

In summary, the workgroups recommended precision and reliability of diagnoses, inclusion of all sources of clinical information, pre-established criteria, and a hierarchy for diagnoses.

**Genetic Models for Schizophrenia and Linkage Parameters.** The second workshop discussed genetic models for schizophrenia.

Table 1. Inclusion and exclusion criteria for a hierarchy of “schizophrenia” diagnoses (Workgroup B)

| Inclusion hierarchy | 1. Chronic schizophrenia *(DSM-III-R)*
|----------------------|----------------------------------|
| 2. + schizoaffective *(DSM-III-R)*
| 3. + paranoid psychosis and other nonaffective psychoses *(DSM-III-R)*
| 4. + schizotypal personality disorder |

| Uncertain conditions | 1. + paranoid personality disorder
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>2. + affective disorder with mood incongruent psychotic features</td>
<td></td>
</tr>
</tbody>
</table>

| Exclusion hierarchy | 1. Personality disorders other than schizotypal and paranoid (schizoid? borderline?)
|---------------------|----------------------------------|
| 2. Chemical dependency
| 3. Neuroses (anxiety disorders, somatoform)
| 4. Unipolar (mood congruent)
| 5. Bipolar (mood congruent) |

1American Psychiatric Association (1987).
Workgroup A. Drs. Reich and Stephen Faraone were discussion leaders for workgroup A. Dr. Reich began by noting that, because linkage results are model dependent, it is necessary to consider which models are viable and which variables should be included in the model. He pointed out that in formulating a model, it is important to take into account the many complex characteristics of psychiatric phenotypes—for example, sex effects; age of onset effects; secular trends, such as the disappearance of catatonic schizophrenia in most Western societies; assortative mating; genetic and phenotypic heterogeneity; ill-defined diagnostic thresholds; diagnostic instability; comorbidity; natural selection, such as the low reproductive rate of schizophrenic patients; ascertainment bias; complex gene-environment interactions; and cultural transmission. Other participants noted that it is probably impossible to take all these effects into account when selecting and analyzing families, especially since some of these effects are hard to measure. The methods of analysis (affected-sib-pair method, affected-pedigree-member method [Weeks and Lange 1988], lod score analysis) are probably differentially robust to the above effects since the affected relative methods, which use mainly information on affected members, require far fewer assumptions about the disease than the traditional lod score method, which uses information on all members of the pedigree. The affected-pedigree-member method is robust to assortative mating, age effects, and natural selection, but these may reduce power.

Based on their discussion of segregation analyses and linkage studies, workgroup A drew the following conclusions:

- Simple genetic models, even including incomplete penetrance and sporadic cases, generally do not fit the family data for psychiatric disorders.
- Except possibly for AD and X-linked major depressive disorder, the existence of major locus forms of psychiatric disorders is not confirmed.
- Genetic heterogeneity is likely but not proven.
- The usual lod score threshold of 3.0 to establish linkage is not valid if multiple models with different definitions of affection status or mode of inheritance are “tried out.”
- The design of linkage studies is often insufficiently defined to permit precise replication, so it is often impossible to determine whether the failure to replicate simply represents a failure to use correct methods.
- Multiple approaches to the design and analysis of linkage studies are needed to achieve our goals.

It is important to collect extended pedigrees, as well as numerous small families for sib-pair and pedigree-member analyses. It is quite difficult first to ascertain large pedigrees and then to extend them by identifying additional affected relatives. If nonsystematic approaches are used, segregation analyses will not be possible. Dr. Reich is currently working on a method for systematic extension of pedigrees, using a modification of a method, suggested by Cannings and Thompson (1977), that will enable both segregation and linkage analysis to be performed on extended pedigrees. The goal of this research is to increase the size of pedigrees by preferentially accessing informative branches, thereby efficiently increasing the power of the extended pedigree.

Workgroup B. Group discussion leaders for workgroup B were Drs. C. Robert Cloninger and Neil Risch. Dr. Risch reviewed recurrence risk observations in the relatives of schizophrenic probands and the implications of such observations for the plausibility of proposed genetic models (Risch 1990). The pattern of observed recurrence risks among relative classes—MZ cotwins, offspring of dual matings, DZ cotwins, first-degree relatives, second-degree relatives, third-degree relatives—is inconsistent with a single major locus (SML) model with no multifactorial background and no environmental effects (McGue and Gottesman 1989). Heterogeneity involving additive SML’s is also not likely. Models that cannot be excluded are the usual multifactorial model with threshold effect and a mixed model of SML and polygenic background; in the latter model, however, the SML would not have a large effect. A likely model, which is consistent with the pattern of recurrence risks, is a two- or three-locus system acting epistatically (i.e., alleles at the different loci interacting to influence the phenotype). This model would fit the recurrence risk pattern for schizophrenia because the greater the interaction effects, the sharper the drop-off in recurrence risks over relative classes.

The ensuing discussion centered on the implications of the proposed model for linkage analyses of schizophrenic pedigrees. Dr. Risch noted that if schizophrenia is etiologically homogeneous and is due to an epistatic system with no more than two or three contributing loci, at least one of such loci may be detectable. Dr. Ott mentioned that it is important to incorporate such epistatic effects in linkage analysis programs, as Dr. Lodewijk Sandkuijl
has already started to do. Under this epistatic model, multipoint mapping would help greatly in the detection of linkage. Moreover, it would be clearly advantageous to study a large number of families so that the results would be less sensitive to misclassification and more generalizable. Dr. Risch pointed out that a large number of small families would be best with respect to these concerns. Large, highly loaded families present problems because the results might be difficult to replicate. A question was raised about how to handle the fact that penetrance is correlated within families; while this could affect the estimate of recombination, however, it would not pose problems for linkage detection.

The workgroups reconvened. At the outset of the plenary session when the two workgroups reconvened, workgroup B presented the following conclusions:

- The SML model with nonfamilial environmental factors is unlikely.
- Heterogeneity involving multiple independent SML's is also unlikely.
- Oligogenic (i.e., several major genes, ranging between 2 and 10) models with epistasis or multifactorial models involving gene-gene or gene-environment interactions are more likely than the models above.
- Power calculations indicate that, to detect linkage, larger samples are needed than have been planned. The collection and analysis of such large samples will require both standardization and collaboration.

Even if the true mode of inheritance is oligogenic, linkage might still be detected by analyzing under an SML model. However, this might require doing "brute force" exclusion mapping first, and it is questionable whether that is possible yet. Detection of linkage may depend on whether there is a major gene. Another problem is that only 10 to 20 percent of families of schizophrenic probands are multiplex families, and it is uncertain whether conclusions about schizophrenia in general can be drawn from such a small subset of families.

Some participants suggested that, if positive results are obtained in a linkage study, the search should not stop there but should continue looking elsewhere in the genome for other genes that may be contributing to the etiology of schizophrenia. Other members of the group argued that, once a gene is detected, the most important effort is to identify its role since understanding and fine mapping of one gene may add power to the search for a second.

Selection and Ascertainment of Pedigrees. The third workshop, on selection and ascertainment of pedigrees, was led by Drs. Raymond Crowe and Lodewijk Sandkuijl in workgroup A and by Drs. Jurg Ott and Lynn DeLisi in workgroup B.

Workgroup A. Workgroup A discussed four topics involving selection and ascertainment:

- Whether to use genetically isolated populations or broader, more heterogeneous populations. It was concluded that if larger populations are available, genetic isolates are not preferable since results based on the isolate may not be generalizable to other populations. Also, if the genetic isolate is quite small, there may not be sufficient informative meioses to fine map a gene, even if loose linkage can be detected.
- Whether to use nuclear families or large, extended pedigrees. Although it was generally agreed that the bigger the pedigree, the better and that under homogeneity the best pedigree is one in which schizophrenia is segregating down more than one branch from a common founder, several drawbacks to large extended pedigrees were mentioned. For example, such pedigrees are often difficult to find (but see workgroup B's discussion). In addition, if the disease is quite heterogeneous, a large pedigree may be difficult to analyze because two different disease genes may be segregating in the same pedigree. It was also noted that conclusions based on a single large pedigree may not be generalizable to the disease as a whole.
- What inclusion criteria to formulate. The recommended inclusion criterion was core schizophrenia in the proband and in at least one other family member. Criteria suggested for exclusion of a pedigree included (1) presence of bipolar disorder, especially if it occurs in early generations; and (2) assortative mating, although it was not clear how far back in the pedigree to check to determine that schizophrenia is being transmitted on just one side of the family.
- What ascertainment scheme to use to ensure replicability. No specific ascertainment scheme was decided on, although members of workgroup A agreed that the ascertainment method needs to be systematic and reportable so it can be replicated. For schizophrenia, it may be necessary to sample a larger number of families than expected because there will be losses of potential probands owing to the high suicide rate among schizophrenic individuals. Also, potentially informative families may be lost because of uncooperative probands and their relatives.

Workgroup B. Workgroup B considered four issues. The first was the availability of multiplex families with
schizophrenia (i.e., families with at least two affected individuals). Dr. DeLisi commented that it is very difficult to find such families and referred to a report at the First World Congress on Psychiatric Genetics in which it was noted that a large catchment area in Maryland had yielded relatively few multiplex families. Several other members of workgroup B disagreed about the scarcity of appropriate families and pointed to their successes in identifying multiplex families in various geographic areas (Dr. Arthur Falek in Georgia, especially Appalachia; Drs. Cloninger and Reich in Missouri; Dr. Bassett in the Canadian Maritime Provinces; Dr. Vincent Raymond in Quebec). On the other hand, Dr. Kendler stated that he had ascertained only 20–25 multiplex families in an Irish county with a population base of 60,000 (Assuming a 1% rate of schizophrenia, this would mean that only 3–4% of schizophrenic patients in this population had an affected family member). Dr. McGuffin noted that there are clear regional differences in the rate of multiplex families. For example, he has found it much easier to ascertain multiplex families in Wales than in London.

The second issue, which workgroup B discussed at some length, was whether ascertainment rules are important for linkage studies. Whereas several members argued that systematic ascertainment is very important and that ascertainment strategies need to be standardized, others felt that, given the difficulties of locating large numbers of multiplex families, systematic ascertainment may not be an efficient approach. Dr. Gershon did not believe that a volunteer population poses problems for linkage detection, although he acknowledged that such a population does present problems for replication and parameter estimation. It was noted, though, that volunteer families may often have multiple disturbances and thus be less informative than is desired. Dr. Cloninger called attention to the fact that inconsistencies in the linkage results for schizophrenia and bipolar disorder may be at least partly attributable to differences in ascertainment strategies.

Workgroup B was concerned about ensuring that families under study are free of other psychiatric disorders. Ideally, in a linkage study of schizophrenia, the rate of other disorders in the pedigrees under study should not exceed the rates in the general population. For example, since alcoholism rates are high in many populations, investigators can expect to find alcoholism in families that have schizophrenic members. However, the alcoholism rate in such families should not exceed that of the general population. Although there was some disagreement about the relationship between schizophrenia and bipolar disorders, Dr. Cloninger concluded that the weight of evidence suggests that these two disorders are not related. However, in the best of situations, schizophrenia families in which bipolar disorder is being transmitted should be discarded since the risk of diagnostic misclassification of spectrum cases is increased in such families.

The problem of unilateral versus bilateral families was also discussed. As a rule, unilateral families are preferred because bilaterality can obscure transmission patterns. Investigators should obtain a detailed family history from someone who knows each side of the family. In addition, all members on the side with the illness should be interviewed. However, because many apparently unilateral families turn out to be bilateral on further examination, it is important to investigate carefully those parts of the family that are seemingly free of schizophrenia. Only if the family appears to have unilateral transmission of schizophrenia should it be included for study.

The workgroups reconvened. When the two workgroups reconvened on the topic of selection and ascertainment, workgroup A presented the following conclusion: the ability to replicate is essential. This requires that the mode of ascertainment be precisely defined and replicable, that protocols be well described and available to other investigators, and that rules for extending the family be explicit. Again, because of the need for replication, families from small genetic isolates are not desirable. Moreover, families must include at least one other individual with core schizophrenia besides the proband and must exclude bipolar members and married-in individuals with schizophrenia or bipolar disorder. Finally, the family data must be usable for segregation analysis.

Much of the ensuing discussion of the two workgroups centered on a debate about the need for replication and the consequent need for systematic ascertainment versus how many and what type of families are actually available in the field. Several participants reiterated that high priority should be given to collecting data sets in which segregation analyses, as well as linkage analyses, can be conducted. According to these participants, random approaches to ascertainment may suffice if the genetics of the disease are simple. If the genetics are complicated, however, as in schizophrenia, the ascertainment scheme is very important. Participants arguing against stringent...
rules of ascertainment commented that strict rules may compromise the ability to find any families when, in fact, many families are needed. No resolution was reached on this topic, which remains a critical problem in practice.

Future Extensions of Molecular Genetic Strategies. Workshop four was conducted as a joint session, led by Drs. Gilliam, Maja Buca, and Charles Kaufmann. Dr. Kaufmann began the session with a discussion of three approaches to linkage analysis: searches using anonymous DNA markers, favored loci, and candidate genes. The anonymous marker approach systematically scans the entire genome by testing a large number of regularly spaced DNA markers. This is by far the most labor-intensive approach, possibly requiring the testing of over 300 markers, and it is hindered by the fact that certain regions of the genome are not yet spanned by a set of sufficiently close DNA markers.

The favored locus approach seeks to identify areas of chromosomes that may be involved in the etiology of a disease. For example, screening of dysmorphic-affected individuals for chromosomal abnormalities might yield cytological evidence implicating particular chromosomal regions. However, these cases of schizophrenia could reflect a nonspecific effect of the cytological abnormalities, as appears to be true of a certain subset of patients with mental retardation. Moreover, this approach could be confounded by cytological abnormalities induced by certain antipsychotic medications.

The candidate gene approach focuses on those genes whose abnormal function could be presumed to lead to the pathophysiology underlying the disease. This approach is attractive because it can be very easy to exclude a particular candidate locus through linkage analysis. Additionally, a positive lod score below the traditional threshold of 3.0 may suffice to establish linkage since one is presumably searching a very limited number of loci. However, since approximately 50 percent of human genes may be expressed in the nervous system, there are so many candidate genes to consider that this reduced lod score threshold would no longer be appropriate. A potential disadvantage of candidate genes is that often the DNA markers associated with them are not very polymorphic, presumably due to evolutionary pressure to conserve protein-coding sequences. This can be overcome by the use of more highly polymorphic flanking markers. Dr. James Kennedy pointed out that, when reporting linkage to a complex disease, it has already become the accepted standard to carry out a linkage analysis with flanking markers.

Referring to the example of the dopamine D2 receptor, Dr. Kaufmann raised an additional caveat about exclusion mapping of candidate genes that belong to gene families. Since there appears to be a family of D2 receptors, exclusion of a particular D2 receptor does not exclude another D2 receptor at a different locus. The general problem of heterogeneity was also discussed, and Dr. Gershon cited methods for performing power calculations for candidate gene studies under the assumption of heterogeneity.

One scheme for identifying possible candidate genes was presented by Dr. Kaufmann, who suggested post-mortem comparisons of MZ twins discordant for schizophrenia with respect to messenger ribonucleic acid (mRNA) expression in regions of the brain that are thought to be implicated in the disease. A difference in level of expression for particular species of mRNA could reflect a significant cellular response to some prior environmental event that may have triggered the onset of the disease in one twin but not the other. Unfortunately, other effects, such as drug treatment and differences in age at time of death, could also affect the results, and one might end up with a very large number of candidate genes.

Next, Dr. Gilliam discussed ways to increase detectable polymorphisms, including analysis of dimorphic repeats, use of denaturing gradient gels, and use of computer pattern recognition technology to analyze complex hybridization patterns. Dr. Ott mentioned use of the polymerase chain reaction on individual sperm to map markers at a distance of 0.1 cM. The consensus was that a 1-cM resolution map should be available within the next 5 years. Dr. Gilliam then asked the group how such a fine map would influence linkage analysis. Dr. Ott pointed out that a very fine map eliminates the need to estimate recombination fractions because it is sufficient to assume that the disease-susceptibility gene lies in the midpoint of any interval. However, it could be quite difficult to establish the position of the disease locus within a group of such closely linked markers. Several people commented that the mapping of schizophrenia would not be limited by the resolution of the genetic map but by other factors, such as uncertain diagnoses, reduced penetrance, epistasis, phenocopies, and possible oligogenic etiologies.

Dr. Gershon noted that using markers spaced approximately 20 cM apart was the most efficient approach for an anonymous marker
search. He presented a graph of simulated data showing the relationship between the number of families needed to detect linkage and the average spacing between markers. Below 20 cM, the number of families required did not decrease very much, but the workload increased markedly since every halving of the average distance between markers approximately doubles the number of markers needed.

Next, Dr. Bućan discussed bridges between genetic and physical mapping. She began with methods for saturating a given region with probes, including the use of microdissected chromosomes, somatic cell hybrids, and chromosome-specific libraries. New probes could be quickly mapped through hybridization to somatic cell hybrids and with David Ward's (Lichter et al. 1990) technique of fluorescent in situ hybridization. In addition, physically close probes could be grouped easily by pulse field gel (PFG) analysis, with the most polymorphic then being used for linkage analysis. Pooling of probe data among collaborating labs could be coordinated readily through the distribution of some standard PFG blots from the region of interest.

Discussing the development of new probes through directional cloning, Dr. Bućan mentioned jumping libraries and preparative PFG made from somatic cell hybrids, with the band of interest being excised from the gel and used as a source of DNA to clone either the entire excised fragment or its ends. Currently, areas of less than 1,000 Kb between flanking markers have been spanned successfully with cloned probes. For example, 200 Kb were spanned in the study of Duchenne's muscular dystrophy, 250 Kb in CF, and 600 Kb in HD. The increased use of yeast artificial chromosomes as cloning vectors should expand the size of the regions that can be feasibly covered.

Once a region of interest has been cloned, the next step is to isolate the genes in that area. Regions of DNA rich in cytosine-guanine (CG) often signal the beginnings of genes, so restriction enzymes specific for CG patterns can be helpful in finding genes. Additionally, checking for sequences that are conserved across species by hybridization of probes to "zoo blots" can identify genes.

The next problem is to determine whether an isolated gene is actually responsible for the disease. Dr. Kaufmann suggested adopting Koch's (1891) postulates, which require that the putative gene mutation be consistently present in patients, be absent in normal individuals, and be demonstrated to transmit the disease. The third criterion will be the hardest to satisfy, although one possible solution is to construct transgenic animals with an analogous mutation incorporated into their genome. Questions were raised about even being able to satisfy the first two criteria in the presence of reduced penetrance, phenocopies, and heterogeneity.

The session concluded with a discussion about when to move from linkage studies into physical mapping and gene isolation. No simple answer was forthcoming from any of the participants. Since heterogeneity may hinder the rapid replication of a finding in other populations, Dr. Gilliam recommended waiting until the finding had been strengthened by extension of the original pedigrees and additional marker analysis. But Dr. Gershon questioned the value of pursuing a finding in a single large pedigree since it would be likely that the affected individuals would have inherited a large identical region of DNA around the disease locus, lowering the power of the analysis. Dr. Ott remarked that in the absence of any clear statistical criteria for proceeding, the decision to proceed with physical mapping would largely depend on how much one believed in a given linkage finding.

Directions for Future Collaboration. As the discussion leader in the final plenary workshop about directions for future collaborations, Dr. Bassett noted that collaboration is needed because it would increase power, replicability, and generalizability, all of which were recurrent concerns throughout the meeting. Drs. Carlos Pato and David Shore commented on the recent effort of the National Institute of Mental Health (NIMH) to establish cooperative diagnostic centers to ascertain families. Sites for three diseases—schizophrenia, bipolar disorder, and AD—have been selected as cooperating centers based on their demonstrated capability to identify pedigrees. The problem now is how to achieve a consensus on recruitment approaches, diagnostic assessments, and analytic approaches. Dr. Shore asked what Federal mechanisms would facilitate collaboration (in addition to the NIMH initiative for cooperative centers). Several participants stated that such collaboration can be facilitated at many different levels. The simplest level would involve meetings at which the investigators might simply discuss their differences or, more productively, share information arising from their own experiences. At a different level, investigators might exchange anonymous case materials or 2-point lod scores. Investigators in a collaboration might also be willing to obtain additional information to complete a standardized diagnostic
system agreed upon by the collaboration.

The international collaboration on neurofibromatosis was mentioned as a possible model of collaboration. The neurofibromatosis investigators pooled 2-point lod scores, which permitted the rapid construction of exclusion maps. In addition, they exchanged DNA marker data to construct a multipoint map. Another model of collaboration is the network on schizophrenia and bipolar disorder established by the European Science Foundation, which is aimed at obtaining a large sample of families. In addition to adopting a standardized set of diagnostic assessments for schizophrenia and bipolar disorder, the network has established guidelines for sharing cell lines and has developed a standardized method for storing data. It is possible that North American investigators might set up a voluntary collaborative network similar to these two collaborative models, with exchange of raw diagnostic data for consistency checks at first and establishment of a raw data pool later. The MacArthur Foundation has asked Dr. Cloninger to create a network of investigators working on affective disorders; he expects this new network to be operational in 1990. However, Dr. Reich noted that some research groups have already been collecting families according to ascertainment approaches that would not satisfy the criteria that a collaborative network would probably adopt. Despite the foregoing problem, there was general agreement that collaboration is greatly needed among the investigators working on genetic linkage studies of psychiatric disorders.

Summary. The workshop participants identified several issues about which there were differing viewpoints and several others on which there was total or near total agreement.

Unresolved questions. One unresolved issue was how to balance the need for strict ascertainment rules for genetic linkage studies of schizophrenia versus the need for flexibility, given how few useful families may actually be available. Ascertainment rules were seen by some as essential for replicability and generalizability of linkage findings, an ultimate goal of the field. In addition, segregation parameters may be estimated most accurately only when a specified ascertainment strategy is employed. Rules for extending a pedigree were also thought to be important and contingent upon ascertainment schemes. However, other participants believed that adherence to strict ascertainment rules might impede progress in identifying susceptibility loci for schizophrenia. For example, such rules could preclude use of families discovered adventitiously outside the sampling frame, while the designated sampling frame itself might yield very few families. These participants acknowledged that without an ascertainment method segregation analyses are difficult or impossible, but they argued that the goal of linkage studies is to find susceptibility genes, not to do segregation analyses. They favored studying as many families as possible without worrying unduly about ascertainment.

Another unresolved issue concerned the range of diagnostic phenotypes that would be acceptable as affected cases among the secondary cases in families of schizophrenic probands. Although a diagnostic hierarchy was thought to be valuable, there was disagreement as to which diagnoses should be included and exactly where they should be placed in the hierarchy. It also was not clear how relatives with diagnoses lower in the hierarchy should be handled in the linkage analyses. Among the possibilities offered were (1) to leave uncertain cases out of the analyses—that is, to use a phenotype of "unknown"; (2) to treat such cases as unaffected; (3) to treat hierarchical categories as continuous variables by assigning values to each; or (4) to quantify the probability of being a true genetic case based on such variables as long-term stability of diagnosis, treatment response, etc.

Other unresolved issues involved questions about the most useful type of population and the best type of family and analytic method for linkage studies. While some participants argued that results based on genetically isolated populations were not as generalizable as those based on population-based samples, others stated that the use of an isolate may reduce the possibility of nonallelic heterogeneity. Additionally, it was unresolved which size of family is best for linkage studies: small families (but many of them) or large, extended families (but fewer of them). It was also unresolved whether it would be best to use affected-sib pairs, other pedigree-member pairs, or large pedigrees. A final unresolved question concerned the value of searching for cytological anomalies as possible clues to a chromosomal region contributing to the etiology of schizophrenia.

Consensus. The workshop members agreed on the following issues:

- In each family, the proband and at least one relative should meet strict diagnostic criteria. Most participants agreed that the diagnoses of probands should be core schizophrenia although the addition of schizo-
affective disorder, mainly schizophrenia, was also proposed.

- Diagnostic criteria for a schizophrenia spectrum should be derived from independent family studies, not from the linkage studies themselves.

- Diagnoses should be made according to pre-established criteria and should follow a pre-established hierarchy. All sources of clinical information should be considered when making a diagnosis.

- It would be worthwhile to establish a minimum set of diagnostic data to be collected in all studies, as the European Scientific Foundation Network has done, to enhance the possibilities of future collaboration.

- While an SML may be responsible for schizophrenia in some families, the SML model does not fit all schizophrenia.

- Although there was some disagreement about the best type of family and analytical method for studying linkage (as noted above), the group concluded that it is desirable to use more than one approach. Thus, linkage studies should incorporate both large and small families and employ several methods of analysis (affected-sib-pair method, affected-pedigree-member method, and lod scores).

- After an initial positive linkage finding, it would be prudent to wait before publication until the finding has been strengthened by extension of the pedigree(s) and the use of additional markers.

References


Acknowledgments

We thank Dr. Herbert Pardes for making available support for the workshop from the Department of Psychiatry of Columbia University and the New York State Psychiatric Institute, and Dr. John Rainer for making available support from the Franz and Helly Kallmann Memorial Fund.

The Authors

All of the authors are members of the Department of Medical Genetics, New York State Psychiatric Institute, New York, NY. Daniel E. Weeks, Ph.D., and Linda Brzustowicz, M.D., are Research Scientists; Elizabeth Squires-Wheeler, Ph.D., is a Senior Research Scientist; Barbara
Cornblatt, Ph.D., is a Senior Research Scientist and Assistant Professor, Department of Psychiatry, Columbia University, New York, NY; Thomas Lehner, Ph.D., and Michael Stefanovich, M.D., are Research Scientists; T. Conrad Gilliam, Ph.D., is Head of the Molecular Genetics Laboratory and Assistant Professor, Departments of Psychiatry and Neurology, Columbia University, New York, NY; Jurg Ott, Ph.D., is Head of the Statistical Genetics Unit and Professor, Departments of Psychiatry and Genetics, Columbia University; L. Erlenmeyer-Kimling, Ph.D., is Director of the Division of Developmental Behavioral Studies, and Professor, Departments of Psychiatry and Genetics, Columbia University, New York, NY. Anne Bassett, M.D., is a Clinical Research Fellow, Department of Psychiatry, Columbia University, New York, NY, and is affiliated with the Queen Street Mental Health Centre and The Clarke Institute, Toronto, Canada.

**List of Participants**

Anne Bassett, M.D.
Queen Street Mental Health Centre and The Clarke Institute
Toronto, Canada

Linda Brzustowicz, M.D.
New York State Psychiatric Institute
New York, NY

Maja Bučan, Ph.D.
Wistar Institute
University of Pennsylvania
Philadelphia, PA

William Byerley, M.D.
University of Utah
Salt Lake City, UT

C. Robert Cloninger, M.D.
Washington University
St. Louis, MO

Barbara Cornblatt, Ph.D.
College of Physicians and Surgeons
Columbia University and New York State Psychiatric Institute
New York, NY

Raymond Crowe, M.D.
University of Iowa
Ames, IA

Lynn DeLisi, M.D.
State University of New York at Stony Brook
Stony Brook, NY

Jean Endicott, Ph.D.
College of Physicians and Surgeons
Columbia University
New York, NY

Nikki Erlenmeyer-Kimling, Ph.D.
College of Physicians and Surgeons
Columbia University and New York State Psychiatric Institute
New York, NY

Arthur Falek, M.D.
Georgia Mental Health Institute and Emory University
Atlanta, GA

Stephen Faraone, Ph.D.
Harvard University
Boston, MA

Elliot Gershon, M.D.
National Institute of Mental Health
Bethesda, MD

T. Conrad Gilliam, Ph.D.
College of Physicians and Surgeons
Columbia University
New York, NY

Irving Gottesman, Ph.D.
University of Virginia
Charlottesville, VA

Joachim Hallmayer, M.D.
Stanford University
Stanford, CA

Charles Kaufmann, M.D.
College of Physicians and Surgeons
Columbia University
New York, NY

Kenneth Kendler, M.D.
Medical College of Virginia
Richmond, VA

James Kennedy, M.D.
Yale University
New Haven, CT

Thomas Lehner, Ph.D.
College of Physicians and Surgeons
Columbia University and New York State Psychiatric Institute
New York, NY

Peter McGuffin, M.D.
University of Wales, Cardiff, Wales

Arno Motulsky, M.D.
University of Washington
Seattle, WA

Jurg Ott, Ph.D.
College of Physicians and Surgeons
Columbia University and New York State Psychiatric Institute
New York, NY

Herbert Pardes, M.D.
College of Physicians and Surgeons
Columbia University
New York, NY
Carlos Pato, M.D.
National Institute of Mental Health
Rockville, MD

John Rainer, M.D.
College of Physicians and Surgeons
Columbia University and New York State Psychiatric Institute
New York, NY

Vincent Raymond, M.D., Ph.D.
Centre de Recherche
Laval Robert-Giffard
Beauport, Quebec, Canada

Theodore Reich, M.D.

Washington University
St. Louis, MO

Neil Risch, Ph.D.
Yale University,
New Haven, CT

Lodewijk Sandkuijl, M.D.
Consultant
Delft, The Netherlands

David Shore, M.D.
National Institute of Mental Health
Rockville, MD

Elizabeth Squires-Wheeler, Ph.D.

New York State Psychiatric Institute
New York, NY

Michael Stefanovich, M.D.
New York State Psychiatric Institute
New York, NY

Cathy Walsh, M.B., MRC (Psych.)
Medical Research Council: Clinical Research Centre
Harrow, England

Daniel Weeks, Ph.D.
College of Physicians and Surgeons
Columbia University and New York State Psychiatric Institute
New York, NY