

THE HUMAN GENOME PROJECT

*Francis S. Collins, M.D., Ph.D.;
and Leslie Fink*

The Human Genome Project is an ambitious research effort aimed at deciphering the chemical makeup of the entire human genetic code (i.e., the genome). The primary work of the project is to develop three research tools that will allow scientists to identify genes involved in both rare and common diseases. Another project priority is to examine the ethical, legal, and social implications of new genetic technologies and to educate the public about these issues. Although it has been in existence for less than 6 years, the Human Genome Project already has produced results that are permeating basic biological research and clinical medicine. For example, researchers have successfully mapped the mouse genome, and work is well under way to develop a genetic map of the rat, a useful model for studying complex disorders such as hypertension, diabetes, and alcoholism. **KEY WORDS:** *genome; genetic mapping; DNA; applied research; molecular genetics*

The Human Genome Project is an international research project whose primary mission is to decipher the chemical sequence of the complete human genetic material (i.e., the entire genome), identify all 50,000 to 100,000 genes contained within the genome, and provide research tools to analyze all this genetic information. This ambitious project is based on the fact that the isolation and analysis of the genetic material contained in the DNA¹ (figure 1) can provide scientists with powerful new approaches to understanding the development of diseases and to creating new strategies for their prevention and treatment. Nearly all human medical conditions, except physical injuries, are related to changes (i.e., mutations) in the structure and function of DNA. These disorders include the 4,000 or so heritable “Mendelian” diseases that result from mutations in a single gene; complex and common disorders that arise from heritable alterations in multiple genes; and disorders, such as many cancers, that result from DNA mutations acquired during a person’s lifetime. (For more information on the genetics of alcoholism, see the articles by Goate, pp. 217–220, and Grisel and Crabbe, pp. 220–227.)

Although scientists have performed many of these tasks and experiments for decades, the Human Genome Project is unique and remarkable for the enormity of its effort. The human genome contains 3 billion DNA building blocks (i.e., nucleotides), enough to fill approximately one thousand 1,000-page telephone books if each nucleotide is represented by one letter. Given the size of the human genome, researchers must develop new methods for DNA analysis that can process large amounts of information quickly, cost-effectively, and accurately. These techniques will characterize DNA for family studies of disease, create genomic maps, determine the nucleotide sequence of genes and other large DNA fragments, identify genes, and enable extensive computer manipulations of genetic data.

FRANCIS S. COLLINS, M.D., PH.D., is director and LESLIE FINK is communications officer at the National Center for Human Genome Research at the National Institutes of Health, Bethesda, Maryland.

FOCUS OF THE HUMAN GENOME PROJECT

The primary work of the Human Genome Project has been to produce three main research tools that will allow investigators to identify genes involved in normal biology as well as in both rare and common diseases. These tools are known as positional cloning (Collins 1992). These advanced techniques enable researchers to search for disease-linked genes directly in the genome without first having to identify the gene’s protein product or function. (See the article by Goate, pp. 217–220.) Since 1986, when researchers first found the gene for chronic granulomatous disease² through positional cloning, this technique has led to the isolation of considerably more than 40 disease-linked genes and will allow the identification of many more genes in the future (table 1).

Each of the three tools being developed by the Human Genome Project helps bring the specific gene being sought into better focus (see sidebar, pp. 192–193). The first of these tools, the *genetic* map, consists of thousands of landmarks—short, distinctive pieces of DNA—more or less evenly spaced along the chromosomes. With this tool, researchers can narrow the location of a gene to a region of the chromosome. Once this region has been identified, investigators turn to a second tool, the *physical* map, to further pinpoint the specific gene. Physical maps are sets of overlapping DNA that may span an entire chromosome. These sets are cloned and frozen for future research. Once the physical map is complete, investigators will simply be able to go to the freezer and pick out the actual piece of DNA needed, rather than search through the chromosomes all over again. The final tool will be the creation of a *complete sequence* map of the DNA nucleotides, which will contain the exact sequence of all the DNA that makes up the human genome.

To make all this information available to researchers worldwide, the project has the additional goal of developing computer methods for easy storage, retrieval, and manipulation of data. Moreover, because researchers often can obtain valuable information about human genes and their functions by comparing them with the corresponding genes of other species, the project has set goals for mapping and sequencing the genomes of several important model organisms, such as the mouse, rat, fruit fly, roundworm, yeast, and the common intestinal bacterium *E. coli*.

TECHNOLOGICAL ADVANCES IN GENOMIC RESEARCH

The need for large-scale approaches to DNA sequencing has pushed technology toward both increasing capacity and decreasing instrument size. This demand has led, for example, to the development of automated machines that reduce the time and cost of the biochemical processes involved in sequencing, improve the analysis of these reactions, and facilitate entering the information obtained into databases. Robotic instruments also have been developed that expedite repetitive tasks inherent in large-scale research and reduce the chance for error in several sequencing and mapping steps.

Miniaturization technology is facilitating the sequencing of more—and longer—DNA fragments in less time and increasing the portability of sequencing processes, a capability that is particularly important in clinical or field work. In 1994, for exam-

¹For a definition of this and other technical terms used in this article, see central glossary, pp. 182–183.

²Chronic granulomatous disease is an inherited disease of the immune system.



Figure 1 Artist's rendering of the DNA molecule from a single cell.

ple, the National Institutes of Health (NIH), through its National Center for Human Genome Research (NCHGR), began a new initiative for the development of microtechnologies to reduce the size of sequencing instrumentation and thereby increase the speed of the sequencing process. NCHGR also is exploring new strategies for minimizing time-consuming sequencing bottlenecks by developing integrated, matched components that will help ensure that each step in the sequencing process proceeds as efficiently as possible. The overall sequencing rate is only as fast as its slowest step.

Other developments in DNA sequencing have aimed to reduce the costs associated with the technology. Through refinements in current sequencing methods, the cost has been lowered to about \$0.50 per nucleotide. Research on new DNA sequencing techniques is addressing the need for rapid, inexpensive, large-scale sequencing processes for comparison of complex genomes and clinical applications. Further improvements in the efficiency of current processes, along with the development of entirely new approaches, will enable researchers to determine the complete sequence of the human genome perhaps before the year 2005.

APPLICATIONS OF THE HUMAN GENOME PROJECT

The detailed genetic, physical, and sequence maps developed by the Human Genome Project also will be critical to understanding the biological basis of complex disorders resulting from the interplay of multiple genetic and environmental influences, such as diabetes; heart disease; cancer; and psychiatric illnesses, including alcoholism. In 1994, for example, researchers used genetic maps to discover at least five different chromosome regions that appear to play a role in insulin-dependent (i.e., type 1) diabetes (Davies et al. 1994). Analyses to identify the genetic components of these complex diseases require high-resolution genetic maps and must be conducted on a scale much larger than was previously possible. Automated microsatellite marker technology³ now makes it possible to determine the genetic makeup (i.e., the genotype) of enough subjects so that genes for common diseases can be mapped reliably in a reasonable amount of time. NCHGR is planning a technologically advanced genotyping facility to assist investigators in designing research studies; performing genetic analyses; and developing new techniques for analyzing common, multigene diseases.

³Microsatellite markers are short DNA sequences that vary in length from person to person. The length of a particular marker is inherited from one's parents, allowing researchers to track the markers through several generations of the same family.

Table 1 Disease Genes Identified Using Positional Cloning

Year	Disease
1986	Chronic Granulomatous Disease Duchenne's Muscular Dystrophy Retinoblastoma
1989	Cystic Fibrosis
1990	Wilms' Tumor Neurofibromatosis Type 1 Testis Determining Factor Choroideremia
1991	Fragile X Syndrome Familial Polyposis Coli Kallmann's Syndrome Aniridia
1992	Myotonic Dystrophy Lowe's Syndrome Norris's Syndrome
1993	Menkes' Disease X-Linked A gammaglobulinemia Glycerol Kinase Deficiency Adrenoleukodystrophy Neurofibromatosis Type 2 Huntington's Disease von Hippel-Lindau Disease Spinocerebellar Ataxia I Lissencephaly Wilson's Disease Tuberous Sclerosis
1994	MacLeod's Syndrome Polycystic Kidney Disease Dentatorubral Pallidolouysian Atrophy Fragile X "E" Achondroplasia Wiskott Aldrich Syndrome Early Onset Breast/Ovarian Cancer (<i>BRCA 1</i>) Diastrophic Dysplasia Aarskog-Scott Syndrome Congenital Adrenal Hypoplasia Emery-Dreifuss Muscular Dystrophy Machado-Joseph Disease
1995	Spinal Muscular Atrophy Chondrodysplasia Punctate Limb-Girdle Muscular Dystrophy Ocular Albinism Ataxia Telangiectasia Alzheimer's Disease Hypophosphatemic Rickets Hereditary Multiple Exostoses Bloom Syndrome Early Onset Breast Cancer (<i>BRCA 2</i>)

GENETIC MAPS PROVIDE BLUEPRINT FOR HUMAN GENOME

A primary focus of the Human Genome Project is to develop tools that will enable investigators to analyze large amounts of hereditary material quickly and efficiently. The success of this project hinges on the accurate mapping of each chromosome. The Human Genome Project is using primarily three levels of maps, each of which helps to increase understanding not only of the construction of individual genes but also of their relation to each other and to the entire chromosomal structure.

Genetic Mapping

Genetic mapping, also called linkage mapping, provides the first evidence that a disease or trait (i.e., a characteristic) is linked to the gene(s) inherited from one's parents. Through genetic mapping, researchers can approximate the location of a gene to a specific region on a specific chromosome; the process is like establishing towns on a road map (figure 1). For example, Interstate 10 runs from Florida to California. It would be difficult to find a landmark along that highway if the only cities mapped were Jacksonville and Los Angeles. It would be much easier, however, to pinpoint the landmark if one knew that it was located between markers that are closer together (e.g., El Paso and San Antonio).

Genetic mapping begins with the collection of blood or tissue samples from families in which a disease or trait is prevalent. After extracting the

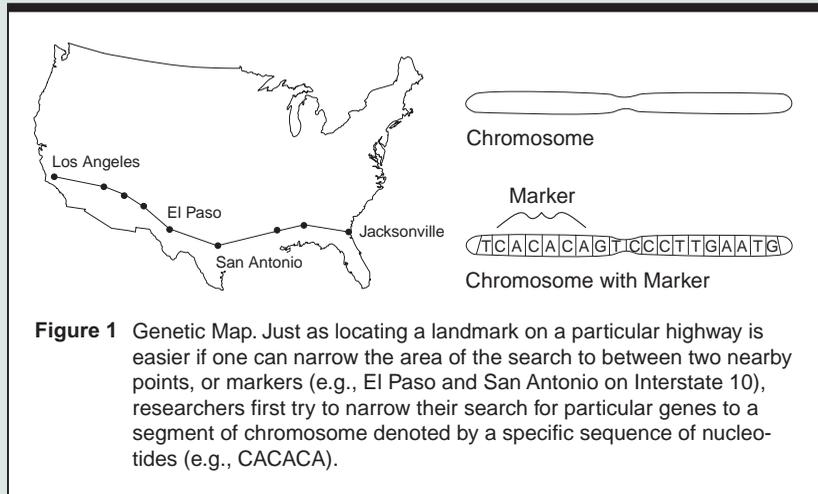


Figure 1 Genetic Map. Just as locating a landmark on a particular highway is easier if one can narrow the area of the search to between two nearby points, or markers (e.g., El Paso and San Antonio on Interstate 10), researchers first try to narrow their search for particular genes to a segment of chromosome denoted by a specific sequence of nucleotides (e.g., CACACA).

DNA from the samples, researchers track linearly the frequency of a recurring set of nucleotides (represented, for example, by the letters "CACACA") along a region of a chromosome. If this sequence is shared among family members who have the disease, the scientists may have identified a marker for the disease-linked gene. Mapping additional DNA samples from other people with and without the disease allows researchers to determine the statistical probability that the marker is linked to the development of the disease.

Physical Mapping

Physical mapping generates sets of overlapping DNA fragments that span regions

of—or even whole—chromosomes. These DNA fragments, which can be isolated and stored for future analysis (figure 2), serve as a resource for investigators who want to isolate a gene after they have mapped it to a particular chromosome or chromosomal region. The physical map allows scientists to limit the gene search to a particular subregion of a chromosome and thus zero in on their target more rapidly.

One early goal of the physical mapping component of the Human Genome Project was to isolate contiguous DNA fragments that spanned at least 2 million nucleotides. Considerable progress has been made in this area, with sets of contiguous DNA

Molecular Medicine

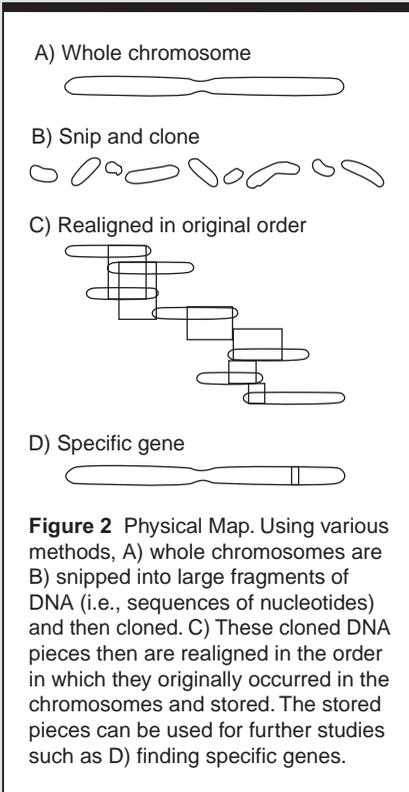
Efforts to understand and treat disease processes at the DNA level are becoming the basis for a new molecular medicine. The discovery of disease-associated genes provides scientists with the foundation for understanding the course of disease, treating disorders with synthetic DNA or gene products, and assessing the risk for future disease. Thus, by going directly to the genetic source of human illness, molecular medicine strategies will offer a more customized health management based on the unique genetic constitution of each person. Molecular medicine also will increase clinicians' focus on prevention by enabling them to predict a person's risk for future disease and offer prevention or early treatment strategies. This approach will apply not only to classical, single-gene hereditary disorders but also to more common, multi-gene disorders, such as alcoholism.

During the past 3 years, positional cloning has led to the isolation of more than 30 disease-associated genes. Although this number has increased dramatically, compared with the years predating the Human

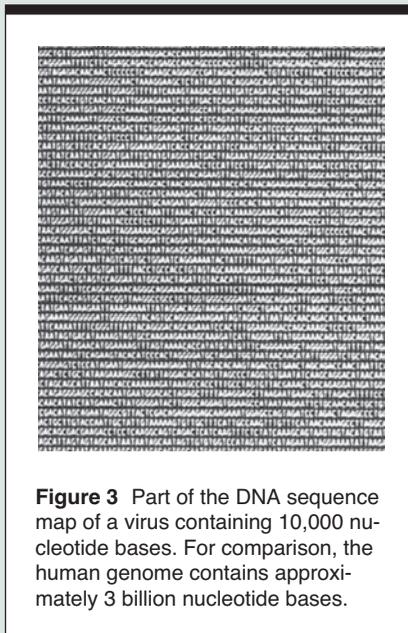
Genome Project, it is still a small fraction of the 50,000 to 100,000 genes that await discovery in the entire genome. NCHGR has helped develop efficient biological and computer techniques to identify all the genes in large regions of the genome. One technique was used successfully last year to isolate *BRCA1*, the first major gene linked to inherited breast cancer. The location of *BRCA1* first was narrowed to a DNA fragment of several hundred thousand nucleotides containing many genes. A process that isolates the protein-coding sequences of a gene (i.e., exon trapping) allowed researchers to identify and examine not only the correct *BRCA1* gene in that region but also several new genes that now serve as disease-gene candidates for future investigations.

Diagnostics

Clinical tests that detect disease-causing mutations in DNA are the most immediate commercial application of gene discovery.



fragments (“contigs”) now frequently ranging from 20 to 50 million nucleotides in length. Because the order of DNA fragments in a physical map should reflect their actual order on a chromosome, correct alignment of contigs also requires a set of markers to serve as mileposts, similar to those of an interstate highway. Genome scientists have developed a physical map that currently contains about 23,000 markers, called sequence tagged



sites (STS’s). Scientists likely will meet their ultimate goal of establishing 30,000 STS markers on the physical map—one every 100,000 nucleotides—within the next year or two. This detailed STS map will allow researchers to pinpoint the exact location of any gene within 50,000 nucleotides of an STS marker.

Researchers also are attempting to use fragments of expressed genes known as expressed sequence tags (EST’s), which are made from complimentary DNA, as markers on the physical genome map. By using EST’s, they hope to increase the power of maps for finding specific genes. A recent collaboration between Merck and Co. (a major pharmaceutical corpora-

tion) and researchers at Washington University in St. Louis, Missouri, will provide a resource for placing tens of thousands of such markers derived from actual genes on the physical map.

Marker development to be used in creating both the linkage and the physical maps also takes into account the need for connectivity between these two types of maps. Information learned from one stage of the gene-finding process must be easily translatable to the next.

The DNA Sequence Map

The Human Genome Project’s most challenging goal is to determine the order (i.e., sequence), unit by unit, of all 3 billion nucleotides that make up the human genome. Once the genetic and physical maps are completed, a sequence map can be constructed, which will allow scientists to find genes, characterize DNA regions that control gene activity, and link DNA structure to its function.

To date, the technology for this work has been developed and implemented primarily in model organisms. For example, researchers now have sequenced 25 million DNA nucleotides from the roundworm—about 25 percent of the animal’s genome—and, in the process, have increased their annual sequencing rate to 11 million nucleotide bases (figure 3). The investigators expect to finish sequencing the roundworm genome by the end of 1998. The complete DNA sequence of yeast and *E. coli* genomes will be determined even sooner.

—Francis S. Collins
and Leslie Fink

These tests may positively identify the genetic origin of an active disease, foreshadow the development of a disease later in life, or identify healthy carriers of recessive diseases such as cystic fibrosis.⁴ Genetic tests can be performed at any stage of the human life cycle with increasingly less invasive sampling procedures. Although DNA testing offers a powerful new tool for identifying and managing disease, it also poses several medical and technical challenges. The number and type of mutations for a particular disease may be few, as in the case of achondroplasia,⁵ or many, as in the case of cystic fibrosis and hereditary breast

⁴For a recessive disease to develop, a person must inherit two altered gene copies, one from each parent. People who inherit only one altered gene copy usually are healthy (i.e., they do not show symptoms of the disease); these people are called asymptomatic carriers.

⁵Achondroplasia is a disorder that results in defective skeletal development in the fetus and dwarfism. Affected children often die before or within their first year of life.

cancer. Thus, it is essential to establish for each potential DNA test how often it detects disease-linked mutations and how often and to what degree detection of mutations correlates with the development of disease.

Therapeutics

Gene discovery also provides opportunities for developing gene-based treatment for hereditary and acquired diseases. These treatment approaches range from the mass production of natural substances (e.g., blood-clotting factors, growth factors and hormones, and interleukins and interferons⁶) that are effective in treating certain diseases to gene-therapy strategies. Gene therapy is designed to deliver DNA carrying a functional gene to a patient’s cells or tissues and thereby correct a genetic alteration.

⁶Interleukins and interferons are substances that stimulate and regulate the immune system.

Currently, more than 100 companies conduct human clinical trials on DNA-based therapies (Pharmaceutical Research and Manufacturers of America [PRMA] 1995). The top U.S. public biotechnology companies have an estimated 2,000 drugs in early development stages (Ernst and Young 1993). Since 1988, NIH's Recombinant DNA Advisory Committee has approved more than 100 human gene-therapy or gene-transfer protocols (Office of Recombinant DNA Activities, NIH, personal communication, April 1995). Seventeen gene-therapy products are now in commercial development for hereditary disorders, cancer, and AIDS (PRMA 1995).

ETHICAL, LEGAL, AND SOCIAL CONCERNS OF THE HUMAN GENOME PROJECT

Implications for Disease Detection

The translation of human genome technologies into patient care brings with it special concerns about how these tools will be applied. A principal arena in which psychosocial issues related to these technologies are being raised is the testing of people who may be at risk for a genetically transmitted disease but who do not yet show the disease's symptoms (i.e., are asymptomatic). These concerns stem largely from the delay between scientists' technical ability to develop DNA-based diagnostic tests that can identify a person's risk for future disease and their ability to develop effective prevention or treatment strategies for the disorders those tests portend. In the meantime, people who undergo genetic tests run the risk of discrimination in health insurance and may have difficulty adapting to test results—particularly in families in which hereditary disease is common—regardless of whether a test indicates future disease. When no treatment is available and when no other medical course of action can be taken on the basis of such tests, the negative social, economic, and psychological consequences of knowing one's medical fate must be carefully evaluated in light of the meager medical benefits of such knowledge.

To help ensure that medical benefits are maximized without jeopardizing psychosocial and economic well-being, the Human Genome Project, from its beginning, has allocated a portion of its research dollars to study the ethical, legal, and social implications (ELSI) of the new genetic technologies. A diverse funding program supports research in four priority areas: the ethical issues surrounding the conduct of genetic research, the responsible integration of new genetic technologies into the clinic, the privacy and fair use of genetic information, and the professional and public education about these issues.

Because of the many unresolved questions surrounding DNA testing in asymptomatic patients, in 1994 NCHGR's advisory body released a statement urging health care professionals to offer DNA testing for the predisposition to breast, ovarian, and colon cancers only within approved pilot research programs until more is known about the science, psychology, and sociology of genetic testing for some diseases (National Advisory Council for Human Genome Research 1994). The American Society of Human Genetics and the National Breast Cancer Coalition have issued similar statements. More recently, the NIH-DOE [Department of Energy] Working Group on ELSI launched a task force to perform a comprehensive, 2-year evaluation of the current state of genetic testing technologies in the United States. The task force will examine safety, accuracy, predictability, quality assurance, and counseling strategies for the responsible use of genetic tests.

In a related project, NCHGR's ELSI branch spearheaded a new group of pilot studies shortly after researchers isolated *BRCA1* and

several genes for colon cancer predisposition. These 3-year studies are examining the psychosocial and patient-education issues related to testing healthy members of families with high incidences of cancer for the presence of mutations that greatly increase the risk of developing cancer. The results will provide a thorough base of knowledge on which to build plans for introducing genetic tests for cancer predisposition into medical practice.

Implications for Complex Traits

Research in human genetics focuses not only on the causes of disease and disability but also on genes and genetic markers that appear to be associated with other human characteristics, such as height, weight, metabolism, learning ability, sexual orientation, and various behaviors (Hamer et al. 1993; Brunner et al. 1993). Associating genes with human traits that vary widely in the population raises unique and potentially controversial social issues. Genetic studies elucidate only one component of these complex traits. The findings of these studies, however, may be interpreted to mean that such characteristics can be reduced to the expression of particular genes, thus excluding the contributions of psychosocial or environmental factors. Genetic studies can also be interpreted in a way that narrows the range of variation considered "normal" or "healthy."

Both reducing complex human characteristics to the role of genes and restricting the definition of what is normal can have harmful—even devastating—consequences, such as the devaluation of human diversity and social discrimination based on a person's genetic makeup. The Human Genome Project must therefore foster a better understanding of human genetic variation among the general public and health care professionals as well as offer research policy options to prevent genetic stigmatization, discrimination, and other misuses and misinterpretations of genetic information.

PROGRESS ON GENETIC AND PHYSICAL MAPS

In the United States, NCHGR and DOE, through its Office of Environmental Health Research, are the primary public supporters of major genome research programs. In 1990, when the 15-year Human Genome Project began, NCHGR and DOE established ambitious goals to guide the research through its first years (U.S. Department of Health and Human Services and U.S. Department of Energy 1990). After nearly 6 years, scientists involved in the Human Genome Project have met or exceeded most of those goals—some ahead of time and all under budget. Because scientific advances may rapidly make the latest technologies obsolete, a second 5-year plan was published in 1993 (Collins and Galas 1993) to keep ahead of the project's progress. Already, further technological advances make it likely that a new plan will be needed, perhaps as early as this year.

In 1994, an international consortium headed by the Genome Science and Technology Center in Iowa published a genetic map of the human genome containing almost 6,000 markers spaced less than 1 million nucleotides apart (Cooperative Human Linkage Center et al. 1994). This map was completed more than 1 year ahead of schedule, and its density of markers is four to six times greater than that called for by the 1990 goals. This early achievement is largely a result of the discovery and development of micro-satellite DNA markers and of large-scale methods for marker isolation and analysis.

In a related project, technology developed so quickly that a high resolution genetic map of the mouse genome was completed in just 2

years. NCHGR is now helping to coordinate an initiative with other NIH institutes, particularly the National Heart, Lung, and Blood Institute and the National Institute on Alcohol Abuse and Alcoholism, to develop a high-resolution genetic map of the rat, a useful model for studying complex disorders such as hypertension, diabetes, and alcoholism.

The original 5-year goal to isolate contiguous DNA fragments that span at least 2 million nucleotides was met early on; soon, more than 90 percent of the human genome will be accounted for using sets of overlapping DNA fragments, each of which is at least 10 million nucleotides long. Complete physical maps now exist for human chromosomes 21, 22, and Y. Nearly complete maps have been developed for chromosomes 3, 4, 7, 11, 12, 16, 19, and X.⁷

As the end of the first phase of the Human Genome Project draws near, its impact already is rippling through basic biological research and clinical medicine. From deciphering information in genes, researchers have gained new knowledge about the nature of mutations and how they cause disease. Even after someday identifying all human genes, scientists will face the daunting task of elucidating the genes' functions. Furthermore, new paradigms will emerge as researchers and clinicians understand interactions between genes, the molecular basis of multigene disorders, and even tissue and organ function.

The translation of this increasing knowledge into improved health care already is under way; however, the value of gene discovery to the promising new field of molecular medicine will be fully realized only when the public is secure in the use of genetic technologies. ■

REFERENCES

BRUNNER, H.G.; NIELEN, M.; BREAKEFIELD, X.O.; ROPERS, H.H.; AND VAN OOST, B.A. Abnormal behavior associated with a point mutation in the structural gene for monoamine oxidase A. *Science* 261:321-327, 1993.

COLLINS, F.S. Let's not call it reverse genetics. *Nature Genetics* 1:3-6, 1992.

COLLINS, F.S., AND GALAS, D. A new five-year plan for the U.S. human genome project. *Science* 262:43-46, 1993.

Cooperative Human Linkage Center (CHLC): MURRAY, J.C.; BUETOW, K.H.; WEBER, J.L.; LUDWIGSEN, S.; SCHERPBIER-HEDEMA, T.; MANION, F.; QUILLEN, J.; SHEFFIELD, V.C.; SUNDEN, S.; DUYK, G.M.; ET AL. A comprehensive human linkage map with centimorgan density. *Science* 265:2049-2054, 1994.

DAVIES, J.L.; KAWAGUCHI, Y.; BENNETT, S.T.; COPEMAN, J.B.; CORDELL, H.J.; PRITCHARD, L.E.; REED, P.W.; GOUGH, S.C.; JENKINS, S.C.; PALMER, S.M.; ET AL. A genome-wide search for human type 1 diabetes susceptibility genes. *Nature* 371:130-136, 1994.

ERNST AND YOUNG. *Biotech 94: Long-term Value, Short-term Hurdles. The Industry Annual Report*. 1993. p. 31.

HAMER, D.H.; HU, S.; MAGNUSON, V.L.; HU, N.; AND PATTATUCCI, A.M.L. A linkage between DNA markers on the X chromosome and male sexual orientation. *Science* 262:578-580, 1993.

National Advisory Council for Human Genome Research. Statement on use of DNA testing for presymptomatic identification of cancer risk. *Journal of the American Medical Association* 271(10):785, 1994.

Pharmaceutical Research and Manufacturers of America. *Biotechnology Medicines in Development*, March 1995.

U.S. Department of Health and Human Services and U.S. Department of Energy. *Understanding Our Genetic Inheritance. The U.S. Human Genome Project: The First Five Years*. NIH Publication No. 90-1590. Bethesda, MD: National Institutes of Health, 1990.

⁷Of the 23 chromosome pairs in human cells, 22 pairs are numbered according to their size, with chromosome 1 being the largest and chromosome 22 being the smallest chromosome. The gender-determining chromosomes are referred to as X and Y.

ADOPTION STUDIES

Remi J. Cadoret, M.D.

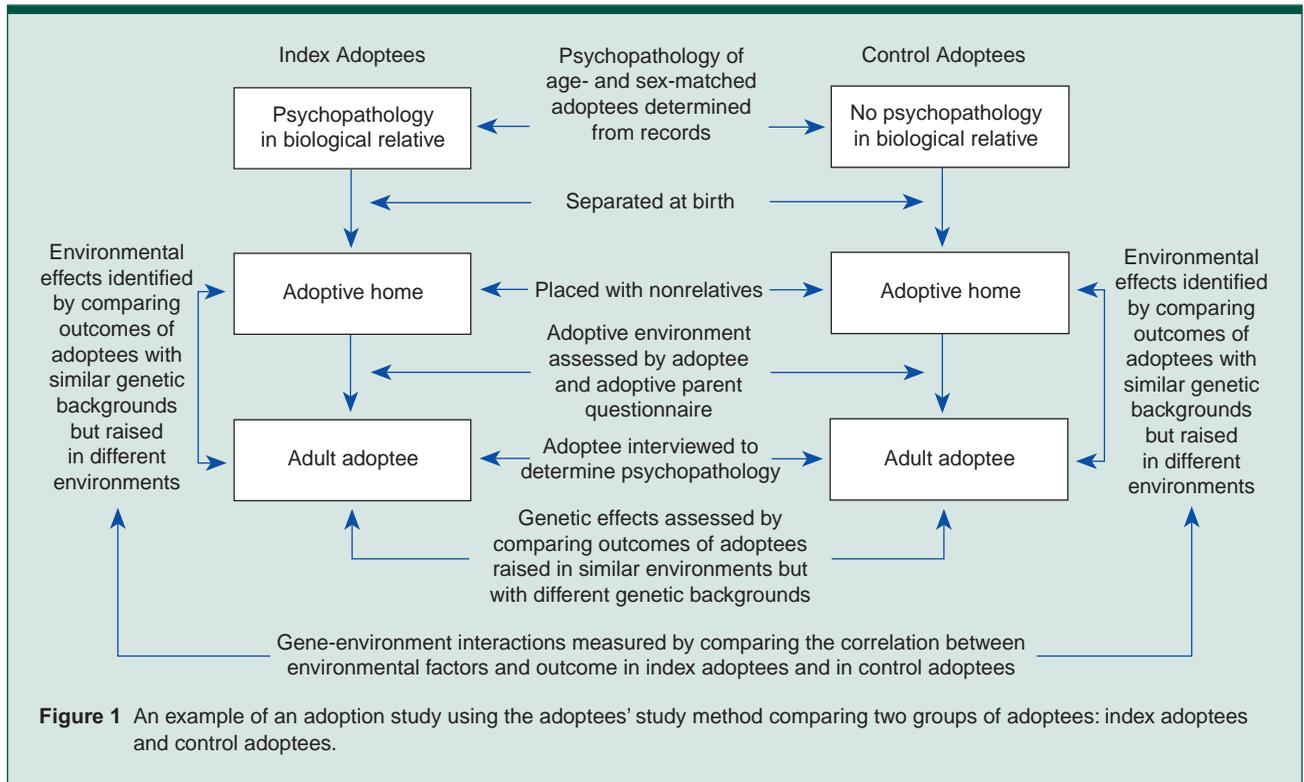
Researchers use adoption studies to determine the contributions of genetic and environmental factors to the development of alcohol problems. These studies generally compare the outcomes of adoptees who have biological parents with alcohol problems and who grow up in various adoptive environments with the outcomes of adoptees without such family backgrounds but raised in similar environments. Using certain statistical approaches, adoption studies also allow for the evaluation of specific gene-environment interactions in determining an outcome such as alcoholism. To obtain data that allow meaningful and generalizable conclusions, however, scientists must select a representative group of study subjects, obtain valid information about these subjects from a wide variety of sources, and consider biases inherent in adoption practices. KEY WORDS: adoption study; AODR (alcohol and other drug related) problems; hereditary factors; environmental factors; research and evaluation method; behavioral problem; gene

Adoption studies are a powerful tool for evaluating the interactions of genetic and environmental factors in eliciting human characteristics, such as intelligence (i.e., IQ), and disorders, such as alcoholism. The relative importance of "nature" (i.e., genetic inheritance) versus "nurture" (i.e., the rearing environment) in human behavior was first debated at the beginning of this century. Simultaneously, some techniques were developed that are still used to study the inheritance of behaviors, including the family study; the twin study (see the article by Prescott and Kendler, pp. 200-205); and statistical methods, such as regression analysis. One pioneer of human genetics, Sir Francis Galton, used these techniques in his studies. Galton concluded from his investigations that "nature prevails enormously over nurture" (Pearson 1914-30). In 1912, one year after Galton's death, another researcher, L.F. Richardson, proposed to study children who had been separated from their birth parents in order to investigate the inheritance and development of intelligence (Richardson 1912-13).

Concurrent social changes led to greater public acceptance of adoption and also improved researchers' access to adoptees. For example, founding societies and orphanages promoted adopting orphans or children born out of wedlock into foster families who were mostly nonrelatives. Adoptive parents usually received little information about the adoptees' biological parents. The lack of information may have been attributable to the belief at that time in the environment's overwhelming importance on a child's development. In addition, having a child out of wedlock was considered shameful, and consequently, confidentiality protected the birth mother. These "closed" adoptions were advantageous for conducting adoption studies because they clearly separated the biological and environmental influences on the adoptee.

In contrast, during the past two decades, a movement has occurred toward more "open" adoptions, in which biological and adoptive parents receive information about each other. Furthermore, this type of adoption may encourage continuing contact of the birth parents with both the adoptee and the adoptive family. In addition, social changes have drastically reduced the number of infant adoptees. For example, most unwed mothers now keep their children rather than give them up for adoption. These developments

REMI J. CADORET, M.D., is a professor of psychiatry at the University of Iowa, College of Medicine, Iowa City, Iowa.



have increased the practical problems involved in finding and recruiting suitable adoptees for studies.

Between the 1930's and 1950's, most adoption studies examined the heritability and effects of environmental influences on IQ. For example, during the 1930's, Skodak and Skeels (1949) demonstrated increases in IQ in certain environments using an adoption paradigm.¹ Since the 1960's, however, adoption studies have been used primarily to demonstrate the importance of genetic factors in psychopathological disorders, such as schizophrenia, alcoholism, or depression (for review, see Cadoret 1986). This article briefly examines some of the principles of adoption studies and the considerations required for their effective evaluation.

INFLUENCES ON ADOPTEES' BEHAVIOR

The strength of the adoption design—separating genetic from environmental influences on a person's development—results from removing the child (ideally at birth) from the birth parents and their environment into a different environment with biologically unrelated adoptive parents. Thus, adoption studies assess “real-world” influences on the adoptee's development while allowing for the separation of genetic and environmental factors that are confounded when children are reared to adulthood by their birth parents.

The adoptee's development and behavioral outcome result from multiple influences exerted by the birth parents and their environment and by the adoptive parents and their environment (for more information on these influences, see sidebar, p. 199). Determining the contributions of these different influences is a multivariate statis-

tical problem. Several statistical techniques, such as multiple regression analysis and log-linear analysis, can address such problems and have been used in evaluating adoption studies. Bohman, Cloninger, and their research group pioneered the use of multivariate approaches for studying the genetics of alcoholism in their analyses of Swedish adoption data (Bohman et al. 1982; Cloninger et al. 1982; Sigvardsson et al. 1982). Using these methods, the investigators assessed the contributions of both genetic and environmental factors on the development of alcoholism in the adoptees.

Selective Placement and Other Confounding Factors

To allow valid conclusions about the relative influences of genes and environment on adoptee outcome, it is essential that factors originating from the birth parents and their environment are unrelated to, and do not interact with, factors originating from the adoptive environment. This condition could be fulfilled by randomly placing infants in adoptive homes. However, adoption usually is not a random process. Adoption agencies carefully screen adoptive parents, and practical placement decisions frequently result in the selection of older, more stable families; families in higher socioeconomic brackets; and intact, rather than single-parent, families. Conversely, families that give up children for adoption commonly are single-parent, low-income ones.

In addition, adoptees may be matched to prospective adoptive parents depending on a variety of factors. For example, at one time adoptees often were matched with adoptive parents based on physical characteristics, such as hair and eye color. Other, more subtle matchings could depend on psychosocial characteristics. For example, an adoption agency might estimate a child's “potential” from birth-parent characteristics (e.g., education or socioeconomic level) and place the child according to some expectation of future performance.

¹For a definition of this and other technical terms used in this article, see central glossary, pp. 182–183.

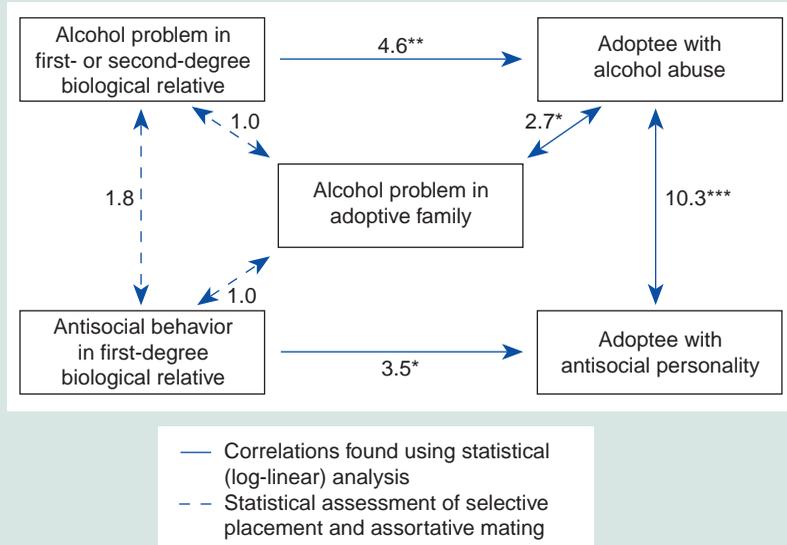


Figure 2 Results of an adoptees' study method adoption paradigm based on 160 male adoptees and their biological and adoptive families assessed for alcoholism, antisocial personality disorder, and other psychological parameters. The numbers next to the arrows are odds ratios.¹ (For example, an adoptee with first- or second-degree biological relatives with alcohol problems is 4.6 times more likely to abuse alcohol than an adoptee without such a family background.)

* $p < 0.05$
 ** $p < 0.01$
 *** $p < 0.001$
¹An odds ratio is a measure of association between two variables.
 SOURCE: Adapted from Cadoret et al. 1987.

the pathology of the high-risk group and are considered “normal.” In the adoptees’ study design, researchers usually compare the outcome of adoptees with contrasting biological backgrounds (e.g., alcoholic versus nonalcoholic birth parents); further control can be obtained by matching the proband and control birth parents on variables such as socioeconomic level or age. In the adoptees’ family design, the study compares the biological backgrounds of proband adoptees with those of control adoptees, who usually have been selected for normality. In addition, the adoptees may be matched on variables such as age, gender, and socioeconomic level.

A typical adoptees’ study design compares so-called index adoptees—adult adoptees who have backgrounds of psychopathology (e.g., alcoholism) in their biological families—with age- and sex-matched control adoptees who have no family histories of psychopathology. (For a more detailed description of the design of an adoptees’ study paradigm, see figure 1.) An adoption study by Cadoret and colleagues (1987) illustrates how the contributions of several genetic and environmental factors to the development of alcoholism can be determined using this method (figure 2). In the study, 160 male adoptees, their biological relatives, and their adoptive families were analyzed regarding alcohol problems, antisocial behavior, and other psychological variables. The study found that a genetic influence, such as alcohol

problems in first-degree (i.e., parents) or second-degree (i.e., grandparents) biological relatives, increased an adoptee’s risk for alcohol problems 4.6-fold. Similarly, an environmental influence, such as alcohol problems in a member of the adoptive family, resulted in a 2.7-fold higher risk for alcohol problems in the adoptee, compared with adoptive families without alcohol problems.

DESIGN AND EVALUATION OF ADOPTION STUDIES

Because the adoption agencies often were aware of both alcoholism and antisocial behavior in the biological parents, these factors could have influenced placement decisions and correlated with the environmental factor of adoptive family alcohol problems. To control for such potential selective placement effects, the correlations between alcohol problems or antisocial behavior in the biological family and alcohol problems in the adoptive family also were assessed in the statistical analysis (figure 2). The study found no evidence of selective placement based on the factors shown: As indicated by the odds ratios² of 1.0, the likelihood of a member of the adoptive family having alcohol problems was the same whether or not biological relatives of the adoptee displayed alcohol problems or antisocial behavior.

Adoption studies generally can be classified based on whether the adoptees or the birth parents are the probands (i.e., the initial subjects) of the study (Rosenthal 1970). In the adoptees’ study method, researchers identify proband birth parents with a certain characteristic (e.g., alcoholism) and then examine the outcome of these probands’ adopted-away children. A contrasting design is the adoptees’ family method, in which researchers identify proband adoptees with a certain characteristic (e.g., alcoholism or depression) and subsequently examine both the birth and adoptive parents. Both designs have been used to demonstrate the importance of genetic factors in the development of alcoholism. Whether the adoptees’ study method or the adoptees’ family method is used often depends on certain considerations, such as practicality and the ease of recruiting probands and gathering information about them.

Most adoption studies have used a design comparing high-risk probands (i.e., adoptees or birth parents) having certain characteristics (e.g., alcoholism) with a control group of subjects who lack

problems in first-degree (i.e., parents) or second-degree (i.e., grandparents) biological relatives, increased an adoptee’s risk for alcohol problems 4.6-fold. Similarly, an environmental influence, such as alcohol problems in a member of the adoptive family, resulted in a 2.7-fold higher risk for alcohol problems in the adoptee, compared with adoptive families without alcohol problems.

Because the adoption agencies often were aware of both alcoholism and antisocial behavior in the biological parents, these factors could have influenced placement decisions and correlated with the environmental factor of adoptive family alcohol problems. To control for such potential selective placement effects, the correlations between alcohol problems or antisocial behavior in the biological family and alcohol problems in the adoptive family also were assessed in the statistical analysis (figure 2). The study found no evidence of selective placement based on the factors shown: As indicated by the odds ratios² of 1.0, the likelihood of a member of the adoptive family having alcohol problems was the same whether or not biological relatives of the adoptee displayed alcohol problems or antisocial behavior.

Assortative Mating

Another factor that can affect a child’s development and behavior is assortative mating (i.e., the nonrandom choice of a partner

²An odds ratio is a measure of association between two variables.

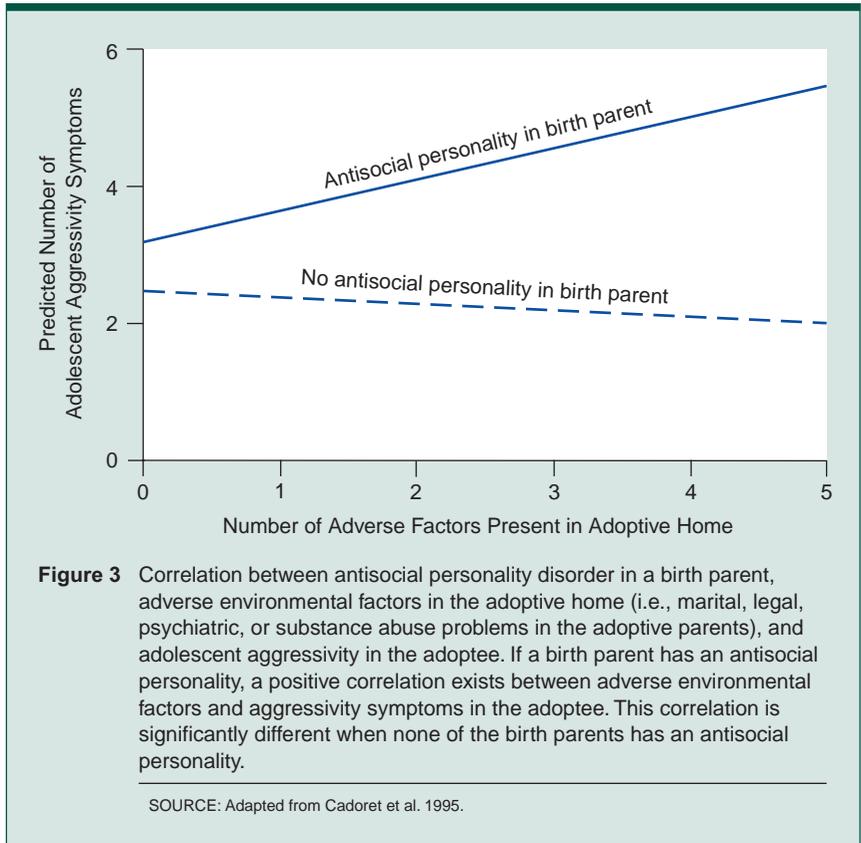
based on personal characteristics). For example, an alcoholic person may be more likely than a nonalcoholic person to have an antisocial or alcoholic partner, possibly because of shared traits or behaviors. The combination of two genetic predispositions may enhance the predisposition of the offspring to develop any psychopathology. Multivariate statistical analyses can help control for the effects of assortative mating if relevant information is available on both birth parents. Similar analyses also can be used to control for the genetic predisposition for two disorders (e.g., alcoholism and antisocial personality disorder) within one person.

Alternative Evaluation Methods

Simpler statistical analyses also have been used to evaluate the results of adoption studies. For example, when the assessment of genetic influences is the main objective, a common strategy is to demonstrate that the environmental influences are the same for adoptees from high-risk backgrounds (i.e., with alcoholic biological family members) and low-risk backgrounds (i.e., without alcoholic biological family members). Comparable environmental factors for both groups would indicate that no selective placement occurred that could confound the study results. Using this method, Goodwin and colleagues (1973) demonstrated the importance of a genetic predisposition to the development of alcoholism. However, although environmental influences may be similar when averaged over high- or low-risk adoptee groups, considerable environmental variability still exists among the members of each adoptee group that could affect the outcome of individual adoptees and which should be assessed by multivariate statistical approaches.

Gene-Environment Interactions

In determining the contributions of genetic factors to an outcome such as alcoholism, it is important to know whether a genetic factor exerts its effect only in the presence of a specific environmental condition or does so independently of environment. The adoption paradigm is a powerful tool for evaluating the interaction of specific genetic factors with specific environmental factors that affect adoptee outcome (DeFries and Plomin 1978). For example, researchers and clinicians have long recognized that both conduct disorder and aggressivity predispose an affected person to alcohol and other drug abuse (see figure 2). Adoption studies also have demonstrated that antisocial personality disorder in birth parents predisposes adopted-away offspring to both conduct disorder (Cadoret and Cain 1981; Cadoret 1986) and aggressivity (Cadoret et al.1995). In the latter study, however, the genetic predisposition inherited from a birth parent with antisocial personality disorder increased conduct disorder and aggressivity only in adoptees raised in an environment with additional adverse factors (e.g., an adoptive parent suffering from a psychiatric or marital problem) (figure 3) (Cadoret et al. 1995).



Findings from the study of this type of gene-environment interaction may suggest points of intervention, thereby helping to prevent behavior leading to alcoholism. For instance, in the above example, modifications of the environment (e.g., treatment of the adoptive parents' problems) could affect the adoptee's outcome even in the presence of a genetic predisposition.

FACTORS INFLUENCING STUDY QUALITY

Obtaining Valid Information

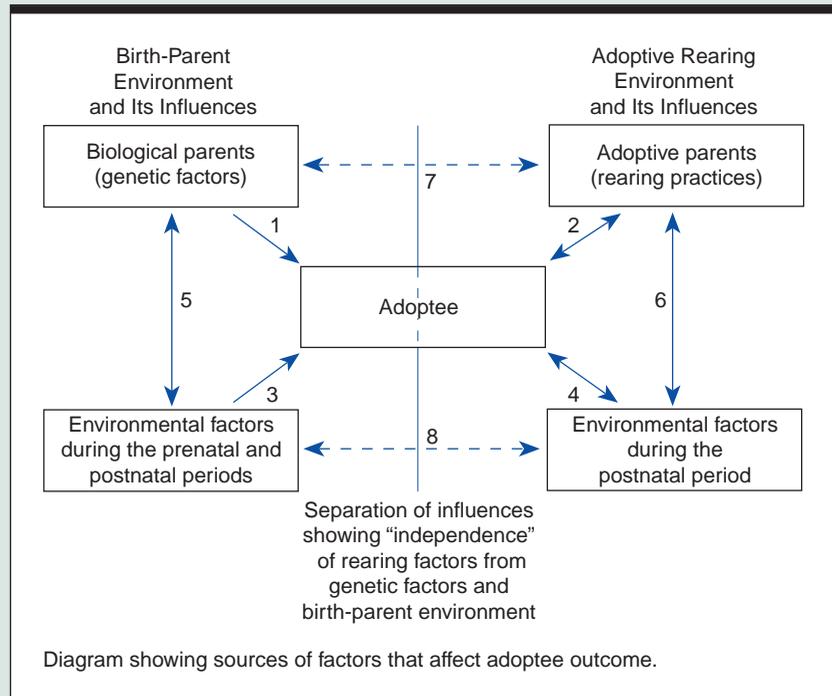
Valid information about the birth parents, the adoptive parents, and the rearing environment is crucial when using adoption studies to assess the influences of genetic and environmental factors on behavior. This information must address the four important sources of influences on the adoptee: the genetic and environmental factors from the birth parents, the parental influences from the adoptive parents, and the adoptive family environment. Thus, a major technical difficulty in adoption studies is arranging for data collection from a wide range of sources, some of which are protected by confidentiality.

Information about the birth parents and their behaviors is necessary to determine which adoptee characteristics may represent phenotypes of a genetic predisposition inherited from the parents (e.g., genes predisposing the adoptee to develop alcoholism). This information can be obtained from the records of the adoption agency, hospitals, social services, and similar sources. In studies of adoptees born out of wedlock, reliable information about birth fathers frequently is lacking. However, recent laws requiring writ-

SOURCES OF INFLUENCES AFFECTING ADOPTEE OUTCOME

A multitude of influences on the adoptee play a role in determining the adoptee's development and behavioral outcome. The left side of the diagram (the vertical line represents the separation of biological- and adoptive-family factors) indicates the influences affecting the adoptee during pregnancy, delivery, and the immediate neonatal period, including genetic predispositions inherited from the birth parents (arrow 1) and prenatal and neonatal environmental influences (e.g., maternal alcohol consumption during pregnancy; arrow 3). These genetic and environmental factors also interact with each other, as represented by arrow 5 (e.g., genetically determined antisocial personality disorder or depression in the mother may contribute to her alcohol consumption).

The factors on the right side of the diagram represent the postnatal influences on the adoptee (which, in turn, are influenced by the child) following placement with nonrelatives. Adoptive-parent characteristics are the most important influences affecting the adoptee (arrow 2). The two-headed arrow indicates that the child-parent relationship is an interaction of many factors (e.g., child temperament and parenting skills of adoptive parents). Arrow 4 indicates the correlation between the adoptee and environmental influences. Factors such as friends outside the family influence the adoptee, but the adoptee often simultaneously



exerts an influence by seeking out those friends in the first place. Finally, adoptive-parent characteristics and environmental factors also interact with each other (arrow 6). Parent characteristics influence factors such as socioeconomic status. Environmental factors, in turn, can influence parents (e.g., financial stressors may affect parenting behavior by causing depression and irritability).

In addition, adoptees may be matched to a certain extent to prospective parents

based on a variety of factors that can lead to correlations between the biological and the adoptive environments (broken arrows). For example, the educational levels of birth parents and adoptive parents could be used as the basis for matching (broken arrow 7). Similarly, a correlation could exist among environmental factors (e.g., both birth parents and adoptive parents live in rural areas; arrow 8).

— Remi J. Cadoret

ten permission from biological fathers to release children for adoption may improve information collection. For example, if a birth father's name is available, archival information from hospitalizations, incarcerations, or other records (e.g., death certificates) can be obtained provided that the confidentiality required for such records can be maintained.

Adoption agencies usually can provide information about pregnancy and delivery (i.e., influences of the birth-parent environment). Similarly, agency records can supply a large amount of personal information about the adoptive parents and the rearing environment. This information is especially of interest because adoption studies can measure the influences of specific environmental effects as effectively as the influences of genetic

effects. Information about the adoptees themselves also is readily available in most cases.

Ideally, adoption studies would include information obtained by personal interviews with all the people who primarily affect the adoptee's outcome (i.e., the birth parents, the adoptive parents, the adoptee, and friends of the adoptee). Data collected solely from institutional records, however, such as those from the central registries in Scandinavian countries, also can provide valuable information and, at the very least, be used to identify subjects for direct study. Long-term followup of the adoptees, their birth parents, and their adoptive families would result in even more valid information about behaviors that tend to change over time, such as conduct disorders, alcohol abuse, or depression. Such longitudinal

studies could considerably increase the identification of psychopathological behaviors that might go undetected in a study relying only on information gathered during one time period.

Proband Recruitment

How the probands are recruited also can affect the quality of a study's conclusions. One potential source of bias is the influence of environmental factors on the selection of proband adoptees in the adoptees' family method. For example, psychological or social problems in an adoptive family may contribute to the adoptee's psychopathology. Simultaneously, these problems may prompt the family and the adoptee to seek more treatment and thus increase their chances of being included in a sample of adoptees recruited from a clinic population. Factors such as these may compromise the representativeness of the sample.

Similarly, refusal rates among potential study participants could influence the quality of the data obtained. For example, it is possible that adoptees and their families who refuse to participate in a study as a group are distinguished by certain qualities (e.g., personality characteristics). Consequently, their refusal could reduce the representativeness of the study sample.

GENERALIZABILITY OF ADOPTION STUDIES

Whether the findings from adoption studies can be used to draw general conclusions about the contribution of both genetic and environmental factors to the development of alcoholism depends largely on how representative the adoptee sample is. Representativeness, in turn, is determined by variables, such as the criteria for proband selection. Although many of these variables can be controlled for or at least recognized, the inherent biases in adoption practices (e.g., selective placement and predominant recruitment of adoptive families from certain population groups) limit generalizability.

SUMMARY

Despite the existing limitations and the technical problems associated with conducting adoption studies, the adoption paradigm provides important information about the significance of specific genetic and environmental factors in human behavior. In addition, adoption studies allow researchers to identify specific genetic-environmental interactions that could be relevant for designing early interventions for behaviors that predispose a person to alcohol abuse and dependence. ■

REFERENCES

BOHMAN, M.; CLONINGER, C.R.; SIGVARDSSON, S.; AND VON KNORRING, A.-L. Predisposition to petty criminality in Swedish adoptees: I. Genetic and environmental heterogeneity. *Archives of General Psychiatry* 39:1223-1241, 1982.

CADORET, R.J. Adoption studies: Historical and methodological critique. *Psychiatric Developments* 1:45-64, 1986.

CADORET, R.J., AND CAIN, C. Environmental and genetic factors in predicting adolescent antisocial behavior in adoptees. *Psychiatric Journal of the University of Ottawa* 6:220-225, 1981.

CADORET, R.J.; TROUGHTON, E.; AND O'GORMAN, T.W. Genetic and environmental factors in alcohol abuse and antisocial personality. *Journal of Studies on Alcohol* 48:1-8, 1987.

CADORET, R.J.; TROUGHTON, E.; YATES, W.; WOODWORTH, G.; AND STEWART, M.A. Genetic-environmental interaction in the genesis of conduct disorder and aggressivity. *Archives of General Psychiatry* 52:916-924, 1995.

CLONINGER, C.R.; SIGVARDSSON, S.; BOHMAN, M.; AND VON KNORRING, A.-L. Predisposition to petty criminality in Swedish adoptees: II. Cross-fostering analysis of gene-environment interaction. *Archives of General Psychiatry* 39:1242-1247, 1982.

DEFRIES, J.C., AND PLOMIN, R. Behavioral genetics. *Annual Review of Psychology* 29:473-515, 1978.

GOODWIN, D.W.; SCHULSINGER, F.; AND HERMANSEN, L. Alcohol problems in adoptees raised apart from alcoholic biologic parents. *Archives of General Psychiatry* 28:238-43, 1973.

LEWONTIN, R.C.; ROSE, S.; AND KAMIN, L.J. *Not in Our Genes: Biology, Ideology, and Human Nature*. New York: Pantheon Books, 1984.

PEARSON, K. *The Life, Letters, and Labours of Francis Galton, Vols. I, II, IIIA and IIIB*. Cambridge, United Kingdom: Cambridge University Press, 1914-30.

RICHARDSON, L.F. The measurement of mental "nature" and the study of adopted children. *Eugenics Review* 4:391-394, 1912-13.

ROSENTHAL, D. *Genetic Theory and Abnormal Behavior*. New York: McGraw-Hill, 1970.

SKODAK, M., AND SKEELS, H.M. A final follow-up study of one hundred adopted children. *Journal of Genetic Psychology* 75:85-125, 1949.

SIGVARDSSON, S.; CLONINGER, C.R.; BOHMAN, M.; AND VON KNORRING, A.-L. Predisposition to petty criminality in Swedish adoptees: III. Sex differences and validation of the male typology. *Archives of General Psychiatry* 39:1248-1253, 1982.

TWIN STUDY DESIGN

*Carol A. Prescott, Ph.D.,
and Kenneth S. Kendler, M.D.*

By studying human pairs of twins, researchers can learn the relative contributions of genetic and environmental factors to the development of alcoholism. Identical (i.e., monozygotic, or MZ) twins share 100 percent of their genes, whereas fraternal (i.e., dizygotic, or DZ) twins generally share only 50 percent of their genes. Using certain techniques and theoretical models, researchers can compare the two types of twin pairs for how often alcoholism occurs in both members of a twin pair. If alcoholism occurs more often in both members of MZ twins, genetic factors are implicated in the origin of the disorder. Twin research also has been applied to studies of differences between men and women in their genetic contribution to alcoholism.
KEY WORDS: twin study; AOD dependence; hereditary factors; environmental factors; applied research; gender differences

Humans are biologically similar, sharing almost all of their genetic material. Many questions relevant to alcoholism,¹ however, concern how people differ. Why do some people abuse alcohol? Why do some heavy drinkers, but not others, become physiologically addicted to alcohol? Twins are a unique resource for identifying the genetic and environmental

¹ The term "alcoholism" is used in this article to encompass all levels of problem alcohol use and does not refer to a particular diagnostic system. The term "alcohol dependence" is used when referring to research that used this diagnosis as defined by the *Diagnostic and Statistical Manual of Mental Disorders, Third Edition, Revised* of the American Psychiatric Association (1987).

sources of these differences. The study of the causes of differences (i.e., variation) among people is a major goal of human behavioral genetic research, of which twin studies form a subset.

Twin studies have been used to address important questions about the causes of alcoholism, including the following:

- How important to the development and course of alcoholism are genetic and environmental influences?
- Do genetic influences differ in importance for different groups of people (e.g., males and females or different age groups)?

Based on partial answers from these investigations, new twin studies are seeking answers to even more specific questions, such as the following:

- Do the same sets of genes (or the same environments) produce the same effects in males and females?
- What environmental characteristics are most hazardous for people at high genetic risk for developing alcoholism?
- What environmental characteristics *protect* against alcoholism among people who have a strong family history of the disorder?
- To what degree is the frequently seen overlap between alcoholism and other disorders (such as depression and anxiety) a result of the same sets of genes influencing both disorders?

These questions have important implications for understanding the causes of alcoholism and ultimately providing knowledge to help design treatment and prevention strategies for people at high risk of developing the disorder. This article describes how twin studies are designed, offers examples of research questions for which twin studies are useful, and reviews their limitations.

THE TWIN STUDY DESIGN

The Value of Twins

Because of their unique genetic status, twins play a valuable role in teasing apart genetic and environmental influences on people's development of alcoholism. Genes are segments of DNA that form the blueprints for the development of the human body. Genetic expression that influences behavior is only partly preprogrammed; the biological processes regulated by genes are modified in com-

CAROL A. PRESCOTT, PH.D., is co-director of the Stress & Coping Twin Project and assistant professor of the Psychiatric Genetics Research Program, Department of Psychiatry, Medical College of Virginia/Virginia Commonwealth University, Richmond, Virginia.

KENNETH S. KENDLER, M.D., is director of the Psychiatric Genetics Research Program and is Rachel Brown Banks Distinguished Professor in the departments of psychiatry and human genetics, Medical College of Virginia/Virginia Commonwealth University, Richmond, Virginia.

This article was prepared with support from National Institute on Mental Health grants M-40828 and MH-49492 and National Institute on Alcohol Abuse and Alcoholism grant AA-09095.

plex ways by experiences with the environment. Thus, neither genes nor experiences operate independently of one another. In this sense, the often-cited dichotomy of "genes versus environment" is false; both are required for the development and expression of human characteristics. Knowledge of how genetic and environmental factors act both separately and together to influence the development of alcoholism can inform strategies (e.g., biological, individual, or societal) for intervention and prevention.

When researchers study genetic influences on behavioral variation, they focus on the small percentage of genes that differ among people. Fraternal, or dizygotic (DZ), twin pairs, like ordinary siblings, have, on average, 50 percent of these genes in common, whereas identical, or monozygotic (MZ), twin pairs have 100 percent of these genes in common. By comparing MZ with DZ twin pairs, researchers can use this difference in the twins' degrees of genetic resemblance to estimate the relative importance of genetic and environmental influences on the development and course of disorders such as alcoholism.

Liability Models

When searching for genetic and environmental influences on alcoholism, researchers can only observe a characteristic in a subject, such as the presence or absence of the clinical diagnosis of alcoholism. Researchers, however, are actually trying to estimate the contribution of genetic and environmental effects to the subject's *liability* for alcoholism (the theoretical components contributing to the risk for alcoholism are portrayed in figure 1). Liability may be thought of as the outcome of genetic and environmental risk factors that together produce a person's total risk for developing alcoholism (Falconer 1965). Liability is an unobserved, or latent, characteristic that varies in the population.

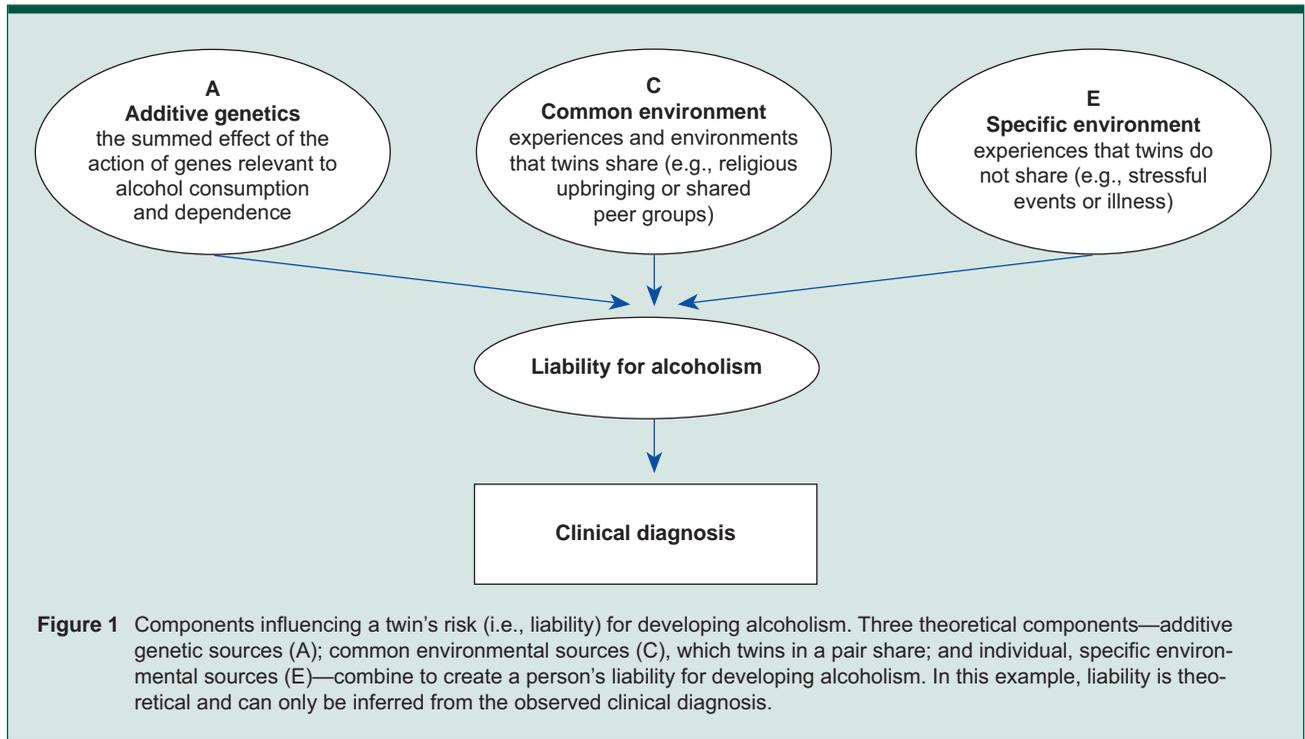
The three components of liability usually identified in twin research are as follows:

- Additive genetic liability, or the summed effect of many genes relevant to alcohol consumption and dependence
- Common environmental effects that twins share (e.g., parental influence, intrauterine environment, and sociocultural factors)
- Environmental effects that twins do not share (e.g., experiences such as marital problems or job loss). (For further discussion, see sidebar, pp. 204-205.)

Researchers assume that the three components listed above combine to produce total liability. Only people who exceed a specific level, or threshold, of liability will manifest the clinical aspects of the disorder. Thus, people with low genetic liability require higher levels of environmental risk before they are likely to develop alcoholism, whereas those with high genetic liability may develop alcoholism in response to lower levels of stressful events.

ESTIMATES FROM TWIN STUDIES

At conception, MZ twins are genetically identical. Although mutations in some genes are likely to occur, and adult MZ twins may differ in which genes currently are active (because of unshared environmental influences, such as disease or nutrition), these differences are believed to be small. MZ twins are assumed in studies to be perfectly correlated for genes relevant to the development of



alcoholism. Because first-degree relatives, including DZ twins, have, on average, one-half their genes in common, DZ twins have only one-half the degree of genetic correlation seen between MZ twins. Thus, if members of MZ twin pairs are much more alike for liability for alcoholism than are DZ twins, genetic influences can be implicated as a major contributor to liability. Conversely, if MZ and DZ twin pairs are equally similar for alcoholism liability, then any similarity is attributed to shared environmental factors. Researchers use equations based on the MZ–DZ comparison to estimate the proportions of liability resulting from the genetic and environmental components (for further discussion, see sidebar, pp. 204–205).

APPLYING THE TWIN DESIGN

The twin design frequently has been applied to studying differences in the relative importance of genetic and environmental influences among different groups, such as males and females or different age sets (i.e., cohorts). For example, the estimates of genetic influence from studies of male twin pairs could be compared with the estimates resulting from female twin pair studies to identify differences between the groups. Other variations on traditional twin study designs could address important questions about differences in the mechanisms of genetic influence (discussed below).

Sex Differences

Although the risk for alcoholism may vary somewhat in different demographic groups (e.g., by age, ethnicity, or socioeconomic level), the difference between sexes in prevalence of alcoholism is particularly dramatic. Because the genetic differences between males and females are larger than between any other groups, comparing the sexes can provide more specific information about genetic mechanisms of risk for alcoholism.

Sources of Liability. Studying male and female same-sex pairs allows the calculation of separate estimates of genetic and environmental effects for each sex. Although the data can show if the sexes differ in the relative importance of genetic and environmental sources of liability for alcoholism, the data cannot determine whether these genetic and environmental influences are the same in both sexes. Researchers attempt to solve this question by studying opposite-sex twin pairs (which are never MZ pairs).

By using opposite-sex twin pairs, sex differences in genetic and common environmental sources of liability sometimes can be estimated. For example, if DZ opposite-sex twin pairs are less similar in their alcoholism liability than DZ same-sex pairs, this finding would be evidence that the two sexes differ either in the genes or environments (or both) relevant to the development of alcoholism. For statistical reasons, the twin design cannot always reveal the degree to which these differences result from genetic or environmental distinctions between the sexes. Evidence that such differences exist, however, provides an important starting point for researchers studying sex differences in alcoholism. Thus far, the available opposite-sex twin data are limited (for review, see McGue 1994), but two large twin studies that include male–female pairs are underway in Australia (A. Heath and colleagues, ongoing research) and in Virginia (K. Kendler and colleagues, ongoing research).

Transmission of Liability. Other applications of the twin design address the following questions about sex differences in the mechanisms responsible for transmitting liability for alcoholism between generations:

- Do mothers and fathers differ in the amount of genetic liability they pass to their children?
- Once genetic influences are accounted for, do parents contribute any further environmental risk by their behavior?

- Do mothers and fathers differ in the magnitude or mechanism of environmental liability they pass to their children?
- Does an alcoholic parent contribute different levels of risk to sons versus daughters?
- Does the importance of parent-offspring transmission differ in different age cohorts?

These questions require the use of twin-family designs, in which both twins and their parents are studied. Results from one such study of female twins and their parents indicated that mothers and fathers were equally likely to transmit liability for alcoholism to their daughters (Kendler et al. 1994).

HETEROGENEITY OF EXPRESSION

Another major application of the twin model is in the study of the different ways the same genetic liability can be expressed, often called *phenotypic heterogeneity*. For example, the same genetic liability may produce different clinical presentations (or *phenotypes*) in the sexes: A woman may become depressed, but a man may become alcohol dependent. In another example, genetic liability may be expressed as alcoholism in one generation and drug dependence in another. Studying opposite-sex twin pairs or twins from different age cohorts can address these important issues regarding the origins of alcoholism and its overlap with other disorders.

Another type of difference in expression of liability is called *etiologic heterogeneity*—when the same clinical disorder (e.g., alcoholism) occurs through multiple pathways. For example, women with alcoholism are at higher risk for anxiety and depression than are alcoholic men, whereas alcoholic men are more likely than women to have antisocial characteristics (Helzer and Pryzbeck 1988). These data, in combination with results from family studies, have led researchers to propose that alcoholism in men is primary (i.e., it occurs first) and is often associated with sociopathy, whereas alcoholism in women is more commonly secondary to other emotional disorders (see, for example, Cloninger and Reich 1983; McGue et al. 1992).

By examining the types of disorders present in the co-twins of alcoholics (e.g., antisocial personality versus depression), researchers also can better understand the clusters of symptoms and subtypes that constitute alcoholism. Furthermore, these studies can show whether the basis for the overlap is genetic, environmental, or both. For example, Kendler and colleagues (1994) found that the co-occurrence of alcohol dependence and depression among female twins was attributable primarily to shared genetic, not shared environmental, risk factors.

Genetic liability for alcoholism also may interact and combine with genetically influenced characteristics other than sex (e.g., personality type) or with environmental influences (e.g., choice of peer group or religious or cultural prohibitions), resulting in varying outcomes for different people. More sophisticated behavioral genetic models are being developed that can address how these factors may mediate the development of alcoholism.

LIMITATIONS OF TWIN STUDIES

As with any research design, results from twin studies are generalizable only to the extent that twins are representative of the entire population. Evidence indicates that the prevalence of psychiatric

symptoms among twins does not differ from the prevalence in the general population (Kendler et al. in press). People willing to participate in research studies, however, may differ from random samples (e.g., by the severity of their disorder or by their socioeconomic level), leading to potential biases in the study population.

Some studies obtain alcoholic twin subjects from treatment settings, whereas other studies assess twins from the general population. Because treatment-based studies typically identify only a minority of the affected cases in the general population, researchers must determine whether these cases differ in severity or symptomatology from those not included in the research. A particularly important consideration for twin research is whether twins in studied pairs are more likely to be similar than pairs in the general population. For example, twin pairs in which both twins are alcoholic may be more likely to participate in treatment or attract the attention of researchers (and thus be included in studies) than pairs in which only one member is alcoholic, thereby leading to biases in estimates of twin similarity (see, for example, Prescott and Kendler 1995). Similarly, studies based on twins in the general population are subject to volunteerism biases. Females and identical twins are more likely to participate in research than are males or fraternal twins, and people with alcohol-related problems are less likely to be willing to participate in research; they also can be difficult to locate. The effects of such biases can be examined using statistics, but only if the biases are identified and measured.

Other limitations of the twin method arise from the assumptions required to estimate genetic and environmental effects. The standard twin design assumes the absence of *assortative mating* (i.e., the tendency for people to choose mates who are similar to themselves with respect to the characteristic being studied). Alcohol consumption, however, is an integral aspect of social activity for many people, and spouses tend to resemble each other regarding alcohol use and misuse (see, for example, Jacob and Bremer 1986). When assortative mating for alcoholism occurs, offspring who are DZ twins share more than 50 percent of alcoholism-relevant genes, leading to an underestimation in twin studies of the magnitude of genetic influence. It is possible, however, to correct the results from twin studies by using estimates from other studies of spousal similarity with respect to alcoholism.

Another crucial assumption that can limit the accuracy of twin studies is the *equal-environment assumption*. This is the assumption that MZ and DZ twin pairs are equally similar in their shared environments. To the extent that this is not true, the importance of a shared environment will be underestimated and genetic influence overestimated. For example, if peers are an important influence on drinking behavior, and MZ twins are more likely to share peers with their twin than are DZ twins, the greater MZ similarity will be attributable to environmental as well as genetic causes but will be ascribed solely to genetic influences.

Attempts to address the equal-environment assumption with respect to alcohol use and misuse have yielded mixed results (Kendler et al. 1992; Lykken et al. 1990; Rose et al. 1990), with some studies finding that twins who spend more time together (often MZ pairs) are more alike in their drinking habits. Even after statistically controlling for the effects of greater environmental similarity among MZ twin pairs, however, substantial genetic influences on alcohol intake remain (see, for example, Kendler et al. 1992).

A further limitation of the twin study design is that the separation of liability into distinct genetic and environmental components may be invalid if sizable gene-environment *correlations* or *interactions* exist. A gene-environment *correlation* occurs if persons with high genetic liability for alcoholism are more likely to experience

DETERMINING LIABILITY

To estimate the relative importance of genetic and environmental sources of a characteristic, such as a person's liability for developing alcoholism, researchers use mathematical equations in twin studies. In the most common model, liability for alcoholism is attributed to three unobserved, or latent, sources: additive genetic (A); common environmental (C); and individual, specific environmental (E). Each type is described below.

Additive Genetic Sources. Humans have 46 chromosomes,¹ a number that includes 22 pairs plus the sex chromosomes, XX or XY. The genes in a pair may be similar but not identical; a gene for a particular trait on one chromosome may vary slightly from the corresponding gene on the paired chromosome. Such gene variants are called *alleles*.

The effects of alleles on observed characteristics combine in three major ways. Nonadditive genetic mechanisms include *dominance* (a term that describes any interaction between paired alleles other than a simple summing of their effects) and *epistasis* (wherein alleles at different sites alter the effects of each other). However, the most common mechanism of allelic action for complex behavioral traits is believed to be through *additive genetic action*. In the case of alcoholism, this would mean that all

alleles relevant for the development of alcoholism combine and are not suppressed or magnified by alleles on the paired chromosome or any other location. The A component estimated in twin studies represents the total additive genetic action relevant to a person's risk for developing alcoholism. Most twin models assume that dominance and epistasis are negligible. For example, genes that predispose a person to have a higher risk for alcoholism (e.g., by increasing alcohol's euphoric effects) might combine with other genes that reduce risk (e.g., by producing a flushing response) to produce a moderate level of risk. Identical, or monozygotic (MZ), twins have 100 percent of their genes—including those that influence risk for alcoholism—in common, whereas fraternal, or dizygotic (DZ), twins share (on average) only 50 percent of the genes that vary in the population (see figure).

Common Environmental Sources. Common environmental influences on problem alcohol use might include parental drinking habits and teachings about alcohol, shared peer groups, and sociocultural influences. For twins (but not other siblings), these components also include the prenatal environment. In studies, researchers assume these factors correlate perfectly for twins or siblings reared together, shown by a correlation of 1.0 in the figure for the C components.

Individual, Specific Environmental Sources. These include any experiences

or environmental influences not shared by siblings, such as unshared peers and stressful life events (e.g., job loss, marital problems, or physical illness). The estimate of this term usually includes random measurement error (such as that from misdiagnosis), which cannot be separated from E in studies in which twins are measured only once. Because the members of a twin pair do not share specific environmental sources, no correlation exists between the E components in the figure shown.

Liability Models

Under the usual liability model, the three components described above are assumed to combine. The liability (L) for an individual (i) can be expressed as: $L_i = a(A_i) + c(C_i) + e(E_i)$. The uppercase letters A, C, and E represent theoretical "scores" on each of the three components; these scores vary across people. The lowercase letters a, c, and e designate the amount each component contributes to the outcome, and these are constant for everyone in a group.

The similarity, or correlation, in liability for two people is estimated as the sum of the components that they have in common. Because members of MZ twin pairs are identical for both their A and C components, their correlation is estimated as $a^2 + c^2$. Members of DZ twin pairs share one-half their genes and all their common environment; therefore, the correlation for them is $0.5(a^2) + c^2$. With just these

¹For a definition of this and other technical terms used in this sidebar, see central glossary, pp. 182–183.

alcoholism-promoting environments (this may occur among children of alcoholics, who may obtain both genetically and environmentally transmitted risk for alcoholism). Gene-environment *interactions* occur when persons with high genetic risk are particularly sensitive to the effects of alcoholism-promoting environments. Twin data can be applied to these issues if a range of environmental conditions is available for study. Data from studies of adopted children also are useful for testing for gene-environment interactions, but these studies have other methodological limitations. (For further discussion of adoption studies, see the article by Cadoret, pp. 195–200.)

SUMMARY

Twin studies can be a powerful tool for addressing important questions about the development and course of alcoholism, including the role of genetic and environmental factors, as well as sex differences in the mechanisms and the magnitude of genetic and environmental influences. The results from opposite-sex twin studies, along with

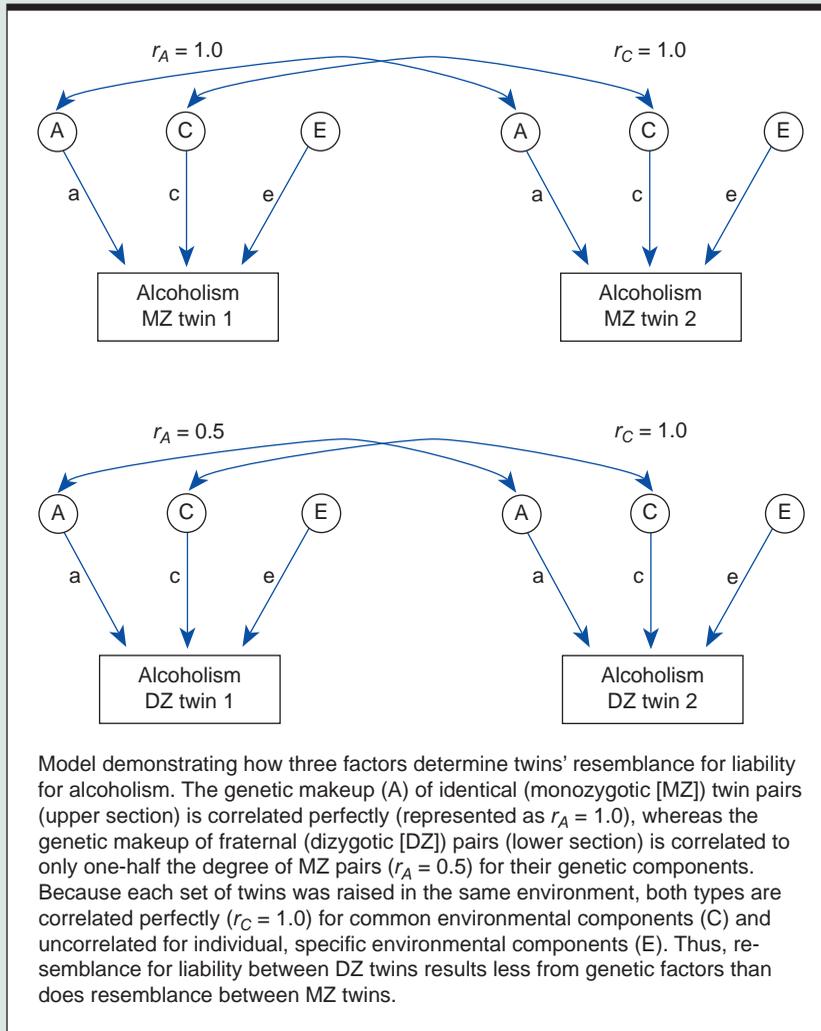
those from adoption and twin-family studies, promise to improve our understanding of the roles of genetic factors and of experiences and their interactions in the causes of alcoholism. ■

ACKNOWLEDGMENT

The authors thank J.J. McArdle and two anonymous reviewers for their comments on an earlier draft of this article.

REFERENCES

- American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders, Third Edition, Revised*. Washington, DC: the Association, 1987.
- CLONINGER, C.R., AND REICH, R. Genetic heterogeneity in alcoholism and sociopathy. In: Kety, S.S.; Rowland, L.P.; Sidman, R.L.; and Matthyse, S.W., eds. *Genetics of Neurological and Psychiatric Disorders*. New York: Raven Press, 1983. pp. 145–166.



two correlations, the values of a, c, and e can be determined using algebraic rules.

For a simplified illustration of these formulas at work, imagine that a study finds that MZ twin pairs correlate at +0.7 for alcoholism or not, whereas DZ pairs correlate at +0.4 for this characteristic. This pattern would occur if MZ pair members were much more likely than DZ pair members to both be alcoholic or both be nonalcoholic. Inserting these numbers into the equations above [$0.7 = (a^2 + c^2)$; $0.4 = (0.5(a^2) + c^2)$] results in estimates of $a^2 = 0.6$, $c^2 = 0.1$, and $e^2 = 0.3$. This means that 60 percent of the differences among people in the tendency to develop alcoholism would be attributable to differences in their genotypes, 10 percent would be attributable to environmental factors shared with siblings, and 30 percent would be attributable to unshared environmental effects.

Note that the scores A, C, and E are not estimated; standard twin studies provide estimates of a, c, and e, and the overall contribution of these components for a whole population, but usually do not provide estimates of the genetic and environmental levels of risk for individual persons. Researchers in the field of behavior genetics, however, have begun to include measured aspects of the environment in their models and, as molecular genetic technology improves, will be able to include measures of specific genetic regions (see the article on quantitative trait loci by Grisel and Crabbe, pp. 220–227).

— Carol A. Prescott
and Kenneth S. Kendler

FALCONER, D.S. The inheritance of liability to certain diseases estimated from the incidence among relatives. *Annals of Human Genetics* 29:51–76, 1965.

HELZER, J., AND PRYZBECK, T.R. The co-occurrence of alcoholism with other psychiatric disorders in the general population and its impact on treatment. *Journal of Studies on Alcohol* 49(3):219–224, 1988.

JACOB, T., AND BREMER, D.A. Assortative mating among men and women alcoholics. *Journal of Studies on Alcohol* 47:219–222, 1986.

KENDLER, K.S.; HEATH, A.C.; NEALE, M.C.; KESSLER, R.C.; AND EAVES, L.J. A population based twin study of alcoholism in women. *Journal of the American Medical Association* 268(14):1877–1882, 1992.

KENDLER, K.S.; NEALE, M.C.; HEATH, A.C.; KESSLER, R.C.; AND EAVES, L.J. A twin-family study of alcoholism in women. *American Journal of Psychiatry* 151(5):707–715, 1994.

KENDLER, K.S.; MARTIN, N.G.; HEATH, A.C.; AND EAVES, L.J. Self-report psychiatric symptoms in twins and their non-twin relatives: Are twins different? *Neuropsychiatric Genetics*, in press.

LYKKEN, D.T.; MCGUE, M.; BOUCHARD, T.J.; AND TELLEGEN, A. Does contact lead to similarity or similarity to contact? *Behavioral Genetics* 20:547–561, 1990.

MCGUE, M. Genes, environment and the etiology of alcoholism. In: Zucker, R.; Boyd, G.; and Howard, J., eds. *The Development of Alcohol Problems: Exploring the Biopsychosocial Matrix of Risk*. National Institute on Alcohol Abuse and Alcoholism Research Monograph No. 26. NIH Pub. No. (ADM) 94–3495. Bethesda, MD: the Institute, 1994. pp. 1–40.

MCGUE, M.; PICKENS, R.W.; AND SVIKIS, D.S. Sex and age effects on the inheritance of alcohol problems: A twin study. *Journal of Abnormal Psychology* 101:3–17, 1992.

PRESCOTT, C.A., AND KENDLER, K.S. Gender and genetic vulnerability to alcoholism. In: Hunt, W.A., and Zakhari, S., eds. *Stress, Gender, and Alcohol-Seeking Behavior*. National Institute on Alcohol Abuse and Alcoholism Research Monograph No. 29. NIH Pub. No. 95–3893. Bethesda, MD: the Institute, 1995. pp. 23–46.

ROSE, R.J.; KAPRIO, J.; WILLIAMS, C.J.; VIKEN, R.; AND OBEMSKI, K. Social contact and sibling similarity: Facts, issues and red herrings. *Behavioral Genetics* 20:763–778, 1990.

GENETIC ENGINEERING IN ANIMAL MODELS

Susanne Hiller-Sturmhöfel, Ph.D., Barbara J. Bowers, Ph.D., and Jeanne M. Wehner, Ph.D.

Multiple genetic and environmental factors influence the development of alcoholism. To evaluate the contributions of individual genes to the development of alcoholism in living organisms, rather than in tissue-culture experiments, researchers have begun to use new genetic technologies in laboratory animals. These techniques include generating transgenic mice, in which a foreign gene is inserted permanently into the animal's genetic material; generating knockout mice, in which a gene is permanently inactivated; and using antisense ribonucleic acid (RNA) treatment, which allows the temporary inactivation of individual genes. Although not yet widely used in alcohol research, these technologies may allow researchers to study important questions and gain new insights into the causes and consequences of alcoholism. KEY WORDS: animal model; animal strains; AOD dependence; research; genetics and heredity; gene; RNA

The use of animal models to elucidate causes and mechanisms of human diseases and to develop new treatment approaches is a mainstay of modern biological research, including alcohol research. Laboratory animals, such as rats and monkeys, are used to model drinking behavior, to study alcohol-induced damage to different organs, and to analyze the brain chemistry involved in mediating alcohol's effects.

An important focus of current alcohol research is to identify genes that contribute to alcohol drinking behavior and its consequences. The functions of these genes, that is, the mechanisms through which the genes exert their effects, are difficult to study. Genetic studies with isolated cells and tissues, which have been successful in analyzing the causes and consequences of other diseases, cannot reflect behavioral responses, such as those that occur in alcoholism. Conversely, genetic studies in whole organisms, whether laboratory animals or humans, are difficult to conduct and interpret because scientists and researchers believe that alcoholism is influenced by both genetic and environmental factors. In addition, alcoholism is a polygenic disease (i.e., many genes play a role in its development), making it hard to determine the contribution of each individual gene.

Recent progress in employing genetic engineering technologies to develop new animal models eventually may enable researchers to overcome some of these difficulties and to study the role of individual genes and their products in the development of alcoholism. These technologies allow the insertion of foreign genes, the permanent inactivation of specific genes, and the temporary elimination of

particular gene products in a living organism. Using these approaches, researchers can evaluate the impact of individual genes on the development of a disorder such as alcoholism.

This article describes three of these new technologies: transgenic mice, knockout mice, and treatment with antisense ribonucleic acid (RNA). Although these approaches have not been used widely in alcohol research, their use in other research areas indicates both their potential applications and limitations in the alcohol field.

TRANSGENIC MICE

In transgenic animals, a foreign gene is integrated permanently into the animal's genetic material, the DNA,¹ in both the reproductive (i.e., germ) cells and the nonreproductive (i.e., somatic) cells, leading to the expression and propagation of the gene across future generations. This technique has been used primarily to evaluate the role of specific genes during fetal development or to mimic human diseases in animals. In the latter case, scientists introduce into an animal a human gene that is known to cause a disorder and then study the animal's development of the disease. Examples of human diseases that have been studied in transgenic animals include cystic fibrosis (Snouwaert et al. 1992) and muscular dystrophy (Vincent et al. 1993). Studying the mechanism of the development of a disease in more detail may enable scientists to devise better prevention or treatment approaches and subsequently evaluate them in these animals.

Similarly, genes that are known or suspected to contribute to alcoholism could be introduced into transgenic animals. It may be more difficult, however, to measure a gene's effect on a behavior, such as alcohol consumption, than on a specific bodily function or biochemical process. Consequently, alcohol researchers are still evaluating the potential of transgenic animals for their studies.

Creating a Transgenic Mouse

Although researchers can use several mammalian species to create transgenic animals, they primarily use mice. Mice can be bred easily, they have a short generation period, they bear many pups per litter, their embryos can be manipulated easily during experiments, and their genes have been studied extensively (Lovell-Badge 1985).

Before being introduced into a mouse or any other animal, the foreign gene must first be identified and isolated from its original organism (e.g., from the DNA of human cells). Next, many identical copies of the gene are chemically synthesized, which then are injected into mouse embryos.

To create mouse embryos, eggs of laboratory mice are fertilized in a test tube with mouse sperm. The fertilized egg, or embryo, contains two sets of DNA, each of which is contained in a separate structure called a pronucleus. One DNA set comes from the mother (i.e., the female pronucleus); the other set comes from the father (i.e., the male pronucleus). At this stage, the foreign gene is added by injecting the DNA directly into the male pronucleus with a very fine glass needle (figure 1) (also see Gordon et al. 1980). Although this procedure may seem straightforward, its challenges become obvious when one considers that the whole embryo is only about 0.1 mm in diameter.

SUSANNE HILLER-STURMHÖFEL, PH.D., is a science editor of Alcohol Health & Research World.

BARBARA J. BOWERS, PH.D., is a research associate in the Department of Psychiatry, Indiana University, Indianapolis, Indiana.

JEANNE M. WEHNER, PH.D., is a professor at the Institute for Behavioral Genetics at the University of Colorado at Boulder, Boulder, Colorado.

This work was supported by National Institute on Alcohol Abuse and Alcoholism (NIAAA) grant AA-03527 and NIAAA Research Scientist Development Award AA-00141 to Jeanne M. Wehner.

¹For a definition of this and other technical terms used in this article, see central glossary, pp. 182-183.

In the 50 to 90 percent of the injected embryos that survive this procedure (the success rate depends on the scientist's skill), the foreign gene integrates into the embryo's DNA, and the embryo continues to develop normally. The two pronuclei fuse, and the cell begins to divide. The embryos are implanted into surrogate mothers, where about 10 to 30 percent of the embryos develop to term. The mouse pups then can be tested to determine whether they have integrated the foreign gene into their DNA and whether they actually express the gene (i.e., synthesize the protein that is the gene product). (For more information on some of the basic processes of gene expression, see sidebar, p. 208.)

The foreign gene product usually can be detected in 10 to 35 percent of the pups. These numbers demonstrate that the creation of transgenic mice is a somewhat inefficient process (Mann and McMahon 1993). Of 100 embryos injected with a foreign gene, only a few will become viable pups that can grow into adult transgenic animals.

Although biochemical tests can determine if the mice that develop from injected embryos (known as first generation, or F_1 , animals) express the foreign gene, the tests cannot determine if all cells of the animal, particularly the germ cells, have integrated the foreign gene. To ensure the gene's presence in the germ cells, the F_1 animals are mated with each other. Only F_1 animals with the foreign gene in their germ cells can pass it on to their offspring (i.e., second generation, or F_2 , animals). The F_2 animals with the foreign gene will carry it in all their cells. These animals are used to study the gene's function.

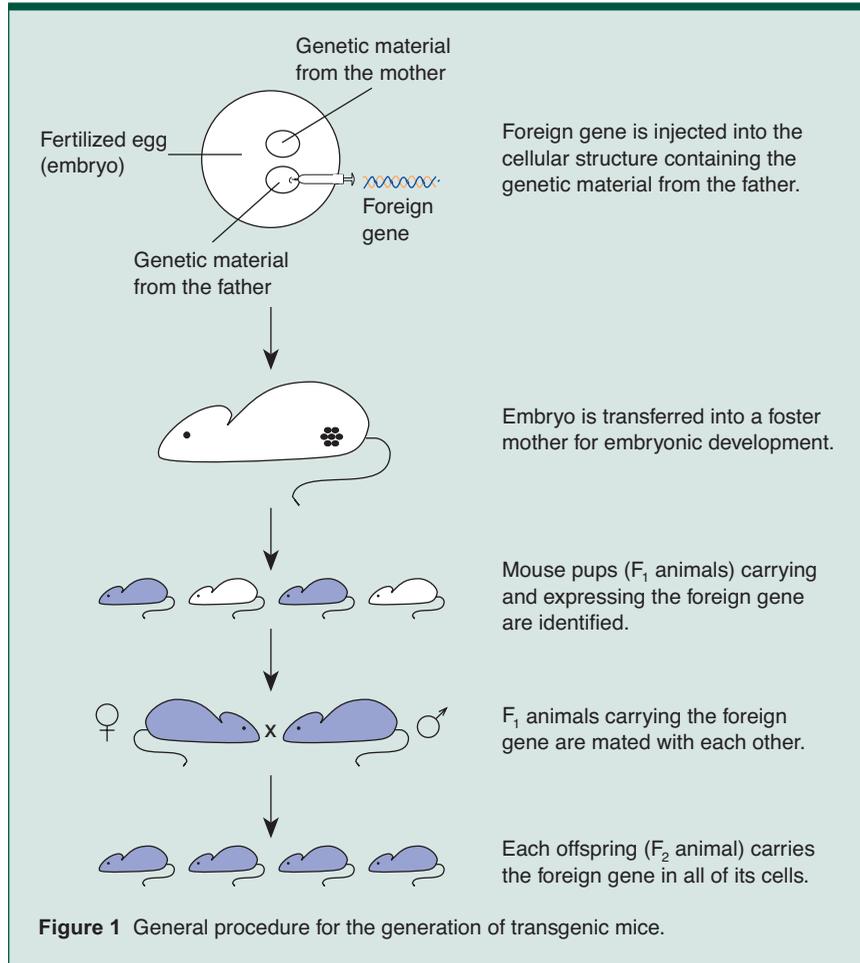


Figure 1 General procedure for the generation of transgenic mice.

APPLICATIONS OF TRANSGENIC MICE

As mentioned previously, transgenic animals have not yet been used specifically for alcohol research. The study of transgenic mice created for unrelated research projects, however, has produced unexpected results relevant to alcohol research that offer a glimpse of the potential of this technology.

Transforming Growth Factor Alpha

Hilakivi-Clarke and colleagues (1992, 1993) studied transgenic mice that express the gene for human transforming growth factor alpha (TGF-alpha) and synthesize excess amounts of TGF-alpha protein. TGF-alpha plays a role during fetal development; excess TGF-alpha, however, can cause liver tumors. The researchers noticed that male transgenic mice expressing human TGF-alpha not only developed liver tumors but also were more aggressive than male nontransgenic animals and displayed altered function of the brain chemical serotonin.

As part of an experiment to study the correlation between aggression and the development of liver tumors in TGF-alpha mice, the animals also received alcohol. When the researchers

analyzed different effects of alcohol on the behavior of the transgenic mice, they found that the transgenic mice differed from the nontransgenic mice in their sensitivity to some of alcohol's effects (Hilakivi-Clarke et al. 1993). For example, the transgenic mice were more sensitive to alcohol's effects on aggression.

Corticotropin Releasing Factor

Another existing transgenic mouse line that may have applications in alcohol research carries the rat gene for corticotropin releasing factor (CRF) (Stenzel-Poore et al. 1994). CRF is a hormone involved in the organism's response to stress. One of the body's physiological reactions to stress is the synthesis of hormones called glucocorticoids. CRF synthesis is the first step in the chain of events leading to glucocorticoid production. Consequently, CRF is studied in many animal experiments as a physiological measure for the behavioral or emotional condition of stress.

Alcohol alters the body's reaction to stress. For example, alcohol has been shown to increase the glucocorticoid levels released in response to stress (e.g., Zgombick and Erwin 1988). The CRF mice therefore might provide insight into the mechanisms by which alcohol affects glucocorticoid production levels under stressful conditions.

Both examples illustrate how transgenic animals created for studies unrelated to alcohol research can be used to study phenomena such as physiological sensitivity to alcohol or alcohol's effects on the response to stress. Other existing transgenic mouse strains might prove to be equally valuable tools in investigating alcohol's effects and mechanisms of action on different organs.

KNOCKOUT MICE

Whereas transgenic mice carry an additional foreign gene in their DNA, knockout mice are characterized by the targeted elimination of one of their own genes and, consequently, gene products. This approach allows inferences about the function of the deleted gene

FROM DNA TO PROTEIN: HOW GENETIC INFORMATION IS REALIZED

All the genetic information that is necessary to create and maintain an organism is encoded in long, thread-like DNA molecules in the nucleus of each of the organism's cells. But how is this information converted into the proteins that compose a significant portion of the cell's components and which drive most chemical reactions in the body? This conversion, which also is referred to as gene expression, is a complex biochemical process that consists of several steps occurring in the cell nucleus and in the cytoplasm. To better understand how gene expression works, it helps to review briefly the chemical structure of DNA. The characteristic design of DNA molecules is the basis for the reactions involved in gene expression.

The building blocks of DNA, the nucleotides, are sugar molecules linked to organic bases. DNA includes four different organic bases: adenine (represented by the letter A), cytosine (represented by the letter C), guanine (represented by the letter G), and thymine (represented by the letter T). The order in which they are arranged specifies which amino acids will be linked to form a protein. Because more than four amino acids exist and are necessary to produce a protein, a triplet of three nucleotides represents (i.e., codes for) one specific amino acid in the final protein. For example, the nucleotide triplet ATG codes for the amino acid methionine, and the triplet TGG codes for the amino acid tryptophan. The section of a DNA molecule containing the information needed to make one specific protein is called a gene.

DNA is a double-stranded molecule: Two chains of nucleotides face each other and are connected through specific bonds (see figure 3 of the main article). Because of the nature of these

bonds, each nucleotide can bind to only one other particular nucleotide. For example, the nucleotide containing A always pairs with the nucleotide containing T, and the nucleotide containing C always pairs with the nucleotide containing G. The composition of the second strand therefore depends on the composition of the first strand. Accordingly, the strands are called complementary. This also means that if one knows the nucleotide sequence of one strand, one can automatically infer the sequence of the second strand.

Transcription

To convert the information encoded in the DNA of one gene into a protein, the first step is to copy, or transcribe, one of the DNA strands into another nucleic acid molecule called messenger ribonucleic acid (mRNA). This process is performed by specific enzymes in the cell nucleus.

There are different kinds of RNA in the cell that have different functions but the same chemical structure. RNA molecules are similar in their chemical composition to DNA molecules. The main differences are that the sugar component differs between DNA and RNA and that the organic base thymine, which is present in DNA, is replaced by the base uracil (represented by letter U) in RNA. In addition, RNA molecules are single stranded; unlike DNA, they do not have a complementary strand.

During transcription, the DNA sequence representing one gene is converted into mRNA. Only one strand of the double-stranded DNA molecule, however, serves as a template for mRNA synthesis. RNA nucleotides are guided to the DNA sequence that is being transcribed and temporarily bind to it. Again, only one specific RNA nucleotide can bind to each DNA nucleotide (e.g., the RNA nucleotide containing A pairs with the DNA nucleotide containing T, and the RNA nucleotide containing C pairs with the DNA nu-

cleotide containing G). This specificity guarantees that the genetic information contained in the DNA is accurately converted into mRNA. As with the DNA template, the sequence of a triplet of nucleotides in the RNA codes for one amino acid in the final protein.

After all the information for one gene has been copied into an mRNA molecule, the DNA and mRNA molecules separate. The mRNA then undergoes some additional modifications in the cell's nucleus before it is transported to the cytoplasm for the next step, the translation into the protein product.

Translation

In the cell's cytoplasm, macromolecules called ribosomes attach to, and slide along, the mRNA. In this manner, the ribosomes "read" the sequence of the mRNA's nucleotide triplets. According to that sequence, the ribosomes recruit a second kind of RNA, the so-called transfer RNA (tRNA) molecules, which guide the amino acids needed for protein synthesis to the mRNA-ribosome complex. One end of each tRNA molecule has a region that recognizes one specific nucleotide triplet on the mRNA. Another region of each tRNA molecule is attached to a specific amino acid. Thus, by recruiting tRNA molecules that recognize the nucleotide sequence of the mRNA, the ribosomes also retain the right amino acids in the right order to form the protein encoded by the gene represented in the mRNA. Specific enzymes then connect the amino acids until the complete protein is synthesized. Because each mRNA molecule can be read consecutively by several ribosomes, many protein molecules can be derived from just one mRNA template.

—Susanne Hiller-Sturmhöfel

by comparing the phenotype (i.e., the appearance or behavior) of the knockout mice with that of normal mice. In extreme cases, when the deleted gene has a vital function during embryonic development, knockout mice lacking the gene will not develop beyond a certain embryonic stage (Li et al. 1994; Klein et al. 1993). In the more desirable scenarios, only one specific aspect of metabolism or behavior will be eliminated or modified.

Creating a Knockout Mouse

As with the starting material for transgenic mice, the gene to be altered in knockout mice also must be identified, isolated, and copied first. In most knockout mice, the targeted gene is not actually deleted from the genome; instead, changes, or mutations, are introduced into the isolated gene, preventing it from encoding a functional gene product. The mutated gene then is transferred into mouse cells.

In contrast with the procedure for transgenic mice, the mutated gene is not introduced directly into an embryo but into a cell type, called embryonic stem (ES) cells (figure 2) (also see Capecchi 1989). These cells can differentiate into all the different cell types and tissues that compose the animal. Once the mutated gene has entered the ES cells, it can exchange places with the cells' normal gene through a process known as homologous recombination. ES cells expressing the mutated gene are identified by growing the cells in a petri dish in a specific medium in which only the modified cells can survive.

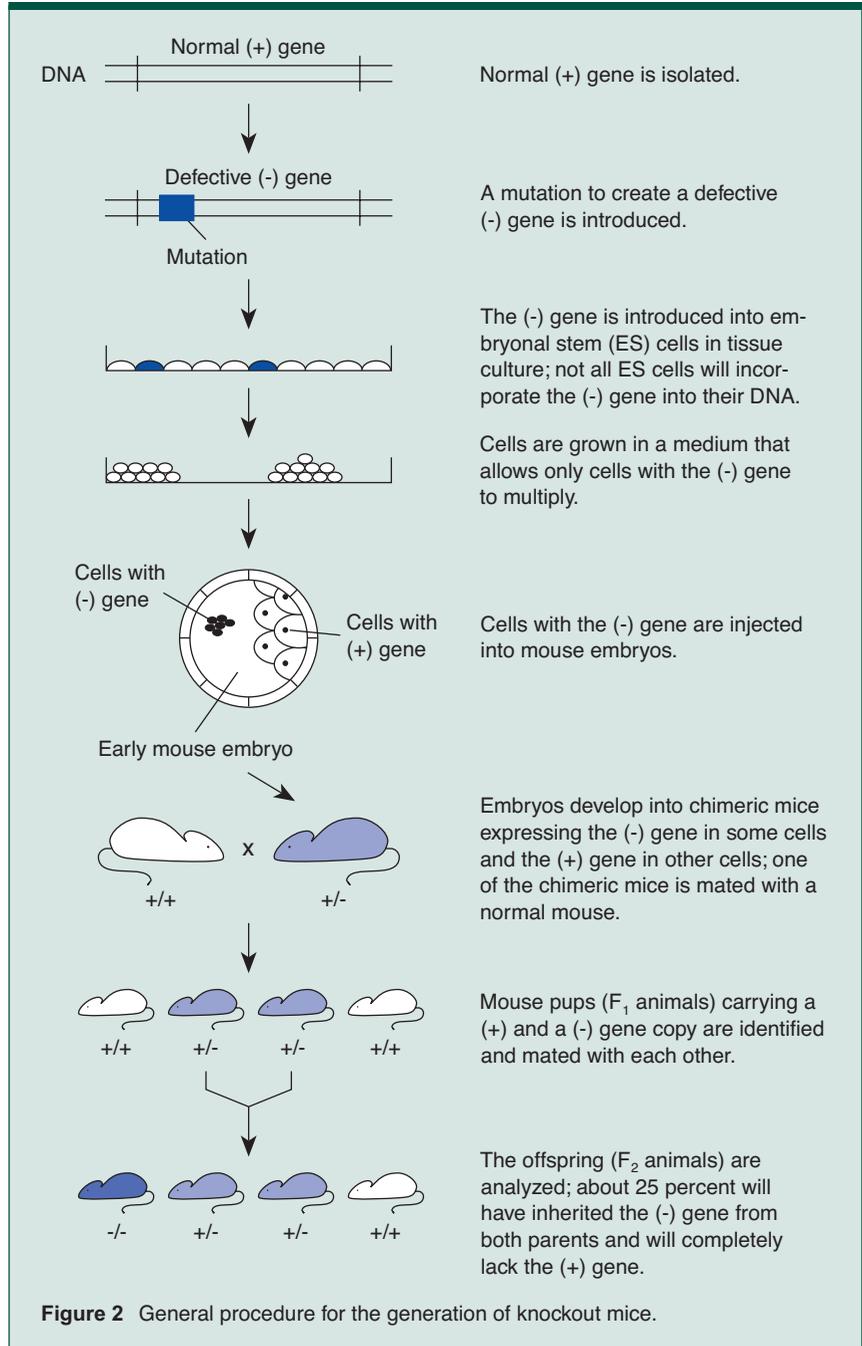
These modified cells then are injected into a mouse embryo at an early stage of development, and the embryo is implanted into a surrogate mother. The injected ES cells and the embryo's own cells develop together into an intact mouse in which some cells are derived from the ES cells and consequently contain the mutated gene, whereas other cells are derived from the embryo's own cells and contain the normal gene. Such a mouse is called a chimera and is only the first step in the creation of the real knockout mouse. In the chimera, the mutated gene is not yet present in all the animal's cells, nor is it integrated in both the maternal and the paternal copies of the gene.

To determine if the chimeras have incorporated the mutated gene into their germ cells, they are mated with normal mice. Pups (i.e., F₁ animals) in which the mutated gene can be detected must have inherited this gene from the chimeric parent. In addition, they still have an intact gene copy that they have inherited from the normal parent. To replace this copy with a mutated gene, two F₁ pups are mated with each other. According to the basic laws of

inheritance, 25 percent of the F₂ offspring from this mating will carry two copies of the mutated gene. These F₂ animals are the knockout mice that are the goal of this long and arduous procedure and whose function or behavior will be analyzed.

Applications of Knockout Mice

Thus far, the technology of knockout mice rarely has been applied to alcohol research, partly because only a few genes have been identified that may contribute to the susceptibility to alcoholism or that



may mediate alcohol's effects. As researchers' knowledge increases, however, new applications for knockout mice in alcohol research may become apparent. Already some promising candidate genes have been identified, including neurotransmitters and other molecules that transmit chemical or electrical signals within cells. These substances may play a role in alcohol's effects on the brain.

Gamma-Protein Kinase C. Protein kinases, including protein kinase C (PKC), are enzymes that activate or deactivate the function of proteins by attaching phosphate groups to the proteins. The activation and deactivation of specific proteins is one component of a signaling mechanism through which chemical signals are relayed from the cell's surface to its interior (Kikkawa et al. 1989). Gamma-PKC is one of several PKC molecules that exist in the body (Kikkawa et al. 1989) and thus also is involved in intracellular signal transmission.

Several lines of evidence suggest that PKC function correlates with alcohol's effects on the brain. For example, PKC may modify and thus affect the function of a receptor that is located on nerve cells and is activated by the neurotransmitter gamma-aminobutyric acid (GABA). Alcohol researchers have studied the GABA receptor intensely, because it may be responsible for some of alcohol's effects in the brain. In addition, researchers have found that brain cells containing gamma-PKC are sensitive to alcohol (Palmer et al. 1992).

To study whether gamma-PKC-dependent signal transmission is affected by alcohol, Harris and colleagues (1995) used knockout mice lacking functional gamma-PKC. Because gamma-PKC is not essential during development, the knockout mice developed and reproduced normally. The PKC knockout mice were less sensitive to alcohol's sleep-inducing effects than were their normal littermates. Also, alcohol's enhancement of GABA-induced reactions was reduced in brain extracts from knockout mice compared with brain extracts from normal mice. These results indicate that gamma-PKC may be involved in determining an organism's sensitivity to alcohol.

LIMITATIONS OF TRANSGENIC AND KNOCKOUT MOUSE TECHNOLOGY

Although the creation of transgenic and knockout mice is a powerful tool that eventually may provide scientists with a better understanding of human drinking behavior and of alcohol's effects on the brain and other organs, some limitations exist to the usefulness and validity of such experiments. These limitations are not specific to alcohol research but apply to all research areas analyzing bodily functions and disorders that depend on the cooperation of several genes or that manifest themselves through behavior rather than biochemical reactions.

First, the individual gene being studied must have a large enough impact on the development of the disease (e.g., alcoholism) so that the effect of the gene's overexpression or deletion can be detected reliably among all other factors contributing to the disorder. Furthermore, other genes sometimes compensate naturally for the function of the deleted gene in knockout mice, thereby masking the deletion's effect. Accordingly, not all genes and their products can be analyzed using these technologies; the functions of some genes may be identified more readily than the functions of others.

Second, the gene product studied must not be vital to embryonic development; otherwise, too much or too little of it will interfere with normal development. In that case, the transgenic or knockout mouse embryos will not develop to term, and the desired effects (e.g., those resulting from alcohol) cannot be studied.

Third, in transgenic mice, the integration of foreign DNA into mouse DNA to date cannot be targeted to a specific region of the mouse DNA. Consequently, the foreign DNA may integrate in the middle of another gene and thus disturb that gene's function (Lacy et al. 1983). Researchers should be aware of this possibility when interpreting their findings.

Fourth, similar to the site of foreign DNA integration, the amount of the foreign gene and its product in each cell cannot be predicted. The effect of the foreign gene on the transgenic animal may vary, depending on how much of the gene product is made. Similarly, in many cases the foreign gene will be expressed in all tissues, in contrast with the normal physiological condition, in which the gene may be expressed only in selected cells (Lacy et al. 1983). To circumvent this problem, scientists currently are developing methods that target the expression of foreign genes to the cells in which they normally are active. This process uses specific regulators that allow gene expression only in specific cell types or tissues.

Fifth, the choice of mouse lines used to create transgenic or knockout mice may affect the experiments' results. Just as two people may respond differently to alcohol although they share some of the same genes that mediate alcohol's effects, two mouse lines may react differently to the addition or deletion of the gene to their genetic material (Grant et al. 1992; Liu et al. 1993).

Nonetheless, these caveats do not diminish the potential that transgenic and knockout mice hold for alcohol research; they illustrate, however, that the results obtained with genetically engineered animals should be interpreted cautiously, because they may not be as straightforward as they initially seem.

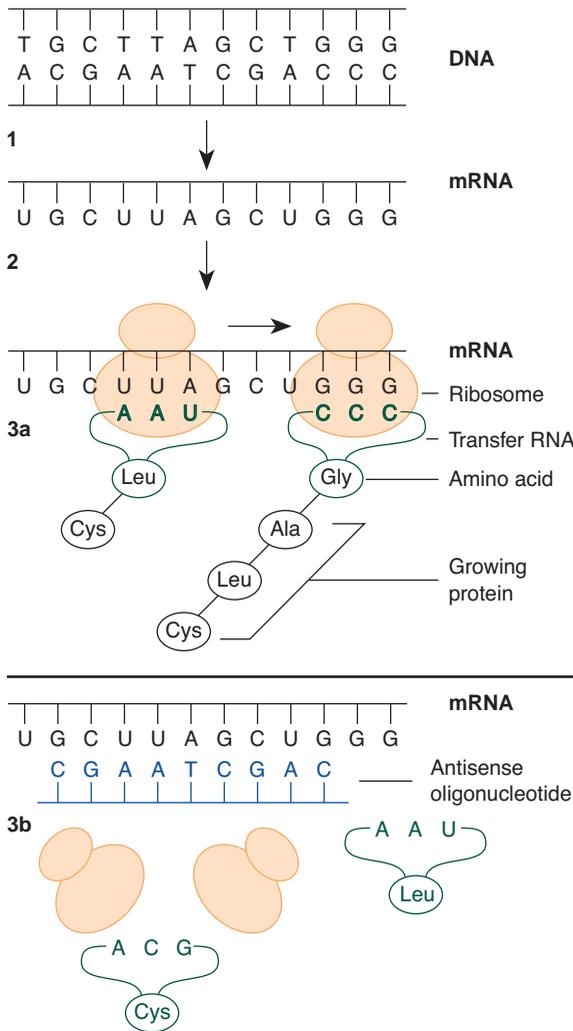
ANTISENSE RNA STRATEGIES

Another method for exploring the role of specific genes in mediating alcohol's effects in a living organism is antisense RNA technology (Crooke 1992). Similar to the approach using knockout mice, antisense RNA technology reduces or prevents the expression of a specific gene. In contrast with the knockout mice, however, this modification usually is not complete or permanent, because the antisense RNA can be administered to the animals temporarily.

Antisense RNA technology has several potential advantages compared with transgenic and knockout mice. For example, antisense RNA treatment is faster and cheaper than creating and breeding genetically modified mice. Because it is not a permanent modification, antisense RNA treatment also avoids some of the limitations of knockout and transgenic mice, such as the difficulty of modifying genes that are critical during development. With antisense RNA technology, the animals can develop normally before their gene expression is manipulated. Finally, antisense RNA may have some therapeutic potential in humans, for example, by targeting the receptors for certain neurotransmitters in the brain, provided that adequate amounts of RNA can be delivered to brain cells (Crooke 1992).

What Is Antisense RNA Treatment?

The conversion of the genetic information encoded in the DNA into a protein product is a complex process. Briefly, through a process called transcription, the DNA information in the cell's nucleus is copied into an intermediary molecule, the messenger RNA (mRNA). The mRNA is transported to the cell's cytoplasm, where it serves as a template for the synthesis of a protein in a process referred to as



(1) The nucleotide sequence (gene) of one of the two DNA strands is copied (transcribed) into messenger RNA (mRNA).

(2) The mRNA moves from the cell's nucleus into the cytoplasm.

(3a) Ribosomes move along the mRNA, recruiting carrier molecules (transfer RNA's, or tRNA's) that each carry a specific amino acid. The amino acids are linked to form the protein.

(3b) In antisense RNA treatment, oligonucleotides (short, synthetic DNA molecules) bind to the mRNA, forming a double-stranded DNA/RNA hybrid to which ribosomes cannot attach.

Figure 3 The conversion of genetic information into protein without and with antisense RNA treatment. Steps 1–3a show the usual way in which the information on a DNA strand serves as a blueprint for generating proteins. In antisense RNA treatment (3b), a “dummy” sequence of DNA prevents ribosomes from carrying out the process of making proteins. Using this technique, researchers may be able to investigate the link between genes and alcohol-related problems. For example, certain proteins may be needed to manufacture neurotransmitters involved in the desire to consume alcohol; if blocking the creation of one of those proteins would change alcohol consumption, the gene(s) responsible for making that protein might be involved in the urge to drink alcohol.

translation (for more information on the steps involved in transcription and translation, see figure 3 and the sidebar on p. 208).

Antisense RNA technology aims to inhibit the translation of an mRNA into its respective protein. As with transgenic and knockout mice, the technique requires that the gene to be studied has been isolated and copied and that its exact DNA sequence is known. During antisense RNA treatment, however, the gene itself is not the target of the procedure. Instead, the target is the mRNA transcribed from the gene.

To block translation of the mRNA, scientists make a short synthetic DNA molecule called an oligonucleotide, which can bind to the end of the mRNA molecule at which translation begins. This oligonucleotide is called the “antisense” molecule because it is complementary (i.e., like a reverse image) to a section of the “sense” information encoded in the mRNA. When the DNA oligonucleotide is introduced into a cell, it can attach itself to the complementary region of the mRNA, creating a region that is a DNA–RNA hybrid (figure 3). This DNA–RNA hybrid region interferes with translation of the mRNA, because the proteins required for this process

bind only to the free RNA molecule, not to a DNA–RNA hybrid. In addition, the DNA–RNA molecules degrade rapidly in the cell. Consequently, no protein is synthesized from the mRNA.

This mechanism works in isolated cells, tissues, and whole animals. The oligonucleotide, which is readily taken up by individual cells, can be delivered to cells and tissues by various methods (Wahlestedt 1994). It can be administered to animals by a one-time injection or over a longer period by chronic transfusion. Cells in specifically targeted organs or regions of the body incorporate the oligonucleotide. By varying the length or frequency of oligonucleotide administration, researchers can modify the duration for which the targeted protein is not being synthesized. They also can study the effects of this manipulation at different developmental stages and under different experimental conditions (e.g., in the presence or absence of alcohol). Consequently, antisense RNA technology allows a flexibility in studying the effects of individual genes that cannot be achieved with transgenic or knockout mice in which the genetic material is modified permanently.

Applications of Antisense RNA Technology

Antisense RNA strategies have not yet been applied specifically to alcohol research. However, the functions of several neurotransmitters and their receptors, both of which may contribute to alcohol's effects on the brain, have been studied under experimental conditions not related to alcohol.

N-methyl-D-aspartate Receptor. One important neurotransmitter in the brain is the amino acid glutamate, which binds to a receptor on the nerve cell surface called the N-methyl-D-aspartate (NMDA) receptor. Alcohol researchers study this receptor because its response to glutamate binding is thought to contribute to alcohol withdrawal seizures.

The NMDA receptor also has other effects on the brain that are not related to alcohol. For example, its response to glutamate binding may be involved in causing cell death after ischemic strokes in the brain. An ischemic stroke occurs when obstruction of a blood vessel disrupts the blood flow to an area of the brain. Such a stroke can lead to the death of the brain cells in that area.

When isolated brain cells were treated with antisense oligonucleotides that could bind to the NMDA receptor mRNA, the number of NMDA receptors on the cells decreased, and cell death could be induced only to a lesser extent than in untreated cells (Wahlestedt et al. 1993). Similarly, in animals treated with the antisense oligonucleotides, the number of cells that died after the obstruction of a brain artery also was reduced significantly (Wahlestedt et al. 1993). These findings demonstrate that the NMDA receptor does in fact contribute to cell death after ischemic strokes. In addition, the example shows that antisense RNA technology can affect the expression of genes both in isolated cells and in whole animals, thereby allowing researchers to study the function of these genes in both experimental systems.

LIMITATIONS OF ANTISENSE RNA TECHNOLOGY

One major advantage of antisense RNA technology—that it is a temporary, easy-to-manipulate pharmacological approach—also contributes to the technology's limitations. Some of these limitations are as follows:

- Delivering sufficient amounts of antisense oligonucleotides to tissues in living animals, especially to the brain, can be diffi-

cult. As a result, the inhibition of a gene's function in an animal or even in a specific tissue rarely is complete.

- Antisense oligonucleotides only have a limited lifespan in the body before they are broken down (Whitesell et al. 1993). Researchers are trying to modify the chemical structure of the oligonucleotides to extend that lifespan, but a specific treatment protocol probably will need to be determined empirically for each gene or gene product to be studied.
- The lifespans of the mRNA and the protein being studied must be established for each experiment, because they determine the duration of the antisense treatment required to produce meaningful results.
- Researchers must ensure that the oligonucleotides bind only to the desired mRNA and do not interfere with the translation of other mRNA's in the cell. Such an unwanted interaction could distort the results and lead to misinterpretations of the gene's function.
- Oligonucleotide treatment in some instances can be toxic to cells or to living animals (Wahlestedt 1994).

SUMMARY

This article presents three exciting new technologies that allow scientists to study the functions of single genes in the context of the living organism. Because these techniques specifically examine the functions of individual genes, they appear particularly well suited for the analysis of genes involved in the development of polygenic disorders such as alcoholism.

So far, these technologies have not been applied systematically to studying the causes and effects of alcoholism. This partly is attributable to the fact that researchers have isolated few candidate genes that may contribute to the development of alcoholism or to mediating alcohol's effects. As some of the examples in this article demonstrate, however, even seemingly unrelated experiments may produce results relevant to alcohol research. At the very least, the examples illustrate the potential of new genetic engineering technologies in animal models to help scientists answer pressing questions in many research areas. ■

REFERENCES

- CAPECCHI, M.R. Altering the genome by homologous recombination. *Science* 244:1288–1292, 1989.
- CROOKE, S.T. Therapeutic applications of oligonucleotides. *Annual Review of Pharmacology and Toxicology* 32:329–376, 1992.
- GORDON, J.W.; SCANGOS, G.A.; PLOTKIN, D.J.; BARBOSA, J.A.; AND RUDDLE, F.H. Genetic transformation of mouse embryos by microinjection of purified DNA. *Proceedings of the National Academy of Sciences USA* 77:7380–7384, 1980.
- GRANT, S.G.N.; O'DELL, T.J.; KARL, K.A.; STEIN, P.L.; SORIANO, P.; AND KANDEL, E.R. Impaired long-term potentiation, spatial learning, and hippocampal development in *fyn* mutant mice. *Science* 258:1903–1910, 1992.
- HARRIS, R.A.; MCQUILKIN, S.J.; PAYLOR, R.; ABELIOVICH, A.; TONEGAWA, S.; AND WEHNER, J.M. Mutant mice lacking the gamma isoform of protein kinase C show decreased behavioral actions of ethanol and altered gamma-aminobutyrate type A receptors. *Proceedings of the National Academy of Sciences USA* 92:3658–3662, 1995.
- HILAKIVI-CLARKE, L.A.; ARORA, P.K.; SABOL, M.-B.; CLARKE, R.; DICKSON, R.B.; AND LIPPMAN, M.E. Alterations in behavior, steroid hormones and natural killer cell activity in male TGF alpha mice. *Brain Research* 588:97–103, 1992.

PRIMATES IN ALCOHOL RESEARCH

J. Dee Higley, Ph.D.

The genetic similarity to humans of nonhuman primates makes them well suited to serve as models of complex human disorders such as alcoholism. Like humans, nonhuman primates vary with respect to their alcohol consumption, even within the same species. Studies of the origins of high consumption among nonhuman primates have suggested that both genetic and environmental factors play a role in their drinking. In fact, researchers have found some support for multiple subtypes of alcoholism among nonhuman primates. KEY WORDS: animal model; AOD dependence; research; hereditary factors; environmental factors; Cloninger's typology

Because of their close genetic similarity to humans and their complex social behaviors, nonhuman primates have been widely used to study a number of psychiatric syndromes (for a review, see McKinney 1988). Although they have been used less often in alcohol research, nonhuman primates are ideally suited as subjects for psychobiological studies of excessive alcohol consumption, because the similarities in human and other primate DNA result in their sharing many physiological and behavioral processes. In addition, nonhuman primates have genetic and rearing backgrounds that can be carefully planned so as to allow for observation in closely controlled settings. These two attributes allow researchers to extrapolate results from studies of nonhuman primates to humans more readily than the results from less closely related animals.

This article reviews examples of ongoing research that uses nonhuman primates to investigate how genetic predispositions influence alcohol consumption. It also examines how a genetic predisposition may interact with other factors, such as environment and exposure to stress, to produce greater voluntary alcohol consumption.

PRIMATE DRINKING

One reason nonhuman primates are not used frequently in alcohol research is that they rarely consume alcohol in the amounts necessary to create models of human alcoholism. Some researchers have suggested that this is because the animals find the initial taste of alcohol aversive (Higley et al. in press *b*). Most humans also initially find the taste of alcohol unpleasant, particularly when the alcohol concentration in a solution is high. Consequently, alcohol is rarely consumed in its pure state. Instead, it is consumed in solutions with low concentrations; the taste often is disguised using colas, fruit juices, and other flavorings. Similarly, in at least some species of nonhuman primates, when the solution is palatable and the concentration of alcohol is under 15 to 20 percent, most animals will consume alcohol at rates producing pharmacological effects. Only about 10 to 20 percent of subjects who have not been subjected to stressful environmental conditions that induce drinking, however, will freely consume palatable alcohol solutions at rates that consistently produce blood alcohol levels greater than the "legal" level of intoxication for most States (Ervin et al. 1990; Higley et al. in press *b*; Kraemer and McKinney 1985).

J. DEE HIGLEY, PH.D., is a research psychologist at the National Institute on Alcohol Abuse and Alcoholism, Laboratory of Clinical Studies, Primate Unit, Poolesville, Maryland.

- HILAKIVI-CLARKE, L.; DURCAN, M.; AND GOLDBERG, R. Effect of alcohol on elevated aggressive behavior in male transgenic TGF alpha mice. *NeuroReport* 4:155-158, 1993.
- KIKKAWA, U.; KISHIMOTO, A.; AND NISHIZUKA, Y. The protein kinase C family: Heterogeneity and its implications. *Annual Review in Biochemistry* 58:31-44, 1989.
- KLEIN, R.; SMEYNE, R.J.; WURST, W.; LONG, L.K.; AUERBACH, B.A.; JOYNER, A.L.; AND BARBACID, M. Targeted disruption of the *trkB* neurotrophin receptor gene results in nervous system lesions and neonatal death. *Cell* 75:113-122, 1993.
- LACY, E.; ROBERTS, S.; EVANS, E.P.; BURTONSHAW, M.D.; AND COSTANTINI, F.D. A foreign beta-globin gene in transgenic mice: Integration at abnormal chromosomal positions and expression in inappropriate tissues. *Cell* 34:343-358, 1983.
- LI, Y.; ERZURUMLU, R.S.; CHEN, C.; JHAVERI, S.; AND TONEGAWA, S. Whisker-related neuronal patterns fail to develop in the trigeminal brainstem nuclei of NMDAR1 knockout mice. *Cell* 76:427-437, 1994.
- LIU, J.-P.; BAKER, J.; PERKINS, A.S.; ROBERTSON, E.J.; AND EFSTRATIADIS, A. Mice carrying null mutations of the genes encoding insulin-like growth factor I (*Igf-1*) and type 1 IGF receptor (*Igf1r*). *Cell* 75:59-72, 1993.
- LOVELL-BADGE, R.H. Transgenic animals: New advances in the field. *Nature* 345:628-629, 1985.
- MANN, J.R., AND MCMAHON, A.P. Factors influencing frequency production of transgenic mice. In: Wassarman, P.M., and DePamphilis, M.L., eds. *Guide to Techniques in Mouse Development*. Vol. 255. Boston, MA: Academic Press, 1993. pp. 771-782.
- PALMER, M.R.; HARLAN, J.T.; AND SPUHLER, K. Genetic covariation in low alcohol-sensitive and high alcohol-sensitive selected lines of rats: Behavioral and electrophysiological sensitivities to the depressant effects of ethanol and the development of acute neuronal tolerance to ethanol *in situ* at generation eight. *Journal of Pharmacology and Experimental Therapeutics* 260:879-886, 1992.
- SNOUWAERT, J.N.; BRIGMAN, K.K.; LATOUR, A.M.; MALOUF, N.N.; BOUCHER, R.C.; SMITHIES, O.; AND KOLLER, B.H. An animal model for cystic fibrosis made by gene targeting. *Science* 257:1083-1088, 1992.
- STENZEL-POORE, M.P.; HEINRICH, S.C.; RIVEST, S.; KOOB, G.F.; AND VALE, W.W. Overproduction of corticotropin-releasing factor in transgenic mice: A genetic model of anxiogenic behavior. *Journal of Neuroscience* 14:2579-2584, 1994.
- VINCENT, N.; RAGOT, T.; GILGENKRANTZ, H.; COUTON, D.; CHAFEY, P.; GREGOIRE, A.; BRIAND, P.; KAPLAN, J.-C.; KAHN, A.; AND PERRICAUDET, M. Long-term correction of mouse dystrophic degeneration by adenovirus-mediated transfer of a minidystrophin gene. *Nature Genetics* 5:130-134, 1993.
- WAHLESTEDT, C. Antisense oligonucleotide strategies in neuropharmacology. *Trends in Pharmacological Sciences* 15:42-46, 1994.
- WAHLESTEDT, C.; GOLANOV, E.; YAMAMOTO, S.; YEE, F.; ERICSON, H.; YOO, H.; INTURRISI, C.E.; AND REIS, D.J. Antisense oligodeoxynucleotides to NMDA-R1 receptor channel protect cortical neurons from excitotoxicity and reduce focal ischemic infarctions. *Nature* 363:260-263, 1993.
- WHITESSELL, C.; GESELOWITZ, D.; CHAVANY, C.; FAHMY, B.; WALBRIDGE, S.; ALGER, J.R.; AND NECKERS, L.M. Stability, clearance, and disposition of intraventricularly administered oligodeoxynucleotides: Implications for therapeutic application within the central nervous system. *Proceedings of the National Academy of Sciences USA* 90:4665-4669, 1993.
- ZGOMBICK, J.M., AND ERWIN, V.G. Ethanol differentially enhances adrenocortical response in LS and SS mice. *Alcohol* 5:287-294, 1988.

Because few animals consume alcohol at rates higher than the legal intoxication levels, when using nonhuman primates, such as monkeys, researchers must have access to a sizable population of subjects. The cost of maintaining nonhuman primates prevents most research facilities from supporting populations large enough to study excessive alcohol consumption.

Most nonhuman primates that have been tested for alcohol preference are Old World species (e.g., baboons, vervet monkeys, and many macaques), although at least one study investigated squirrel monkeys, a New World species that frequently is used in the laboratory (Kaplan et al. 1981). Although consumption rates vary somewhat between species, when an alcohol solution is palatable and freely available, most Old World species consume alcohol in amounts that produce pharmacological effects. Nevertheless, wide differences exist between individuals within a species. However, as discussed below, other factors can increase the amount of alcohol that individual subjects consume.

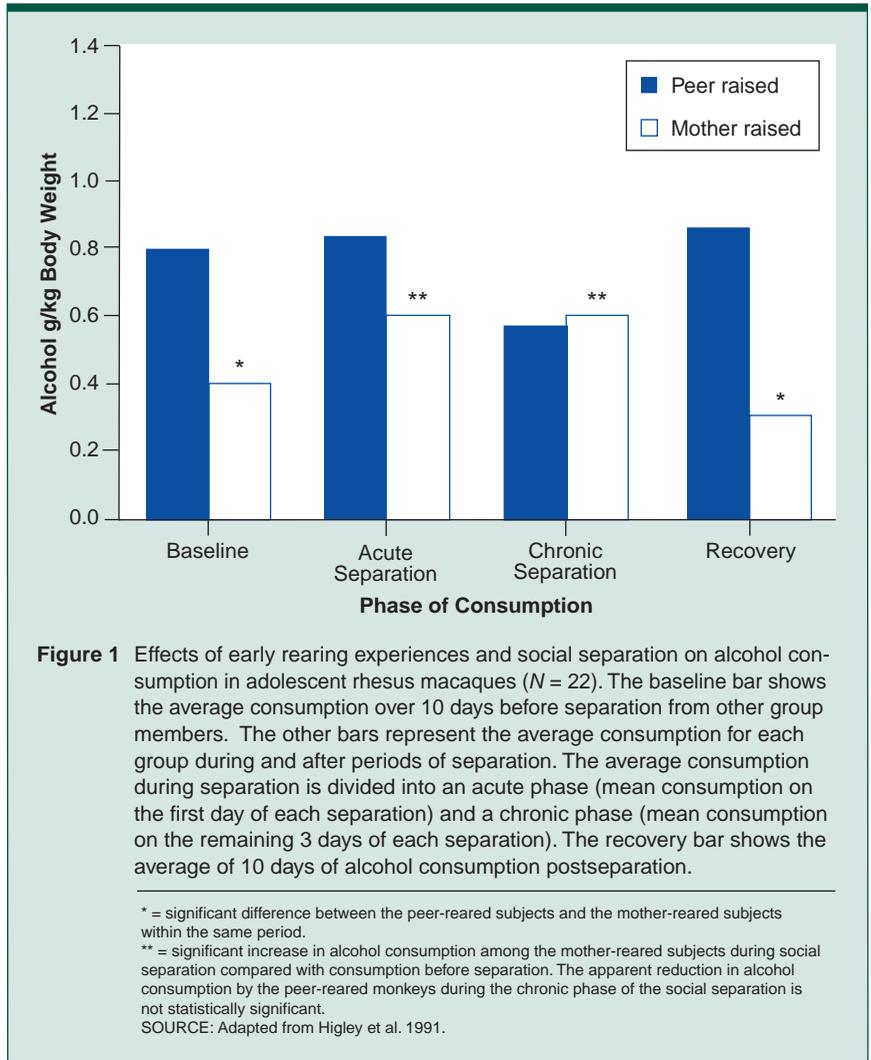
STUDIES OF INDIVIDUAL DIFFERENCES

Researchers studying alcohol consumption in nonhuman primates have made significant progress in developing an animal model of alcohol abuse¹ by focusing research on subjects that show high rates of voluntary alcohol consumption. Studies show that once primates' rates of consumption stabilize, average differences in consumption between subjects are markedly consistent over time (Higley et al. 1991; Kraemer and McKinney 1985). A similar phenomenon has been observed in humans. This traitlike alcohol-consumption pattern indicates the potential for nonhuman primates to model the etiological mechanisms of high alcohol preferences.

Applying Subtypes to Nonhuman Primates

Researchers investigating alcohol problems now widely recognize that alcohol abuse and alcoholism have multiple etiologies (Cloninger 1987; Gilligan et al. 1987). Thus, categorizing alcoholism into subtypes has developed as a method for recognizing patterns in the symptoms and causes of the disorder. Cloninger delineated two of the best known subtypes (Cloninger 1987). The first subtype, anxiety-related alcoholism, Cloninger labeled as type I. This subtype typically is less severe than the second subtype, develops during adulthood, and has been associated with adverse experiences in childhood. The second subtype, which Cloninger labeled type II alcoholism, is characterized by aggression, antisocial behavior, and the onset of alcohol prob-

¹In this article, the term "alcohol abuse" refers to patterns of problem drinking that have resulted in adverse effects on health, social functioning, or both. The term "alcoholism" refers to abnormal alcohol-seeking behavior that leads to impaired control over drinking.



lems early in life. Both subtypes appear to be influenced by inherited, or genetic, factors. Nonhuman primates have only recently begun to be used to investigate the underlying mechanisms that produce type II alcoholism. They have been employed with increasing frequency, however, in studies of the type I, or anxiety-related, disorder.

Research Methods. Studies of the factors that affect drinking rates in nonhuman primates often involve comparing subjects exposed to stressful environments during infancy with those raised under normal laboratory conditions. For example, rhesus macaques that spend the first 6 months of their lives only in the constant company of their age peers (i.e., peer reared) and are chronically absent from their mothers behave differently, displaying less curiosity and more fear, than those raised by their mothers (i.e., mother reared). In adolescence and in adulthood, drinking rates in the two groups can be compared during additional stress-provoking challenges, such as separating subjects from other monkeys for certain periods of time (i.e., social separation) (Higley et al. 1991). The macaques' responses to the challenges differ, depending at least in part on their rearing experiences (discussed below).

Type I. Studies using adolescent rhesus macaques have shown that during nonstressful periods, such as routine life in subjects' home environments, individual rates of alcohol consumption are positively correlated with characteristics of anxiety and fearfulness among animals (Higley et al. 1991). Such characteristics may be either learned or inherited. For example, these studies show that early rearing experiences placing subjects at risk for high levels of anxiety, such as depriving an infant of its parents, may increase alcohol consumption. In fact, the majority of rhesus macaque subjects that are reared in peer-only groups consume alcohol during nonstressful periods at rates producing average blood alcohol concentrations that exceed most legal intoxication limits (Higley et al. 1991). High rates of alcohol consumption also can be induced in many mother-reared subjects by increasing the amount of stress to which they are subjected. For example, in response to a social-separation stressor, mother-reared adolescent subjects increase their rates of alcohol consumption to match the rates of peer-reared subjects (figure 1) (Higley et al. 1991; Kraemer and McKinney 1985).

Nevertheless, not all mother-reared macaques increase their alcohol consumption rates in response to such stressors, and some peer-reared subjects do not develop high consumption rates following their early rearing experiences. These animals also display lower levels of anxiety, and their moderate responses may result from genetic predispositions. Within both the peer- and the mother-reared groups, individual differences in alcohol consumption are directly related to behavioral measures of anxiety, such as clasping themselves or behaviorally withdrawing, and biological measures, such as levels of plasma corticotropin, a hormone secreted by the pituitary gland that increases during periods of most types of stress. These findings suggest that increased levels of anxiety, whether resulting from prior early rearing experiences or from temporary challenges such as social separation, can increase alcohol consumption. These results are similar to those observed in Cloninger's type I alcoholism (Cloninger 1987).

Type II. Higley and colleagues recently have begun a series of studies designed to investigate parallels between Cloninger's type II alcoholism and correlates of high alcohol consumption in adolescent and young adult rhesus macaques living in groups. These studies have been reported in preliminary form (Higley et al. 1994; King et al. 1993).

The findings show that high rates of alcohol consumption are observed in subjects with infrequent social interactions, less competent social behaviors, and high rates of aggression. These subjects are violent, impulsive, and eventually ostracized by their peers (Higley et al. 1993; Mehlman et al. 1994). Their antisocial-like characteristics appear to be related in part to early rearing experiences. Subjects that are peer reared show impaired functioning of the brain chemical serotonin (as measured by cerebrospinal fluid [CSF] concentrations of 5-hydroxyindoleacetic acid [5-HIAA], a product of serotonin metabolism) beginning

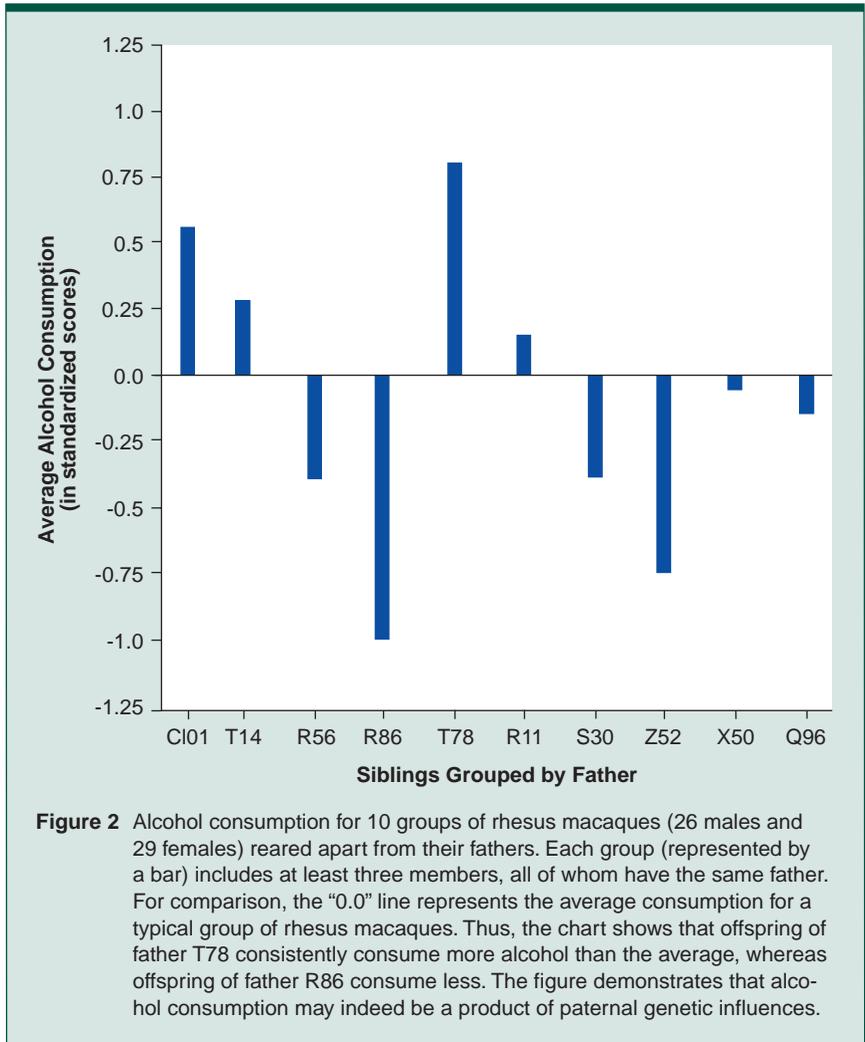


Figure 2 Alcohol consumption for 10 groups of rhesus macaques (26 males and 29 females) reared apart from their fathers. Each group (represented by a bar) includes at least three members, all of whom have the same father. For comparison, the "0.0" line represents the average consumption for a typical group of rhesus macaques. Thus, the chart shows that offspring of father T78 consistently consume more alcohol than the average, whereas offspring of father R86 consume less. The figure demonstrates that alcohol consumption may indeed be a product of paternal genetic influences.

early in infancy and continuing into adulthood, suggesting that 5-HIAA levels are influenced by rearing experiences. Serotonin is believed to help control impulses and mood. Adult peer-reared subjects have low CSF 5-HIAA concentrations and exhibit inept social behaviors. Frequently, they must be removed from their social groups for excessive aggression and deviant behavior.

Impaired serotonin function also results from genetic background (Higley et al. 1993), and some subjects show impaired behavior regardless of rearing experience. Within both mother- and peer-reared groups, subjects with low CSF 5-HIAA concentrations exhibit reduced rates of social interaction, low social dominance rankings, and high rates of violence (Higley et al. in press *a*) and alcohol consumption (King et al. 1993). The increased alcohol consumption, inappropriate aggression, and incompetent social behaviors seen in adolescent and adult nonhuman primate subjects with low levels of CSF 5-HIAA parallel findings in human alcoholics with low serotonin functioning (Virkkunen et al. 1994*a,b*).

With some exceptions, these findings generally are consistent with predictions from Cloninger's type II model of excessive alcohol consumption in men, among whom Cloninger has found low CSF 5-HIAA levels (Cloninger 1987; Virkkunen et al. 1994*a,b*).

Because the nonhuman primates that exhibit type II features of alcoholism also have high levels of anxiety, these findings also suggest a partial overlap between type I and type II alcoholism, an intimation that may merit further study in humans.

GENETIC INFLUENCES

Results of the type I and type II studies using nonhuman primates have hinted that genetic influences play a role in rates of alcohol consumption. Emerging evidence from nonhuman primate sibling studies also suggests a genetic influence on alcohol consumption in rhesus macaques. Rates of alcohol consumption were studied in 55 paternal half-sibling adolescent and young adult subjects that were reared apart from their fathers. Although the results are preliminary because of the small sample size, more than 50 percent of the variation between subjects in rates of alcohol consumption appeared to result from paternal genetic contributions, suggesting that genetic factors greatly influence alcohol intake rates. Figure 2 shows the average rates of consumption grouped statistically by father. Higley and colleagues currently are investigating genetic-environmental interactions and maternal genetic contributions to alcohol intake in the rhesus macaques. By selectively breeding subjects for similar neurogenetic traits, such as low CSF 5-HIAA, and rearing them in different highly controlled environments, Higley and colleagues plan to investigate the observed behaviors (i.e., phenotypes²) and underlying genetic traits that lead to excessive alcohol consumption.

EFFICACY AS MODELS

Although nonhuman primates with high rates of alcohol consumption have been shown to exhibit aspects of type I and type II alcoholism, a number of features of alcoholism are yet to be demonstrated adequately in these animals to produce a model for all aspects of alcoholism. For example, in contrast to human data (Cloninger 1987), Higley and colleagues (in press *b*) found no gender differences in rates of alcohol consumption among rhesus macaques. In addition, long-term voluntary consumption at rates that produce physiological withdrawal and prolonged social deficits—such as failure to parent adequately or to maintain social bonds—has not been demonstrated routinely in nonhuman primates.

CONCLUSIONS

A phenotype for high rates of alcohol consumption is present in at least some species of nonhuman primates. Subjects that exhibit this behavior will consume alcohol at rates producing intoxication. Adverse experiences early in life appear to exaggerate genetic predispositions for alcohol consumption. Underlying etiological mechanisms and biobehavioral correlates of the high-consumption phenotype—such as stress, anxiety, antisocial-like behaviors, and impaired serotonin functioning—parallel many of the predictions of Cloninger's neurogenetic model of alcoholism.

These examples from nonhuman primate research demonstrate its value for studying aspects of human alcoholism. Thus, it is surprising that nonhuman primates have been used so infrequently in this field. Their use in alcohol consumption research is a relatively recent development. Fewer than 100 research papers on

this topic have been published during the past 25 years, compared with the more than 380 articles published over the past 3 years alone on alcohol consumption in mice and rats. Because humans and other primates share a large percentage of their genetic material (Stone et al. 1987), studies of nonhuman primates will help clarify the genetic and environmental interactions that contribute to the development of alcoholism in humans. ■

REFERENCES

- CLONINGER, C.R. Neurogenetic adaptive mechanisms in alcoholism. *Science* 236(4800):410–416, 1987.
- ERVIN, F.R.; PALMOUR, R.M.; YOUNG, S.N.; GUZMAN-FLORES, C.; AND JUAREZ, J. Voluntary consumption of beverage alcohol by vervet monkeys: Population screening, descriptive behavior and biochemical measures. *Pharmacology Biochemistry and Behavior* 36(2):367–373, 1990.
- GILLIGAN, S.B.; REICH, T.; AND CLONINGER, C.R. Etiologic heterogeneity in alcoholism. *Genetic Epidemiology* 4(6):395–414, 1987.
- HIGLEY, J.D.; HASERT, M.F.; SUOMI, S.J.; AND LINNOILA, M. Nonhuman primate model of alcohol abuse: Effects of early experience, personality, and stress on alcohol consumption. *Proceedings of the National Academy of Science USA* 88(16):7261–7265, 1991.
- HIGLEY, J.D.; THOMPSON, W.T.; CHAMPOUX, M.; GOLDMAN, D.; HASERT, M.F.; KRAEMER, G.W.; SUOMI, S.J.; AND LINNOILA, M. Paternal and maternal genetic and environmental contributions to CSF monoamine metabolite concentrations in rhesus monkeys (*Macaca Mulatta*). *Archives of General Psychiatry* 50:615–623, 1993.
- HIGLEY, J.D.; HASERT, M.F.; DODSON, A.; SUOMI, S.J.; AND LINNOILA, M. Diminished central nervous system serotonin functioning as a predictor of excessive alcohol consumption: The role of early experiences. *American Journal of Primatology* 33:214, 1994.
- HIGLEY, J.D.; KING, S.T.; HASERT, M.F.; CHAMPOUX, M.; SUOMI, S.J.; AND LINNOILA, M. Stability of interindividual differences in serotonin functions and its relationship to aggressive wounding and competent social behavior in rhesus macaque females. *Neuropsychopharmacology*, in press *a*.
- HIGLEY, J.D.; SUOMI, S.J.; AND LINNOILA, M. Progress toward developing a nonhuman primate model of alcohol abuse and alcoholism. *Social Science and Medicine*, in press *b*.
- KAPLAN, J.N.; HENNESSY, M.B.; AND HOWD, R.A. Oral ethanol intake and levels of blood alcohol in the squirrel monkey. *Pharmacology Biochemistry and Behavior* 17:111–117, 1981.
- KING, S.T.; HIGLEY, J.D.; DODSON, A.; HIGLEY, S.B.; SUOMI, S.J.; AND LINNOILA, M. Alcohol consumption in rhesus monkeys: Effects of genetics, gender, early rearing experience, and social setting. *American Journal of Primatology* 30:323, 1993.
- KRAEMER, G.W., AND MCKINNEY, W.T. Social separation increases alcohol consumption in rhesus monkeys. *Psychopharmacology* 86(1–2):182–189, 1985.
- MCKINNEY, W.T. *Models of Mental Disorders: A New Comparative Psychiatry*. New York: Plenum Medical Book Company, 1988.
- MEHLMAN, P.T.; HIGLEY, J.D.; FAUCHER, I.; LILLY, A.A.; TAUB, D.M.; SUOMI, S.; AND LINNOILA, M. Low CSF 5-HIAA concentrations and severe aggression and impaired impulse control in nonhuman primates. *American Journal of Psychiatry* 151:1485–1491, 1994.
- STONE, W.H.; TREICHEL, R.C.S.; AND VANDEBERG, J.L. Genetic significance of some common primate models in biomedical research. *Progress in Clinical and Biological Research* 229:73–93, 1987.
- VIKKUNEN, M.; KALLIO, E.; RAWLINGS, R.; TOKOLA, R.; POLAND, R.E.; GUIDOTTI, A.; NEMEROFF, C.; BISSETTE, G.; KALOGERAS, K.; KARONEN, S.L.; AND LINNOILA, M. Personality profiles and state aggressiveness in Finnish alcoholic, violent offenders, fire setters, and healthy volunteers. *Archives of General Psychiatry* 51:28–33, 1994a.
- VIKKUNEN, M.; RAWLINGS, R.; TOKOLA, R.; POLAND, R.E.; GUIDOTTI, A.; NEMEROFF, C.; BISSETTE, G.; KALOGERAS, K.; KARONEN, S.L.; AND LINNOILA, M. CSF biochemistries, glucose metabolism, and diurnal activity rhythms in alcoholic, violent offenders, fire setters, and healthy volunteers. *Archives of General Psychiatry* 51:20–27, 1994b.

²For a definition of this and other technical terms used in this article, see central glossary, pp. 182–183.

MOLECULAR BIOLOGY

Alison M. Goate, D.Phil.

Recent advances in molecular biology techniques permit scientists to identify genetic contributions to alcoholism. Two main types of technology are commonly used to identify genes that cause or predispose a person to a disease: positional cloning techniques and candidate gene techniques. Positional cloning techniques allow disease genes to be identified based solely on their location within the genome without prior knowledge of the gene's function. Techniques for confirming the role of candidate genes rely on sufficient prior understanding of the disease process to implicate possible disease-related genes. Scientists use cloning techniques or the application of certain enzymes to reproduce a candidate gene in sufficient quantity for study. As the human genome project progresses and the gene map becomes increasingly complete, more and more disease genes will be identified through a combination of positional cloning and the candidate gene approach. **KEY WORDS:** AOD dependence; gene; research; laboratory method; molecular genetics; genome; genetic mapping; genetic linkage; environmental factors; etiology

Alcoholism is a complex disease caused by a matrix of biological, psychological, and social factors. Although family, twin, and adoption studies have established a genetic contribution to alcoholism, the nature of this contribution is unknown. Researchers are using molecular biology techniques to identify and elucidate the mode of action of genes that predispose people to alcoholism.

GENES AND DISEASE

The genetic "blueprint" that determines the structure and composition of any living organism resides within the DNA molecule. DNA forms the backbone of each of the 23 pairs of chromosomes in the nucleus of every cell in the human body. A DNA molecule comprises long chains of chemical subunits called nucleotides.

Genes are specific sequences of nucleotides that provide the code for particular genetic traits. Each gene directs the synthesis of a different protein from chemical subunits called amino acids. The sequence of nucleotides in a gene determines the order of different amino acids in the finished protein and, hence, the nature of the protein.¹

Some proteins form structural components of cells and tissues. Other proteins (e.g., enzymes) perform vital functions. A normal range of genetic variability contributes to diversity among the human population (e.g., racial characteristics and variation in eye color). Variation at a key site within a gene, however, may render the gene defective. Defective genes produce defective proteins; the resulting structural or functional abnormality forms the basis of genetic disease.

Chromosomes are inherited in pairs, one set of 23 from each parent. Consequently, each cell contains two copies of each gene. Diseases caused by a defect in a single gene are described as Mendelian, after the Austrian monk Gregor Mendel, who formulated

many of the principles of inheritance in the mid-1800's. Mendelian diseases fall into three classes: (1) dominant diseases, in which a single defective copy of one gene, inherited from either parent, is sufficient to cause the disease; (2) recessive diseases, in which both copies of the gene must be defective to cause the disease; and (3) sex-linked diseases, such as hemophilia, which are caused by a defective gene on the X (i.e., the female) chromosome.

Few diseases are inherited in straightforward Mendelian fashion; rather, many common diseases are complex in origin, caused by the effects of several genes interacting with the environment. Such diseases include heart disease, schizophrenia, manic-depressive disorder, and alcoholism. Researchers commonly use two approaches to identify genes that cause or predispose a person to a disease: positional cloning and the candidate gene approach.

POSITIONAL CLONING TECHNOLOGY

Positional cloning comprises a group of techniques that allow disease genes to be identified based solely on their location within the subject's total genetic material (i.e., genome)² without any prior knowledge of the gene's function. Genes so identified can then be copied (i.e., cloned) for study. This approach was used to identify the genes for both Huntington's disease (Group 1993) and cystic fibrosis (Rommens et al. 1989). More recently, positional cloning has been used to identify chromosomal regions predisposing to diabetes (Davies et al. 1994). Positional cloning requires both a collection of families in which more than one member is affected by the disease under investigation and a comprehensive map of polymorphic (i.e., variable) genes to serve as reference points evenly spaced throughout the genome (see discussion of markers below). Researchers obtain DNA from certain cells in the subjects' blood.

A common positional cloning technique uses restriction enzymes derived from bacteria. These enzymes cut DNA at sites marked by specific sequences of four to six nucleotide pairs (i.e., restriction sites), producing DNA fragments of different sizes.³ Because restriction sites are polymorphic, the size distribution of restriction fragments varies from person to person. This variation in fragment size is known as restriction fragment length polymorphism (RFLP).

Restriction fragments can be sorted by size using a procedure called Southern blotting. First, the DNA is applied to one end of a gel and subjected to an electric current, a technique called electrophoresis. The fragments migrate across the gel under the influence of the current, with smaller fragments moving faster. The DNA is then blotted from the gel onto a nylon membrane to facilitate identification of individual fragments. The membrane is soaked in a solution containing radioactively labeled DNA (i.e., a probe) designed to bind to a specific nucleotide sequence. The membrane is exposed to x-ray film to detect the polymorphic fragments, which appear as different banding patterns on the film (figure 1). If the location of a given band is consistently associated (i.e., linked) with a disease, then this pattern can be used as a marker for that disease.

²For technical terms not defined in the text, see central glossary, pp. 182-183.

³Because DNA is double-stranded, the distance along the chromosome is measured in nucleotide pairs. Only one strand contains the genetic code per se. The double-strandedness of the DNA molecule is a separate phenomenon from the existence of paired genes and chromosomes.

¹There are four types of nucleotides; individual amino acids are encoded by specific sequences of three nucleotides.

Another method for detecting polymorphisms is the polymerase chain reaction (PCR). The enzyme DNA polymerase can copy a section of DNA if a short piece of DNA called a primer has been bound to each end of the section. The primer directs the DNA polymerase to produce multiple copies of the section in a chain reaction called amplification. The DNA section, generally amplified more than a million-fold, can be fragmented by a restriction enzyme and separated by electrophoresis. The creation of multiple DNA copies by amplification eliminates the need for a sensitive detection system, such as a radioactive probe.

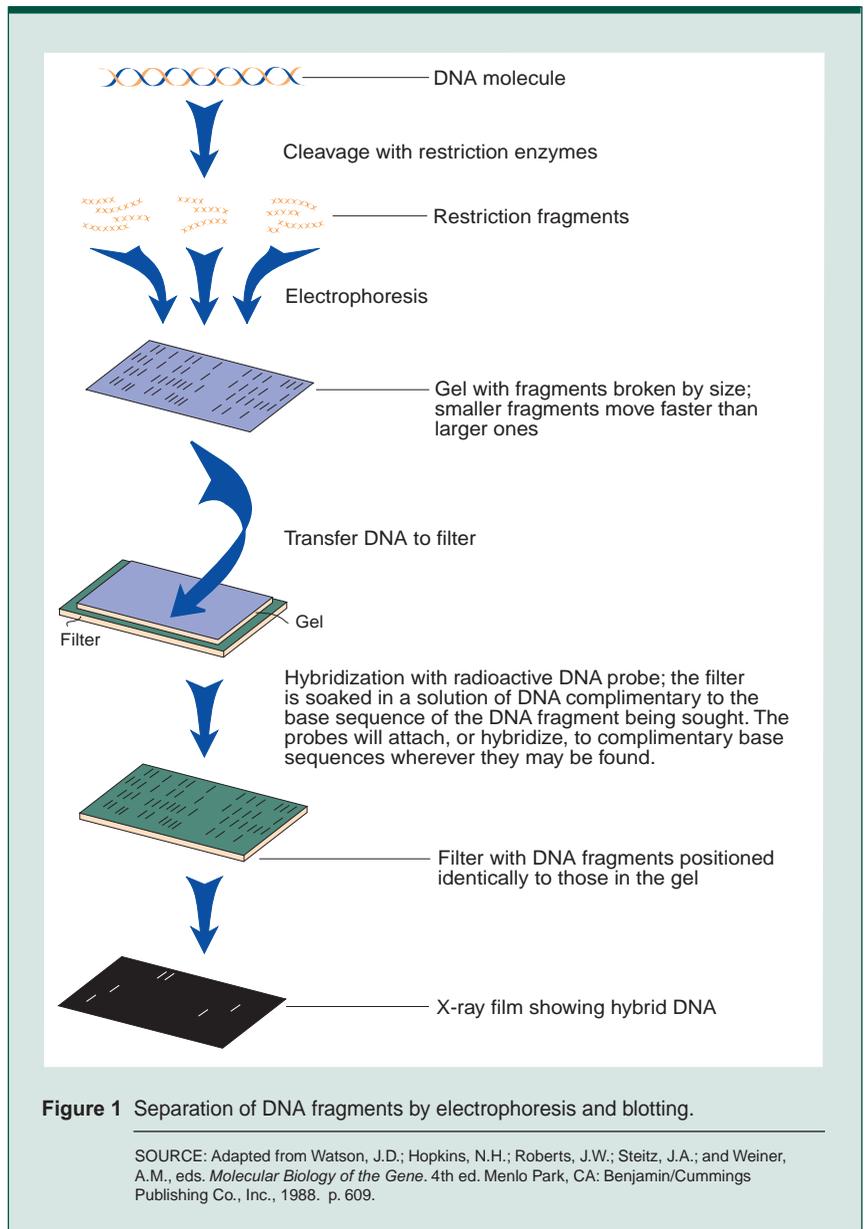
The main drawback with RFLP's is that most restriction sites exhibit only two possible variations—the site is either present or it is not. Consequently, only a small number of RFLP's are genetically informative, permitting an unambiguous determination of which alleles (i.e., variants of a given gene) in the offspring are inherited from each parent.

Today the use of RFLP's has been replaced largely by processes using microsatellite repeat markers. These markers are usually repeating sequences of two, three, or four nucleotide pairs randomly spaced throughout the genome. They are highly polymorphic, having as many as 10 alleles, and can be detected by PCR (Weber and May 1989). The polymorphism reflects variation in the length of the marker and can be detected using PCR amplification, gel separation, and radioactive or fluorescent primers. Radioactively labeled products are detected by exposure to x-ray film (Weber and May 1989), and fluorescently labeled products are detected by laser technology (Davies et al. 1994). Unlabeled PCR products also can be detected by blotting the gel onto a nylon membrane and then probing the filter with a radiolabeled PCR primer (Weissenbach et al. 1992). By using several fluorescent dyes, up to 15 different polymorphisms can be assessed simultaneously.

Using statistical methods, researchers can determine whether a marker and a disease gene are close enough together on a chromosome to be inherited as a unit (a condition referred to as link-

ALISON M. GOATE, D.PHIL., is an associate professor in the Department of Psychiatry, Washington University School of Medicine, St. Louis, Missouri.

The author is the recipient of the National Institutes of Health Career Development Award AG00634-01 and is a participant in the Collaborative Study on the Genetics of Alcoholism.



age). Polymorphic microsatellite markers are usually selected for study such that 300 to 400 markers can cover the entire genome. Maps containing several thousand such markers are now available (Weissenbach et al. 1992). If markers are spaced too widely along the genome, there might not be any marker close enough to the disease gene to exhibit linkage. When a linkage is detected (generally by statistical analysis of inheritance patterns), additional microsatellite polymorphisms from the same region of the chromosome are tested to confirm co-inheritance with the disease and to determine the precise location of the disease gene with respect to the polymorphic markers. A variety of techniques can then be used to identify genes within the region where the putative disease gene (i.e., candidate gene) is located (Group 1993). Each

gene will then be tested as a candidate gene, as described in the following section. If a highly likely candidate gene is already known to occur within the linked region, it can be tested immediately after linkage has been detected.

CANDIDATE GENE APPROACH TECHNOLOGY

Techniques for confirming the role of candidate genes rely on sufficient prior understanding of the disease process to implicate possible disease-related genes. Formerly, scientists used cloning techniques to reproduce a candidate gene in sufficient quantity for study. The cloning process generally involved propagating fragmented human DNA within a bacterium. The gene fragment of interest would then be isolated and its nucleotide sequence determined.

Currently, if the nucleotide sequence of a gene is already known from another species, the gene can often be purified from a batch of total human DNA by selective amplification using DNA primers capable of binding to the specific DNA sequence of the gene.⁴ The technique of amplification has revolutionized molecular biology by speeding up many processes and by reducing the need to clone genes.

Association Techniques

The association approach has been used to identify gene variants that might predispose a person to a disease. Association tests whether a particular polymorphism is observed more frequently in a group of test subjects, who have the disease, than in a group of control subjects, who do not have the disease.

Polymorphisms can be detected using a variety of techniques. For example, the DNA sequence of interest can be amplified by PCR and treated with a restriction enzyme to produce RFLP's. If the polymorphism does not create or destroy a restriction enzyme site, then several alternative approaches can be used. One such method is single-stranded conformation analysis, which compares a DNA sequence from a test subject with the corresponding sequence from a control subject based on the strands' migration through a gel. Differences in the migration of test DNA and the control sequence provide evidence that the nucleotide sequence of the test DNA differs from that of the control DNA.

Two common problems exist with the association approach. First, a spurious result may occur if the control subjects are not properly matched to the disease subjects by ethnic and other factors that reflect genetic composition. Second, association alone does not rule out the possibility that the DNA variant is merely located very close to the disease gene on the chromosome. Biochemical experiments are necessary to demonstrate that a DNA sequence variant actually contributes to the development of the disease.

An association approach recently has been used to identify a polymorphism that may increase a person's risk of developing Alzheimer's disease; the presence of multiple copies of the disease allele increases the risk of developing the disease (i.e., dosage effect) (Corder et al. 1993). Another allele of the same gene, however, appears to decrease the risk for the disease (Corder et al. 1994). Together, these findings support the idea that the polymorphism is associated with the disease, although the precise biological effect of the sequence variants is not known.

⁴This technique relies on the substantial similarity of many genes across species.

A similar approach was used to study alcoholism. One allele of a polymorphism within the D₂ dopamine receptor gene⁵ was reportedly associated with an increased risk of developing alcoholism (Blum et al. 1990). Further research, however, suggests that this result may be spurious (Gelernter et al. 1993).

Linkage

An alternative approach is to study the inheritance pattern of a polymorphism within a candidate gene in families in which more than one member is affected by the disease under investigation. If the candidate gene is located near the disease gene on the same chromosome (or if it is the disease gene), then one allele of the polymorphism will be co-inherited with the disease more often than chance alone would predict within a family.⁶ In these circumstances, the polymorphism and the disease gene are said to be linked (Hsiao et al. 1989; Goate et al. 1991).

Screening for Polymorphisms

If evidence from genetic linkage studies in families or genetic association studies in populations suggests involvement of a candidate gene in causing a disease, then the next step is to sequence the DNA of the gene to determine the variation in the nucleotide sequence that produces a change in the amino acid sequence of the protein (Hsiao et al. 1989; Goate et al. 1991). The most comprehensive way to screen a candidate gene is to sequence the DNA from affected and unaffected subjects. However, if the gene is large, this process may involve considerable work. Techniques like single-stranded conformation analysis can analyze genes in small pieces (i.e., a few hundred nucleotide pairs). Only a small piece of the DNA then need be sequenced to confirm the presence of a DNA sequence variant.

Although these techniques are rapid and enable screening of large numbers of samples, their main drawback is that they may not detect all DNA sequence variants. In addition, if a sequence variant is identified, it could represent either a normal DNA polymorphism or a defective (i.e., disease) allele.

Variation in the DNA sequence between individuals is extremely common but rarely leads to disease: Many stretches of DNA along the chromosome do not code for anything, and the DNA code itself is highly redundant. With four different nucleotides available, 64 possible 3-nucleotide combinations can code for the 20 amino acids found in proteins. Therefore, some polymorphism can occur in a coding sequence without altering the protein.

In a dominant disease, the defective allele will be present in one or both copies of the gene in affected subjects but not at all in normal subjects. In a recessive disease, the defective allele will be present in both copies of the gene in an affected subject and in one copy of the gene in carriers. In complex diseases, the alleles predisposing to disease also may be present in unaffected subjects, making it more difficult to determine whether a particular variant influences risk for disease. However, any sequence variant that influences risk for the disease would still be expected to show association with the disease.

⁵This gene directs the synthesis of a brain protein thought to be involved in certain aspects of the addiction process.

⁶The specific allele co-inherited with the disease may differ between families.

CONCLUSION

With the increasing number of polymorphic markers identified on the genetic map, positional cloning has become an extremely powerful tool, permitting the linkage and identification of a large number of disease genes. As the human genome project progresses and the gene map becomes as dense as the microsatellite polymorphism map, scientists will identify more disease genes through a combination of positional cloning and the candidate gene approach. ■

REFERENCES

BLUM, K.; NOBLE, E.; MONTGOMERY, A.; RITCHIE, T.; JAGADEESWARAN, P.; NOGAMI, H.; BRIGGS, A.; AND COHN, J. Allelic association of human dopamine D₂ receptor gene in alcoholism. *Journal of the American Medical Association* 263:2055–2066, 1990.

CORDER, E.; SAUNDERS, A.; STRITTMATTER, W.; SCHMECHEL, D.; GASKELL, P.; SMALL, G.; ROSES, A.; HAINES, J.; AND PERICAK-VANCE, M. Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. *Science* 261:921–923, 1993.

CORDER, E.; SAUNDERS, A.; RISCH, N.; STRITTMATTER, W.; SCHMECHEL, D.; GASKELL, P.; RIMMLER, J.; LOCKE, P.; CONNEALLY, P.; SCHMADER, K.; SMALL, G.; ROSES, A.; HAINES, J.; AND PERICAK-VANCE, M. Protective effect of apolipoprotein E type 2 allele for late onset Alzheimer disease. *Nature Genetics* 7:180–183, 1994.

DAVIES, J.; KAWAGUCHI, Y.; BENNETT, S.; COPEMAN, J.; CORDELL, H.; PRITCHARD, L.; REED, P.; GOUGH, S.; JENKINS, S.; PALMER, S.; BALFOUR, K.; ROWE, B.; FARRALL, M.; BARNETT, A.; BAIN, S.; AND TODD, J. A genome-wide search for human type 1 diabetes susceptibility genes. *Nature* 371:130–136, 1994.

GELERNTER, J.; GOLDMAN, D.; AND RISCH, N. The A1 allele at the D₂ dopamine receptor gene and alcoholism: A reappraisal. *Journal of the American Medical Association* 269:1673–1677, 1993.

GOATE, A.; CHARTIER HARLIN, M.C.; MULLAN, M.; BROWN, J.; CRAWFORD, F.; FIDANI, L.; GIUFFRA, L.; HAYNES, A.; IRVING, N.; AND JAMES, L. Segregation of a missense mutation in the amyloid precursor protein gene with familial Alzheimer's disease. *Nature* 349:704–706, 1991.

GROUP, H.D.C.R. A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. *Cell* 72:971–983, 1993.

HSIAO, K.; BAKER, H.F.; CROW, T.J.; POULTER, M.; OWEN, F.; TERWILLIGER, J.D.; WESTAWAY, D.; OTT, J.; AND PRUSINER, S.B. Linkage of a prion protein missense variant to Gerstmann-Straussler syndrome. *Nature* 338:342–345, 1989.

ROMMENS, J.; IANNUZZI, M.; KEREM, B.; DRUMM, M.; MELMER, G.; DEAN, M.; ROZMAHEL, R.; COLE, J.; KENNEDY, D.; HIDAKA, N.; ZSIGA, M.; BUCHWALD, M.; RIORDAN, J.; TSUI, L.C.; AND COLLINS, F. Identification of the cystic fibrosis gene: Chromosome walking and jumping. *Science* 245:1059–1065, 1989.

WEBER, J.L., AND MAY, P.E. Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *American Journal of Human Genetics* 44:388–396, 1989.

WEISSENBACH, J.; GYNAPAY, G.; DIB, C.; VIGNAL, A.; MORISSETTE, J.; MILLASSEAU, P.; VAYSSEIX, G.; AND LATHROP, M. A second-generation linkage map of the human genome. *Nature* 359:794–801, 1992.

QUANTITATIVE TRAIT LOCI MAPPING

Judith E. Grisel, Ph.D., and John C. Crabbe, Ph.D.

Researchers interested in the physical locations of genes that influence a person's alcohol-related behaviors can use a method known as quantitative trait loci (QTL) mapping to identify the approximate locations of genes in the genome. QTL mapping can use recombinant inbred mouse strains, which are sets of inbred strains derived from cross-breeding the offspring of two genetically distinct parent strains. The inbred strains exhibit different patterns of the parent strains' genes. QTL mapping involves comparing alcohol-related behaviors in these strains and identifying patterns of known genetic markers shared by strains with the same behaviors. The markers allow the identification of probable locations of genes that influence alcohol-related behaviors. These locations can then be verified using other tests, and specific genes can be sought there. KEY WORDS: AOD use behavior; animal strains; hereditary factors; genetic mapping; genetic markers

Alcohol's complex impact on behavior is demonstrated by studies that attempt to relate problems associated with alcohol abuse, such as dependence on alcohol or tolerance to its effects, to the biological and molecular mechanisms underlying those effects. Alcoholism¹ has been described as a polygenic disorder, that is, one that is influenced by many genes located in different areas, or loci, of a person's or animal's DNA (McClearn et al. 1991; Goldman 1993). The genetically influenced characteristics, or traits, thought to underlie responses to alcohol (e.g., sensitivity to its effects) are called quantitative traits, and many genes influence the overall characteristic, each to a certain extent. Thus, within a population, a quantitative trait differs in the degree to which individuals possess it (e.g., height) rather than in the kind of trait they possess (e.g., eye color). Accordingly, a section of DNA thought to contribute to a quantitative trait is called a quantitative trait locus (QTL). Susceptibility to cancer is another example of quantitative traits that are determined by the combined contributions of several QTLs. Quantitative traits are said to be continuously distributed in a population, because individuals exhibit them to different degrees. Because of this distribution, quantitative traits are much more difficult to study than qualitative traits, such as eye color or blood type, which are determined by a single locus and which therefore are distributed discretely.

One way to study the contributions of individual QTLs to a quantitative trait such as sensitivity to alcohol is for researchers first to locate in the genome (i.e., an organism's entire genetic material) the genetic information that encodes these traits. A technique for finding this information is called QTL mapping. After a QTL has been identified, the gene can be isolated and its functions studied in more detail. Thus, QTL analysis provides a means of locating and measuring the effects of a single QTL on a trait, or phenotype.² This article provides a brief overview of the

¹In this article, the term "alcoholism" refers to the criteria for alcohol abuse and dependence as outlined in the American Psychiatric Association's *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*.

²For a definition of this and other technical terms used in this article, see central glossary, pp. 182–183.

methods involved in QTL mapping. It also discusses the advantages and limitations of QTL analysis and includes several examples demonstrating the application of the technique.

MAPPING THE GENOME

A QTL is a small section of DNA on a chromosome thought to influence a specific trait. Scientists search different areas of the genome for locations (i.e., loci) they can associate with the trait. The gene included in each QTL exists in more than one form, or allele, and can differ between individuals in a population. One person can carry two different alleles of a gene, one inherited from the mother and one from the father. The effect of one QTL is often fairly small. Thus, the collective impact of many genes located at several QTLs provides the genetic influence on different behavioral and physical phenotypes, such as those related to alcohol abuse (figure 1).

For technical reasons, QTL mapping for alcohol-related traits is more commonly performed in animal models than in humans; however, because of the common evolutionary history of all mammals, large regions of our genomes have a common identity. This method is therefore a potentially rich source of information about the genes associated with alcohol abuse in humans (Copeland et al. 1993).

Genetic Markers

To identify the location of QTLs, researchers determine the degree to which a phenotype, such as a person's initial sensitivity to alcohol's effects, is associated with a known genetic marker. These markers are DNA segments known to occupy particular places on the chromosomes.³ Each marker is polymorphic—that is, it exists in several different variants that can distinguish individuals or strains of laboratory animals from one another (for a further discussion of genetic markers, see the article by Anthenelli and Tabakoff, pp. 176–181). The association of a genetic marker with a certain trait allows researchers to estimate both the location of a QTL and the magnitude of its contribution to the trait (Tanksley 1993). For example, if all persons with one specific marker are particularly sensitive to alcohol, then this marker constitutes, or is located close to, a QTL contributing to the trait "sensitivity."

A marker can be a gene, but it also may be derived from a region of the genome that does not produce a functional product, that is, does not underlie the production of proteins. In the latter case, researchers assume that the marker is located near an unmapped gene. Information on these marker locations is helpful in finding new genes. Advances in molecular biology in the past 15

³Chromosomes are composed of strings of nucleotides; a gene is a specific sequence of nucleotides that influences a phenotype.

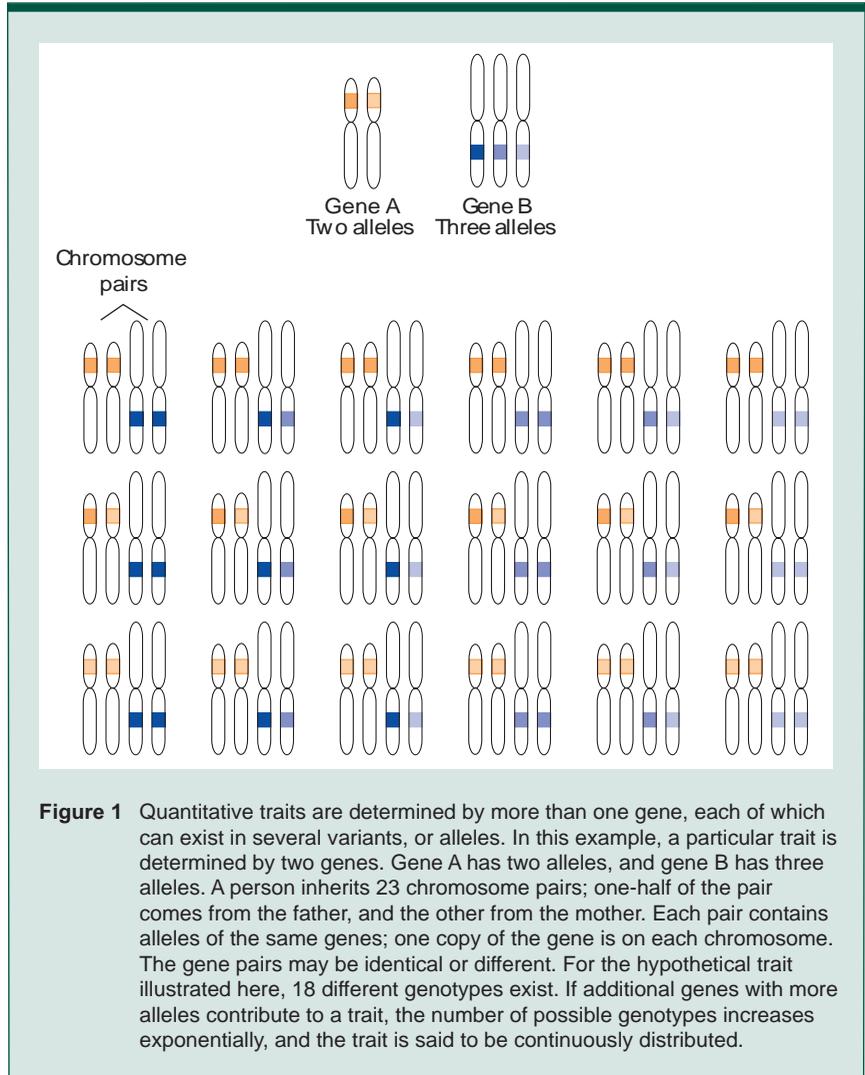


Figure 1 Quantitative traits are determined by more than one gene, each of which can exist in several variants, or alleles. In this example, a particular trait is determined by two genes. Gene A has two alleles, and gene B has three alleles. A person inherits 23 chromosome pairs; one-half of the pair comes from the father, and the other from the mother. Each pair contains alleles of the same genes; one copy of the gene is on each chromosome. The gene pairs may be identical or different. For the hypothetical trait illustrated here, 18 different genotypes exist. If additional genes with more alleles contribute to a trait, the number of possible genotypes increases exponentially, and the trait is said to be continuously distributed.

years have led to the identification of thousands of genetic markers; increasingly dense maps are being constructed for human, mouse, and rat genomes using QTL and related mapping techniques.

BXD RECOMBINANT INBRED STRAINS

Most QTL analyses are performed in laboratory animals, specifically in mice. For their experiments, researchers prefer to com-

JUDITH E. GRISEL, PH.D., is a postdoctoral fellow and JOHN C. CRABBE, PH.D., is both a research career scientist at the Department of Veterans Affairs Medical Center and a professor at the School of Medicine, Oregon Health Sciences University, Portland, Oregon.

This work was supported by a grant from the Department of Veterans Affairs and by National Institute on Alcohol Abuse and Alcoholism (NIAAA) grant AA08621. Dr. Judith E. Grisel is supported by NIAAA grant T32-AA07468.

pare different strains of mice in which all animals of a strain have identical genetic material. One type of animal fulfilling this requirement is the so-called recombinant inbred (RI) mouse strain. RI strains are derived from a population of second generation, or F_2 , offspring of two genetically distinct parental, or progenitor, inbred strains (Bailey 1971). In the RI method, the members of a single strain are inbred to be genetically identical (i.e., they all have the same alleles at all loci). Figure 2 shows, for a single chromosome pair, the process by which RI strains are developed.

Researchers have extensively used one such group of animals, the BXD RI strains, to investigate the genes that influence alcohol-related effects. These RI strains are derived from the F_2 offspring of two progenitor strains called C57BL/6J (C57) and DBA/2J (DBA). Through recombination of the genetic material, each F_2 mouse inherits a distinctive combination of genes from the two progenitor strains. Sister-brother pairs of F_2 mice are mated to begin to “fix” this unique pattern of recombinations (see figure 2). Their progeny are inbred for many more generations, ultimately resulting in an RI strain of mice that are identical (homozygous) at each locus. For any gene, each mouse possesses two copies of the allele from either the C57 or the DBA progenitor.

Several different strains can be generated by inbreeding different sister-brother pairs from the original group of F_2 animals. Such a family, or set, of RI strains allows researchers to compare strains that each have distinct patterns of the same marker set in their genomes.

Currently, 26 BXD RI strains exist, each with a unique genotype. These animals are especially useful for studying the effects of alcohol because their progenitors, the C57 and DBA mice, differ widely with respect to many alcohol-related traits. Other RI sets exist, but only the long-sleep x short-sleep (LS x SS) set, discussed below, also has been used to study alcohol-related phenotypes (Markel et al. in press).

Locating QTLs in RI Strains

QTL mapping correlates the differences in alleles of particular genetic markers with differences in phenotypes in a population. The methodology relies on statistical correlations. In the C57 and DBA mouse strains and in most of the BXD RI strains derived from them, for example, researchers have identified numerous markers. Alcohol-related phenotypes also have been ascertained and recorded for both of the progenitor strains and the battery of BXD RI's. All this information has been collected in a database. By comparing the phenotypes for each strain with the pattern of genetic markers for that strain, researchers can determine if relationships exist between particular phenotypes and markers.

For example, assume that every strain possessing a DBA allele of markers X and Y and a C57 allele of marker Z on the genome shows high sensitivity to alcohol's depressant effects (figure 3). Strains with C57 alleles at X and Y and a DBA allele at Z, in contrast, show low sensitivity to these effects. Thus, the pattern of alleles at these loci predicts alcohol sensitivity. At other loci, whether an animal possesses a DBA or C57 allele may not affect sensitivity; therefore, no significant correlation would exist for these alleles. Using statistical methods, correlation coefficients⁴ are obtained for each phenotype and for each marker that has a different allele for C57 and DBA mice. High correlation coefficients, when statistically significant, indicate that the phenotype may be influenced by the genetic information at this particular locus.

⁴Correlation coefficients provide a measure of the likelihood that animals with a specific phenotype carry a particular marker in their genome.

One key to successful QTL mapping is the identification of a large array of markers that are well distributed over all 20 mouse chromosomes. As mentioned previously, the strength of the BXD RI strains comes primarily from the large number of markers that have been identified for them, more than 1,500 to date. QTL analysis also is being conducted in less densely mapped RI sets, but the statistical procedures in these cases are somewhat more complex (Johnson et al. 1992; Markel et al. in press).

Applying QTL Analysis

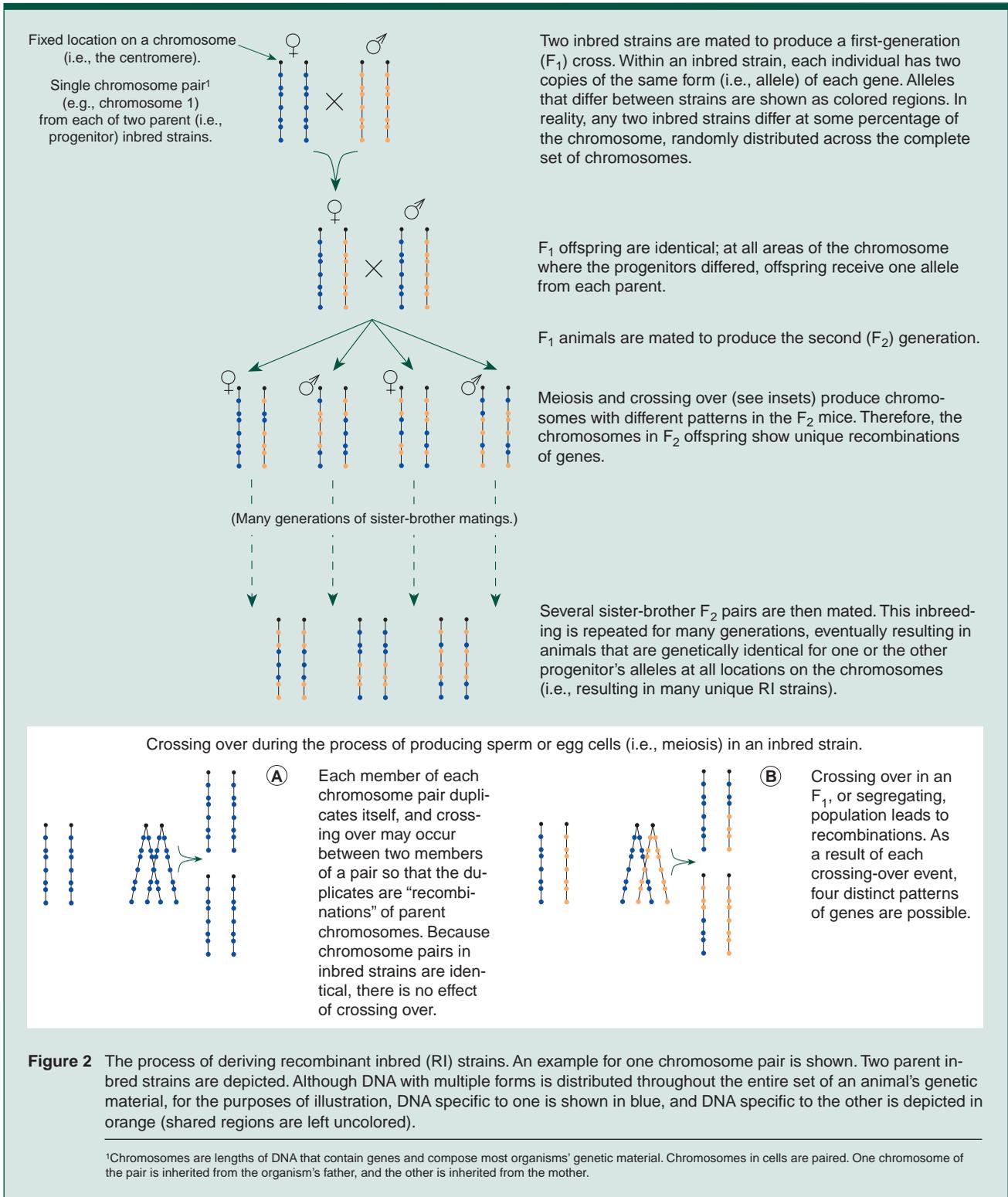
The association of a marker with a certain phenotype does not necessarily mean that the marker *is* the gene contributing to the phenotype. The marker may just be located close enough to the QTL that it is not separated by genetic recombination processes (for more information on genetic recombination and linkage, see sidebar, p. 224). Ultimately, researchers hope to find markers ever closer to the locus that actually produces an effect on a phenotype until a gene that influences the trait eventually is identified. Although the process appears lengthy and tedious, QTL studies begin with no information about the genes in question and rapidly narrow the location to a relatively small piece of DNA. The procedure is like looking for a single family in the entire United States, even narrowing its location to a particular county is extremely helpful to the search effort.

Once QTL analyses have identified a chromosome region containing a gene that may affect a certain phenotype (i.e., a candidate gene), other tests can be performed to determine the magnitude of the gene's influence. For instance, researchers can use pharmacological or molecular methods to alter the protein that the gene produces and then measure the effects of this manipulation on the phenotype (for further discussion of such methods, see the article by Hiller-Sturmhöfel and colleagues, pp. 206–213). The results can then be used to substantiate the gene's effect.

A study by Crabbe and colleagues (1994b) of alcohol-induced hypothermia (i.e., the drop in body temperature that occurs after alcohol administration) provides one example of this kind of QTL analyses. The study found that hypothermic sensitivity to alcohol was associated with the marker *D1Byu7*. This marker is located near a gene called *Ltw-4* on mouse chromosome 1. *Ltw-4* codes for a prevalent protein expressed in brain, liver, and kidney tissues. The *Ltw-4* locus also is associated with the amount of alcohol an animal will drink, amphetamine-induced hyperthermia, and withdrawal from some central nervous system depressants, indicating that a single gene might influence various drug-related phenotypes (Crabbe et al. 1994b). The mechanism by which this protein may influence hypothermic sensitivity and other traits currently is not understood, but now that *Ltw-4* has been identified as a candidate gene, these questions can be addressed in more detailed analyses.

Benefits and Limitations of the RI-QTL Approach

Benefits. The RI-QTL approach has numerous advantages. It can provisionally identify candidate genes for a particular phenotype without any prior knowledge of the biological mechanisms that influence such phenotypes. Simply determining the correlation coefficients between known markers and a quantitative trait in several RI strains readily yields information about the possible location of genes influencing that trait. Furthermore, all information obtained in QTL studies, including data on individual animal strains, is both cumulative and comparable, providing a rich resource and incentive for collaborative efforts. This benefit exists because researchers can obtain the computer-stored



information, which makes the determination of genetic correlations for different traits possible.

Limitations. One unfavorable consequence of analyses in which a large number of correlations are evaluated is that some correlations occur by chance rather than because of a real relationship between a marker and a phenotype. This problem makes further analysis of each QTL necessary (discussed below). Although the statistical tests used to determine correlations can be made more stringent, this strategy may miss some genuine relationships. A more practical approach, therefore, is to use a second population that also has been screened for the proposed QTLs and phenotype to confirm QTLs identified in the first set of animals. Only those associations seen in both analyses are considered to represent QTLs affecting the trait.

One instance of this confirmation approach involves the example of the *Ltw-4* gene mentioned previously. Before *Ltw-4* was implicated as a candidate gene using QTL approaches, Goldman and colleagues (1987) noted an association of the *Ltw-4* locus with alcohol consumption in a set of standard inbred strains. The independent confirmation through QTL analysis of a relationship between this locus and responsiveness to alcohol in RI strains provided further evidence that *Ltw-4* may be important for influencing alcohol-related phenotypes. A recent study by Rodriguez and colleagues (1995), however, found no association between *Ltw-4* and another measure of alcohol consumption in the BXD RI's. Although differences in methodology may account for these discrepancies, this example demonstrates the value of QTL analysis for generating hypotheses as well as the need for subsequent testing of those hypotheses in different animal populations and experimental settings. Several other approaches to confirming QTLs have been discussed in recent reviews (Tanksley 1993; Lander and Schork 1994).

THE ROLE OF LINKAGE IN INHERITANCE

One basic concept underlying quantitative trait loci analysis is that of linkage, or the tendency of genes near each other on a chromosome to be inherited together. When cells undergo meiosis (i.e., the specialized process of cell division that generates sperm and egg cells), genetic material on one chromosomal strand changes places with the corresponding material on its pair strand, a process called crossing over (see figure 2 in text).¹ Genes that are on different chromosomes are inherited indepen-

dently of each other, whereas genes from the same chromosome generally show increasing dependence the closer they are to one another along the length of the chromosome. Consequently, genes near each other are more likely to remain together after crossing over than those farther from each other, because distance increases the chances that the chromosomal break will occur between the two genes in question. For mapping studies, the unit of chromosome length, the centimorgan (cM), is defined as the length of chromosome within whose span a 1-percent likelihood exists of crossing over.

— Judith E. Grisel and John C. Crabbe

¹Crossing over occurs only between chromosome pairs, which contain the DNA encoding the same set of genes. Thus, when pieces of a chromosome change places, each chromosome may receive new forms (i.e., alleles) of certain genes, but neither chromosome loses any genetic information.

Because most RI sets consist of only a limited number of strains, many QTL studies are limited to detecting QTLs with relatively large effects (Belknap 1992). Undoubtedly, important QTLs with smaller effects are missed using this method. In the future, after researchers have identified more of the loci having major effects, the loci with smaller effects will be easier to determine.

Furthermore, QTL analysis assumes that no interactions occur in which some genes alter the expression or function of others; rather, the premise is that each locus affects a phenotype independently of other loci. This limitation can be addressed by producing so-called congenic mice, in which the QTL of interest is bred into a known inbred strain; in this way only a small chromosomal segment containing the potential QTL differentiates the congenic strain from the inbred strain (Bailey 1981). For example, assume

Strain	Marker			Phenotype (sensitivity)	More than 1,500 markers are typed for each strain.
	X	Y	Z		
1	0	1	1	Low	Strain marker assignments: allele from DBA/2J = 0 allele from C57BL/6J = 1
2	1	0	1	Low	
3	1	1	0	High	
4	1	1	1	Low	
5	1	0	1	Low	
⋮					Each strain has a unique distribution pattern of alleles for the identified set of markers.
26	1	1	0	High	

Figure 3 A chart relating markers X, Y, and Z and strain traits (i.e., phenotypes). Strains 3 and 26 have a phenotype of high sensitivity to alcohol's effects. These mice have the form of a gene (i.e., an allele) from the C57BL/6J parent at markers X and Y and an allele from the DBA/2J parent at marker Z. In this manner, markers correlated with the phenotype can be identified.

that results of QTL analysis suggest that 50 percent of the genetic influence on the phenotype arises from the identified locus, whereas tests in congenic mice demonstrate an influence of only 20 percent. Such a disagreement indicates that other genes, in conjunction with the locus in question, are likely to have influenced the phenotype in the QTL experiment.

EXAMPLES OF APPLYING QTL ANALYSIS

Following are a few examples that illustrate the procedures involved in the QTL approach and demonstrate the potential of these analyses.

Hypothermia: Several QTLs Determine One Trait

When Crabbe and colleagues (1994b) performed QTL analysis to identify loci contributing to alcohol-induced hypothermia, they found a correlation with markers other than *D1Byu7* on chromosome 1 (mentioned above). They also discovered that the animals' hypothermic sensitivity correlated with two closely associated markers on chromosome 9. This example underscores the strength of the QTL approach for simultaneously identifying genetic loci in completely different regions of the genome that contribute to the same trait.

The two hypothermia-associated markers on chromosome 9 are located in the same region as the gene for a neurotransmitter receptor, the serotonin 5HT_{1B} receptor (Crabbe et al. 1994b) (for a discussion of different neurotransmitters, see box). To test the possibility that serotonin activity at 5HT_{1B} receptors plays a role in the hypothermic response, knockout mice (i.e., mice whose genes have been experimentally altered) (for further information on knockout mice, see the article by Hiller-Sturmhöfel and colleagues, pp. 206–213) lacking 5HT_{1B} receptors were evaluated for their sensitivity to alcohol. These mice were found to be less sensitive to alcohol-induced hypothermia than normal mice were, thus providing further support for the hypothesis, generated through QTL analysis, that activity at this receptor might play a role in the hypothermic response to alcohol.

The Acrg Locus: One QTL May Affect Several Traits

Not only can several QTLs influence one trait; one QTL also may affect several traits. One gene that demonstrates this capability is the

ALCOHOL-RELATED RESPONSES IN THE CENTRAL NERVOUS SYSTEM

Drinking alcohol produces several immediate effects that vary among people, including changes in behavior, such as losing one's inhibitions or becoming suddenly aggressive. But how do these behavioral changes occur, and why do they differ from person to person? Scientists believe that many individual responses to alcohol originate with genes active in the central nervous system, which controls behavior. Some of these genes are thought to produce substances essential to the brain's ability to relay messages between nerve cells.

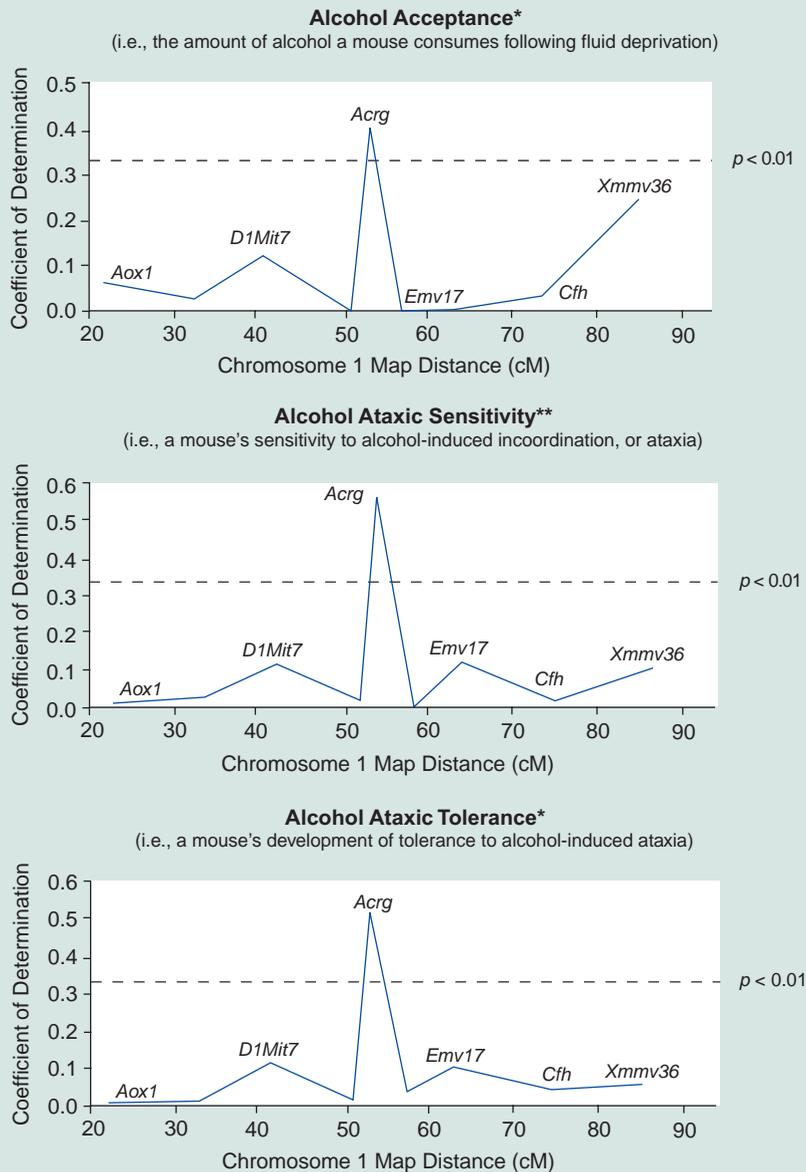
All brain activities depend on cell-to-cell communication. Nerve cells grow long projections during development to connect with each other in a vast and complex network that makes up brain tissue. Signals, the language of the brain, travel along these projections, moving from one cell to the next in the network. Microscopic gaps exist between the ends of many nerve cells' outgoing projections and the beginnings of other cells' incoming projections. These gaps are called synapses, and they must be bridged—using chemicals called neurotransmitters—for a signal to travel between nerve cells. Neurotransmitters are believed to be important in many alcohol-related responses, and different forms (i.e., alleles) of genes that encode neurotransmitter molecules could contribute to individual reactions to alcohol.

A nerve cell sending a signal across a synapse usually releases a neurotransmitter into the synaptic space. Across the synapse, the receiving cell has protein molecules called receptors embedded in its message-receiving surface. The receptors are designed to recognize and bind specific neurotransmitters; several types of receptors can exist for a single neurotransmitter. The binding of neurotransmitter molecules to the appropriate receptors sets off a series of chemical reactions inside the receiving nerve cell that ultimately results in information being processed and transmitted in the brain. Some of the neurotransmitters and their receptors that are mentioned in the articles in this issue have the following general functions:

- *Acetylcholine* is a common neurotransmitter that usually excites receiving cells. Its receptor is composed of several building blocks, or subunits. One is the gamma subunit, which is encoded by the *Acrg* gene. This subunit may play a role in several alcohol-related responses.
- *Catecholamines* are a class of neurotransmitters, one of which is *dopamine*. Dopamine is involved in motor, reward, and cognitive functions; emotions; and aggression. Alcohol's effects on catecholamines may contribute to alcohol tolerance.
- *Gamma-aminobutyric acid (GABA)* and its receptors in the brain inhibit receiving nerve cells. These systems are important for sensory processing and coordination of motor control. GABA may mediate some of alcohol's effects in the brain.
- *Glutamate* is a protein building block (i.e., an amino acid) that also functions as a neurotransmitter and excites receiving nerve cells. Glutamate activity that occurs at the *N*-methyl-D-aspartate (NMDA) receptor may contribute to alcohol-withdrawal seizures.
- *Serotonin* and its receptors affect mood, sleep, drug consumption, the development of tolerance to alcohol and other drugs, higher cognitive functions, and the sensation of pain. Serotonin activity has been connected with alcohol-related responses. For example, one of serotonin's receptors, the 5HT_{1B}, appears to play a role in the drop in body temperature that occurs after alcohol administration.

— Kathryn Ingle

Ltw-4 gene discussed earlier. Another is a QTL near the gene *Acrg*, located on chromosome 1, which codes for a portion of the receptor for the neurotransmitter acetylcholine. Rodriguez and colleagues (1995) reported that this locus is associated with several alcohol-related responses. Their QTL analysis indicated, for example, that



*Rodriguez et al. 1995; **Gallaher et al. in press.

Figure 4 Schematic depiction of mouse chromosome 1. Map distances in centimorgans (cM) from the centromere¹ (at cM = 0) are shown for several markers, whose names are given in italics (e.g., *Aox1*). The Y-axis (i.e., the coefficient of determination) indicates the proportion of trait (i.e., phenotypic) variance between strains accounted for by each marker. The dotted line indicates the proportion of variance required for statistical significance ($p < 0.01$). To have a significant effect on a phenotype, a location on a chromosome (i.e., a locus) must have a coefficient of determination value higher than the significance line. Three phenotypes are shown; all are significantly associated with the marker *Acrg*.

¹A centromere is a fixed location on a chromosome that controls the movement of the chromosome during cell division. A centimorgan is the length of the chromosome within whose span there exists a 1-percent likelihood of crossing over (see figure 2, insets).

Acrg correlates strongly with alcohol acceptance⁵ across the RI strains.

In separate studies investigating the same QTL, Gallaher and colleagues (in press) demonstrated an association between the *Acrg* locus and a different measure of sensitivity to alcohol, alcohol-induced ataxia (i.e., incoordination of voluntary movements). The *Acrg* locus correlated with the initial ataxic sensitivity to alcohol as well as with the development of the tolerance to ataxia that accrues following repeated alcohol injections (figure 4). The results demonstrated that *Acrg* or a nearby gene may contribute between 40 and 60 percent of the genetic influence on these phenotypes.

Given the concordance in results across two different studies and three phenotypes—alcohol acceptance, initial ataxic sensitivity, and tolerance to ataxia—*Acrg* is likely to be an important QTL. As always, however, this finding should be considered preliminary until a relationship between *Acrg* and these behaviors can be confirmed using other approaches. For example, it is currently unknown whether the part of the acetylcholine receptor encoded by *Acrg* mediates these alcohol effects or whether a nearby gene influences these characteristics. Nevertheless, the combined QTL studies have led to a specific hypothesis, and the *Acrg* locus now can be further evaluated for its influence on alcohol drinking and incoordination.

Withdrawal: Followup Experiments Confirm the Significance of a QTL

Using QTL methods, researchers have identified several loci that may influence the severity of alcohol withdrawal (Belknap et al. 1993; Crabbe et al. 1994a). One of these is a site near the *b* locus on mouse chromosome 4. To confirm the hypothesis that this QTL contributes to the severity of withdrawal, individual F₂ offspring derived from C57 and DBA parents were tested for acute alcohol withdrawal and the presence of genetic markers in this region. This followup study determined a significant association between alcohol withdrawal and the *b* locus. This result, obtained in a statistically independent population of animals, confirms that a QTL near the *b* locus accounts for about 40 percent of the genetic contribution to acute alcohol withdrawal (Crabbe et al. 1994a).

⁵Alcohol acceptance was measured by the amount of 10-percent alcohol in water a mouse consumed in the 24-hour period following 24 hours of fluid deprivation.

LS x SS RI Strains: QTL Analyses in Animals Bred for Alcohol-Related Characteristics

The principles of QTL analysis also can be applied to the offspring of other genetic crosses, such as those derived from lines that have been bred selectively for their sensitivity to specific alcohol effects. Such analyses may enable researchers to identify more easily the QTLs related to the particular sensitivity. For example, RI's were developed from the LS and SS mice, which were bred to differ with respect to alcohol-induced loss of righting reflex, that is, hypnotic sensitivity to alcohol.⁶ Erwin and Jones (1993) have examined 26 LS x SS RI strains for potential genetic influences on their hypnotic, hypothermic, and locomotor responses to alcohol. The researchers' QTL analysis results suggest influences of several genes on these behaviors, with some of the genes appearing to influence more than one behavior.

In a similar project, Johnson and colleagues at the University of Colorado have begun to identify QTLs underlying the differential sensitivity to alcohol in the LS x SS RI's. They recently found several potential QTLs that influence initial sensitivity to alcohol (personal communication; Markel et al. in press) and that currently are being evaluated in inbred LS and SS F₂ mice.

SUMMARY

The use of QTL analysis has greatly extended current understanding of the particular genes involved in the expression of alcohol-related traits. Through the use of these statistical and molecular biological tools, made practical by an expanding library of known genetic markers, the complex genetic substrates of alcoholism can be examined even in the absence of prior knowledge or specific hypotheses. The suggested QTLs can then be evaluated further for their possible role in the expression of these traits. ■

ACKNOWLEDGMENT

The authors would like to thank J.K. Belknap for providing suggestions on the manuscript.

REFERENCES

BAILEY, D.W. Recombinant inbred strains. *Transplantation* 11(3):325-327, 1971.

BAILEY, D.W. Recombinant inbred strains and bilineal congenic strains. In: Foster, H.L.; Small, J.D.; and Fox, J.G., eds. *The Mouse in Biomedical Research*. Vol. I. New York: Academic Press, 1981. pp. 223-239.

⁶Following an injection of alcohol, mice lose their righting reflex (i.e., the reflex to point their bellies toward the ground and their backs up). This reflex is regained at different time periods after injection and is used as a measure of hypnotic sensitivity.

BELKNAP, J.K. Empirical estimates of Bonferroni corrections for use in chromosome mapping studies with the BXD recombinant inbred strains. *Behavior Genetics* 22(6):677-684, 1992.

BELKNAP, J.K.; METTEN, P.; HELMS, M.L.; O'TOOLE, L.A.; ANGELI-GADE, S.; CRABBE, J.C.; AND PHILLIPS, T.J. Quantitative trait loci (QTL) applications to substances of abuse: Physical dependence studies with nitrous oxide and alcohol in BXD mice. *Behavior Genetics* 23(2):213-222, 1993.

COPELAND, N.G.; JENKINS, N.A.; GILBERT, D.J.; EPPIG, J.T.; MALTAIS, L.J.; MILLER, J.C.; DIETRICH, W.F.; WEAVER, A.; LINCOLN, S.E.; STEEN, R.G.; STEIN, L.D.; NADEAU, J.H.; AND LANDER, E.S. A genetic linkage map of the mouse: Current applications and future prospects. *Science* 262:57-66, 1993.

CRABBE, J.C.; BELKNAP, J.K.; AND BUCK, K.J. Genetic animal models of alcohol and drug abuse. *Science* 264(5166):1715-1723, 1994a.

CRABBE, J.C.; BELKNAP, J.K.; MITCHELL, S.R.; AND CRAWSHAW, L.I. Quantitative trait loci mapping of genes that influence the sensitivity and tolerance to ethanol-induced hypothermia in BXD recombinant inbred mice. *Journal of Pharmacology and Experimental Therapeutics* 269(1):184-192, 1994b.

ERWIN, V.G., AND JONES, B.C. Genetic correlations among ethanol-related behaviors and neurotensin receptors in Long Sleep (LS) x Short Sleep (SS) recombinant inbred strains of mice. *Behavior Genetics* 23(2):191-196, 1993.

GALLAHER, E.J.; JONES, G.E.; BELKNAP, J.K.; AND CRABBE, J.C. Identification of genetic markers for initial sensitivity and rapid tolerance to ethanol-induced ataxia using quantitative trait loci 7 analysis in BXD recombinant inbred mice. *Journal of Pharmacology and Experimental Therapeutics*, in press.

GOLDMAN, D. Genetic transmission. In: Galanter, M., ed. *Recent Developments in Alcoholism*. Vol. 11. New York: Plenum Press, 1993. pp. 231-248.

GOLDMAN, D.; LISTER, R.G.; AND CRABBE, J.C. Mapping of a putative genetic locus determining ethanol intake in the mouse. *Brain Research* 420:220-226, 1987.

JOHNSON, T.E.; DEFRIES, J.C.; AND MARKEL, P.D. Mapping quantitative trait loci for behavioral traits in the mouse. *Behavior Genetics* 22(6):635-653, 1992.

LANDER, E.S., AND SCHORK, N.J. Genetic dissection of complex traits. *Science* 265:2037-2048, 1994.

MCCLEARN, G.E.; PLOMIN, R.; GORA-MASLAK, G.; AND CRABBE, J.C. The gene chase in behavioral science. *Psychological Science* 2(4):222-229, 1991.

MARKEL, P.D.; FULKER, D.W.; BENNETT, B.; CORLEY, R.P.; DEFRIES, J.C.; ERWIN, V.G.; AND JOHNSON, T.E. Quantitative trait loci for ethanol sensitivity in the LSXSS recombinant inbred strains: Interval mapping. *Behavior Genetics*, in press.

RODRIGUEZ, L.A.; PLOMIN, R.; BLIZARD, D.A.; JONES, B.C.; AND MCCLEARN, G.E. Alcohol acceptance, preference, and sensitivity in Mice. II. Quantitative trait loci mapping analysis using BXD recombinant inbred strains. *Alcoholism: Clinical and Experimental Research* 19(2):367-373, 1995.

TANKSLEY, S.D. Mapping polygenes. *Annual Review of Genetics* 27:205-233, 1993.