Public Health Genomics at CDC
Accomplishments & Priorities 2004

Introduction

New developments in human genomics present challenges and opportunities for CDC. Public health research is underway throughout the agency in areas including newborn screening, genetic testing, pharmaco- and toxicogenomics, analysis of gene-environment interaction in susceptibility to infectious and chronic diseases, and use of family history in public health programs.

In order to document CDC's priorities, accomplishments, and future directions in human genomics, OGDP developed a summary briefing book in January 2005 that highlights each of CDC's scientific components. Although similar in overall style to conventional program review books, it is intended to provide a general overview with selected highlights, rather than a comprehensive inventory. The briefing book is organized by coordinating centers and contains a brief overview from each CIO.

Office of Genomics and Disease Prevention

The Office of Genomics and Disease Prevention (OGDP) was established in 1997 and has since provided national leadership while building partnerships with other federal agencies, public health organizations, professional groups, and the private sector. The mission of OGDP is to integrate genomics into public health research, policy, and programs and to improve population health and prevent disease through the application of genomic information

Top priorities

- Integrating genomics into public health research.
- Assessing the value of family history and genomic tests for population health.
- Incorporating genomics into public health practice.

Major Accomplishments

Integrating Genomics into Public Health Research

Acute Public Health Investigations
OGDP formed a multidisciplinary Acute Public Health Investigation (APHI) Working Group in collaboration with the Council of State and Territorial Epidemiologists (CSTE) to develop a plan and tools for incorporating genomics into APHIs. A workshop convened on May 12-13, 2004 addressed "*The Role of Human Genomics in Acute Public Health Investigations: Current Practice and Future Strategies."* Workshop participants reviewed the rationale, state of the science, challenges and opportunities for incorporating human genomics into acute public health investigations. They outlined next steps for prioritizing investigations and developing the necessary capacity at CDC to incorporate human genomics. Input from the workshop will inform a new initiative to develop criteria for incorporating human genomics into APHIs and develop standard tools and protocols for field and laboratory workers.

NHANES Working Group
In collaboration with the National Institutes of Health (NIH), a CDC-wide team is measuring population variation in selected genes using stored DNA samples collected during the third National Health and Nutrition Examination Survey (NHANES) III. The goal of this collaboration is to develop genotype prevalence estimates based on a nationally representative sample of the U.S. population. These data will add another dimension to the analysis of clinical, physical, and lifestyle information collected by NHANES, creating a resource for analysis of genotype-phenotype correlations and gene-environment interactions.

HuGENet

OGDP established the Human Genome Epidemiology Network (HuGeNet™, http://www.cdc.gov/genomics/hugenet) as a global collaboration of individuals and organizations committed to assessing the role of human genome variation in population health and the potential of genomics for improving health and preventing disease. HuGENet promotes publication of systematic reviews of population-based data on genotype prevalence, gene-disease associations, and gene-environment interactions. Accomplishments in 2004 include publication of a book, *Human Genome Epidemiology*, and co-sponsorship of a workshop on meta-analysis by the Public Health Genetics Unit in Cambridge, UK (http://www.cgkp.org.uk/work/activities.html#syst_rev).

Assessing the value of family history and genomic tests for population health

Family History Public Health Initiative

The *Family History Public Health Initiative* (http://www.cdc.gov/genomics/info/factshts/famhist.htm), launched in 2002, has produced a new family history tool, Family Healthware™, which is a web-based and self-administered. It collects information on family health history and personal risk factors. It calculates familial risk using algorithms built into the software and generates a report that includes a family health pedigree and tailored prevention messages and screening recommendations. A resource manual being developed for healthcare providers will include an explanation of familial risk levels and possible genetic conditions underlying high risk, recommendations for further risk assessment, and suggested interventions and resources. In 2005, the clinical utility of Family Healthware™ will be evaluated in a randomized controlled trial conducted by three academic medical centers.

On November 8th, 2004 the Surgeon General held a national press conference to encourage Americans to make Thanksgiving Day the first annual National Family History Day and to direct them to the Surgeon General's web site, which includes *My Family Health Portrait,* an abbreviated version of the CDC family history tool. OGDP has also launched a new family history website for the general public that includes fact sheets, case studies, news articles, and presentations.

Evaluation of Genetic Testing

In fall 2004, OGDP launched *Evaluation of Genomic Applications in Practice and Prevention* (*EGAPP),*a new project whose goal is to develop and evaluate a coordinated process for systematic assessment of genetic tests in transition from research to clinical and public health practice in the U.S.. A key element for achieving success in this project will be the development and maintenance of partnerships and collaborations with stakeholders (e.g., US Preventive Services Task Force, CDC Community Guide), other HHS agencies, the international health technology assessment community, and other relevant projects and advisory groups.

On September 27-28th, 2004, OGDP held a short course, *Public Health Assessment of Genetic Tests for Screening and Prevention,* on systematic approaches to evidence-based assessment of genetic tests. Participants included public health and health care professionals, policymakers and other who are involved in the development, evaluation, utilization and reimbursement of genetic tests for screening and prevention.

Incorporating genomics into public health practice

Training Public Health Professionals in Genomics

The Michigan Center for Genomics and Public Health, in collaboration with Genomics Centers at the University of North Carolina and University of Washington, OGDP, and the Chronic Disease Directors, completed a web-based training program for public health professionals called *Six Weeks to Genomic Awareness*. The program is available as a series of six presentations designed to help public health professionals understand how genomic advances are relevant to public health. The series introduces the user to genomic concepts such as the human genome and heredity, genetic variation in populations, genetic testing, and ethical, legal, and social issues from a public health perspective. It is capped off by an overview of state and national resources for public health professionals.

OGDP and the three Centers for Genomics and Public Health also created *Genomics for Public Health Practitioners,* a 45-minute introductory presentation for public health practitioners. It describes the application of genomics to public health, dispels some myths and identifies challenges.

Future Directions

1. Develop a CDC-wide plan to enhance capacity in public health genomics research, evidence-based assessment and integration of genomics into practice.

2. Develop a public-private initiative to use NHANES DNA bank for intramural and extramural public health genomics research.

3. Launch the Evaluation of Genomic Applications in Practice and Prevention (EGAPP) initiative.

4. Develop and implement guidance for the systematic evaluation of gene-disease associations.

5. Develop guidance for the systematic design and reporting of results from genomic cohort ("biobank") studies.

6. Conduct a national workshop on methods for applying family history tools in preventive medicine and public health.

Publications

Khoury MJ, Millikan R, Little J, Gwinn M. The emergence of epidemiology in the genomics age. Int J Epidemiol 2004;33:936-44.

McCusker ME, Yoon PW, Gwinn M, Malarcher AM, Neff L, Khoury MJ. Family history of heart disease and cardiovascular disease risk-reducing behaviors. Genet Med 2004;6:153-8.

Khoury MJ, Yang Q, Gwinn M, Little J, Dana Flanders W. An epidemiologic assessment of genomic profiling for measuring susceptibility to common diseases and targeting interventions. Genet Med 2004;6:38-47.

Scheuner MT, Yoon PW, Khoury MJ. Contribution of Mendelian Disorders to Common Chronic Disease: Opportunities for Recognition, Intervention, and Prevention. *Am J Med Genet* 2004;125C:50-65.

Lindegren ML, et al. Applying public health strategies to primary immunodeficiency disorders: a model approach to genetic disorders. MMWR 2004 (RR01);53:1-29.

Khoury MJ. The case for a global human genome epidemiology initiative. Nat Genet 2004;36:1027-1028.

Yoon, PW, Scheuner MT, Gwinn M, Khoury MJ, Jorgensen C, Hariri S, Lyn S. Awareness of family health history as a risk factor for disease, United States, 2004. MMWR 2004; 53(44):1044-1047

Myers, M, Jorgensen CJ, et al. Genetic Testing for Breast and Ovarian Cancer Susceptibility: Evaluating Direct-to-Consumer Marketing—Atlanta, Denver, Raleigh-Durham, and Seattle, 2003. MMWR July;53(27):603-606.

Human Epidemiology: A Scientific Foundation for Using Genetic Information to Improve Health and Prevent Disease was published by Oxford University Press in 2004. It provides a scientific foundation to help researchers, policy makers, and practitioners integrate genomics into medical and public health practice.

Publication of the first annual OGDP report titled, "*Genomics and Population Health: United States 2003*." The report includes a timely and practical collection of vignettes of the status of genomics and population health in the United States, and is intended for public health professionals who are interested in integrating genomics into health promotion, disease prevention and healthcare.

Publications from the ACCE Project included *Clinical sensitivity of prenatal screening for cystic fibrosis via CFTR carrier testing in a United States panethnic population* (Palomaki et al., Genetics in Medicine 6:405-414), and *Prenatal screening for cystic fibrosis: An early report card (editorial)* (Palomaki, Genetics in Medicine 6:115). The American College of Medical Genetics utilized data from the ACCE review on cystic fibrosis carrier testing to revise recommendations for the mutation panel (Watson et al, *Cystic fibrosis population carrier screening: 2004 revision of American College of Medical Genetics mutation panel.* Genetics in Medicine 6:387-91). Papers in press at the end of 2004 were *An evaluation of BRCA1 and BRCA2 founder mutations penetrance estimates for breast cancer among Ashkenazi Jewish women* and *Adjusting the estimated proportion of breast cancer cases associated with BRCA1 and BRCA2 mutations: Public health implications* (McClain et al., Genetics in Medicine).

Chronic Disease Prevention and Genomics
2004 Activity Snapshot

The National Center for Chronic Disease Prevention and Health Promotion (NCCDPHP) continues to look at the important links between chronic disease and genetics. Augmenting the work of the Office of Genomics and Disease Prevention, NCCDPHP is growing the science around some of the leading causes of death – diseases such as heart disease, cancer, and diabetes. Today, we know that many chronic diseases are preventable and will become more fully understood only when both the genetic and environmental contributions to their etiology are known. CDC is committed to advancing the science to improve health and prevent and control chronic diseases. The following document highlights CDC supported chronic disease/genomics activities, including epidemiologic studies, planning and policy activities, and dissemination and education efforts.

2004 Major Accomplishments

Exploring the Links between Chronic Diseases and Family History

One of the primary activities, a collaborative effort of OGDP and NCCDPHP, is the Family History Public Health Initiative, which began in 2002 with an extensive literature review and a subsequent paper that examined the idea of using family history for disease prevention. NCCDPHP participates in OGDP's Family History Work Group, which was formed to conduct research and develop a family history tool for disease prevention. The multidisciplinary group has established criteria for diseases to include in a family history tool, reviewed the literature for nearly 40 diseases, and developed a prototype tool with 16 diseases. Ongoing work includes pilot studies to further refine the tool, development of algorithms to assess risk, development of a resource manual for primary care providers, and design and funding of studies to evaluate the validity and utility of the approach.

Family History of Cancer and Cancer Screening

Little data are available on the use of screening tests among individuals with a family history of cancer. Using data from the 2000 National Health Interview Survey (NHIS), a logistic regression model is being used to examine the association between first-degree family history of cancer and receipt of a recent screening test (mammography, prostate specific antigen, and colonoscopy or flexible sigmoidoscopy) among persons who had at least one screening test, adjusting for age, race, gender, education, income, marital status, and health insurance coverage.

Genetic Testing for Breast and Ovarian Cancer Susceptibility: Evaluating Direct-to-Consumer Marketing

Breast and ovarian cancer are the second and fifth leading causes of cancer death, respectively, among women in the United States. Mutations in two genes, BRCA1 and BRCA2 (BRCA1/2), are associated with predisposition for inherited breast and ovarian cancer and are identified in 5% to 10% of women with breast or ovarian cancer (BOC). Since 1996, genetic testing for these mutations has been available clinically; however, population-based screening is not recommended because of the complexity of test interpretation and limited data on clinical validity and utility. Despite the test's limited applicability in the general population, the U.S. provider of clinical BRCA1/2 testing (Myriad Genetic Laboratories, Inc.,

Salt Lake City, Utah) conducted a pilot direct-to-consumer (DTC) marketing campaign in two cities (Atlanta, Georgia, and Denver, Colorado) during September 2002–February 2003. Although DTC advertisements have been used to raise consumer awareness about pharmaceuticals, this was the first time an established genetic test was marketed to the public. To assess the impact of the campaign on consumer behaviors and healthcare provider practices, CDC and the respective state health departments for the pilot cities and two comparison cities (Raleigh-Durham, North Carolina, and Seattle, Washington) surveyed consumers and providers. This report summarizes results of those surveys, which indicated that consumer and provider awareness of BRCA1/2 testing increased in the pilot cities and that providers in these cities perceived an impact on their practice (e.g., more questions asked about testing, more BRCA1/2 tests requested, and more tests ordered). However, in all four cities, providers often lacked knowledge to talk to patients about inherited BOC and testing. These findings underscore the need for evidence-based recommendations on appropriate use of genetic tests and education of providers and the public to achieve maximum individual and public health benefit from genetic testing. This project is a collaboration between NCCDPHP and OGDP.

Genomics and Chronic Disease: What Project Officers Need to Know
CDC, in conjunction with the Universities of Michigan, North Carolina, and Washington, presented "Genomics and Chronic Disease: What Project Officers Need to Know" in June 2004. This workshop was designed to provide an introduction to genomics to NCCDPHP project officers and help them better understand how existing chronic disease programs – particularly in the areas of diabetes and cancer – have incorporated genomics into their existing programs.

Hereditary Hemochromatosis
CDC and its partners continue to support an Internet-based teaching module on Hereditary Hemochromatosis for health professionals and the public, including the development and dissemination of education materials, which include the patient brochure *Iron Overload and Hemachromatosis: Information for Patients and Family.*

Johnston County Osteoarthritis Project
This ongoing community-based cohort study of rural white and black persons to determine the prevalence, incidence, and factors associated with the occurrence or progression of hip and knee osteoarthritis includes a genomics component for examining genes (HH, COMP, COL2A, others to be determined) that may be linked to osteoarthritis and related conditions. The study is conducted by the University of North Carolina with CDC funding and collaboration.

Oregon Sudden Unexplained Death Study
Despite current medical knowledge, sudden cardiac death – or the sudden failure of heart rhythm – is unexplained in up to 15 percent of cases. The Oregon Sudden Unexplained Death Study (SUDS) Group, funded by CDC, is conducting research on the mechanisms of sudden cardiac death. This prospective study of all patients suffering sudden cardiac death in Multnomah County, Oregon, will develop a comprehensive database describing patients' medical histories, including clinical findings, pathologic data, and genetic analysis.

State Capacity Grants for Integrating Genomics into Chronic Disease Prevention Programs
State and community health agencies recognize the need to augment existing genetics expertise in maternal and child health and newborn screening capacity to integrate genomics into a wider range of disease control and prevention programs. To develop this capacity, NCCDPHP established cooperative agreements with state health departments in Michigan, Minnesota, Oregon, and Utah to strengthen

leadership and promote coordination in programs for genomics and chronic disease prevention. Through these agreements, those states are working on a variety of projects over a five year period from 2003 to 2008. These projects include integrating genomics and family history into ongoing and new population-based strategies for identifying and reducing the burden of specific chronic, infectious, and other diseases; enhancing planning and coordination for integrating genomics into core state public health specialties (such as epidemiology, laboratory activities, and environmental health); and facilitating the use of family history and new knowledge about gene-environment interactions to enhance chronic disease prevention. This project is a collaboration between NCCDPHP and OGDP.

Stroke Prevention in Young Women
A population-based case control study examining gene-environment interactions in the risk of stroke, the study will determine the prevalence of genetic polymorphisms in the population, gene-environment interactions in stroke risk. The study also will determine the population-attributable fraction associated with genetic risk factors and gene-environment interaction. The study is conducted in collaboration with University of Maryland.

Future Directions: Opportunities Ahead for Genomics and Chronic Disease Prevention

In 2005 and beyond, NCCDPHP will continue its work in genomics in the following priorities areas:

- Foster efforts to use family history to promote health – particularly through support and expansion of CDC's family history project.
- Support state based grants to build capacity in states to understand the links and importance of family history in chronic disease prevention and health promotion.
- Promote communication and policy efforts to foster the public's understanding of genomics and the role in health promotion.

Public Health Genomics at the National Center on Birth Defects and Developmental Disabilities

Top Priorities

The National Center on Birth Defects and Developmental Disabilities (NCBDDD) has four main priorities related to genetics activities:

- Investigate susceptibility genes for birth defects
- Support, develop, and evaluate newborn screening programs for genetic disorders
- Improve quality of life of children with potentially disabling single gene disorders and their families
- Prevent morbidity and mortality related to genetic disorders of hemostasis by:
    - Investigating susceptibility genes for thrombotic diseases
    - Characterizing the types of health care services needed to care for people with thrombotic disease
    - Investigating the role of genetic mutations on development of inhibitors in hemophilia

Birth Defects Research
Birth defects etiologic research is a major and well-established NCBDDD activity. Numerous investigations are ongoing, but the most genomic-related project is the National Birth Defects Prevention Study (NBDPS). In 1996, Congress directed CDC to establish the Centers for Birth Defects Research and Prevention. This directive was formalized with the passage of the Birth Defects Prevention Act of 1998. Currently, NCBDDD has established centers in Arkansas, California, Iowa, Massachusetts, New York, North Carolina, Texas, and Utah. The centers were established in states with existing birth defect programs that had nationally recognized expertise in birth defects surveillance and research. Each of these centers is a site for the NBDPS. CDC coordinates the centers and participates in the NBDPS as the ninth study site.

The NBDPS has three components: review and classification of records of infants with birth defects by clinical geneticists, maternal interviews of eligible participants, and collection and processing of cheek swabs from study participants and their parents. Genomic DNA from buccal cells is stored both at the individual centers and at a central repository at CDC. The information gathered from the interviews, combined with the availability of DNA, provides an invaluable resource for the study of genetic susceptibility to environmental exposures for a broad range of carefully classified birth defects. The unprecedented statistical power from this collaborative study has enabled investigators to study the epidemiology and genetics of both common and rare birth defects, and the compiled data and banked DNA will facilitate future research as new hypotheses and improved technologies emerge.

Newborn Screening for Genetic Conditions
The first newborn screening program was implemented in the 1960s to provide early diagnosis of phenylketonuria (PKU). This early diagnosis allows dietary management as a means to prevent the development of mental retardation in children with PKU. While all states now screen universally for three disorders (PKU, galactosemia, and congenital hypothyroidism), considerable variation still exists in testing for other conditions. States also differ in specific laboratory techniques, storage and use of residual blood spots, and follow-up and integration of screening practices. The newborn screening system involves more than laboratory analysis, and recently a multiagency Newborn Screening Task Force has called for recognition of the

newborn screening system as a comprehensive public health program that includes screening, short-term and long-term follow-up, diagnosis, management, and evaluation. New technologies introduced in recent years have also led to interest in expert guidance and new funding for newborn screening.

Children with Potentially Disabling Single Gene Disorders
Pediatric single gene disorders account for about 13% of inpatients in pediatric hospitals, and for about 3% to 5% of pediatric deaths. Individually, single gene disorders are rare and often get "lost" in the overall picture. In addition to health and disability concerns, genetic conditions raise additional psychosocial and familial issues. Families, health care providers, and public health officials need to become familiar with the unique challenges raised by genetic conditions. Lessons learned from public health activities in single gene disorders can be applied to complex disorders as they become elucidated. The goals of NCBDDD are to:

- Develop surveillance systems that meet challenges of single gene disorders and that are expandable. Data from surveillance systems are necessary to determine the following:
    - How common is the condition?
    - Is it equally common in different racial and ethnic groups?
    - What is the natural history of the condition?
    - What factors affect disease progression?
    - Does the type of care received affect the progression?
    - Do different populations receive different care?
    - What services are families receiving?
    - What barriers to services are families facing?
- Improve screening and diagnosis.
- Improve services to patients and families.

Duchenne/Becker Muscular Dystrophy (DBMD): DBMD is the first disorder to be investigated using this approach. Duchenne muscular dystrophy (DMD) affects about 1 in 4,000 males and is the most common form of muscular dystrophy in children. In the absence of newborn screening, DMD is usually diagnosed when a boy is 3 to 6 years of age. Early signs include failure to walk by 18 months of age, frequent falling, difficulty getting up from a sitting or lying position, and a waddling gait. As muscle deterioration progresses, children with DMD become unable to walk around age 12 years. The disease is fatal in the teens or early 20s, because of severe respiratory and heart problems. A milder form of the disease, Becker muscular dystrophy, is caused by mutations in the same gene. The combined spectrum is referred to as Duchenne/Becker muscular dystrophy (DBMD). Standard birth defects monitoring systems in the United States do not detect children with DBMD because these children do not have recognizable signs or symptoms at birth. Consequently, existing birth defects monitoring systems would need to be supplemented with additional ascertainment sources to find all cases of DBMD. NCBDDD is working with four states to set up DBMD surveillance systems called the Muscular Dystrophy Surveillance Tracking and Research Network (MD STARnet). The states are Arizona, Colorado, Iowa, and western New York state. The goal of the project is to find all patients with DBMD in the areas and collect information about them from their medical records. Because many patients with DBMD are seen in Muscular Dystrophy Association (MDA) clinics, the researchers are working closely with the MDA clinics in their states. In addition, the MD STARnet researchers will be searching for patients with DBMD through other neuromuscular clinics, emergency rooms, pathology laboratories, orthopedists, and muscular dystrophy associations to ensure that all patients with DBMD are included in the project. The states have worked together to develop a common system to find patients and collect information. Families that are identified in these areas will be contacted yearly to collect up-to-date information. NCBDDD is also sponsoring projects to identify the service needs of families with DBMD, pilot newborn screening for DMD, pilot infant screening for DMD, and identify cardiac care issues in carrier females.

Thrombotic Diseases
Hereditary defects in one or more of the clotting factors in blood can cause the formation of potentially dangerous blood clots (thrombosis). Approximately 5% to 8% of the U.S. population has one of these clotting disorders, collectively called *thrombophilia*. Thrombophilia is a propensity for blood clotting in which a genetic defect can be identified that often results in thrombosis. More than 60,000 Americans die each year from venous thromboembolism, the blockage of a blood vessel caused by a clot dislodged from a vein; in addition, nearly half of patients with deep vein clots experience long-term health consequences that adversely affect their quality of life. Evidence suggests that thrombophilia is related to adverse pregnancy outcome, including thrombosis in pregnancy and potentially recurrent fetal loss, intrauterine growth restriction of the fetus, and preterm delivery. A coordinated, standard approach to managing the care of these patients has not been established among health care providers.

Despite the substantial percentage of people with clotting disorders, thrombophilia is not easily recognized. Identifying the predisposing genetic risk factors and interacting external or environmental factors could help determine which people are most susceptible to clotting and help prevent the resulting complications. Moreover, identifying and understanding the modifiable risk factors associated with the risk of thrombotic disease will facilitate preventions and intervention efforts.

Hemophilia
Hemophilia is an inherited bleeding disorder that affects approximately 18,000 people (primarily males) in the United States. The disorder results from deficiencies in blood clotting factors and can lead to spontaneous internal bleeding and bleeding following injuries or surgery. These bleeding episodes can cause severe joint damage; neurological damage; damage to other organ systems involved in the hemorrhage; and, in rare cases, death. Treating the bleeding episodes involves the prompt and proper use of clotting factor concentrates.

As many as one-third of people with hemophilia will develop an antibody (inhibitor) to the intravenous antihemophilic factor products that are infused to stop or prevent a bleeding episode. Although most of these inhibitors are transient and will resolve, about 5% to 7% of the hemophilia population have a clinically significant long-term inhibitor. An inhibitor renders the treatment product ineffective in controlling bleeding. The public health costs associated with inhibitors are staggering. People with hemophilia with inhibitors are twice as likely to be hospitalized for a bleeding complication. In addition, the cost of hospital care for a bleeding complication is an average of 10 times greater for those with inhibitors compared with those without an inhibitor. Incidence rates of inhibitors appear to vary according to the defect on the Factor VIII or IX gene. However, less than 7% of people with hemophilia enrolled in Universal Data Collection (hemophilia surveillance) have undergone an analysis of their genetic defect, primarily due to the high cost and lack of insurance reimbursement.

Major 2004 Accomplishments of NCBDDD Programs in Genetics

National Birth Defects Prevention Study (NBDPS) Biologics Summit and Epidemiologic Findings
After 5 years of data collection, in 2003 and 2004 sufficient power was achieved for several NBDPS birth defect groups to allow epidemiologic analyses to begin. In 2004, data were published in the *Morbidity and Mortality Weekly Report* in response to a public health concern, raised by a scientific study showing a possible association between the drug loratadine, also sold under the brand name Claritin® and recently approved for over-the-counter use, and a male genital defect (hypospadias). The study showed no association between use of loratadine in early pregnancy and the occurrence of hypospadias. This study did not examine genetic risk factors for hypospadias because of the specific concern about a medication

exposure, but numerous other proposals for molecular analyses have been made to the NBDPS Data Sharing Committee. In April 2004, a Biologics Summit was held in Atlanta, Georgia to discuss strategies for optimal use of collected DNA. Speakers from NBDPS centers and other institutions discussed cutting-edge techniques such as whole genome amplification, high-throughput genotyping, and haplotype analyses. Immediate, practical solutions to improve sample collection and DNA yields were also proposed. Pilot testing of some of these suggestions has been instituted as a result of this meeting.

Cystic Fibrosis Screening Recommendations
NCBDDD convened a workshop on newborn screening for cystic fibrosis (CF) in Atlanta, Georgia, in November 2003, cosponsored by the Cystic Fibrosis Foundation. The workshop brought together national and international experts and state health department staff to review evidence of benefits and harms of such screening. Presentations from the workshop will be published in a special issue of the *Journal of Pediatrics* in 2005. A writing group of seven people conducted an evidence review of relevant studies and prepared an *MMWR Recommendations and Reports* that was published on October 15, 2004. The main recommendation from that review was that, based on evidence of moderate benefit and low risk of harm, states should consider adding CF to screening panels.

Colorado Project for Tracking Affected Children Identified Through Newborn Screening and Collecting Data Relating to a Variety of Long-Term Outcomes
Through an NCBDDD cooperative agreement with the Colorado Department of Public Health and Environment, this project will provide aggregate data relating to surveillance, short-term management, and long-term follow-up of infants with metabolic conditions and hemoglobinopathies identified through newborn screening. The primary objectives for the Colorado health department have been to integrate data from newborn blood spot and hearing screenings and to ascertain information from specialty clinics for children with hemoglobinopathies, metabolic disorders, and other conditions identified by newborn screening. Data collection using a health department-developed electronic instrument in specialty clinics began in 2004, with plans to integrate data in the upcoming year.

Storage and Use of Residual Dried Blood Spots From State Newborn Screening Programs
Residual blood spot specimens have historically been used for quality control and evaluation of new screening tests and clinical or forensic testing. In recent years, numerous investigators have also used these stored blood spots for a variety of other purposes. The number and visibility of these types of studies have increased the awareness of the uses of stored blood spots among researchers, public health officials, ethicists, and policymakers. In 2004, investigators from NCBDDD collaborated with the Office of Genomics and Disease Prevention, the National Center for Environmental Health, and the Association of Public Health Laboratories to compile and present 2003 survey data regarding state practices on storage and use of blood spots.

Muscular Dystrophy Surveillance Tracking and Research Network (MD STARnet) Implementation
Data collection for this project began in April 2004. Families who are identified through the MD STARnet system will be asked to take part in interviews with public health workers to provide information related to DBMD that might not be found in the medical records. The types of information that will be collected include basic demographic information (such as race and ethnicity), the types of treatments that have been received, the types of clinics that the care was received in, and any medical problems associated with DBMD.

Newborn Screening for Duchenne Muscular Dystrophy (DMD) Workgroup
On March 12, 2004, NCBDDD sponsored a one-day meeting in Atlanta, Georgia, with experts from around the world to look at newborn screening for DMD. At the meeting, past and

present DMD newborn screening programs were discussed, as well as known and potential risks and benefits of such programs. A report for the general public was released in September 2004 (http://www.cdc.gov/ncbddd/duchenne/NBS_Lay_Report.pdf). A second report will be submitted for publication in a peer-reviewed journal. Highlights from the meeting were presented at the 2nd National NCBDDD Conference in July 2004.

Long-Term Follow-Up of Children with Metabolic Disorders Identified Through Tandem Mass Spectrometry-Based Newborn Screening
Newborn screening programs have traditionally been limited to conditions that are serious, treatable or controllable, and with a natural history that is understood. In the 1990s, the technology of tandem mass spectrometry was introduced for population-based newborn screening. This technology allows for a more accurate measurement of metabolites associated with a broader range of conditions than was previously available. For some conditions, such as phenylketonuria, the benefits of newborn screening and early treatment are generally accepted. However, for other conditions, such as some of the fatty acid oxidation disorders, evidence regarding long-term benefit from screening and intervention is lacking. For certain fatty acid oxidation disorders, there is also currently a lack of data regarding the distinction between pathologic levels of some metabolites and natural variations in levels. Population-based tracking and follow-up studies of children identified through tandem mass spectrometry are needed to assess the public health impact of newborn screening for many of the disorders identified with this technology. In 2004, an electronic system was developed and tested to collect and pool data on children in Idaho, Iowa, and Oregon with selected metabolic disorders. The data collected include collection of information related to treatment options, treatment compliance, and long-term outcomes.

Ongoing Initiatives Related to Genetic Disorders of Hemostasis
The Division of Hereditary Blood Disorders (DHBD), NCBDDD, is collaborating with several academic institutions to conduct case-control studies investigating the association of gene polymorphisms with thrombotic outcomes, evaluating health-related services directed toward the reduction or prevention of complications of thrombosis and thrombophilia, and investigating inhibitor development in people with hemophilia.

The Role of the Maternal Hemostatic System and Maternal Coagulation Genes in Intrauterine Growth Restriction
Disruption of uteroplacental circulation resulting in decreased nutrient exchange between mother and fetus has been implicated as a potential cause of intrauterine growth restriction (IUGR). The DHBD, NCBDDD, is conducting a multisite case-control study to evaluate the presence and interaction of genetic polymorphisms of the coagulation, fibrinolytic, and inflammatory pathways and their association with IUGR. Study enrollment began in February 2000 and is expected to continue through 2005. DNA samples are being collected from both mother and fetus to assess the role of both the maternal and fetal hemostatic systems in IUGR. Findings from this research will help to further elucidate the role of these genes and pathways in IUGR pathophysiology. Results might also help to identify women and fetuses most at risk for IUGR and to develop targeted preconception care, as well as management for IUGR pregnancies.

The Genetic Attributes and Thrombosis Epidemiology (GATE) Study
The GATE study is a case-control study to evaluate the genetic and environmental risk factors for venous thromboembolism (VTE). Current analyses have focused on associations between VTE and polymorphisms in coagulation, fibrinolytic, and inflammatory pathways. Early analyses have been published in the scientific literature and presented at international meetings. Findings from this research will help to elucidate the role of gene polymorphisms and

interactions with lifestyle factors in VTE pathophysiology. Results could also help to identify people at risk for VTE and to develop targeted preventive care.

Health Care Services for People With Thrombotic Disease
The Division of Hereditary Blood Disorders, NCBDDD, has cooperative agreements with eight Hemostasis and Thrombosis Centers (pilot sites) in the United States to provide health-related services, directed toward the reduction or prevention of complications of thrombosis, thrombophilia, and bleeding disorders. The initial objectives of this collaboration are to characterize the patients treated at pilot Hemostasis and Thrombosis Centers and the types of services provided. To achieve these goals, a Web-based data registry has been created and implemented at the pilot sites. Through the registry, information is collected for the purposes of: (1) determining the demographics of the patient population; (2) characterizing referral patterns to the centers; (3) elucidating the experiences of the patients at the centers (extent of care, encounters with health care professionals, and educational materials received); (4) describing the medical history of these patients, laboratory and radiological tests performed, and treatment prescribed; and 5) investigating and characterizing use of the center through follow-up and return visits. Since August 2003, 895 patients have been enrolled in the registry. Long-term goals are to evaluate the effectiveness of an integrated health care model in preventing and treating coagulation disorders and their complications. Data collection across pilot sites will facilitate future collaborative clinical and epidemiologic investigations, health services evaluations, and quality of life assessments.

Inhibitors and Hemophilia
Beginning in 2005, DHBD will begin a pilot study to perform post-marketing surveillance of treatment products for inhibitors as part of an ongoing surveillance project already established in the bleeding disorders community. Hemophilia genetic testing is important to the success of this study for several reasons. First, hypotheses concerning the risk for inhibitor development and specific genetic defects can be tested. Second, the possible confounding effects of certain genetic defects on product-specific risk can be assessed. Third, should the results of the study confirm that patients with certain defects are at higher risk of inhibitors, interventions designed to minimize this risk can be developed and tested. For the study, mutations in FVIII will be detected by direct nucleotide sequencing. In the first year of the study, genetic sequencing will be performed on 500 patients.

Future Directions

Based on the four main priorities outlined in the introduction, investigators at the National Center on Birth Defects and Developmental Disabilities and NCBDDD grantees have specific goals for future activities:

Investigate Susceptibility Genes for Birth Defects

- Begin analyses of National Birth Defects Prevention Study biologic data to investigate for gene–environment and gene–gene interactions in the causation of specific birth defects.
- Initiate alternative and novel procedures to improve collection of biologic samples from NBDPS participants.
- Determine strategies for optimal analyses of collected NBDPS samples through a genetic analysis working group.

Support, Develop, and Evaluate Newborn Screening Programs for Genetic Disorders

- Support existing and new efforts to link newborn screening records with postnatal data sources, such as clinic treatment records, to ensure optimal outcomes.

- Study the risks and benefits of newborn and other early screening programs for conditions that do not meet the traditional criteria for newborn screening in the United States.
- Enhance long-term evaluation of children identified through tandem mass spectrometry-based newborn screening by increasing the number of states participating in long-term follow-up studies.
- Conduct economic evaluations of newborn screening tests.

Improve Quality of Life of Children with Potentially Disabling Single Gene Disorders and Their Families

- Identify preventable risk factors for secondary complications.
- Decrease age of diagnosis.
- Improve health care and related services for families.
- Decrease barriers to health care and related services.

Prevent Morbidity and Mortality Related to Genetic Disorders of Hemostasis

- Complete, analyze, and disseminate results from case-control studies in the Division of Hereditary Blood Disorders, NCBDDD. Determine risk factors and describe potential intervention measures.
- Evaluate the effectiveness of the specialized health care system to improve health outcomes and quality-of-life measures for people with thrombophilia.
- Develop hereditary blood disorder educational materials for patients and health care providers.
- Develop rapid screening methods to detect risk factors for thrombosis.
- Extend genetic testing for people with hemophilia and inhibitors.

GENETIC TESTING PRACTICE ACTIVITIES
DIVISION OF PUBLIC HEALTH PARTNERSHIPS-LABORATORY SYSTEMS
NATIONAL CENTER FOR HEALTH MARKETING

Who we are

The genetics group within the National Center for Health Marketing's Division of Public Health Partnerships includes two board-certified laboratory geneticists and three ORISE fellows with expertise in genetics, genetic counseling, nursing, and molecular technology.

What we do

We engage the public and private sector to promote the quality of genetic testing in clinical and public health practice. We accomplish this by conducting activities that

- monitor and evaluate current and emerging laboratory practices associated with genetic testing
- assure appropriate ordering, reporting, and use of genetic tests and results
- assure availability and access to a variety of quality control (QC) materials
- assure availability and access to quality genetic testing for rare diseases
- assure a competent workforce
- develop and evaluate voluntary and regulatory standards, guidelines, and policies

---

Our partners and customers

Much of our work includes ongoing collaborations with a diversity of partners that include

- State and local public health professionals and laboratories
- Clinical laboratory services
- Providers of genetic testing services
- Users of genetic testing services
- Policy makers that address genetic testing issues
- International public health organizations

Monitoring and evaluation activities

We are working with external partners to prepare reports on the current and emerging state of laboratory practices relevant to genetic testing. These efforts inform program development and evaluation for CDC, other federal and State agencies, the private sector, and international efforts.

Major Accomplishment, 2004:
Co-authored "Ensuring the Quality of Genetic Testing in the United States," and "Carrier Testing for Cystic Fibrosis: Transition from Research to Clinical Practice," chapters in the *CDC Report on Genomics and Public Health 2003,* published March 2004.

Ordering, reporting and use of genetic tests

We are facilitating the development of laboratory and health provider networks to address public health concerns in assuring appropriate ordering, reporting, and use of genetic tests and results. We are also working with the international community to address similar issues pertaining to cross-border flow of genetic test referrals.

Major Accomplishments, 2004:

- As a result of a May 2003 national conference held in collaboration with Mt. Sinai's College of Medicine, efforts are underway to create more effective requisitions for genetic testing services and improve test reports by assuring critical test result information is clearly and concisely presented.
- We partnered with the European Union cystic fibrosis network to perform a comparative analysis of U.S. and European reporting practices. Results from this study will assist policy makers in determining where international consensus may be achieved.
- We accepted a co-leadership position in addressing international test result reporting issues as part of our involvement with the Organization for Economic Cooperation and Development's (OECD) Steering Group on Quality Assurance and Proficiency Testing in Molecular Genetic Testing.

QC materials

We are spearheading a community process to implement a sustainable mechanism to provide appropriate quality control materials and guidelines for use in routine genetic testing, inter-laboratory comparisons of quality assessment, and test development.

Major Accomplishments, 2004:

- Together with key members of the genomic/molecular diagnostic community we hosted two conferences entitled "QC Materials for Genetic Testing." The conferences were held to develop recommendations relative to current and future needs for genetic testing QC materials and establish a sustainable, practical process to make QC materials available to genetic testing laboratories.
- Based on the recommendations from the above conferences, a framework has been established to collect, store, verify, and distribute certain cell lines that may be used for genetic quality control and assessment procedures. A CDC-based Coordination Program will facilitate the collection and verification of and access to the materials; private sector repositories will store and distribute them.

Rare disease genetic testing

We are addressing an urgent need to assure and improve availability of quality genetic testing for rare diseases.

Major Accomplishment, 2004:
In collaboration with other HHS agencies, CDC CIOs, professional organizations, patient advocacy groups, and Emory University, we convened a conference entitled "Promoting Quality Laboratory Testing for Rare Diseases: Keys to Ensuring Quality Genetic Testing." The conference brought together a broad scope of stakeholders to address crucial issues surrounding the availability of and access to quality genetic testing for rare diseases, which affect 25 million (or 1 in every 12) people in the United States.

Workforce competencies

Working with our external partners, educational programs are developed, taught, and evaluated to promote workforce competency of users and providers of genetic testing services.

Major Accomplishments, 2004:

- With assistance from the American College of Medical Genetics and the American Medical Association, we have distributed over 4,000 copies (1,415 copies in 2004) of our award-winning interactive CD-ROM entitled "Genetics in Clinical Laboratory Practice: A Team Approach." This interactive multimedia educational tool, developed through a cooperative agreement with Dartmouth Medical School, emphasizes the uniqueness of the genetic testing process and increases the knowledge of healthcare providers using genetic tests in clinical practice.
- Since September 2002, a webcast produced by the Division of Laboratory Systems entitled "A New Era in Newborn Screening - Saving Lives, Improving Outcomes," has been viewed by a total of 690 health professionals (249 in 2004). The satellite program, also available on CD ROM, provides information to improve recognition, detection, and diagnosis leading to early intervention and effective management of metabolic disease, endocrine disorders, and hemoglobinopathies in newborns.

Standards, Guidelines, and Policies

We are working with federal and State agencies, professional organizations and international groups to develop voluntary and regulatory standards, guidelines, and policies to improve the quality of laboratory practice.

Major Accomplishments, 2004:

- We continued to provide the Centers for Medicare & Medicaid Services technical consultation on the development of quality and personnel requirements for genetic testing under the Clinical Laboratory Improvement Amendments of 1988 (CLIA).
- We participated as U.S. representatives to the OCED's Steering Group's meeting on Quality Assurance and Proficiency Testing for Molecular Genetic Testing.

Future activities

In 2005 we plan to

- Conduct health services research through a cooperative agreement with the Wadsworth Center, New York State Department of Health, entitled "Improving the Quality of Genetic Testing and Assuring its Appropriate Integration into Clinical and Public Health Practice"
- Convene a conference entitled "Best Practices in Ordering and Reporting Genetic Tests and Results"
- Publicly launch the QC Materials Coordination website, which will facilitate information exchange and communication about QC materials needs and availability
- Disseminate validated cell lines for genetic testing quality control and quality assessment activities
- Facilitate development of public/private processes to improve genetic testing for rare diseases
- Collaborate with public health laboratories to identify the resources and training needs to ensure continued capacity to support public health services
- Continue to collaborate with other federal agencies, professional organizations, and users and providers of laboratory services to assure quality genetic testing services and protect patient safety
- Continue to collaborate with OECD, Eurogentest, ISO and other international groups to harmonize and promote quality and good practices of genetic testing services worldwide

---

For more information

For additional information, please contact Dr. Joe Boone by calling (770) 488-8080, or by e-mail at djb2@cdc.gov

---

COORDINATING CENTER FOR INFECTIOUS DISEASES (CoCID)

CoCID Genomics Working Group: In 1996, CDC's National Center for Infectious Disease (NCID) established the Genetic Working Group (GWG) with the purpose of addressing issues concerning the role of genetics in infectious diseases and promoting collaboration between investigators. This working group was further expanded in 2003 to coordinate the genomic research interests of NCID, National Immunization Program (NIP), and National Center for HIV, STD, TB Prevention (NCHSTP) and the Office of Genomics and Disease Prevention (OGDP). The Coordinating Center for Infectious Diseases (CoCID, created May 2004) includes NCID, NIP and NCHSTP.

The overall mission of this working group includes:

a) Identifying and conducting investigations of host genes associated with infectious diseases and vaccine safety that have public health relevance and to which interventions and preventions can be targeted;
b) Building laboratory and epidemiological capacity for genomics;
c) Training public health professionals in laboratory and epidemiological aspects of host genomics and infectious diseases;
d) Communicating information about genomics and infectious diseases to the public through meetings, seminars, web sites, and publications.

Top Priorities

Seven high priority research areas have been identified for host genetics in CoCID. These reflect areas where CoCID and its partners are making or can make a significant impact in research that leads to new control and prevention methods or development of control and prevention strategies.

1. ***Infectious Diseases:*** resistance and susceptibility to and severity of infectious diseases.
2. ***Chronic disease etiology:*** clarifying host-pathogen interactions in the causation and pathophysiology of chronic diseases such as chronic fatigue syndrome (CFS), arthritis, cancer, and cardiovascular diseases.
3. ***Vaccines:*** response, failure, adverse events; using host genetic information for vaccine design.
4. ***Antimicrobial and adjunctive therapy response:*** failure, adverse events and development of new therapies.
5. ***Prevention:*** identification of genetically at-risk populations for infectious diseases; targeting of prevention strategies to these populations.
6. ***Development of rapid molecular tools*** based on DNA, mRNA and protein expression patterns that can be used in surveillance, studies monitoring response to therapy, and basic-research.

7. ***Drug metabolism:*** defining pharmacogenetic factors associated with response to drug therapy, drug interactions, and disease outcome in tuberculosis and other infectious diseases.

Major Accomplishments, 2004

A total of sixteen NCID GWG projects were funded internally with annual funding ranging from $30,000 to $60,000 per year from 1998 to 2003. A number of key scientific observations have been published from these studies. In 2004, the GWG hired a Career Development fellow, who has significant experience in human genetics, to support genetic investigations and provide technical guidance on study design and data analysis for investigators in NCID and other centers.

The CoCID GWG conducts cross cutting work to promote host genomics research that includes:

- Collaborations with OGDP in a CDC-wide effort to integrate genomics research into acute public health investigations that involved an expert meeting in May 2004 to formulate a CDC wide strategic plan. This includes developing model consent forms for IRB pre-approval, facilitating storage capability for long-term storage of specimens, technical guidance on study design and methods, technical support for genotyping and data analysis, and outlining criteria to prioritize outbreaks for incorporation of host genomics.
- Ongoing contributions to the CDC-wide effort to use the National Health and Nutrition Survey (NHANES) III DNA Bank to identify the frequency of different mutations in the immunologically relevant host genes and support for data analysis related to infectious disease outcomes. CDC/OGDP has an IRB-approved proposal to generate nationally representative frequency data on variants of genes of public health significance using NHANES III DNA samples. NHANES III is a nationally representative sample of the US population and DNA specimens are available from ~ 7,000 persons that reflect the racial composition of the U.S. population. In 2003, CDC developed a protocol to determine the prevalence of 57 genes of public health significance by looking at 87 different polymorphisms. When complete, this will generate the most comprehensive database of the background frequencies of these genes in a healthy population that is unparalleled in size in the United States. Of the 57 genes chosen for the NHANES project, more than 30 are relevant to infectious diseases. The GWG will assist with the data analysis of infectious diseases related issues. (More detail at: http://www.cdc.gov/genomics/activities/ogdp/2003/chap01.htm).
- Assistance in development and testing of candidate gene selection software tools for genomics research as part of CDC-wide effort.
- Provision of technical and data analysis support for research investigations related to host genomics projects at CoCID.
- Organization and promotion of research and educational seminars in the genomics of infectious diseases.

Ongoing Initiatives

Important ongoing initiatives are listed by title and Center in Tables 1–3. A description of each listed project is provided in the text that follows, organized by genomics priority research area.

Table 1. **Genomics Research Projects in NCHSTP**

Title	Division
Infectious Diseases: Resistance and Susceptibility to Infection and Disease, and Severity of Infectious Diseases	
• Genomics approach to identify drug targets for a variety of pathogen	DHAP
• Influence of host genetics on HIV resistance, disease progression and effect of virological failure on HAART	DHAP
• Role of gene polymorphisms in susceptibility to TB in Botswana	DHAP/DTBE
• Prospective evaluation of immunogenetic risk factors for susceptibility to TB infection and progression to TB disease	DTBE/DHAP
Pharmacogenetics of Drug Metabolism and Genetic Risk Factors Associated with Treatment Outcome	
• N-Acetyl Transferase Type II genotype and TB treatment outcome with isoniazid and rifapentine and rifabutin regimens	DTBE/DHAP
• 2C19 polymorphisms and metabolism of nelfinavir	DTBE/DHAP
• Evaluation of genetic risk factors for TB treatment outcome	DTBE/DHAP

Table 2. **Genomics Research Projects in NIP**

Title	Division
Genetic risk factors associated with vaccine related adverse outcomes	
• Establishing registry of clinically significant adverse events and related clinical data and a repository of biological specimens	ESD
• Nested case control study of Smallpox vaccination and myo/pericardial injury and inflammation	ESD
• Retrospective cohort study to evaluate genetic predisposition to developing rheumatoid arthritis in persons receiving Hepatitis B vaccine	ESD

Table 3. **Genomics Research Projects in NCID**

Title	Division
Integrating genomics into acute public health investigations of infectious	

diseases	
• Role of genetic risk factors in Leptospirosis outbreak	DBMD
• Outbreak investigations involving Severe Acute Respiratory Syndrome (SARS)	DVRD
• Analysis of genetic risk factors for fatal influenza in children	DVRD
• Searching for a flavivirus susceptibility or resistance gene in West Nile virus (WNV) outbreak	DVRD
Infectious Diseases: Resistance and susceptibility to infection and disease and severity of infectious diseases	
• Genomics research studies in malaria	DPD
• Genetic risk factors associated with development of lymphedema in lymphatic filariasis	DPD
• Genetic factors in meningococcal and pneumococcal diseases	DBMD
• Molecular signatures of cervical neoplasia	DVRD
• Integration of gene expression, clinical, and epidemiological data to characterize chronic fatigue syndrome	DVRD
• Host gene expression profiles that precipitate post-infective and chronic fatigue syndromes in response to common infections	DVRD
• Exercise responsive genes measured in peripheral blood of women with chronic fatigue syndrome and matched control subjects	DVRD
• Integration of gene expression and clinical data from CFS andnon-fatigued subjects	DVRD
• Development of a text mining tool that provides gene information in specific disease and biological context	DVRD
Infrastructure	
• Setting up a high throughput genotyping capacity	SRP
• Global amplification of sense RNA: a novel method to archive and replicate RNA for microarrays	DVRD
• Beta-test site for emerging genomic technologies	DVRD
• Automated tools for genomic research	DVRD/SRP
• Development of a gene expression microarray database for flexible data processing, analysis and archiving	DVRD

Integrating Genomics into Acute Public Health Investigations of Infectious Diseases

Host genes play an important role in influencing infectious disease outcomes. Integration of host genomics investigations into outbreak investigations can help identify particular host factors relevant for severe disease outcomes and transmissibility of diseases. Some examples of acute public health investigations involving genomics research have included the following.

- **Role of genetic risk factors in Leptospirosis outbreak** (CDC/NCID/Division of Bacterial and Mycotic Disease (DBMD)/Meningitis and Special Pathogens Branch): The first success story of integrating host genomics into an outbreak investigation at CDC[16]occurred during a leptospirosis outbreak involving triathletes exposed to lake water. HLA-DQ6-positive triathletes had increased risk of laboratory-confirmed leptospirosis (OR=2.8, P=0.04) compared to DQ6-negatives. DQ6-positive triathletes swallowing lake-water had greatest risk (OR 8.46, P< or =0.001). This investigation is the first report of a genetic risk factor affecting susceptibility to leptospirosis and of documented gene-environment interaction (DQ6 and swallowed water) that affected infectious disease susceptibility.

- **Outbreak investigations involving Severe Acute Respiratory Syndrome (SARS)** (NCID/Division of Viral and Rickettsial Diseases (DVRD)/Respiratory and Enteric Viruses Branch)**:** A pilot study to evaluate the role of immunologic and host genetic factors in SARS infection is ongoing using specimens collected as part of the 2003 SARS outbreak.

- **Analysis of genetic risk factors for fatal influenza in children** (CDC/NCID/DVRD/Influenza Branch)**:** A case-control study integrated with multistate population-based surveillance for influenza-related hospitalizations in children during the 2004-2005 season to identify genetic variants that may put certain children at a greater risk of mortality associated with influenza.

- **Searching for a flavivirus susceptibility or resistance gene in West Nile virus (WNV) outbreak** (CDC/NCID/Division of Vector-borne Infectious Diseases (DVBID)/Arboviral Diseases Branch): CDC scientists are currently collaborating with Georgia State University to search for a flavivirus susceptibility/resistance gene in persons who were hospitalized for WNV disease during the 2003 WNV outbreak in Colorado and to study genetic polymorphisms that confer flavivirus resistance in mice, humans and horses with either asymptomatic or overt WNV disease.

Infectious Diseases: Resistance and Susceptibility to Infection and Disease, and Severity of Infectious Diseases

A better understanding of the role of various genetic factors in the outcome of infectious diseases will lead to development of novel therapies and better prevention efforts. Recent progress in high throughput genotyping and decoding of the human genome has led to new optimism in identifying such genetic risk factors. The availability of well-

characterized large cohorts from different populations will be critical in identifying such risk factors. CDC has access to epidemiologically well-characterized different infectious disease cohorts. These cohorts are unique resources in identifying such genetic risk factors. Several investigations at CDC that have integrated genomics research are highlighted below.

- **Genomics factors associated with resistance and susceptibility to malaria** (CDC/NCID/Division of Parasitic Diseases (DPD)/Malaria Branch): By taking advantage of a unique community based birth cohort study conducted in Kenya, researchers have demonstrated that sickle cell traits provide significant protection against mortality and morbidity, especially in infancy, before the onset of clinical immunity in areas with intense transmission of malaria[1]. A novel polymorphism in the inducible nitric oxide synthase gene confers protection against severe malaria anemia[13]. In addition, Fc receptor IIA genetic polymorphism was found to be associated with protection against high-density malaria infection in children[30]. This project has led to several collaborative studies involving various academic institutions including the National Institutes of Health (NIH). The ongoing studies focus on identifying the genetic risk factors associated with severe malaria.

  A collaborative study with the Malaria Research Center of India and the Morehouse School of Medicine in Atlanta involves using host genomic approaches to identify potential biomarkers and genetic risk factors associated with cerebral malaria and its associated neurological deficits in collaboration.

  Researchers demonstrated that Fc receptor IIA genetic polymorphism is associated with protection against placental malaria infection in HIV positive women and perinatal HIV infection[5, 6].

- **Genetic risk factors associated with development of lymphedema in lymphatic filarisis** (CDC/NCID/DPD): Using a filarial study of family cohort, investigations are underway to determine the genetic and environmental factors that affect the risk of developing lymphedema due to *Wuchereria bancrofti* infection.

- **Genetic factors associated with meningococcal and pneumococcal diseases** (CDC/NCID/DBMD): Using blood spot specimens collected at birth for state-based newborn screening programs (NBS), efforts are underway to identify genetic risk factors associated with the risk of and susceptibility to severe meningococcal and pneumococal disease, particularly among populations of infants and young children enrolled in routine immunization programs. CDC is collaborating with the Wisconsin EIP (Emerging Infections Program) and NBS program, Wadsworth Lab of New York State University, Emory University, University of Washington, and Molecular Staging.

- **Genomics approach to identify drug targets for a variety of pathogens** (CDC/NCHSTP/Division of HIV/AIDS Prevention (DHAP) and NCID): Using gene

trap technique to identify host cell proteins that are critical for survival of pathogens within host cells, about 200 candidate host genes that potentially play critical roles in the life cycles of Marburg (MBG) virus, Ebola (EBO) virus, HIV-1 and HIV-2, influenza A, or retrovirus have been identified. One candidate gene, Rab9, has been explored as a potential target for anti-viral therapy.

- **Molecular signatures of cervical neoplasia** (CDC//NCID/DVRD): As part of the National Cancer Institute's (NCI) Early Detection Research Network a genomics project to detect and validate biomarkers that can be used to improve the sensitivity and specificity of cervical cancer screening is ongoing. This project focuses on gene expression profiling of exfoliated cervical cells and proteomic profiling of cervical mucous samples using SELDI technology. Protocols were established for optimizing molecular quality of archived biologic samples and the method was published for extracting RNA and DNA from exfoliated cervical cells and for archiving RNA allowing sense global amplification and making cDNA in biorepositories a renewable resource. Candidate biomarkers that were identified are being explored further.

- **The chronic fatigue syndrome (CFS) program to integrate genomics into chronic infectious diseases and illness** (CDC/NCID/DVRD/Viral Exanthems and Herpes Virus Branch (VEHB)): The CFS Molecular Epidemiology Program was established in 1997. The CFS Program was designed to apply rapidly evolving cutting-edge genomics, proteomics, and bioinformatics technology to epidemiologic studies whose objective is CFS prevention and control. Its aim is to characterize CFS at a systems biology level by integrating surveillance, case definition, and clinical studies with genomics, proteomics and bioinformatics. The effort includes data from population-based and clinical studies. Examples of each of these are described below.

    - **Integration of gene expression, clinical, and epidemiological data to characterize CFS**
        - Integrated the peripheral blood gene expression results with epidemiological and clinical data to determine whether CFS is a single or heterogeneous illness
        - Using statistical tests and cluster analysis to distinguish CFS subjects and identify differentially expressed genes
        - The latest results suggest that CFS is a heterogeneous illness. The differentially expressed genes imply fundamental metabolic perturbations that will be further investigated and illustrates the power of microarray technology for furthering our understanding CFS[37]

    - **Host gene expression profiles that precipitate post-infective and chronic fatigue syndromes in response to common viral and rickettsial infections**
        - An example of a CFS gene expression study that is based on model systems

- Studying host gene expression profiles following acute infection with Epstein Barr Virus (EBV), *Coxiella burnetti* (the causative agent of Q fever) and Ross River Virus (RRV) in collaboration with the University of New South Wales in Sydney, Australia**.** Some observations from this longitudinal study include: a) severity of the acute illness is a powerful predictor of the likelihood of development of post-infective fatigue syndrome (PIFS) at three and six months, b) although the pattern and severity of symptoms in the acute illness were correlated with production of pro-inflammatory cytokines, these relationships did not persist through to the PIFS phase of the illness, c) identified several novel gene expression correlates of individual symptoms using microarray gene expression profiling

- **Exercise responsive genes measured in peripheral blood of women with chronic fatigue syndrome and matched control subjects**
  - Measuring peripheral blood gene expression profiles of women with CFS and matched controls before and after exercise challenge to search for markers of CFS-associated post-exertional fatigue, differential expression of exercise-responsive genes classified in chromatin and nucleosome assembly, cytoplasmic vesicles, membrane transport, and G protein-coupled receptor ontologies between CFS patients and controls

- **Integration of gene expression and clinical data from CFS and non-fatigued subjects enrolled in a two day clinical evaluation in Wichita, Kansas**
  - Following a well-characterized cohort of people with CFS over a four-year period to determine if unique gene expression profiles are associated with symptom occurrence or persistence of illness**.** Evaluated each subject and the corresponding multiple samples using a 40,000-gene microarray. Data analysis is in progress. This will also serve as the dataset for C^3, the CFS Computational Challenge, (described in section F 'Major Conferences')

- **Development of a text mining tool that provides gene information in specific disease and biological context**

  - Pioneered a number of genomic and bioinformatics technologies at CDC's VEHB (see CDC MAdB in Infrastructure section below)
  - Simultaneous assessment of tens of thousands of genes using high-throughput technology, such as gene expression profiling using microarrays
  - Key to getting specific digital information associated with genes is mining text in the appropriate biological, clinical and epidemiological context. Development of a text-mining tool that will provide biologic

and disease relevant information for genes identified as important by microarray gene expression profiling

- **Influence of host genetics on HIV resistance, disease progression and effect of virological failure on highly active anti-Retroviral therapy (HAART)** (CDC/NCHSTP/DHAP/Laboratory Branch): Researchers have focused on the influence of host genetics on HIV resistance, disease progression, and more recently on the effect of virological failure on HAART therapy in diverse populations. They found for the first time that CCR5 human haplogroup E was associated with an accelerated CD4 count decline to <200 cells/μL in an injecting drug user cohort from Thailand, providing evidence that the CCR5 human haplogroup E accelerates the decline of the CD4 cell count and may lead to accelerated disease progression in an Asian population. Study of the effect of host genetics on virological failure of HAART using a HIV-infected adult cohort with well-documented treatment history is currently undertaken.

- **The role of host genetic polymorphisms in susceptibility to *M. tuberculosis* infection and progression to TB disease** (CDC/NCHSTP/DHAP/Laboratory Branch and Division of TB Elimination (DTBE)):

  Tuberculosis (TB) is the leading infectious cause of death among adults worldwide, with nearly 2 million deaths annually from TB worldwide. Although an estimated one-third of the world's population is infected with *Mycobacterium tuberculosis*, only approximately 10% of infected persons will advance to TB disease, with the remainder maintaining latent infection. The low ratio of disease to infection despite almost universal exposure in certain regions and reports of ethnic differences in susceptibility to disease suggest that host-specific factors may play a major role in progression to TB disease. Identifying genetic risk factors for susceptibility to TB would permit TB programs in other countries to target costly interventions such as contact investigation, treatment for latent TB infection, directly observed therapy, and intensive follow-up towards the 10% of exposed persons truly at risk of developing TB. This targeted approach would result in considerable cost savings, and ultimately, improved TB prevention and control.

  In recent years a role of host genetic polymorphisms in susceptibility to *M. tuberculosis* infection and progression to TB disease has been reported. Based on these reports of genetic associations with susceptibility to TB, polymorphisms in at least 10 genes may influence susceptibility to TB infection and/or progression to TB disease. Despite these single gene associations, no study has reported the role of mutiple genes in a single study. Moreover, all studies to date have been conducted in TB-endemic countries, where multiple TB exposures and use of BCG vaccine limits the possibility of evaluating genetic risk factors for TB infection.

  - **Role of gene polymorphisms in susceptibility to TB in Botswana** (CDC/NCHSTP/ DHAP and NCHSTP/DTBE)

- Investigating the potential role of multiple immune response genes in infection with *M. tuberculosis* and disease progression in a cohort of patients reporting to a respiratory clinic in Gaborone, Botswana in 1998 – 2000
- Assessing the role of polymorphisms in the N-acetyl transferase 2 (NAT2) gene on drug metabolism and treatment failure also in this study
- Significant associations of polymorphisms in the IL-10 gene promoter -819/-592 haplotype and the NOS2 -954 with TB disease and an association of the IFNγ +874 with latent TB infection

- **Prospective evaluation of immunogenetic risk factors for susceptibility to TB infection and progression to TB disease** (CDC/NCHSTP/DTBE, NCHSTP/DHAP, and TB Epidemiologic Studies Consortium)
  - Launched a prospective study of polymorphisms of 17 different candidate genes in an epidemiologically well-characterized population of contacts exposed to infectious tuberculosis patients at 11 sites in the United States and Canada
  - Genotypes under investigation individually and in combination
  - Candidate genes were selected based on literature reports, biologic plausibility, and data from animal models. Genotype and haplotype frequencies for exposed contacts who do and do not develop latent TB infection will be compared. TB registry matches will be conducted annually for 5 years after enrollment is completed to assess progression to TB disease among enrolled contacts. A total of 1300 contacts have already been enrolled into this 5–year prospective study. Enrollment is planned through December 2006

Pharmacogenetics of Drug Metabolism and Genetic Risk Factors Associated with Treatment Outcome (CDC/NCHSTP/DTBE, NCHSTP/DHAP, and the TB Trials Consortium)

A series of pharmacokinetic studies undertaken have sought to investigate aspects of the pharmacogenetics of TB therapy. The first efforts involved incorporation of NAT-2 genotypes into the evaluation of isoniazid PK and its relation to treatment failure and relapse. A second effort has sought to assess the impact of possible polymorphisms of a particular cytochrome p450 isoform on the interaction between rifamycins and the anit-HIV protease inhibitor nelfinavir. Additional studies are being launched to evaluate genetic risk factors for TB treatment outcome.

- **N-acetyl transferase type II genotype and TB treatment outcome with once weekly isoniazid and rifapentine and twice-weekly rifabutin regimens**. Emergence of resistance during treatment with once weekly regimens including isoniazid is associated with isoniazid acetylator phenotype, occurring almost exclusively among patients who metabolized isoniazid quickly (so-called “fast

acetylators") and who therefore had a shorter period of drug exposure. Thus, isoniazid acetylator phenotype may be an important factor in the emergence of rifamycin resistance in the rifapentine trial. Isoniazid is metabolized by a polymorphic group of enzymes called N-acetyl transferases. The primary locus involved in INH metabolism is called NAT-2 and methods exist to identify the precise genotypes present in a given individual.

- **2C19 polymorphisms and metabolism of nelfinavir** (DTBE/NCHSTP, DHAP/NCHSTP and the TB Trials Consortium). The P450 isozymes CYP3A4, CYP2C19, CYP2D6 and CYP2C9 are responsible for nelfinavir metabolism *in vitro*. Unlike the other protease inhibitors, which are predominantly metabolized by CYP3A4, approximately one-half of nelfinavir clearance is achieved by CYP2C19. Several minor metabolites and one major oxidative metabolite of nelfinavir are found in plasma. The major oxidative metabolite, NFV hydroxy-t-butylamide (M8), has HIV activity comparable to the parent compound in vitro. The contribution of M8 to the clinical efficacy of nelfinavir is uncertain. Transformation of nelfinavir to M8 appears to occur solely at 2C19, as genetic polymorphism at 2C19 (poor metabolism at CYP2C19 is seen in 3-5% of Caucasians), or drug inhibition of 2C19 resulted in absence of M8 and increased plasma concentrations of nelfinavir.

Evaluation of genetic risk factors for TB treatment outcome (DTBE/NCHSTP, DHAP/NCHSTP, and the TB Trials Consortium). An evaluation of genetic risk factors for TB treatment outcome has been integrated into an ongoing TB clinical trial evaluating nucleic acid amplification tests and epidemiological factors as potential surrogate markers to predict relapse of tuberculosis and to monitor the effectiveness of treatment. Surrogate markers may help to tailor treatment in individual patients and might also permit more rapid and efficient conduct of therapy trials, shortening significantly the total duration of these long and costly studies. The study is co-enrolling patients in TBTC TB treatment trials. Each subject is thoroughly described clinically, radiographically, and by standard TB microbiology. A total of 130 patients have been enrolled to date, out of a planned sample size of 300. The existence of a bank of serial plasma, buffy coat, and sputum specimens from a well-characterized group of patients provides an opportunity also to evaluate possible genetic and serologic markers of treatment outcome. Genotypes to be tested include polymorphisms of MDR1, UGT, Toll-like receptor 2, and IL12B receptor.

Genetic risk factors associated with vaccine related adverse outcomes (NIP)

Pre-licensure clinical safety trials and post-licensure signal detection have been the benchmarks to measure vaccine safety. Mainstays of vaccine safety surveillance are the Vaccine Adverse Event Reporting System (VAERS) and the Vaccine Safety Data link (VSD). VAERS is a national passive surveillance system and receives reports of suspected adverse event following receipt of any U.S. licensed vaccine. Its goals are to detect new, unusual, or rare vaccine adverse events; monitor increases in known adverse

events; and determine potential patient risk factors for particular types of adverse events. VAERS was established in 1990 and is jointly administered by CDC and the Food and Drug Administration (FDA). Although VAERS receives approximately 20,000 reports per year, the clinical and laboratory data collected cannot be routinely validated or standardized due to the limitations of a passive surveillance system,

VSD is a collaboration of researchers from private and public sectors who study the safety of vaccines administered to all age groups by conducting epidemiological research using a large linked database. Although, the collective VSD data includes active surveillance of approximately seven million people (2.5% of the total U.S. population), it does not investigate or manage vaccine adverse events on an individual level for the purpose of systematically collecting and evaluating those incidents.

Post-licensure surveillance must be augmented by innovative research that addresses the fundamental pathogenesis of adverse reactions to vaccine. To address these unmet research and surveillance needs, CDC established the Clinical Immunization Safety Assessment (CISA) Network in 2001. CISA unites CDC's expertise with highly regarded national experts in infectious disease, clinical investigation, vaccine studies, and public policy. CISA provides leadership, clinical research experience, and scientific credibility for the novel assessment and monitoring vaccine safety from scientific and clinical perspectives.

A comprehensive vaccine safety surveillance system requires closure of existing research and surveillance gaps that can identify at-risk groups. CISA addresses these gaps through research, which (1) standardizes collection of clinical data; (2) creates a centralized database (registry) to include clinical data, treatment regimes and patient outcomes; and (3) creates of a specimen bank (repository) to store biological specimens from patients who have experienced post-vaccination Adverse Events (AEs). Ongoing efforts include the following.

- **Establishing registry of clinically significant AEs and related clinical data and a repository of biological specimens from patients who have experienced serious post-vaccination AEs.** Adverse events following immunization may be a result of interacting host, pathogen and environmental factors. Because most serious AEs are relatively rare, standardized clinical assessment, treatment recommendations and understanding of the pathogenesis for adverse events following immunization (AEFIs) are lacking.

  The Registry / Repository may be used to guide the development of treatment plans and serve as sources of data and specimens for future research studies designed to test specific hypotheses regarding the causal relationship between AEs and particular vaccines. For example, specimens in the repository may be used for future studies of cytokine responses, pharmacodynamics and the relationship of vaccine strain subtypes, gene expression profiles and gene polymorphisms to AEs from the vaccine under question. Specimens stored in the repository will be linked to related

epidemiologic data (demographic, clinical, exposure history and risk factor). The registry and repository will be valuable resources for future, hypothesis-driven research studies. Such studies will increase our understanding of AEFIs and guide development of re-immunization guidelines for patients who might benefit from further vaccinations but may be at higher risk for AEs.

- **Nested case control study entitled, "Smallpox Vaccination and Myo/pericardial Injury / Inflammation"** in collaboration with CISA network, Department of Defense, University of Washington and National Institutes of Health. Objectives are 1) To determine genetic diversity associated with smallpox vaccination and myopericardial injury and disease (clinically overt or subclinical disease).This will include identifying and evaluating polymorphisms in specific genes that may be associated with risk.identify statistically significant association(s) between risk for clinically overt smallpox vaccine-associated myocarditis and polymorphisms in immune response genes, and 2) To correlate the discovery and understanding of the genetic diversity with the immune response in order to improve the ability to identify risk factors for adverse clinical responses to smallpox vaccine.determine whether this association(s) is also true for subclinical cases.

- **The Vaccine Safety Datalink (VSD) project is currently conducting a retrospective cohort study entitled, "***Is there a genetic predisposition to developing Rheumatoid Arthritis in persons receiving Hepatitis B vaccine*?" This study goal is to determine whether exposure to Hepatitis B vaccine increases the risk of Rheumatoid Arthritis (RA) in adults. Using the automated data from the VSD, clinic visits for health plan members enrolled during the study period are collected. Exposure status is determined from the NCK immunization database and is supplemented by chart reviews and a telephone interview sample. Charts are reviewed for all potential cases to determine case status, intervals between first automated clinic visit for RA and actual onset date as well as exposure status. This VSD adverse events study is still ongoing.

  The RA cases identified are invited to participate in a genetic substudy of the aforementioned VSD study. This substudy seeks to determine if there is a genetic predisposition to developing Rheumatoid Arthritis (RA) following Hepatitis B vaccine (HBV). Consenting patients provide blood specimen and complete a questionnaire on their medical history, exposure status, and demographics. Blood specimens are sent to CDC for HLA typing where mononuclear cells will be isolated and DNA extracted. This genetics study is currently in the data collection phase.

Infrastructure

- **Setting up a high throughput genotyping capacity** (CDC/NCID/Scientific Resources Program (SRP)): The mapping of the human genome has led to many changes in the field. This advance has revolutionized the genotyping technology leading to high throughput genotyping as a standard tool for the future. High

throughput genotyping methods are robust, rapid and cost effective. Currently, limited infrastructure and personnel resources for large-scale high throughput genotyping facility exist in CoCID. Hiring the Career Development Fellow for the GWG has facilitated the group's ability to focus on developing infrastructure for high throughput genotyping. This includes exploring ABI SNPlex and MassARRAY™ system from Sequenom for high throughput genotyping.

- **Candidate gene selecting tool and genotyping database:** The CoCID GWG and OGDP are currently developing tools and materials to assist CDC programs in identifying candidate genes and gene pathways to assess the role of genomics in epidemiologic studies and acute public health investigations. The end result of this project will be a suggested approach to identifying candidate genes, criteria for inclusion, pathways of interest, priorities, and an initial list of genes single nucleotide polymorphisms (SNPs). To reach this goal, more focused research and consolidation of data and literature must be performed. To facilitate the needs of large-scale genotyping, efforts on developing genotyping databases across numerous projects are also made.

- **Global amplification of sense RNA: a novel method to archive and replicate RNA for microarrays** (CDC/NCID/DVRD): Biorepositories from population-based epidemiological studies are increasingly important resources for cancer biomarker discovery and validation. The finite amount of mRNA available from each sample and the viable nature of RNA during long-term storage significantly limit the number of studies that can be supported by these priceless collections. Several clever amplification approaches have been developed that allow gene expression profiling to be performed on even the few cells available from microdissection. However, these methods do not address the problem of the viability of RNA during long-term storage, nor the compatibility of the RNA with the wide variety of gene expression profiling platforms and approaches.

  Researchers developed a novel procedure to amplify mRNA into sense RNA (sRNA) [27]. The cDNA intermediate forms a stable biorepository capable of regenerating the complex mRNA profile in the original sample. Because the amplified RNA is in the 5'→3' orientation, it is a synthetic equivalent of mRNA and can be used as the template for any platform or approach to gene expression analysis. The procedure exploits the template-switching activity of reverse transcriptase to incorporate RNA polymerase binding site upstream of single stranded cDNA. sRNA prepared from RNA extracted from human cell lines, tissues, blood and fixed exfoliated cervical cells performed satisfactorily in microarray and real-time RT-PCR assays. sRNA preparation preserved the relative differences in mRNAs spiked at concentrations spanning 5 orders of magnitude (0.00001-0.1%). This reflects the high fidelity of sRNA for mRNA species present at concentrations as low as 0.3 copies/cell. From 40 nanograms of input RNA, one round of sRNA amplification resulted in an approximately million-fold amplification of mRNA, and yielded highly reproducible microarray results (Pearson correlation coefficient 0.97 using MWG 10K arrays and

Resonance Light Scattering technology for signal detection). Global amplification of sRNA should find applications in the RNA archiving of population-based biorepositories for biomarker discovery and validation studies for detection and stratification of cancers and other diseases. This RNA amplification and archiving protocol may also find applications in pathogen discovery projects.

- **Beta-test site for emerging genomic technologies** (NCID/DVRD/VEHB): Viral Exanthems and Herpes Virus Branch served as the beta-test site for two important and emerging genomic technologies. Both technologies aim to improve the reproducibility, sensitivity and specificity of glass microarray hybridizations. In collaboration with Ventana Medical Systems (http://www.ventanamed.com/, Tucson, AZ), VEHB has evaluated a robotic platform for automated microarray hybridization. This automated platform enables simultaneous hybridization of 20 microarrays. Consultations are underway with Ventana Medical Systems on use of the automated hybridization system for high-throughput *in situ* hybridization.

  In collaboration with Genicon Sciences (recently purchased by Invitrogen), VEHB evaluated resonance light scattering gold particles (RLS System) for improved signal detection on microarrays and was the first to demonstrate that as little as 500 ng of total RNA can be labeled with biotin hybridized to glass microarrays and detected with resonance light scattering gold particles resulting in several publications (Ojaniemi H et al, 2003; Habis et al, 2004).

- **Automated tools for genomic** research (CDC/NCID/DVRD and the Scientific Resources Program (SRP)): In 1999, a high-performance microarrayer for spotting biological samples into arrays and a laser-scanner for microarray image detection were purchased. This equipment was transferred to the SRP Core Facility to allow access of custom microarrays to any interested NCID user. Today, the SRP Core Facility produces custom microarrays for many NCID intramural investigators and several extramural investigators (e.g., Emory University).

- **Development of a gene expression microarray database for flexible data processing, analysis and archiving (CDC/NCID/DVRD):** For the past 5 years, VEHB has custom-built a microarray database and analysis package, called CDC MAdB. This custom microarray database allows integration of microarray gene expression data with epidemiologic and clinical datasets (non-gene data). Large-scale analysis of gene expression data obtained from microarrays is at the cutting edge of biomedical research. There are many commercial and academic microarray analysis packages available. Most are designed for analysis of specific microarray formats and none have integrated epidemiologic and clinical data. Because of this shortcoming, CDC developed a MAdB to serve as a data warehouse and also a flexible and powerful front-end microarray data processing tool. Because of CDC MAdB's flexible and modular design, it is amenable to integration of non-gene data. Integration of epidemiologic and clinical common data elements with microarray expression data will allow for selection, grouping and/or stratification of microarray

data based on variables deemed to be important risk factors of the illness. Data integration will also allow for exploration of clinical and epidemiologic variables that may have a novel affect on gene expression. Operationalization of this integration for one disease will provide the model integration template for other diseases. CDC MAdB was designed to allow for global microarray data exchange and sharing and is available to all CDC microarray users.

Training

- **Leadership Management Institute Training:** CoCID GWG members have been selected to receive CDC's Leadership and Management Institute training (2004-2005). This program will provide training for improving team effort, nurture leadership qualities, and promote collaborative spirit. As part of this training, members are working on developing a strategic plan for CoCID to integrate genomics research for public health application.

- **Consultation:** Provided consultation for various investigators on study design issues, methodological aspects of genotyping consent form development, and data analysis.

- **Seminars and journal club:** Organized select seminars and journal clubs to highlight significant new developments in this field.

Major conferences

- **Integrating Disparate Data to Simulate Lymphocyte Function (CDC/NCID/DVRD):** The CDC CFS Research Program sponsored a workshop, *Integrating Disparate Data to Simulate Lymphocyte Function*, at the Banbury Center, Cold Spring Harbor Laboratory, on September 19-22, 2004. The objective was to discuss current knowledge concerning lymphocyte function and to identify means by which computational modeling could be used to understand how this complex biologic system functions in persons with CFS. The workshop brought together experts in immunology, molecular biology, computer sciences, and molecular modeling. Specific aims were to 1) define the types of laboratory and clinical data involved in the current concept of lymphocyte function in normal and abnormal states; 2) present approaches for integrating genomic, proteomic, clinical, and epidemiologic data in such models; and 3) define the level of abstraction and types of assumptions necessary to create the next generation of molecular models.

- **C^3: The CFS Computational Challenge:** The CFS Research Program is hosting a CFS Computational Challenge (C3). The results of this challenge will help elucidate the pathophysiology of CFS, identify markers of CFS (or subsets of CFS) that may be useful for effective diagnosis and treatment of CFS, and formulate hypotheses to test in future studies.

The CFS Research Program has conducted a 2-day in-patient clinical study of 227 persons identified with CFS, other unexplained chronically fatiguing illnesses, and randomly selected non-fatigued controls from the general population of Wichita, Kansas. Subjects were carefully evaluated medically and psychiatrically. Investigators obtained measurements of their neuroendocrine status, cytokine profiles, sleep, cognitive function, and evaluated their lifetime stress history and coping mechanisms. To classify parameters of CFS, they evaluated disability, fatigue characteristics, and the impact of cumulative symptoms. Finally they measured expression levels of 40,000 genes in peripheral blood cells.

The challenge will engage computer scientists, bioinformaticians, statisticians, biologists and clinicians to mine biologically and clinically meaningful information relevant to diagnosis and therapeutic intervention of CFS from the Wichita Clinical Study data set. Participants will be organized into teams. The challenge will begin with a 1-day workshop where an introduction to CFS will be given along with a description of the dataset for C^3. Each teams results will be presented as a paper and judged for biological and mathematical soundness by an expert panel. All participants will present their results at the Banbury Center, Cold Spring Harbor Laboratory, September 18-21, 2005.

- **CoCID GWG contributed to and attended CDC sponsored conference**: *Incorporating Genomics into the Acute Public Health Response*. Atlanta, GA - May 2004.

Future directions

CDC is in a unique position to integrate and advance host genomics research by virtue of its access to valuable population based cohorts, outbreak investigations, the repository of specimens associated with vaccine adverse events, and expertise of epidemiology and laboratory science. Although the concept of targeting genetically at risk populations for appropriate interventions to improve their health is not new (e.g.: sickle cell patients owing to mutations in the hemoglobin gene are targeted for antibiotic propohylaxis and this has led to improved life span for these patients), recent advances in genomics research provides novel tools for identifying genetically at risk populations for various risks of infectious diseases and vaccine adverse events. Thus, CDC can promote the research that will help to identify genetic risk factors associated with various outcomes of infectious diseases and to gather knowledge from the published studies to identify genetically at risk individuals for better prevention efforts. In addition, this field is likely to open up ways to identify genetically at risk individuals for vaccine adverse events and therapeutic failure especially for chronic infectious diseases such as AIDS and TB. CDC will have to develop ways to integrate these developments for prevention efforts. In order to accomplish this goal increased efforts are needed in the following areas.

- Resources: Laboratory infrastructure not only to conduct intramural genomics research but also establishing core competencies in evaluating reliable technologies for accurate genotyping of U.S. population. Dedicated funding will be required to support genomics research. Awards by competitive selection for the best proposals for integrating new genomics research of programmatic relevance into ongoing epidemiologic projects from investigators within the three CoCID Centers. Awards are planned annually, and will fund specimen collection and testing costs associated with adding a genomics component to ongoing epidemiologic field studies.
- Personnel: Rapid development in genomics research has led to novel and complicated laboratory methods and data analysis. In addition, various bioinformatics tools have become available for data analysis and information extraction. It is important to bring appropriately trained individuals in these areas to the CDC work force. Training in genetic epidemiology will also be relevant in the future.
- Training: CDC epidemiologists and laboratory scientists will benefit from dedicated training funds to update their laboratory training in new areas and to participate in scientific meetings to increase their awareness of latest development in this field. Research Fellowships can also be helpful to bring in young talent interested in this area of research.

Key scientific observations (Publications from 2000-2004)

1) Aidoo M, McElroy PD, Kolczak MS, Terlouw DJ, ter Kuile F, Nahlen BL, Lal A, Udhayakumar V. Tumor necrosis factor promoter variant 2(TNF-2) is associated with preterm delivery, infant mortality, and malaria morbidity in western Kenya. Asembo Bay Cohort Project IX. **Genetic Epidemiology 2001** 21: 201-211

2) Aidoo M, Terlouw DJ, Kolczak MS, McElroy PD, ter Kuile F, Kariuki S, Nahlen BL, Lal A, Udhayakumar V. Protective effects of the sickle cell trait gene against malaria morbidity and mortality. **Lancet** 2002 359: 1311-1312

3) Aikhionbare FO, Hodge T, Kuhn L, Bulterys M, Abrams EJ, Bond VC. Mother-to-child discordance in HLA-G exon 2 is associated with a reduced risk of perinatal HIV-1 transmission. **AIDS** 2001 15:2196-8.

4) Bridges SL, Jenq G, Moran M, Kuffner T, Whitworth WC, McNicholl J. Single-nucleotide polymorphisms in tumor necrosis factor receptor genes. Definition of novel haplotypes and racial/ethnic differences. **Arthritis Rheum** 2002 46:2045-50.

5) Brouwer KC, Lal RB, Mirel LB, Yang C, Van Eijk AM, Ayisi J, Otieno J, Nahlen BL, Steketee R, Lal AA, Shi YP. Fc(receptor IIa (CD32) polymorphism is associated with perinatal transmission of HIV-1 in western Kenya. **AIDS** 2004 18: 1187-1194.

6) Brouwer KC, Lal AA, Mirel LB, Lal RB, Van Eijk AM, Ayisi J, Otieno J, Steketee R, Nahlen BL, Shi YP. Association between genetic polymorphism in Fc

receptor IIa for IgG (Fc(RIIa) and placental malaria in western Kenya. **J Infect Dis** 2004 196: 1192-8

7) Campbell C, Vernon SD, Nisenbaum R, Reeves WC, Unger ER. Daily and weekly periodicity of peripheral blood gene expression. **Disease Markers** 2002 18:201-206

8) Correa PA, Whitworth WC, Kuffner T, McNicholl J, Anaya JM. HLA-DR and DQB1 gene polymorphism in the northwestern Colombian population. **Tissue Antigens** 2002 59:436-9.

9) Downer MV, Hodge T, Smith DK, Qari SH, Schuman P, Mayer KH, Klein RS, Vlahov D, Gardner LI, McNicholl JM. Regional variation in CCR5-Δ32 gene distribution among women from the US HIV epidemiology research study (HERS). **Genes and Immunity** 2002 3:295-8.

10) Hader SL, Hodge TW, Buchacz KA, Bray RA, Padian NS, Rausa A, Slavinski SA, Holmberg SD. Discordance at human leukocyte antigen-DRB3 and protection from human immunodeficincy virus type 1 transmission. **J Infect Dis** 2002 185:1729-35.

11) Habis AH, Vernon SD, Lee, Unger ER. Optimization of RNA extraction from exfoliated cervical cells. **Cancer Epidemiology, Biomarkers and Prevention** 2004 13:492-6

12) Haukim N, Bidwell JL, SmithAJP, Keen LJ, Gallagher G, Kimberly R, Huizinga T, McDermott MF, Oksenberg J, McNicholl J, Pociot F, Hardt C, Alfonso SD. Cytokine gene polymorphism in human disease: on-line databases, supplement 2. **Genes and Immunity** 2002 3:313-30.

13) Hobbs MR, Udhayakumar V, Levesque MC, Booth J, Tkachuk AN, Pole A, Roberts JM, Kariuki S, Nahlen BL, Mwaikambo ED, Lal AA, Granger DL, Anstey NM, and Weinberg JB. A new NOS2 promoter polymorphism associated with protection from severe malaria in children from Tanzanian and Kenyan children. **Lancet** 2002 360:1468-75.

14) Kuffner T., Whitworth W., Jairam M., and McNicholl J. HLA class II and TNF genes in African Americans from the southeastern United States: Regional differences and allele frequencies. **Human Immunol** 2003 64:639-47.

15) Kuhn L, Abrams EJ, Palumbo P, Bulterys M, Aga R, Louie L, Hodge T, and the Perinatal AIDS Collaborative Transmission Study (PACTS). Maternal verses paternal inheritance of HLA class I alleles among HIV-infected children: consequences for clinical disease progression. **AIDS** 2004 18:1281-1289.

16) Lingappa J, Kuffner T, Tappero J, Whitworth W, Mize A, Kaiser R, McNicholl J. HLA-DQ6 and ingestion of contaminated water: possible gene-environment interaction in an outbreak of Leptospirosis. **Genes and Immunity** 2004 5: 197-202

17) McNicholl JM, M. Downer, V. Udhayakumar, CA. Alper, and DL Swerdlow.

Host-Pathogen interactions in emerging and re-emerging infectious diseases: A genetic perspective of tuberculosis, malaria, human immunodeficiency virus infection, hepatitis B, and cholera. **Annual Review of Public Health**, 2000 21:15-46.

18) McNicholl JM, Downer MV, Aidoo M, Hodge T, Udhayakumar V. Public health assessment of genetic susceptibility to infectious diseases: Malaria, TB, and HIV**.** In **Genetics and Public Health in the 21st Century: Using Genetic Information to Improve Health and Disease**. Khoury MJ, Burke W, Thomson EJ. (Eds), Oxford University Press, 2000 pages: 173-202.

19) McNicholl JM, Promadej N. Insights into the role of host genetic and T-cell factors in resistance to HIV transmission form studies of highly HIV-exposed Thais. **Immunologic Res** 2004; 29:161-74.

20) McNicholl JM, Lal RB, Kaslow R, HIV Infection, Genetics. In **Encyclopedia of the Human Genome**. Nature Publishing Group, London, England, (In Press).

21) Nicholson AC, Unger ER, Mangalathu R, Ojaniemi H, Vernon SD. Exploration of neuroendocrine and immune gene expression in peripheral blood mononuclear cells. **Brain Res Mol Brain Res** 2004 129:193-7.

22) Nguyen L, Chaowanachan T, Vanichseni S, McNicholl JM, Mock PA, Nelson R, Hodge TW, van Griensven F, Choopanya K, Mastro TD, Tappero JW, Hu DJ. Frequent human leukocyte antigen class I alleles are associated with higher viral load among HIV type 1 seroconverters in Thailand. **J Acquir Immune Defic Syndr** 2004 37:1318-1323.

23) Nguyen L, Li M, Chaowanachan T, Hu DJ, Vanichseni S, Mock PA, van Griensven F, Martin M, Sangkum U, Choopanya K, Tappero JW, Lal RB, Yang C. CCR5 promoter human haplogroups associated with HIV-1 disease progression in Thai injection drug users. **AIDS** 2004 18:1327-1333.

24) Rajeevan MS, Dimulescu IM, Vernon SD, Verma M, Unger ER. Global amplification of sense RNA: A novel method to replicate and archive mRNA for gene expression analysis. **Genomics** 2003 82:491-497.

25) Ojaniemi H, Evengard B, Lee DR, Unger ER, Vernon SD. Impact of RNA extraction from limited samples on microarray results. **BioTechniques** 2003 35:968-973.

26) Ranamukhaarachchi DG, Unger ER, Vernon SD, Lee DR, Rajeevan MS. Gene expression profiling of dysplastic differentiation in cervical epithelial cells harboring human papillomavirus 16. **Genomics**, in press.

27) Rajeevan MS, Vernon SD, Taysavang N, Unger ER. Validation of array-based gene expression profiles by real-time (kinetic) RT-PCR. **J Mol Diag** 2001 3:26-31.

28) Rajeevan MS, Ranamukhaaraachchi D, Vernon SD, Unger ER. Use of real-time quantitative PCR to validate the results of cDNA array and differential display-PCR technologies. **Methods in Enzymology** 2001 25:443-451

29) Ranamukhaaraachchi D, Rajeevan MS, Vernon SD, Unger ER. Modifying Differential Display-PCR to detect relative changes in gene expression profiles. **Anal.Biochem** 2002 343-346

30) Shi, YP, Nahlen, BL, Urdahl, K, McElroy, P, Jacqueline MR, Lal, AA. Fc(receptor IIa (CD32) polymorphism is associated with protection of infants against high density *Plasmodium falciparum* infection: VI. Asembo Bay Cohort Project. **J Infect Dis** 2000 184: 107-11.

31) Steinau M, Unger ER, Vernon SD**,** Jones JF, Rajeevan MS. Differential display PCR of peripheral blood for biomarker discovery in chronic fatigue syndrome. **Journal of Molecular Medicine** 2004 82: 750-5.

32) Terlouw A. Aidoo M, Udhayakumar V, Kolczak, Oloo AJ, Kager PA, Lal AA, Nahlen BL and ter Kuile F. Increased Efficacy of Sulfadoxine-pyrimethamine in the treatment of uncomplicated falciparum malaria among children with sickle cell trait in western Kenya. **J Infect Dis** 2002 186: 1661-8.

33) Vernon SD, Unger ER, Rajeevan M, Dimulescu IM, Nisenbaum R, Campbell CE. Reproducibility of alternate probe synthesis approached for gene expression profiling with arrays. **J Mol Diag** 2000 2:124-127.

34) Vernon SD, Unger ER, Dimulescu IM, Rajeevan MS, Reeves WC. Utility of the Blood for Gene Expression Profiling and Biomarker Discovery in Chronic Fatigue Syndrome. **Disease Markers** 2002 18:193-199

35) Vernon SD, Shukla SK, Conradt J, Unger ER, Reeves WC. Analysis of 16S rRNA gene sequences and circulating cell-free DNA from plasma of chronic fatigue syndrome and non-fatigued subjects. **BMC Microbiology** 2002 2:39

36) Whistler T, Jones JF, Unger ER, Vernon SD. Exercise responsive genes measured in peripheral blood of women with Chronic Fatigue Syndrome and matched control subjects. **BMC Physiology** in press.

37) Whistler T, Vernon SD, Unger ER, Nisenbaum R, Reeves WC. Empirically stratifying chronic fatigue syndrome by integration of gene expression, symptom and epidemiologic data. **J Translational Medicine** 2003 1:1-10.

38) Weiner M, Burman W, Vernon A, Benator D, Peloquin CA, Khan A, Weis S, King B, Shah N, Hodge T; Tuberculosis Trials Consortium. Low isoniazid concentrations and outcome of tuberculosis treatment with once-weekly isoniazid and rifapentine. **Am J Respir Crit Care Med**. 2003 1167: 1341-7.

39) Weiner M, Benator D, Burman W, Peloquin CA, Khan A, Avernon A, Jones B, Silva-Trigo C, Zhao Z, Sterling TR, Kernodle D, Weis S, Lahart C, Hodge T, Tuberculosis Trials Consortium. The association between acquired rifamycin resistance and the pharmacokinetics of rifabutin and isoniazid among patients with HIV-related tuberculosis. **Clin Inf Dis** submitted.

40) Yang C, Li M, Limpakarnjan L, Young NL, Hodge T, Butera .T, McNicholl JM, Mastro T., and Lal RB. Polymorphisms in the CCR5 coding and noncoding

regions among HIV type 1-exposed persistently seronegative sex-workers from Thailand. **AIDS Res Human Retroviruses** 2003 19:661-665.

Environmental & Occupational & Injury Prevention Coordinating Center

Top priorities:

Newborn Screening Program

Newborn screening, a major public health responsibility, is the largest genetic testing effort in the nation. Effective screening of newborns begins with the accuracy and reliability of the screening test, followed by diagnostic studies and treatment. For more than 25 years, the Division of Laboratory Sciences (DLS) at CDC's National Center for Environmental Health with its cosponsor, the Association of Public Health Laboratories, has conducted research on materials development and assisted laboratories with quality assurance (QA) for these screening tests. CDC's Newborn Screening Quality Assurance Program (NSQAP) is designed to help screening laboratories achieve excellent technical proficiency and maintain confidence in their performance while processing large volumes of specimens daily.

The NSQAP provides QA services for congenital hypothyroidism, phenylketonuria, galactosemia, congenital adrenal hyperplasia, maple syrup urine disease, homocystinuria, biotinidase deficiency, galactose-1-uridyltransferase (GALT) deficiency, cystic fibrosis, and hemoglobinopathies (including sickle cell disease). QA services are also provided for fatty acid oxidation and organic acid disorders detected by tandem mass spectrometry (MS/MS). The program prepares and distributes more than 500,000 dried blood spots (DBS) to national laboratories every year. These materials must simulate as closely as possible the actual specimens for the assay systems and are certified for homogeneity, accuracy, stability, and suitability for all assays from different commercial sources. NSQAP is the sole source of these comprehensive QA services worldwide.

National Health and Nutrition Examination Survey (NHANES) DNA Bank

The National Health and Nutrition Examination Survey (NHANES) is a program of periodic examination surveys conducted by CDC's National Center for Health Statistics (NCHS) that provides nationally representative information on the health and nutritional status of the U. S. population. CDC's NHANES DNA bank was conceived by scientists in NCEH in 1988 in response to the newly established Human Genome Project to map and sequence the human genome and to the emergence of new technologies such as the polymerase chain reaction (PCR) and automated DNA sequencing.

The intent in establishing CDC's NHANES DNA bank was to create a representative genetic sample of the U.S. population for intramural and extramural research. Immortalized cell lines were established because current technology would require large amounts of DNA, and these cell lines provided a virtually unlimited source of DNA. The DNA Bank consists of 8,000 immortalized B-lymphocyte cell lines. Approximately 14,000 white cell samples that were not immortalized were stored in liquid nitrogen and

could be used as a back up for immortalization or as a direct source of DNA. In addition, approximately 12,000 blood clots have also been frozen that also could be used as a source of DNA.

Prevalence of gene variants that code for enzymes involved in nicotine and carcinogen metabolism

Cotinine, a metabolite of nicotine, is a reliable measure of exposure to cigarette smoke. People convert the harmful chemicals in cigarette smoke into potentially harmful substances differently, depending on their genetic make-up.

To address the question of whether or not genetically-based differences in nicotine metabolism account for differences in cotinine levels among individuals or groups, the Molecular Biology Branch of NCEH's Division of Laboratory Sciences will genotype approximately 7,300 samples from the NHANES III DNA Bank for 14 polymorphic genes which code for enzymes that are associated with the metabolism of nicotine and other toxicants in cigarette smoke or are associated with smoking characteristics. Each of the proposed gene variants have been implicated in smoke-related and other cancers or in smoking habits among people.

This study will compare the body burden of cotinine and other parent chemicals from cigarette smoke among people and among racial or ethnic groups by genotype. In addition, the number of cigarettes smoked per day and smoking history will also be examined.

This information will help explain the potential benefits and risks associated with having one or more of the gene variants as these variants relate to the effects of smoking on health, smoking habits (including the ability of people to successfully stop smoking), and to targeted screening and intervention programs. The results of this study will help explain the documented differences in cotinine levels in the U.S. population and specific ethnic groups in that population

Churchill County Leukemia Cluster Investigation

As part of its response to an investigation of a cluster of cases of acute lymphocytic leukemia among children in Churchill County, Nevada, the Nevada State Health Division (NSHD) requested technical assistance from the Centers for Disease Control and Prevention (CDC). The purpose of the subsequent collaborative investigation was to conduct a cross-sectional exposure assessment to identify contaminants unique to the Churchill County community.

One part of that investigation involves using DNA obtained from study participants to look for differences among genes that may affect susceptibility to leukemia. For example, a gene that directs the production of the protein called MTHFR has at least two

different forms in the United States population. MTHFR is important in the body's production of folic acid, a vitamin needed to prevent birth defects and that may also be

important in preventing cancer and heart disease. One form of MTHFR, although it does not cause birth defects, can increase the risk for birth defects, such as spina bifida, in the presence of environmental factors and may be related to the risk for cancer.

Some scientific evidence shows that having one of these forms of MTHFR may protect people against developing acute lymphoblastic leukemia. MTHFR is only one of about 15 genes that CDC plans to study for this type of variation among people that, in the presence of certain environmental exposures, such as exposure to pesticides, may affect the risk for leukemia. This work is being done to determine whether or not the Churchill County children with leukemia or their family members are more likely to have one form of the gene, whereas children and their families in the control group are more likely to have another form of the gene.

The National Birth Defects Prevention Study Centralized Laboratory

The National Birth Defects Prevention Study (NBDPS) is one of the largest case-control studies of birth defects ever conducted in the United States. The goal of this ongoing multi-center study is to identify environmental and genetic risk factors for birth defects. Information on environmental risk factors is collected through a maternal interview, and DNA is collected from the infant and both parents for evaluation of genetic risk factors.

Cheek cell (buccal) specimens are obtained because of their low cost and ease of use for sampling infants without the direct involvement of health-care workers. However, unlike whole blood, buccal specimens yield lower amounts of DNA that is often of poorer quality. Because of these limitations, the best DNA banking and quality-control practices must be employed in establishing buccal DNA collections.

Initially, each Center was responsible for DNA isolation and quality control. However, the laboratory proficiency testing program demonstrated that considerable variation existed in the technical ability of the laboratory staff of the various centers participating in the study, and there was also variation in laboratory results. Moreover, standardization of methods for isolating DNA and for ensuring quality control was lacking.

To ensure the integrity of the biological data, the decision was made to consolidate the biologic processing to a single laboratory dedicated to this activity. In 2003, the NBDPS Centralized Laboratory was established in the Molecular Biology Branch of the Division of Laboratory Sciences, NCEH. Key features of this laboratory include a staff with extensive expertise in DNA banking and performance of laboratory operations that satisfy CLIA regulations, an extensive and integrated data-management system, and an array of state-of-the-art instrumentation. This activity is further enhanced by an active research program in emerging laboratory technologies.

Genetics of Kidneys in Diabetes (GoKinD) Study

Diabetes Mellitus is the leading cause of treated end-stage renal disease (ESRD), accounting for 43% of new cases. About 40% of those with type 1 diabetes (T1D) develop severe kidney disease and ESRD by the age of 50. The health-care expenditures for the renal complications of diabetes were $1.9 billion for 2002. This burden of illness and cost reinforces the need for strategies to prevent the development of renal disease in those who are at risk.

The role of genetic risk factors in the etiology of renal disease is supported by 1) evidence for family clustering of renal disease in a wide variety of ethnic groups with various predisposing conditions; 2) the fact that hyperglycemia alone is not sufficient to cause renal disease; and 3) family histories that indicate that hypertension, cardiovascular disease, and possibly insulin resistance predispose to nephropathy. In various populations, renal disease clusters in families independent of predisposing conditions, such as diabetes, lupus erythmatosis, or HIV infection, indicating that genetic risk factors for diabetic renal disease may differ from those for diabetes itself.

The GoKinD Study is a collaboration of the Juvenile Diabetes Research Foundation (JDRF); its Coordinating Center, consisting of the Joslin Diabetes Center and George Washington University Biostatistics Center; and CDC. GoKinD is designed as a collection of samples from clearly delineated case participants with type 1 diabetes who also have renal disease and from control participants with T1D who do not have renal disease but who excrete very small amounts of albumin. The intent of the collaboration is to facilitate the study of genetic risk factors for diabetic renal disease. Half of the participants were recruited by the Joslin Diabetes Center and half by clinical centers associated with the George Washington University Biostatistics Center. The clinical centers and the University of Minnesota provided extensive clinical data that characterized these participants, and the Molecular Biology Branch in the Division of Laboratory Sciences at NCEH provided genetic data on major T1D risk factors. CDC is also conducting additional research on the collection. Recruitment is complete, and sample distribution to other approved researchers will begin in 2005.

Case-control study of environmental exposures and genetic susceptibility among individuals with multiple sclerosis living in three geographic areas

Although multiple sclerosis (MS) is the most common neurologic disease disabling young adults in the United States, the cause of this disease is unknown. Evidence indicates that it is a complex disease with a multifactorial etiology determined by both environmental factors and genetic susceptibility. However, previous research examining potential causes of MS has focused on either the role of environmental exposures or the role of genetic factors but not on a combination of the two. This study examines the joint role of selected environmental exposures and candidate genes as potential risk factors for MS.

In November 2004, study investigators began enrolling 1500 individuals (500 case and 1000 control subjects). Case subjects will include individuals who have been diagnosed with MS and were identified through the prevalence studies that were recently completed in Ohio, Missouri, and Texas by local and state health departments under a cooperative agreement with ATSDR. Control subjects will be selected from patients who attended the same neurologist's office from which the cases arose. The diagnoses for control selection will be limited to those for which there is no prior indication of a relation with the exposures under study.

Each participant in this study will complete a questionnaire and provide a blood sample. Information collected in the questionnaire will include exposure to potential environmental factors thought to be associated with MS such as exposure to heavy metals, solvents, and other toxic chemicals, as well as a complete residential history, school history (location of), occupational history, and hobbies/lifestyle exposures. Investigators also will collect information on family medical, history, reproductive, and smoking histories. Case subjects will complete one extra section about the course of their disease. A blood sample will be obtained from all participants that will be genotyped for specific genes thought to be associated with MS, including human leukocyte antigen, immunoglobulin heavy chain, T-cell receptor, tumor necrosis factor, and myelin basic protein. A metabolomic analysis will also be conducted on a sub-set of blood samples.

Major accomplishments, 2004

Newborn Screening Program

The QA services provided by CDC's Newborn Screening Program primarily support newborn screening tests performed by state laboratories; however, the program also accepts other laboratories and international participants. All laboratories in the United States that screen newborns participate voluntarily in NSQAP. NSQAP participants can use the program's Web site (www2.cdc.gov/nceh/newbornScreening) to report their data and receive performance evaluations. Over the last few years, NSQAP has grown substantially, both in the number of participants and in the scope of global participation. In 2004, 400 laboratories in 60 countries (at least one laboratory per country) were active program participants. NSQAP currently provides QA services for 35 disorders.

In January 2003, the program began distributing five-specimen panels for Type 1 Diabetes composed of dried blood spots from a validated specimen library of sequenced patient samples. Four research laboratories that do population-based testing participate in the pilot PT to assure comparability of data from the different research sites. The program has now added a DNA confirmatory testing component to the testing materials for cystic fibrosis, producing our second phenotype/genotype combination DBS. Additionally, NSQAP is investigating the development of specimens for detecting toxoplasmosis antibodies in DBS using serum from infected individuals and is conducting pilot testing to determine the feasibility of using these specimens to establish a proficiency testing program.

National Health and Nutrition Examination Survey (NHANES) DNA Bank

In 1999, DNA was made available to the research community. The Division of Laboratory Sciences, in collaboration with the National Center for Chronic Disease Prevention and Health Promotion (NCCDPHP), conducted a pilot study to assess the prevalence and penetrance of mutations related to hereditary hemochromatosis. This disease is mainly associated with mutations in a gene called *HFE*, which helps regulate the amount of iron absorbed from food. The defect can result in iron overload that can lead to organ damage. If a diagnosis is made early in the course of disease, organ failure can be prevented by inducing a mild anemia through the periodic removal of blood. Hereditary hemochromatosis is one of the most common genetic disorders in the United States. It most often affects Caucasians of Northern European descent, although other ethnic groups are also affected. The pilot study provided the first estimate of the prevalence of these mutations in a representative sample of the United States population. This information is for making policy decisions about screening health populations for hereditary hemochromatosis.

DLS and NCCDPHP collaborators published the first studies using NHANES III DNA bank samples, the aforementioned work related to prevalence of mutations that cause hereditary hemochromatosis, and one on the association of the HFE mutations with levels of iron in the blood. The results of this study show that although the presence of the mutations is associated with elevated iron levels, a substantial proportion of people with the mutations do not have high levels of iron. The study reaffirms the need for additional information about the risk of iron overload and chronic disease associated with HFE mutations, as well as other genetic and environmental factors that modify this risk in order to make informed decisions regarding genetic screening for hemochromatosis.

Steinberg KK, Cogswell ME, Chang JC, Caudill SP, McQuillan GM, Bowman BA, Grummer-Strawn LM, Sampson EJ, Khoury MJ, Gallagher ML. The prevalence of C282Y and H63D mutations to the hemochromatosis (HFE) gene in the United States. *JAMA*. 2001;285(17):2216-2222.

Cogswell ME, Gallgher ML, Steinberg KK, Caudell SP, Looker A, Bowman B, Gunter E, Franks AL, Khoury MJ, Grummer-Strawn LM. The HFE genotype and transferrin saturation in the United States. Genetics in Medicine. 2003;5(4) 5(4):304-10.

NHANES 1999-2003 Survey:
The current NHANES survey began in 1999. We have obtained DNA from whole blood specimens obtained from 8,000 subjects aged 20 yrs and older. These additional 8,000 DNA samples will significantly expand the existing NHANES DNA bank and will be an important resource for intramural (CDC) and extramural genetic research programs.

Prevalence of gene variants that code for enzymes involved in nicotine and carcinogen metabolism

Major accomplishments for 2004 include the development of robust, high throughput assays which will be used to determine the genotypes of the selected variants present in the 14 polymorphic genes. Each assay is subjected to a rigorous validation process to verify the accuracy of the assay. Extensive quality control procedures were developed. Both of these components of the study will ensure that the highest quality of data is produced.

Churchill County Leukemia Cluster Investigation

MTHFR and several other genes have been tested for the presence of known DNA polymorphisms (gene variants) and analysis of the results is in progress. Elevated levels of several metals were found in the Churchill County population. Although these metals have not been associated with an increased risk of leukemia, several enzymes that may play a role in formation of blood cells are influenced by metals such as tungsten. The genes that produce three of these enzymes were studied in depth in order to discover all of DNA polymorphisms present in these genes. A number of new gene variants were identified and analysis of the results is in progress.

The National Birth Defects Prevention Study Centralized Laboratory

Over 5,000 specimens were received last year. Since its inception, the NBDPS Centralized Laboratory has received at least 6,000 buccal specimens, and 95.1 % of DNA samples have successfully passed the quality-control process. This high success rate will enhance the utility of this resource in investigations of genetic risk factors for birth defects.

Genetics of Kidneys in Diabetes Study (GoKinD)

GoKinD has recruited over 2,700 participants including 243 case subjects with two parents (trios), 283 control-subject trios, 559 case-subject singletons, and 565 control-subject singletons. In all, this collection includes 802 individuals with both T1D and renal disease and 848 individuals with T1D and no renal disease. It will also include 1650 individuals with T1D and 526 T1D trios. CDC authors have published four papers and made seven presentations at major meetings on GoKinD. A draft of the baseline paper for the study is complete and ready for clearance and submission, and analysis of the HLA data for DRB1, DQA1, and DQB1 and the insulin gene data is in progress.

CDC has completed one half of a pilot study of HLA matched case and control subjects using the Affymetrix 10,000 single nucleotide polymorphism micoarray to identify regions of the genome associated with renal disease. In addition, CDC is conducting a study of type 1 diabetes and renal disease candidate genes and is extending the HLA genotyping to DPB1 and HLA Class I B.

Future directions:
In addition to continued work on the programs and studies that are current top priorities in the Coordinating Center, the following future directions have been identified to meet priority needs.

Asthma

Asthma is a major public health problem that affected 20.3 million Americans in 2001. The estimated cost of treatment of those younger than 18 years of age is $3.2 billion per year. Asthma disproportionately affects low-income populations, minorities, and children living in inner cities. Currently there are no preventative measures or cure for asthma. Despite intensive investigation, the factors that confer susceptibility are not well understood, but it is clear that both environmental and genetic factors each account for about 50% of the risk. A better understanding of the interaction of environmental risk factors with genetic risk factors is crucial to developing prevention strategies and interventions to decrease the impact of this disease.

A case-control study of sufficient size involving patients drawn from representative populations with carefully described phenotypes of asthma, supporting clinical data, and well-documented environmental exposures would offer a valuable means of advancing this field. Such a study would allow genetic and environmental risk factors to be studied objectively to determine the relative size of their effects and how they interact.

Clusters

CDC has long recognized the need to respond to public concern about cancer clusters, while recognizing the low probability of finding causes of cancer in community clusters. Over a decade ago, CDC organized The National Conference on Clustering of Health Events and published the proceedings of that meeting (Neutra, 1990). CDC also released "Guidelines for Investigating Clusters of Health Events" (CDC, 1990), and variations of these guidelines are still used in cluster investigations. These deliberations also identified deficiencies including lack of documentation of exposures.

Today, CDC has added state-of-the-science laboratory methods for comprehensive investigations of cancer clusters. These methods have enabled CDC to document and assess environmental exposures by comparing them to exposures in the general population. CDC is also using molecular methods to study the effects of documented environmental exposures. For example, for one cluster of acute lymphoblastic leukemia in children, CDC documented excessive exposures to environmental chemicals and recommended steps to avoid those exposures. CDC then looked for genetic differences between case and control children that might make some more susceptible than others to the effects of these toxicants and possibly alter susceptibility to leukemia.

In the future, CDC will introduce molecular methods to study differences in susceptibility to environmental toxicants. Further, CDC will work to assure collection of important diagnostic information on molecular lesions in patient's bone marrow before the patients are treated with chemotherapy. These lesions, including translocations and mutations, have been reported to be associated with specific leukemias. This work will be done in an effort to look for common features that may add to evidence of a common exposure and perhaps someday help to identify "signature" molecular lesions for classes of exposure.

Newborn Screening

As states adopt new screening tests to detect additional diseases, DLS must maintain QA services for these new technologies. DLS plans to establish a counterpart of NSQAP, which will be called the Newborn Screening Translational Research Laboratory 1) to identify promising tests currently used in research settings that could be applied to public health newborn screening; 2) to develop laboratory methods for detecting in newborns medical conditions and risk factors that ultimately could benefit children's' health; 3) to evaluate new technologies, particularly high-throughput nanoscale methods, that could be applied to newborn screening; 4) to conduct pilot programs in collaboration with state newborn screening programs to assess and refine the practical application of new laboratory tests; 5) to develop QA/QC tools for new tests before they are implemented on a population basis; and 6) to assist epidemiologists in using tests on newborn DBS for population-based assessment of risk factors and disease prevalence.

Several ongoing activities will include developing QA/QC materials for disorders, such as genotyping DBS materials for cystic fibrosis; supporting research programs that use newborn screening to assemble cohorts for longitudinal studies on type 1 diabetes and related autoimmune disorders; developing a screening test for severe combined immunodeficiency (SCID); developing multiplexed ligand-binding assays to identify risk factors for neurobehavioral developmental disorders such as autism and Fragile X syndrome; and evaluating reverse microarray technology for high-throughput, low-cost genotyping.

Toxicogenomics: Use of gene expression arrays to document exposure to environmental toxicants with signature profiling of chemicals or to predict toxicity from these exposures.

Currently, scientists are doing basic research to identify phenotypic changes associated with the altered patterns of gene expression and to link those phenotypic changes to conventional measures of toxicity. Scientists must also define toxic doses and the latency period between expression of such signature genes and outcomes. This problem is especially difficult because of the multiple exposures that occur in communities. The work must first be done in animals with exposures to single chemicals.

NCEH is poised to translate the basic research findings of the National Center for Toxicology and others to define mechanisms of toxicity in order to predict the toxicity of environmental exposures. In cases where scientists identify signature molecular lesions or patterns of expression, CDC will carry out hypothesis-driven, epidemiologic research in exposed populations to confirm the predictive value of these expression profiles/lesions.

National Institute for Occupational Safety and Health

Top Priorities

NIOSH has three priorities in the area of genetics.

1. Conduct of research on the interaction of occupational exposures and genetics factors in the development of disease.

2. Use of genetics markers to assess health effects of occupational exposures

3. Development of guidance information for the use of genetics in occupational safety and health

Gene environment interactions
Despite the strong causal associations that have been detected in many occupational studies, there are differences in disease incidence between groups of workers that cannot be accounted for by differences in exposures, work practices, or lifestyle. Genetic polymorphisms are likely to be responsible for some of these differences. In the early history of occupational epidemiology, it was generally accepted that genetic influences were accounted for by controlling for confounding by race and gender. Today, as many occupational exposures are lowered, the importance of genetic information as a source of variability in risk estimates is increasing. This is not to imply that occupational etiologies will be replaced with genetic etiologies; rather, polymorphisms, which might modify exposure–disease associations, should be included as relevant variables in study design and analysis. Although the individual risk associated with a genetic polymorphism may be relatively low, the population-attributable risk may be large, thus indicating the public health importance of this research. Genetic research that evaluates the role of specific genes in occupational disease or injury contributes to understanding the mechanisms and helps to disentangle which causal factors may pose the greatest threats to workers. NIOSH has looked specifically at gene environment interactions for chronic beryllium disease, Parkinson disease, and brain cancer.

Gene-environment interactions in chronic beryllium disease are being investigated in a case control study of beryllium manufacturing workers. This study is superior to previous studies in its improved exposure assessment, its increased size and because it represents a longitudinal follow-up. Previous studies have identified certain common genetic variants of a major histocompatibility complex gene, HLA-DPB1, to be associated with susceptibility to chronic beryllium disease (CBD). Many research questions remain of the more than 100 known HLA-DPB-1 variants, precisely, which are associated with disease? Is the association with beryllium sensitization? Do some alleles covey greater risks than others? Is latency of CBD genetically determined? A sophisticated PCR-DNA sequencing assay has been developed to more precisely identify specific HLA-DPB-1 haplotypes. Worker education and case control recruitment is

proceeding in collaboration with DRDS. To date this project has identified certain HLA-DPB1 that are associated with CBD and beryllium sensitization, it has identified homozygotes of those variants to be at higher risk of disease than heterozygotes (OR = 3.1, 95%CI = 1.5-6.1). Molecular epidemiology, together with molecular modeling, appears to show that HLA-DPB1 alleles that code for HLA molecules with greatest negative charge (-9) convey significantly more risk than those coding for HLA molecules that are moderately negatively charged (-7) (OR = 3.2, 95%CI = 1.9-5.4). Worker education efforts continue, which includes a two-day meeting at NIOSH, Morgantown, where workers now participate in making presentations.

Brain cancer is the fourth leading cause of cancer death among middle-aged men in the United States, and incidence and mortality rates are steadily increasing. One plausible hypothesis for the increasing incidence is increased exposure to etiologic agents or exposure to newly introduced carcinogens, such as pesticides and other chemicals. Even though farmers experience a lower overall cancer mortality than the general population, several studies indicate that individuals working on a farm or in the agricultural industry have an excess risk of brain cancer. NIOSH is conducting a population-based case-control study to determine if specific exposures prevalent among farmers and people living near farms are associated with increased risk for developing primary intracranial gliomas and to identify genetic susceptibility factors associated with brain cancer. Employing DNA specimens from 325 cases and 579 controls, NIOSH has analyzed over 90 gene variants in genes potentially related to brain cancer, such as polymorphisms in genes involved in activation and detoxification metabolism, DNA repair, cell cycle control, cell-to-cell communication, and immune function. In 2004, NIOSH evaluated the associations of polymorphisms in genes important in estrogen metabolism with the risk of glioma in women, since gender differences in brain tumor incidence and an analysis of questionnaire-reported reproductive and hormonal data from the study population suggested that hormonal factors might play a role in the etiology of these tumors. NIOSH found that polymorphisms in estrogen metabolism genes did not appear to have a strong association with glioma risk in women, although they have been linked to susceptibility in other cancers in women

Health effects of toxic exposures

Genetic monitoring is the evaluation of an exposed population for genetic damage over time and involves the detection of biomarkers of the effect of exposure. Much of this has grown out of the effort to assess workers and populations exposed to nuclear weapons and nuclear medicine techniques. Somatic mutations, DNA and protein adducts, and other cytogenetic changes have frequently been used as biological measures of exposure and in some cases as biomarkers of effect. The evaluation of changes in genetic material is usually part of research studies that investigate the effects of exposure or can be part of periodic medical examinations performed specifically for genotoxic agents in the workplace. Specific studies conducted by NIOSH are described below.

A previous NIOSH study showed an increase in sister chromosome exchanges in hospital workers exposed to ethylene oxide, an IARC Group I carcinogen, as compared to unexposed hospital workers. NIOSH will reassess exposed and unexposed former

workers with baseline cytology for levels of chromosomal damage, and evaluate them for susceptibility markers (polymorphic metabolic & DNA repair genes).

1-Bromopropane (1-BP) is an alternative to ozone-depleting chlorofluorocarbons that has a variety of potential applications as a degreasing agent for metals and electronics, and as a solvent vehicle for spray adhesives. As part of two NIOSH Health Hazard Evaluations (HHEs), DNA damage was assessed in peripheral leukocytes from workers with occupational exposure to 1-BP. Start-of- and end-of-workweek blood and urine samples were collected from workers at two facilities where 1-BP was used as a solvent for spray adhesives in foam cushion fabrication. Urinary bromine (Br) levels served as a biomarker of exposure. The comet assay was used to estimate DNA damage. In 1-BP exposed workers, start-of- and end-of-workweek comet endpoints were stratified based on job classification. Data were analyzed by combining the data sets from both facilities, log transformation of 1-BP exposure indices, Pearson correlation analysis, and the use of multiple linear regression models for each combination of exposure index and the level of DNA damage. Pearson correlation analysis indicated DNA strand breaks were positively associated with urine ($r= 0.288$, $P= 0.026$) and serum ($r=0.259$, $P=0.044$) Br concentrations, but stratification of workers into exposure quartiles did not reveal any significant differences and supported the conclusion that workplace exposure to 1-BP was not associated with increased DNA damage in leukocytes at these two facilities

The use of genetic information in occupational health studies

Because of these scientific advances, genetics has begun to transform research questions and study designs in the applied sciences of public and occupational health. Genetic studies provide new ways to study "risks" by evaluating genes and gene–environment interactions. The incorporation of genetics into occupational safety and health research generally requires collecting biological specimens from participating workers, analyzing those specimens, and developing test and study results. High throughput technology, such as microarrays, presents a number of challenges in terms of validity, data reduction and summarization, and analysis and interpretation. Analysis of these large datasets will also amplify the challenges that already exist when trying to relate genetic information and environmental factors. NIOSH has begun a discussion of study design and data interpretation issues that affect occupational health studies. A guidance document is currently being developed to address genetic issues in the workplace.

Major Accomplishments, 2004

Gene-Environment Interactions

Eleven papers in the area of gene environment interactions have been published this year.

Snyder, J., Weston, A., Tinkle, S.S., Demchuk, E.: Electrostatic potential on human leukocyte antigen: implications for putative mechanism of chronic beryllium disease. *Environmental Health Perspectives,* 111: 1827-1834, 2003.

McCanlies, E.C., Ensey, J., Schuler, C., Kreiss, K., Weston, A.: The Association Between *HLA-DPB1*^{Glu69}, Chronic Beryllium Disease, and Beryllium Sensitization. *American Journal for Industrial Medicine,* 46: 95-103, 2004.

Keshava, C., McCanlies, E.C., and Weston, A.: *CYP3A4* Polymorphisms in breast and prostate Cancer: A HuGE Review. *American Journal of Epidemiology*, 160: 825-841, 2004.

Gwinn, M.R., Whipkey D. L., Weston, A.: the effect of oxythioquinox exposure on normal human mammary epithelial cell gene expression: A microarray analysis study. Environmental Health, 3: 9 – 19, 2004.

Keshava, C., Divi, R., Whipkey, D.L., Frye, B.L., McCanlies, E.C., Kuo, M., Poirier, M.C., and Weston A.: Induction of CYP1A1 and CYP1B1 and formation of carcinogen-DNA adducts in normal human mammary epithelial cells treated with benzo[a]pyrene. *Cancer Letters*, 2004, in press.

Keshava, C., Whipkey, D.L., and Weston, A.: Transcriptional signatures of environmentally relevant exposures in normal human mammary epithelial cells: benzo[a]pyrene. *Cancer Letters*, 2004, in press.

Mahadevan, B., Keshava, C., Musafia-Jeknic, T., Pecaj, A., Weston, A., Baird, W. M.: Altered gene expression patterns in MCF-7 cells induced by the urban dust particulate complex mixture SRM 1649a. *Cancer Research*, 2004, in press.

Weston, A., Snyder, J., McCanlies, E.C., Schuler, C. R., Kreiss, K., Demchuk, E.: Immunogenic factors in chronic beryllium disease. *Mutagenesis*, 2004, in press.

Tinkle, S.S., Weston, A.: Beryllium toxicity and disease. Encyclopedia of Toxicology, Springer-Verlag. Third Edition. 2004, in press.

Weston, A., Ensey, J. S., Frye, B.L.: DNA-Sequence Determination of a Novel HLA-DPB1 Allele, HLA-DPB1*0403. *DNA Sequence*, 2004, in press.

Yuan, B-Z., Jefferson, A.M., Popescu, N.C., Reynolds, S.H. (2004) Aberrant gene expression in human non-small cell lung carcinoma cells exposed to demethylating agent 5-aza-2'-deoxycytidine. Neoplasia 6 (4):412-419.

Health effects of toxic exposures

Toraason, M., Singh, N. and Lynch, D.W. 2004. DNA damage in human leukocytes induced in vitro by 1- or 2-bromopropane. Toxicological Sci. 78(S-1):31-32.

Weston, A., Poirier, M.C.: Carcinogen DNA-adducts and DNA Repair. Encyclopedia of Toxicology, Springer-Verlag. Third Edition. 2004, in press.

Hooven, L.A., Mahadevan, B., Keshava, C., Perira, C., Desai, D., Amin, S., Weston, A., Baird, W.M.: effects of suberoylanilide hydroxamic acid and trichostatin A on induction of cytochrome P450 enzymes and benzo[a]pyrene DNA adduct formation in human cells. *Bioorganic & Medicinal Chemistry Letters*, 2004, in press.

Clark, A.M., Reynolds, S.H., Anderson, M. Wiest, J.S. (2004) Mutational activation of the MAP3K8 protooncogene in lung cancer. Genes, Chrom. & Cancer 41:99-108.

The use of genetic information in occupational health studies

Three papers were published this year that discussed the use of genetic information in epidemiological studies and practice.

Schulte PA. Some implications of genetic biomarkers in occupational epidemiology and practice. Scan J Work Environ Health 30 (1): 71-79, 2004.

Vineis P, Schulte PA, Carreon T, Bailer AJ, Medvedovic M. Issues in design and analaysis of gene-environment interactions. In P Buffler, J Rice, R Baan, M Bird, P Boffeta (eds). Mechanisms of carcinogenesis. IARC, Lyon France. IARC Sci Publ 157: 417-36, 2004.

Schulte PA. Interpretation of genetic data for medical and public health uses. In G Arnason, S Nordal, V Arnason (eds.) Blood and Data: Ethical Legal and Social Aspects of Human Genetic Databases. University of Iceland Press , Reykjavik, 2004.

Future Directions

NIOSH is proposing to develop a Laboratory for Occupational Genomics (NIOSH-LOG). The objectives of the proposed NIOSH-LOG are to 1) build an occupational biological sample base (DNA, blood, buccal smears, paraffin blocks, other) with which to perform molecular epidemiological genetic association studies to identify genetic risk factors of occupational disease. Samples will be available for investigators within the laboratory as well as NIOSH and non-NIOSH investigators that develop an appropriate study design. 2) Foster partnerships with academia, industry, other government and international agencies. 3) Provide a framework for data analysis that will be available for NIOSH investigators, academics and regulatory agencies. 4) Explore ethical challenges surrounding the use of genetic data in occupational health research. The focus, initially, will be on diseases currently being evaluated, such as silicosis, asthma, chronic beryllium disease, AZT treatment for HIV, irritant dermatitis, genetic susceptibility to lung cancer, and deterioration of lung function in fire fighters.

Autism is a developmental disability that is characterized by repetitive behavior, impairment in reciprocal social interaction, and difficulty communicating with and intuiting the feelings of others. A diagnosis of autism is generally made between the ages of two and three when there is a noticeable delay in language skills. It occurs most often in boys, but occurs in all social, racial, and ethnic groups. The prevalence has been estimated at 15-34 per 10,000. Genetic and epidemiologic research have shed some light

on the etiology of the disease, but further research on the potential role of environmental and occupational factors is needed. In a proposed study an investigation of the relationship between parental occupation, as defined by a job exposure matrix, and autism will be assessed.

Future gene expression studies using ejaculate sperm may yield important mechanistic insights regarding the effects of toxicants on male fertility. A pilot methods development project to assess gene expression in human sperm is underway. The recent discovery of spermatozoal coding RNAs that are delivered to the oocyte on fertilization has generated hypotheses regarding the role of sperm gene expression in human fertility and the potential effects of toxicants on sperm gene expression. The aim of this project, entitled "The Feasibility of Examining Gene Expression in Sperm," is to develop NIOSH field and laboratory methods to evaluate sperm gene expression, enabling investigators to potentially incorporate this parameter in future fertility and toxicant studies.