Encyclopedia of Epidemiology

Family Studies in Genetics

Contributors: Scott R. Diehl & Robert P. Erickson
Edited by: Sarah Boslaugh
Book Title: Encyclopedia of Epidemiology
Chapter Title: "Family Studies in Genetics"
Pub. Date: 2008
Access Date: December 11, 2015
Publishing Company: SAGE Publications, Inc.
City: Thousand Oaks
Print ISBN: 9781412928168
Online ISBN: 9781412953948
DOI: http://dx.doi.org/10.4135/9781412953948.n153
Print pages: 384-386

Family studies may be considered a key entry point for research into the role of inherited genetic variation in disease. There are three major kinds of family studies: (1) evaluations of the extent to which a disease or other trait of interest aggregates or clusters within families, and how it is transmitted across the generations; (2) linkage analyses aimed at pinpointing the specific location on one of the chromosomes of a gene containing a mutation that has a major or moderate effect on disease risk; and (3) association studies aimed at finding common gene variants that have smaller but still medically important effects on disease severity or disease incidence. All three of these kinds of family studies are considered in this entry.

Analysis of Aggregation or Clustering

One of the first questions that investigators need to ask when considering genetic studies of a disease or some other trait of interest is 'What is the evidence that inherited genetic variation has an important influence on the trait?' A necessary but not sufficient condition required to demonstrate the importance of genetic variation is the occurrence of *familial aggregation* of the trait. We know from the simple rules of Mendelian inheritance that family members tend to share genes in common. For example, siblings share 50% of their genes inherited identical by descent from their parents, and cousins share 12.5% of genes inherited from their common grandparents. Therefore, if genetic variation really is important for the development of a disease (incidence) or its severity, then we would expect to find the disease co-occurring or 'clustering' among family members more often than among randomly drawn unrelated individuals in the population. One way this is often formally tested in research studies is to compare the frequency of disease in relatives of persons with the disease compared with the frequency found in relatives of matched healthy controls. If disease frequency is not at all elevated in relatives of cases, then it is unlikely that most cases of the disease in the population have a substantial genetic basis. On the other hand, familial aggregation of disease is evidence that the trait has a genetic basis but still does not constitute definitive proof. This is because in addition to sharing genes, family members also usually share similar environments (diet, exposure to toxins, etc.), and it is possible that the aggregation of disease in relatives is caused by their common environments rather than by shared genes. Investigators can measure environmental exposures that are suspected to be risks for the disease in cases, in controls, and in their relatives and attempt to statistically adjust for these effects in data analyses. Alternatively, if the disease is sufficiently common, studies of monozygotic and dizygotic twin families offer a very powerful design that can provide very powerful capability to distinguish between environmental and genetic causes of variationindiseaserisk.

For some forms of very severe single-gene (Mendelian) disorders there may not be a positive family history if transmission is dominant and clinical symptoms onset at a young age. People with such diseases are unlikely to reproduce so patients with the disease frequently have arisen via a new mutation not present in their ancestors. On the other hand, with recessive diseases, the parents, not surprisingly, usually do not know that they are carriers for the recessive mutation and the disorder may be new to the family. An exception to this rule is for recessive diseases occurring among cultures with consanguinity (e.g., first-cousin marriages) where occurrence of the disease may not be surprising. Studies of consanguineous families can be very useful for gene-mapping studies.

It should be noted, however, that even if no evidence of familial aggregation or heritability is obtained from family studies, this does not rule out the possibility that a small subset of disease cases (e.g., 1% to 5%) might be caused by a mutation in a gene that causes a major increase in disease risk. In fact, strong familial aggregation may exist for this small genetically based subset of cases, but this is obscured by the majority (95% to 99%) of disease cases for which genetic variation has little or no influence on disease risk. For example, most cases of breast cancer lack familial aggregation, and in twin studies there is little evidence of heritability, but relatively rare mutations in BRCA1 gene and other genes have a very major effect on cancer risk in individuals who inherit these mutations. Furthermore, sometimes the same genes that are involved in the rare inherited forms of a disease are mutated somatically in nonhereditary cases. Therefore, understanding the biological mechanisms involved in rare hereditary forms may have great importance for developing improved methods of diagnosis, prognosis, and therapy for both hereditary and nonhereditary forms of the disease.

When a disorder shows familial aggregation that appears not attributable to shared environmental exposures, a statistical method called *segregation analysis* can be used in an attempt to estimate the mode of inheritance—autosomal recessive, dominant, or codominant; X-linked dominant or recessive. This technique has been successfully applied to many simple (single gene) disorders, but it has only limited value in studies of complex diseases where multiple disease susceptibility genes interact to influence disease risk. Segregation analyses usually need to assume *homogeneity*, meaning that the same type of gene is responsible for causing the disease in all families included in the study. If, in fact, some families have inherited a gene that acts dominantly while other families in the data set have inherited mutations at either the same gene or a different gene where risk is recessively transmitted, segregation analysis will be unreliable. Even with relatively simple disorders, the method has serious limitations. First, one must be aware of and appropriately adjust for the way the families and family members were selected for study (ascertainment bias). Second, there is the problem of unrecognized shared environments (noted above), and, for quantitative traits, deviations from assumptions of normality can lead to incorrect inferences about the mode of transmission. The method has been modified in recent years in an attempt to address these weaknesses, but it has nonetheless been largely supplanted in genetic epidemiological research by family studies that incorporate DNA markers.

Linkage Analysis

When a single gene has a major effect (e.g., > 10-fold increase) on the risk of developing a disease, and when the disease is relatively uncommon in the population, then the method of linkage analysis can rapidly lead to successful gene identification. *Linkagemapping* families are evaluated for the cosegregation of polymorphic DNA markers (either short tandem repeats or single-nucleotide polymorphisms) with the disease phenotype. Linkage mapping depends on the fact that recombination during meiosis occurs only rarely between markers that are located physically close (linked) to the disease gene. Recombination occurs increasingly more often as markers are located farther away (> 10 Mbp) on the same chromosome as the disease gene or located on a different chromosome altogether. It is possible to detect linkage with as few as 12 to 16 informative individuals when the disorder is highly penetrant (i.e., nearly every person who inherits the mutation gets the disease), when there are few

phenocopies (i.e., hardly anyone who does not inherit the mutation gets the disease), and if the density of markers is sufficiently high. Most often, a single sufficiently large extended family is not available, so several unrelated families may need to be combined, especially when attempting to map a recessive trait. In some special circumstances, consanguineous (inbred) families may allow investigators to use an approach called *homozygosity mapping* to localize a recessive disease gene. For example, in offspring of first cousins, about one sixteenth of the genome is expected to be homozygous. The specific homozygous genome regions would be random in affected offspring of different sets of first cousins, except for the region that contains their recessive disease gene. This region would be homozygous in the offspring of all the offspring, so by evaluating only a limited number of such offspring, a disease gene can be mapped. Linkage analysis has had some success for oligogenic diseases (i.e., those with only a few genes involved). Unfortunately, despite major investments of resources and years of effort, studies of complex disorders that are likely to involve multiple genes of smaller effect (e.g., twofold increase in risk) and potentially involving gene-gene and gene-environment interactions have usually been disappointing.

Association Analysis

Mathematical analyses and several recent disease studies have shown that *association mapping* methods can provide good statistical power for identifying genes that underlie complex diseases while requiring much smaller numbers of patients and their relatives than would be required for linkage analysis. Association mapping can be performed using either small or large families or unrelated cases and controls. The only catch is that instead of needing only a few hundred polymorphic DNA markers to cover the entire human genome, the association strategy requires several hundred thousand such markers assayed on each subject. Fortunately, molecular genetic technologies have been developed that can meet the challenge of producing these massive amounts of data, and the International HapMap Project ('HapMap' being an abbreviation of 'haplotype map') has cataloged this variation in several human populations and made it freely available online for researchers wishing to tap into this rich genomic treasure chest. In the first phase of this project, the frequency of DNA variants was measured at more than 1 million locations distributed across the human genome in European, African, and Asian subjects. DNA variants that are located near each other on thechromosomeoftenarecorrelatedwitheachother, so if an investigator determines the DNA sequence for a subject at one position, the DNA bases at the neighboring variant positions often can be predicted with a high degree of confidence. This phenomenon is known as *linkage disequilibrium*, and the Hap Map Project has determined where these patterns of correlation among neighboring DNA variants exist for a large portion of the human genome. Armed with this information, investigators interested in studying inherited variation at a candidate gene for their disease (or searching through all genes in the entire genome) need not undertake the large effort of conducting assays for all known variants in their clinical subjects. Instead, they can use computer algorithms on data derived from the HapMap Project to measure most of the inherited variation present in the genome at a substantially reduced cost by identifying an optimized subset of DNA variants that serve as statistical 'tags' for many other variants that are not actually assayed in the laboratory. The National Institutes of Health and other organizations responsible for support of biomedical research are currently developing plans for a major expansion of whole-genome association studies to a wide range of diseases and to drug side effects and therapeutic responses (pharmacogenomics). Family studies are certain to play an important role for these exciting initiatives in the

future of genomic medicine.

Ethical Issues

There are many important and complex ethical issues that arise when performing family studies. Members of families need to be carefully educated about the risks, both social and cultural, of participating in family studies, which include providing information on their relatives. Although investigators will have obtained approval from an institutional review board responsible for protecting research subjects, such review boards generally focus on possible outcomes from the genetic (biological) information that will be obtained and may not always fully consider the possibility of altered family dynamics that may arise as a consequence of participating in the study. Anticipatory counseling of prospective families can enhance participation rates and minimize stressful effects on family dynamics.

- genes
- linkage analysis
- disease
- mutation
- family studies
- aggregation
- DNA

Scott R. Diehl, and Robert P. Erickson
http://dx.doi.org/10.4135/9781412953948.n153
See also

- Gene-Environment Interaction
- Genetic Disorders
- Genetic Epidemiology
- Genetic Markers
- Genomics

Further Readings

Arar, N. H., Hazuda, H., Steinbach, R., Arar, M., and Abboud, H. Ethical issues associated with conducting genetic family studies of complex disease. Annals of Epidemiology 15 (9) (2005). 712–719.http://dx.doi.org/10.1016/j.annepidem.2004.09.010

Elston, R. C., and Spence, A. M. Advances in statistical human genetics over the last 25 years. Statistics in Medicine 25 (2006). 3049–3080.http://dx.doi.org/10.1002/sim.2650

Morton, N. E. Fifty years of genetic epidemiology, with special reference to Japan. Journal of Human Genetics 51 (2006). 269–277.http://dx.doi.org/10.1007/s10038-006-0366-9

Savage, S. A., and Chanock, S. J. Genetic association studies in cancer: Good, bad or no longer ugly? Human Genomics 2 (2006). 415–421.