

Genetics of complex diseases: From mouse to man and back

CRAIG H WARDEN PHD, JEROME I ROTTER MD

IDENTIFICATION OF GENES UNDERLYING complex (non-Mendelian) traits, which include diseases such as inflammatory bowel disease (IBD), has been an important and difficult problem in genetics. However, new tools

CH WARDEN, JI ROTTER. Genetics of complex diseases: From mouse to man and back. Can J Gastroenterol 1995;9(3):169-174. Identification of genes underlying complex traits has been difficult, but combined application of novel methods and mouse models provides new hope. Rare monogenic syndromes, and candidate gene and biochemical approaches are sometimes useful, but each of these approaches also has limitations. Some problems that prevent identification and isolation of genes underlying complex disease can be avoided by the use of whole genome mapping of mouse crosses or of human families. Mice have many advantages for the study of complex disease, including an extensive genetic map. A generic method has recently been developed and applied for detection of quantitative trait loci (QTLs) using whole genome maps of mouse crosses. Availability of more than 200 congenic strains provides another incentive for studies in mice. Congenic strains provide a rich, but previously unexploited, resource for the rapid identification of genes causing complex diseases. A congenic mouse strain is genetically identical to a background strain, except for a small chromosomal region derived from a donor strain. Thus, comparison of a phenotype in a congenic strain with the phenotype in its background strain allows study of the effects of single genes derived from the donor strain, isolated from the effects of other donor strain genes. Application of all or several techniques to complex disease studies in mice and in humans may lead to the identification and understanding of complex diseases whose etiology is currently unknown. (*Pour résumé, voir page 170*)

Key Words: Complex trait, Congenic, Mouse model, Quantitative trait loci (QTLs), Whole genome maps

Department of Medicine and Department of Microbiology and Molecular Genetics, University of California; and Division of Medical Genetics and IBD Center, Cedars-Sinai Medical Center, Los Angeles, California, USA

Correspondence: Dr Craig H Warden, Department of Medicine, Division of Cardiology, UCLA, Los Angeles, CA 90024-167917, USA. Telephone (310) 206-0133, Fax (310) 794-7345, e-mail warden@biovxl.biology.ucla.edu

This paper was presented at the Basic Research and Clinical Implications in IBD meeting, April 6 to 9, 1994, held in Victoria, British Columbia. This paper has also been published in Sutherland LR, et al, eds. *Inflammatory Bowel Disease: Basic Research, Clinical Implications and Trends in Therapy*. Boston, Dordrecht and London: Kluwer Academic Publishers, 1994

and methods have brought on a new era in the genetic analysis of complex traits. While identification of the genes underlying common diseases will never be easy, we believe that generic approaches to analysis of common diseases are now practical. An outline of the approaches to analysis of complex diseases is given in Figure 1. The six sections of this review are: the role of rare monogenic syndromes in the analysis of complex diseases; candidate gene and biochemical approaches to complex diseases; complete linkage maps in humans and animal models; positional cloning of genes underlying complex diseases in animal models; testing hypotheses by genetic modifications in animals; and testing whether genes identified in animal models also underlie human complex disease.

Theories and methods recently developed to detect novel genes underlying complex traits are illustrated using the authors' studies of the genes causing atherosclerosis (1-3). The general conclusion from these studies is that systematic linkage maps are a powerful approach to the study of complex diseases.

RARE MONOGENIC SYNDROMES

Rare genetic syndromes can provide biochemical and genetic information that leads to the identification of genes underlying complex human diseases.

tion or family study. The first striking example of the utility of the candidate gene approach was the demonstration that structural polymorphisms of apolipoprotein (apo) E affect cholesterol levels and are associated with type III hyperlipidemia. Subsequent work revealed the biochemical basis of the polymorphisms: altered binding of apo-containing lipoproteins to cellular receptors (10).

Thirty-five of the 41 genes underlying diseases in mice that have been identified at a molecular level were identified by a candidate gene approach (11). The candidate gene approach can also be applied to complex traits. Studies of the nonobese diabetic (NOD) mouse have already used comprehensive linkage maps constructed in genetic crosses to demonstrate that a defective Fc receptor for immunoglobulin G is linked to autoimmune diabetes in NOD mice (12).

There are many potential candidate genes for IBD, including genes coding for the human leukocyte antigen complex, complement proteins, immunoglobulins, cytokines and additional molecules. These candidate genes could be tested for their association with, and linkage to, IBD by using specific assays in human families and in association studies. The role of each of these many candidate genes in IBD could be tested more quickly by using complete linkage maps in family studies.

COMPLETE LINKAGE MAPS

Animal models: Quantitative trait loci (QTL) mapping is a general method that has recently been developed to map genes underlying complex disease in animal models (13,14). It does not require any previous knowledge regarding the underlying physiology of the disease being studied.

QTL mapping involves four steps. First, two different inbred strains are crossed to produce F2 or backcross progeny. Next, the progeny are individually genotyped for markers that span the genome at 10 to 20 centiMorgan (cM) intervals (a cM is a measure of genetic linkage distance, indicating frequency of recombination between

the genetic loci of 1%). Third, each of the progeny is phenotyped for the traits of interest and, finally, QTLs are located by a statistical approach, such as that provided by MAPMAKER/QTL (14).

Mice are a powerful resource for analysis of complex traits (7). The advantages of using mice are: many diverse inbred strains are available; mice can be used for hypothesis testing by construction of transgenics and knockout mice; the number of markers in the genetic map of the mouse is second only to that of humans; finally, many powerful genetic resources are available, such as congenics and recombinant inbred strains.

Example of mapping complex traits in a mouse model – apoA-II: The principles of QTL mapping can be illustrated with a mouse backcross that has been analyzed for genes underlying obesity and atherosclerosis. A backcross was performed by crossing F1 females, resulting from the cross of female C57BL/6J with male *Mus spretus* mice, with male C57BL/6J mice. These backcross progeny are designated BSB mice (15). Linkage of these genes with the quantitative traits has been determined by ANOVA and by logarithm of odds (LOD) score analysis with the MAPMAKER/QTL program (2).

The authors measured plasma apoA-II levels in BSB mice. A peak LOD score of 4.0 was revealed on mouse chromosome I centred around the apoA-II gene locus. This QTL includes approximately 10 cM in the 90% confidence interval and thus includes many diverse genes. However, the coincidence of the QTL for plasma apoA-II levels and the apoA-II gene suggested that the apoA-II gene directly controls plasma apoA-II levels. This hypothesis was subsequently tested successfully in transgenic mice and in humans.

Animal models for IBD: There are several animal models that may be used for studies of IBD. For instance, peptidoglycan-polysaccharide injection of Lewis rats leads to granulomatous hepatitis and anemia. However, Buffalo and F344 rats are genetically resistant to this injection. Thus, crosses of Lewis rats with Buffalo or F344 rats could be used to identify genes involved

in this response to peptidoglycan-polysaccharide injection.

Complete linkage maps in humans: Theories and methods to detect novel genes underlying complex traits with systematic linkage maps in humans have been developed (13,16). The major requirement of these methods is the generation of systematic linkage maps. Systematic linkage maps require that hundreds of markers be typed in each member of families with IBD. The simple sequence repeat polymerase chain reaction (PCR) markers are ideal for this purpose. There are more than 2000 simple sequence repeat PCR markers available for human linkage mapping (17,18). These markers can be typed rapidly and at low cost, are highly polymorphic and are spaced at an average of 1.5 cM (there are 3000 cM in the human genome).

Studies with the apoA-II locus, discussed below, suggest that random genetic markers may exhibit significant linkage to quantitative traits 5 to 20 cM from the genes underlying quantitative traits. For instance, the author and colleagues (1) found that D1S74, 5.7 cM from the apoA-II structural locus, can detect a highly significant linkage ($P < 0.045$) to serum levels of free fatty acids (FFA) in multiplex coronary artery disease families. A complete linkage map capable of detecting novel genes underlying quantitative traits could theoretically be constructed by genotyping just 150 markers. The actual number of markers needed will depend on the complexity of the disease, the heterozygosity of the markers, the number of families and the number affected in each family (16).

POSITIONAL CLONING

Positional cloning (also known as reverse genetics) is a powerful approach for identification and characterization of genes for monogenic disorders in which no biochemical or candidate gene can be identified. Positional cloning involves: identifying genetic markers linked to a disease gene by testing for cosegregation of the disease phenotype with genetic markers spanning various parts of the genome; physically cloning regions of the ge-

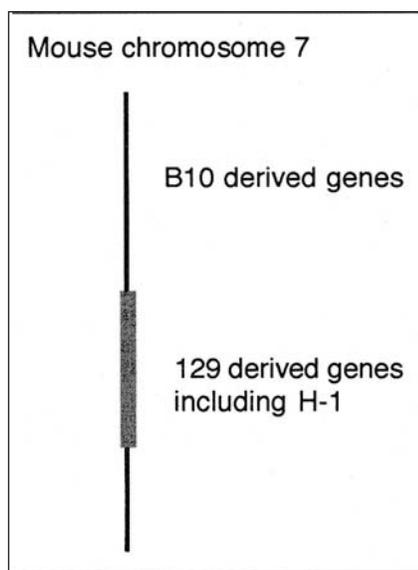


Figure 2) Approximate appearance of the B10.129 (H-1) congenic mouse strain. All other chromosomes in this congenic strain would be derived from the C57BL/10SnJ background strain

nome surrounding the genetic markers; and testing whether genes present in the cloned regions have properties consistent with the disease gene, such as lack of recombination with the disease gene or alterations in expression in affected individuals. Thus, a reverse genetics approach does not depend on an understanding of the molecular mechanisms contributing to the disease. The past several years have witnessed the cloning of several genes causing human disease by a 'positional genetics' approach in humans, including those for familial adenomatous polyposis (19,20) and colon cancer.

Unfortunately, reverse genetic approaches are difficult to use in human studies of multifactorial disorders such as atherosclerosis and IBD. First, it is difficult to identify the responsible genetic loci because multiple independently segregating loci contribute to the disorders. This problem is exacerbated by genetic heterogeneity, in which a similar disorder is caused in different individuals by different genetic defects. Other problems with human genetic studies include incomplete penetrance and gene interactions (synthetic traits).

Use of congenics to facilitate positional cloning: Identification of QTLs is just

the first step in the study of genes underlying atherosclerosis in mice. Biochemical and physiological studies require identification of the specific genes underlying complex traits. One method to identify the specific gene underlying a QTL is to use, or to produce, congenic mouse strains. A congenic mouse strain is genetically identical to a background strain, except for a small chromosomal region derived from a donor strain. Thus, one can study the effects of a single donor gene on traits in the background strain, isolated from the effects of other donor strain genes (21). For example, a congenic with C57BL10/J as the background strain and strain 129 as the donor strain for the mouse chromosome 7 H-1 locus would then have this very approximate appearance (Figure 2).

Congenic strains can be bred to their background strain and the resulting F2 progeny used to isolate the underlying genes. For instance, positional cloning of atherosclerosis genes could be guided by progeny of a cross between a congenic strain containing a single atherosclerosis gene with the background strain. This method can work because a single gene would be responsible for genetic variation in the trait. Several methods are also available that specifically identify restriction fragment length variants, which can aid in the positional cloning of the underlying genes, for the donor strain of a congenic (22).

TESTING HYPOTHESES BY GENETIC MODIFICATION IN ANIMALS

As shown above, a QTL for serum apoA-II levels at the locus surrounding the mouse apoA-II gene has been demonstrated. The author next used transgenic mice to test the hypothesis that the apoA-II gene controls plasma apoA-II levels. The author and associates (23) generated mice that are transgenic for the mouse apoA-II gene. Transgenic mice were constructed using a genomic clone of the mouse apoA-II gene containing several kilobases of 5' flanking and 3' flanking DNA. Southern blot analysis revealed that approximately 10 copies of the

apoA-II gene were integrated into the genome in this line.

The studies revealed that transgenic mice that overexpress mouse apoA-II had elevated high density lipoprotein (HDL) cholesterol concentrations but, nevertheless, exhibited increased atherosclerotic lesion development compared with normal mice (3). Laboratory strains of mice do not develop significant aortic fatty streak lesions when maintained on chow diets. ApoA-II transgenic mice developed aortic lesions on a chow diet despite the fact that they had high concentrations of HDL and low concentrations of LDL and very low density lipoprotein. Lesion development was observed in transgenic mice of both sexes. The HDL in the transgenic mice was larger and an increased ratio of apoA-II:apoA-I was present. Thus, both the composition and amount of HDL appear to be important determinants of atherosclerosis. These results clearly suggest that the apoA-II gene underlies the QTL observed in the BSB cross.

TESTING WHETHER GENES CONTRIBUTE TO HUMAN DISEASE

Extension of mouse mapping results to humans is possible because comparative gene mapping has demonstrated that linked homologous genes are found in at least 101 conserved autosomal segments (11). To test whether variations of the apoA-II gene influence plasma lipid metabolism in humans, the author studied 306 individuals in 25 families enriched for coronary artery disease. The segregation of the apoA-II gene was followed using an informative simple sequence repeat in the second intron of the gene (24) as well as two nearby genetic markers. Robust sib-pair linkage analysis was performed on members of these families using the SAGE linkage programs (25).

The results demonstrated linkage between the human apoA-II gene and a gene controlling plasma apoA-II levels. Plasma apoA-II levels were also significantly correlated with plasma FFA levels. Moreover, the apoA-II gene exhibited significant linkage with a gene

controlling FFA levels. Evidence for nonrandom segregation was seen with markers as far as 6 to 12 cM from the apoA-II structural locus (1).

SUMMARY

Animal models are useful to identify genes underlying complex diseases, to test hypotheses and to study functions of proteins involved in complex traits (Figure 1). Thus, transgenic models recently have been used to examine questions relating to atherosclerosis. Targeted disruption of genes in mice are potentially very informative. Similarly, several novel models of IBD have been identified by targeted mutagenesis in mice. Many studies have suggested that IBD patients have hyper-responsiveness to normal gut constituents; however, they have failed to define the reasons at cellular or molecular levels (26). Thus, the discovery of ulcerative colitis-like disease in mice with a disrupted interleukin (IL)-2 gene suggested a primary role for the im-

mune system in the etiology of ulcerative colitis (27). Spontaneous development of IBD has also been reported in mice with mutations in the T cell receptor alpha mutant, T cell receptor beta mutant or class II major histocompatibility mutant mice (28).

These results suggested that dysfunction of the immune system may underlie pathogenesis of IBD. It has also been reported that IL-10-deficient mice suffer from chronic enterocolitis (29). The results suggest that bowel inflammation results from uncontrolled immune responses and that IL-10 is an essential regulator (29). All totalled, these gene-targeted mice represent novel models of IBD. The presumed final common pathway is a disorder in the normal control mechanisms that normally down-regulate responses to mucosal antigens. Finally, it seems possible that these models can be used to develop therapeutic strategies for IBD, even though the pathogenesis of the

mouse and human diseases are not identical (26).

CONCLUSIONS

The author and colleagues' work with mice strongly suggests that complete linkage maps are very likely to identify genes underlying complex traits, such as IBD. One of the biggest advantages of the systematic linkage approach is that one can narrow the list of candidate genes for a disease to those that are present in the loci underlying complex traits. Furthermore, systematic linkage maps may also identify novel genes or loci that are important in the etiology of the disease.

ACKNOWLEDGEMENTS: Supported in part by grants from the Crohn's and Colitis Foundation of America, National Institutes of Health Program project grant DK46763, the Stuart Foundation and the Cedars-Sinai Board of Governors' Chair in Medical Genetics.

REFERENCES

1. Warden CH, Daluiski A, Bu X, et al. Evidence for linkage of the apolipoprotein A-II locus to plasma apolipoprotein A-II and free fatty acid levels in mice and humans. *Proc Natl Acad Sci USA* 1993;90:10886-90.
2. Warden CH, Fislis JS, Pace MJ, Svenson KL, Lusic AJ. Coincidence of genetic loci for plasma cholesterol levels and obesity in a multifactorial mouse model. *J Clin Invest* 1993;92:773-9.
3. Warden CH, Hedrick CC, Qiao JH, Castellani LW, Lusic AJ. Atherosclerosis in transgenic mice overexpressing apolipoprotein A-II. *Science* 1993;261:469-72.
4. Goldstein JL, Brown MS. Familial hypercholesterolemia. In: Stanbury JB, Wyngaarden JB, Fredrickson DS, Goldstein JL, Brown MS, eds. *The Metabolic Basis of Inherited Disease*. New York: McGraw-Hill, 1983:672-712.
5. Hobbs HH, Russell DW, Brown MS, Goldstein JL. The LDL receptor locus in familial hypercholesterolemia: mutational analysis of a membrane protein. *Ann Rev Genet* 1990;24:133-70.
6. Yang H, Rotter JI. The genetics of inflammatory bowel disease: genetic predispositions, disease markers, and genetic heterogeneity. In: Targan SR, Shanahan F, eds. *Inflammatory Bowel Disease: From Bench to Bedside*. Baltimore: Williams and Wilkins, 1994:32-64.
7. Warden CH, Daluiski A, Lusic AJ. Identification of new genes contributing to atherosclerosis: The mapping of genes contributing to complex disorders in animal models. In: Lusic AJ, Rotter JI, Sparkes RS, eds. *Monographs in Human Genetics: Molecular Genetics of Coronary Artery Disease*. Basel: Karger, 1992:419-41.
8. Brown MS, Goldstein JL. A receptor-mediated pathway for cholesterol homeostasis. *Science* 1986;232:34-47.
9. Mehrabian M, Lusic AJ. Genetic markers for studies of atherosclerosis and related risk factors. In: Lusic AJ, Rotter JI, Sparkes RS, eds. *Monographs in Human Genetics: Molecular Genetics of Coronary Artery Disease*. Basel: Karger, 1992:363-418.
10. Rail SC Jr, Mahley RW. The role of apolipoprotein E genetic variants in lipoprotein disorders. *J Intern Med* 1992;231:653-9.
11. Copeland NG, Jenkins NA, Gilbert DJ, et al. A genetic linkage map of the mouse: Current applications and future prospects. *Science* 1993;262:57-66.
12. Prins J-B, Todd JA, Rodrigues NR, et al. Linkage on chromosome 3 of autoimmune diabetes and defective Fc receptor for IgG in NOD mice. *Science* 1993;260:695-8.
13. Lander ES, Botstein D. Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* 1989;121:185-99.
14. Paterson AH, Damon S, Hewitt JD, et al. Mendelian factors underlying quantitative traits in tomato: Comparison across species, generations, and environments. *Genetics* 1991;127:181-97.
15. Warden CH, Mehrabian M, He K, et al. Linkage mapping of 40 randomly isolated liver cDNA clones in the mouse. *Genomics* 1993;18:295-307.
16. Lander ES, Botstein D. Mapping complex genetic traits in humans: new methods using a complete RFLP linkage map. *Cold Spring Harbor Symposia on Quantitative Biology LI, Cold Spring Harbour, 1986:49-62*.
17. Weissenbach J, Gyapay G, Dib C, et al. A second-generation linkage map of the human genome. *Nature* 1992;359:794-801.
18. Hudson TJ, Engelstein M, Lee MK, et al. Isolation and chromosomal assignment of 100 highly informative human simple sequence repeat polymorphisms. *Genomics* 1992;13:622-9.
19. Groden J, Thilveris A, Samowitz W, et al. Identification and characterization of the familial adenomatous polyposis coli gene. *Cell* 1991;66:589-600.
20. Kinzler KW, Nilbert MC, Su L-K, et al. Identification of FAP locus genes

- from chromosome 5q21. *Science* 1991;253:661-5.
21. Mehrabian M, Qiao J-H, Hyman R, Ruddle D, Laughton C, Lusic AJ. Influence of the ApoA-II gene locus on HDL levels and fatty streak development in mice. *Arterioscler Thromb* 1993;13:1-10.
22. Lisitsyn NA, Segre JA, Kusumi K, et al. Direct isolation of polymorphic markers linked to a trait by genetically directed representational difference analysis. *Nature Genet* 1994;6:57-63.
23. Hedrick CC, Castellani LW, Warden CH, Puppione DL, Lusic AJ. Influence of mouse apolipoprotein A-II on plasma lipoproteins in transgenic mice. *J Biol Chem* 1993;268:20676-82.
24. Weber JL, May PE. Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *Am J Hum Genet* 1989;44:388-96.
25. Wilson AF, Elston RC, Tran LD, Siervogel RM. Use of the robust sib-pair method to screen for single-locus, multiple-locus, and pleiotropic effects: application to traits related to hypertension. *Am J Hum Genet* 1991;48:862-72.
26. Strober W, Ehrhardt RO. Chronic intestinal inflammation: an unexpected outcome in cytokine or T cell receptor mutant mice. *Cell* 1993;75:203-5.
27. Sadlack B, Merz H, Schorle H, Schimpl A, Feller AC, Horak I. Ulcerative colitis-like disease in mice with a disrupted interleukin-2 gene. *Cell* 1993;75:253-61.
28. Mombaerts P, Mizoguchi E, Grosby MJ, Glimcher LH, Bhan AK, Tonegawa S. Spontaneous development of inflammatory bowel disease in T cell receptor mutant mice. *Cell* 1993;75:275-82.
29. Kühn R, Löhler J, Rennick D, Rajewsky K, Müller W. Interleukin-10-deficient mice develop spontaneous chronic enterocolitis. *Cell* 1993;75:263-74.
-
-



Hindawi
Submit your manuscripts at
<http://www.hindawi.com>

