

FEBRUARY 12, 2001

Human Disease Gene Survey Yields Underlying Principles

By compiling and categorizing 923 genes that malfunction in inherited diseases, Howard Hughes Medical Institute (HHMI) researchers have discerned patterns that indicate that this approach might be a powerful new tool for understanding genetic contributions to human diseases.

Future analyses of newly discovered human disease genes revealed through the human genome sequencing projects, say the scientists, will give rise to new approaches to understanding and treating disease based on fundamental principles.

"These insights represent only the beginning of the kinds of discoveries of general principles that can be made as a comprehensive list of disease genes is developed and compared to the list of all human genes."

— David L. Valle

The results of the survey were published in the journal *Nature* on February 12, 2001, by HHMI investigator David L. Valle and colleagues Gerardo Jimenez-Sanchez and Barton Childs at The Johns Hopkins University School of Medicine. The article is part of a collection of articles published by *Nature* that discusses the implications of sequencing the human genome.

"Most of us who have studied genetic disease have focused on one particular disease, or perhaps a small set of related diseases," said Valle. "However, this is a classic example of stepping back to take a look at the forest instead of specific trees. We decided to collect a list of disease genes, to correlate those genes with various aspects of the disease and see if we could begin to discern general patterns."

The scientists began by assembling a list of 923 genes that cause inherited disease, drawing the list from the seventh edition of *Metabolic and Molecular Bases of Inherited Diseases* and from the *Online Mendelian Inheritance in Man* database. Next, they categorized each disease gene based on the function of its protein product. This strategy yielded four major categories:

genes that encode enzymes; genes that encode proteins that influence the function of a second protein for example, stabilizing, activating or folding another protein; genes that encode receptors; and genes that encode transcription factors.

Valle and his colleagues next scored each gene in these four major categories based on the clinical features of the disease including age of onset, mode of inheritance, frequency, severity, extent of tissue involvement and association with malformations.

The analysis yielded a number of striking general insights, said Valle. For example, the scientists found that genes that encoded transcription factors are over-represented among genes that cause malformation diseases that begin *in utero*. This frequency, said Valle, reflects the importance of transcription factors in embryonic development.

The analyses also indicated that "an extraordinarily high" fraction of diseases with onset in the first year of life are caused by defects in genes that encode enzymes. This finding was expected, wrote the scientists, because the mother's metabolic system protects the fetus from enzyme deficiencies until birth. Metabolic deficiencies are usually identified after birth.

The scientists found that diseases caused by genes that encode enzymes are primarily recessive, while those caused by genes whose proteins influence other proteins are evenly split between dominant and recessive. Also, diseases caused by transcription factors are more likely to be dominant, found the researchers.

Each of the four major functional categories of gene showed a different peak age of onset, found the scientists. Diseases due to transcription factors peaked *in utero*; those due to enzymes in the first year; those for receptors between one year and puberty, and those due to protein modifiers in early adulthood.

"These insights represent only the beginning of the kinds of discoveries of general principles that can be made as a comprehensive list of disease genes is developed and compared to the list of all human genes," said Valle. "Clearly, the human genome projects will enable us to make this comparison, and we find the prospect very exciting."

Valle added that insights arising from such comparisons will likely alter the fundamental view of human disease. "In the past, physicians have viewed disease as something that is visited upon us from outside. But the biologically correct perspective is that disease is really a consequence of living. As researchers continue to identify these patterns of disease, we will gain a more global view of the biology of disease as a normal consequence of natural selection and the evolution of our species.

"So, instead of having one group of people, namely physicians, interested in disease and another, namely biologists, interested in evolution and normal biology, we will come to understand that both groups are working on two aspects of the same population. We will probably find that each group will be

better served by learning from the other."