Genetic disease and the genome

Rita Shiang
Department of Human Genetics

The focus in the laboratory is to identify and characterize genes that cause human genetic disease. The diseases studied in the laboratory include the craniofacial disorder, Treacher Collins syndrome, the mental retardation syndrome, Cri du Chat and Wolfram Syndrome, a neurodegenerative disorder whose main symptoms include diabetes insipidus, diabetes mellitus, optic atrophy and deafness. These disorders have been localized to human chromosomes 5 and 4q. Projects include positional cloning of a second locus for Wolfram Syndrome and identification of mouse homologues of genes deleted in Cri du Chat. A large focus of the laboratory is the characterization of the Treacher Collins syndrome (TCOF1) gene.

Bioinformatics information is important in every step of human disease gene identification and characterization. Identification of a disease gene's chromosomal location includes the use of known polymorphic markers or the characterization of new ones in linkage analysis. This is the first step in the cloning of the disease gene. Once a region of the chromosome has been targeted, the use of sequence information from the human genome project aids in the identification and characterization of candidate genes. Known and predicted genes can be identified, expression data of genes extracted and homologues in other species with known functions or protein motifs can be identified to aid in determination of good candidate genes. Mutation detection is simplified by the knowledge of intron/exon boundaries and possible splice products of a transcript.

Once a gene has been identified to cause a particular disorder, the function of the gene needs to be elucidated. Bioinformatics information can also aid in this endeavor. Structure and homology information will give insight into important protein domains. For example, the Treacher Collins syndrome protein, treacle, was predicted to have phosphorylation and nuclear and nucleolar localization signals. The protein has since been confirmed to be a nucleolar phosphoprotein by localization studies using GFP-fusion constructs and phosphorylation studies. In addition, the protein is phosphorylated in a cell-cycle dependent manner. Identification of homologues of the gene of interest in model organisms will aid in modeling the disorder for further study, which can later be used in possibly therapies. No homologues of the TCOF1 gene were identified in lower organisms such as yeast, the worm, *C. elegans* or *Drosophila*. Thus functional analysis in these organisms is not an option. The mouse homologue of the TCOF1 gene has been identified. The developmental expression of the gene has been delineated using whole-mount *in-situ* hybridization and the gene is expressed in an appropriate temporal and spatial manner. A conditional knockout mouse is currently being constructed to study the function of the gene in development.