TABLE 82-5 GENES AND LOCI INVOLVED IN MONO- AND POLYGENIC FORMS OF DIABETES

Disorder	Genes or Susceptibility Locus	Chromosomal Location	Other Factors
Monogenic permanent neonatal diabetes mellitus	*KCNJ11* (inwardly rectifying potassium channel Kir6.2)	11p15.1	AD
	GCK (glucokinase)	7p15-p13	AR
	INS (insulin)	11p15.5	AR, hyperproinsulinemia
	ABCC8 (ATP-binding cassette, subfamily c, member 8; sulfonylurea receptor)	11p15.1	AD or AR
	GLIS3 (GLIS family zinc finger protein 3)	9p24.2	AR, diabetes, congenital hypothyroidism
Maturity-onset diabetes of the young (MODY): Monogenic forms of diabetes mellitus			
MODY 1	*HNF4α* (hepatocyte nuclear factor 4α)	20q12-q13.1	AD inheritance
MODY 2	*GCK* (glucokinase)	7p15-p13	
MODY 3	*HNF1α* (hepatocyte nuclear factor 1α)	12q24.2	
MODY 4	*IPF1* (insulin receptor substrate)	13q12.1	
MODY 5 (renal cysts, diabetes)	*HNF1β* (hepatocyte nuclear factor 1β)	17cen-q21.3	
MODY 6	*NeuroD1* (neurogenic differentiation factor 1)	2q32	
MODY 7	*KLF1* (Kruppel-like factor 1)	19p13.13-p13.12	
MODY 8	*CEL* (carboxyl ester lipase)	9q34.3	
MODY 9	*PAX4* (paired box transcription factor 4)	7q32	
MODY 10	*INS* (insulin)	11p15.5	
MODY 11	*BLK* (B-lymphocyte-specific tyrosine kinase)	8p23-p22	
Diabetes mellitus type 2; loci and genes linked and/or associated with susceptibility for diabetes mellitus type 2	Genes and loci identified by linkage/association studies		Heavily influenced by diet, energy expenditure, obesity
	PPARG, KCNJ11/ABCC8, TCF7L2, IGF2BP2, CDKAL1, SLC30A8, CDKN2A/B, HHEX, FTO, HNF1B, NOTCH2, THADA, ADAMSTS9, JAZF1, CDC122/CAMK1D, KCNQ1, TSPAN8/LGR5, IRS1, DUSP9, PROX1, BCK11A, G6PC2, GCKR, ADCY5, SLC2A2, WFS1, ZBED3, DGKB/TMEM195, GCK, KLF14, TP53INP1, GLIS3, TLE4, ADRA2A, CENTD2, CRY2, FADS1, MADD, MTNR1B, HMGA1, HNF1A, IGF1A, IGF1, C2CD4B, PRC1, VPS13C, ZFAND6, GIPR		

Abbreviations: AD, autosomal dominant; AR, autosomal recessive; MODY, maturity onset diabetes of the young.

The identification of genetic variations and environmental factors that either predispose to or protect against disease is essential for predicting disease risk, designing preventive strategies, and developing novel therapeutic approaches. The study of rare monogenic diseases may provide insight into some of the genetic and molecular mechanisms important in the pathogenesis of complex diseases. For example, the identification of the genes causing monogenic forms of permanent neonatal diabetes mellitus or maturity-onset diabetes defined them as *candidate genes* in the pathogenesis of diabetes mellitus type 2 (Tables 82-2 and 82-5). Genome scans have identified numerous genes and loci that may be associated with susceptibility to development of diabetes mellitus in certain populations. Efforts to identify susceptibility genes require very large sample sizes, and positive results may depend on ethnicity, ascertainment criteria, and statistical analysis. Association studies analyzing the potential influence of (biologically functional) SNPs and SNP haplotypes on a particular phenotype are providing new insights into the genes involved in the pathogenesis of these common disorders. Large variants ([micro]deletions, duplications, and inversions) present in the human population also contribute to the pathogenesis of complex disorders, but their contributions remain poorly understood.

Linkage and Association Studies There are two primary strategies for mapping genes that cause or increase susceptibility to human disease: (1) classic linkage can be performed based on a known genetic model or, when the model is unknown, by studying pairs of affected relatives; or (2) disease genes can be mapped using allelic association studies (Table 82-6).

GENETIC LINKAGE *Genetic linkage* refers to the fact that genes are physically connected, or linked, to one another along the chromosomes. Two fundamental principles are essential for understanding the concept of linkage: (1) when two genes are close together on a chromosome, they are usually transmitted together, unless a recombination event separates them (Figs. 82-6); and (2) the odds of a crossover, or recombination event, between two linked genes is proportional to the distance that separates them. Thus, genes that are farther apart are more likely to undergo a recombination event than genes that are very close together. The detection of chromosomal loci that segregate with a disease by linkage can be used to identify the gene responsible for the disease (*positional cloning*) and to predict the odds of disease gene transmission in genetic counseling.

Polymorphisms are essential for linkage studies because they provide a means to distinguish the maternal and paternal chromosomes in an individual. On average, 1 out of every 1000 bp varies from one person to the next. Although this degree of variation seems low (99.9% identical), it means that >3 million sequence differences exist between any two unrelated individuals and the probability that the sequence at such loci will differ on the two homologous chromosomes is high (often >70–90%). These sequence variations include variable number of tandem repeats (VNTRs), short tandem repeats (STRs), and SNPs. Most STRs, also called *polymorphic microsatellite markers*, consist of di-, tri-, or tetranucleotide repeats that can be characterized readily using the polymerase chain reaction (PCR). Characterization of SNPs, using DNA chips or beads, permits comprehensive analyses of genetic variation, linkage, and association studies. Although these sequence variations often have no apparent functional consequences, they provide much of the basis for variation in genetic traits.

In order to identify a chromosomal locus that segregates with a disease, it is necessary to characterize polymorphic DNA markers from affected and unaffected individuals of one or several pedigrees. One can