TABLE 82-1 SELECTED DATABASES RELEVANT FOR GENOMICS AND GENETIC DISORDERS

Site	URL	Comment
National Center for Biotechnology Information (NCBI)	*http://www.ncbi.nlm.nih.gov/*	Broad access to biomedical and genomic information, literature (PubMed), sequence databases, software for analyses of nucleotides and proteins Extensive links to other databases, genome resources, and tutorials
National Human Genome Research Institute	*http://www.genome.gov/*	An institute of the National Institutes of Health focused on genomic and genetic research; links providing information about the human genome sequence, genomes of other organisms, and genomic research
Catalog of Published Genome-Wide Association Studies	*http://www.genome.gov/GWAStudies/*	Published high-resolution genome-wide association studies (GWAS)
Ensembl Genome browser	*http://www.ensembl.org*	Maps and sequence information of eukaryotic genomes
Online Mendelian Inheritance in Man	*http://www.ncbi.nlm.nih.gov/omim*	Online compendium of Mendelian disorders and human genes causing genetic disorders
Office of Biotechnology Activities, National Institutes of Health	*http://oba.od.nih.gov/oba*	Information about recombinant DNA and gene transfer; medical, ethical, legal, and social issues raised by genetic testing; medical, ethical, legal, and social issues raised by xenotransplantation
American College of Medical Genetics and Genomics	*http://www.acmg.net/*	Extensive links to other databases relevant for the diagnosis, treatment, and prevention of genetic disease
American Society of Human Genetics	*http://www.ashg.org*	Information about advances in genetic research, professional and public education, social and scientific policies
Cancer Genome Anatomy Project (CGAP)	*http://cgap.nci.nih.gov/*	Information about gene expression profiles of normal, precancer, and cancer cells
GeneTests	*http://www.genetests.org/*	International directory of genetic testing laboratories and prenatal diagnosis clinics; reviews and educational materials
Genomes Online Database (GOLD)	*http://www.genomesonline.org/*	Information on published and unpublished genomes
HUGO Gene Nomenclature	*http://www.genenames.org/*	Gene names and symbols
MITOMAP, a human mitochondrial genome database	*http://www.mitomap.org/*	A compendium of polymorphisms and mutations of the human mitochondrial DNA
International HapMap Project	*http://www.hapmap.org/*	Catalogue of haplotypes in different ethnic groups relevant for association studies and pharmacogenomics
ENCODE	*http://www.genome.gov/10005107*	*Encyclopedia of DNA Elements*; catalogue of all functional elements in the human genome
Dolan DNA Learning Center, Cold Spring Harbor Laboratories	*http://www.dnalc.org/*	Educational material about selected genetic disorders, DNA, eugenics, and genetic origin
The Online Metabolic and Molecular Bases of Inherited Disease (OMMBID)	*http://www.ommbid.com/*	Online version of the comprehensive text on the metabolic and molecular bases of inherited disease
Online Mendelian Inheritance in Animals (OMIA)	*http://omia.angis.org.au/*	Online compendium of Mendelian disorders in animals
The Jackson Laboratory	*http://www.jax.org/* *http://www.informatics.jax.org*	Information about murine models and the mouse genome Mouse genome informatics

Note: Databases are evolving constantly. Pertinent information may be found by using links listed in the few selected databases.

Saccharomyces cerevisiae, *Caenorhabditis elegans*, and *Drosophila melanogaster*; bacteria (e.g., *E. coli*); and Archaea, viruses, organelles (mitochondria, chloroplasts), and plants (e.g., *Arabidopsis thaliana*). Genomic information of infectious agents has significant impact for the characterization of infectious outbreaks and epidemics. Other ramifications arising from the availability of genomic data include, among others, (1) the comparison of entire genomes (*comparative genomics*), (2) the study of large-scale expression of RNAs (*functional genomics*) and proteins (*proteomics*) to detect differences between various tissues in health and disease, (3) the characterization of the variation among individuals by establishing catalogues of sequence variations and SNPs (HapMap Project), and (4) the identification of genes that play critical roles in the development of polygenic and multifactorial disorders.

CHROMOSOMES The human genome is divided into 23 different chromosomes, including 22 autosomes (numbered 1–22) and the X and Y sex chromosomes (Fig. 82-1). Adult cells are *diploid*, meaning they contain two homologous sets of 22 autosomes and a pair of sex chromosomes. Females have two X chromosomes (XX), whereas males have one X and one Y chromosome (XY). As a consequence of meiosis, germ cells (sperm or oocytes) are haploid and contain one set of 22 autosomes and one of the sex chromosomes. At the time of fertilization, the diploid genome is reconstituted by pairing of the homologous chromosomes from the mother and father. With each cell division (mitosis), chromosomes are replicated, paired, segregated, and divided into two daughter cells.

STRUCTURE OF DNA DNA is a double-stranded helix composed of four different bases: adenine (A), thymidine (T), guanine (G), and cytosine (C). Adenine is paired to thymidine, and guanine is paired to cytosine, by hydrogen bond interactions that span the double helix (Fig. 82-1). DNA has several remarkable features that make it ideal for the transmission of genetic information. It is relatively stable, and the double-stranded nature of DNA and its feature of strict base-pair complementarity permit faithful replication during cell division. Complementarity also allows the transmission of genetic information from DNA → RNA → protein (Fig. 82-2). mRNA is encoded by the so-called sense or coding strand of the DNA double helix and is translated into proteins by ribosomes.

The presence of four different bases provides surprising genetic diversity. In the protein-coding regions of genes, the DNA bases are arranged into codons, a triplet of bases that specifies a particular amino acid. It is possible to arrange the four bases into 64 different triplet codons (4^3). Each codon specifies 1 of the 20 different amino acids, or a regulatory signal such as initiation and stop of translation. Because there are more codons than amino acids, the genetic code is degenerate; that is, most amino acids can be specified by several different codons. By arranging the codons in different combinations and in