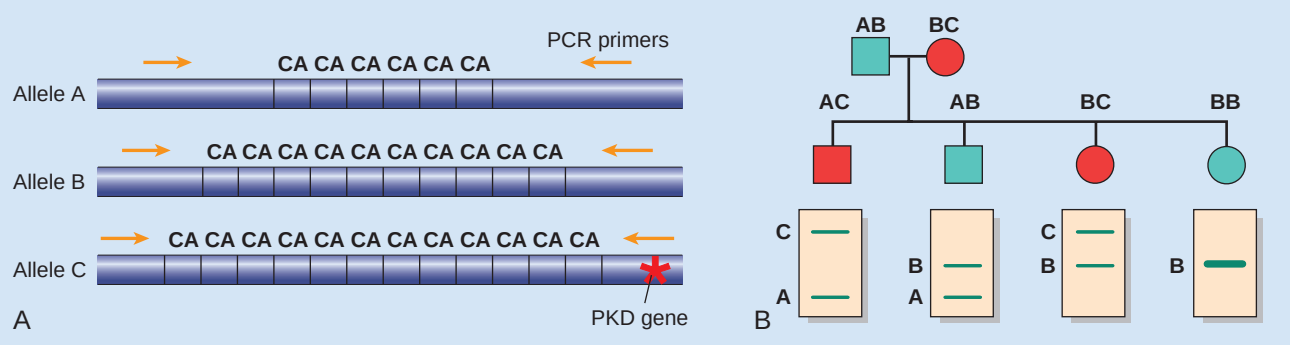

Figure 5-32 DNA polymorphisms resulting from a variable number of CA repeats. The three alleles produce PCR products of different sizes, thus identifying their origins from specific chromosomes. In the example depicted, allele C is linked to a mutation responsible for autosomal dominant polycystic kidney disease (PKD). Application of this to detect progeny carrying the disease-related gene (red symbols) is illustrated in one hypothetical pedigree. Males (squares); females (circles).

the same chromosome are almost certain to cosegregate during meiosis, due to the extremely low chance of a crossover event happening between them. Thus, the closer two loci are, the safer it is to assume that they will travel together in family pedigrees. In the event of a challenging or unknown pathogenic allele, a diagnostics lab can instead choose simply to examine nearby marker loci in the context of the family pedigree, as a surrogate approach. **The two types of genetic polymorphisms most useful for linkage analysis are SNPs (described earlier) and repeat-length polymorphisms known as minisatellite and microsatellite repeats.**

Human DNA contains short repetitive sequences of DNA giving rise to what are called repeat-length polymorphisms. These polymorphisms are often subdivided on the basis of their length into microsatellite repeats and minisatellite repeats. Microsatellites are usually less than 1 kilobase and are characterized by a repeat size of 2 to 6 base pairs. Minisatellite repeats, by comparison, are larger (1 to 3 kilobases), and the repeat motif is usually 15 to 70 base pairs. It is important to recall that the number of repeats, both in microsatellites and minisatellites, is extremely variable within a given population, and hence these stretches of DNA can be used quite effectively to establish genetic identity for linkage analysis. Microsatellites and the smaller minisatellites can be readily distinguished by using PCR primers that flank the repeat region. Figure 5-32 depicts the application of microsatellite linkage analysis to the *PKD1* gene (historically very difficult to sequence) for the familial diagnosis of adult polycystic kidney disease. It can be seen that the longest microsatellite allele is linked in the family to the disease allele and can be used to track transmission.

Assays to detect genetic polymorphisms are also important in many other areas of medicine, including in the determination of relatedness and identity in transplantation, cancer genetics, paternity testing, and forensic medicine. Since microsatellite markers are scattered throughout the human genome and have such a high level of polymorphism, they are ideal for differentiating between two individuals and to follow transmission of the marker from parent to child. Panels of microsatellite marker PCR assays have been extensively validated and are now routinely used for determining paternity and for criminal investigations. Since PCR can be performed even with highly degraded biologic samples, DNA technology is critical in forensic identifications. The same assays are regularly applied to the detection and quantification of transplant chimerism in allogeneic hematopoietic stem cell transplant patients, by looking for relative amounts of both donor- and host-specific microsatellite markers in host blood and blood cell subsets.

Polymorphisms and Genome-Wide Analyses

Beyond the clinic, linkage-based analysis has a long and storied history as a critical tool for discovery in the research laboratory. Many Mendelian diseases (including cystic fibrosis) were originally localized to candidate chromosomal locations using family pedigrees, testing a variety of candidate marker loci in a search for linkage disequilibrium, followed by subsequent refinement and testing of new nearby markers. Linkage studies have been similarly invaluable for identifying genes responsible for various phenotypes in laboratory animal models. However, similar analyses of complex (multifactorial) disorders have been unsuccessful since conventional linkage studies lack the statistical power to detect variants with small effects and low penetrance, which are thought to contribute to complex disorders.

To address this problem, researchers have utilized SNP genotyping array technology to perform large-scale linkage studies of complex diseases (e.g., type 2 diabetes, hypertension), which are termed genome wide association studies (GWAS). **In GWAS, large cohorts of patients with and without a disease (rather than families) are examined across the entire genome for common genetic variants or polymorphisms that are overrepresented in patients with the disease.** This identifies regions of the genome that contain a variant gene or genes that confer disease susceptibility and provides a springboard for further targeted research to find the true causative factor.

In addition to polygenic diseases, GWASs also have led to the identification of genetic loci that modulate common quantitative traits in humans, such as height, body mass, hair and eye color, and bone density. There has been much dispute about the value of GWAS studies, with criticisms often focusing on their underlying hypothesis that common disease risk could be explained by examining associations with common genetic variants. In support of that criticism, the combined relative risk from the associated variants that