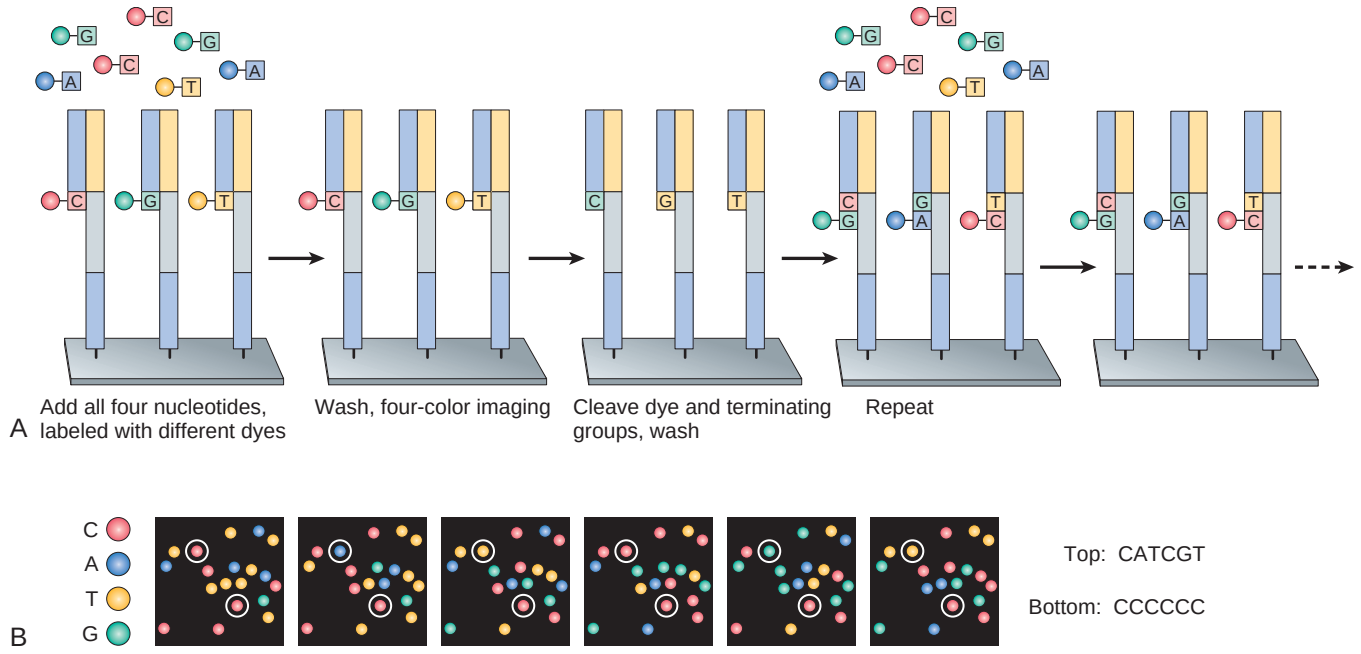

Figure 5-33 Principle of next-generation sequencing. Several alternative approaches currently are available for "NextGen" sequencing, and one of the more commonly used platforms is illustrated. **A**, Short fragments of genomic DNA ("template") between 100 and 500 base pairs in length are immobilized on a solid phase platform such as a glass slide, using universal capture primers that are complementary to adapters that have previously been added to ends of the template fragments. The addition of fluorescently labeled complementary nucleotides, one per template DNA per cycle, occurs in a "massively parallel" fashion, at millions of templates immobilized on the solid phase at the same time. A four-color imaging camera captures the fluorescence emanating from each template location (corresponding to the specific incorporated nucleotide), following which the fluorescent dye is cleaved and washed away, and the entire cycle is repeated. **B**, Powerful computational programs can decipher the images to generate sequences complementary to the template DNA at the end of one "run," and these sequences are then mapped back to the reference genomic sequence in order to identify alterations. (Reproduced with permission from Metzker M: Sequencing technologies—the next generation. Nat Rev Genet 11:31-46, 2010, © Nature Publishing Group.)

and a detailed discussion is beyond the scope of this text. However, it is worth describing the basic steps necessary to process this type of data in a generic human DNA context:

- *Alignment.* Alignment is the process by which the sequencing reads from a sample (which are individually uninformative) are mapped onto the appropriate reference genome, where they can be viewed and interpreted in context.
- *Variant calling.* This process involves "walking" across the reference genome and evaluating all of the sequence data that mapped to each position and comparing it with the reference sequence. The more reads that cover a particular location (sequencing depth), the more likely that a variant will be detected if present. If a locus shows sufficient evidence of a difference from the reference sequence, a variant call will be made.
- *Variant annotation and interpretation.* Called variants can be annotated with various sources of information (e.g., gene names, coding changes and protein effects predictions, SNP IDs, information from databases of both benign and pathogenic variants, clinical information). In the clinical laboratory the data may then be ready for interpretation and reporting.

Clinical Applications of NGS DNA Sequencing

As discussed, any DNA sample can be analyzed by NGS. However, the DNA needs to first be prepared for sequencing, and the choice of technique determines what data will be collected. In order to allow for binding and focal amplification on the instrument, input DNA must be in the form of small segments (<500 bp) with instrument-compatible linker oligonucleotide sequences incorporated at the ends. For whole genome sequencing this often involves automated genomic DNA fragmentation followed by ligation of adapter oligonucleotides. The final product is termed a *library*. There are many methods for NGS library preparation from genomic DNA, depending on the question and desired result. Currently in the clinical labs, most applications of NGS are targeted toward constitutional genetic disease and cancer diagnostics, using a few different basic approaches:

- *Targeted sequencing.* Most clinical laboratory NGS tests today fall in this category. With targeted sequencing just one gene or a panel of genes minimizes sequencing costs as well as the time and expense required for manual interpretation and clinical reporting. Sample preparation can be performed either by subselecting relevant clones from a whole genome library via custom complementary probes or by alternate preparations from genomic DNA, such as multiplex PCR. Single gene assays are more common for constitutional testing, and many laboratories that previously performed costly whole gene analyses by Sanger sequencing (i.e., comprehensive *CFTR* sequencing) are now moving those assays to NGS platforms. Analysis of large panels of genes (up to 100 or more) was previously unfeasible, but is now a commonplace approach for children with