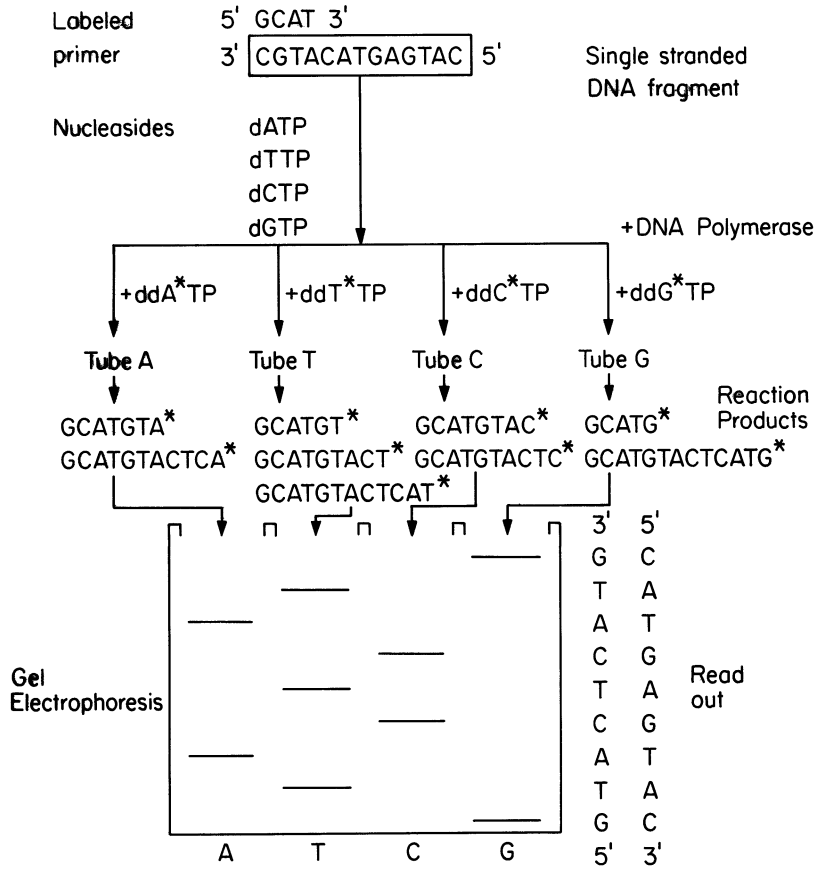

Figure 8.9. Example of a sequencing gel to obtain the nucleotide sequence of a DNA fragment. See text for details.

approach described above. For the purposes of automation a fluorescent dye is used for each ddNTP. Four different dyes are used to distinguish A, T, C, and G. These labelled fragments are then read in automated sequencing machines. These machines rely on 104 glass capillaries; capillary electrophoresis is used to separate the fragments by size. As the fragments exit the capillary, a laser beam detects the color of the dye at the end of the fragment. The information is led to a computer and the 500-letter sequence is determined. One company, Celera, expects to use such machines to read 100 million letters of DNA sequence per day. These sequences are stored in a computer; computer alogrithms are then used to align the overlapping sequences.

Because of the large number of fragments, this information processing is very challenging. For the human genome 70 million separate sequences are necessary to achieve sufficient overlap to reconstruct the whole human genome (about 3 billion letters). This technique cannot do a perfect reconstruction, but one that is effectively complete. One