Some of the other algorithms and concepts that can be found in more technical literature in the field of bioinformatics are described in the paragraphs to follow.

Gollery (2008) and many other sources describe Hidden Markov Models (HMM) and related methods, like the HMMER package, Sequence Analysis Method (SAM), and the PSI-BLAST algorithm. This area has its own set of HMM databases with names like Pfam, PANTHER, COG, and metaSHARK. We won’t discuss these programs here; we mention them only to whet your appetite, in case bioinformatics is an area you would like to study further. Googling any of these terms will provide a plethora of information.

Shui Qing Ye (2008) describes some additional DNA and genome analysis tools, some for phylogenetic analysis, SNP analysis, haplotype analysis, and regulation of gene expression. SAGE is one of the tools described. SNP, pronounced “snip,” stands for Simple Nucleotide Polymorphism. A polymorphism is a substitution in one base of the code of DNA and is what makes some people have blue eyes and other people have brown eyes, for example, among many examples of polymorphisms we could list. One of the co-authors spent his doctoral candidate years looking for polymorphisms in blood enzymes of mice; this was years ago, shortly after Watson and Crick discovered the DNA molecule. So, you can see that basic genetics and polymorphisms, DNA, and RNA sequences are still in active service in the twenty-first century as we continue to work out the details of bioinformatics. Haplotype analysis involved looking at the haploid genotype. Complete genetic complements of most organisms contain two of each type of chromosome. A set of one of each chromosome is called a haploid, with one haploid contributed by the male and other by the female. The complements of genes in a haploid are called a haplotype. Haplotypes are inherited as units; they consist of a combination of alleles at different markers along the same chromosome.

Serial Analysis of Gene Expression, or SAGE, was developed in the mid-1990s. It provides an overview of a cell’s complete gene activity. The technique captures RNAs and then allows a quantitative analysis of the transcripts of information that are made from these RNAs. It has the potential of developing a catalog of not only the mRNAs present in a cell, but also their prevalence.

For those who would like to gain a practical understanding of what bioinformatics is all about, we recommend Shui Qing Ye’s book *Bioinformatics: A Practical Approach*, as it contains not only some unique biocomputing tools including use of Perl and R languages, but also useful web sites and database listings. Particularly, we refer you to “Tutorials” in Ye (2008), like the one on page 352, which takes you step by step through analyzing a protein sequence, and the one on page 456, which again takes you step by step through the process of creating a Perl script.

In her book on bioinformatics, Parida (2007) clearly presents the opinion that most bioinformatic success stories require algorithmic and statistical ingenuity. As such, she discusses and develops a considerable number of algorithms for better pattern discovery in the bioinformatic field. These include such unique names as

- Prim’s algorithm
- Fitch’s algorithm
- Discovery algorithm