nature immunology

A hornet's nest for $1,000

Thirty years ago, two DNA-sequencing methods ushered in the study of genomes. Are we ready for the next step?

This year James Watson received his sequenced genome on two DVDs: one vast jumble of nucleotides transposed into another vast jumble of digitized optical information. Producing Watson's genome sequence cost approximately one million dollars, a fraction of that spent elucidating the first human genome sequences. Two other sobering facts about the event of Watson's completed genome are the rapid pace of technological change since efficient sequencing methods were published and the inevitability, now, that anyone will soon be able to have their genome sequenced relatively cheaply. When that day arrives, great insight into comparative genomics will no doubt follow.

But for those who seek information about their genome, so too will anxiety follow. Those who find their genome contains mutations associated with Crohn's disease, arthritis, Huntington's disease or breast cancer, for example, will face difficult choices; hopefully, genetic counseling might alleviate some of that concern. Yet that day could also bring the arrival of insurance companies requiring, in an effort to reduce risk, sequence information on 'blacklisted' mutations. The ever-increasing sophistication and affordability of DNA-sequencing technology has (perhaps) made imminent a 'fool's paradise'.

It has been a short 30 years since two new methods for determining the nucleotide sequence of DNA were introduced in 1977. Allan Maxam and Walter Gilbert published "a chemical procedure that breaks a terminally labeled DNA molecule partially at each repetition of a base" (*Proc. Natl. Acad. Sci. USA.* **74**, 560–564 (1977)), whereas Frederick Sanger and Alan Coulson published a method based on the "use of the 2′,3′-dideoxy and arabinonucleoside analogues of the normal deoxynucleoside triphosphates, which act as specific chain-terminating inhibitors of DNA polymerase" (*Proc. Natl. Acad. Sci. USA.* **74**, 5463–5467 (1977)). It would be difficult to overstate the importance of these two methods for determining the nucleotide sequence of a contiguous piece of DNA, as previous methods by both Sanger and Gilbert yielded hard-earned and very limited sequence. In 1973, for example, Gilbert and Maxam, using a combination of methods, published the seventy-two base pair sequence of the *lac* operon (*Proc. Natl. Acad. Sci. USA.* **70**, 3581–3584 (1973)).

As often occurs with 'competing' inventions, however, the new methods of Sanger-Coulson and Maxam-Gilbert would not enjoy similar fates. Unlike the Maxam-Gilbert method, which required relatively toxic chemical reactions to cleave DNA at consecutive bases, the Sanger 'dideoxy' method used random chain termination through the incorporation of modified nucleotides lacking the 2′ hydroxyl groups required for the addition of the next base. Sanger's method thus caught on and is still widely in use.

Today, sequencing a piece of DNA often involves simply generating a product polymerase chain reaction that is then sent off to a sequencing facility, where a modified Sanger method is used to produce a sequence in a day or two. This is a far cry from the early days of 'in-laboratory' sequencing, when four dideoxy nucleotide–containing 'cocktails' (one for each base) and single-stranded M13 phage clones of the DNA of interest had to be prepared, and days were spent assembling a few hundred bases of contiguous sequence by reading 14- × 17-inch autoradiograms.

But such is the story of science and progress through technical achievements. Today's 'next-generation' sequencing technologies, such as 454 Life Sciences (Roche Diagnostics), have catapulted DNA sequencing to places not dreamed possible in 1977. These new methods can achieve upward of 20 million nucleotide bases of sequence in only a few hours, and an entire genome of any organism can be sequenced by a single researcher with only one preparation of DNA (assembling the first human genome sequences, in contrast, required 60 million sample preparations, 12 years and $2.7 billion dollars, for the National Institutes of Health–based effort). It was, in fact, the 454 technology that produced Watson's genome sequence for around $1 million dollars, a mark that will no doubt be surpassed by the present drive to produce a single genome sequence for as little as $1,000 (*Science* **311**, 1544–1546 (2006)).

Such prospects are intriguing. But what will be done with the information contained in peoples' genomes? Do people want to know which diseases loom in their future? Even worldly James Watson had qualms about the information in his genome; he released nearly all the data to the public so comparisons could be made with the previous genome sequences, yet he balked at supplying information about his apolipoprotein E gene—mutations in which are known to predispose a person toward Alzheimer's disease.

Another possible drawback in gaining information about their genome is that people may increasingly think of their lives as being determined by their genes, lives in which choice is seen as secondary to 'biological destiny'. Such a scenario is not too difficult to imagine. People today talk about the genetic inevitability of 'intelligence', 'bad behavior', 'obesity', 'rowdiness' and 'attention deficit'. Often less discussed is the implicit assumption that such 'biological determinism' necessarily entails the diminished expectation of personal responsibility—would this be even more so in a future with much more genomic information?

Reflecting on 'beginning of modern sequencing' in 1977 honors the scientists who made lasting contributions to the understanding of genomes. Yet we hope that the reflection also stirs a thought or two about how genomic information will be used. 'Knowledge is power' has perhaps few better examples than the day of the $1,000 genome. The status of key genes could have considerable consequences for peoples' lives, with the knowledge that diseases such as arthritis, lupus and leukemia are 'encoded' by their DNA. How much about this do people really want to know?