Future Implications of Genetic Engineering

John Corrigan
Department of Life Sciences
Indiana State University
Terre Haute, Indiana 479809

To understand the basis for future developments implicit in genetic engineering, it is applicable to review basic concepts and discoveries which support this area. Over the past decade, molecular genetic research has demonstrated that genes are able to function in almost any cellular environment and that this function persists even after substantial modification of gene sequences. Genetics and Biochemistry used to be considered as separate fields but in 1972, Jackson, Symonds and Berg at Stanford University fashioned a circlet of DNA containing a piece from a bacteriophage (bacterial virus) to which was chemically linked part of the DNA of a simian tumor virus also found as a circular gene (1). This work drew upon many decades of prior research conducted in the USA and internationally which commenced with isolation of the first bacterial viruses in 1915, a finding subsequently confirmed in many laboratories. By the 1960s, studies on the mechanism of various bacterial viruses had shown that certain types incorporate into the host genes after infection there to remain dormant for a variable period, while others multiply rapidly in a bacterial host, resulting in death of the host and release of virus progeny. This research also demonstrated the existence of circular genetic pieces in some bacteria which were extrachromosomal (outside the main body of DNA) and capable of autonomous duplication when a cell duplicated its main DNA. The circular pieces are called plasmids and the 1972 report from Berg's laboratory signified that a plasmid had been assembled which was of experimental rather than natural origin.

In the same year Cohen, Chang and Hsu also at Stanford, published a method for introducing circular plasmid DNA into the common laboratory bacillus *E. coli*, which exists in nature in the large intestine of humans (2). The following year, Cohen and Chang collaborating with Boyer and Helling at the University of California in San Francisco, announced that a plasmid carrying genes for resistance to antibiotics had been designed and inserted into *E. coli* where it was increased in quantity as the bacterium underwent its numerous cycles of doubling (3). This process of increasing copies of an engineered plasmid by inserting it into a fast-growing cell is one definition of the much-used term cloning. The early reports of recombinant-DNA technology, as this type of work is designated, include a demonstration that segments of complex cell DNA (eucaryotic cells), in this case from the African toad *Xenopus laevis*, could be enzymatically linked to plasmid genes of laboratory origin and cloned in simple cells (procaryotic cells) such as *E. coli* (4).

These pioneering discoveries are the foundations of a large body of research discoveries and biomedical applications utilizing gene construction or modification, followed by insertion of the novel gene into a cell where it multiplies (*replicates*) and its information may subsequently be *transcribed* into RNA, including stands of messenger RNA. Cells containing these synthetic genes can often be treated in a fashion which results in *translation* of the messenger RNA, resulting in a protein or peptide product. Readers who wish to review DNA replication, RNA synthesis (transcription) and protein synthesis (translation) will find a non-technical treatment in any college-level text of biology for non-majors. Figure 1 summarizes these relationships.

A list of hormones or regulatory molecules produced for medical or veterinary use by recombinant-DNA technology includes human insulin, human growth hormone (somatotropin), porcine growth hormone, interferon (a natural protein which inhibits viral invasion of animal cells) and somatostatin, involved in antagonizing the action of somatotropin. Many more peptide or protein factors are becoming available. These are

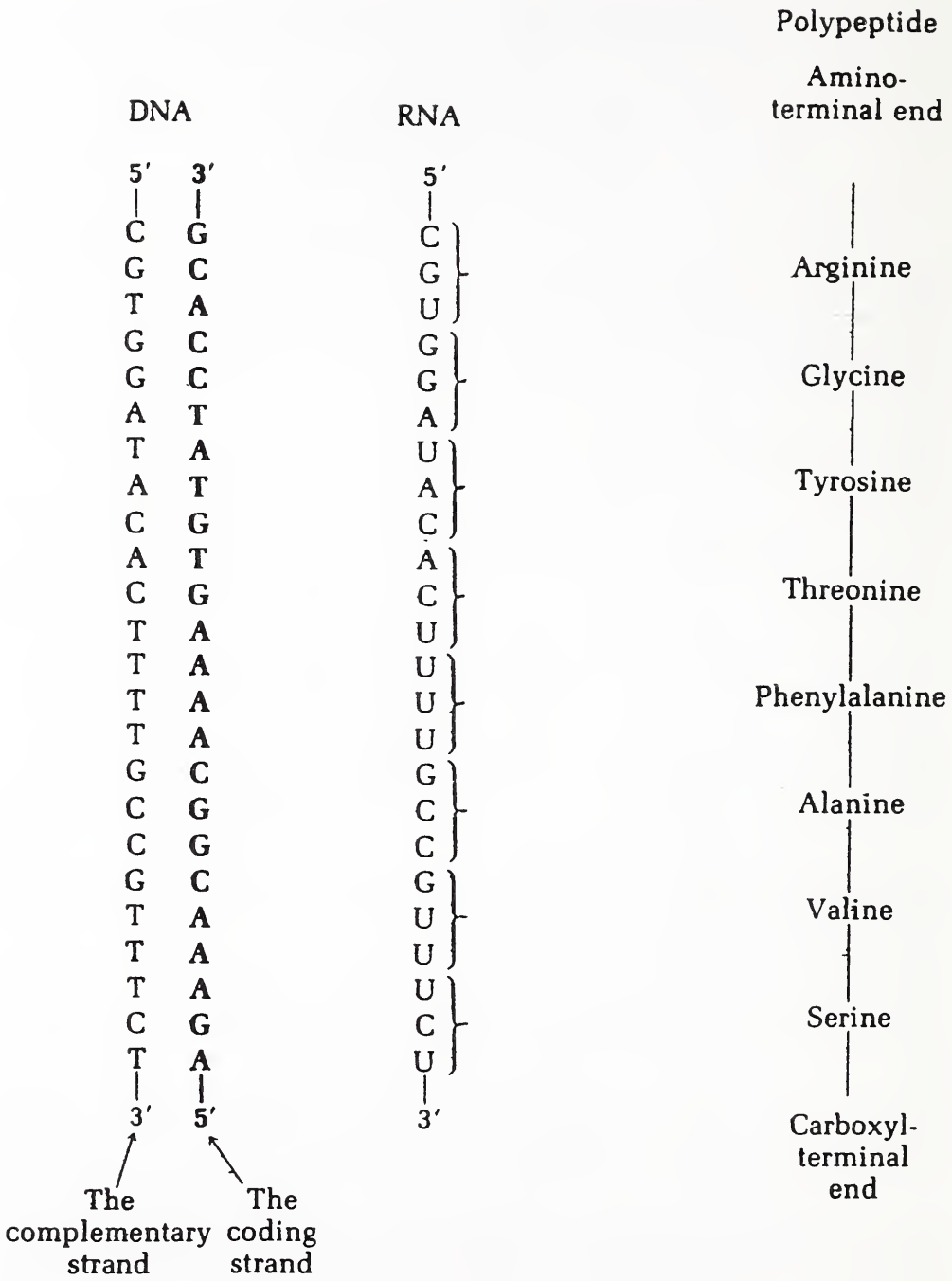

FIGURE 1. DNA Double Strand, Messenger RNA Transcript (complementary to Coding Strand) and translated Octapeptide. Triplets of nucleotides (bases) in DNA determine mRNA triplets which when translated determine the amino acid sequence in protein.

produced by introduction of enzymatically-modified plasmids into bacteria which subsequently multiply and synthesize the regulatory agent coded for by DNA in the appropriate plasmid.

In 1983, Brinster and collaborators at the School of Veterinary Medicine of The University of Pennsylvania described the incorporation of the gene for human growth hormone into the region of mouse DNA which codes for the synthesis of a regulatory protein stimulated by metals in the diet such as zinc. The synthetically produced fragment

containing DNA for human growth hormone, was then injected into the nucleus of a fertile mouse egg and the egg was implanted into the uterus of a laboratory mouse. When the mice were born they were fed a diet containing small amounts of zinc. Those which carried the synthetic DNA in their body cells responded by growing to a size larger than their litter mates which did not receive the human growth-hormone gene. One mouse from an egg which was injected with two copies of the recombined DNA grew to twice the size of its litter mates (5). Any animal altered by incorporating into its cells novel or alien genetic information is called a *transgenic* animal. The establishment of transgenic animals, where specific alien genes are incorporated into the germ line of fertile eggs, represents one of the most potent aspects of genetic engineering. This technology includes the possibility of establishing laboratory animal colonies carrying genetic anomalies identical to those which afflict humans. This is a bright prospect for future therapy for genetic diseases which comprise a large portion of the catalog of human suffering and tragedy. A comprehensive review of developments in this relatively new field has been published by Jaenisch at the Massachusetts Institute of Technology (6). Figure 2 illustrates the principal steps used for transgenic animal research.

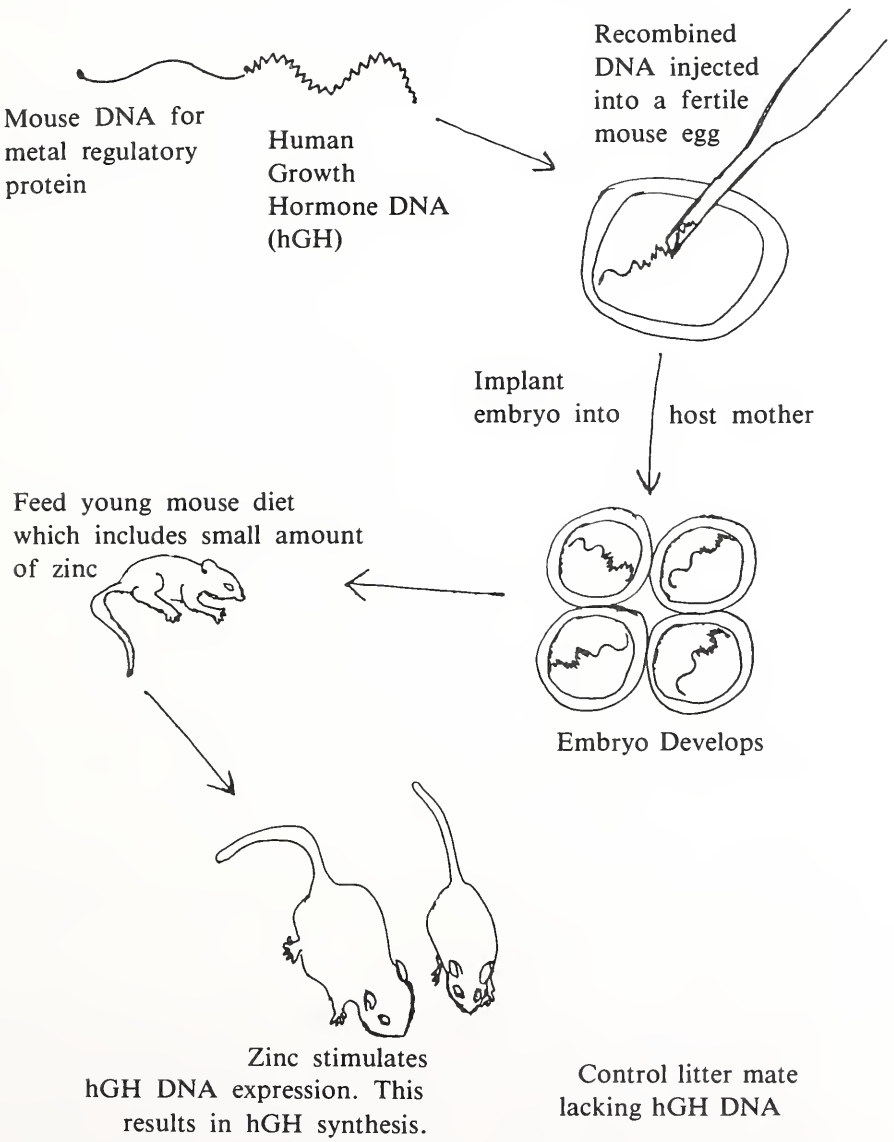

FIGURE 2. Transgenic animal production.

The same enzymes used to make specific cuts in DNA of various species, along with other enzymes used to ligate or splice the DNA fragments together, can be utilized for the identification of marker regions of DNA near genes which code for specific observable characteristics in an individual organism. Readers familiar with genetics will recognize that all the observable characters in an individual comprise the *phenotype*. On a chemical basis, the phenotype of an individual consists of all the proteins synthesized during translation. Controlling the individual proteins are the individual sequences of DNA, each containing the four chemical nucleotides or bases A, C, G, and T in linear array. These sequences code for the linear sequences of amino acids in each protein. The linear amino acid sequences in the DNA are the *genotypes* or genotypic features. Figure 3 shows the nucleotide code and its amino acid equivalents. Another standard term is *trait* which is used for any variaiton of an inherited character or characteristic. Eye color is a trait. When referring to variations in eye color, each color is called an *allele*. Every allele of a trait represents a mutation in the DNA.

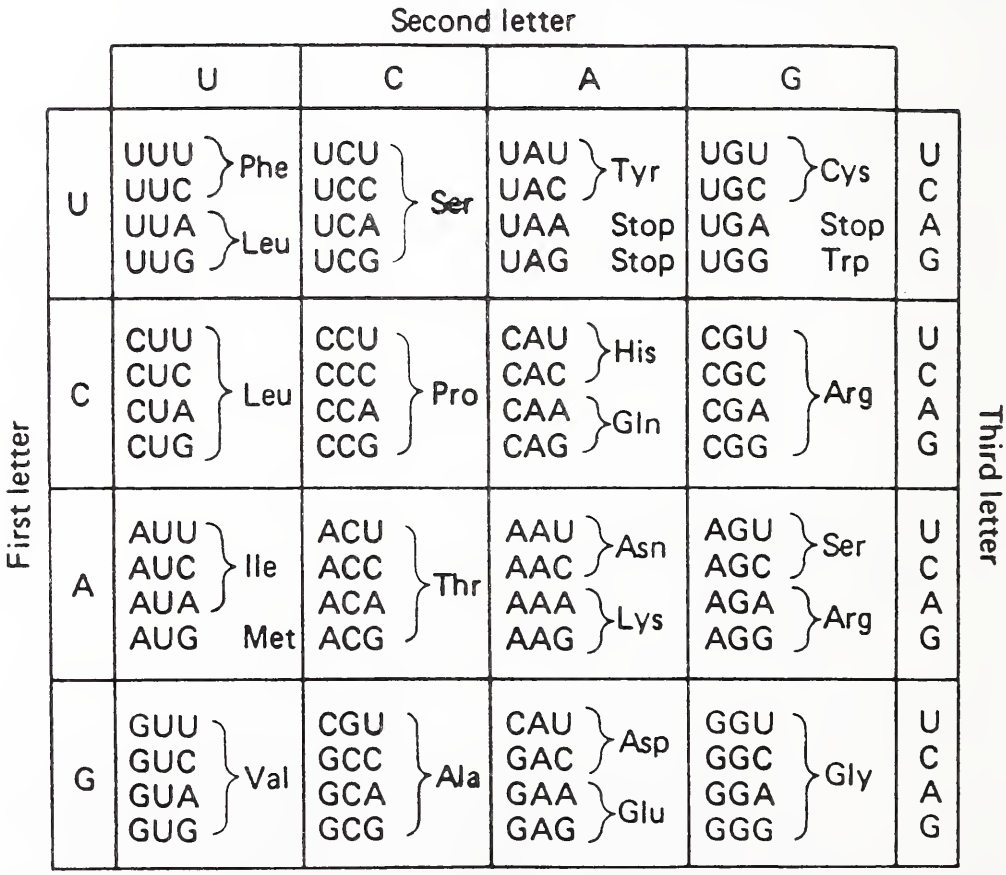

FIGURE 3. The Genetic Code in its RNA form. AUG also serves as a start codon. Three of the sixty-four triplet condons are stop codons. A few codons differ in the mitochondria. Otherwise the code appears to be universal.

In humans, the total collection of genetic information or *genome* contains between 50,000 and 200,000 genes, most of which either code for specific proteins or for RNA molecules which are either translated or play a role in translation (7). The thousands of genes are distributed over 46 pairs chromosomes and estimates are that there are some 3 billion bases (nucleotides) in double strands of complementary DNA, which make up the entire genome. Possibly only 10 percent of all the bases are ever expressed. Recent

work has revealed large sections of DNA present in repeat or redundant sequences and other sections contained in pseudogenes which are defective and can never be expressed (8). Over the past decade, several laboratories have concentrated on applying recombinant-DNA methods to the identification of markers associated with human genetic disease. Dr. Helen Donis-Keller of Collaborative Research in Bedford Massachusetts has reviewed progress in this area and states that, as of 1985, two hundred and forty-nine human genes had been assigned to specific chromosomes (9).

In many cases, the allele which calls attention to the existence of a specific gene is involved in a genetic disease, due to a mutation in the DNA which interferes with or prevents expression of the normal trait. By use of enzymes known as restriction endonucleases, since they make specific cuts in DNA without completely breaking it down into its four constituent bases, a number of enzyme-sensitive sites have been detected in human DNA (as well as that from other species). Many restriction endonucleases are now available from commerical suppliers and with these, a large number of non-overlapping restriction sites have been reported in human DNA. Dr. Donis-Keller points out that as of 1987, some 376 specific sites known as Restriction Fragment Length Polymorphisms or RFLPs have been assigned to specific human chromosomes (9). This means that enzymatically sensitive sites are mapped as RFLP loci and these provide the opportunity for cutting the huge pieces of chromosomal DNA down to more manageable fragments. It is important to emphasize that the RFLP loci or markers have no causal connection with the genetic diseases, but rather serve as signposts to aid in locating the disease-associated region for further study.

Figure 4 shows the main procedures used for RFLP identification. In the figure, white blood cells or other body cells are obtained from the parents, an affected child if there is one, and the fetus at risk for inheriting the condition or disease. Fetal cells are obtained either from aminotic fluid drawn at 16-18 weeks of gestation or chorionic villus samples obtained at 8 to 10 weeks of gestation. The cells are cultured, the genomic DNAs isolated and then each DNA isolate is treated with the same restriction endonucleases. The fragments of DNA are separated by size using an electric current passed through a sieve-like gel. After separation, the fragments are transferred to a nylon-like membrane to which they are made to bind tightly using specific environmental conditions. Then a radioactively labeled complementary DNA piece called a probe is incubated with the DNA fragments now bound to the membrane as single strands. The probe is from regions near the disease gene. After this incubation, the non-bound probe pieces are washed away and the gel is exposed to X-ray film. Within one to two weeks the film is developed and examined to see if RFLP fragments of the child at risk bind the probe made from regions associated with the disease. This method depends upon RFLPs which are spatially close enough to migrate (or during sex-cell production to cosegregate) with the gene that causes the disease. In figure 4, the fetus at risk has alleles 3 and 4 and is free of the disease. The mother and father (M and F on the autoradiograph) each carry the RFLP marking different alleles for the disease. In this illustration, the alleles numbered 1 and 2, if present individually, confer a carrier state on each parent. The markers are both present only when the child is homozygous (carries a marker from each parent). Accordingly, the fetus in question is not a carrier of either marker allele. Donis-Keller points out that a useful RFLP must be closely linked to the disease locus, where the criterion is not over 10 centimorgans away from the actual disease allele. (One centimorgan is 1 million nucleotides or base pairs). It is also important that the RFLPs flank the disease locus.

It has been estimated that sequence differences in the DNA of individuals occur on average, every 50 to 100 nucleotides. This permits large variation to exist when DNA is treated with a specific set of restriction endonucleases. Applications of this heterogeneity are now appearing in forensic analysis of blood, semen or tissue samples recovered where violent crimes have taken place. This application has already resulted in jury convictions

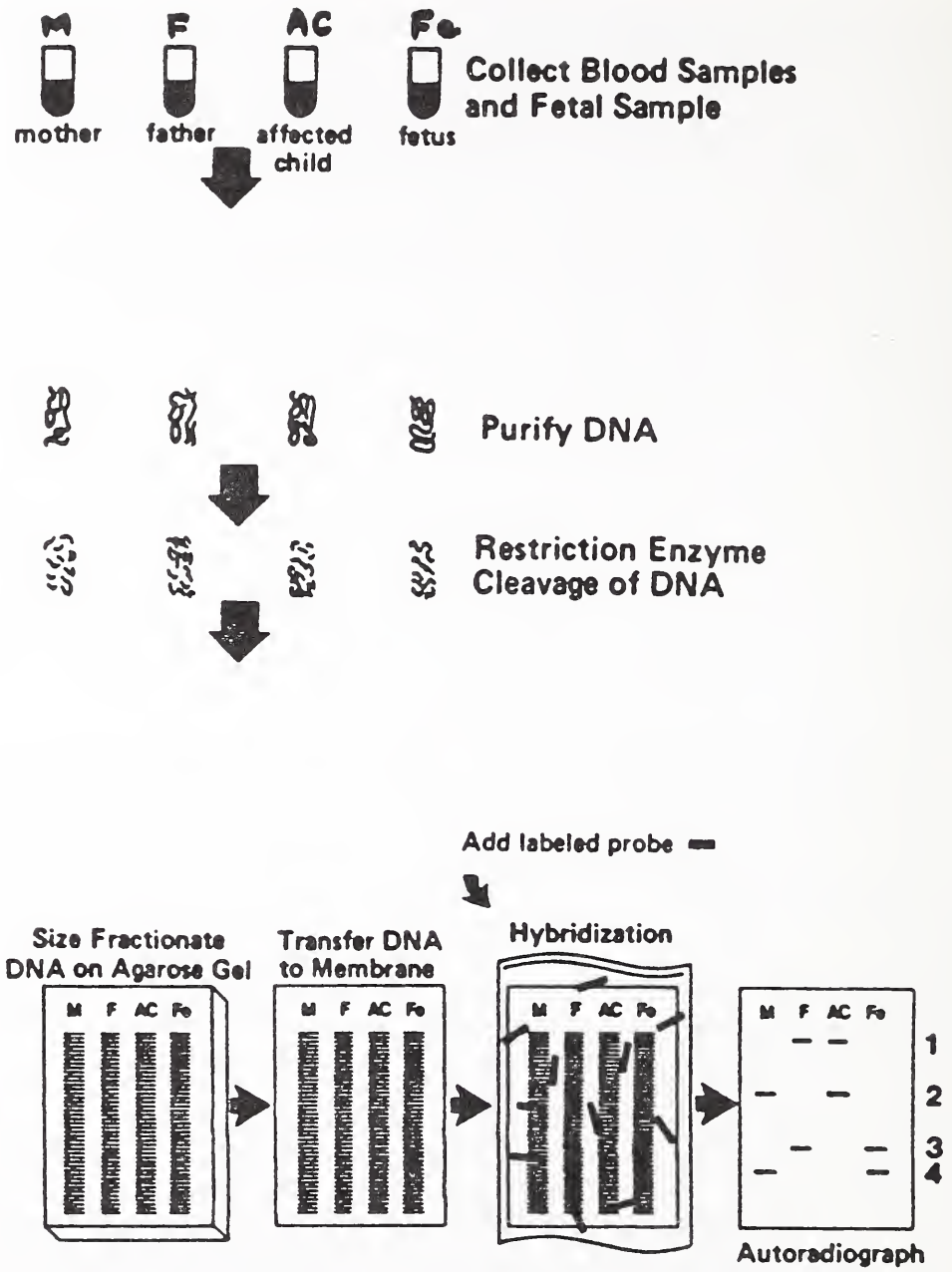

FIGURE 4. Restriction fragment length polymorphism (RFLP) identification steps.

of suspects in England and America. But the most future-directed application is the current work described by Donis-Keller where the goal is discovering genetic markers for various diseases and assigning these to specific chromosomes. A relatively complete map has been made of chromosomes 7, 16 and X. Partial maps exist for 12, 13 and 21. The first RFLP marker used for a probe was linked to the cystic fibrosis gene, at a distance of 15 centimorgans away, on chromosome 7. Since that report in 1985 (9), the list has increased and by 1987, RFLP probes were available for more than 40 inherited disorders. While this work is progressing, a lively debate has ensued over a proposal for federal support for mapping the human genome (10). This project would require an increase in funds and a commitment to a long-term program the ultimate goal being a complete sequence of the human genome. One of the liveliest proponents is Dr. Leroy Hood of the California

Institute of Technology. Recently Dr. Hood presented the argument to the medical research community (11). Hood has greatly improved machine methods for sequencing DNA and believes that the "proposition to map and sequence the human genome . . . is undoubtedly the most ambitious project ever conceived by biologists". Perhaps no development has more potential to bring about projects associated with this proposal than the increase in aged in America, Northern Europe, the Soviet Union and Japan. Currently, over 12 percent of all Americans are 65 or over, and by the year 2025 this number will have reached about 20 percent along with similar numbers in Northern Europe. At that time Sweden and West Germany will lead the world with 22 percent of their population aged 65 and over (12). The relationship to gene mapping and sequencing arises from the growing body of experimental work which shows genes in human DNA with the potential for causing transformation of cells to a cancerous state. Further, some biologists postulate that control of aging and the diseases of aging are all of genetic origin. Mapping and sequencing of the human genome, utilizing genetic engineering technology, offers a unique opportunity pertinent to these issues.

Literature Cited

1. Jackson, D. A., Symons, R. H., and Berg, P. (1972). Proc. Natl. Acad. Sci. USA *69*, p. 2904.
2. Cohen, S. N., Chang, A. C. Y. and Hsu, L. (1972). Proc. Natl. Acad. Sci. USA *69*, p. 2110.
3. Cohen, S. N., Chang, A. C. Y., Boyer, H. W., and Helling R. B. (1973). Proc. Natl. Acad. Sci. USA *70*, p. 3240.
4. Morrow, J. F., Cohen, S. N., Chang, A. C. Y., Boyer, H. W., Goodman, H. M. and Helling, R. B. (1974). Proc. Natl. Acad. Sci. USA *71*, p. 1943.
5. Palmiter, R. D., Norstedt, G., Gellinas, R. E., Hammer, R. E., And Brinster, R. L. (1983). Science, *222*, p. 809.
6. Jaenisch, R. (1988). Science, *240*, p. 1468.
7. Cummings, M. R. (1988). *Human Heredity*, 476 pages, Reference on p. 166-173. West Pub. Co., N. Y. City, New York.
8. Aktipis, S. (1986). "DNA: The Replicative Process and Repair". Chapter 17 in *Textbook of Biochemistry*, Editor, Devlin, T. M. Pages 625-695. Reference on p. 655-663. J. Wiley and Sons, N. Y. City, New York.
9. Donis-Keller, H., (1987). "Disease Diagnosis Using Restriction Fragment Length Polymorphisms" in *Genetic Engineering Technology: The Principles and Applications*, Editor, Tabor, J. M., Marcell Dekker Pub. Inc., N. Y. City, N. Y.
10. Roberts, L. (1987). Research News: "Agencies Vie over Human Genome Project". Science *237*, p. 486
11. Hood, L. (1988). "Biotechnology and Medicine of the Future" Journ. Am. Med Assoc. *259*, p. 1837.
12. "Aging America - Trends and Projections", 1987-88 Edition. Prepared by the U.S. Senate Special Committee on Aging with the American Association of Retired Persons, The Federal Council on Aging and The U.S. Administration on Aging. Printed by The U.S. Dept. of Health and Human Services, 1988.