Research instrumentation for the 21st century: progress toward complete genomic maps and sequence data bases, and indexes of protein gene products

NORMAN G. ANDERSON, N. LEIGH ANDERSON and J.-P. HOFMANN
Large Scale Biology Corporation, 7503 Standish Place, Rockville, MD 20855, U.S.A.

Abstract The instrumentation requirements for agricultural research in the 21st century hinge on whether or not complete analyses of animal and plant cells will be required, on the cost of performing the analyses, and on the cost and complexity of the analytical instrumentation needed. It is emphasized that nature's response to modern agriculture is to evolve pests and diseases better fitted to attack new genetically uniform plants, or to escape pesticides or herbicides. For all future history, new strains of crop plants, new herbicides and pesticides, and new animal strains will be required periodically. Existing native plant varieties now provide genes for resistance which are empirically discovered, and bred into high-yield strains by conventional methods. It is suggested that this process will not be endlessly successful, and that at some future time the required genes will not be found. Then, true genetic engineering with rational design and optimization will be required. This means the complete sequencing of plant and animal genomes, the identification and characterization of all gene products, computer modelling of cell function, and ultimately the introduction of sets of modified or totally synthetic genes. We review the status of some of the instruments required, and suggest future directions for their development.

Introduction

In 1986 a large fraction of the scientific community concluded that it was technically feasible to map and completely sequence the entire human genome (3.4 billion base pairs) (Anderson and Anderson, 1985; Palca, 1986; Lewin, 1986; Martin and Davies, 1986; Wada and Soeda, 1987) and a smaller fraction believed that the project should be started now. It has been previously proposed that nearly all human protein gene products could be separated and indexed (Anderson and Anderson, 1982a, b and 1984; Taylor, Anderson, Scandora et al., 1982), and methods have gradually developed which allow nearly all known metabolites to be chromatographically separated and identified. The medical importance of such efforts is obvious. The questions we ask here are these: Will similar analytical efforts be required in agriculture, what will motivate such projects, what instruments or systems will be required, and will the results justify the costs?

The discovery of new elements, of new subnuclear particles, or of new planets has turned out not to be endless. Biological systems, in contrast, have been found to be so complex that few have wished to consider, even briefly, the question of whether living cells could, even in theory, be so completely understood that predictable and successful modifications of

them could be made. Fortunately the organizers of this symposium have provided both a license and a forum for discussing these questions – at least once.

Will a complete understanding of living cells be necessary?

At present, new enzymes, new biochemical pathways, new structural proteins, and new chemicals which affect plants and animals are being discovered at an increasing rate. Genes for plant proteins are being isolated and sequenced, and transferred from one strain or species of plant to others. The astonishing rate of acceleration of research, however, obscures the fact that only a very tiny fraction of any plant genome has been sequenced, and the structure and function of an almost equally small fraction of the protein gene products of plant genes are known. It is unpopular to attempt to measure or indicate how much potentially useful knowledge we lack. However we cannot escape the fact that we have hardly begun to amass the bodies of knowledge which a continued human food supply will, we believe, ultimately require.

The first biotechnology revolution gave rise to agriculture and animal domestication, made possible civilization, and ultimately facilitated a rapid expansion of the human population. Most of the advances involved understanding and manipulation at the whole organism level, including the modification of human behavior to match agricultural requirements.

The second biotechnology revolution has started within the memory of most practising scientists who, incidentally, also constitute the majority of all scientists who have ever lived (De Sola Price, 1986). Nature itself will provide us with the motivation to keep this revolution going. Herbicides provide means for selectively killing undesirable plants in crop areas, and nature has and will continue to respond by providing new mutations for herbicide resistance in weed and non-crop plant populations. Pesticides are and will continue to stimulate selection of new resistant strains of insects, and antifungal agents will give us resistant fungi. As antibiotics become more widely used, antibiotic resistant bacteria have emerged and spread, holding the promise that in time all presently used antibiotics may become useless. As genetically homogeneous crop plants are cultivated, nature provides, often within as little as ten years, "improved" pests and diseases to attack them. In man and animals immune systems have provided marvelously versatile responses to viral and bacterial infections. Nature's latest response is the AIDS virus which attacks the immune system itself.

These reponses may be viewed as nature's attempts to prevent any one

species from reaching its full reproductive potential, and overgrowing others. Once we have started, through modern technology, to break free of nature's control mechanisms, which include starvation, disease, and predation, then we must attempt to counter nature's responses *in perpetuo*. For the rest of human life on earth, barring a catastrophic decline in the human population, we will require new crop varieties, improved agricultural animals, new herbicides and pesticides, and new antibiotics. We are barely into this new era, but we have already passed the point of no return.

While the green revolution now provides sufficient world food, it would appear that cycles of new pests and diseases, followed by development of new resistant strains will occur for ensuing centuries, possibly on shorter and shorter cycles. We must ask, can our response be as varied and resourceful as nature's attacks?

Evolution in our time

Nature's responses to modern agriculture, animal breeding and husbandry, and to medicine may be described as *evolution in our time*, based on the random generation and selection of new or previously unnoticed mutations. Our response thus far is almost equally random, involving the search for existing genes for resistance or adaptation, with the work carried out by individuals or groups which by accident (possibly involving social mutations?) happened to be interested in a particular problem.

Even with the most comprehensive and systematic collection of strains and species related to present crop plants obtained from all the Vavilov centers of the world, there is no guarantee that resistance genes exist and will be found to fill all future needs, or that useful new mutations will occur by chance when needed. Thus the time required to breed new resistant strains may be expected to increase, while the chances of success may decrease.

Note that nature's responses are not fine-tuned. More species are extinct than now survive on earth, and there is no guarantee of long-term survival of any species, including man. While these considerations may appear pessimistic, prudence suggests that now is the time to begin to consider how the role of chance can be reduced in crop breeding, in herbicide and pesticide development, and in the development of antibiotics and antiviral agents. The only answer appears to be massive acquisition of new knowledge in the biological sciences, and serious attempts to understand and to model living systems in such detail that they may be engineered, that is, modified and improved in a rational and predictable manner.

Barriers to progress in biotechnology

In the past plant and animal breeders have systematically altered the genetic makeup of plants and animals by every means that they possessed, and have proliferated new varieties freely. Recombinant DNA techniques, however, have been viewed as allowing new and unique combinations of genes to be made which would never and could never exist in the wild, and hence might constitute new, unique, and dangerous threats to existing balances of nature.

If the genetic makeup of plants is fixed or shielded in such a way that gene transfer between non-interbreeding species *never* occurs in nature, and if many (or any) new combinations of genes can be made artificially which are better adapted to survive than are native ones, then the popularized dangers of genetic engineering are potentially real.

Gene transfer, however, appears to occur widely in nature. Many viruses are known which infect not only across species barriers, but between different classes and phyla. In addition many viruses readily transfer bits of DNA between hosts, providing a system for widely transfecting genes. While cancer virology has long emphasized the importance and universality of transfection, the concept appears to have escaped popular notice. If indeed miscellaneous bits of genetic material freely float around the plant and animal kingdoms, transported by a viral mailing system, one might expect that a significant amount of excess DNA would accumulate in host cells which is not expressed, as is indeed the case.

If there are few barriers in nature to gene transfer, and if it occurs widely and naturally, then it would appear that very few transferred genes confer advantage, and nearly all the advantageous (in nature) collections of genes already exist (Anderson, 1970).

The following simple analogy, for which I apologize in advance, makes the point. Some aircraft have almost as many parts as there are different structural genes in a higher plant or animal. If one learned how to take a few parts from one type of plane and put them into another, then one might (provided one was sufficiently ignorant of the details) propose that there was a danger that a few fighter parts could convert an airliner into a super fighter, better than either original craft. This, of course, is ludicrous because modern aircraft are designed from scratch and are carefully designed, highly-integrated systems of parts that fit and match. We believe that the same will be found to be true of living systems. In practice, cultivation and selection have always produced new varieties which were *less well* adapted to survival in the wild than were the native varieties. Otherwise we may well have overgrown the earth

with some plant or other by now. Effective biological warfare – the limiting case against genetic engineering – turns out to be very difficult to do. Given tons of natural vectors spread over the earth's surface, mega-trillions of plants, and eons of time, nature can and does do experiments on a scale and of such a range of complexity as to relegate to total insignificance work done in all genetic engineering laboratories in the world. We are not trespassing on a pristine nature, rather we are under assault by a nature whose technical versatility and repertoire of tricks dwarfs our imagination. We do not believe that all the plant breeders or genetic engineers together could or will produce a better weed. These are not merely theoretical or philosophical arguments. Much of the future of agricultural research rests on the answers.

Progress in genome sequencing

New DNA and protein sequences are being recorded at a logarithmically increasing rate (Table 1 and Fig. 1). The strategy and tactics for mapping and sequencing the human genome are now subjects of intense discussion in the biomedical community. Sequencing the human genome at a cost of approximately $1 per base pair would cost over three billion dollars. However, with robotic sequencing systems now planned, costs could drop by one or two orders of magnitude, making it feasible to contemplate completion of plant and animal genomes. Given genomic libraries widely available on optical discs, and expert systems running on

Table 1. Genetic sequence data bank database summary GENBANK (R) Release 48.0 16 February 1987

Major groups of entries	Reports	Entries	Bases
1. PRIMATE Sequences	1,891	1,337	1,602,436
2. RODENT Sequences	2,307	1,612	1,460,441
3. Other MAMMAlian Sequences	361	310	325,323
4. Other VERTebrate Sequences	612	513	440,085
5. INVERTebrate Sequences	793	685	542,525
6. PLANT Sequences	828	689	819,740
7. ORGANELLE Sequences	498	423	840,194
8. BACTerial Sequences	1,321	916	1,268,642
9. Structural RNA Sequences	756	659	72,838
10. VIRAL Sequences	1,848	1,160	1,684,316
11. BacterioPHAGE Sequences	350	169	283,756
12. SYNTHETIC Sequences	287	251	78,943
13. UNANNOTATED Sequences	2,192	2,189	1,542,126
Summary:	13,774	10,913	10,961,365

Plant sequence data is approximately 0.75% of the total. 6.8% of plant sequences are for maize, and 4.5% for wheat. Less than 1/500 000th of the maize genome has been sequenced, while approximately 0.05% of the human genome has been sequenced. (Data kindly provided by Dr. Howard Bilofsky of Bolt, Beranek and Newman, Inc., Cambridge, Mass.).

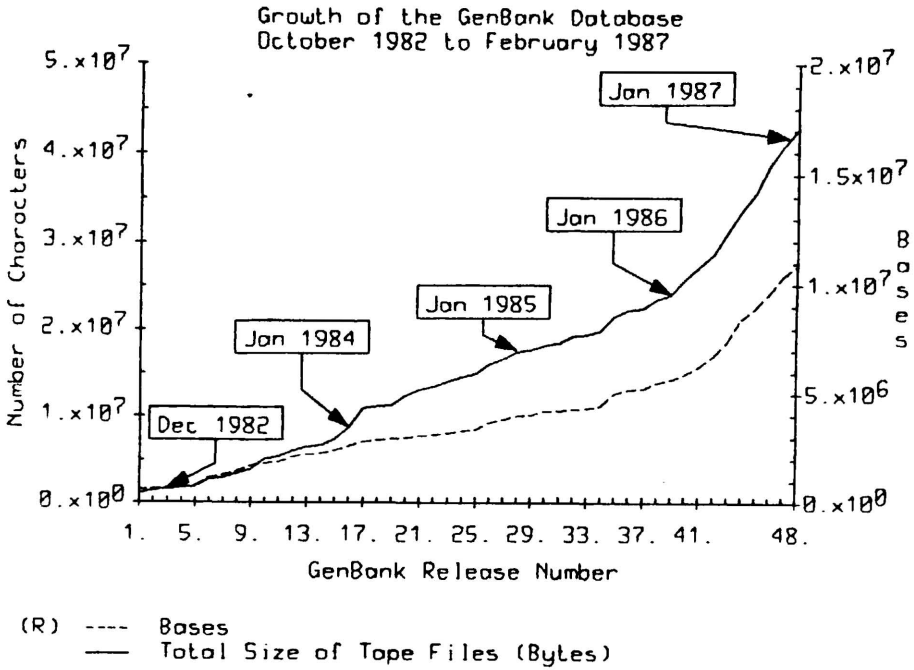

Figure 1. Rate of acquisition of DNA sequence data in GENBANK (Data courtesy Dr. Howard S. Bilofsky, Bolt, Beranek and Newman, Inc., Cambridge, Mass.).

currently available work stations, graduate students will be able to roam through and intercompare entire genomes to find interesting and useful differences, and to deduce functions of individual genes or effects produced by sets of genes. Courses in genomic architecture, on integration of gene systems, and on theoretical genomic biology will be commonplace. The instrumental requirements include the development of routine large-scale chromosome sorting, systems and methods for producing and separating large DNA fragments, robotic systems for routinely producing and cloning small fragments from larger ones, completely robotic and computerized systems for sequencing DNA and for integrating results back into complete chromosome structures, and computerized methods for performing model experiments and for keeping track of experimental results. All of these are in some stage of planning, development, or use now.

Information explosions

The scientific literature is now unmanageably large, and its sheer size severely inhibits interdisciplinary research. Radically new methods for evaluation, summarizing, transmitting, presenting and manipulating information will be required based on agreed-upon architectures and

matching hardware. The key appears to be sophisticated work stations which provide real time staged access to an ever enlarging data base.

A DNA-based index to biology

It is technically feasible to develop complete data bases of plant or animal DNA with masses of useful data electronically linked to each coding region. On a computer-generated sequence map, one would merely point to a gene of interest, and the computer would provide menus listing the types of information available which relate to that gene, i.e., all known properties of the gene product including its name, function, known genetic variants, and keys to literature citations. The difficulty with DNA-based indexing of biological knowledge, however, is that long sequences are not easily remembered. This suggests that genes be named for their products, which are nearly all proteins. While at present many genes are "named" for a phenotypic effect or a genetic disease, this is most likely a temporary situation which will last until the relevant gene product (protein, or in a few instances an RNA) has been isolated, identified, characterized, and named. It appears that the index to any genome will ultimately be based on a list of proteins.

Protein-based indexing of biochemical knowledge

While it may ultimately be possible to deduce the three-dimensional structure and function of a protein from its primary sequence, which in turn may be deduced from the nucleotide sequence of the gene, this cannot be done now. The unifying concept is of a protein data base which is the index to a centralized DNA data base such as GENBANK.

Hence some way of separating and cataloguing proteins is required which also allows each protein to be linked to its DNA sequence. The protein data base should contain all available pertinent information on each protein organized in such a manner that one may retrieve specific information starting with a name, sequence, function, E.C. number, disease, or activity. The problem has been to evolve a protein analysis system capable of resolving thousands of proteins, and of developing image analysis and data base systems to match the complexity of living cells.

High-resolution two-dimensional electrophoresis resolves thousands of proteins (O'Farrell, 1975; Anderson and Anderson, 1978a and b, 1979, 1982, a and b, 1984) and provides the results in the form of maps containing large numbers of spots (Young, 1984). These maps may be scanned, processed, and displayed (Anderson, Taylor, Scandora et al., 1981; Anderson, Taylor, and Anderson, 1983) with both spacial (Taylor,

Anderson, and Anderson, 1983) and quantitative precision (Anderson, Nance, Tollaksen et al., 1985), and individual spots used as entry points into large data bases (Taylor, Anderson, Scandora et al., 1982). Linked data may include the location of the corresponding gene, the name of the protein (or subunit), the protein's function, structure, a description of genetically variant forms of the protein, the intracellular location, the cell types in which found, association with disease, and literature references. The development of a human protein index based on these concepts has been described in detail (Anderson and Anderson, 1982a). With the development of methods for partially sequencing proteins recovered from 2-D gels (Aebersold, Teplow, Hood et al., 1985), a systematic technique for linking large numbers of proteins to their respective genes by reverse translation has become available.

Gene product (protein) data bases

It is well to describe in some detail how image-linked data bases work, and how they lend themselves to research and analysis. The data bases are built around image spot files which include all of the proteins resolved from a particular class of sample. These files can be used to construct a "master" map such as the one we have prepared for corn (Fig. 2). Individual samples are analyzed with reference to this master pattern. One may ask whether a new sample contains any protein not previously seen, whether some proteins usually present are absent, and whether quantitative differences are detected relative to one standard reference pattern or to a reference set of patterns. The key point is that the results (differences) are displayed on a video screen either as changes in grey levels, pseudocolor differences, or by flickering. One sees directly spots related to the answers to experimental questions. When a cursor is placed on a spot identified as interesting, the computer provides its identifying number, and any additional information required from a separate menu (Anderson and Anderson, 1982a),.

Thus the proteins which differ between different parental strains of hybrid corn may appear on the screen in different and identifying colors. Or wheat endosperm proteins which correlate with superior baking quality may be similarly identified. One may then ask specific questions relative to either individual proteins found to be interesting, or to groups of proteins. For example the set of all mitochondrial proteins may be highlighted (from information in the data base), and compared with a set of proteins whose expression may be affected by a herbicide or plant hormone. One may ask for the intersection of these two sets to be displayed (i.e., all proteins which are both mitochondrial and are affected

by the agent studied). Examples of such studies on lymphocytes have been previously published (Anderson and Anderson, 1982b). A fundamental objective is to arrange to manage complex data quantitatively and rapidly, to do statistical studies in real time, and to very rapidly relate the results to what is already known. Representative of agriculture-related studies which have been done include those on wheat endosperm proteins (Anderson, Tollaksen, Pascoe et al., 1985; Anderson and Anderson, 1987), wheat plant tissue (Zivy, de Vienne, and Hofmann, 1983), Douglas fir (Barhman, de Vienne, Thiellement et al., 1985), milk (Anderson, Powers, and Tollaksen, 1982) and muscle tissue (Giometti, Anderson, and Anderson, 1979).

We have concluded that high resolution computer-processed and CRT displayed color protein maps provide a unique solution to the problem of presenting and intercomparing complex data, as a point of entry into the rapidly growing data base of protein-related information, and as a link to DNA sequence information.

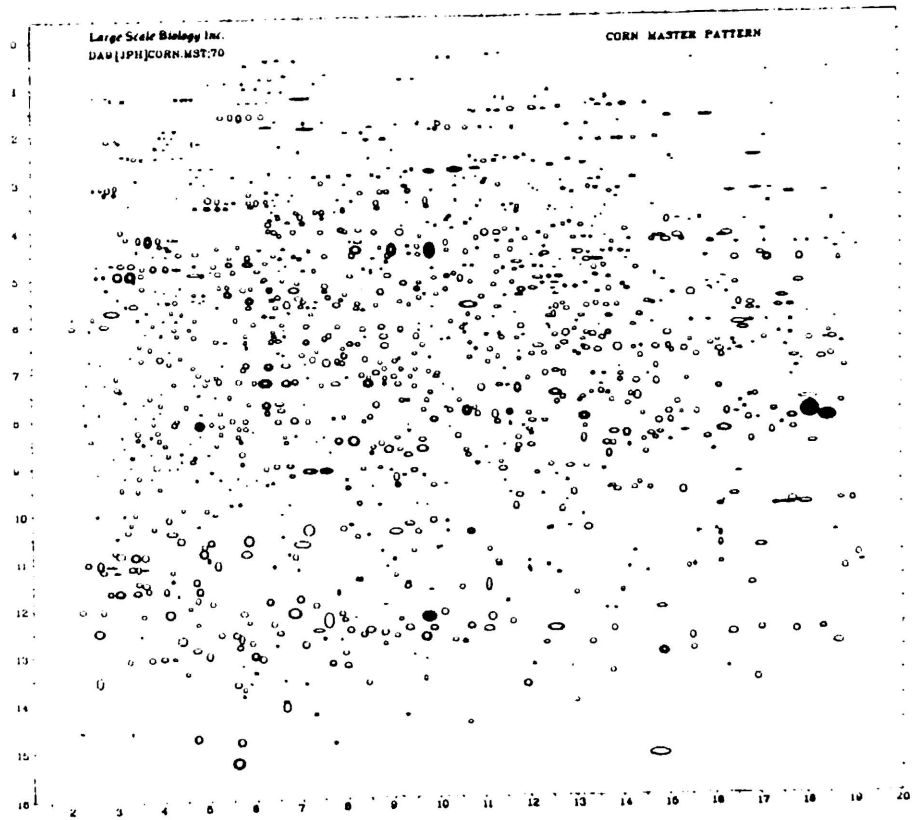

Figure 2. Master plot of corn seedling proteins. (Prepared by Dr. Jean-Paul Hofmann, LSB Corporation in collaboration with the Argonne National Laboratory).

The use of 2-D mapping in plant breeding is now under intensive investigation in several laboratories, and will become routinely feasible as the cost of the analysis drops due to automation and the use of robots.

Instruments for the 21st century

The basic instrumentation for resolving living cells includes systems for cell separation or cell sorting, for high-resolution subcellular fractionation, for separation of cellular metabolites by HPLC, for cloning, amplifying and sequencing DNA, for sequencing proteins, for peptide and oligonucleotide synthesis, and for resolving very complex protein mixtures. High resolution two-dimensional electrophoresis, as originally developed by O'Farrell (1975) and others, combined with sophisticated image analysis and data base systems such as the TYCHO system, allow the data provided by all of the instruments mentioned to be stored, organized, managed, and accessed. With the development of methods for partially sequencing proteins recovered from 2-D gels spots, efficient methods for linking proteins in 2-D maps with DNA data bases have become a reality. Since protein recovered from 2-D gels can also be used to produce antibodies, it is feasible to develop libraries of antisera against specific proteins to be used to isolate undenatured proteins, to identify associated subunits, and to cross-identify proteins between different laboratories (Anderson and Pearson, 1980).

All of these systems will doubtless be vastly improved before the end of this century.

The structure of research in the 21st century

While centralized instrumentation facilities are emerging in many universities (Hunkapiller, Kent, Caruthers et al., 1984) their reduplication and enlargement may not solve the massive data acquisition requirements we have discussed. Many scientists are now disturbed over the possibility that much of biological research will become "big science", with central organization and loss of individual initiative.

We propose that there will not be enough scientists in government, private, or university laboratories to acquire data on the scale required. Hence large centralized automated or robotic facilities for analytical data acquisition will be essential. They will perform an analytical function analogous to the synthetic and distribution function performed by large biochemistry supply houses. It is less than a generation since biochemists generally prepared most of their own reagents such as ATP.

Two trends follow from the views we have expressed. The first is

enrichment of biological research by disciplines and technologies now somewhat foreign including robotics and sophisticated electronics.

The second is an increase in the importance of individual investigator-initiated research. Biology is now becoming so complex that many central problems cannot now be attacked by individual investigators in even well equipped laboratories because of the high cost and complexity of analytical instrumentation. With masses of centrally developed data available on optical discs, specific reagents available from large banks, and access to sophisticated analytical systems by overnight mail, the individual investigator at his computer work station *is* the man to explore arcane aspects of genomic architecture, to find genes for resistance to specific diseases, to pinpoint factors controlling yield, and ultimately to begin the design of new and unique genes. The challenge is to provide talented individuals who have questions with organized and useful data in real time, and with sophisticated tools that carry out instructions. He need not do manually tasks robots and computers do better.

It has been said that agricultural research in the United States has been too successful, and has worked itself out of a job. We propose that in the 21st Century the situation will be reversed, and agricultural leadership and productivity will be closely tied to very high research technology.

References

1. Aebersold, R.H., D.B. Teplow, L.E. Hood and S.B.H. Kent. 1985. *Electroblotting onto activated glass: High efficiency preparation of proteins from analytical SDS-polyacrylamide gels for direct sequence analysis.* J. Biol. Chem. 261: 4229-38
2. Anderson, N.G. 1970. *The evolutionary significance of virus infection.* Nature 227: 1346-7.
3. Anderson, N.G. and N.L. Anderson. 1978a. *Analytical techniques for cell fractions. XXI. Two-dimensional analysis of serum and tissue proteins: Multiple isoelectric focusing.* Anal. Biochem. 85: 331-340.
4. Anderson, N.G. and N.L. Anderson. 1979. *Molecular anatomy.* Behring Inst. Mitt. 63: 169-210.
5. Anderson, N.G. and N.L. Anderson. 1982a. *The human protein index.* Clin. Chem. 28: 739-748.
6. Anderson, N.G. and N.L. Anderson. 1982b. *The human protein index project and the molecular pathology data base.* Medical Laboratory 11: 75-94.
7. Anderson, N.G. and N.L. Anderson. 1984. *Introduction: Some perspectives on two-dimensional protein mapping.* Clin. Chem. 30: 1898-1905.
8. Anderson, N.G. and N.L. Anderson. 1985. *A program and plan for biotechnology.* (Editorial) Amer. Biotechn. Laboratory, Sept. issue.
9. Anderson, N.G. and N.L. Anderson. 1987. *The application of high resolution two-dimensional electrophoresis to the analysis of wheat, milk, and other agricultural products. In* New Directions in Electrophoretic Methods, Jorgenson, J.W., and M. Phillips, eds. American Chemical Society Symposium 335: 132-142.
10. Anderson, N.G., M.T. Powers and S.L. Tollaksen. 1982. *Proteins of human milk. I. Identification of major components.* Clin. Chem. 28: 1045-1055.

11. Anderson, N.G., S.L. Tollaksen, F. Pascoe and N.L. Anderson. 1985. *Two-dimensional electrophoretic analysis of wheat seed proteins.* Crop Sciences 25: 667–674.
12. Anderson, N.L., and N.G. Anderson. 1978. *Analytical techniques for cell fractions. XXII. Two-dimensional analysis of serum and tissue proteins: Multiple gradient slab gel electrophoresis.* Anal. Biochem. 85: 341–354.
13. Anderson, N.L., S.L. Nance, S.L. Tollaksen, F.A. Giere and N.G. Anderson. 1985. *Quantitative reproducibility of measurements from Coomassie-Blue-stained two-dimensional electrophoresis gels. Analysis of mouse liver protein patterns and a comparison of BALB/c and C57 strains.* Electrophoresis 6: 592–599.
14. Anderson, N.L., and T.W. Pearson. 1983. *Use of high-resolution two-dimensional electrophoresis for analysis of monoclonal antibodies and their specific antigens.* Methods in Enzymology 92: 196–200.
15. Anderson, N.L., J. Taylor and N.G. Anderson. 1983. *Image analysis and interactive data base manipulation applied to two-dimensional maps of human proteins.* Proc. National Computer Graphics Association, June 26–30, 1983, pp 69–76.
16. Anderson, N.L., J. Taylor, A.E. Scandora Jr., B.P. Coulter and N.G. Anderson. 1981. *The TYCHO system for computerized analysis of two-dimensional gel protein mapping data.* Clin. Chem. 27: 1807–1820.
17. Barhman, N., D. de Vienne, H. Thiellement and J.-P. Hoffman. 1985. *Two-dimensional gel electrophoresis of proteins for genetic studies in Douglas Fir (Pseudotsuga menziesii).* Theoret. Appl. Genet. 23: 247–255.
18. De Solla Price, D.J. 1986. *Little Science, Big Science . . . and Beyond.* Columbia University Press. N.Y. pp 30.
19. Giometti, C.S., N.G. Anderson and N.L. Anderson. 1979. *Muscle protein analysis. I. Development of high resolution two-dimensional electrophoresis of skeletal muscle proteins for analysis of microbiopsy samples.* Clin. Chem. 25: 1877–1884.
20. Hunkapiller, M., S. Kent, M. Caruthers, W. Dreyer, J. Firca, C. Giffin, S. Horvath, T. Hunkapiller, P. Tempst and L. Hood. 1984. *A microchemical facility for the analysis and synthesis of genes and proteins.* Nature 310: 105–111.
21. Lewin, R. 1986. *Proposal to sequence the human genome stirs debate.* Science 232: 1598–1600.
22. Martin, W.J. and R.W. Davies. 1986. *Automated DNA sequencing: Progress and prospects.* Bio/Technology 4: 890–895.
23. O'Farrell, P.H. 1975. *High resolution two-dimensional electrophoresis.* J. Biol. Chem. 250: 4007–4021.
24. Palca, J. 1986. *Department of Energy on the map.* Nature 321: 371.
25. Pearson, T. and N.L. Anderson. 1980. *Analytical techniques for cell fractions. XXVIII. Dissection of complex antigenic mixtures using monoclonal antibodies and two-dimensional gel electrophoresis.* Anal. Biochem. 101: 377–386.
26. Taylor, J., N.L. Anderson, A.E. Scandora Jr., K.E. Willard and N.G. Anderson. 1982. *Design and implementation of a prototype human protein index.* Clin. Chem. 28: 861–866.
27. Taylor, J., N.L. Anderson and N.G. Anderson. 1983. *Numerical measures of 2-D gel resolution and positional reproducibility.* Electrophoresis 4; 338–346.
28. Wada, H. and E. Soeda. 1987. *Strategies for building an automatic and high speed DNA sequencing system.* Proc. 4th Congr. Fed. Asina and Oceanian Biochemistr, in press.
29. Young, D.A. 1984. *Advantages of separations on "giant" two-dimensional gels for detection of physiologically relevant changes in the expression of protein-gene products.* Clin Chem. 30: 2104–2108.
30. Zivy, M., D. de Vienne and J.-P. Hofmann. 1983. *Studies on nuclear and cytoplasmic genome expression in wheat by two-dimensional gel electrophoresis. I. First results in 18 alloplasmic lines.* Theoret. Appl. Genet. 66: 1–7.