HGCA

RESEARCH REVIEW No. 9

MODERN METHODS FOR CEREAL VARIETY IDENTIFICATION

NOVEMBER 1988

Price £12.50

CONTENTS

HGCA RESEARCH REVIEW No. 9

Modern Methods for Cereal Variety Identification

Robert J Cooke,
Chemistry and Quality Assessment Department
National Institute of Agricultural Botany
Huntingdon Road
CAMBRIDGE CB3 OLE

ABSTRACT

This review is concerned with what might be called modern methods for cereal variety identification, that is laboratory-based methods which do not involve field growing-out procedures and morphological assessments. Such methods are becoming increasingly important, not only in commercial situations (i.e. milling and malting) but also in seed certification and testing and in the statutory assessment of new varieties for distinctness, uniformity and stability.

The principal modern method, which is currently very widely used in a range of situations, is electrophoresis of proteins and enzymes. The term electrophoresis in fact covers a range of analytical techniques, all of which separate molecules such as proteins by their differential rates of migration through a porous gel support medium. Many kinds of electrophoretic analysis have been utilised for the separation of cereal seed proteins. These are especially useful for variety identification, since the precise molecular composition of the seed protein fraction is a reflection of the underlying genetic constitution. As proteins are the primary products of the genetic material, their composition is unaffected by environmental conditions. Thus the use of electrophoresis to reveal this composition provides a powerful and accurate means of variety identification, in the form of a gel protein banding pattern or 'finger-print'. The degree of resolution achieved between varieties is impressive, with over 85% of a collection of 155 UK wheat varieties, for instance, being uniquely different on the basis of their finger-print. The methods are also relatively rapid (results can be obtained in 1-2 days) and cheap.

Particularly for commercial applications, there is a need for standardisation of methodology and equipment and for research into the automated interpretation of gel banding patterns. However, electrophoresis is the method of choice for cereal variety identification at present.

An alternative method of determining protein composition is to utilise high-performance liquid chromatography (HPLC). Again seed protein profiles can be used successfully for cereal variety identification, with the degree of discrimination between varieties being broadly equivalent to that achieved by electrophoresis. The major advantages of HPLC are that it allows a quantitative determination of individual proteins, lending an additional layer of precision to discrimination, and that the data produced can be readily stored and manipulated by computer. The equipment necessary is rather expensive, however, and although individual analyses are rapid (about 1 hour), the throughput of samples does not equal that obtained by electrophoresis. The possibility of analysing mixtures of varieties by precise quantification of proteins present is worthy of further research, as is the development of procedures allowing shorter analysis times.

Apart from using such analytical methods to analyse seed protein composition, it may be possible to utilise immunological methods (antigen-antibody reactions) to detect specific proteins very rapidly and cheaply. Monoclonal antibodies to various cereal seed proteins exist and might be used for variety identification or classification. However, it does not seem likely at present that the use of antibodies can provide other than a rapid screening method, giving a broad indication of variety. More probable is the detection of particular quality types in seed lots by an appropriate selection of antibody tests. The rapidity and low cost of such a system would seem to warrant its further investigation to some degree.

Instead of analysing protein compositions as a means of varietal identification, it is possible to investigate the variability of the genetic material (DNA) itself and hence to discriminate between varieties at the fundamental level. Using restriction enzymes to cut DNA molecules, it is possible to examine by electrophoresis the

variability in the size of the fragments produced. This variability can then be utilised to distinguish between individuals, which could be, for instance, cereal varieties. Such technology represents an immensely powerful means of recognising and differentiating varieties. However, at the moment it is expensive, time-consuming and requires highly skilled staff and specialist facilities. There seems little doubt that this situation will change in the medium to long term. Many plant breeders are already actively using this restriction fragment length polymorphism (RFLP) approach to characterise their breeding lines and to identify specific genes. The simplification of the methodology would make it attractive as a tool for routine variety identification and this is clearly an area in which future research effort should be targeted.

The final modern method of variety identification differs from the preceding ones in that it is not biochemically based but involves a more classical taxonomic approach, albeit achieved through the use of sophisticated technology. Machine vision (which refers to the acquisition of data via a video camera or similar system and the subsequent computerised analysis of these data following suitable processing) provides a powerful means for the classification of plant material by the measurement of shape and size. The potential of this technology is being investigated with regard to cereal variety identification and the preliminary results appear very encouraging. Wheat varieties can be distinguished on the basis of differences in various aspects of their shape as assessed via machine vision. Although the equipment needed is expensive, the analysis is very rapid (10 minutes) and might be particularly suitable for quality control at factory intake points. More work would be advantageous in this area, not only to improve the equipment used but also to consider other species.

The various different modern approaches which can be taken to cereal variety identification all have their own advantages and disadvantages. In the short to medium term, electrophoresis will continue to be the most widely applied of the methods, perhaps augmented by HPLC in certain situations. Machine vision offers possibilities for more rapid evaluation of seed lots, but the questions of the precise

identification of samples and the characterisation of varietal mixtures need to be addressed. Immunological methods are best suited to the separation of seed by quality rather than strictly by variety, but would be rapid and cheap to operate. Perhaps the ultimate identification system is to utilise RFLP analysis. Whilst this is a long-term prospect at the moment, it is reasonable to predict that simpler and cheaper methodology will emerge.

This review, completed in November 1988 and with 112 pages in the full article, was funded by the HOME-GROWN CEREALS AUTHORITY, Hamlyn House, Highgate Hill, London N19 5PR, from whom copies may be obtained at a price of £12.50 each (postage and packing included).

GLOSSARY OF TERMS AND ABBREVIATIONS

alleles -

alternative genes carried at a genetic locus.

amino acids -

the sub-unit components of protein molecules.

ampholytes -

a mixture of basic and acidic compounds which is used in isoelectric focusing techniques to create a pH gradient within the electrophoresis gel.

buffer -

a solution which will minimise changes in acidity or alkalinity.

chromatography -

a method of separating and analysing mixtures of chemical substances by differential partitioning between two phases.

chromosomes -

the structures present in all living cells containing the genetic code. Most higher plants e.g. barley are diploid (have two sets of chromosomes, one paternal and one maternal); durum wheat is tetraploid (two sets of chromosomes from each parent) and bread wheat is hexaploid (three sets from each parent).

DNA (deoxyribonucleic acid) -

the genetic material of all organisms, consisting of two strands, wound in a double helix. Chromosomes are composed of DNA.

EBC -

European Brewery Convention.

EEC -

European Economic Community.

ELISA -

enzyme - linked immunosorbent assay; a type of sensitive immunological assay system.

electrophoresis -

a method of separating and analysing charged molecules (proteins, nucleic acids) by their differential rates of movement under the influence of an electric field.

electrophoregram -

the pattern of protein or nucleic acid bands resulting from the electrophoresis of a sample; a 'finger-print'.

enzyme -

one of a group of proteins produced by living cells which acts as a catalyst in specific biochemical reactions.

GAFTA -

Grain and Feed Trade Association

gel -

the inert porous matrix used for the electrophoretic separation of proteins or DNA. The concentration of the compound comprising the gel can be altered within large limits. With polyacrylamide gels, the total concentration is referred to as T and the concentration of cross-linking material as C.

gene -

the basic unit of inheritance, comprising a DNA sequence which codes for the production of a specific protein molecule.

genome -

the complete set of genes of an organism.

HPLC -

high performance liquid chromatography; a type of chromatography using specialised column packings which require high pressures to force the eluting solvents through at a reasonable rate.

hybridoma -

a hybrid cell-line produced by the fusion of a normal lymphocyte (spleen cell) with a myeloma cell; used in the production of monoclonal antibodies.

IEF -

isoelectric focusing; a type of electrophoresis in which the gel contains ampholytes, creating a gradient of pH.

ICC -

International Association for Cereal Science and Technology.

ISO -

International Standards Organisation.

ISTA -

International Seed Testing Association.

image analysis -

the extraction of numerical data from an image or series of images acquired via machine vision.

isozyme -

an enzyme which exists in two or more forms, differing in their electrophoretic mobility.

locus -

any site on a chromosome which has been defined genetically; a locus may be a gene, part of a gene, or a set of genes.

machine vision -

the acquisition of data regarding shape and size etc. via a video camera or similar means, and the subsequent manipulation and analysis of these data.

monoclonal antibody -

an antibody preparation which contains only a single type of antibody molecule.

myeloma -

a tumour of the immune system used in the production of monoclonal antibodies.

NIAB -

National Institute of Agricultural Botany.

nucleic acid -

a DNA (or RNA) molecule.

nucleotides -

the sub-unit components of nucleic acids.

PAGE -

polyacrylamide gel electrophoresis; a type of electrophoresis employing gels made from the synthetic material, acrylamide.

pH -

a term used to describe the acidity or alkalinity of a system.

pI -

isoelectric point; the pH at which protein molecules carry no net electrical charge.

phase -

used in chromatography to describe the two elements of the separation process, the stationary phase (column packing) and the mobile phase (eluting solvent), between which partitioning occurs.

phenotype -

the observable characters of an organism, as opposed to the genotype, its genetic constitution.

polymorphism -

existing in many different forms.

probe -

a specific DNA (or RNA) sequence which has been labelled (usually radioactively); they are used to detect specific complementary DNA sequences in RFLP analysis.

prolamins -

the alcohol-soluble storage proteins of seeds, known as gliadins in wheat and hordeins in barley.

protein -

a high-moleculr weight compound composed of a range of amino acids; they are the products of genes. Seeds generally contain four kinds of proteins of differing solubility properties - storage proteins (prolamins and globulins), enzyme proteins (albumins) and structural proteins (glutenins).

RFLP -

restriction fragment length polymorphism; a method of detecting variability in DNA sequences between organisms.

RP-HPLC -

reversed phase high performance liquid chromatography.

restriction enzyme -

a compound which will bind to and cut DNA molecules at a specific sequence of nucleotide bases.

retention time -

in chromatography, the time taken to elute a particular compound using a specified system.

SDS -

sodium dodecyl sulphate.

UPOV -

International Union for the Protection of New Varieties of Plants.

Modern Methods for Cereal Variety Identification

INTRODUCTION

That it is essential to have a rapid and reliable means of distinguishing between varieties (cultivars) of cereals is now thoroughly established within the milling, baking, malting and brewing industries. However, this is a comparatively recent phenomenon. Only some 25 years ago, for example, the bulk of the home-grown wheat in the UK was used for cake and biscuit flour or as animal feed, with grain being imported for use in bread-making. This meant that there was very little need for millers to be able to distinguish between varieties of wheat entering the factory. Today the situation is completely reversed and the greater part of the grist for a bread-making flour will be of UK origin. Although other factors are involved, the primary reasons for this are two-fold. Following the accession of the UK to the European Economic Community in 1973, the Common Agricultural Policy began to take serious effect, with financial inducements being placed on millers to buy and use more home-grown wheat. This in turn lead to millers utilising a far wider range of varieties than previously. The proliferation of available varieties was due to the passing, in 1964, of the Plant Varieties Rights Act, legislation which enabled plant breeders to obtain what in effect are royalties on the sale of seed of their varieties. This Act transformed plant breeding into a large national (and frequently multi-national) business venture, with the ultimate production of a wide range of high-yielding wheat (and other crop) varieties of varying quality. The millers, and to a lesser extent maltsters, now had the problem of adjusting to these different varieties and of trying to monitor the quality of the grain being utilised. Varietal identity, which is one of the most important quality characteristics, could not at this time (early 1970s) be readily checked. This was especially true for wheat, where the grain possesses few distinctive morphological features which are characteristic of varieties. A further problem is that the economic life-time of even a successful variety can be as short as 3-4 harvest years. Thus the industries were faced with being able to distinguish between and identify a constantly changing spectrum of cereal varieties. The ability to discriminate between varieties is also

important in other areas of activity, not the least of which is the operation of the statutory variety testing and certification procedures. Before new varieties can be marketed, they have to undergo testing to determine their distinctness from already existing varieties, their uniformity and their genetic stability (the so-called DUS procedures). Varieties which successfully complete DUS testing and are shown to have suitable agronomic value are added to the National List and become eligible for a grant of Plant Breeders' Rights. Detailed observation and description of various features of the growing and flowering plant form the basis for distinctness testing and in practice the procedure is very effective when applied to cereals. However, it can be a time-consuming process and requires large areas of land and highly-skilled personnel. Also many of the morphological descriptors used are multi-genic quantitative or continuous characters, the expression of which can be altered by environmental factors. There are thus compelling reasons to find more rapid and cost-effective procedures which could augment this morphologically-based testing.

Seed certification is another area in which the determination of varietal identity and purity is crucial. Certification can be envisaged as a quality control and consumer protection operation, in which the seed being multiplied prior to commercial release is monitored to ensure that it complies with certain statutory standards. Again it is clearly imperative to be able to discriminate between varieties to enable this process to be effective. Field inspections form an important part of certification but there is also a need for rapid laboratory-based methods which can be used for identification purposes.

Both farmers and seed merchants are interested in variety determination. Merchants are required by law to supply seed of the stated variety. Furthermore, the plant breeders are also in need of methods which can be used to discriminate between their breeding lines and perhaps also to help in the selection of those lines containing appropriate desirable characteristics without recourse to field trials. It is thus evident that all sectors of the cereal seed industry, from plant breeders through to the variety testing authorities, the certification agencies, seed merchants, farmers and, ultimately, the

end-users of the grain, would benefit from the existence of rapid, accurate, and preferably reasonably cheap methods for the determination of varietal identity and purity.

This article is concerned with what might be called the modern approaches to this problem, which are largely biochemically or machine-based. Currently, the prime method of this type is electrophoresis, but other techniques are being used, such as chromatography of different kinds. An alternative is to use an immunological approach, with the aim of producing antibody test kits. The possibility of examining the genetic material (DNA) of varieties is being actively pursued. Finally, the potential for machine vision (or image analysis) as a means of distinguishing between varieties needs to be considered. Each of these methods has advantages and disadvantages and these are also discussed, along with areas where further study would seem to be desirable. The use of these methods, which are being increasingly referred to as 'New Technology' in this area, offers exciting possibilities for all sectors of the cereal industry. This article reviews these possibilities with the aims of focusing attention on them and of highlighting future activities.

Chapter 1:

CEREAL VARIETY IDENTIFICATION BY ELECTROPHORESIS

Of the techniques which are considered in this article, the various different kinds of electrophoresis have been by far the most widely-used and successful so far for cereal variety identification. This chapter will firstly outline some of the basic theory of electrophoresis methods and why they are so useful for identification purposes. Then, the various methods which have been applied to cereals will be summarised, with attention being focused on the efforts to standardise the methodology. The degree of discrimination achieved between varieties will be compared and finally some of the practical applications of electrophoresis will be discussed.

What is electrophoresis?

Electrophoresis is a technique used to separate charged particles under the influence of an electric field. It is important to realise that it is not a new analytical technique, having been used by one of the pioneers of separation science, Tiselius, as long ago as 1925. These early experiments were carried out in free solution, which makes the separation difficult to observe and it was not until the introduction of different types of support medium in about 1960 that the techniques began to be exploited widely by scientists in various disciplines. The support media are used to hold the charged particles whilst they are undergoing separation and thus remove problems caused by diffusion. There are many types of support available, including paper and cellulose acetate, but the most commonly used media are gels of different kinds. Agar, agarose and hydrolysed starch have been extensively employed, but the most popular gels are those composed of the synthetic polymer polyacrylamide. This gives rise to the frequently occurring abbreviation, PAGE (polyacrylamide gel electrophoresis).

Electrophoresis is most useful for separating mixtures of large molecules such as proteins or nucleic acids, which in solution behave as charged particles. In this chapter, only electrophoresis of proteins is considered. To understand the rationale behind the application of electrophoresis for variety identification, it is necessary to have a basic understanding of protein structure.

Proteins are composed of long chains of amino acids, which have the generic formula R-CH (NH_3^+) COO^-. There are about 20 amino acids found in proteins and they have different types of chemical group comprising the R-part of the molecule. Some of these R-groups are capable of ionisation i.e. they can become charged. The precise order or sequence of different amino acids in a given protein and the number of amino acids which comprise the protein are determined genetically. In other words, the composition of the genetic material of an individual determines the composition of its proteins. There are many thousands of different kinds of protein, all with unique amino acid sequences. Because of this, and the varying lengths of the amino acid 'chains', proteins differ both in the electrical charge which they can carry and in their size (molecular weight). Both of these parameters are utilised in electrophoretic separations.

Consider a solution of a mixture of protein molecules, of varying sizes and charges (see Figure 1.1). If this is applied to the top of a gel and an electrical field is placed across the gel, the proteins begin to move. The rate at which they move depends primarily on two factors. Firstly, and perhaps predictably, molecules with a high charge migrate more quickly than those of lower charge. Secondly, the rate of movement varies according to the size of the molecule. This is probably not so predictable and arises from the structure of the gel used as a support medium. A gel can be envisaged as a series of holes, or pores, of a defined size, which can vary and which is determined by the concentration of material (starch, acrylamide) used in the gel. The pores exhibit a 'molecular sieving' effect, such that small molecules can pass easily through them whilst larger molecules, which may be approaching the size of the pores, are restricted and hence slowed down. Thus the mixture of proteins in Figure 1.1 can be separated into discrete bands, depending on the size and charge of the

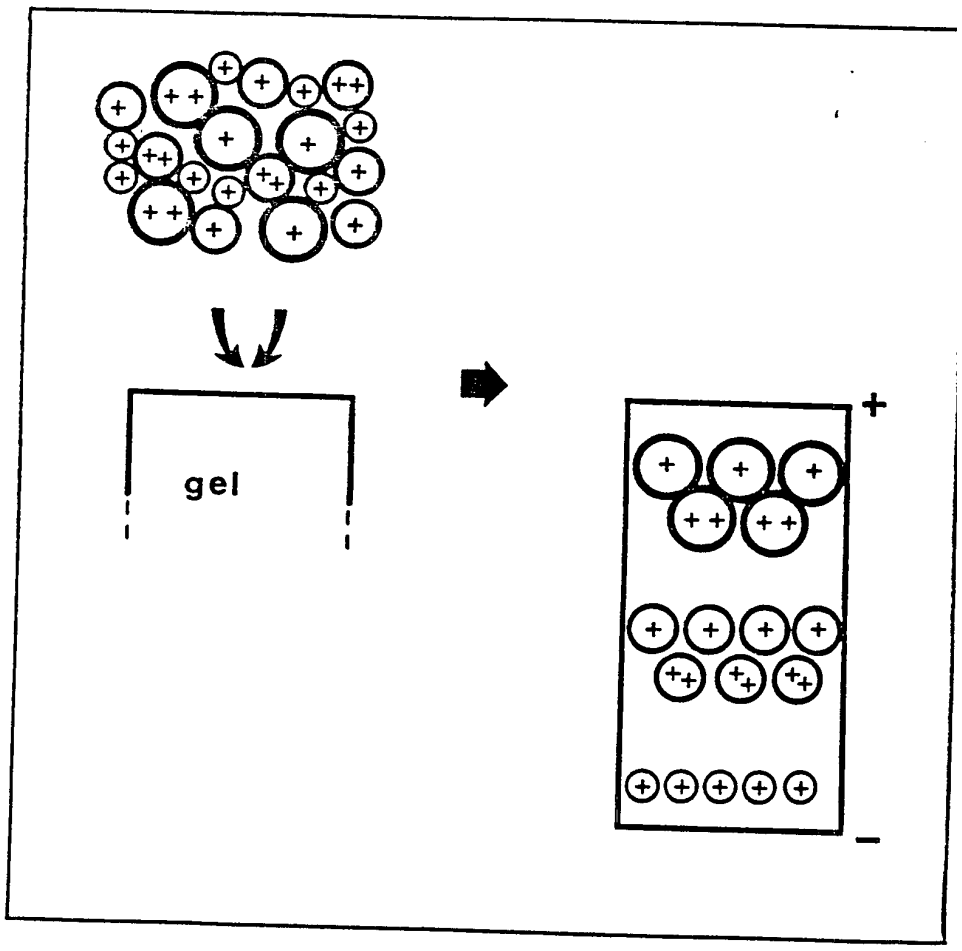
molecules. Following such an electrophoretic separation, it is necessary to visualise the proteins in some way, as they cannot be seen, and so various staining procedures are used.

This then is the basis, albeit much simplifed, of all electrophoretic separations. However, there are a number of different kinds of electrophoretic methodology and it is useful to have some understanding of the differences.

Types of electrophoresis

There are many versions of electrophoretic techniques and several text-books of methodology and theory are in print. All of the techniques share certain common features. For instance, the gel, of defined dimensions but variable pore size, is held between glass plates or in a tube and generally incorporates a buffer of some kind (known as the gel buffer). The ionic strength and pH of this buffer can be varied to suit particular types of proteins. An electrophoresis or tank buffer is also usually used to carry the electrical charge across the gel. The separation can be carried out in either a horizontal or vertical apparatus. Usually some means of controlling the temperature of the gel and buffer during separation is used. There are so many different types of equipment available, both commercially and home-made, that it is not really worthwhile to attempt to describe or recommend any particular one, except in terms of standardisation of methodology (see below). However, there are three principal and distinct kinds of electrophoresis which ought to be recognised.

Figure 1.1 A generalised illustration of the basis for electrophoretic separation of proteins. A mixture of proteins of differing sizes (molecular weights) and charges is applied to the top of a gel. When an electric field is applied across the gel, the proteins migrate. The rate of movement is determined by both the size and charge of the proteins and leads to the separation of protein bands from the original mixture.



1) Native PAGE (or native starch gel electrophoresis)

The method illustrated in Figure 1.1 is an example of a so-called 'native' electrophoresis method in which separation is based on both charge and size. There are versions of this which employ the same buffer in the gel as in the electrophoresis tank, the continuous methods. On the other hand, there are versions in which not only are the gel and tank buffers different but also a two-phase gel system is used, comprising the main or resolving gel on top of which is placed a large-pore stacking gel. These are known as discontinuous systems. Many different buffers and gel concentrations have been devised for both continuous and discontinuous systems. For the latter, the 'Ornstein-Davis' buffers are especially popular. Sample preparation for native PAGE is generally straightforward and involves crushing the tissue and extracting the proteins of interest with a suitable solvent.

2) PAGE in the presence of sodium dodecyl sulphate (SDS-PAGE)

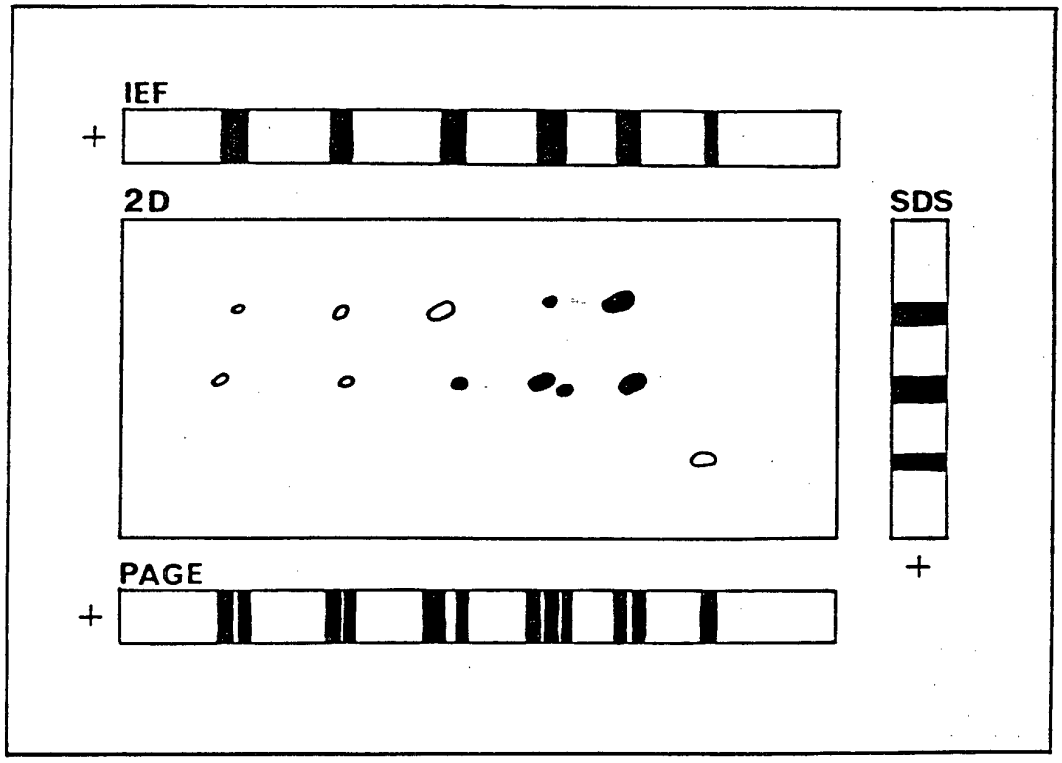
SDS-PAGE is a very commonly used type of electrophoresis, in which all of the gel buffers, tank buffers, sample buffers etc.are made up including SDS, a powerful anionic (negatively charged) detergent molecule. It has been shown that all proteins, when in a reduced form, will bind SDS in essentially the same proportion (1.4g of SDS per g of protein, which is about 1 SDS molecule per 2 amino acid residues). Because SDS has a strong negative charge, it effectively overcomes the inherent charge of proteins and as a result, electrophoretic separations of SDS-treated proteins are based only on the size of the proteins. A widely used SDS-PAGE method is the discontinuous system originally devised by Laemmli, although other formulations are also employed.

3) Isoelectric focusing (IEF)

The third kind of electrophoresis differs from native- and SDS-PAGE in that generally gel and tank buffers are not employed. Instead, a pH gradient is created within the gel, usually by the addition to the gel mixture of chemicals known as carrier ampholytes. Now although as was pointed out above, protein molecules carry an electrical charge, the size and indeed the sign (+ or -) of the charge depends on the pH of the solution in which the protein is dissolved. This is because of the structure of amino acids, which contain chemical groups which can be positively ($-NH_3^+$) or negatively ($-COO^-$) charged. For a given protein, there is a pH value at which it carries no net charge and this is known as the isoelectric point, or pI. At pH values below the pI, a protein will be positively charged whilst at pH values above the pI it will be negatively charged. In IEF methods, a mixture of proteins is loaded on to the gel containing the pH gradient and when the electric field is applied, the proteins migrate in the gel until they reach the point where the external pH equals their pI value. There, because they no longer carry a charge, the proteins cannot move and in effect become 'focused'. Because proteins differing only slightly in pI can be readily resolved, IEF is a powerful way of separating protein mixtures.

All of the above three types of electrophoresis have been employed more or less successfully for cereal variety identification. A fourth kind of method also needs to be considered. This is the two-dimensional (2D) approach, in which a sample is first fractionated by a particular technique and the resultant gel is then turned through 90°and forms the sample for another separation, carried out at right angles to the first. This is illustrated in Figure 1.2. Such 2D methods are amongst the most powerfully resolving electrophoretic systems, especially when used in conjunction with very sensitive staining procedures. Various combinations of methods have been used for 2D electrophoresis, but perhaps the most common is IEF + SDS-PAGE, sometimes known as the O'Farrell technique, after its originator.

Figure 1.2 An illustration of two-dimensional (2D) electrophoresis. The protein sample is first separated by electrophoresis in the normal way. The gel is then turned through 90° and forms the sample for a second separation, carried out at right angles to the first.



Of necessity, this can only be a brief description of the various possible kinds of electrophoresis. Why is it, though, that electrophoresis can be used for cereal variety identification?

The rationale for the use of electrophoresis

The success of electrophoresis in identifying and distinguishing between varieties relies on the fact that proteins are the primary products of genes. Proteins can thus be regarded as 'markers' for the structural genes (i.e. segments of DNA) that encode them. Hence, a comparison of the protein compositions of individuals or populations can be equated, albeit one stage removed, to a comparison of the underlying variation in gene expression. As genes are connected into genetic systems, protein markers can be used to label these systems, which might be a set of genes, part or all of a chromosome, or the genome as a whole. By considering a sufficient number of protein markers, a large part of the genetic material can be covered. Crop varieties can be envisaged as collections of germplasm which differ in their genetic expression. A comparison of the composition of particular proteins or enzymes occurring in these collections can thus be a means of 'typing' or characterising the material. Clearly, electrophoretic methods can provide an excellent way of carrying out such a comparison. In variety identification work, investigations have concentrated on the electrophoretic examination of seed protein composition, although the composition of particular enzymes from seeds or vegetative tissues has also been used. Seed proteins can be considered to be of four types (see below), but the most useful for identification purposes are generally the storage proteins. In almost all species, storage proteins have been found to be extremely polymorphic, that is they exist in many molecular forms, which differ with regard to charge, size or both parameters. Furthermore, storage proteins are genetically encoded at several points, or loci, throughout the genome, are present in comparatively large amounts and are readily extracted. Hence the electrophoretic examination of storage protein composition provides a powerful and convenient way of characterising

plant genotypes of many kinds and has proved to be extremely useful for cereal variety analysis (see reviews 7,10,34). As an alternative to staining gels for total protein activity, the use of specific stains to reveal the multiple molecular forms of particular enzymes (isozymes) has also been advocated (10,34) although not as widely in cereals as in other crops.

Seed proteins

The nomenclature of seed proteins is based largely on the pioneering observations of the American chemist T B Osborne, who demonstrated that seeds contain proteins differing in their solubility properties. Although this 'Osborne fractionation' can be criticised as being insufficiently rigorous in modern molecular biological terms, it is still widely referred to and provides a convenient system for everyday use in the laboratory. There are considered to be four types of seed proteins: (a) albumins, which are water-soluble and comprise mostly enzymic proteins; (b) globulins, which are soluble in dilute salt solutions and occur in membrane-bound protein bodies i.e. are storage proteins in the strict sense; (c) prolamins, which are soluble in aqueous alcohol solutions and are also true storage proteins; (d) glutelins, which are soluble in acid or alkaline solutions and are probably mainly structural or storage proteins, although some may have metabolic functions. The amino acid composition of each of these categories of proteins varies in a characteristic way. Moreover, the proportion of each type of protein which is found varies from species to species. For instance, the seeds of cereals such as wheat, barley, maize or rye contain high levels of prolamin-type proteins, whereas other cereals (oats, rice) have high levels of globulins. Leguminous seeds (eg beans, peas) also characteristically contain high levels of globulin-type proteins. Much of what follows is concerned with the electrophoretic separation of cereal prolamins and glutelins, although fractionations of albumins and/or globulins have also been used for variety identification purposes.

Historical aspects

The development of electrophoretic methods for cereal variety identification concurred with the need for such an approach within the milling trade, although this was probably fortuitous. In the Introduction, it was mentioned that prior to about 1970, there was little or no reason for millers or maltsters to be too concerned about variety identification. However, there had been a substantial amount of research previous to this, primarily aimed at investigating the nature of cereal storage proteins. Thus, for instance, Jones et al. (22) in 1959 reported one of the first electrophoretic fractionations of wheat gluten, using a free boundary system with low pH aluminium lactate buffers. This was adapted for use in starch gels by Elton and Ewart (20) at Chorleywood in 1962, who demonstrated that there were variations in the gluten proteins of eight wheat varieties. They also prepared different fractions of wheat proteins, including the albumins, globulins and prolamins (known as gliadins). In addition, they examined other cereal species and showed that the barley prolamin proteins (hordeins) could also be analysed by this method. These findings were confirmed by workers in different countries, including France, the Netherlands and Australia, over the course of the next few years. The research was extended to include a wider range of varieties and techniques, such as the use of polyacrylamide in place of starch gels (25). Although the earlier work had utilised flour made from different varieties, it was soon demonstrated that proteins could be successfully extracted and analysed from single grains. It was thus clear that gel electrophoresis of grain proteins possessed considerable potential for variety identification, especially of wheat. The first systematic approach was that of Ellis, working at the NIAB in 1971 (19), who used starch gel electrophoresis and other tests (such as the phenol test, grain hardness and coleoptile colour) in combination to distinguish between wheat varieties. The whole process required five days and could not be readily applied to varietal mixtures.

As research continued, it became clear that the gliadin proteins were probably the most useful for wheat variety identification when separated by starch gel electrophoresis. Workers at the Wheat Research Unit in Sydney and at the INRA station in Paris both produced detailed

schemes for variety identification, based on the analysis of single grains, at more or less the same time (1973-75) (3,33). Their method, and its subsequent modifications, provided the most widely used electrophoretic system for wheat variety identification for many years and was extensively utilised in the UK (18). It has only recently been generally superseded by PAGE techniques. The details of the method have been very well described (6,34). Briefly, gliadins are extracted from individual crushed wheat seeds or from wheat flour using 25% (v/v) 2-chloroethanol and separated using gels containing 10-12% (w/v) hydrolysed starch, 0.5M urea and 0.2% (v/v) aluminium lactate (buffered to pH 3.2 with lactic acid). Following electrophoresis at 8V/cm of gel, using an aluminium lactate-lactic acid buffer at pH 3.2, the gels are sliced and stained in a mixture of nigrosin and trichloroacetic acid. By this procedure, some 30-40 individual gliadin bands can be separated. Different varieties possess characteristically different combinations of these gliadin components and this forms the basis for being able to discriminate between and identify varieties. This procedure can be impressively powerful. For example, at the time of its use at the NIAB (1980-1984), it was possible to distinguish uniquely almost all of the wheat varieties contained on the UK National List. A number of authors published systematic keys, based on the presence/absence of particular gliadin bands and their relative staining intensities, to assist in identification (3,6,18,34). Starch gel electrophoresis methods, becoming available in the mid-1970s as a relatively routine procedure, had an immediate effect on grain trading in the UK. It was now possible to define the varietal specifications in contracts and moreover to monitor how well the contract was being honoured. The first UK company to adopt the method for wheat trading was Rank Hovis McDougall (RHM) in 1976. The impact of the technology on the quality of grain being milled has been well described by Ellis (17). With home-grown supplies, the number of deliveries which contained an undesirable admixture fell from 1 in 7 in 1976-77 to less than 1 in 50 in 1983-84. When imported loads were considered as well, the improvement in varietal quality was even greater.

However, there are certain disadvantages to the routine use of starch gels. For instance, the quality and purity of the starch used to prepare gels is crucial and can vary considerably from batch to batch, even from a reliable supplier. The preparation of the gels depends to some degree on the skill of the individual operator and the method is also rather slow, requiring a whole day to prepare, pre-run and run gels. The gels need to be sliced before being fixed and must be stained overnight. Also starch gels are not easily handled and cannot be stored satisfactorily. The number of simultaneous analyses possible is rather limited and concentrated protein samples are required. Furthermore, the method of loading the samples onto the gel (using filter paper squares to absorb the protein extracts) generally precludes the duplication of analyses. For these and other reasons, many laboratories have adopted PAGE methods to separate gliadins and identify wheat varieties.

With regard to barley, attempts were made to use similar procedures for discrimination between varieties. However, these were not generally very successful, probably because unsuitable techniques were being used to extract the hordeins. The systematic evaluation of hordein extraction and analysis by Shewry and colleagues at Rothamsted (30) laid the foundations for the application of a wide range of PAGE methods.

PAGE methods for cereal variety identification

There are many different types of PAGE which have been suggested for wheat and barley variety identification. Many of these use a version of PAGE, originally proposed by the Canadian workers Bushuk and Zillman (5) for the analysis of gliadins. This is essentially an adaptation of the starch gel method outlined above. Several other methods have been published for the fractionation of gliadins or hordeins, usually at acid pH. In addition, both SDS-PAGE and IEF have been utilised for the analysis of gliadins or hordeins and also of the glutelin fraction, usually known as glutenins. Tables 1.1 and 1.2 are lists of the main PAGE methods which can be used to fractionate seed proteins as an aid

to wheat or barley variety identification. References to the original publications have not been given, but can be found in reviews (7,10). It is clear that a wide range of methodology is available. The techniques listed in Tables 1.1 and 1.2 can, however, be conveniently divided into four types:

a) those using PAGE in lactate buffers
b) those using PAGE at acid pH in buffers other than lactate
c) SDS-PAGE methods
d) others, including PAGE at alkaline pH and IEF.

The majority of methods analyse gliadins or hordeins whilst the remainder will extract 'total' seed proteins, but in fact, under the analytical conditions employed, primarily resolve gliadins or hordeins and glutenins. With regard to protein staining, almost all of the methods use Coomassie Brilliant Blue type dyes, often in conjunction with 10-15% (v/v) trichloroacetic acid to remove the need for de-staining of the gels. It ought to be pointed out that not all of the methods listed in Tables 1.1 and 1.2 were designed primarily for variety identification, although all can be used. In particular, some of the SDS-PAGE methods applied to wheat are principally utilised for the purposes of genetic analysis or for the investigation of the bread-making quality of wheat breeding lines.

This wide choice of available methodology raises serious problems, not the least of which is comparing results from different laboratories. Overall, with both wheat and barley, the most commonly used routine methods are those involving the separation of gliadins or hordeins by PAGE at an acid pH. These methods have the advantage of combining relatively simple seed extraction and gel preparation procedures with comparatively rapid analysis times. However, there are still many methods to choose from and thus various organisations have begun to try to standardise analytical methods.

Standardisation of PAGE methods

It is clearly desirable to have standard analytical methods available, if for no other reason than allowing laboratories to evaluate and compare their own procedures. Also, they provide a means of arbitration in cases of dispute. Of the bodies which are involved in the area of testing food and foodstuffs, several have begun to examine the possibility of standardising PAGE methods for cereal variety identification. The International Standards Organisation (ISO) is discussing both starch gel electrophoresis of gliadins and a modified version of the Bushuk and Zillman PAGE procedure. The International Association for Cereal Science and Technology (ICC) has decided to adopt starch gel electrophoresis and a PAGE method using ready-made gradient gels and sodium lactate buffer at pH 3.1 (Table 1.1, method 17) as standard methods. In the laboratory at the NIAB, we have found that the methods using sodium and/or aluminium lactate buffers can be slightly unreliable, due in part to the varying purity of the chemicals used. Also, the 6% acrylamide gels often recommended, for example in the Bushuk and Zillman method, can be rather fragile and difficult to handle. Furthermore, the accumulation of metal salts on the platinum electrodes of the electrophoresis tank shortens their life-time considerably. For these reasons, the method now routinely used at the NIAB for wheat and barley variety identification is that recently adopted by the International Seed Testing Association (ISTA) as a standard reference method (9,14). The European Brewery Convention (EBC) also recommend this method for barley and the International Union for the Protection of New Varieties of Plants (UPOV) is actively investigating its use for the discrimination of wheat varieties.

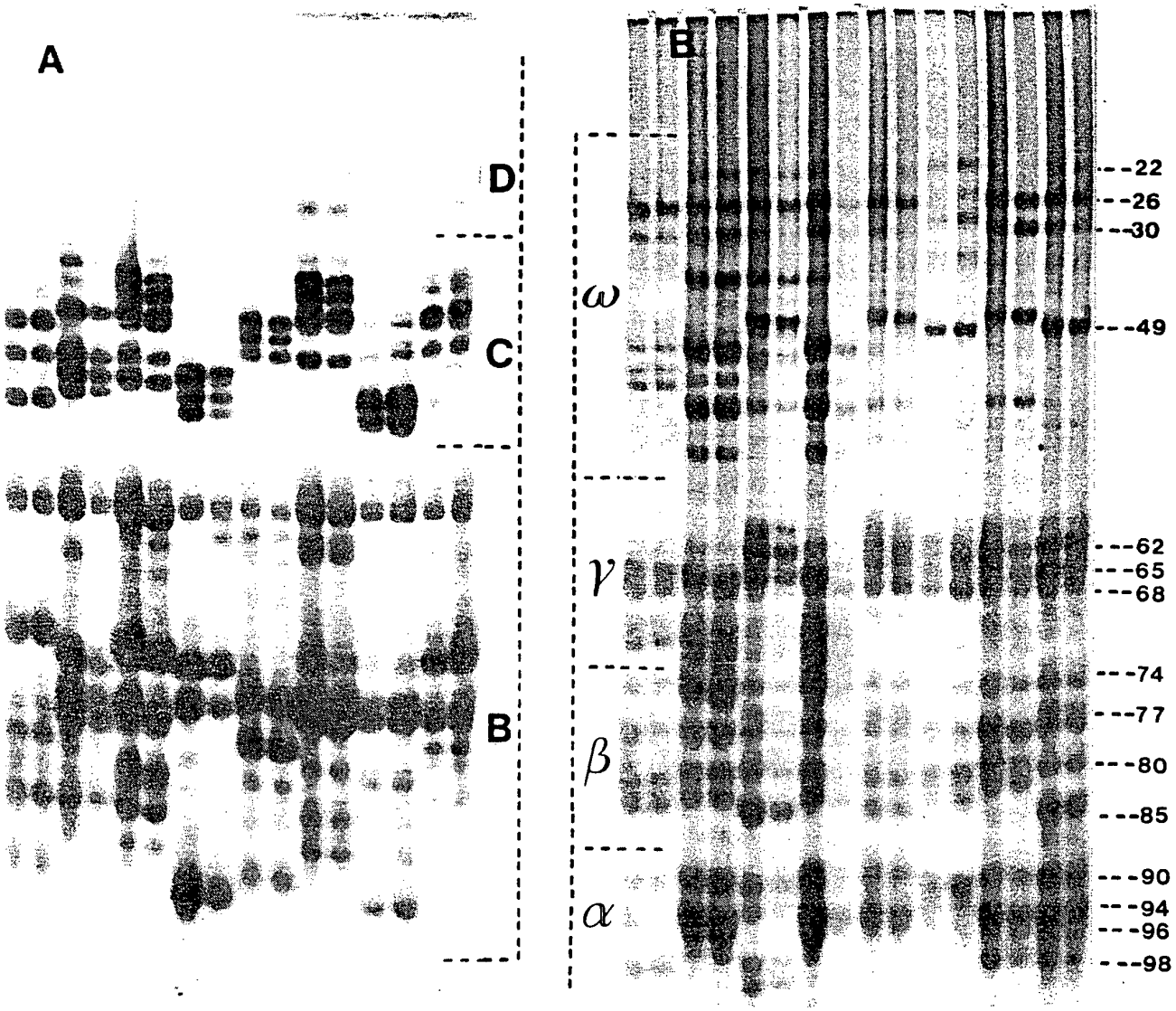
In this method, 9% acrylamide gels containing urea are used, with a glycine-acetic acid buffer at pH 3.2, to separate the chloroethanol-soluble proteins from seeds. For barley, the extracting solution also includes a reducing agent. Using the Pharmacia GE-2/4 apparatus and custom-made well-forming combs, it is possible to analyse 72 samples simultaneously in about 2 hours. The protein profiles of varieties of bread wheat and barley obtained using this method are shown in Figure 1.3. Such profiles, or electrophoregrams, are

generally representative of those generated by the acid gel methods listed in Tables 1.1 and 1.2. Gliadin and hordein can be fractionated into many components (about 40 for gliadin and 20 for hordein) by such electrophoretic procedures. Different cultivars have characteristically different combinations of these components, and the protein banding patterns can be used to discriminate between and to identify varieties. Because of their proximity to the primary genetic information, the protein patterns are for the most part unaffected by environmental factors. Severe sulphur deficiency can alter the intensity of the expression of certain groups of gliadins and hordeins, but does not affect the overall qualitative pattern. Otherwise, the protein composition of a given variety appears to be strictly a function of its genotype and hence independent of environmental influence (7,10,34). This allows the gliadin or hordein pattern to be used for identification purposes with more confidence than many morphological characters.

Nomenclature of gliadins and hordeins

Three systems of nomenclature have been proposed for the nomenclature of gliadin patterns produced by acid PAGE procedures. The first is an adaptation of the scheme originally used to name gliadins separated by moving boundary or starch gel electrophoresis and divides the gliadin electrophoregram into four regions, called α, β, γ and ω in decreasing order of mobility (see Figure 1.3). Within each region, individual protein bands are referred to by numerals. To provide additional flexibility, a series of over-and-under-linings, sub-scripts and dots is used to describe band intensities and slight variations in mobility. This system has been particularly used in Leningrad by Professor Konarev, to describe varieties and landraces of various cereal species and their wild relatives (23). The second system of nomenclature assigns numbers to each gliadin band, based on their mobility within the gel. In the scheme originally proposed by Autran and Bourdet (3) for starch gels, a reference band is given a mobility of 65, whereas Bushuk and Zillman (5) assign a reference band an arbitrary mobility of 50. Other band mobilities are then calculated relative to these. The staining intensity of the gliadin bands can be incorporated into this system, by using crosses (Autran) or numerical values (Bushuk). A

Figure 1.3 The protein patterns of (A) barley and (B) wheat varieties, following analysis by acid PAGE. The varieties are (left to right, in pairs): A - Patty, Igri, Triumph, Digger, Keg, Pirate, Marko, Goldspear; B - Chinese Spring, Conveyor, Norman, Ambassador, Maris Huntsman, Broom, Brimstone, Avalon. Also shown are different methods of nomenclature for hordeins and gliadins.



third and more sophisticated system of naming gliadin bands has been devised and used by Professor Sozinov and colleagues in Moscow. This assigns varieties a formula based on the gliadin alleles (patterns of bands) expressed in so-called 'allelic blocks' (26) and is perhaps more suited to genetic applications rather than routine varietal identification. In Table 1.3, the electrophoretic profile of the variety Chinese Spring is presented, according to the three systems. Chinese Spring is very widely used in genetic studies of wheat and was a reference variety in the collaborative experiments organised by the ISTA when testing their standard method.

The Konarev system as described above for gliadins can also be applied to hordein patterns. However, probably the simplest and most commonly used system of nomenclature for hordeins is illustrated in Figure 1.3 and recognises the existence of different groups of hordein polypepetides, usually termed A,B,C and D in decreasing order of mobility but also known as 'upper' and 'lower' groups or identified by numbers (10). Various types of pattern for B - and C - hordeins in particular have been catalogued. The groups are evident following hordein separation by either acid - PAGE or SDS-PAGE. In the system which is being considered for adoption within the EBC and ISTA, the groups of C - hordein patterns are designated by letters and B - hordein patterns by numbers. Thus for instance the varieties Natasha and Grit are designated A1 and A4, indicating that they have the same C - hordeins (group A), but different B - hordeins (groups 1 and 4). Similarly, Plaisant (B6) and Igri (D6) have the same pattern of B - hordeins (group 6), but differ in their C - hordeins (groups B and D). Other authors have assigned relative mobility numbers to hordein bands, based either on a reference band in a particular variety having an arbitrary mobility of 0.50 or on a commonly occurring band being given a mobility of 100. The relative staining intensities of the bands can be described by using a system of crosses or a series of numeric intensity values.

Discrimination between varieties

Using these types of nomenclature, authors in several different countries have produced catalogues of wheat and barley varieties and taxonomic keys to assist in identification (see 7,10,34 for references). For gliadins, the use of computerised data collection, storage and retrieval has been suggested (28) which would aid the compilation of an international catalogue of wheat varieties, provided that agreement could be reached on a standard reference method of analysis and nomenclature. The discriminating power of acid PAGE used for gliadin analysis and subsequent variety identification is often very impressive. For instance, Sapirstein and Bushuk (28) reported that out of 122 Canadian bread and durum wheat varieties, 16 pairs had similar gliadin patterns, the remaining 90 being uniquely different. In the UK, it has been shown that 138 out of 155 varieties can be readily distinguished (9).

Regardless of the electrophoresis method used for hordein analysis, the degree of discrimination achieved between barley varieties is not as great as that observed in wheat. Thus, Shewry and co-workers reported the classification of 164 mostly European varieties into 32 groups using SDS-PAGE (31). By IEF, 77 French varieties could be separated into 21 clearly distinguishable groups (29). Using the ISTA standard reference acid PAGE method, Cooke and Morgan (14) recognised 41 different groups in a sample of 191 varieties of diverse origins. There are three main reasons for this somewhat less successful electrophoretic discrimination. Firstly, barley is a diploid species, whereas wheat is a hexaploid. The triplication of the chromosomes in wheat increases the number of protein - encoding genetic loci, leading to the occurrence of more gliadin bands, greater possibilities for mutational divergence and hence more potential differences between genotypes. Again, many modern barley varieties share a common or narrowly based ancestry and so are genetically close to one another. Finally, the genes which encode for the major groups of hordeins are found at separate, but linked loci on the same arm of the same chromosome, limiting the possibilities for recombination between the groups. For example, using SDS-PAGE, workers at Rothamsted reported the existence of 17 patterns in the B-hordein group and 8 patterns of

C-hordeins (4). However, because of genetic linkage, the patterns are not randomly associated and the total number of groups of B- and C-hordeins together is 34 compared to the theoretical maximum of 136.

To improve the discrimination between groups of wheat or barley cultivars, it is often possible to use electrophoretic methods in combination. Thus IEF or PAGE at pH 4.6 of hordein can subdivide some of the groups of varieties which are identical following SDS-PAGE (10,34). Modifications of SDS-PAGE extraction methods and gel conditions can enhance discrimination, as can the use of a basic pH gel system following initial analysis of hordein at acid pH (4,10). Again, SDS-PAGE of glutenins can be useful in distinguishing between wheat varieties which have identical gliadin spectra (10,34). IEF has not found general favour for wheat variety identification, even though it has been used to demonstrate differences in gliadin spectra not observed by other methods. The major disadvantge is cost, as IEF gels are much more expensive than other types. The use of ultra-thin layer (UTL) gels can reduce this disadvantage, but laboratories have still not adopted the methodology widely.

IEF has been successfully used as the first step in various two - dimensional (2-D) electrophoresis mapping techniques. Thus IEF in combination with starch gel electrophoresis at pH 3.2, with PAGE at pH 8.9 or with SDS-PAGE has been used to fractionate wheat and barley seed proteins and to demonstrate distinct differences between closely-related varieties and lines (7,10,34). In wheat, another 2-D method which uses separations at two different pH values (3.2 followed by 9.2) has also been developed (24). These techniques can resolve at least 60 major protein components from wheat or barley seeds and anything up to 100 minor ones. The use of very sensitive silver-staining procedures can increase the number of major protein spots detectable to something approaching 200. Clearly, such very powerful methods have considerable potential for the demonstration of differences between varieties. However, the expense and time involved and the level of expertise which is required both for the successful

analysis and for the interpretation of gels are severe limitations to the routine application of 2-D methods for the present. Progress in the simplification of 2-D technology and improvement of the reproducibility of the data produced could be very significant for seed and variety work.

Another way of enhancing the discrimination between particular varieties or groups of varieties is to analyse electrophoretically various isozymes. The great success of the seed protein approach has limited the need for alternative techniques. However, some authors have examined the potential of isozyme analysis. Thus starch gel electrophoresis, PAGE and IEF have all been used to separate different isozymes from seeds and/or leaves of both wheat and barley (for references, see 7,10,34). A range of isozymes has been utilised, including esterases, peroxidases, α- and β- amylases and acid phosphatases. Useful classifications of barley varieties have been produced by Nielsen and Frydenberg (27) who used the distribution of α- amylase and esterase isozymes from seeds or seedlings to group 107 varieties and by Andersen (1), who carried out the same analysis for the 47 varieties comprising the Danish National List. At the NIAB, we have shown that IEF of seed esterases is a useful way of distinguishing between certain barley varieties which are identical following PAGE or SDS-PAGE (unpublished).

A different approach to the improvement of electrophoretic discrimination, which is particularly applicable to barley, is to combine electrophoresis with a small number of morphological characters, preferably those which can be observed relatively easily from the grain. One suggested scheme proposes the examination of rachilla hair type, aleurone colour, anthocyanin pigmentation and the presence or absence of spicules and ventral furrow hairs, in conjunction with hordein composition. This enables all of the varieties currently encountered commercially within the UK to be distinguished. It would be interesting to evaluate this, and alternative schemes, within a broader, European context.

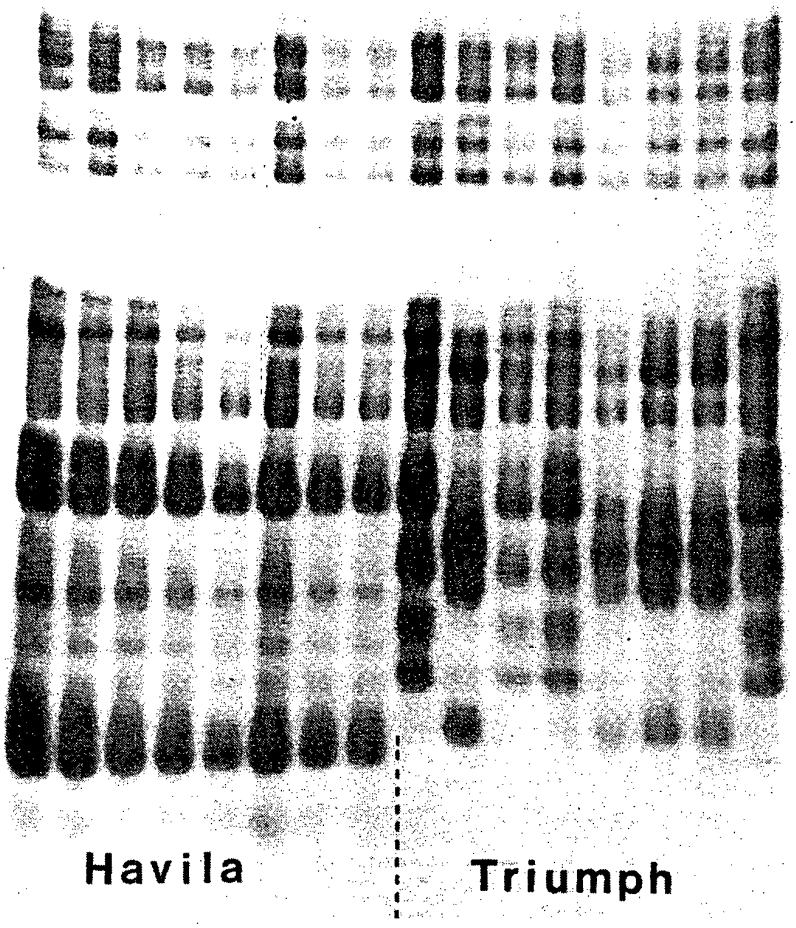
Uniformity of varieties

Because they are self-pollinated, varieties of wheat and barley are very highly inbred and hence demonstrate a high level of phenotypic (morphological) uniformity. Depending on how many generations of selfing have been performed before the final production of the variety, they are also largely genetically homozygous, i.e. all individuals carry the same alleles. Thus the gliadin, hordein or isozyme composition of plants within a given variety is normally repeatable and uniform. However, wherever this has been thoroughly investigated, a small number of varieties are found which contain two or more electrophoretic lines or biotypes (1,3,4,6,7,9,10,14,18,27,31,34, see also Figure 1.4). These lines can only be adequately observed if a sufficient number of individual seeds of a variety are analysed, which does not always seem to have been appreciated by investigators. Overall, about 10-15% of western European cereal varieties are known to be mixed electrophoretically. In the UK, biotypes have been reported in 14 out of 191 barley varieties (14) and in 6 out of 155 wheat varieties (9), using acid PAGE. The existence of biotypes is not entirely unexpected and arises from the lack of selection by breeders for electrophoretic homogeneity. The method of breeding can itself give rise to a mixed structure if a variety is multiplied from a range of 'families' of plants, with an identical morphological phenotype but differing in the distribution of alleles at the protein-encoding loci. Provided that the lines are recognised and catalogued, their existence presents no real problems and does not detract from the overall power of electrophoresis for wheat or barley variety identification.

Analysis of other cereals

One of the attractions of electrophoretic identification procedures is that essentially the same methods can be used for investigation of a range of species. Thus for example, the ISTA standard reference procedure, devised for wheat and barley, can be applied with only minor modifications for the study of varieties of oats, durum wheat, triticale and certain other cereals (11). This is illustrated in Figure 1.5. Methods such as SDS-PAGE can also be used for different

Figure 1.4 An example of the uniformity and non-uniformity of electrophoretically - revealed protein patterns in barley varieties. The first 8 tracks are individual seeds of the variety Havila, which is uniform electrophoretically. The remaining tracks are individual seeds of Triumph, which has two biotypes, differing in the B-hordein region.



cereal species (7,10). These other cereals have not been as extensively researched as wheat and barley, but classifications of oat varieties have been published. It has proved possible using acid PAGE to divide 50 varieties which are available in the UK into 14 groups on the basis of the composition of their prolamin proteins (avenins). Further discrimination between these groups can be achieved by sequential analysis with SDS-PAGE and IEF (12). Isozyme methods have also been utilised for oat variety identification (10,34).

Practical applications of electrophoresis

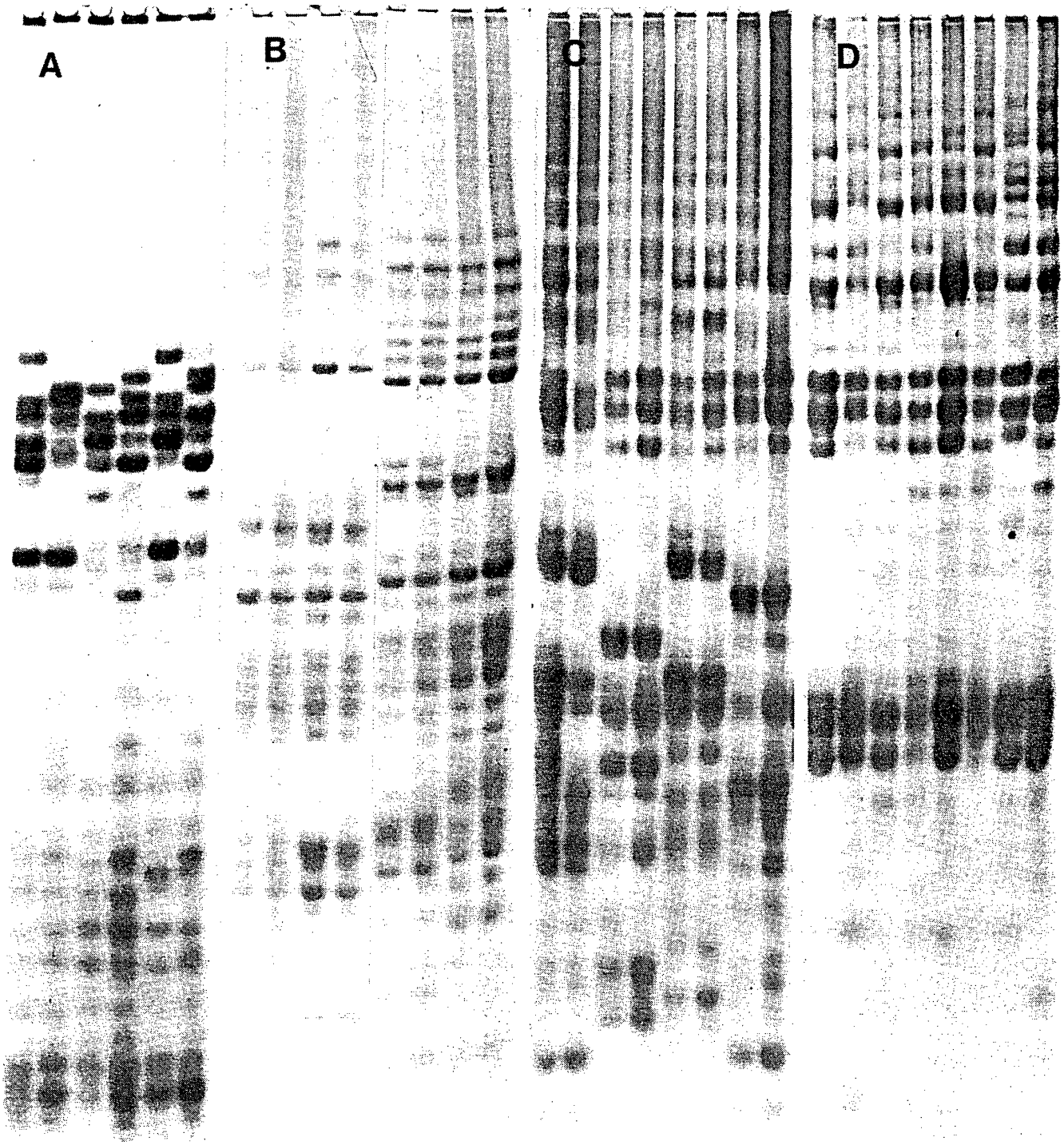
There are several practical ways in which it can be envisaged that electrophoresis methods could be applied. These include: a) determination or confirmation of varietal identity and/or purity; b) distinctness testing; c) seed certification procedures; d) quality control e.g. in milling and malting; e) species identification; f) documentation of genetic resourses. This subject has been considered previously in various publications (8,10,11).

For several of these purposes, electrophoresis can often be used comparatively, i.e. does the protein pattern of the sample correspond to that of the known reference variety? For this to be successful, two things are vital. Firstly, the user must have access to the authentic stocks of varieties, usually maintained by statutory licensing authorities. Secondly, the electrophoretic method must give reliably repeatable results. The best way to achieve this is to use one of the well-proven standard reference methods, such as that recommended by the ISTA.

Distinctness testing and seed certification

Given the discriminating power of electrophoretic gliadin and hordein analysis, it is evident that it could have considerable potential for distinctness testing purposes and the statutory registration of new varieties. However, at the present time, UPOV does not recognise the use of electrophoresis in determining distinctness. The situation is kept under constant review and an exercise is being conducted by UPOV to assess the possibilities presented by electrophoresis for the

Figure 1.5 The use of acid PAGE to separate proteins from varieties of (A) oats, (B) durum wheat, (C) triticale and (D) rye.



testing of winter wheat varieties. Whilst there may well be both technical and administrative difficulties associated with the recognition of electrophoresis as a means of determining distinctness, it does appear rather perverse not to use these comparatively rapid, cheap and powerful methods. In the UK electrophoresis of candidate varieties of wheat and barley is used as an unofficial means of confirming the distinctness (or otherwise) in specific problematical cases. The potential is clear and several instances have arisen at the NIAB in a range of crop species whereby electrophoretic analysis of proteins or enzymes would allow the discrimination between what appear to be morphologically identical varieties (15). It would be possible to make the protein profile of a variety part of the official description used for registration purposes. The electrophoretic spectra of wheat, barley and oats, as produced by acid PAGE, are included in the booklet 'Detailed Descriptions of Varieties' published annually by the NIAB. A combination of electrophoresis and a small number of key morphological characters ought to provide a highly discriminating system. It may be more useful if 2-D patterns were used in this way, as such maps would go a long way towards uniquely characterising a given variety.

As seed certification is concerned with varietal identity and purity, electrophoresis has some evidently obvious applications in this area. Apart from cases of identification of seed lots where mis-labelling etc. has occurred, some rather more interesting possibilites are available. A good example is the analysis of wheat seeds following a phenol test, a grain colouration test which is generally used to determine the uniformity of seed lots but which also gives limited information as to identity. Some seed lots give a mixed phenol reaction in some seasons and certain varieties are variable in any case. It is possible to check the identity of apparently aberrant seeds by electrophoresis. Soaking the seeds to remove excess phenol allows the gliadins to be extracted and analysed by acid PAGE. Another interesting situation arises when varieties are morphologically identical and can only be distinguished by means of an additional 'special test'. The oats varieties Cabana and Leanda, provide such a case. These are distinguishable only by a seedling disease resistance test, the former being resistant to an isolate of oat mildew race 2

whereas the latter is susceptible. This can cause problems in certification, when the need for a rapid and unambiguous identification arises, since the disease test has to be specially commissioned, is lengthy, expensive and somewhat unreliable. Although the two varieties have identical avenin patterns following acid PAGE by the ISTA method, they can be unequivocally distinguished by an IEF method (15). It is also possible to use electrophoresis to characterise to some degree the morphological variants which are noted during seed production and certification. A systematic approach to evaluating the origin of off-types in Australian wheat varieties using starch gel analysis of gliadins showed that four different types of variant could be categorised. These were a) mechanical admixture with known or unknown varieties; b) seed arising from natural out-crossing; c) segregants of the original cross that produced the variety; d) phenotypic variants, which are usually environmentally induced and have no genetic basis.

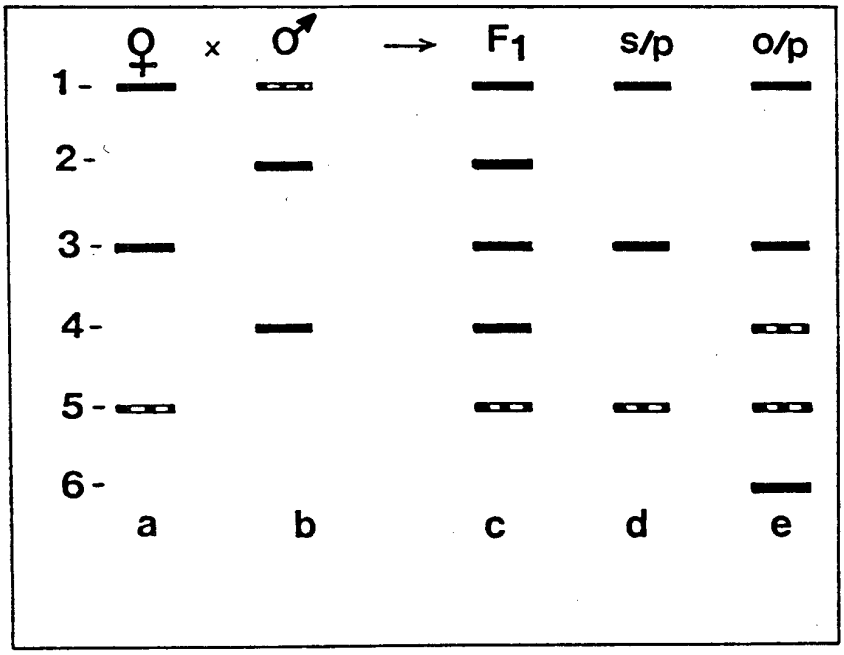
A somewhat specialised, but valuable, application of electrophoresis in the area of seed certification is the determination of hybrid purity. The exploitation of heterosis in the production of F_1 hybrids is outside the scope of this chapter. Essentially, however, hybrid production involves the crossing of two highly inbred parental lines, with the resultant offspring (the F_1 hybrid seed) displaying enhanced performance and uniformity. For this to be effective, it is essential that the selected female line is, or can be rendered, male-sterile i.e. produces no fertile pollen. The advent of chemical hybridising agents has enabled this to be relatively easily achieved and plant breeders are now able to produce F_1 hybrid varieties of crops such as wheat. This has long been an objective of plant breeders, following on from the enormous success and profitability of the hybrid maize business, especially in the USA. An important part of hybrid seed production is concerned with evaluating the genetic purity of seed lots and it is here that electrophoresis can play a major role. The great power of the techniques lies in the nature of the genetic expression of electrophoretically-revealed proteins. Most bands are inherited in a simple Mendelian fashion and expressed co-dominantly. What this means can be seen in Figure 1.6. Consider that two inbred parental lines, A and B, are to be used for F_1 hybrid seed production and that they have

different protein compositions. The female line A is made male-sterile and pollen from the male line B used for fertilisation. The resultant F_1 hybrid seed (C) has the protein pattern of both parents and hence can be recognised. Moreover, the two predominant sources of contamination, namely seed arising from self-pollination of the female (sibs) and seed arising from pollination with another male, can both be detected (D and E in Figure 1.6). The use of a suitable electrophoretic method to analyse proteins or enzymes from individual seeds can thus easily and rapidly give a measure of the purity of F_1 hybrid seed. These methods are widely used within the USA and elsewhere by hybrid maize seed companies (10). At the NIAB, we have successfully used acid PAGE of gliadins to investigate the purity of F_1 hybrid wheat seed lots, in both experimental and commercial seed production situations.

Quality control

An increasingly important application of electrophoresis of gliadins and hordeins is in checking the varietal purity of the grain being processed by millers and malsters. The impact of starch gel electrophoresis in this area has already been mentioned and indeed this technique is still used by some companies. The quality control managers at many major British mills are now skilled in the use of electrophoresis for variety identification. A similar situation is found in other major grain trading countries, such as Australia. Electrophoresis of hordeins has had a similar, though perhaps less widespread, effect on the malting and brewing industry. In western Europe, routine monitoring of the malting barley in trade has only been in progress since about 1980, but has already led to a substantial improvement in the quality of the grain utilised. Professor Schildbach (Berlin) has described this improvement as greater than that brought about by any technological changes in processing and by plant breeding. At the NIAB, all of the malting barley being exported from the UK under contract with the Grain and Feed Trade Association (GAFTA) is analysed for varietal purity by a combination of visual grain inspection and acid PAGE of hordeins. Financial penalties are incurred by merchants whose cargoes fail to meet the required standards. Much of the malting barley from the Irish Republic is likewise monitored.

Figure 1.6 An illustration of the use of electrophoresis to measure the purity of F_1 hybrid seed. If the parental lines (A and B) differ in their electrophoretically - revealed protein composition, then the F_1 hybrid seed (C) represents the summation of the banding patterns of the two parents. Seed arising from self-pollination of the female parent (D) can be detected, as can seed arising from the pollination of the female by an unknown male (E).



This is in fact a common occurrence, particularly within the member countries of the EBC. Of additional interest is the fact that it is possible, in some cases, to identify the variety of barley grain which has been malted. SDS-PAGE methods of hordein analysis seem to be best suited for this purpose, especially if the extraction medium is modified by the use of increased concentrations of reducing agent and by the incorporation of dimethylformamide.

The importance of electrophoretic quality control should not be underestimated. The technology has had a wide-ranging impact on the processing of cereal grains, due to the impressive levels of varietal discrimination that can be attained. The major disadvantages of the methods are that they require skilled analysts, especially for gel interpretation, and they are not really suited for rapid intake testing. Despite many efforts to reduce the analysis time, an electrophoretic separation still takes some four hours to perform, including protein extraction and gel staining times. Improvements in electrophoretic equipment may reduce this further, but it is difficult to envisage the time ever being less than two hours. Future research efforts could perhaps be more profitably directed at automatic gel evaluation, since this is still the most difficult and operator-dependent stage. The standardisation of methodology now being undertaken should help and may also assist in improving sample throughput if gels could be obtained commercially of the correct composition. This would largely remove the need for a clean preparation area within intake laboratories.

Species identification

Many of the foregoing electrophoretic methods which are used for the identification of cereal varieties can also be used to distinguish between cereal species. Two examples of particular relevance to seed testing are the identification of wild oats in seed crops and the discrimination between species of Triticum and related genera.

Several Avena (oat) species are found as important competitive weeds in agricultural crops. For instance, in the UK Seeds Regulations, the presence of three species - A. fatua (common wild oat), A. ludoviciana (winter wild oat) and A. sterilis (wild red or animated oat) - in seed lots is controlled by specific standards. Thus purity checks on seed lots entered or accepted for certification are required to identify these species. Whilst morphological characters can be used, identification is often very difficult, or even impossible, when seeds have been commercially harvested and the characteristic external structures of the seeds damaged or removed. In these circumstances, electrophoretic methods can often be used for the identification of wild oat species. The most comprehensive demonstration of this is the report of Cooke and Draper (13) who used acid PAGE to separate the avenins of both wild and cultivated oats. By examining multiple accessions of the wild species, to allow for any intra-species variation, it was possible to use PAGE to distinguish between the wild species and also to distinguish them from varieties of cultivated oats.

Seeds of related species can be difficult to identify categorically. This is well illustrated by considering bread wheat (Triticum aestivum), durum wheat (T. durum), rye and triticale. Some of the more modern varieties of triticale appear more wheat-like than earlier types. A system of biochemical characterisation would thus be advantageous and various forms of electrophoresis of seed proteins and enzymes have been shown to provide such characterisation. For instance, it is evident that the various kinds of acid PAGE or SDS-PAGE used to analyse prolamins provide distinctive electrophoregrams for different species (see Figures 1.3 and 1.5). A systematic evaluation of electrophoretic methods used for discriminating between wheats, rye and triticale has been carried out at the NIAB (16). Two methods are recommended for general use, either singly or in combination. These are acid PAGE of prolamins and IEF of grain esterases. The second of these is perhaps more suited for routine application, since there is a very limited degree of varietal variation in esterase composition compared to gliadin composition. Either method will readily discriminate between bread wheat and durum wheat, since the latter lacks the gliadin or esterase components encoded by the D-genome of

hexaploid bread wheat. The protein components of rye are also characteristic and, furthermore, rye is an open-pollinating species, so that varieties generally display considerable heterogeneity in their protein composition. Triticales contain proteins donated by both of the parental species and can also be distinguished by their banding patterns.

This type of analysis has an application in determining the composition of flour. It is important that the durum flour used for making semolina and pasta should not be mixed with bread wheat, which would affect both the colour and texture of the final cooked product. The presence of wheats can be readily detected by acid PAGE of the gliadins extractable from the flour, since such wheats possess characteristic low mobility ω-gliadins (16). Adulteration of pasta flour with as little as 5% of bread wheat can be easily ascertained by PAGE of the ethanol-soluble protein fraction. The converse situation - durum wheat contamination of flour intended for bread-making - is more difficult to uncover. However, gliadin analysis by IEF in an immobilised pH gradient in gels containing 5M urea allows the two wheat species to be differentiated and a 2% admixture of durum wheat in bread wheat to be detected (21).

Documentation of genetic resources

The formation of the International Board for Plant Genetic Resources (IBPGR) in 1974 acted as a stimulus for the collection of a diverse range of wild and cultivated plants. There are now many such collections, housed in 'gene-banks' throughout the world. These genepools are important, not least because they provide sources of variation which can be utilised as raw material for the improvement of agricultural and horticultural crops. A major problem exists in trying to evaluate and characterise these collections, which can contain many thousands of accessions. However, it is imperative to know the amount of genetic variation represented. Morphological and agronomic descriptors can be used for this purpose, but such evaluation is a lengthy and costly process. Clearly, electrophoresis of proteins and isozymes could be of fundamental importance in this respect.

This has been recognised and an excellent review by Simpson and Withers (32), published by the IBPGR, summarises the potential uses of protein analysis for gene-banks and contains over 750 references of value. The electrophoretic analysis of storage proteins from primitive wheat and barley lines and related species has been used to evaluate material collected in Nepal and Yemen. The application of electrophoretic methods would allow the compilation of catalogues of genetic resources on a world-wide basis, provided that a standardised system of analysis and protein band nomenclature can be agreed.

Conclusions

This chapter has reviewed and summarised the use of electrophoresis of protein and enzymes to identify cereal varieties. Electrophoretic methods are already widely applied in various sectors of the cereals industry and have had a particularly large impact in the areas of seed and variety testing and quality control. Electrophoresis is also widely used in plant breeding, as has been reviewed recently (10). The attractions of the methods - principally speed, low cost, a high degree of discrimination and freedom from environmental interactions - ensure that electrophoresis will continue to be utilised within the industry.

The efforts of various bodies to standardise methodology are to be welcomed and should particularly assist laboratories who are newly adopting this technology. There is still a requirement, especially for routine large-scale operations such as intake control or assessment of varietal purity for seed certification purposes, to simplify the methods and equipment somewhat. For instance, the commercial availability of suitable gels would be most helpful. The most difficult and skillful task remains the evaluation of gel patterns and research into automatic or machine interpretation would be appropriate. French workers recently reported the development of an automated soft-laser scanning densitometry and computer-aided analysis system for wheat variety identification (2). However, they specifically stated that the system was not intended for the identification of varietal mixtures or for distinguishing closely similar varieties and was most suitable for use by well-trained personnel, already familiar with protein band patterns and nomenclature. Again, since not all varieties

of wheat and barley can be uniquely identified by a single technique, further studies of alternative procedures might also be profitable.

Electrophoresis methods for cereal variety identification have been available for several years now. However, the application of the methods overall is still increasing as more people in different sectors of the industry become aware of the possibilities and advantages offered. Although it is generally unwise to attempt to predict the future, it seems probable that electrophoresis will continue to be the method of choice for most circumstances, certainly in the short-to-medium term.

References

(1) Andersen, H.J. (1982). Isoenzyme characters of 47 barley cultivars and their application in cultivar identification. Seed Science and Technology 10, 405-413.

(2) Autran, J.C. and Abbal, P. (1988). Wheat cultivar identification by a totally automatic soft-laser scanning densitometry and computer analysis of protein electrophoregrams. Electrophoresis 9, 205-213.

(3) Autran, J.C. and Bourdet, A. (1975). L'identification des variétés de blé; établisement d'un tableau général de détermination fondé sur la diagramme électrophorétique des gliadines du grain. Annales del'Amélioration des Plantes 25, 277-301.

(4) Bunce, N.A.C., Shewry, P.R. and Miflin, B.J. (1984). The identification of barley cultivars by analysis of hordein polypeptides. in Draper, S.R. and Cooke, R.J. (eds) Biochemical Tests for Cultivar Identification, ISTA, Zurich, pp 69-77.

(5) Bushuk, W. and Zillman, R.R. (1978). Wheat cultivar identification by gliadin electrophoregrams. I. Apparatus, method and nomenclature. Canadian Journal of Plant Science 58, 505-515.

(6) Clydesdale, A. and Draper, S.R. (1982). A revised classification of wheat gliadin electrophoregrams. Journal of the National Institute of Agricultural Botany 16, 53-60.

(7) Cooke, R.J. (1984). The characterisation and identification of crop cultivars by electrophoresis. Electrophoresis 5, 59-72.

(8) Cooke, R.J. (1986). Gel electrophoresis - a role in agriculture. in Dunn, M.J. (ed.) Electrophoresis '86, VCH Veslagsgesellschaft, Weinheim, pp 203-217.

(9) Cooke, R.J. (1987). The classification of wheat cultivars using a standard reference electrophoresis method. Journal of the National Institute of Agricultural Botany 17, 273-282.

(10) Cooke, R.J. (1988). Electrophoresis in plant testing and breeding. Advances in Electrophoresis 2 (in press).

(11) Cooke, R.J. (1988). The standardisation of electrophoresis methods for variety identification. Proc. ISTA Symposium, Leningrad (in press).

(12) Cooke, R.J. and Cliff, E.M. (1984). The characterisation of oat cultivars by electrophoresis. Journal of the National Institute of Agricultural Botany 16, 415-430.

(13) Cooke, R.J. and Draper, S.R. (1986). The identification of wild oat species by electrophoresis. Seed Science and Technology 14, 157-167.

(14) Cooke, R.J. and Morgan, A.G. (1986). A revised classification of barley cultivars using a standard reference electrophoresis method. Journal of the National Institute of Agricultural Botany 17, 169-178.

(15) Cooke, R.J., Parnell, A. and Draper, S.R. (1984). The application of different electrophoresis methods to cultivar identification. in Draper, S.R. and Cooke, R.J. (eds) Biochemical Tests for Cultivar Identification, ISTA, Zurich, pp 32-40.

(16) Cooke, R.J., Smith, T.M. and Ainsworth, C.C. (1986). Discrimination between bread wheat, durum wheat, rye and triticale by electrophoresis of seed proteins and enzymes. Seed Science and Technology 14, 693-704.

(17) Ellis, J.R.S. (1984). The cereal grain trade in the United Kingdom : the problem of cereal variety. Philosophical Transactions of the Royal Society of London, series B, 304, 395-407.

(18) Ellis, J.R.S. and Beminster, C.M. (1977). The identification of UK wheat varieties by starch gel electrophoresis of gliadin proteins. Journal of the National Institute of Agricultural Botany 14, 221-231.

(19) Ellis, R.P. (1971). The identification of wheat varieties by the electrophoresis of grain proteins. Journal of the National Institute of Agricultural Botany 12, 223-235.

(20) Elton, G.A.H. and Ewart, J.A.D. (1962). Starch gel electrophoresis of cereal proteins. Journal of the Science of Food and Agriculture 13, 62-72.

(21) Gunther, S., Postel, W., Weser, J. and Gorg, A. (1986). Isoelectric focusing of alcohol soluble proteins with immobilized pH gradients. in Dunn, M.J. (ed) Electrophoresis '86, VCH Verlagsgesellschaft, Weinheim, pp 485-488.

(22) Jones, R.W., Taylor, N.W. and Senti, F.R. (1959). Electrophoresis and fractionation of wheat gluten. Archives of Biochemistry and Biophysics 84, 363-376.

(23) Konarev, V.G., Gavrilyuk, I.P., Gubareva, N.K. and Peneva, T.I. (1979). Seed proteins in genome analysis, cultivar identification and documentation of cereal genetic resources : a review. Cereal Chemistry 56, 272-278.

(24) Lafiandra, D., Kasarda, D.D. and Morris, R. (1984). Chromosomal assignment of genes coding for the wheat gliadin protein components of the cultivars 'Cheyenne' and 'Chinese Spring' by two-dimensional (two-pH) electrophoresis. Theoretical and Applied Genetics 68, 531-539.

(25) Lee, J.W. (1964). Gluten protein composition of individual wheat grains. Nature 201, 79-80.

(26) Metakovsky, E.V., Novoselskaya, A.Y., Kopus M.M., Sobko, T.A. and Sozinov, A.A. (1984). Blocks of gliadin components in winter wheat detected by one-dimensional polyacrylamide gel electrophoresis. Theoretical and Applied Genetics 67, 559-568.

(27) Nielsen, G. and Frydenburg, O. (1972). Distribution of esterase isozymes, α -amylase types and DDT reactions in some European barleys. Zeitschrift für Pflanzenzuchtung 68, 213-224.

(28) Sapirstein, H. and Bushuk, W. (1986). Computer - aided wheat cultivar identification and analysis of densitometric scanning profiles of gliadin electrophoregrams. Seed Science and Technology 14, 489-517.

(29) Scriban, R., Autran, J.C., Strobbel, B. and Nicolaidis, M. (1979). Synthèse des différéntes récherche analytiques sur la chimiotaxonomie des orges et des malts. Proceedings of the 17th European Brewery Convention, pp 571-581.

(30) Shewry, P.R., Hill, J.M., Pratt, H.N., Leggatt, M.M. and Miflin, B.J. (1978). An evaluation of techniques for the extraction of hordein and glutelin from barley seed and a comparison of the protein composition of Bomi and Riso 1508. Journal of Experimental Botany 29, 677-692.

(31) Shewry, P.R., Pratt, H.M., Faulks, A.J., Parmar, S. and Miflin, B.J. (1979). The storage protein (hordein) polypeptide pattern of barley (Hordeum vulgare L) in relation to varietal identification and disease resistance. Journal of the National Institute of Agricultural Botany 15, 34-50.

(32) Simpson, M.J.A. and Withers, L.A. (1986). Characterization using isozyme electrophoresis : a guide to the literature. International Board for Plant Genetic Resources, Rome.

(33) Wrigley, C.W. and Shepherd, K.W. (1974). Identification of wheat cultivars by laboratory procedures : examination of pure samples of grain. Australian Journal of Experimental Agriculture and Animal Husbandry 14, 796-804.

(34) Wrigley, C.W., Autran, J.C. and Bushuk, W. (1982). Identification of cereal varieties by gel electrophoresis of the grain proteins. Advances in Cereal Science and Technology 5, 211-259.

TABLE 1.1 PAGE Methods Used for Variety Identification in Wheat[a]

	Extraction	Gel	Buffer	Type[b]	Time[c]	Samples[d]
	Acid PAGE					
1)	70% EtOH, 1h, room temp, +sucrose[e]	6%T,5%C	8.5mM Al lactate, lactic acid, pH 3.1	H	6.5	10
2)	as 1)	as 1)	3.5mM Na lactate, lactic acid, pH 3.1	H	5.0	10
3)	25% CE, 1h, room temp, +saccharose[e]	as 1)	as 2)	H	5.0	10
4)	as 1) + glycerin[e]	as 1)	as 1) in gels; ÷2 in tank	V	5.25	8
5)	as 4)	as 1)	as 2) or 4)	V	6.5-7.0	8
6)	as 1)	as 1)	as 1)	V	4.0	11
7)	1.5M DMF, 10 min, + sucrose[e]	as 1)	as 1)f	H	6.5	10
8)	as 1)	as 1)	as 2)	V	5.0	10
9)	as 1)	as 1)	as 1)	V	1.3, 0.6 or 0.3[g]	20
10)	as 1)	7.5%T 5%C	8.5mM Na lactate, lactic acid, pH 3.1	V	5.0	6
11)	70% EtOH, 20 min, 40° + sucrose[e]	8%T, 5%C	as 1)	V	3.0	-h
12)	as 1)	as 10)	13mM Na acetate, acetic acid, pH 3.1	V	-h	-h
13)	as 1)	7%T, 2.7% C, 6M urea	12.5mM KOH, 7.2mM lactic acid, pH 3.6, 6M urea	V	5.5	7
14)	1M urea	8%T, 5%C 1M urea	0.4% acetic acid, 0.04% glycine, pH 3.1	V	2.0	14

Table 1.1 (continued)

	Extraction	Gel	Buffer	Type[b]	Time[c]	Samples[d]
15)	6% urea, 1h room temp	as 1)	0.19M acetic acid, glycine, pH 3.1	V	2.0	14
16)	25% CE, 2h room temp	9%T, 5%C 1M urea	0.19M acetic acid, 0.012M glycine, pH 3.1 in gels; ÷2.5 in tank	V	2.0	16
17)	1M urea, 1h room temp	2.5%-13%T 5%C	4.25mM NaOH, lactic acid, pH 3.1	V	2.5	12
18)	40% DMSO, 2% saccharose overnight, room temp	6%T, 6%C	0.019M Tris, lactic acid, pH 7.5 (gel); Tris, 0.02M lactic acid, pH 3.1 (tank)	V	4.5	24
19)	Ethylene glycol, 1-3 days, room temp.	12%T, 3%C	0.125M acetic acid (gel); 0.04M acetic acid (tank), pH 2.7	V	7.25	40
	SDS - PAGE					
20)	55% PrOH, ME, 0.5h, sonicated[j]	17.5%T, 1% C,0.1% SDS[i]	0.125M Tris-borate, pH 8.9 0.1% SDS	V	3.0	14
21)	0.06M Tris-HCl pH6; 2% SDS, 5% ME, 10% glycerol, 2h room temp+ 2 min at 100°	17%T, 1.5% C,0.1% SDS [i]	0.37M Tris-HCl pH 8.8, 0.1% SDS (gel); 0.025M Tris, 0.19M glycine, 0.1%SDS, pH 8.3 (tank)	V	18.0	13
22)	as 21)+ 20% DMF	as 21)	as 21)	V	18.0	18
23)	as 21) no 100°step	10%T, 0.7 %C, 0.1% SDS[i]	as 21)	V	4.0	18
24)	as 21) at 40° over-night	8.4%T, 0.08%C, 0.1% SDS[i]	as 21)	V	2.0	14

Table 1.1 (continued)

	Extraction	Gel	Buffer	Type[b]	Time[c]	Samples[d]
25)	0.12M Tris-borate pH8.9, 0.2% SDS, 1% ME, 0.5h room temp	5%T, 5%C 0.1% SDS	0.12M Tris-borate, pH 8.9, 0.1% SDS	H	4.5	7
26)	8mM Al lactate, 10% glycerol, pH 3.2[j]	7-12%T 5%C, 0.1% SDS[i]	as 21)	V	4.0	25
	Other Methods					
27)	30% CE, 1h, room temp	6%T, 4%C[i]	0.36M Tris-HCl, pH9.1 (gel); 0.04M Tris-GAB, pH 9.2 (tank)	V	3.0	12
28)	as 27	5%T, 4%C 4M urea, 1% CHAPS, 2% ampholytes, pH 6-8	Serva pH3 solution (anode) pH 10 solution (cathode)	H	1.5	20
29)	as 17)	7.5%T, 3%C 2M urea, 2% ampholytes pH 3.5-10	1M H_3PO_4 (anode), 1M NaOH (cathode)	H	3.5	20

Footnotes

a) Abbreviations: CE- 2-chloroethanol, DMF- dimethylformamide; ME- 2-mercaptoethanol; EtOH-ethanol; GAB- γ-amino-butyrate; PrOH- propan-2-ol; DMSO- dimethylsulphoxide
T = total acrylamide concentration,
C = cross-linker (Bis) concentration.
b) Horizontal or Vertical system
c) Time in hours for typical separation.
d) Number of samples per gel.
e) Added prior to electrophoresis.
f) Gels polymerised in water and equilibrated in buffer.
g) Depending on gel thickness and number of gels in parallel.
h) Information not supplied.
i) Stacking gel also used.
j) Samples dried and resuspended in Tris-HCl buffer, pH 6.8 +SDS+ME + glycerol.

TABLE 1.2 PAGE Methods Used for Variety Identification in Barley[a]

	Extraction	Gel	Buffer	Type[b]	Time[c]	Samples[d]
	Acid Page					
1)	55% PrOH 2% MTG, 2h room temp, + sucrose[e]	6%T, 5%C	8.5mM Al lactate, lactic acid, pH 3.1	H	4.0	9
2)	as 1)	as 1)	as 1)	V	3.0	20
3)	as 1) with ME instead of MTG	9%T, 5%C	as 1) (gel); 4.25mM NaOH, lactic acid, pH 3.1 (tank)	V	2.5	16
4)	3M urea, 1%ME	8%T, 5%C, 1M urea	0.4% acetic acid, 0.04% glycine, pH 3.1	V	2.0	14
5)	50% CE	7%T, 2.5% C[f]	KOH/acetic acid, pH 2.9 (gel); glycine/ acetic acid pH 4.0 (tank)	V	3.5	12
6)	20% CE, 3M urea, 1% MTG, over-night, 4°	9%T, 5%C 1M urea	0.19M acetic acid,0.012M glycine,pH 3.1 in gels; ÷ 2.5 in tank	V	2.5	16
7)	1M urea, 2% ME, 1h, room temp	3-27%T, 5%C	as 3)	V	2.0	12
	SDS - PAGE					
8)	55% PrOH, 2% ME, son-icated, 0.5h, 20°[g]	17.5%T 1%C, 0.1% SDS[f]	0.125M Tris-borate, pH 8.9, 0.1% SDS	V	3.0	15
9)	0.06M Tris-HCl, pH 6.8 2% SDS, 8% ME, 15% DMF, 10% glycerol, 3h room temp +2 mins at 100°	17%T 1.5%C 0.1% SDS[f]	0.375M Tris-HCl, pH 8.8 0.1% SDS (gel); 0.025M Tris, 0.19M glycine, 0.1% SDS, pH 8.3, (tank)	V	16.0	12
10)	as 9) but with no room temp extraction	13%T 1.5%C 0.1% SDS[f]	as 9)	V	16.0	12

Table 1.2 (continued)

	Extraction	Gel	Buffer	Type[b]	Time[c]	Samples[d]
11)	as 9) with 5% ME, and no DMF	10%T, 1.5%C, 0.1% SDS, 4M urea[f]	as 9)	V	3.0	12
12)	as 9) with no 100° extraction	10%T, 0.7%C, 0.1% SDS[f]	as 9)	V	4.0	18
13)	as 9), but with 5% ME and 20% DMF	as 9)	as 9)	V	18.0	18
	Other Methods					
14)	50% PrOH, 0.3% ME, 2h room temp[h]	6%T, 3%C, 4M urea, 2.5% ampholytes, pH 5-9	5%H_3PO_4 (anode);5% ethylenediamine (cathode)	H	2.0	12
15)	1M urea, 1% ME, 1h room temp	7.5%T, 3%C, 2M urea, 2% ampholytes, pH 3.5-10	1M H_3PO_4 (anode); 1M NaOH (cathode)	H	3.5	20
16)	30% CE	4%T, 4%C, 5M urea, 20% glycerol, 'Immobilines' pH 6-8	10mM glutamic acid (anode); 10mM NaOH (cathode)	H	16.0	22

Footnotes

a) Abbreviations: MTG-monothioglycerol: ME- 2-mercaptoethanol;
CE-2-chloroethanol; PrOH - propan-2-ol;
DMF-dimethylformamide;
T = total acrylamide concentration,
C = cross-linker (Bis) concentration.

b) Horizontal or Vertical system

c) Time in hours for typical separation

d) Number of samples per gel

e) Added prior to electrophoresis

f) Stacking gel also used

g) Samples are dried, alkylated and reduced before analysis

h) Samples are dried and resuspended in 6M urea before analysis.

TABLE 1.3 Three Systems of Nomenclature for the Gliadin Bands in the Wheat Variety Chinese Spring

(1) 'Konarev' system (see Konarev et al., 1979) (23).

$\alpha\bar{2}\ 4\ \underline{6}\ 7 \quad \beta 2\ 3_1\ 4\ 5 \quad \gamma 2\ 3\ \bar{5} \quad \omega 3\ \ddot{4}\ 5\ 7\ 8\ \underline{9}$

Notes
1) Number over-lined indicates lower band intensity than standard.
2) Number under-lined indicates higher band intensity than standard.
3) Subscript 1 indicates slightly greater mobility than standard.
4) $\ddot{}$ over number indicates a split band.

(2) 'Sozinov' system (see Metakovsky et al., 1984) (26).

1A	1B	1D	;	6A	6B	6D
14	6	3	;	7	3	6

Notes
1) 1A, 1B etc. refer to chromosomal locations.
2) Numbers refer to patterns characteristic of blocks of bands.

(3) 'Autran' system (see Autran and Bourdet, 1975) (3).

26	30	50	52	53	62	65	67	71	74	77	81	83	90	96	99	105
+++	++	++	+	++	+	+++	++	++	++	++	++	++	+(+)	+(+)	+(+)	(+)

Notes
1) Numbers refer to band positions
2) Crosses denote relative band intensities.

Chapter 2:

CEREAL VARIETY IDENTIFICATION BY HIGH PERFORMANCE LIQUID CHROMATOGRAPHY

Introduction

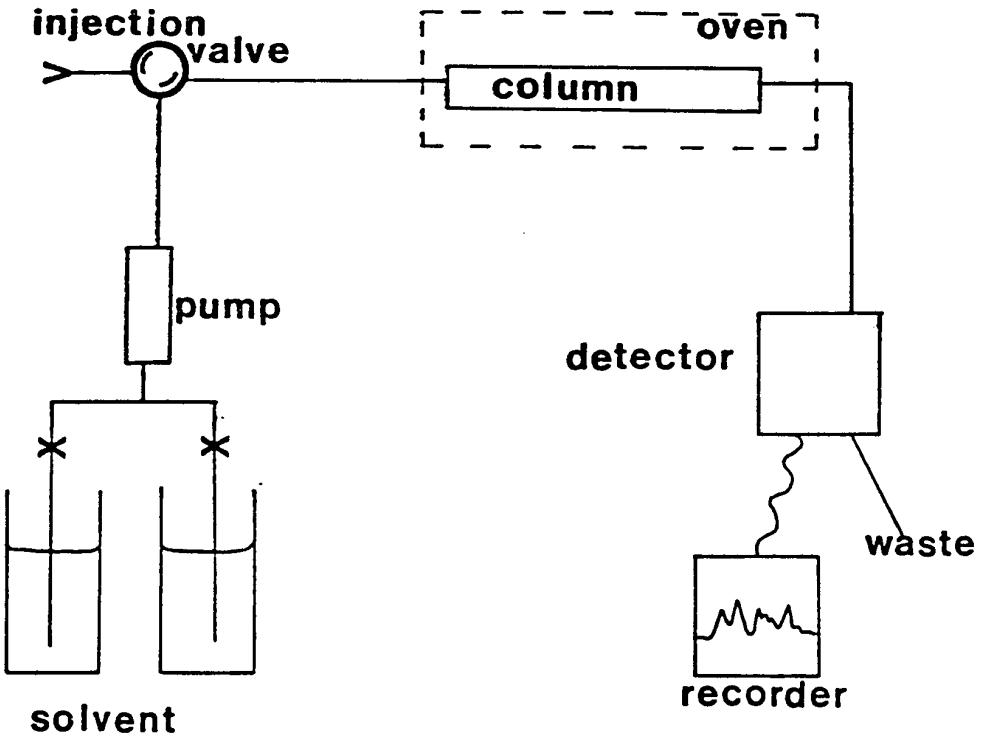
The term chromatography describes a range of analytical techniques for the separation of mixtures of compounds. It is probably true to say that the existence of chromatography provides one of the cornerstones of biochemistry, since it allows the isolation and preparation in a pure form of most of the molecules which are involved in the functioning of cells and organisms. As with electrophoresis, it is not a new procedure. Indeed, the first use of chromatography was described by the botanist M. Tswett in 1903, who was interested in separating plant pigments. All chromatographic techniques are based on the principle that the components of a mixture of chemical compounds may be separated and concentrated into zones by passage through a system consisting of two phases, one stationary and one mobile. These names describe the roles of the phases, as the mobile phase carries the mixture through the static stationary phase. The separation is brought about by the differential affinity of the components of the mixture for the different phases.

Liquid chromatography is characterised by the fact that both the mobile and stationary phases are liquids, although the latter is often bonded onto a solid support medium. The stationary phase is usually contained within a column and in the simplest type, the sample is applied to the top of this column and then the liquid mobile phase is passed through it. The components of the mixture begin to separate as they pass through the column (a process usually known as elution) and partitioning between the two phases ocurs. Finally, the separated zones emerge from the end of the column and must be detected in some way. The power of liquid chromatography resides, firstly, in the array of phases that are available, and secondly in the fact that each particle of the stationary phase (plus

support medium) acts as a partitioning unit. During passage through a column, many thousands of individual partitions occur. A revolution in liquid chromatography occurred in the late 1960's when scientists from the du Pont company showed how by using very small (30 μm diameter) particles of support medium, coated with suitable stationary phases, it was possible to achieve an extremely fine resolution of mixtures. These small particles required the use of high pressures to force the mobile phase through the column at acceptable rates, and so high pressure (or high performance) liquid chromatography (HPLC) was developed (12).

Over the past 20 years, HPLC has become one of the most widely applied techniques in biochemistry and has been shown capable of resolving mixtures of all types of compounds. Modern equipment is extremely sophisticated and versatile. However, the basic requirements are relatively simple and are shown in diagrammatic form in Figure 2.1. The focal point of the equipment is the analytical column and these can be obtained in a variety of sizes and diameters, packed with a wide range of support media of varying particle sizes and coated with any one of dozens of stationary phases. The operating temperature of the column can be altered by means of a thermostatically-controlled heating jacket. The sample to be analysed is generally introduced onto the column by some form of injection device and the mobile phase is carried through the column to elute the components of the sample by a pump arrangement. Modern systems are often equipped with devices for creating mobile phase solvent gradients of various types, to enhance resolution. The separated components emerging from the column are monitored by passage through a detector, usually an ultra-violet (UV) device although other kinds of detection system are available. The detector can be linked to a recorder, to provide a printed trace of the separation and often also to a computing integrator, which will measure the position and size (area) of the various peaks. Increasingly, the whole chromatographic system is becoming automated. For example, automatic sample injection facilities are commonly used, enabling the analysis to be carried out continuously without operator intervention.

Figure 2.1 A schematic representation of the basic requirements of an HPLC system. The sample applied to the analytical column is eluted with solvents (mobile phase) delivered through a pump system. Separation occurs on the column (stationary phase) and the components emerging from the end of the column pass through a detector, linked to a recorder and integrator.



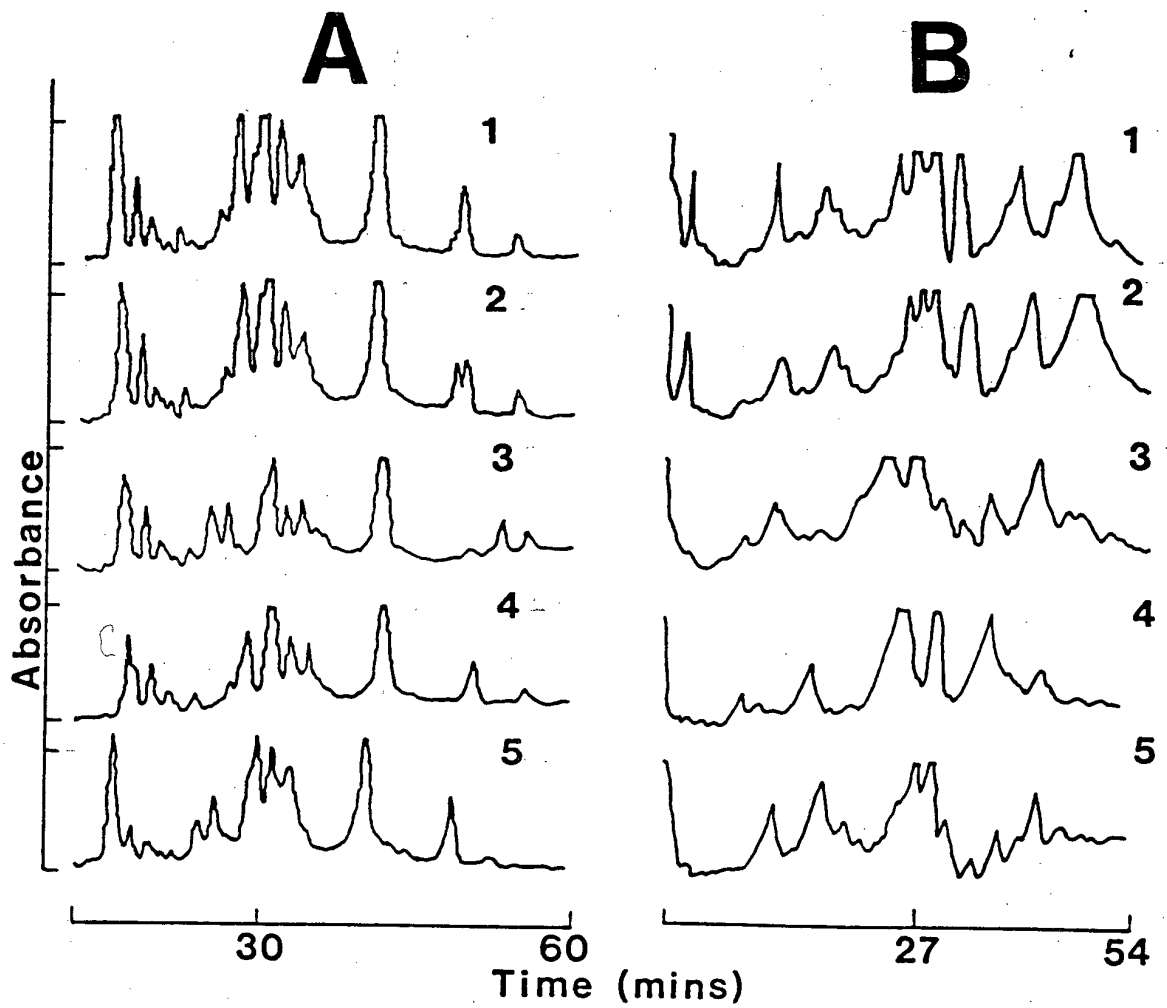
From the point of view of this article, the most important development in HPLC was the success, in the mid-1970's, of attempts to separate protein mixtures (13). This in turn was due primarily to the availability of large pore size silica-based stationary phases, the surfaces of which are chemically modified by the presence of non-polar hydrocarbon chains (or ligands). The large pore size was critical, since it now allowed larger molecules, such as proteins, to penetrate the supports and hence fully interact with the stationary phase. Because this system is in essence a reversal of the normal chromatographic situation, with non-polar stationary phases and polar mobile phases, it is often known as reversed-phase HPLC or RP-HPLC. In RP-HPLC, proteins bind hydrophobically and then separate by partitioning between the stationary and mobile phases. There are many different types of silica-based phases available commercially, which differ in the size and shape of the silica particles, the pore size, and the length of the carbon chain of the bonded ligands, which can vary from 1 to 18 atoms. All of these parameters will influence protein separation. The composition of the mobile phase is also clearly of importance. Proteins will absorb to RP-HPLC stationary phases equilibrated with relatively polar mixtures of water and organic solvents and will begin to elute when the organic solvent concentration is increased. For cereal proteins, mixtures of water and acetonitrile have been found to be particularly useful. The pH of the mobile phase is also important and trifluoroacetic acid (TFA) is usually added, at a low concentration, to adjust the pH to about 2.5. This has the effect of ionising the free amino groups on the proteins, which leads to decreased elution times and increased peak resolution.

HPLC and cereal variety identification

The first successful separation of wheat seed storage proteins by RP-HPLC was reported from the USA in 1983 by Bietz (4). He analysed both gliadins and glutenins and showed that the resolution was equal to, or in some cases better than, that achieved by electrophoretic methods. Subsequent work, both in the USA and elsewhere, has demonstrated unequivocally that HPLC can be used to separate seed proteins and identify varieties of wheat, barley and other

cereals and that the discrimination achieved between varieties is similar to that found by the use of electrophoresis (1,10,17,20). Several different chromatographic systems have been employed and the salient points of some of the methods are summarised in Table 2.1. The type of varietal profile produced by RP-HPLC analysis of seed proteins is illustrated in Figure 2.2. As the same proteins (gliadins, hordeins, glutenins) are utilised in HPLC as in electrophoresis, many of the attractions and advantages considered in Chapter 1 apply equally. For example, the protein profiles are relatively free from environmental influence, due to their proximity to the primary genetic information (DNA). Varieties are distinguishable by the presence or absence of particular peaks of protein detectable at specific points (elution or retention times) on the profiles. There are two important additional points. Firstly, the physical basis for separation by RP-HPLC is chiefly the surface hydrophobicity of the proteins in question. This is different from the situation in electrophoresis, where the proteins are separated on the basis of their size and/or charge. For this reason, it is sometimes possible to achieve discriminations between gliadin or hordein compositions (and hence varieties) by HPLC which are not possible by electrophoresis (1,5,10,18,19). The converse can also be true. Thus HPLC should best be viewed as being a technique which is complementary to electrophoresis, as regards variety identification. The second point also concerns the potentially enhanced discrimination achieved by HPLC and derives from the use of peak heights or areas to make varietal distinctions. This is equivalent to using protein banding intensities following electrophoresis, but can be used with much more confidence, since the computerised integration of peaks allows a far more accurate quantification of protein levels than can be obtained by electrophoresis. This means that quantitative differences, provided that they are reproducible, can readily be used to distinguish between particular varieties following HPLC of their gliadins or hordeins.

Figure 2.2 The protein profiles of different varieties of A) wheat and B) barley produced by typical HPLC procedures (adapted and re-drawn from reference 21 and from unpublished data of Dr M Griffin). The varieties are: A 1-Brimstone, 2-Galahad, 3-Norman, 4-Rapier, 5-Avalon; B 1-Maris Otter, 2-Keg, 3-Patty, 4-Kym, 5-Egmont.



As an example of the discriminating power of HPLC, Allison (1) investigated barley varieties from two of the groups produced by classification of hordein patterns following SDS-PAGE. For one group, 13 electrophoretically identical varieties could be divided into five sub-groups on the basis of qualitative differences in the hordein profiles produced by RP-HPLC. Moreover in one of these sub-groups, all six varieties showed quantitative differences in the proportions of certain hordein components. For another electrophoretically identical group of eight varieties, RP-HPLC revealed the presence of four sub-groups, one qualitative and three based on the relative proportions of hordeins. It must be appreciated, however, that HPLC has not been used to anything like the same extent as electrophoresis for cereal variety identification and no formal classifications of large numbers of diverse varieties have been reported. Again, any seed to seed variability in RP-HPLC profile has been relatively little studied, since most authors use 'flours' from varieties in their studies. It is thus difficult to make valid comparisons of discriminating power. Even so, it is clear that HPLC analysis of seed proteins is an extremely valuable tool for identification purposes.

Advantages and disadvantages of HPLC

As was mentioned above, the fact that HPLC can be considered as another way of separating seed storage proteins means that most of the advantages of electrophoresis discussed in Chapter 1 apply also to HPLC. In addition, the possibility exists for substantial automation of the HPLC separation procedure, by using programmable sample injection systems for instance, and the data produced can not only be readily quantified, which enhances discrimination possibilities, but can also be stored and processed by computer (see below also). However HPLC does inevitably have some disadvantages, the two primary ones being are 1) the capital and operating costs, and 2) the time required for adequate separation of proteins. An HPLC system suitable for protein separation could cost anything up to £25,000 especially when equipped for automation as outlined above. The analytical columns used have a limited life and their performance

declines with time. This could be especially serious if computerised capture and comparison of profiles were being utilised. Although there are ways of protecting the columns, for instance by the use of 'guard' or pre-columns (which act effectively as sieves, removing particulate material), these accordingly lengthen the separation time and increase the costs. There is also some evidence that different columns of apparently the same type can produce different sample resolutions, which again could seriously hamper the automated comparison of profiles (19). Table 2.1 shows that the typical separation time for RP-HPLC is 50-60 minutes per sample. Although this compares favourably with electrophoresis, which can take anything from 1-18 hours, plus gel staining time to produce results, it has to be appreciated that with electrophoresis, multiple samples (up to 100-150) are being analysed simultaneously, which means that the level of sample throughput is much higher. Various authors have suggested using shorter analysis times for RP-HPLC (6,11,16) and indeed it is possible to obtain varietal profiles in 10-15 minutes. However, there is an inevitable loss of resolution of proteins using these rapid times and as a consequence a reduction in discriminating power. Relatively rapid analysis times can also be obtained using ion-exchange, rather than reversed - phase, HPLC (3,21). Again, though, the resolution of protein mixtures, and hence the degree of varietal discrimination, is reduced when compared to alternative and lengthier procedures. A further consideration is that the HPLC equipment represents highly sophisticated technology and care must be taken to ensure that the equipment is adequately maintained. For example, the performance of the pump used to force the mobile phase through the column must be monitored, since small changes in flow rate can result in significant changes in varietal protein profiles.

Practical applications of HPLC

As a technique which complements electrophoresis in many respects, it might be anticipated that HPLC could be of use in many or all of the situations discussed in Chapter 1. Indeed, this would seem to be the case (1,5,10) although again it must be emphasised that because of its recent and continuing development, HPLC has yet to be applied in any of the practical situations where electrophoresis is currently

being utilised. However, HPLC can certainly be used to discriminate between and identify cereal varieties and thus has potential applications in the areas of distinctness testing, seed certification and seed testing. As regards quality control in the milling and malting industries, it is unlikely that HPLC can compete with electrophoresis in terms of sample throughput. However, there is a tendency towards the use of shorter analytical columns, which will reduce the time required for individual separations. Columns can be run in 'series', with an array of columns being operated from the same pump, but with separate detectors. The marketing of cartridge column systems is reducing the level of technical sophistication of HPLC equipment. Even given such improvements, though, it is doubtful that HPLC can ever equal electrophoresis for speed when multiple samples need to be analysed, for example for grain by grain analysis of a seed lot.

The computerised systems for the manipulation of data which are available for HPLC do raise some intriguing possibilities for alternative procedures. The most interesting of these is perhaps the analysis of a bulked sample, or flour, of a particular seed lot, in place of the examination of individual seeds. Then only one (or perhaps two) analyses would be required on a particular cargo of seeds, which, even if it took 50-60 minutes, would be more rapid than electrophoresis. The use of such a system would depend on the ability of the computer software to evaluate mixtures of varietal patterns. Software does exist which enables protein profiles to be compared and differences between them to be highlighted. The actual processing is still rather slow and requires considerable operator interaction. However, this is largely a function of computing power and advances in this area are likely. The potential of this approach has already been demonstrated. Work at Trent Polytechnic, in the laboratory of Dr Griffin, has shown that it is possible to detect the adulteration of durum flour for pasta-making with bread wheat, at levels of about 5%, by comparing the ratios of peak areas of certain characteristic peaks in the protein profiles (unpublished). Such a system might, given sufficient development, be able to identify mixtures of cereal varieties in a flour, at about the same levels of detection. This could be valuable not only for quality control of

the grain entering a mill, but also elsewhere to monitor the composition of the flour being processed. A further application may be in assessing the purity of F1 hybrid wheat seed lots, provided that the parental lines used had sufficiently different protein profiles.

This approach to the use of HPLC for cereal variety identification is clearly a profitable area for future research effort. The advantages it would offer over the electrophoretic analysis of individual grains to determine sample purity would be considerable, at least in terms of speed. Whether or not the resolution achieved, and the level of admixture detectable, would be suitable and acceptable are questions that await an answer.

Conclusions

HPLC analysis of proteins can readily be used to distinguish between and identify cereal varieties. The discrimination between varieties is broadly equivalent to that achieved by electrophoretic techniques, although this has not been studied systematically. HPLC will produce a result for a single analysis within 50-60 minutes, but sample throughput overall does not match that of electrophoresis. The capital equipment and operating costs are fairly high. However, the increased mechanisation and automation of HPLC, both for sample analysis and data capture, storage and retrieval, mean that the operator-dependence of the method, especially as regards variety identification from a given protein profile, is relatively low. Furthermore, the precise quantification and computerised manipulation of data which are possible raise some interesting potential applications in the analysis of varietal mixtures. This is certainly a topic which is worthy of future research. Efforts should also be made to establish agreed protocols of analysis amongst users, particularly if HPLC finds applications in the quality control field. Many of the procedures published so far have utilised the Synchropak RP-P (C18) column (2,6,7,8,9,14,15,18), but this is not necessarily ideal (19,21) and further investigation of alternative types of column would be valuable. HPLC technology, especially in the area of protein analysis, is still developing in terms of column packings,

instrumentation and computer software for the analysis of results. There seems little doubt that such advances will lead to improvements which will ultimately benefit the seed and variety testing authorities and the cereals industry.

References

(1) Allison, M.J. (1988). Reversed phase HPLC as an aid for the identification of barley cultivars. in Linsteens, H. F. and Jackson, J.F. (eds) Modern Methods of Plant Analysis, vol. 7, Springer-Verlag, Berlin pp 22-34.

(2) Allison, M.J. and Bain, H. (1986). The use of reversed-phase high performance liquid chromatography as an aid to the identification of European barley cultivars. Euphytica 35, 345-351.

(3) Batey, I.L. (1984). Wheat varietal identification by rapid ion-exchange chromatography of gliadins. Journal of Cereal Science 2, 241-248.

(4) Bietz, J.A. (1983). Separation of cereal proteins by reversed-phase high performance liquid chromatography. Journal of Chromatography 255, 219-238.

(5) Bietz, J.A. (1985). High performance liquid chromatography: how proteins look in cereals. Cereal Chemistry 62, 201-212.

(6) Bietz, J.A. and Cobb, L.A. (1985). Improved procedures for rapid wheat varietal identification by reversed-phase high-performance liquid chromatography of gliadin. Cereal Chemistry 62, 332-339.

(7) Bietz, J.A., Burnouf, T., Cobb, L.A. and Wall, J.S. (1984). Gliadin analysis by reversed-phase high-performance liquid chromatography: optimization of extraction conditions. Cereal Chemistry 61, 124-129.

(8) Bietz, J.A., Burnouf, T., Cobb, L.A. and Wall, J.S. (1984). Wheat varietal identification and genetic analysis by reversed-phase high-performance liquid chromatography. Cereal Chemistry 61, 129-135.

(9) Burnouf, T. and Bietz, J.A. (1984). Reversed-phase high-performance liquid chromatography of durum wheat gliadins: relationships to durum wheat quality. Journal of Cereal Science 2, 3-14.

(10) Burnouf, T. and Bietz, J.A. (1987). Identification of wheat cultivars and prediction of quality by reversed-phase high-performance liquid chromatographic analysis of endosperm storage proteins. Seed Science and Technology 15, 79-99.

(11) Cressey, P.J. (1987). Identification of New Zealand wheat cultivars by reversed-phase high-performance liquid chromatography. New Zealand Journal of Agricultural Research 30, 125-130.

(12) Done, J.N., Kennedy, G.J. and Knox, J.H. (1972). Revolution in liquid chromatography. Nature 237, 77-81.

(13) Henschen, A., Hupe, K.P., Lottspeich, F. and Voelter, W. (1985). High Performance Liquid Chromatography in Biochemistry. VCH Verlegsgesellschaft, Weinheim, 638p.

(14) Huebner, F.R. and Bietz, J.A. (1985). Detection of quality differences among wheats by high-performance liquid chromatography. Journal of Chromatography 327, 333-342.

(15) Huebner, F.R. and Bietz, J.A. (1986). Assessment of the potential bread-making quality of hard wheats by reversed-phase high-performance liquid chromatography of gliadins. Journal of Cereal Science 4, 379-388.

(16) Lookhart, G.L., Albers, L.D. and Bietz, J.A. (1986). Comparison of polyacrylamide gel electrophoresis and high-performance liquid chromatography analyses of gliadin polymorphism in the wheat cultivar Newton. Cereal Chemistry 63, 497-500.

(17) Lookhart, G.L., Albers, L.D., Pomeranz, Y. and Webb, B.D. (1987) Identification of U.S. rice cultivars by high-performance liquid chromatography. Cereal Chemistry 64, 199-206.

(18) Marchylo, B.A. and Kruger, J.E. (1984). Identification of Canadian barley cultivars by reversed-phase high-performance liquid chromatography. Cereal Chemistry 61, 295-301.

(19) Marchylo, B.A., Hatcher, D.W. and Kruger, J.E. (1988). Identification of wheat cultivars by reversed-phase high-performance liquid chromatography of storage proteins. Cereal Chemistry 65, 28-40.

(20) Smith, J.S.C. (1986). Biochemical fingerprints of cultivars using reversed-phase high performance liquid chromatography and isozyme electrophoresis: a review. Seed Science and Technology 14, 753-768.

(21) Wingad, C.E., Iqbal, M., Griffin, M. and Smith, F.J. (1986). Separation of hordein proteins from European barley by high-performance liquid chromatography: its application to the identification of barley cultivars. Chromatographia 21, 49-54.

TABLE 2.1 HPLC Methods Used for Cereal Variety Identification[a]

Extraction	Column [b]	Time [c]	Reference
WHEAT (including durum)			
70% EtOH or 55% PrOH, 30 mins; centrifuge 27000g, 10 mins (gliadins)	Synchropak RP-P (C18), 30°C	65	7,8,9
0.04M NaCl, 30 mins; centrifuge; extract residue with 70% EtOH, then 70% acidified EtOH; centrifuge; adjust supernatant to pH 6.5, cool, centrifuge and extract residue in SDS solution + DTT(glutenins)	Synchropak RP-P (C18), 30°C	45	14
as 7	Synchropak RP-P (C18); Aquapore RP-300 (C8); Ultrapore RPSC (C3), 70°C	55-65[d]	6
as 7	Synchropak RP-P (C18); Bakerbond RP (C8), 70°C	65 30	15
as 7	Ultrapore RPSC (C3), 70°C	15-25	16
as 7	Aquapore RP-300 (C8), room temp.	55[d]	11
50% PrOH + 1% DTT, 60°C, 30 mins; centrifuge at 15600g and filter supernatant (Millipore)	Aquapore RP-300 (C8) (and others-best was Supelcosil LC-30 (C8)), 30°C	120	19
10% NaCl, centrifuge; extract residue with 70% EtOH, 30 mins; centrifuge and dissolve residue in buffer	Mono-Q, 25°C[e]	10-15	3
BARLEY			
50% PrOH + 1% DTT, 60°C, 30 mins; centrifuge at 8000g	Synchropak RP-P (C18), room temp.	120	18
55% PrOH + 1% DTT, 60°C 30 mins; centrifuge 5000g	Synchropak RP-P (C18), room temp.	80	2

Extraction	Column [b]	Time [c]	Reference
70% EtOH; centrifuge 5000g, 10 mins and filter supernatant (Millipore)	Ultrapore RPSC (C3), room temp.	55	21
as 3 or 21	Protein Pak DEAE-5PW, room temp.[e]	45	21

Footnotes:

a) Abbreviations - EtOH - ethanol; PrOH - propanol; SDS - sodium dodecyl sulphate; DTT - dithiothreitol

b) Column - the commercial name of the column (stationary phase) used is given; the C number indicates the length of the carbon chain of the bonded ligand; the operating temperature is also given. The mobile phase generally used is a gradient of acetonitrile in water, both containing trifluoroacetic acid. The details of the gradient vary.

c) Time - typical time (in minutes) for one analytical separation.

d) Shorter analysis times were also used in these experiments, but lead to a decrease in the resolution obtained. Some authors have suggested that such decreased times may be suitable for specific purposes.

e) These are not reversed-phase procedures, but can be classified as ion-exchange versions of HPLC. Elution of proteins is achieved using a mobile phase consisting of an increasing gradient of sodium acetate.

Chapter 3:

IMMUNOLOGICAL APPROACHES TO VARIETY IDENTIFICATION

Introduction

In the previous sections of this article, two methods for analysing cereal seed storage protein composition have been considered. Another approach to utilising the genetic polymorphism of proteins and hence identifying varieties is to make use of immunological methods. Proteins are antigenic - that is they will elicit the production of antibodies if injected into an experimental animal, by means of the normal immune response of that animal. Antibodies are high molecular weight proteins that recognise and bind to the substance which gave rise to their formation (i.e. the antigen, also known as the immunogen). It is relatively straightforward to produce antibodies to cereal seed proteins. Such antibody preparations have been used by some workers, especially in the USSR, to distinguish between species of cereals (e.g. durum wheat and bread wheat, wild and cultivated oats) (3). The method used is generally an immuno-diffusion technique, in which an agar gel, containing a particular antibody preparation is made. Protein extracts under analysis are placed in wells made in this gel and allowed to diffuse. At the points where the protein (antigen) reacts with or binds to the antibody, a line of precipitation is formed, since the antigen - antibody complex is insoluble. Thus protein extracts from particular cereal species can be identified by comparing the number and intensity of the precipitation lines formed. Unfortunately, there are problems in trying to adapt this type of procedure for variety identification, as opposed to species identification. Most of the difficulties arise from the properties of the cereal seed proteins. In particular, prolamins (gliadins and hordeins), which are the most useful proteins for cereal variety identification, are not especially powerful as antigens and fairly large amounts of them are required to produce a reasonable level of antibody production in experimental animals. Also, the rather unusual solubility properties of the

prolamins can cause problems. The most serious drawbacks, however, are (1) the existence of the many molecular forms of the proteins which makes them so useful for variety identification; (2) the nature of the antibodies produced. A preparation of gliadins injected into a mouse or rabbit will cause the production of many different kinds of antibody, because of the presence of the different gliadin components and the selectivity of the immune system. Moreover, antibodies are produced by the B-lymphocyte cells of animals. Each cell synthesises a unique kind of antibody, which is different from those produced by other cells in the same animal. Thus the antibodies produced in an animal are an extremely heterogeneous collection of molecules, which will differ from one animal to the next. These multiple collections of antibodies, which are sometimes known as polyclonal preparations, will react with a wide range of gliadins and perhaps other proteins and do not allow much, or indeed any, discrimination between varieties. Methods must thus be sought for increasing the specificity of the antibody production. Fortunately, such methods are now in existence.

Monoclonal antibodies

Monoclonal antibodies (MAbs) may be defined simply as an antibody preparation which contains only a single type of antibody molecule. In the present case, this would be a preparation which consists of antibodies to a single gliadin or hordein protein, or a small group of proteins. This can be achieved by growing individual antibody producing cells from immunised animals in tissue culture and isolating the antibodies produced. This technique was developed in the mid 1970's, primarily by Milstein and his group in Cambridge (2), who were carrying out basic research on the functioning of the immune system. The existence of MAbs has led to the proliferation of medical diagnostic kits to detect and quantify specific proteins or hormones in body fluids. In contrast, applications within the agricultural world are relatively few. However, this situation is slowly being remedied.

The production of MAbs can now be considered to be a routine procedure, the major steps of which are outlined in Figure 3.1. The technique is nevertheless specialised and requires considerable expertise. There are a number of laboratories which undertake MAb production on a commercial basis. The essential point in the production is the fusion of antibody-producing (spleen) cells from an immunised animal with a myeloma cell (a myeloma is a malignant tumour of the immune system). The myeloma cell can be grown in culture and hence the fused spleen-myeloma cell (known as a hybridoma) will also grow and reproduce itself. Moreover it will produce antibodies, the nature of which is determined by the spleen cell and hence by the original immunisation procedure. The hydridomas must be screened, to select those which are producing antibodies of interest and the useful ones are allowed to grow and divide in culture (i.e. are 'cloned'). The antibodies which are secreted are then collected.

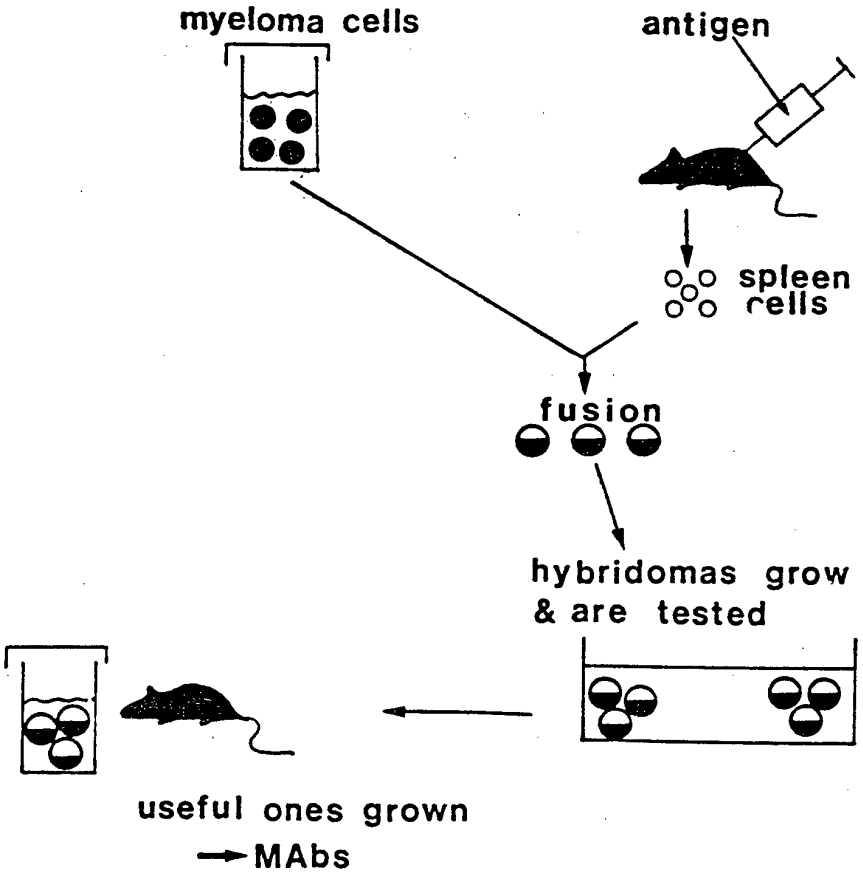
The advantages of monoclonal antibodies are that they offer a high and reproducible specificity, i.e. it is possible to prepare large quantities of an antibody which will always react with only one given antigen (protein or group of proteins). This allows the detection of particular proteins, using suitable types of assay. Because of the nature of antigen - antibody recognition, assays using MAbs are very sensitive, simple to carry out and give rapid results. The potential offered by this technology for cereal variety identification and quality analysis is now beginning to be examined.

Practical applications of monoclonal antibodies in cereal testing

The pioneers of the production of MAbs against cereal seed proteins are the Wheat Research Unit in Sydney, under the directorship of Dr Colin Wrigley. Workers there have successfully prepared a panel or library of MAbs, using gliadins as the antigens (9). An unexpected finding was that unimmunised mice produced antibodies which reacted to several gliadins. This is presumably due to dietary intake of wheat and hence of antigens. Notwithstanding that, the CSIRO group have reported a number of potential applications of the MAbs. For example, it was found that some MAbs would bind to proteins from grain

Figure 3.1

A schematic representation of the production of monoclonal antibodies. Antibody - producing (spleen) cells from an animal injected with the proteins (antigens) of interest are fused with myeloma cells. The hybridomas so formed will grow and produce antibodies of one specific type. Following cloning and screening, the useful ones are grown in culture or in animals. The resultant monoclonal antibodies are collected.



extracts from a wide range of cereal species, whereas others were more specific and would only interact with proteins from particular species. This gives rise to the possibility of several kinds of test, apart from being able to distinguish between species. For instance, MAbs which bind to wheat, barley and rye proteins, but not to oat and rice proteins, would be very useful for detecting gluten in food-stuffs and hence invaluable to sufferers from coeliac-toxin disease (who are allergic to gluten and must exclude it from their diet). Workers in Sydney have identified suitable MAbs (7) and developed and tested a quantitative assay procedure (5,6). Work is now in progress to develop this into a simplified version, as a commercially available kit (10). Other similar kinds of application include the detection of cereals other than barley in brewing mashes and the detection of bread wheat in flours intended for pasta production.

Particular effort is being directed towards the use of MAbs to detect quality types in cereals. Preliminary work appears encouraging. For example, sulphur deficiency in wheat grain can be a serious problem in Australia. It affects the quality of the grain, causing a loss of dough extensibility. This is thought to be due to an alteration in the relative proportions of certain types of gliadins. One of the panel of MAbs was shown to be a good indicator of this, as it was specific for β-gliadins, which are relatively high in sulphur (8). Pedigree-related quality attributes, i.e. those which are due to variety, will be more difficult to detect. However, certain MAbs have been identified which demonstrate varietal differences in their binding affinity and in some instances, these differences could be correlated with quality parameters such as dough resistance (10,12).

This work raises the possibility of using MAbs for variety identification purposes. The simpler genetic structure of barley may be advantageous in developing such tests, since there are fewer individual hordein than gliadin components and less molecular relationship between the components. Antibodies have been identified which bind to certain hordeins and enable some varietal distinctions to be made (1). Attempts are now being made to develop rapid spot-tests which would be suitable for screening large numbers of

samples using the appropriate MAbs. Because the differences in binding capacity are often quantitative rather than qualitative, it is necessary to have an assay which is capable of producing numerical results. The most likely form of assay is one in which the MAbs are absorbed onto a matrix, which could be either nitrocellulose 'paper' or a microtiter plate. Because of the need for quantitation, the MAb must be conjugated to an enzyme. Grain extracts are then applied and the paper or plate is treated with a substrate, which will not only react with the conjugated enzyme, but will also contain reagents which yield a final coloured product. The intensity of this coloured product would be proportional to the degree of binding between the grain extract and the particular MAb employed. These types of assay, which are usually known as ELISA - enzyme linked immuno-sorbent assays - would be very attractive, not least because considerable automation would be possible. Technology is now available which permits the loading of multiple solutions into the wells of micro-titer plates and automatic washing, processing and measurement of resultant colour intensities. This would allow the development of a rapid test, the possible form of which has been described (10). Each sample extract is added along one column (i.e. eight wells) of a microtiter plate. The plates are automatically loaded, with a different MAb-enzyme conjugate, specific for particular proteins, being placed in each of the eight wells. As the plates contain 96 wells in total, 12 samples can be processed at the same time. The plates with the grain extracts plus MAbs are washed automatically, and the enzyme substrate, plus colouring reagent, is added. After a suitable time the colour development is stopped and the absorbance value (colour intensity) of each well is measured, recorded automatically and ranked on a 1-to-10 scale. To evaluate the results, each variety would be coded as an 8-digit number, with each digit corresponding to the relative binding capacity of the proteins of that variety for the specific MAbs. Thus variety A might be 0-1-1-7-6-5-9-2, whereas variety B might be 0-1-1-0-6-5-9-2, indicating a difference in the protein composition between varieties A and B which is expressed as a difference in binding to MAb number four. The 8-digit 'codes' could be stored on computer for each authentic variety and then samples under test could be compared with the stored codes and hence identified.

The great attractions of such assays would be their speed and simplicity. It has been estimated, for example that one operator can analyse 20 microtiter plates per day, that is about 2000 individual assays (10). If MAbs of sufficient specificity were available, it might be possible to use a code with fewer than eight digits as a descriptor. However, this is the major limitation. Because of the antigenic similarity of many hordeins and gliadins, the MAbs available thus far exhibit a high degree of cross-reactivity, that is to say they will bind with several different proteins. Whilst these are usually all proteins of the same type, e.g. a particular group of β-gliadins, this cross-reactivity is probably sufficient to make efforts at definitive varietal discrimination and identification very difficult to achieve in practice. For instance, only two MAbs have so far been identified which demonstrate appreciable varietal differences in binding to gliadin-containing extracts (10).

Conclusions

Immunological methods in general, and especially the use of monoclonal antibodies raised to seed proteins, offer a potentially very attractive system for investigating several aspects of cereal science and technology. A recent review (11) highlights various instances within the malting and brewing industries, for example, where an immunological approach has proved useful. At the Carlsberg laboratory in Denmark, screening for high-lysine lines of barley is carried out using a very simple *in situ* assay, involving abrasion of grains and immersion into a fluorescent antibody reagent. Detection of antibody - antigen reaction is accomplished with a small portable viewer (4). However, such methods may not be suitable for varietal identification. The problems arise mainly from the lack of antibodies to specific gliadin or hordein proteins. It would perhaps be worthwhile investigating a slightly different approach to the manufacture of MAbs. Instead of using crude protein preparations, MAbs could be raised to individual protein components. By selecting gliadin or hordein bands which electrophoretically have been shown to be especially useful for varietal distinctions, it may be possible to build up a more specific library of MAbs. Even so, the extent of cross-reactivity would remain a problem. However, it must be

recognised that relatively few scientists have studied the use of MAbs for variety identification purposes and the technology is still at an early stage of development. There are many difficulties in producing a MAb-based identification system but the speed and simplicity of such a method are very attractive.

It may be more realistic to view MAbs as a means of rapidly screening seed lots for their general quality characteristics. The identification of MAbs where differences in binding correlate with quality parameters (10,12) makes this a viable proposition. Clearly more work is required, both with other MAbs and with other quality measurements, to examine further the feasibility of this approach. An additional factor in favour of this type of scheme would be the possibility of using the MAbs as a preliminary device and then using electrophoresis or HPLC to analyse the more interesting or difficult cases in greater detail. This would require the use of protein extraction techniques which are compatible both with ELISA - type assays and with electrophoretic or chromatographic separation. Fortunately, the use of 1M urea solutions, with or without a reducing agent, ought to allow for this.

A prime requirement in the whole area of MAb production and evaluation in cereals is for more laboratories to be involved. The Wheat Research Unit appears to have a virtual monopoly on research in this area at the moment, particularly as regards the varietal and quality implications. It would seem to be highly desirable to have some UK input into such research and development. Monoclonal antibodies may never be sufficiently specific as to enable them to compete with electrophoresis or HPLC in terms of discriminatory power for cereal varieties, but they may well be able to make a large contribution as a way of initially screening seed lots and determining quality categories.

References

(1) Burbidge, M.J., Batey, I.L., Campbell, W.P., Skerritt, J.H. and Wrigley, C.W. (1986). Distinction between barley varieties by grain characteristics, electrophoresis, chromatography and antibody reaction. Seed Science and Technology 14, 619-629.

(2) Kohler, G. and Milstein C. (1975). Continuous cultures of fused cells secreting antibody of pre-determined specificity. Nature 256, 495-497.

(3) Konarev, V.G., Gavrilyuk, I.P., Gubareva, N.K. and Choroshajlov, H.G. (1981). Electrophoretic and serological methods in seed testing. Seed Science and Technology 9, 807-817.

(4) Rasmussen, U. (1985). Immunological screening for specific protein content in barley seeds. Carlsberg Research Communications 50, 80-93.

(5) Skerritt, J.H. (1985). A sensitive monoclonal-antibody - based test for gluten detection : quantitative immunoassay. Journal of the Science of Food and Agriculture 36, 987-994.

(6) Skerritt, J.H. and Smith, R.A. (1985). A sensitive monoclonal-antibody - based test for gluten detection : studies with cooked or processed foods. Journal of the Science of Food and Agriculture 36, 980-986.

(7) Skerritt, J.H., Diment, J.A. and Wrigley, C.W. (1985). A sensitive monoclonal-antibody - based test for gluten detection : choice of primary and secondary antibodies. Journal of the Science of Food and Agriculture 36, 995-1003.

(8) Skerritt, J.H., Martinuzzi, O. and Wrigley, C.W. (1987). Monoclonal antibodies in agricultural testing : quantitation of specific wheat gliadins affected by sulfur deficiency. Canadian Journal of Plant Science 67, 121-129.

(9) Skerritt, J.H., Smith, R.A., Wrigley, C.W. and Underwood, P.A. (1984). Monoclonal antibodies to gliadin proteins used to examine cereal grain protein homologies. Journal of Cereal Science 2, 215-224.

(10) Skerritt, J.H., Wrigley, C.W. and Hill, A.S. (1988). Prospects for the use of monoclonal antibodies in the identification of cereal species, varieties and quality types. Proceedings of ISTA Symposium, Leningrad (in press).

(11) Vaag, P. and Munck, L. (1987). Immunochemical methods in cereal research and technology. Cereal Chemistry 64, 59-72.

(12) Wrigley, C.W., Batey, I.L., Campbell, W.P. and Skerritt, J.H. (1987). Complementing traditional methods of identifying cereal varieties with novel procedures. Seed Science and Technology 15, 679-688.

Chapter 4:

RESTRICTION FRAGMENT LENGTH POLYMORPHISM

Introduction

The effectiveness of electrophoresis, HPLC and monoclonal antibody tests for variety identification resides in the proximity of the character used - protein composition - to the primary genetic information (DNA). Because of this, environmental effects on the expression of the character are avoided. If there is genetic variation in protein composition, then it follows that there must be variation in the underlying DNA. Indeed, there may well be more variability, since not all alterations in the DNA (mutations) will be expressed as alterations in protein components. Thus a more powerfully discriminating system might be available, if variability in DNA could be detected in different varieties. This would represent the ultimate system for discriminating between varieties which must, by their very nature, differ in the genes which are expressed and hence in their DNA. The technical difficulty lies in finding methods which can be utilised to investigate DNA variations. To understand how this can be done, it is necessary to have some basic understanding of the structure and function of DNA.

The structure of DNA

The recognition that DNA (deoxyribonucleic acid) is concerned with the transmission of genetic information undoubtedly represents the single most important advance in biology of the past 40 years. Along with the elucidation of the structure of DNA, it allowed scientists to explain how hereditary factors (genes) could be transferred from one generation to the next. This led to the creation of a new branch of science, molecular biology, and laid the foundations for the whole 'genetic engineering' revolution which has occurred over the last few years. The story of the Watson-Crick 'double helix' model for the structure of DNA is outside of the scope of this article (11,12). However, it is important to know that DNA consists of two strands, wound into a helix and that each strand in the double helix is composed of a linear polymer of alternating sugar and phosphate groups. Attached to each sugar is a

purine or pyrimidine nucleotide base. The bases 'pair' with the appropriate base in the opposite strand of the helix. Four bases are found in DNA - thymine (T), adenine (A), cytosine (C) and guanine (G). The base pairing is specific, such that T is always found with A and C with G. This specificity is important in that it allows exact copies of DNA molecules to be made from each strand of the helix. Genes can be considered to consist of segments of DNA, which contain a defined number of nucleotide bases in a specific sequence. These gene regions, which might be several hundred bases in length, will have varying functions, but one of the most important is that they control the production of particular proteins. The sequence of amino acids in a protein is determined genetically by the sequence of bases in the DNA. Given that there are only four bases (A,C,G,T), it is not readily apparent how they can code for the 20 or so different amino acids that can occur in proteins. The answer is that sequences of three bases, known as triplets or codons, are necessary to determine the synthesis of a single amino acid. This gives more than enough combinations (4^3=64) to cope with 20 amino acids. A major achievement of the late 1970's in biochemistry was the deciphering of the genetic code, such that it is now established which sequences of three nucleotide bases code for individual amino acids. Thus the codon for glycine is G-G-G, whilst for arginine it is A-G-A and for serine A-G-C. Several features of the code are of interest. For example, it is known to be degenerate, in that there is more than one triplet for each amino acid. Glycine can be coded by G-G-G, G-G-A, G-G-C or G-G-T for instance. Certain triplets do not code for amino acids, but rather signal the beginning or end of protein chains.

Thus anything which alters the sequence of nucleotide bases in a DNA chain, e.g. a mutation or inversion etc., could affect the amino acid composition of a given protein. Depending on the nature of the alteration, the protein may still be functional. Again, if the molecular weight or charge of the protein is altered, it may be possible to detect a difference between the normal and mutant proteins by electrophoresis or HPLC. It is usually assumed that cereal storage proteins have undergone many mutations but, since their function as reservoirs of nutrients for the growing seed will not be generally affected, the mutations can be tolerated. This accounts for the high

degree of polymorphism of gliadins and hordeins. Methods which can uncover the variations in DNA rather than proteins may demonstrate even further polymorphisms. Recent progress and understanding of the techniques of gene manipulation enable this variation to be investigated.

Restriction enzymes

One way of detecting differences in the DNA of different individuals would be to determine the sequence of nucleotide bases. However, DNA molecules have extremely high molecular weights (of the order 10^9) and this would represent an unrealistic task. In order to assist in the sequencing, it is necessary to cut the DNA into smaller fragments. This was not possible until 1970 when an enzyme was recognised and purified which would achieve this. Such enzymes are called restriction endonucleases, or more usually restriction enzymes. These enzymes will restrict (cut) DNA molecules at specific points along the nucleotide base chain, depending on the precise sequence of bases. For instance, the enzyme Eco RI will only act between the G and A bases in a sequence consisting of GAATTC. The enzyme Hind III on the other hand restricts between two adenine bases in the sequence AAGCTT. Incidentally, the names of restriction enzymes are derived from the organism from which they were first isolated. Thus in the examples above, Eco RI was obtained from Escherichia coli RY13 and Hind III from Haemophilus influenzae Rd. There are now over 400 restriction enzymes known. Not all have unique specifities, but even so, there are considerable opportunities for cutting DNA into more manageable pieces at precisely known points. It is by using the fragments of DNA produced (restriction fragments) that variation at the level of DNA can be uncovered and utilised.

Restriction fragment length polymorphism

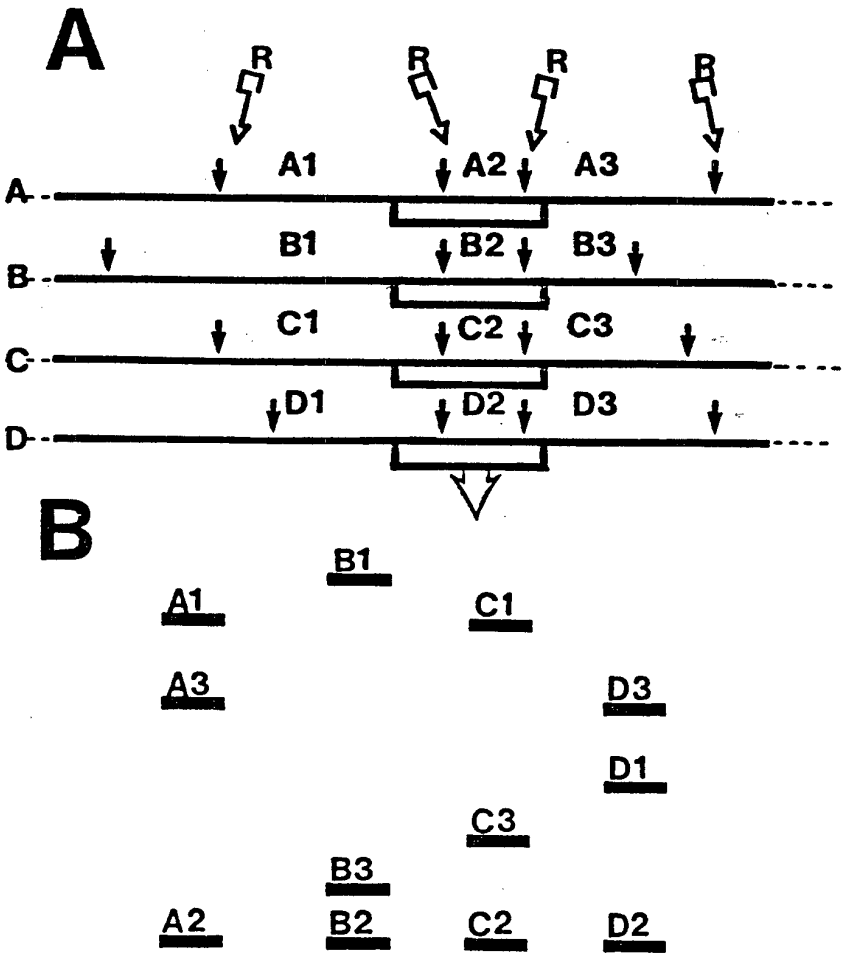
At the simplest level, restriction fragment length polymorphisms (RFLP's) can be seen as variations in the size of restriction enzyme - generated fragments of DNA. The variation is detected by separating the fragments on an electrophoretic gel. To understand this more clearly, an example is useful. Figure 4.1A is a diagrammatic representation of part of the DNA from four inbred lines of a crop species (e.g. varieties of wheat or barley). It is assumed that the sequence of bases within this part of the DNA molecule is different for all four lines, such that when they are restricted with a particular enzyme (R) (arrowed) fragments of varying sizes are produced. The boxed part of the DNA molecule can be taken to be a gene. The restriction site within this gene is identical in all four lines, whilst outside of the gene the sites vary. DNA sequence changes (mutations) within a gene are not likely to be beneficial, whilst changes outside, in the so-called spacer regions, are functionally neutral and can be tolerated. All higher organisms appear to possess these spacer regions (11). The natural variation between lines in these regions is exploited by RFLP technology, although variability of restriction sites within genes also occurs and can be used. Following restriction, the DNA fragments from each line are separated by gel electrophoresis, to give the patterns depicted in Figure 4.1B. In this example, it could be said that the four lines A, B, C and D were clearly different, in that they demonstrated polymorphism for the restriction site specific for enzyme R.

One difficulty with this approach is that in reality, the gel patterns are very much more complex, since hundreds if not thousands of DNA fragments will be produced following restriction. This means that clearly separated bands are not produced and a way has to be found of highlighting bands of interest. The solution is to use probes of various kinds. Probes for use in RFLP analysis are pieces of DNA which are labelled (usually radioactively) in vitro. It is not necessary to know the precise details of how probes are prepared and evaluated. In general terms, DNA is extracted from an organism, digested and the fragments cloned (multiplied) in a bacterial vector. The cloned fragments are screened, to detect those which will recognise and

Figure 4.1

A A restriction map of part of the genome of four individuals (eg wheat or barley varieties), differing in their spacer region restriction sites specific for the enzyme R.

B Gel patterns for the DNA fragments produced from restriction of the individuals shown in A.



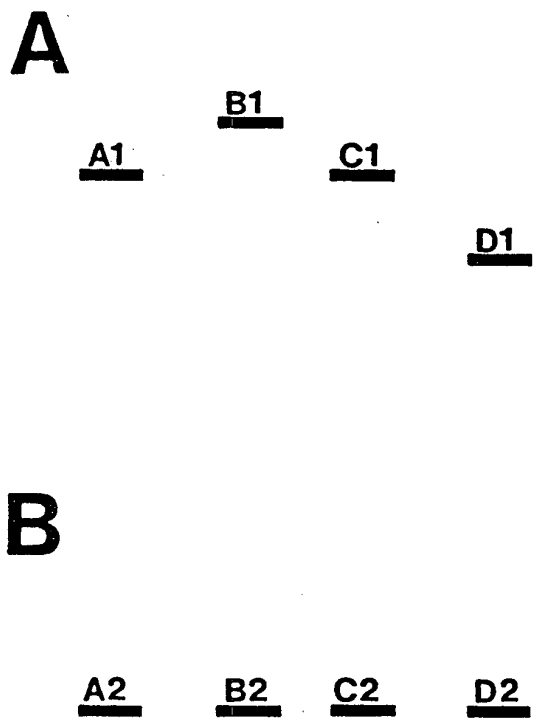
hybridise with DNA fragments produced by the restriction process. Any unique DNA sequence is useful, and it is not necessary to isolate specific areas of DNA or genes. The probe binds to the restricted DNA because of the specific base-pairing (A-T, C-G) inherent in DNA structure. For this to happen, the DNA on the gel has to be denatured, i.e. separated into its two complementary strands. This is achieved *in situ* by immersion of the gel in a basic solution. If the gel shown in Figure 4.1B were probed with a fragment equivalent to A1, then the gel pattern shown in Figure 4.2A would result. On the other hand, if it was probed with fragment A2, the pattern shown in 4.2B would result. This indicates that there is detectable polymorphism at the A1 region of the chromosome, but not at A2. The great power of this methodology lies in (1) the large number of different restriction enzymes available, and (2) the virtually unlimited number of probes. If no variation between lines or varieties is demonstrated by one particular combination of restriction enzyme and probe, it is only necessary to change the enzyme, or use a different probe. In practice, the restricted DNA fragments are transferred (or blotted) to a nylon membrane before being probed. The nature of this membrane is such that once analysed, the probe can be removed from the membrane by washing. The membrane blot can thus be used for repeated analysis with different probes.

Applications of RFLP analysis

In many respects, the use of RFLP analysis can be likened to the use of protein electrophoresis. In either case, two individuals (which could be wheat or barley varieties) are shown to be different by the existence of unique banding patterns on a gel. RFLP's also share many of the advantages of electrophoresis, such as freedom from environmental influence. In addition, any suitable part of a plant (endosperm, embryo, leaf tissue etc.) could be analysed. However, the outstanding benefit of the RFLP technique is that it offers an essentially unlimited number of genetic markers, merely dependent on the availability of the correct combination of restriction enzyme and probe. These markers must exist, if the individuals are genetically distinct. This means that the power of RFLP analysis to demonstrate distinctness or prove the identity of samples (varieties) will inevitably be much greater, even than the

Figure 4.2 A The result of probing the gels shown in Figure 4.1B with a probe specific for the A1 region.

B As above, but using a probe specific for the A2 region.



best protein electrophoretic methods. Thus it ought to be possible to apply RFLP technology to any of the situations where electrophoresis or HPLC of proteins is currently used. The relative novelty of the techniques means that in fact they have not yet been used in these ways. The major effort appears to be being made by plant breeders. This is primarily because RFLP's offer breeders the opportunity of covering entire genomes with molecular markers, as opposed to the rather limited coverage available with protein, isozyme or other markers. It is thus possible for a breeder to have a very detailed map of the genes of an organism. Apart from being useful in labelling and following traits such as disease resistance genes, this enables the breeder to mark the individual component genes of complex, multi-genic, environmentally influenced characters (e.g. yield). The use of RFLP's in selecting for these so-called quantitative traits has been particularly strongly advocated by two Israeli scientists Beckmann and Soller, who have also been prominent in proposing other potential applications of RFLP's, such as variety identification and patent protection (3,4,6,10). The genomes of wheat and barley are currently being mapped by research teams in the United Kingdom using RFLP's. This may well make available (no doubt on a commercial basis) suitable probes, so that organisations such as the NIAB and grain processors (millers, maltsters) could assess the possibilities for varietal identification on a routine basis.

There can be no question that the use of RFLP's would represent an immensely powerful tool for cereal variety identification. Little direct work has been published on this topic. However, Shewry and colleagues at Rothamsted have demonstrated the presence of extensive polymorphism in DNA restriction fragments related to B- and C- hordein genes (9). They have also outlined how RFLP's could be used for barley variety identification (5). This work indicated the power of RFLP analysis, in that two varieties, which were identical when analysed by both one and two-dimensional electrophoretic analysis of hordeins, could be distinguished. In wheat, polymorphism for the spacer region between regions of DNA coding for ribosomal proteins has been demonstrated in different varieties (1,8). Also, restriction fragment variation in the gene family encoding high molecular weight glutenin sub-units has been reported (7). Whilst these are indications of the potential for variety identification, no direct applications were suggested. It seems clear,

however, that the use of RFLP's for variety identification will come about, perhaps in the medium-to-long term rather than immediately. The fact that breeders are actively utilising the methodology at the moment increases the probability that they will suggest that it could be incorporated into statutory distinctness testing. It would seem sensible for some basic research and feasibility studies to be made in advance of this eventuality.

Inevitably there are some disadvantages in applying RFLP's for variety identification. The major drawbacks at the moment are the cost and the fairly sophisticated nature of the analysis. The preparation and evaluation of probes can be problematic and is a lengthy process. However, having once identified a useful probe or set of probes, their usage is unrestricted and they can be readily multiplied. There are companies who will undertake contract probe preparation, but the costs are very high. As an example, a company in the USA was offering a library of probes for use in maize at a fee of some $2 million (c. £1.1 million). This may be an extreme instance, but is indicative of the sort of fees involved. At least two companies have filed patents, in the field of RFLP technology, particularly as applied to proving the co-identity of two samples. It is unclear as to whether or not these patents will be upheld, but their existence demonstrates the interest which these methods have aroused, both scientific and commercial. In the publicity material of one of these companies (which is British based), RFLP analysis of plant varieties is specifically mentioned as one of the areas in which they see applications, both for establishing co-identity of samples and for patenting purposes. It remains to be seen what the scale of charges would be for, say, a barley variety identification. It can be reasonably expected that the cost of the analysis will fall, as the methods become simplified and more amenable to routine application. It is interesting to note that Cellmark, the ICI subsidiary which carries out the DNA 'finger-printing' of humans for forensic and other purposes, charge about £120 per sample. The costs of RFLP analysis for a range of breeding and other purposes have been calculated (2). For variety identification, looking at five individuals per sample, the cost was reckoned to be about $110 (c. £60).

Another possible disadvantage is that the degree of discrimination offered by RFLP's may be too great for routine variety identification purposes. Given a sufficient number of probes, it would undoubtedly be possible to discriminate between all varieties. The problem is that this may well also reveal a high degree of variation within varieties, comparable to (but more widespread than) the existence of electrophoretic biotypes (8). Unfortunately, this cannot be known without evaluating empirically a range of combinations of restriction enzyme and probe. It should be possible to minimise this difficulty by a sensible choice of DNA probe. Instead of using probes which are highly specific and might only recognise a single gene, it would be preferable to utilise probes that are moderately repeated throughout the genome, i.e. would produce several bands on the gel of the restricted DNA. This would help to ensure that the gel patterns produced were more likely to be uniform for a given variety. It might also reduce the number of successive probes necessary in order to elicit complete identification of a sample. An example of this kind of restriction fragment/probe combination can be found in the work of Shewry and co-workers (5).

If RFLP's are used for statutory distinctness testing purposes, then account would have to be taken of the opinions of the plant breeding community. Several objections have been raised previously against the use of protein electrophoresis for distinctness testing, mainly concerning the extra burden which would be placed on breeders in ensuring that their new varieties were uniform for gliadin or hordein composition. This same problem would exist if RFLP analysis were used. However, for breeders already utilising RFLP technology for selection purposes, this would presumably not be insurmountable. A further potential difficulty might be the existence of varieties with identical morphological appearance which differ only in their RFLP patterns. Intra-varietal heterogeneity in RFLP pattern is another possible area of concern. A corollary to the use of RFLP's to establish distinctness is that it would become increasingly difficult for companies to keep the pedigree of their varieties confidential, since much breeding history could be deduced. Such factors are clearly important and need to be discussed, but it is surely reasonable to hope that the legal and technical difficulties are resolved so that full use can be made of

methodology which may well prove invaluable for those involved in distinctness testing.

Conclusions

It seems certain that the application of RFLP and related technologies to plant breeding is going to become increasingly widespread and perhaps even revolutionise the way that breeders operate. Whether or not the methods become suitable for use in routine cereal variety identification is less certain, but the potential is evident. Improvements and refinements to the methodology will make it easier and quicker, which would obviously increase the prospects for routine use and unit costs will decrease. Given this, RFLP techniques could be used in all of the areas where protein electrophoresis is currently utilised or has been advocated, e.g. distinctness testing, seed production and certification, hybrid purity evaluation and quality control in industry. The great advantage would be the ability to distinguish uniquely every variety, given the appropriate combination of restriction enzyme and DNA probes. Hardly any work has been carried out specifically to assess the potential and problems offered by RFLP analysis for cereal variety identification and this is an area where future research effort should be targeted. Only then will it become evident whether or not this particular sector of the biotechnology industry can effectively translate undoubted potential into a practical system of benefit to the agricultural community.

References

(1) Appels, R. and Dvorak, J. (1982). The wheat ribosomal DNA spacer region : its structure and variation in populations and among species. Theoretical and Applied Genetics 63, 338-348.

(2) Beckmann, J.S. and Soller, M. (1983). Restriction fragment length polymorphisms in genetic improvement : methodologies, mapping and costs. Theoretical and Applied Genetics 67, 35-43.

(3) Beckmann, J.S. and Soller, M. (1986). Restriction fragment length polymorphisms and genetic improvement of agricultural species. Euphytica 35, 111-124.

(4) Beckmann, J.S. and Soller, M. (1986). Restriction fragment length polymorphisms in plant genetic improvement. in Miflin B.J. (ed). Oxford Surveys of Plant Molecular and Cell Biology 3, 196-250.

(5) Bunce, N.A.C., Forde, B.G., Kreis, M. and Shewry, P.R. (1986). DNA restriction fragment length polymorphism at hordein loci : application to identifying and fingerprinting barley cultivars. Seed Science and Technology 14, 419-429.

(6) Burr, B., Evola, S.V., Burr, F.A. and Beckmann, J.S. (1983). The application of restriction fragment length polymorphism to plant breeding. in Setlow, J.K. and Hollaender, A. (eds). Genetic Engineering, Principles and Methods, Academic Press, New York, pp 45-59.

(7) Harberd, N.P., Bartels, D. and Thompson, R.D. (1986). DNA restriction fragment variation in the gene family encoding high molecular weight (HMW) glutenin sub-units of wheat. Biochemical Genetics 24, 579-596.

(8) May, C.E. and Appels, R. (1987). Variability and genetics of spacer DNA sequences between the ribosomal - RNA genes of hexaploid wheat (Triticum aestivum). Theoretical and Applied Genetics 74, 617-624.

(9) Shewry, P.R., Bunce, N.A.C., Kreis, M. and Forde, B.G. (1985). Polymorphism at the Hor 1 locus of barley (Hordeum vulgare L.). Biochemical Genetics 23, 389-402.

(10) Soller, M. and Beckmann, J.S. (1983). Genetic polymorphism in varietal identification and genetic improvement. Theoretical and Applied Genetics 67, 25-23.

(11) Watson, J.D. (1970). Molecular Biology of the Gene. W.A. Benjamin Inc., California.

(12) Watson, J.D. and Crick, F.H.C. (1953). Genetical implications of the structure of deoxyribonucleic acid. Nature 171, 964-967.

Chapter 5:

THE USE OF MACHINE VISION

Introduction

In the preceding sections of this article, the use of a range of biochemical techniques to detect underlying genetic variation and hence identify cereal varieties has been considered. These methods represent what can be termed a reductionist approach, getting ever closer to defining varieties in terms of differences in DNA composition. Classical taxonomy in general works in a different direction, with the accurate and detailed morphological description of a wide range of characters forming the basis for systems of classification and differentiation. The extension of this approach to the study of plant varieties has provided the mainstay of distinctness testing and seed certification schemes. Although very satisfactory in practice, this inevitably produces problems. As many morphological characters are continously expressed (quantitative), replicated measurements are required in order to apply statistical means of establishing distinctness. Environmental effects on the expression of various characters again require that considerable replication of sampling and examination is necessary. Usually this requires time-consuming and labour-intensive manual methods of measuring, recording and processing the information. An alternative to this is to make use of machine vision systems. Simply, machine vision refers to the acquisition of data (shape, size etc.) via a video camera or similar system and the subsequent computer analysis of these data following suitable processing. The term image analysis is also used in this context, but more strictly refers to the extraction of numerical data from an acquired image. Apart from providing an automated means of obtaining measurements, the great attraction of machine vision is that it is possible, via the appropriate computer software, to make very detailed comparisons of sets of data, i.e. to operate pattern recognition systems. In recent years, the decreasing costs and increasing sophistication of equipment for machine vision and image analysis have produced a surge of applications, particularly in the biomedical field (anatomy and cytology for instance), but also

in the interpretation of aerial and satellite photographs and in industrial robotics and artificial intelligence systems. The potential of this technology in seeds and plant variety work has now begun to be examined, particularly in the area of cereal grain inspection.

The basic elements of a machine vision system should be capable of the following operations (2):

a) image capture, via a video camera or other electronic system (e.g. charge coupled device);
b) analogue to digital conversion of the image data (this is not necessary if a charge coupled device is used);
c) computerised manipulation of the image data in order to obtain a version suitable for analysis (this might include the removal, via the software, of extraneous and anomolous objects, or 'grey-level partition' to achieve a binary image or silhouette). This is termed image processing;
d) image analysis - the extraction of information from the processed image;
e) pattern recognition, either statistical or syntactic, to sort and compare objects (e.g. cereal varieties). In statistical systems, the results of the image analysis are analysed statistically, whilst in syntactic pattern recognition, the image itself is searched for particular features e.g. angles in close proximity to one another;
f) computerised decision-making and presentation of results, including assessments of significance.

It is also desirable, in a useful practical system, to have an automated or robotic mechanism for movement of the sample (or the camera), so that operator intervention is minimised.

Practical applications

The prospect of using instrumental methods to carry out many of the direct observations made by trained laboratory staff in assessing various aspects of seed quality prompted the first major examinations of machine vision by workers at the Official Seed Testing Station in Cambridge. Using a low-cost image capture and processing system, Draper and Travis demonstrated that it was possible to analyse the shape of seeds and of various vegetative structures (3). The measurements made were relatively simple, comprising length, width, area and perimeter, which enabled two further variates to be derived, namely aspect ratio (width/length) and shape factor (or thinness ratio) (4π.area/perimeter^2). They further showed using seeds of 49 different crop and weed species, that most species could be differentiated on the basis of the shape factor in combination with seed length (8). This clearly indicated the taxonomic potential of the machine vision approach and encouraged the extension of the work to variety identification.

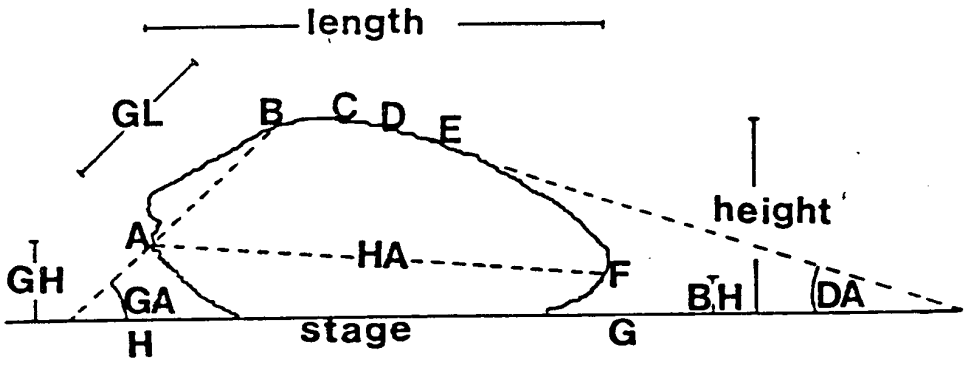
The results of a preliminary study using five wheat varieties illustrated how such identification could be achieved (4). For this work, a rather more sophisticated system was required, comprising a Cambridge Instruments Quantimet 10 (Q10) image analyser and custom - written computer software. Individual seeds were placed crease side down on a horizontal surface, with the longitudinal axis of the seed running parallel to the surface of the camera lens and with the embryo to the left. The seeds were viewed in side elevation using transmitted light. Hence a binary image of the seed and its support was recorded. The support silhouette was removed using the computer software, leaving the image of the seed. The parameters measured are illustrated in Figure 5.1 and are listed, along with the derived variates in Table 5.1. In order to obtain sufficient data for statistical evaluation, ten lots of each of the five varieties were used and two samples of 30 seeds taken from each lot. One sample was taken at random, whereas the other was selected to contain only well-filled grain. Of the measured and derived parameters, seven were considered to be potentially useful for wheat seed taxonomy. These were relative germ height, relative brush height, dorsal tangent,

horizontal axis, high point, relative germ length and germ tangent. A parameter was considered to be useful for differentiating between varieties if there was little or no overlap in the range of values recorded for it between two or more varieties. The measurements made in general characterise the shape of a seed, rather than its absolute dimensions and so differences in seed size caused by environmental and other factors were minimised. The results showed that there was little consistent benefit in selecting samples of well-filled seed, although for two characters (relative brush height and relative germ length), separation between varieties was improved.

This work indicated that measurable and consistent differences occur between wheat varieties in various aspects of their morphology which can be evaluated by machine vision. The authors have now extended the range of varieties to include 22 winter and spring types and have found that about 90% of samples can be a given a correct varietal identity (Draper and Keefe, personal communication). Research is continuing into different methods of statistical evaluation which will improve the discrimination still further. In a refinement to the existing equipment, a custom-built sample presentation device, in the form of a motorised camera gantry controlled by computer, has been installed, to remove excessive operator interation in data acquisition (5).

The potential demonstrated by this research has been evident in work from other laboratories. Thus in the USA, Zayas and co-workers have used image analysis to discriminate between classes of wheat (hard red winter, hard red spring, etc.) and also between eight varieties (9,10). A statistical approach somewhat different from that used at Cambridge utilised two class models to assign an individual seed to its appropriate class (or variety). Different morphological parameters (both measured and derived) were also used. Correct assignment was achieved in 78-85% of cases for the eight varieties tested. At the University of Manitoba, researchers have tried to overcome the constraints imposed by the necessity for precise manual orientation of seeds prior to imaging and also to examine more complex admixtures (7). Initially, they were able to separate successfully mixtures of wheat, oats, barley and rye, with only approximately 1% of

Figure 5.1 Diagram of a wheat seed, illustrating the parameters used for shape description. (Re-drawn from Keefe and Draper (4)).



Germ height (GH): the vertical distance (A-H) of the lower edge of the scutellum above the stage.

Germ length (GL): the length of the scutellar region (A-B).

Germ angle (GA): the angle subtended by a line drawn through A-B to the horizontal.

Length: the total horizontal length of the seed including the embryo.

High point: the horizontal distance from A-C divided by the total length of the seed.

Dorsal angle (DA): as germ angle, but through points D-E.

Height: the total vertical height of the seed.

Brush height (BH): the vertical distance (F-G) of the brush above the stage.

Horizontal axis (HA): germ height/brush height.

over 1100 seeds being incorrectly classified. The samples were randomly orientated before measurement. Non-cereal admixtures (e.g. rapeseed) and weeds (wild oats) could be also be correctly classified. Later work extended the study to consider contrasting types of grain within the same cereal type, i.e. class and varietal identification (6). Although categorisation according to class was generally effective, discrimination of varieties within a class was inconclusive, with correct classification occurring in 96% of cases for some varieties and only 15% of cases for others. This was partly explicable by the parameters chosen for measurement and discrimination, which included some based on absolute size. Environmental factors influencing seed size would inevitably tend to confound the discrimination. This is also, of course, a problem which occurs when subjective human examination of seed is used to assess varietal identity. It is clear that careful selection of measured and derived parameters, as well as effective methods of statistical analysis, are required in order to extract the maximum benefit from machine vision technology.

Conclusions

Although the machine vision approach to cereal variety identification is only at an early stage of development when compared to electrophoresis or HPLC, it represents an area of considerable potential. The great advantages offered include speed and ease of operation, once a system has been established and the appropriate sets of data relating to varietal differences in seed morphology are available. It has been estimated that to measure 400 wheat seeds, using a computer program which moves the camera in a square grid pattern over objects randomly scattered on the sample stage, would take approximately 5-10 minutes. The execution time needed for software evalution of such data is difficult to predict, but might add a further 1-2 minutes (5). This leads to the prospect of having a method sufficiently rapid to be useful for quality control at the point of intake in mills or maltings. Indeed, a prototype version of an instrument designed for wheat variety identification is under construction (details are obtainable from Dr S R Draper, Official Seed Testing Station, NIAB, Cambridge, CB3 0LE). Another possible area of

application would be in examining the purity of cereal seed lots, in seed testing laboratories for instance. This would require an automated method of sample presentation, to enable large numbers of seeds to be examined as rapidly as possible. Adoption of such technology could lead to considerable savings in manpower requirements. The mechanised measurement of characters also raises the possibility of being able to utilise additional characters for distinctness testing or seed certification procedures. This is not, of course, confined to cereal seeds (2,5).

There are, however, certain disadvantages to this approach. One drawback is the probable high capital cost of a machine vision facility. This is inevitable, given the degree of development which is necessary in order to produce a useful and working system and is offset by the low operating costs. Perhaps more significant is the problem of dealing with mixtures of varieties. The great benefit of, e.g. electrophoresis of individual seeds, is that it allows not only identification, but also an assessment of varietal purity. This is vital in quality control operations and also in seed certification, for instance. As configured at present, image analysis of seeds would not be able to assess the purity of samples, or to indicate the identity of any varietal admixtures. This is clearly an area where further research is required.

To date, all of the research effort has been concentrated on bread wheat variety identification. However, there seems little doubt that other cereals such as barley, durum wheat, triticale, rye and perhaps oats could be successfully analysed. Also, maize and particularly rice varieties have been shown to possess differential and observable seed characteristics. There is thus a need for work to be undertaken on these other important cereals. It is probable that only physical and instrumental methods of analysis are suitable for factory gate analysis of cereal seed lots, where results are required within 10-15 minutes. The only other such approach to date appears to be the use of near infra-red reflectance spectroscopy (NIR) (1). Although reasonably rapid, this does require the samples to be milled and has only been studied so far using six varieties. Also, from experience of using NIR for protein and other measurements in cereals, it is

known that the calibration required for the instrument is absolutely crucial and can vary from season to season. It is unlikely that NIR can equal machine vision in terms of either speed or accuracy. The other potential applications of machine vision in distinctness testing and certification procedures would also not be possible with NIR. There thus appears to be no real alternative to the image analysis approach for instrumental cereal varietal identification.

References

(1) Bertrand, D., Robert, P. and Loisel, W. (1985). Identification of some wheat varieties by near infrared reflectance spectroscopy. Journal of the Science of Food and Agriculture 36, 1120-1124.

(2) Draper, S.R. and Keefe, P.D. (1988). Electrophoresis of seed storage proteins and whole-seed morphometry using machine vision: alternative or complementary methods for cultivar identification? Proceedings ISTA Symposium, Leningrad (in press).

(3) Draper, S.R. and Travis, A.J. (1984). Preliminary observations with a computer-based system for analysis of the shape of seeds and vegetative structures. Journal of the National Institute of Agricultural Botany 16, 387-395.

(4) Keefe, P.D. and Draper, S.R. (1986). The measurement of new characters for cultivar identification in wheat using machine vision. Seed Science and Technology 14, 715-724.

(5) Keefe, P.D. and Draper, S.R. (1988). An automated machine vision system for the morphometry of new cultivars and plant genebank accessions. Plant Varieties and Seeds 1, 1-11.

(6) Newman, M., Sapirstein, H.D., Shwedyk, E. and Bushuk, W. (1987). Discrimination of wheat class and variety by digital image analysis of whole grain samples. Journal of Cereal Science 6, 125-132.

(7) Sapirstein, H.D., Neuman, M., Wright, E.H., Shwedyk, E. and Bushuk, W. (1987). An instrumental system for cereal grain classification using digital image analysis. Journal of Cereal Science 6, 3-14.

(8) Travis, A.J. and Draper, S.R. (1985). A computer-based system for the recognition of seed shape. Seed Science and Technology 13, 813-820.

(9) Zayas, I., Lai, F.S. and Pomeranz, Y. (1986). Discrimination between wheat class and varieties by image analysis. Cereal Chemistry 63, 52-56.

(10) Zayas, I., Pomeranz, Y. and Lai, F.S. (1985). Discrimination between Arthur and Arkan wheats by image analysis. Cereal Chemistry 62, 478-480.

TABLE 5.1 **Measured and Derived Variates used for the Description of Wheat Seed Shape**

Measured variates

A - area
L - seed length
H - seed height
P - perimeter
GH - germ height
GL - germ length
GT - the tangent of the germ angle
HP - high point
BH - brush height
DT - the tangent of the dorsal angle

Derived variates

SF - shape factor - $4\pi.A/P^2$
AR - aspect ratio - H/L
RGH - relative germ height - GH/H
RGL - relative germ length - GL/L
RBH - relative brush height - BH/H
HA - horizontal axis - GH/BH

From Keefe and Draper (4).

CONCLUDING SUMMARY

This review has clearly indicated that modern biochemical and computerised instrumental methods - 'New Technology' - can provide a very effective means of cereal variety identification. The new methods offer some considerable advances over more traditional approaches, effective though these have been. For instance, visual inspection of grain is a useful way of identifying at least some barley varieties. It does rely heavily, though, on the experience of the observer and the availability of reference books. For wheat varieties, there are severe limitations to what can be assessed by eye. Field-based growing-on of seed lots will provide a categorical determination of variety identity and purity, but is a lengthy process and requires large areas of land. By contrast, the modern methods can be considered to be rapid. Each of the methods considered has its own advantages and disadvantages. The Table attempts to summarise some of the principal attributes of the new methods and to compare various aspects of their application.

Quality control

Electrophoresis of enzymes and proteins seems likely to remain the method of choice for variety identification, at least in the short to medium term. Electrophoresis is now well established as a quality control test in the seed trade and combines relatively low capital and operating costs with a high degree of discriminatory power. The vast amount of background research which has been carried out on various approaches and classifications of varieties ensures that electrophoresis will not be lightly discarded. The versatility of the techniques and the wide range of species which have been examined are also advantages. By comparison, the other methods are in relatively early stages of development. Nevertheless, it is evident that HPLC, for example, could be useful, not only in its own right, but also as a complement to electrophoresis. The advantages offered by the accurate quantification of protein peaks should not be under-estimated, since this provides an extra dimension for variety characterisation. The possibilities for automation, both of sampling and data storage and evaluation are also interesting. It is probable that the

TABLE A Comparison of the 'New Technology' Methods for Cereal Variety Identification

Feature	Method				
	EP	HPLC	MAb	RFLP	MV
Freedom from environmental interaction	xxxx	xxxx	xxxx	xxxxx	xx
Rapid results	xxx	xx	xxx	x	xxxxx
Low operator interaction	xx	xx	xxx	x	xxxx
Low operator skill required	xx	xx	xxxx	x	xxxxx
Low capital costs	xxx	xx	xxx	x	x
Low operating costs	xxx	xx	xxx	x	xxxxx
Suitability for assigning varietal identity to a seed lot	xxxx	xxxx	xx	xxxxx	xxx
Suitability for estimating varietal purity	xxxx	xxxx	xx	xxxxx	x
Suitability for genotype classification (e.g. distinctness testing)	xxxx	xxxx	xx	xxxxx	xxxx

EP = electrophoresis; HPLC = high performance liquid chromatography; MAb = use of monoclonal antibodies; RFLP = restriction fragment length polymorphisms; MV = machine vision.

A scale of 1-5 crosses is used : more crosses indicates that the feature is exhibited to a higher degree.

immunological approach, i.e. the use of MAbs, will never be sufficiently discriminatory for variety identification purposes, due to the cross-reactivity of the antibodies to cereal storage proteins. However, these techniques may be very useful for separating seed lots into quality categories, which could be all that is required in some circumstances. On the other hand, there will probably always be a necessity for variety identification since certain varieties need to be positively excluded from processing. Also, it is not always desirable, for instance, to malt certain varieties of barley together even though they may share similar quality characteristics.

Ultimately, the separation of varieties on the basis of their DNA composition will provide a definitive method of identification. There is enormous potential in the utilisation of RFLP analysis. The challenge is to translate this potential into a useful, routine system, which may well prove to be very difficult. The fact that plant breeders are already actively using RFLP techniques may hasten the development of suitable routine procedures. A possible difficulty may arise in that the application of RFLP's is actually too discriminating and reveals extensive and wide-spread intra-varietal heterogeneity.

It may be that the only method which is feasible for quality control use at the point of intake in mills or maltings is machine vision. Clearly, there is considerable research needed in this area before it could be unequivocally recommended, but the potential undoubtedly exists. On the other hand, if retrospective testing is satisfactory (seed lots are delivered and then analysed, rather than analysed before delivery is agreed or permitted), then the speed of analysis becomes less crucial and the discriminating power, accuracy of results and ability to identify admixtures are more critical. On these grounds, electrophoresis (or HPLC, or, ultimately, RFLP analysis) is almost always going to be a better choice, provided that sufficiently skilled laboratory staff are available. Although efforts are being made by the manufacturers of electrophoresis equipment to simplify the technology (e.g. Pharmacia have launched a programmable, micro-chip controlled system, with pre-made gels and buffers), electrophoretic analysis is inevitably going to require a higher degree of operator

skill than the use of machine vision equipment. The interpretation of electrophoresis data is also an area requiring highly trained staff. This applies equally to the use of HPLC and, most definitely, to RFLP analysis.

A further consideration must be how flexible the methods are in their ability to cope with changes in the type of grain being traded. These changes can be either short or long term. As an example of short-term changes, this harvest year (1987-88) has seen the UK importing quantities of German milling wheat and also Spanish wheats. A similar situation has arisen with imported French and Irish malting barley varieties. For variety identification by electrophoresis (or HPLC), this presents no real problem, as long as the appropriate authentic reference samples are available (which is not always the case). However, difficulties would be encountered were machine vision analysis being employed, since the data base available would probably not contain information on these imported varieties. Moreover, the information would not be easy to obtain, as different seed lots grown at various locations are required. Long-term changes might include the increased availability of hybrid cereals, especially F1 (or F2) hybrid wheat. Although this would necessitate considerable changes in processing technology, it also creates problems for variety identification. F1 hybrid purity can only realistically be assessed by electrophoresis or HPLC at the moment, although hybrid identity in itself could presumably be readily carried out via machine vision. On the other hand, seed processors would not be receiving F1 hybrid grain, but rather the produce of its self-pollination, that is F2 grain. As this represents, genetically, a maximally segregating generation, it is very heterogeneous and therefore difficult to characterise. Electrophoresis analysis can provide some limited information as to varietal identity and purity, but machine vision would not be particularly useful, at least in its present form where purity analysis is not possible. This situation is not altered if the F2 hybrid seed is the marketed grade, in which case processors would be seeing the F3 generation. None of the techniques considered in this review is ideally suited for this eventuality and clearly millers and maltsters would have to do some hard thinking about how to cope with hybrid grain lots from a quality control viewpoint. It may be

that the separation of quality types using MAb's would be especially valuable here.

It is thus not possible to describe any one method as the 'best' for cereal variety identification in the quality control context. The various features outlined above and in the Table must be balanced against one another, and the final choice will depend on the precise needs and nature of the application.

Other applications

Many of the above arguments apply equally to other possible areas of application, such as statutory distinctness testing of varieties and seed certification. Protein electrophoresis is probably the method most likely to be usefully incorporated into such schemes, although there has been (and is continuing) considerable discussion between UPOV, the testing authorities and the plant breeders as to the best way of utilising electrophoresis. There would seem little problem about its use in seed certification, as a method of confirming varietal identity (and perhaps purity) rapidly and cheaply. The existence of standard reference methods of analysis of cereal varieties is undoubtedly important in this respect. Once electrophoresis is fully integrated into distinctness testing and certification, there could be few objections to the use of HPLC, or even RFLP analysis, although again careful discussions amongst the interested parties would be required. The question of the uniformity of varieties would need to be considered, as would the problem of 'minimum distances' between varieties. As technology allows better discrimination between varieties, based (ultimately) on differences in DNA sequences, so varieties may become ever closer to one another genetically and perhaps agronomically. It would be of little practical advantage to have a National List of wheat varieties which, whilst displaying distinctness due to the differences in RFLP's, were similar or even identical in terms of yield and performance. However, the fact that many breeders are already using RFLP analysis in their selection programmes and that they would not have to concern themselves with incorporating characters into their varieties solely

in order to meet the distinctness testing criteria, make this type of approach potentially very attractive.

Machine vision measurement of shape also has evident applications in distinctness testing. It not only enables existing morphological characters to be assessed more accurately and reproducibly, but it also provides possible new descriptors to enhance levels of discrimination. As this is in effect merely a different way of approaching more 'classical' taxonomy, it may be easier for UPOV and others to accept its incorporation into official systems.

A further area in which the 'Modern Methods' will be important is in the assessment of germplasm resources in gene-banks and other collections. A combination of machine vision, to describe seed (or vegetative) shape and electrophoresis (or RFLP analysis) to describe genotype, would provide a powerful way of cataloguing such valuable material. This could be especially important in developing countries, where no central seed collection may exist and control of authentic plant material resides with the individual breeder.

This review has indicated the many advantages of the different modern methods for cereal variety identification. It is clear that considerable progress has been made over the past 20 years, and that this progress is still continuing at a very rapid rate. We are now able to describe varieties in terms of their biochemical composition (i.e. by electrophoresis and HPLC of proteins) and will soon be able to define them strictly in terms of their genetic material (RFLP analysis of DNA). At the same time, the accurate description of shape by automated machine vision provides more detailed information as to varietal morphology. Nevertheless, there are areas where further research effort is required. These can be summarised as follows:

1) commercially available gels for electrophoresis, prepared according to a well-defined protocol (e.g. the ISTA standard reference method);

2) a system for the automated interpretation of electrophoresis gels, pattern matching and identification;

3) more stream-lined electrophoresis procedures (e.g. faster extraction, gel running and staining);

4) other methods of electrophoresis, to allow enhanced levels of discrimination (e.g. between varieties identical following PAGE of gliadins or hordeins);

5) evaluation of the use of HPLC of flour samples as a means of estimating varietal purity;

6) progress towards the standardisation of HPLC columns and protocols;

7) evaluation of the concept and identity of key 'quality proteins' and their detection using monoclonal antibodies (and also more UK involvement in this whole area);

8) progress towards simplified procedures for RFLP analysis and detection;

9) availability of DNA probes for RFLP analysis, suitable for varietal identification purposes;

10) evaluation of varieties of other cereal species using machine vision;

11) automated (robotic) procedures for sample presentation and image data collection for machine vision analysis;

12) evaluation of random sample presentation devices for variety identification by machine vision.

Thus there are evidently several areas where the available technology needs to be refined and improved before the cereal seed industry in general can be said to have extracted the maximum benefit from modern methods of variety identification. As these methods offer considerable potential savings in terms of speed, discriminating power and cost-effectiveness, it is important that they are evaluated fully and applied, where appropriate.

ACKNOWLEDGEMENTS

Many people kindly provided views, opinions and information which have been useful for this review. In this respect I would especially like to thank Dr M J Allison, Mr M Burbidge, Mr R Cook, Dr S R Draper, Dr P D Keefe, Dr B Oliver, Dr P J Sharpe, Dr P R Shewry, Dr G Wiseman and Dr C W Wrigley. Dr Alan Morgan is thanked for offering comments concerning the review as a whole. The typists at the NIAB are especially thanked for their considerable efforts and G C Gates and A Tiley for the photography. The review was funded by the Home-Grown Cereals Authority.