Computerized Wheat Varietal Identification by High-Performance Liquid Chromatography[1]

M. G. SCANLON, H. D. SAPIRSTEIN, and W. BUSHUK[2]

Cereal Chem. 66(5):439 443

ABSTRACT

A computerized library search procedure is described to perform wheat varietal identification based on reversed-phase high-performance liquid chromatography (RP-HPLC) of gliadins. Normalized peak retention times and relative peak heights (percent basis) comprise the chromatogram fingerprint and were used in the library search as primary and secondary discrimination parameters, respectively. The computer scores peaks as being matched if both the retention time and relative peak heights of compared components are within prescribed difference thresholds. Objective strategies for determining these optimum thresholds using a library of replicate chromatograms are described. The use of a weighted matching procedure in the search minimizes the effect of mismatching small peaks, whose retention times may be altered by noise and peak slope changes. The variety discriminatory power of the program using RP-HPLC is equal to that obtained by standard electrophoretic methods and has the advantage that it is totally automated, requiring no subjective input.

Polyacrylamide gel electrophoresis (PAGE) and reversed-phase high-performance liquid chromatography (RP-HPLC) represent two complementary approaches for the discrimination and identification of wheat varieties (reviewed recently by Wrigley et al [1982] and Bietz [1986], respectively). Despite good resolution and reproducibility, which make PAGE results amenable to automated identification using computer programs (Autran and Abbal 1988, Lookhart et al 1983, Sapirstein and Bushuk 1985b), many manual tasks are involved in preparing a suitably stained PAGE pattern for analysis. In this regard, RP-HPLC is superior, since extensive computer control of the chromatograph and many automated data processing functions, especially quantitation, are possible. The quantitation capability is particularly advantageous, given the complex separations that can be obtained.

[1]Presented in part at the AACC 72nd Annual Meeting, Nashville, TN, November 1987. Publication no. 149, Food Science Department.

[2]Food Science Department, University of Manitoba, Winnipeg, MB, Canada R3T 2N2.

This paper was prepared for electronic processing.

Despite the potential for routine manipulation of chromatographic data, there have been few reports on computer-based methods for cereal protein analysis. Some workers have, for example, strived to use the computer for predicting the proportion of different wheat genotypes in simple known mixtures (Bietz and Cobb 1985, Marchylo et al 1988), for tracing pedigrees (Bietz and Huebner 1987), and for estimating peak multiplicity (Bietz and Huebner 1987). However, automated large-scale comparison of chromatograms (library searches), which is necessary for varietal identification, has not been previously reported. The difficulty in selecting a set of reproducible features from the chromatogram that can be used in a varietal library without operator assistance for comparative purposes is hampered by a number of problems: 1) the complexity of RP-HPLC chromatograms of wheat proteins from different genotypes (Marchylo et al 1988, Wieser et al 1987), 2) the variation of peak quantitation data with location of wheat growth (Huebner and Bietz 1988), and 3) the change in peak retention times with column use (Glajch et al 1987, Marchylo et al 1988, Scanlon et al 1989).

This article describes a computerized approach to the comparison and identification of wheat varieties from their pattern

homologies that resolves some of these problems by implementing normalized peak retention times (Sapirstein et al 1989) and percentage peak heights as the wheat varietal fingerprints. The overall strategy is an adaptation of the approach used for wheat varietal identification by PAGE (Sapirstein and Bushuk 1985b). An objective method is also described to determine the optimum threshold for peak retention time and peak height discrimination to increase the likelihood that accurate peak matches are made.

MATERIALS AND METHODS

Chromatogram Acquisition and Data Base Composition

Apparatus, materials, gliadin protein extraction procedures, and selection of optimal conditions for acquisition of chromatograms have previously been reported (Scanlon et al 1989).

HPLC grade acetonitrile and ethanol were obtained from Fisher Scientific. Sequanal grade trifluoroacetic acid was purchased from Pierce Chemical Co. Water was distilled and then purified with a Millipore Milli-Q system. The varietal data base was constructed from RP-HPLC chromatograms for 36 hexaploid wheat varieties: 24 hard red spring (Apex, Benito, Butte, Canthatch, Canus, Chester, Chinook, Chris, Columbus, Coteau, Era, Garnet, Katepwa, Leader, Manitou, Marquis, Park, Pembina, Pioneer, Polk, Red Fife, Renfrew, Sinton, Thatcher), four hard red winter (Ridit, Sundance, Wasatch, Yogo), one soft red winter (Kent), four soft white winter (Gaines, Genessee, Nugaines, Talbot), one soft white spring (Cascade), and two miscellaneous spring (Bishop, Huron). These test samples were all found to be homogeneous genotypes and were classified as most similar to the variety Neepawa by computer analysis of their gliadin PAGE results by Sapirstein and Bushuk (1985b). In this way the efficacy of the program to discriminate closely related varieties could be tested.

The gliadins were extracted from pulverized grain (four kernels per sample) with a 4:1 ratio of 70% ethanol.

RP-HPLC was performed with a 1090M Hewlett-Packard liquid chromatograph using a wide pore (300Å), C8 Supelcosil column. Solutions of water and acetonitrile (both containing 0.1% trifluoroacetic acid) were made fresh for each, or every second, set of analyses.

The 36 test samples were chromatographed in sets of six or seven; each set included an analysis of a similar extract of the variety Neepawa. The latter represented five standard samples prepared in the same way as the data base samples, so that reproducibility of the procedure could be assessed. The 41 chromatograms so obtained comprised the wheat varietal data base for this study.

The protocol to chromatograph an extract prepared from a sample of Neepawa meal in an intervening fashion between sets of test sample extracts was used for computerized normalization of chromatograms relative to this external Neepawa standard. This significantly improves retention time precision (Sapirstein et al 1989). Chromatograms of 15 meal extracts of the variety Neepawa acquired while the data base was being chromatographed formed a separate (standard chromatogram) data base. This data base was used to evaluate the optimal peak difference threshold parameters for running the variety identification program.

Software Development

Three separate stages were involved in transforming protein elution profiles recorded at 210 nm to varietal identification results.

First, Hewlett-Packard chromatography (79994A) software processed raw chromatogram data files to report files of retention times and percentage peak heights (100 × peak height/sum of the chromatogram's peak heights). The peak width, threshold, and area rejection parameters of this software were set so that peaks were consistently integrated in the same manner in different extracts of the same sample (Scanlon et al 1989). Shoulder detection was not an option available in the software.

In the second stage, software developed in our laboratory condensed the detected peaks in each report file to an integer list of retention times (seconds) and percentage peak heights (× 100), which was termed a signature array. Signature array data were normalized according to the method of Sapirstein et al (1989) and stored as separate entries in a data base file. In this respect, the Hewlett-Packard model 310 computer (used to control the HPLC) provided excellent programming facilities in the form of a Pascal co-resident operating system, and a FORTRAN77 library and compiler (IEM Inc., Fort Collins, CO).

Finally, the same computer environment was used to adapt software originally developed for gliadin PAGE variety identification (Sapirstein and Bushuk 1985b) for RP-HPLC purposes. In the RP-HPLC program, peak retention time was used as the primary search parameter instead of band relative mobility, and percentage peak height was used in place of band intensity as the secondary search parameter. For the HPLC application, the program was improved by incorporating an iterative peak retention time search procedure and automatic calculation of the weighting factor for the data base used.

RESULTS AND DISCUSSION

Data Base Search Algorithm

The program compares peaks in the unknown chromatogram with peaks in a member of the data base, taking one peak at a time, using a combined forward-reverse search strategy (Peysna et al 1976). A peak from the unknown and a peak from one of the data base varieties are scored as matching if the differences in their retention times and peak heights are within prescribed thresholds (see below).

A percentage weighted pattern homology is calculated using the equation given below. This equation is analogous to the equation derived previously for PAGE (Sapirstein and Bushuk 1985b).

$$\text{Weighted pattern homology} = \frac{100\ \Sigma_{i=1}^{m}\ (H_U + H_R)_i / 2H_{AV}}{\Sigma_{i=1}^{m}\ (H_U + H_R)_i / 2H_{AV} + \Sigma_{i=1}^{j}\ (H_U)_i / H_{AV} + \Sigma_{i=1}^{k}\ (H_R)_i / H_{AV} + \Sigma_{i=1}^{l} |(H_U - H_R)|_i / H_{AV}}$$

where H_U = the percentage peak height of a peak in the unknown chromatogram; H_R = the percentage peak height of a peak in a data base (reference) chromatogram; H_{AV} = the weighting factor which is the mean percentage peak height for all varieties in the data base; m = the number of matching peaks; j = the number of peaks present in the unknown, but absent from the reference chromatogram; k = the number of peaks present in the reference but absent from the unknown chromatogram; and l = the number of peaks with similar retention times but significantly different peak heights.

The weighted pattern homology calculation is repeated for each data base entry, so that a list of data base varieties is generated, ranked according to their pattern homologies with the unknown.

Primary Search Parameter (Peak Retention Time) Optimization

As stated by Marshall et al (1987), peak positioning (or retention time) is the prime characteristic of an HPLC chromatogram. Although an appropriate normalization procedure (Sapirstein et al 1989) will correct random and systematic variations in peak retention times, a threshold must be allowed to account for any residual variation when two peaks are compared. This threshold must not be so small that identical peaks are scored as being different (type I error), nor must it be so large that (truly) different peaks are misclassified as matches (type II error). As discussed by Sapirstein et al (1989) the second type of error likely dominates because of the complex heterogeneity of the separation. Ideally both type I and II experimental errors should be minimized (Sapirstein and Bushuk 1985a, Sapirstein et al 1989).

To evaluate objectively the magnitude of the peak retention time difference threshold required to compute chromatogram pattern homologies, the data base of replicate Neepawa chromatograms was analyzed. Theoretically, comparisons among such a group should yield uniformly high pattern homology scores given the selection of an appropriate retention time difference threshold

consistent with the experimental error in the data. The result of this process (Fig. 1) indicated that for 60-min analyses, as carried out in this study, the selected threshold does not need to be increased significantly beyond about 10 sec in order to match common chromatogram peaks. There was virtually no increase in computed pattern homologies beyond 15 sec, which was used as the limit threshold for matching common peaks on the basis of retention time.

Incrementation of the Retention Time Threshold

A fixed difference threshold of 15 sec was found to provide satisfactory performance for computerized varietal identification using the data base of 41 chromatograms evaluated in this study. In practice, however, as the data base increases in size and peak diversity increases, the time separating different chromatogram peaks can be less than 15 sec, as is the case for some minor components in the data base. This would result in false peak matches when using a fixed difference threshold approach. Accordingly, the pattern homology process was further enhanced by using an iterative thresholding strategy.

As implemented in the varietal identification program described below, the difference threshold was set to an initial low value (5 sec). Compared chromatograms were scanned for matching peaks occurring within this retention time window. The threshold was incremented by 5 sec on each successive pass, up to the maximum threshold of 15 sec. Peaks remaining unmatched at the maximum threshold were classified as mismatches.

It is important to note that the effectiveness of this iterative approach may be valid only when the chromatogram data (peak retention times) are very precise. Figure 1 shows that for a data base of normalized peak retention times, the program matches components differing by less than approximately 5 sec. In contrast, however, is the parallel result for the uncorrected peak retention time data base (Fig. 1), where inaccurate pattern homology scores would result, since a significant number of replicate peaks would remain unmatched if a 5-sec difference threshold were to be used. This result provides good evidence that even for short-term acquired data, the ability to comprehensively and accurately compare and find real differences in nonnormalized RP-HPLC chromatograms is limited.

Secondary Differentiative Parameter (Relative Peak Height)

Despite band similarity in terms of relative mobility, Lookhart et al (1983) and Autran and Abbal (1988) considered differences in electrophoretic band intensity the basis for ascribing lesser confidence in the match. The greater the difference in band intensity the more likely it was that the bands were not common protein components. No differences in matching value were used by Sapirstein and Bushuk (1985b), but bands were considered to not contain identical components, despite similar relative mobility, if there were relatively large differences in electrophoretic band intensity. Likewise a pair of peaks with matching retention times was not automatically considered to be a correct match. Therefore, differences in percentage peak heights represented a secondary differentiating parameter for the variety identification program.

Unlike the relatively uniform precision for peak retention times with short-term column use (Sapirstein et al 1989), peak height precision depends on the size of the peak. Ideally an equation that supplies the difference threshold for a given peak height would allow peak height matching decisions to be made. Figure 2 shows the standard deviation of percentage peak height plotted against the mean of percentage peak height for visually matched common components in the Neepawa data base. Only peaks that were detected in 12 or more of the 15 replicates were retained in this analysis.

In view of the large variation in peak size that can occur for extracts of the same wheat variety grown at different locations (Huebner and Bietz 1988), a peak height difference threshold was arbitrarily set at six times the expected standard deviation obtained from the regression relationship given in Figure 2. Although six times the expected standard deviation might seem to allow excessive latitude when comparing peak heights, it can be seen from Figure 2 that the greater majority of Neepawa components have relative peak heights with values less than 5% and standard deviations less than 0.15. The peak height difference threshold evaluated in this way would be less than 1% for these peaks (6 × 0.15), which provides a reasonable level of

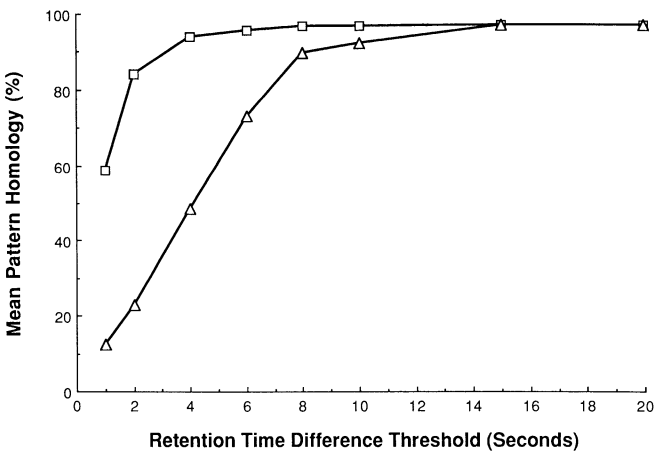

Fig. 1. Mean percentage pattern homology scores at various retention time difference thresholds for the Neepawa data base computed with (□), and without (△), retention time normalization. Each symbol denotes the mean score for pairwise comparisons of the 15 replicates with a test entry selected at random.

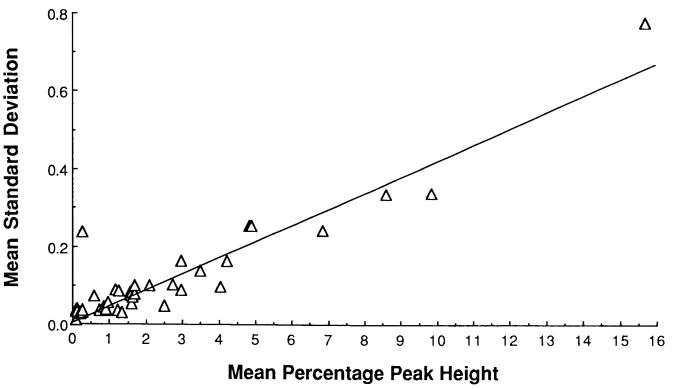

Fig. 2. Mean standard deviation versus mean percentage peak height for peaks that were visually evaluated as identical in 12 or more of 15 Neepawa replicate chromatograms. The line is the linear regression fit to the data described by the equation: Percentage peak height standard deviation = 0.0417 × (mean % peak height) + 0.010.

1

WHEAT CULTIVAR IDENTIFICATION BASED ON GLIADIN CHROMATOGRAMS - I. RANKING BY PATTERN HOMOLOGY

* 41 DATA BASE CULTIVAR PATTERNS ANALYSED
* DATA BASE SEARCH CUTOFF AT 70% PATTERN HOMOLOGY (WEIGHTED BY PERCENTAGE PEAK HT).
* RETENTION TIME DIFFERENCE THRESHOLD - 5 SECONDS TO 15 SECONDS IN INCREMENTS OF 5 SECONDS.
* UNKNOWN (OR TEST) CULTIVAR CHROMATOGRAM CONTAINS 37 GLIADIN PEAKS; TOTAL, WEIGHTED BY PEAK PEAK HT (WPH) = 40.3

			GLIADIN PEAKS IN	MATCHING	DISTRIBUTION OF NON-MATCHING PEAK DATA					
	CULTIVAR	WEIGHTED % PATTERN HOMOLOGY	PATTERN NO. WPH	PEAKS NO. WPH	TOTAL NO. WPH	RET TIME BASIS-R NO. WPH	PEAK HT BASIS NO. WPH	RET TIME BASIS-U NO. WPH	CLASS/TYPE	REGION
1	NEEPAWA.1508	100	37 (40.3)	37 (40.3)	0 (.0)	0 (.0)	0 (.0)	0 (.0)	HRS-SMQ	W.CAN
2	NEEPAWA.1908	94	45 (40.3)	35 (39.0)	11 (2.5)	9 (2.1)	1 (.1)	1 (.3)	HRS-SMQ	W.CAN
3	NEEPAWA.808	93	46 (40.3)	33 (38.8)	17 (3.0)	13 (2.1)	0 (.0)	4 (.9)	HRS-SMQ	W.CAN
4	NEEPAWA.1808	90	39 (40.3)	25 (35.6)	15 (4.1)	3 (1.7)	11 (2.2)	1 (.3)	HRS-SMQ	W.CAN
5	MANITOU	90	45 (40.3)	32 (36.4)	16 (3.9)	11 (2.0)	2 (.8)	3 (1.1)	HRS-SMQ	W.CAN
6	NEEPAWA.1708	88	46 (40.3)	32 (34.6)	17 (4.6)	12 (2.2)	2 (1.0)	3 (1.3)	HRS-SMQ	W.CAN
7	CHRIS	86	46 (40.3)	31 (36.0)	20 (5.7)	14 (3.2)	1 (.4)	5 (2.0)	HRS-SMQ	USA
8	CANTHATCH	86	43 (40.3)	28 (32.4)	19 (5.3)	10 (1.5)	5 (2.1)	4 (1.8)	HRS-SMQ	W.CAN
9	THATCHER	84	47 (40.3)	29 (32.6)	22 (6.0)	14 (2.6)	4 (1.7)	4 (1.7)	HRS-SMQ	W.CAN
10	LEADER	83	44 (40.3)	28 (35.1)	22 (7.3)	13 (3.1)	3 (.7)	6 (3.5)	HRS-SMQ	SAWFLY
11	PARK	81	37 (40.3)	21 (32.5)	23 (7.7)	7 (1.5)	9 (2.6)	7 (3.6)	HRS-SMQ	W.CAN
12	BENITO	81	42 (40.3)	26 (30.5)	21 (7.1)	10 (1.6)	6 (3.0)	5 (2.6)	HRS-SMQ	W.CAN
13	KATEPWA	74	47 (40.3)	28 (27.6)	23 (9.9)	14 (4.5)	5 (4.2)	4 (1.2)	HRS-SMQ	W.CAN
14	ERA	72	40 (40.3)	26 (28.0)	19 (10.8)	8 (2.8)	6 (3.4)	5 (4.6)	HRS-MEMQ	USA

Fig. 3. Variety identification short list ranking for the test extract Neepawa.1508. The "Ret Time Basis-R" column lists the number of peaks in the data base entry chromatogram that differ from the unknown chromatogram on the basis of retention time. The "Ret Time Basis-U" column lists the converse number of peaks for the unknown.

discriminatory capability for peak heights. Thus, if the difference in peak heights was greater than this threshold value, the peaks were scored as being quantitatively dissimilar. The lesser importance of the secondary differentiating parameter is taken into account in the pattern homology equation described previously.

When this secondary differentiating parameter was incorporated into the program, there were no chromatograms in the replicate Neepawa data base with peaks mismatched on the basis of peak height. For the limited sample contained in the Neepawa data base at least, the result indicates that the peak height difference evaluation process performed satisfactorily.

Variety Identification Results

The strategies described above were implemented in the variety identification program that was tested on the varietal data base. Figure 3 shows a typical result for Neepawa (Neepawa.1508) compared with itself and with the rest of the data base. The output format is similar to that for PAGE results (Sapirstein and Bushuk 1985b). The output heading specifies the parameters used in the search, such as the increments used in the retention time difference threshold, the final size of the threshold, and the pattern homology cutoff value. The latter ensures that a full data base report is not printed so that only those varieties that reasonably match the unknown are listed.

Regardless of the absolute number of peaks in the chromatogram, the total weighted number of peaks was computed as 40.3. This value is obtained by dividing 100 (the sum of individual percentage peak heights) by the weighting factor (2.48), which is the mean percentage peak height for the entire data base. In this way, the effect of protein content differences between samples was minimized, with the possible exception of samples with low protein contents, where some components may be present at levels that are too low to permit integration.

The result (Fig. 3) indicates that four out of five of the data base entries that most closely matched the single test unknown (Neepawa.1508) were Neepawa replicates. This confirms the ability of the program to discriminate varieties when a large threshold is used for peak height differences. The effectiveness of employing a weighting procedure for improving library searches, as reported by Peysna et al (1976), is shown for varieties that have large numbers of peaks but that only have a small number of weighted peaks, and so contribute little to the pattern homology score. For example, although Neepawa.808 had 17 peaks mismatched, the weighted number of these peaks was only 3.0, indicating that most were minor components.

In an automated peak assignment system such as used in this study, it is important that the occasional artifacts, such as baseline disturbances that are integrated and incorporated in a data base entry as a peak, are not given the same prominence as true components. The weighting factor will ensure that the computer variety identification results are only minimally affected by these "peaks." Likewise, because the system relies on the integrator to make peak detection decisions, lack of shoulder detection and quantitation can cause problems in regions of the chromatograms where a number of merged peaks are present. Thus chromatograms that appear visually identical may give low pattern homologies due to slight differences in peak slope at points where peaks merge, so that in one chromatogram extra peaks may be assigned, while in the other only shoulders are recognized. Chromatography software that quantifies shoulders in the same manner as separated peaks should eliminate the problem. Moreover the inclusion of additional detected shoulder components should also result in improved discriminatory capability in variety identification, since the shoulders may represent characteristic proteins for certain varieties.

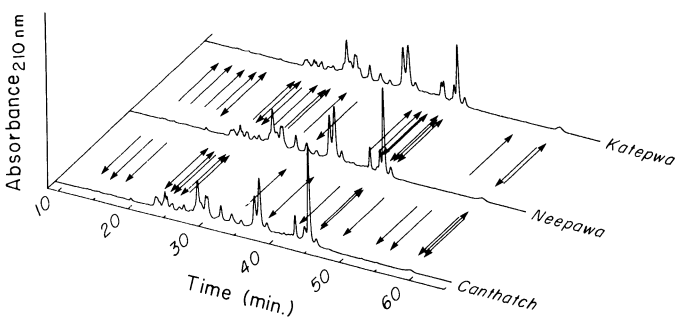

Fig. 4. Comparisons of the chromatogram of Neepawa.1508 with those of Katepwa and Canthatch. The head of the arrow points to the variety that contains the nonmatching peak. Double-headed arrows point to peaks present in Neepawa and in the other variety that are deemed mismatches on the basis of peak height.

Figure 4 shows the chromatograms of Neepawa.1508, Canthatch, and Katepwa after converting peak heights to percentage peak heights. The lines point to differences between peaks of Katepwa and Neepawa, and between Canthatch and Neepawa. Most of the differences are due to minor peaks and so contribute little to differences in pattern homology. Larger peaks that mismatch on the basis of peak height differences do not contribute significantly to pattern homology differences. However, a different extract of the Neepawa standard might show different peaks that would be scored as mismatches, and so change the pattern homology scores and the computed rankings of Figure 3. Consequently each of the Neepawa chromatograms was run against the data base to generate five sets of ranked pattern homology scores similar to those of Figure 3. The mean and standard deviation of pattern homology scores for each variety are shown in Table I. Scores of 100 obtained for each Neepawa run against itself were eliminated to offset the bias towards high pattern homology for the Neepawa standards that this would create. Accordingly the mean scores for Neepawa are means of four results rather than five.

Reproducibility was good; the Neepawa samples retained the tight grouping they had in Figure 3 (Table I). Varieties other than those listed in the table were not considered as matches for Neepawa since they had mean pattern homology scores below 70%. However, because the mean standard deviations for the Neepawa, and non-Neepawa chromatograms are 2.4 and 6.5, respectively, the varieties that are listed would occasionally be classified as being matches for Neepawa on the basis of their chromatographic patterns. This implication is not surprising in view of the common genetic background of many of these varieties.

The extent to which RP-HPLC compared in discriminatory ability with a standard electrophoretic approach was evaluated. Table II lists the weighted pattern homology scores for a PAGE varietal data base search using Neepawa as the test unknown (Sapirstein and Bushuk 1985b) after eliminating varieties that were not chromatographed in this study. Comparing the results in Tables I and II, there is very close agreement between those varieties that match Neepawa by PAGE (i.e., protein size and charge density), and those that match by RP-HPLC (protein surface hydrophobicity). Therefore it appears that 60-min

TABLE I
Means and Standard Deviations of Pattern Homology Scores Comparing Five Neepawa Kernel Extracts with the Data Base

Variety Name	Mean Percentage Homology Score	Standard Deviation
Neepawa.1908	95	1.3
Neepawa.808	94	1.7
Neepawa.1508	92	2.1
Neepawa.1708	91	3.2
Neepawa.1808	91	1.8
Manitou	91	1.9
Chris	87	5.2
Canthatch	84	2.9
Thatcher	84	5.5
Benito	80	1.8
Leader	80	5.2
Park	76	3.8
Era	75	4.2
Katepwa	75	3.7

TABLE II
Electrophoregram Cultivar Listing for Neepawa Taken from the Results of Sapirstein and Bushuk (1985b), After Exclusion of Biotypes

Cultivar Name	Weighted Percentage Pattern Homology
Neepawa	100
Manitou	98
Katepwa	97
Canthatch	96
Thatcher	96
Chris	91
Benito	91
Leader	80
Era	75

chromatographic analyses yield similar results to those obtained by computer analysis of PAGE results (Sapirstein and Bushuk 1985b). The greater spread in the pattern homologies in Table I for the other varieties suggests that the discriminatory power of RP-HPLC is greater than that of PAGE, but the spread in pattern homologies for the Neepawa analyses shows that the other varieties in Table I would not always be differentiated from Neepawa.

Algorithm Improvements

A significant feature of the software is that it can be adjusted to cope with additional peak information, so that extra discriminatory power can be accommodated. For example, the use of diode-array detection allows spectra to be acquired as a peak is eluted, thus adding a purity assignment to the peak (Conlon et al 1987). Each entry in the data base will then have a three-dimensional format: each peak having a retention time, a percentage peak height, and a peak purity assignment. Peaks that match on the basis of retention time and peak height can have their purities compared, possibly improving the discrimination capability of the program. An algorithm to include a third, and possibly additional dimensions, is being developed.

CONCLUSION

A computerized RP-HPLC variety identification program based on wheat prolamin chromatograms was described. The program is fast, comprehensive, and provides quantitative information on matching and nonmatching peaks in objectively compared chromatographic data. The discriminatory power of the program was as good as computer-assisted variety identification results based on gliadin PAGE but has the added advantage that no subjective input is required given the general quantitative capability of the chromatograph. Whereas very satisfactory performance was achieved, it should be noted that the program was tested using a varietal data base constructed from chromatograms acquired over a two-week period only. It is known however, that retention time drifts and selectivity changes do occur with prolonged column use, especially for early eluting gliadin components (Scanlon et al 1989). These may affect peak quantitation even after retention times have been corrected by normalization procedures. What effect long-term column use may have on the variety identification program's discrimination ability remains to be investigated.

ACKNOWLEDGMENTS

The authors are grateful for financial assistance for this project from the Natural Sciences and Engineering Research Council of Canada and thank Denise Lawless for performing some of the chromatographic analyses.

LITERATURE CITED

AUTRAN, J. C., and ABBAL, P. 1988. Wheat cultivar identification by a totally automatic soft-laser scanning densitometry and computer-aided analysis of protein electrophoregrams. Electrophoresis 9:205-213.

BIETZ, J. A. 1986. High-performance liquid chromatography of cereal proteins. Adv. Cereal Sci. Technol. 8:105-170.

BIETZ, J. A., and COBB, L. A. 1985. Improved procedures for rapid wheat varietal identification by reversed-phase high-performance liquid chromatography. Cereal Chem. 62:332-339.

BIETZ, J. A., and HUEBNER, F. R. 1987. Prediction of wheat quality by computer evaluation of reversed-phase high-performance liquid chromatograms of gluten proteins. Page 173 in: Proc. Int. Workshop on Gluten Proteins 3rd. R. Lasztity and F. Bekes, eds. World Scientific: Singapore.

CONLON, R. D., VANDEMARK, F. L., DONG, M. W., and SCHMIDT, G. J. 1987. A high-sensitivity photodiode array LC detector. Am. Lab. 19(10):96-105.

GLAJCH, J. L., KIRKLAND, J. J., and KOHLER, J. 1987. Effect of column degradation on reversed-phase high-performance liquid chromatographic separation of peptides and proteins. J. Chromatogr. 384 81.90.

HUEBNER, F. R., and BIETZ, J. A. 1988. Quantitative variation among gliadins of wheats grown in different environments. Cereal Chem. 65:362-366.

LOOKHART, G. L., JONES, B. L., WALKER, D. E., HALL, S. B., and COOPER, D. B. 1983. Computer-assisted method for identifying wheat cultivars from their gliadin electrophoregrams. Cereal Chem. 60:111-115.

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 Chem. 65:28-40.

MARSHALL, R. J., BLEASBY, A. J., TURNER, R., and COOPER, E. H. 1987. A computer system for analysis of chromatographic data. Chemomet. Intell. Lab. Syst. 1:285-295.

PEYSNA, G. M., VENKATARAGHAVAN, R., DAYRINGER, H. E., and McLAFFERTY, F. W. 1976. Probability based matching system using a large collection of reference mass spectra. Anal. Chem. 48:1362-1368.

SAPIRSTEIN, H. D., and BUSHUK, W. 1985a. Computer-aided analysis of gliadin electrophoregrams. I. Improvement of precision of relative mobility determination by using a three reference band standardization. Cereal Chem. 62:372-377.

SAPIRSTEIN, H. D., and BUSHUK, W. 1985b. Computer-aided analysis of gliadin electrophoregrams. II. Wheat cultivar identification and class comparisons. Cereal Chem. 62:377-392.

SAPIRSTEIN, H. D., SCANLON, M. G., and BUSHUK, W. 1989. Normalisation of HPLC peak retention times for computerised comparison of wheat prolamin chromatograms. J. Chromatogr. 469:127-135.

SCANLON, M. G., SAPIRSTEIN, H. D., and BUSHUK, W. 1989. Evaluation of the precision of high-performance liquid chromatography for wheat cultivar identification. Cereal Chem. 66:112-116.

WIESER, H., MODL, A., SEILMEIER, W., and BELITZ, H. D. 1987. High-performance liquid chromatography of gliadins from different wheat varieties: Amino acid composition and N-terminal amino acid sequence of components. Z. Lebensm. Unters. Forsch. 185:371-378.

WRIGLEY, C. W., AUTRAN, J. C., and BUSHUK, W. 1982. Identification of cereal varieties by gel electrophoresis of the grain proteins. Adv. Cereal Sci. Technol. 5:211-259.

[Received November 30, 1988. Accepted May 1, 1989.]