Journal of Chromatography, 596 (1992) 225–232
Elsevier Science Publishers B.V., Amsterdam

CHROM. 23 986

Characterization and identification of wheat cultivars by multi-dimensional analysis of reversed-phase high-performance liquid chromatograms

Ph. Courcoux* and Th. Serot

École Nationale d'Ingénieurs des Techniques des Industries Agroalimentaires, Domaine de la Géraudière, 44072 Nantes Cedex 03 (France)

C. Larre and Y. Popineau

Institut National de la Recherche Agronomique, Domaine de la Géraudière, 44072 Nantes Cedex 03 (France)

(First received July 18th, 1991; revised manuscript received December 31st, 1991)

ABSTRACT

Reversed-phase high-performance liquid chromatography was used to identify wheat cultivars. A collection of 59 chromatograms of the gliadin fractions of samples from 18 different cultivars was analysed through multi-dimensional statistical methods. The discrimination quality of the different varieties studied makes possible a classification of the chromatograms obtained. The great variability of the chromatographic profiles and the correct repeatability of the technique used indicate that multi-dimensional analysis applied to the treatment of reversed-phase chromatograms is a reliable and relevant method for cultivar identification.

INTRODUCTION

Since Bietz's work [1] showing that cereal storage proteins can be characterized by reversed-phase high-performance liquid chromatography (RP-HPLC), many papers on the use of this method for wheat: identification [2,3], quality prediction [4–6] and genetic analysis of cereals [7] have been published. Studies have also been carried out on the computer processing of the chromatograms obtained [8]. More recently, algorithms allowing identification by comparison of an unknown sample with a database [9] have been published. Recognition is achieved by homology of the retention times of identified peaks. This method originates from work on the identification of wheat cultivars by electrophoresis [10,11].

Statistical methods and techniques of multi-dimensional data analysis are frequently used to study sets of continuous signals, especially in infrared spectrometry [12]. The application of such methods to gel permeation HPLC data has proved relevant for discriminating and identifying varieties [13]. In this work, multi-dimensional methods were used to characterize and classify different wheat cultivars by RP-HPLC of the gliadin fraction, wich has been shown to have good potential for varietal identification [14].

EXPERIMENTAL

Chemicals

Protein sequencing grade trifluoroacetic acid (TFA) was supplied by Sigma Chimie (France) and HPLC-grade acetonitrile (ACN) by Carlo Erba (Milan, Italy). HPLC-grade water was prepared with a Milli-Qplus water purification system (Millipore, Molsheim, France).

0021-9673/92/$05.00 © 1992 Elsevier Science Publishers B.V. All rights reserved

Wheat samples

Different wheat samples were supplied by INRA Plant Breeding Station (Clermont Ferrand, France) and consisted of eighteen cultivars of common wheat: Andain (AN), Arminda (AR), Capitole (CA), Castan (CS), Courtot (CO), Fanion (FA), Festival (FE), Florence–Aurore (FL), Gavroche (GA), Hardi (HA), Lutin (LU), Maître Pierre (MP), Pistou (PI), Prinqual (PR), Récital (RE), Rex (RX), Tarasque (TA) and Thésée (TH).

Sample prepartion

After grinding the grains with a refrigerated Dangoumeau grinder, gliadins were extracted from flour with 70% ethanol and then purified by gel permeation on S300 Sephacryl gel in 0.028 *M* aluminium lactate (pH 3.6) to remove glutenin. Gliadin fractions were then dialysed against 0.1% acetic acid and freeze-dried.

The different samples corresponding to α-, β-, ω- and γ-gliadins were obtained by fractionating crude gliadins by ion-exchange chromatography on SP-Trisacryl M [15] and by hydrophobic interaction chromatographgy [16]. The fractions were dialysed, freeze-dried and characterized by polyacrylamide gel electrophoresis at pH 3.1. For chromatography, gliadins were weighed, dissolved in water–ACN (72:28) + 0.088% TFA and filtered through a 0.2-μm Millipore membrane before injection.

HPLC

Chromatography was carried out with an LDC/ Milton Roy Series 4000 system. Chromatographic data were stored in an IBM-compatible microcomputer through Labnet software (Spectra-Physics).

A Vydac (Cluzeau, France) C_{18} column (15 cm × 0.46 cm I.D.) of 300 Å porosity was used to separate gliadins.

A 100-μl protein sample (5 mg/ml) was injected on to the column. The proteins were eluted with a linear gradient from water–ACN (72:28) + 0.088% TFA to water–ACN (52.5:47.5 + 0.080% TFA in 60 min. The flow-rate was 1 ml/min and protein detection was performed at 226 nm.

Two methods were used. Method A: solvent A ACN–water (15:85) +0.1% TFA and solvent B ACN–water (80:20) +0.06% TFA [1,2], gradient from 28% to 47.5% ACN in 60 min. Method B: solvent A water +0.1% TFA and solvent B ACN +0.06% TFA, gradient from 28% to 47.5% ACN in 60 min.

Statistical treatments

Prior to treatment, the chromatographic signals were truncated in order to retain only the useful part situated between 13 min and 59.8 min after the injection peak. The baseline of these chromatograms was then subtracted and their areas were adjusted to the same reference. After this reduction phase, the chromatograms all had the same surface area.

The chromatograms were then integrated by sections. On the basis of electrophoresis, thirteen sections were retained, corresponding to zones enriched in one or two gliadin components. Therefore, each chromatographic profile is described by thirteen variables corresponding to the areas of the integration zones.

The processed file is made up of 59 observations (repetitions on the eighteen cultivars) and thirteen variables (integration areas). An analysis of variance (ANOVA) and a canonical discriminant analysis (DISCRIM) were carried out in order to describe the most discriminant factors, together with a hierarchical clustering (CLUSTER with Ward's minimum variance method) allowing a tree diagram (or dendrogram) of these cultivars to be elaborated.

These treatments were performed with SAS statistical software (SAS Institute, Cary, NC, USA) on a SUN workstation and with specifically developed software.

RESULTS AND DISCUSSION

Chromatographic reproducibility assays

Preliminary tests with crude gliadins from the cultivar Hardi were carried out in order to determine the chromatographic conditions allowing the best reproducibility. In the first experiments, gliadin elution was effected according to method A, in which prediluted solvents were used. For the twelve major peaks, the average retention times, the standard deviations and the relative standard deviations were calculated for seven repetitions (Table I). It is interesting that as long as the gradient was performed with the same series of solvents A and B, the calculated standard deviation was near 0.1 min. However, the use of new solvents A and B, even

TABLE I

COMPARISON OF THE TWO CHROMATOGRAPHIC METHODS

Averages, standard deviations (S.D.) and relative standard deviations (R.S.D.) for the retention times of the twelve major peaks for seven repetitions on the Hardi cultivar. For methods A and B, see Experimental.

Peak No.	Method A			Method B		
	Average retention time (min)	S.D. (min)	R.S.D. (%)	Average retention time (min)	S.D. (min)	R.S.D. (%)
1	30.59	0.61	1.86	30.37	0.10	0.34
2	31.84	0.58	1.72	31.54	0.12	0.38
3	33.17	0.58	1.65	32.82	0.10	0.30
4	34.10	0.56	1.56	33.71	0.11	0.32
5	35.57	0.56	1.48	35.12	0.11	0.31
6	36.31	0.57	1.50	35.83	0.10	0.29
7	37.10	0.58	1.49	36.65	0.12	0.32
8	37.78	0.55	1.38	37.41	0.11	0.28
9	39.39	0.53	1.29	39.06	0.12	0.30
10	42.86	0.53	1.18	42.33	0.11	0.27
11	44.38	0.53	1.14	43.93	0.13	0.29
12	51.06	0.51	0.95	51.32	0.12	0.23

prepared under closely controlled conditions, decreased the reproducibility considerably. This led to standard deviations up to 0.51 min as indicated in Table I. This lack of reproducibility has to be related to variations in the preparation of solvents. In order to avoid such variations, another method (method B) in which solvents A and B were used pure, was applied in a second series of experiments. The relative standard deviations obtained (Table I) were six times lower than those obtained with method A. These results are in agreement with previous reports for RP-HPLC of gliadins [14] with automatic analysis.

Variability of chromatographic profiles

The subsequent elution profiles were all obtained using method B. Depending on the cultivar, about twenty peaks resulted (Fig. 1). The chromatograms can be roughly split into four zones corresponding to the four gliadin groups. Referring to the profiles of purified gliadins (Fig. 2), the less hydrophobic ω-gliadins were eluted between 13 and 24 min, the α- and β-gliadins between 24 and 40 min and the γ-gliadins between 40 and 45 min. In that zone, the γ-46 gliadin peak was identified at 43 min and that of γ-44 gliadin at 44 min. Proteins eluted after 45 min were not identified. The retention times and the acetonitrile percentages (Table II) at which the different types of gliadins were eluted are similar to those obtained by Lookhart and Albers [17].

Some chromatograms exhibited sufficient qualitative differences to be easily distinguishable. These differences mainly concerned the ω- and α-gliadin fractions.

Statistical analysis

Peak retention time and peak size are the prime characteristics in HPLC. For further analysis, each pattern was normalized. This reduction of the data minimized the protein content differences between samples. Thirteen integration sections were defined (Fig. 3) and their areas were considered as the variables of the analysis of variance. The results of this analysis carried out on these thirteen variables for the 59 observations are summarized in Table III. The sections which best discriminate the different cultivars are S1, S2, S3 and S8 with highly significant tests. From the localisation of the different groups of gliadins on a chromatogram, the sections S1, S2, S3 correspond to ω-gliadins, while S8 correspond essentially to α-gliadins. These results are in agreement with the observations of Branlard and Dardevet [18], who showed that among the four groups of gliadins, the amounts of the ω- and

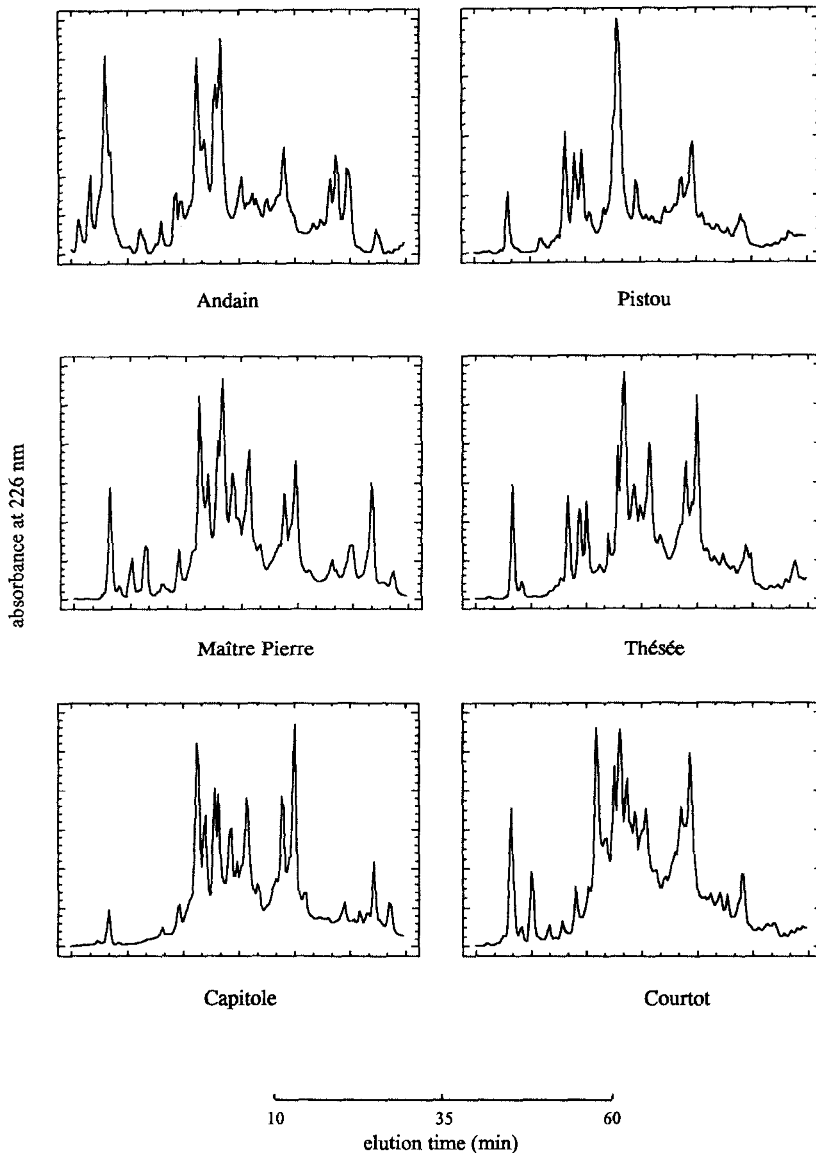

Fig. 1. RP-HPLC elution profiles of the wheat gliadin fraction of six cultivars (part situated between 13 and 59.8 min). The gradient is performed according to method B (ACN from 28% to 47.5% in 60 min); flow-rate, 1 ml/min; detection at 226 nm. The baseline has been subtracted and the surfaces adjusted to the same reference.

α-components were the most variable in 70 cultivars. Nevertheless, they observed also that the amount of each group of gliadin was strongly related to the protein content of the grains, whereas in our experiment the influence of protein content was minimized. Moreover, Autran and Bourdet [19], using electrophoretic techniques for varietal determination, pointed out the discriminating power of ω-gliadins. Nothing has been mentioned previously about α-gliadins.

Fig. 4 shows the first plane of the discriminant analysis on cultivars (this plane explains 63.8% of the total inertia). The projection nearness of profiles corresponding to repetitions of each cultivar con-

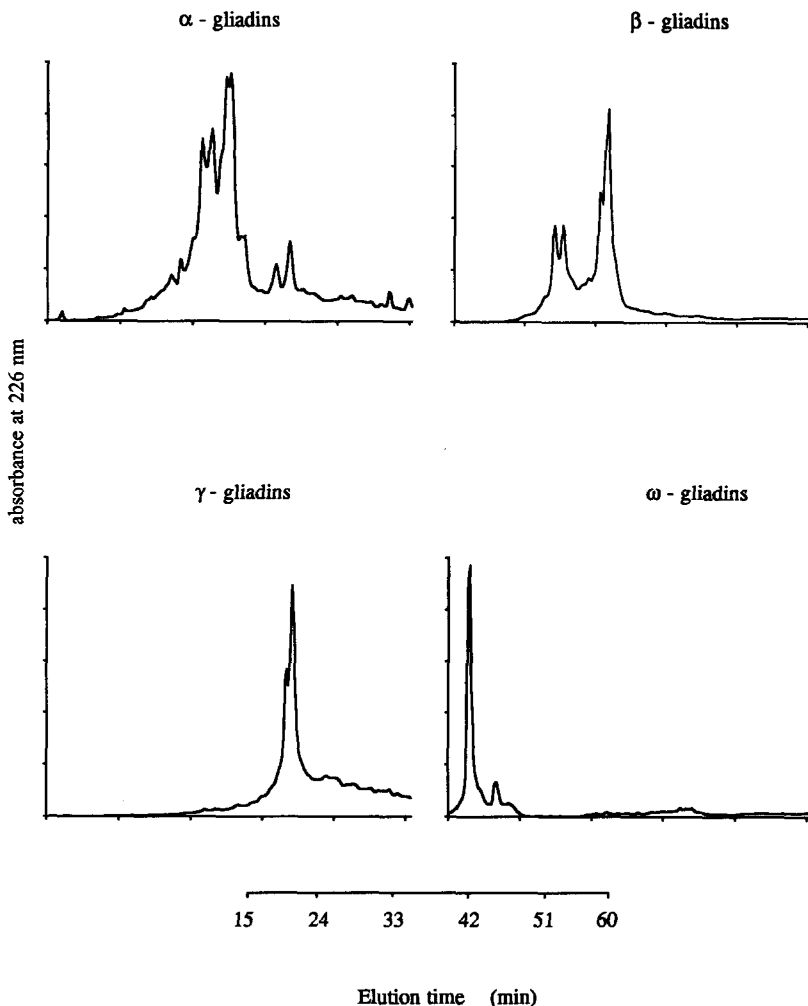

Fig. 2. Chromatograms of purified fractions. Elution conditions as in Fig. 1.

firms the very good repeatability of the chromatograms. On the other hand, great variability of the profiles can be seen, confirming visual observations with particular cultivars (Pistou, Andain) and close relationships between others (*e.g.* Maître Pierre and Fanion).

The discrimination of cultivars by DISCRIM analysis is perfect because, after resubstitution, all observations are correctly assigned to their original cultivar. This is the consequence of the good repeatability of the chromatographic technique used and of the great variability of gliadins. These results confirm the possible use of these statistical methods for the analysis and identification of French wheats based on the gliadin chromatographic profile.

Fig. 5 represents the projections of original variables on the plane of the two first canonical factors. Sections S8 and S12 are very correlated with canonical factor 1 and S1 is very correlated with canonical factor 2. These results are in agreement with those obtained previously by analysis of variance. The proteins eluted in section S12 have not been identified but, in work on the characterization of Hard Red Spring and Hard Red Winter wheats, Endo *et al.* [20] pointed out the discriminating power of the last-eluted peak. The diagram allows a more precise

TABLE II

ELUTION CONDITIONS OF THE DIFFERENT GLIADIN TYPES FOR THE STUDIED CULTIVARS

Operating conditions according to method B (see Experimental).

Peak No.	Retention time (min)	ACN (%)	Gliadin type
1	15.0	23	ω
2	19.4	34.30	ω
3	23.0	35.47	β
4	24.4	35.93	β
5	25.5	36.28	β
6	29.0	37.42	β
7	30.5	37.9	β
8	32.5	38.56	α
9	34.5	39.2	α
10	39.5	40.8	γ-16
11	40.7	41.22	γ-44
12	45–48	42.62–43.6	Not identified

interpretation of the canonical plane previously introduced. The position of the Andain cultivar on the first discriminant plane can then be explained by the wide surfaces of zones S1 and S6 and the positioin of the Pistou cultivar by the high contributions of zones S4 and S5.

Clustering of cultivars

The dendrogram resulting from a hierarchical clustering on the chromatographic profiles of the different cultivars is shown in Fig. 6. This representation helps to examine distances between cultivars and to find natural partitioning in the studied collection. The horizontal axis of the dendrogram represents a distance index measuring the similarities between cultivars. The more the linkage occurred near the leaves of the tree, the more similar are the chromatograms of the cultivars.

The originality of the Pistou and Andain cultivars is verified on the hierarchical tree produced by the CLUSTER procedure and the proximities noticed previously are confirmed. Cutting the tree at the right level allows partitioning of all cultivars into five groups the individuals of which have close profiles:

g1: Pistou;
g2: Andain;
g3: Maître Pierre, Fanion;
g4: Castan, Arminda, Thésée, Lutin;

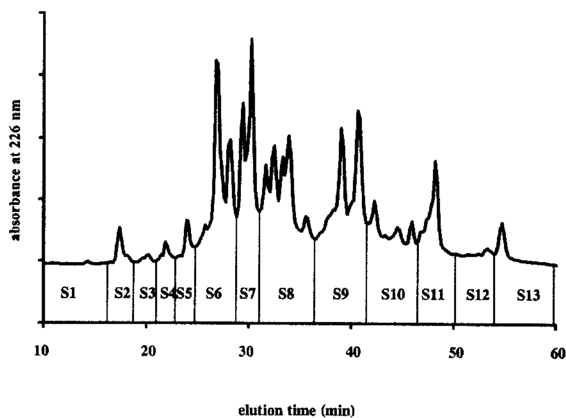

Fig. 3. Separation of the elution profile of Hardi cultivar into thirteen integration sections. Elution conditions as in Fig. 1.

g5: Capitole, Récital, Tarasque, Festival, Rex, Gavroche, Florence, Courtot, Hardi, Prinqual.

This partitioning shows good agreement with the positions of the cultivars on the first plane of the discriminant analysis shown previously. This dendrogram results only from the polymorphism of gliadin as revealed by RP-HPLC. It will be of interest to compare these results with genetic proximities between cultivars.

TABLE III

ANALYSIS OF VARIANCE ON THE 18 CULTIVARS

Influence of the cultivar on the areas of the 13 different sections of the 58 chromatograms.

Section	Mean square	$F(17,41)$[a]
S1	29.39	414
S2	1.73	234
S3	1.78	203
S4	2.30	25
S5	6.42	113
S6	21.96	142
S7	11.43	102
S8	38.06	240
S9	22.45	44
S10	9.19	29
S11	10.00	139
S12	8.92	158
S13	2.26	5

[a] F test calculated with 17 and 41 degrees of freedom (d.f.).

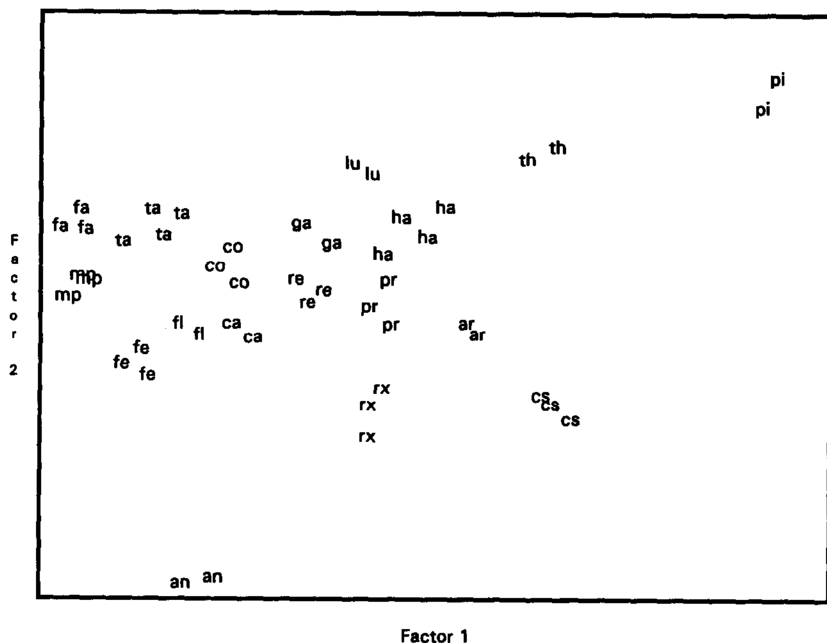

Fig. 4. First factorial plane of the discriminating analysis of cultivars. Projection of the 59 samples.

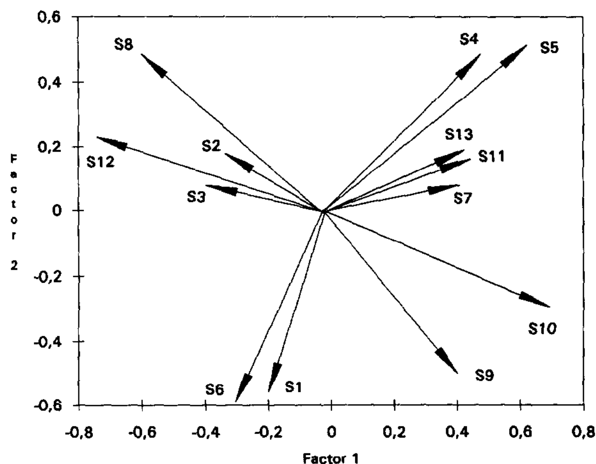

Fig. 5. Projection of the original variables (areas of the thirteen integration section) on the plane of the two first canonical factors of the discriminating analysis.

CONCLUSIONS

The different statistical analyses performed on RP-HPLC data led to coherent conclusions. Therefore, multi-dimensional analysis applied to chromatograms obtained by RP-HPLC is a method particularly suitable for the discrimination and classification of wheat cultivars. This confirms the possibility of applying statistical methods currently used for the treatment of other continuous signals, as in infrared spectrometry [12]. However, well controlled chromatographic conditions are necessary to perform statistical analysis. Nevertheless, this method needs to be validated taking into account growing year and location; such studies are in progress.

ACKNOWLEDGEMENTS

We thank Mr. J. Koenig (INRA, Clermont-Ferrand) for supplying cultivar samples and Dr. G. Branlard (INRA, Clermont-Ferrand) for stimulating discussions.

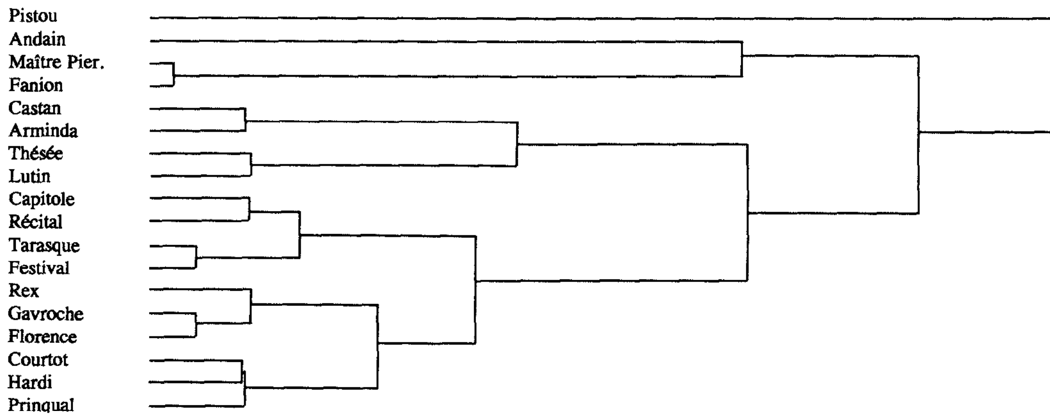

Fig. 6. Dendrogram resulting from a cluster analysis (Ward's method) on the thirteen integration sections of the eighteen cultivars.

REFERENCES

1 J. A. Bietz, *J. Chromatogr.*, 255 (1983) 219–238.
2 T. Burnouf and J. A. Bietz, presented at the *2nd International Symposium on Biochemical Approaches to Identification of Cultivars and Their Properties, Braunschweig, May 5–9, 1985.*
3 M. G. Scanlon, H. D. Sapirstein and W. Bushuk, *Cereal Chem.*, 66 (1989) 112–116.
4 F. R. Huebner and J. A. Bietz, *J. Chromatogr.*, 327 (1985) 333–342.
5 F. R. Huebner and J. A. Bietz, *Cereal Chem.*, 64 (1987) 15–20.
6 B. A. Marchylo, J. E. Kruger and D. W. Hatcher, *J. Cereal. Sci.*, 9 (1989) 113–130.
7 J. A. Bietz and T. Burnouf, *Theor. Appl. Genet.*, 70 (1985) 599–609.
8 J. A. Bietz and F. R. Huebner, in R. Lasztity and F. Bekes (Editors), *Proceedings of the 3rd International Workshop on Gluten Proteins*, World Scientific, Singapore, 1987, pp. 413–427.
9 M. G. Scanlon, H. D. Sapirstein and W. Bushuk, *Cereal Chem.*, 66 (1989) 439–443.
10 H. D. Sapirstein and W. Bushuk, *Cereal Chem.*, 62 (1985) 372–377.
11 J. C. Autran and P. Abbal, *Electrophoresis*, 9 (1988) 205–213.
12 D. Bertrand, P. Robert and W. Loisel, *J. Sci. Food Agric.*, 36 (1985) 1120–1124.
13 D. Bertrand, P. Courcoux, J. C. Autran, R. Meritan and P. Robert, *J. Chemometr.*, 4 (1990) 143–151.
14 B. A. Marchylo, D. W. Hatcher and J. E. Kruger, *Cereal Chem.*, 65 (1988) 28–40.
15 Y. Popineau and F. Pineau, *Lebensm. Wiss. Technol.*, 18 (1985) 133–135.
16 Y. Popineau and F. Pineau, *J. Cereal Sci.*, 3 (1985) 363–368.
17 G. L. Lookhart and L. D. Albers, *Cereal Chem.*, 65 (1988) 222–227.
18 G. Branlard and M. Dardevet, *J. Cereal Sci.*, 3 (1985) 329–343.
19 J. C. Autran and A. Bourdet, *Ann. Amelior. Plantes*, 25 (1975) 277–301.
20 S. Endo, K. Okada, S. Nagao and B. L. D'Appolonia, *Cereal Chem.*, 67 (1990) 480–485.