R&D REPORT NO. 64

Evaluation of Image Analysis Method for Basmati Rice

Martin Whitworth, Tom Fearn and Sam Millar

1998

Campden & Chorleywood
Food Research Association

Campden & Chorleywood
Food Research Association

Director-General
Prof. C. Dennis
BSc, PhD, FIFST

Chipping Campden
Gloucestershire
GL55 6LD UK
Tel: +44 (0) 1386 842000
Fax: +44 (0) 1386 842100

R&D Report No 64
Final Report on MAFF Project Ref: AN1140

Evaluation of Image Analysis Method for Basmati Rice

Martin Whitworth, Tom Fearn* and Sam Millar

*Department of Statistical Science, University College London, Gower Street, London, WC1E 6BT

Abbreviations used in this report

CCFRA..... Campden & Chorleywood Food Research Association

CEERI ..... Central Electronics Engineering Research Institute

CIRAD..... Centre de coopération internationale en recherche agronomique pour le développement

FMBRA..... Flour Milling and Baking Research Association

GAFTA..... Grain and Feed Trade Association

IA..... Image Analysis

IRRI ..... International Rice Research Institute

MAFF ..... Ministry of Agriculture, Fisheries and Food

NIR ..... Near Infrared (spectroscopy)

RP3 ..... Rice Parameter 3

RVA..... Rapid Viscoanalyser

Summary

The incentive for adulteration of rice inadvertently or by dishonest traders has been recognised by MAFF. Particular authenticity issues concern the adulteration of Basmati rice with other types, and of US long grain rice with rice originating from other countries. As part of previous studies, an image analysis method was identified as suitable for identifying Basmati rice, although it was unsuitable for identifying long grain rice. A full protocol for this method was developed. This included a statistical analysis procedure in which comparison of a simple parameter with a cut-off value indicated the likely presence or absence of admixture at or above approximately the 20% level.

A more complex analysis procedure was also suggested in which the Basmati and non-Basmati components of a sample were modelled by normal distributions. By fitting a weighted sum of such distributions to measured data, their proportions may be determined, providing a quantitative estimate of the admixture. This method was demonstrated for 50-grain samples, for which the uncertainty of determination is high. 500-grain samples were instead recommended. However, reference data were only available for 46-grain samples, which were insufficient to ascertain the validity of the model when applied to larger samples.

The purpose of the current project was to measure further reference samples in larger quantities and to use these data to establish the possibility for quantitative estimation of the level of adulteration of Basmati rice, based on statistical methods such as that described above. 30 new samples were collected, including many of Indian origin. Measurements were made of at least 500 grains of each of 21 Basmati samples and 39 non-Basmati samples selected from the newly collected and existing reference samples.

The new data established that the distributions for reference samples were not normal and that the more complex statistical analysis identified above could not be applied. Further, the variability in the shape of the reference sample distributions indicated that any similar approach to the quantitative assessment of the level of adulteration would lead to inherently large margins of error and would represent little improvement on the simpler pass/fail criterion already developed.

The previously identified simple test criterion correctly classified all of the new reference Basmati samples and 32 of the 39 non-Basmati samples. Of the remaining 7 samples, 6 were of Indian origin. On the basis of the new reference data, the simple criterion was made more sensitive to mixtures. To assess its performance, mixtures were made of Basmati and non-Basmati samples in 9 combinations, each at 4 levels of admixture. 500 grains of each of these samples were tested. All of the mixtures were detected at levels of 20% addition or above, and 2 of the 9 mixtures were detected at a level of 10% non-Basmati content. However, the level of detection will vary depending on the characteristics of the adulterant rice.

The test would be useful as a screening test for rice traders in combination with other methods as a method of quality assurance. Although it is insufficiently definitive to be used as the basis of prosecutions, it could also be used by MAFF or others for survey purposes.

Contents

1. Introduction

Basmati rice is a class of rice grown in the Punjab region of India and Pakistan, and is recognised as a market sector in its own right. It can only be grown once a year with a yield half that of other rices and its eating quality cannot be duplicated by growing the same seed in other regions. As a consequence of its scarcity and its popularity in the UK, Basmati commands a high price compared to other rices. The incentive for adulteration was recognised by the Ministry of Agriculture, Fisheries and Food (MAFF) who in 1990 commissioned work at FMBRA (now CCFRA) to develop methods for determining the authenticity of Basmati rice, and by the Grain and Feed Trade Association (GAFTA) whose Rice Standards Section introduced a Code of Practice in January 1992.

The work commissioned by MAFF on the separation of rice proteins by polyacrylamide gel electrophoresis (Scanlon et al., 1991) was only partially successful and it was found that, unlike the gliadin and hordein proteins commonly used to identify varieties of wheat and barley, the major rice proteins do not give fingerprint patterns in gel electrophoresis or HPLC. Attempts were made to derive patterns from the minor prolamine fraction of rice proteins. However, this requires very high resolution and sensitivity which is difficult to reproduce in routine operation. Furthermore, Basmati is defined by origin as well as genotype. Gene-specific components would be expected to be consistent across samples of the same genotype from different origins. As a result of these considerations, further research focused on objective physical methods.

In further work, CCFRA evaluated the methods of image analysis (IA), the Rapid Viscoanalyzer (RVA) and Near Infrared spectroscopy (NIR) for the determination of Basmati rice authenticity (Osborne *et al.*, 1993). 132 samples of milled rice of known provenance were collected, of which 89 were tested by all three techniques to provide a database against which samples of unknown provenance could be compared. All three methods showed some ability to discriminate the Basmati samples from the other types in the database.

Additional work (Whitworth *et al.*, 1995) based on measurements of a further 95 samples showed that the RVA method was insufficiently sensitive, particularly for testing of mixed samples, and that the NIR calibrations previously developed were not robust when applied to the new samples. The IA method, however, continued to show success, although a bias in the results had occurred between the two studies. Combination of IA results with NIR or RVA data gave little improvement over the IA results alone. Further work (Whitworth *et al.*, 1996) was conducted to improve the IA method and to evaluate its performance in more detail. The analysis was transferred to a more reliable IA system and a validation protocol was introduced to provide control charts that would provide warning of any bias such as that previously seen. A simple statistical parameter was developed which, when compared with a threshold value, provided a pass / fail discrimination between authentic Basmati samples and others. This was shown to detect adulteration with a typical detection limit of about 20%, although the sensitivity varied depending on the type of adulterant and some types would not be detected at all. By varying the threshold, the sensitivity could be increased, but only at the risk of an increased probability of false positives.

Within this study, a further statistical analysis technique was also tested, based on representing a test sample as a mixture of a pure Basmati component and an adulterant component modelled by normal distributions. Based on reference measurements of pure samples, determinations were made of the distributions and relative proportions of these components which provided the best fit to the measured data. This method was capable of giving a quantitative estimate of the admixture, but could only be applied with validity to the testing of 50 grain samples, corresponding to the size of reference samples which had been measured. For samples of this size, the uncertainty in determining the level of admixture is high, and it was considered that 500 grain samples would provide a better compromise between adequate sampling and speed of testing.

The aim of the current project was to obtain measurements of 500 grains for each of at least 50 reference samples, to use these to develop the quantitative statistical analysis for application to 500 grain test samples, and to test its performance by using it to test mixed samples of known composition.

2. Materials and methods

2.1 Collection of authentic rice samples

Table 1 - Rice samples collected during this project

Source	Growing region	CCFRA Sample number
IRRI	Philippines	**PSM97/28**
Ricetec Inc., USA	USA	PSM97/61
	USA	PSM97/62
	USA	PSM97/63
	USA	PSM97/64
Directorate of Rice Research, Hyderabad, India	-	CMS/97/213
	-	CMS/97/214
	-	CMS/97/215
	-	CMS/97/216
	-	CMS/97/217
	-	CMS/97/218
	-	CMS/97/219
	-	CMS/97/220
	-	CMS/97/221
Gafta	Australia	PSM96/91
	Australia	**PSM96/92**
	Australia	**PSM96/93**
	Amritsar (Punjab), India	**PSM96/94**
	India	**PSM96/95**
	India	**PSM96/96**
	Amritsar (Punjab), India	**PSM96/97**
	India	**PSM96/98**
	Haryana	**PSM96/99**
	India	**PSM96/100**
	Haryana / Uttar Pradesh, India	**PSM96/101**
	India	**PSM96/102**
	India	**PSM96/103**
	India	**PSM96/104**
	-	PSM96/105
	-	PSM96/106
	Haryana / Uttar Pradesh	**PSM96/107**

* This sample was infested when received and was therefore disposed of.

Samples shown in bold were analysed within this project.

A set of authentic rice samples were available from previous projects (Whitworth *et al.*, 1995; Whitworth *et al.*, 1996). In order to extend this, requests were sent to several research institutes and companies for additional samples. A list of the new samples received during this project is shown in Table 1. Most of the samples received in response to requests made within this project were received too late to be included in the list of samples tested, most of which were instead selected from among samples collected during previous studies. Table 1 also lists a set of samples received from GAFTA during 1996. Many of these samples were in the form of paddy rice and were therefore milled at the Natural Resources Institute, Chatham Maritime, Kent to provide samples suitable for testing. From the rice samples available, the samples listed in Tables 2a and 2b were selected for measurement within this study. 500 grains of each of these samples were selected randomly and tested as pure reference samples.

Table 2a - Basmati rice samples tested in pure form

Sample No.	Origin	Rice type	Date of analysis
MR94/432	India	Basmati	22/11/96,28/11/96
MW337	Pakistan	Basmati	22/11/96,28/11/96
MR94/434	India	Basmati	22/11/96,28/11/96
MW339	Pakistan	Basmati	25/11/96,28/11/96
MW335	Pakistan	Basmati	25/11/96,28/11/96
MW338	Pakistan	Basmati	25/11/96,28/11/96
MW336	Pakistan	Basmati	25/11/96,17/01/97
MR94/431	India	Basmati	13/12/96,17/1/97
MR95/143	India	Basmati	23/1/97
MR95/140	India	Basmati	23/1/97
MW324	Philippines	Basmati	24/1/97
PSM/97/28	Philippines	Basmati	6/3/97
MR/94/433	India	Basmati	12/3/97
PSM/96/94	India	Basmati	27/3/97
PSM/96/95	India	Basmati	27/3/97, 8/4/97
PSM/96/96	India	Basmati	8/4/97
PSM/96/97	India	Basmati	8/4/97
PSM/96/98	India	Basmati	8/4/97, 15/4/97
PSM/96/101	India	Basmati	2/5/97
PSM/96/104	India	Basmati	2/5/97
PSM/96/107	India	Basmati	11/6/97

Table 2b - Non-Basmati rice samples tested in pure form

Sample No.	Origin	Rice type	Date of analysis
MR95/85	Camargue, France	Long grain	25/11/96
MR95/142	India	Non-Basmati	13/12/96
MR95/130	Italy	Non-Basmati	16/1/97
MW801	Spain	Non-Basmati	16/1/97
MW110	Laos	Non-Basmati	16/1/97
MR95/134	Italy	Non-Basmati	17/1/97
MW316	Italy	Non-Basmati	17/1/97
MW104	Camargue, France	Non-Basmati	23/1/97
MR95/150	Italy	Non-Basmati	24/1/97
MW302	Australia	Non-Basmati	24/1/97
MW330	Philippines	Non-Basmati	24/1/97
MR94/1905	Texas	US long grain	20/2/97
MR94/1909	Arkansas	US long grain	20/2/97
MR94/1913	Mississippi	US long grain	14/2/97
MR94/1914	Texas	US long grain	20/2/97
MR94/1917	Louisiana	US long grain	20/2/97
MR94/1919	Arkansas	US long grain	21/2/97
MR94/1923	Mississipi	US long grain	21/2/97
MR94/1933	Mississipi	US long grain	14/2/97,20/2/97
MR94/1924	Texas	US long grain	27/2/97
MR94/1929	Arkansas	US long grain	27/2/97
MR94/1935	Louisiana	US long grain	27/2/97
MR94/1938	Mississippi	US long grain	28/2/97
MR94/1941	Mississippi	US long (Aromatic)	28/2/97
MR94/1942	Texas	US long (Aromatic)	28/2/97
MR94/1946	Arkansas	US long (Aromatic)	28/2/97
PSM/96/102	India	Non-Basmati	10/3/97
PSM/96/103	India	Non-Basmati	10/3/97
PSM/96/93	Australia	Non-Basmati	10/3/97
PSM/96/92	Australia	Non-Basmati	10/3/97
MW705	Thailand	Non-Basmati	11/3/97
MR/94/436	India	Non-Basmati	12/3/97
MR/94/435	India	Non-Basmati	12/3/97
MR/95/141	India	Non-Basmati	21/3/97
MW350	India	Non-Basmati	21/3/97
MR/95/128	Italy	Non-Basmati	21/3/97
MW342	Bangladesh	Non-Basmati	21/3/97
PSM/96/99	India	Non-Basmati	15/4/97
PSM/96/100	India	Non-Basmati	15/4/97

2.2 Preparation of mixed samples

2.2.1 Composition

Additional measurements were made of mixed samples. Each mixture was comprised of one of three Basmati samples mixed with one of three other samples. Details of the samples used are shown in Table 3. A major consideration in the selection of these samples was the availability of sufficient material to make all the required mixtures. However, emphasis was placed on ensuring that the non-Basmati samples included examples from India, which it was thought would be more typical of the types of adulterant likely to be encountered in practice. All nine possible pairings of the three Basmati and three other samples were made at each of four levels (5, 10, 20, 30%) of adulteration, yielding 36 mixtures in total. Details of these mixtures and the sample numbers assigned to them are shown in Table 4.

Table 3 - Samples used to make mixtures

Sample number	Geographical origin	Source	1000 grain weight (g)
Basmati samples			
PSM/96/107	India	Gafta	17.13
PSM/96/97	India	Gafta	14.91
PSM/96/104	India	Gafta	15.38
Non Basmati samples			
PSM/96/100	India	Gafta	19.58
MW801	Spain	Institute Agroquimica	17.44
MR/94/435	India	Directorate of Rice Research, Hyderabad	16.39

2.2.2 Removal of broken grains

The image analysis procedure is unsuitable for the measurement of broken grains and these are therefore removed by hand prior to analysis. If the broken grain content of constituents of a mixed sample differ, such removal would alter the proportions of these samples in the grains actually measured. Therefore, prior to making mixtures, broken grains were removed from the six samples to be mixed.

Table 4 - Mixed rice samples made and tested

Adulterant	% added	Basmati sample					
		PSM/96/107		PSM/96/97		PSM/96/104	
		Sample # CMS 97/..	Date tested	Sample # CMS 97/..	Date tested	Sample # CMS 97/..	Date tested
PSM/96/100	5	533	17/9/97	641	22/9/97	653	29/9/97
	10	534	17/9/97	642	22/9/97	654	29/9/97
	20	535	17/9/97	643	22/9/97	655	29/9/97
	30	632	17/9/97	644	22/9/97	656	29/9/97
MW801	5	633	18/9/97	645	24/9/97	657	30/9/97
	10	634	18/9/97	646	24/9/97	658	30/9/97
	20	635	18/9/97	647	24/9/97	659	30/9/97
	30	636	18/9/97	648	24/9/97	660	30/9/97
MR/94/435	5	637	19/9/97	649	26/9/97	661	1/10/97
	10	638	19/9/97	650	26/9/97	662	1/10/97
	20	639	19/9/97	651	26/9/97	663	1/10/97
	30	640	19/9/97	652	26/9/97	664	1/10/97

2.2.3 Weighing and mixing of samples

After removal of broken grains, the 1000 grain weights of the remaining samples of intact grains were measured using an automatic seed counter and are given in Table 3. Because these values differ for each sample, it is necessary to distinguish whether the samples were mixed by mass or by number of grains. The results generated by the image analysis procedure are produced by number of grains and it was therefore decided to mix samples to consistent proportions of 5%, 10%, 20% and 30% by number of non-Basmati grains. To facilitate simple preparation of the mixtures, the 1000 grain weight of each sample was used to calculate the appropriate mass of the sample required to achieve the desired number of grains in the mixture, these masses were weighed out and the mixtures were made.

For example, to achieve 5% of non Basmati grains in a 500 grain sample of CMS/97/533, the sample should comprise:

$$95\% \times 500 = 475 \text{ grains of PSM/96/107}$$
$$+ \quad 5\% \times 500 = 25 \text{ grains of PSM/96/100}$$

PSM/96/107 has a 1000 grain weight of 17.13g and PSM/96/100 has a 1000 grain weight of 19.58g. Therefore the mixture is made using:

$$475 \times (17.13/1000) = 8.14\text{g of PSM/96/107}$$
$$+ \quad 25 \times (19.58/1000) = 0.49\text{g of PSM/96/100}$$

Ideally, it would be desirable to test directly the effects of sampling from a mixture by preparing large quantities of each mixture and subsampling 500 grains for testing. However, this was not possible due to the limited amounts of each sample available. Therefore samples were mixed to a target of 500 grains per sample, and the entire sample was tested. In practice, it was found that the 1000 grain weights measured using the automatic seed counter tended to underestimate the actual number of grains produced, and the sample quantities weighed out were therefore increased by 10%.

2.3 Analysis of rice samples

Rice samples were tested according to an image analysis procedure previously described in detail by Whitworth *et al.*, 1996. A summary of the procedure is given below.

For each sample, approximately 500 grains were analysed in two batches of approximately 250 grains each.

2.3.1 Cooking

The grains were placed in distilled water maintained at a temperature of at least 96°C by a bath of boiling water. The grains were cooked in this way for 10 minutes. They were then removed from the water bath, any grains which had stuck together or to the beaker were gently separated and the distilled water was decanted off. The grains were then rinsed with boiling distilled water followed by cold distilled water. They were then tipped onto a wet paper towel and analysed.

2.3.2 Image analysis

Grains were manually placed onto a lightbox such that their planes of symmetry were parallel to the surface and such that they did not touch. A silhouette image was captured using a colour CCD camera mounted above the lightbox on a copy stand. Images were analysed using a binary threshold to identify grains. Each grain image was then automatically measured to determine its caliper length, caliper breadth, area and perimeter. From these values, a discriminant parameter, Rice Parameter 3 (RP3) was

calculated for each grain. Further grains were placed on the lightbox and analysed until all grains had been measured.

2.3.3 Calibration

The magnification of the image analysis system was set up consistently on each occasion that it was used. At the start of each analysis, a clear plastic ruler (previously checked against a calibrated metal ruler) was placed on the light box and the distance between two graduations was identified on the image to establish the actual magnification factor. Prior to analysis of each sample, the aperture of the camera was adjusted to give a consistent image brightness. An image of the light box was captured and then used to normalise each image of rice grains prior to analysis, thus ensuring consistent thresholding of the grains from the background.

As a check on the reliability of the measurement system, a set of reference nylon cylinders were measured prior to each set of analyses and at least on every day when analysis was carried out. These cylinders have similar dimensions and optical density to rice grains. They were measured using the image analysis system to give RP3 values which were recorded on a control chart.

3. Results

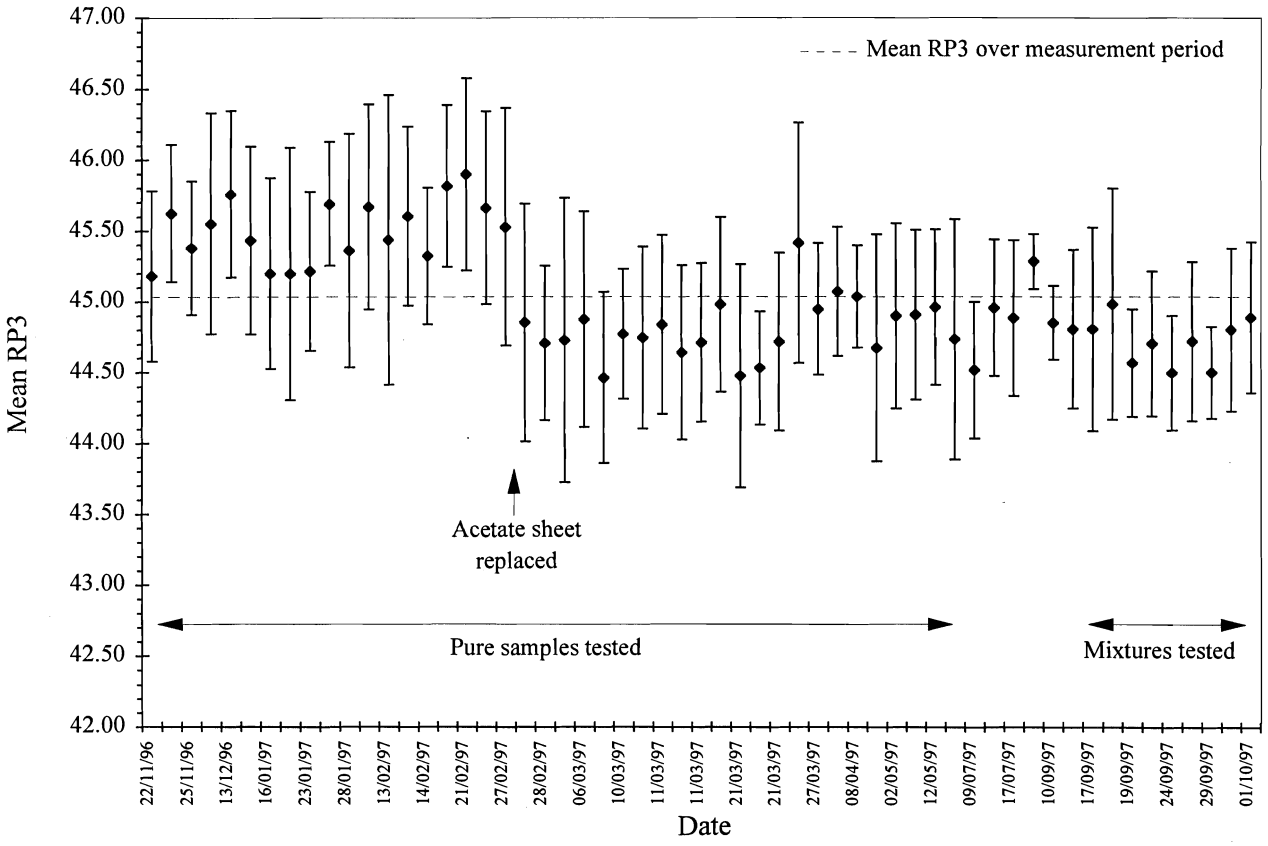
3.1 Calibration

Before analysis of rice samples, the calibration of the image analysis system was tested using nylon cylinders as described by Whitworth *et al.*, 1996. Figure 1 shows all the calibration results recorded over the period 22 November 1996 to 1 October 1997, during which the rice samples for this project were measured (c.f. dates of testing given in Tables 2 and 4). It can be seen that a decrease in the results occurred on 28 February 1997 from a mean RP3 value of 45.50 units over the period before this date to a mean of 44.80 units from this point onwards. On this date, a slight crease had been noticed in the acetate sheet placed on the lightbox to keep it clean, and it had therefore been replaced. It is thought that this is the most likely explanation for the observed effect and that the creased sheet had been causing grains and calibrants to lie slightly clear of the surface of the lightbox, causing a small change in their magnification. Relative to the differences

in RP3 values typically seen for different rice types, the effect is small, and it was considered that replacement of the acetate sheet represented a return to the intended conditions of the test. However, the experience identified the condition of the acetate sheet as a potentially important element of the test, which is now checked regularly. Apart from this small effect, the calibration procedure indicated that the image analysis measurement remained stable over the period of testing.

Figure 1 - Control chart for the image analysis system

(N.B. Bars represent standard deviations of single calibrant results, not standard errors)



3.2 Rice Parameter 3 measurements

3.2.1 Pure samples

Measurements of Rice Parameter 3 (RP3) have been made for between 501 and 520 grains of each of the 21 Basmati samples and 39 non Basmati samples listed in Table 2. Histograms of the RP3 values measured for each sample are given in Figures 2a and 2b. The distribution of RP3 values has been statistically assessed for each of the samples to

Figure 2a - Distributions of Rice Parameter 3 for pure Basmati samples

Each histogram shows the number of grains for each 2 unit interval of RP3.

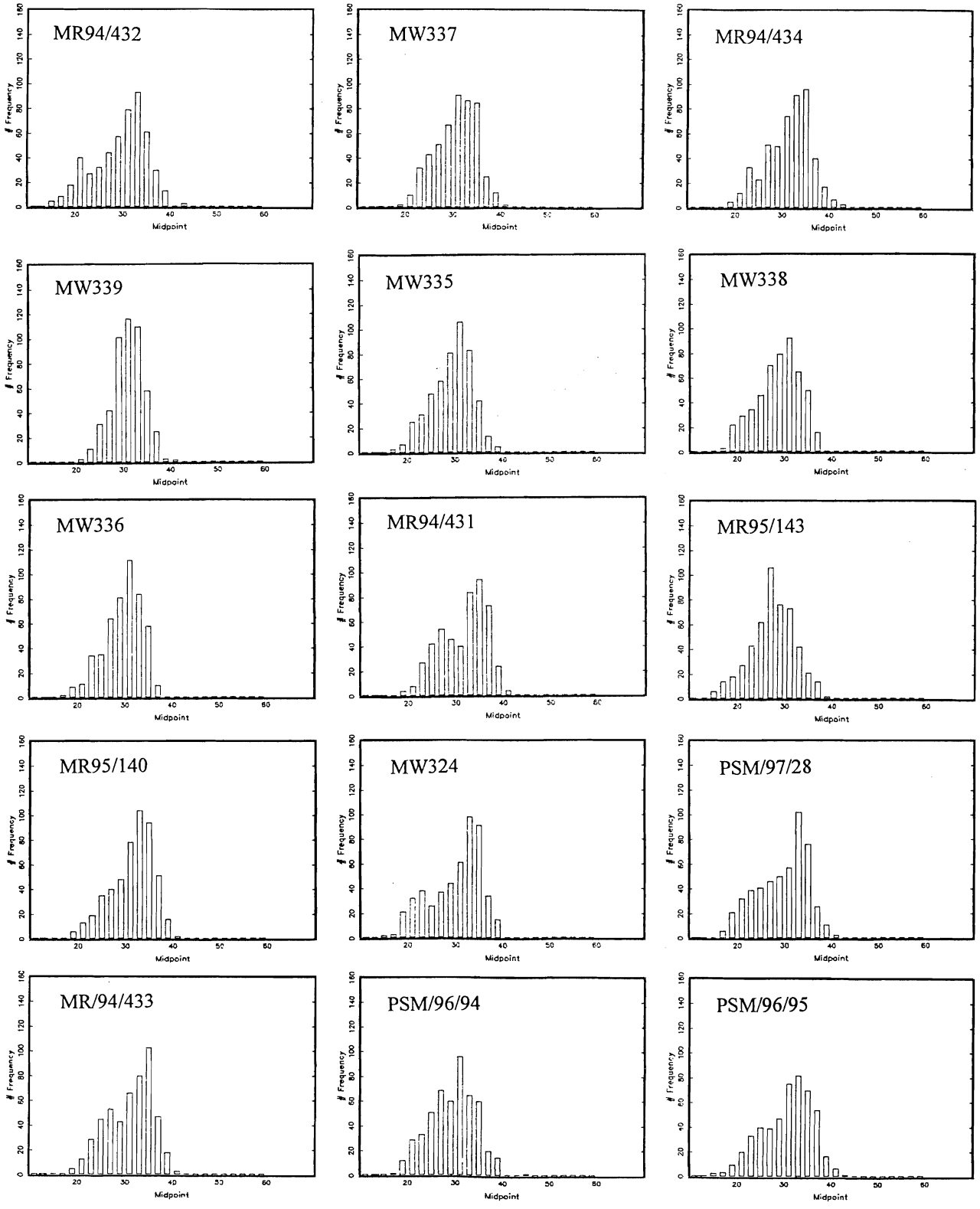


Figure 2a - continued

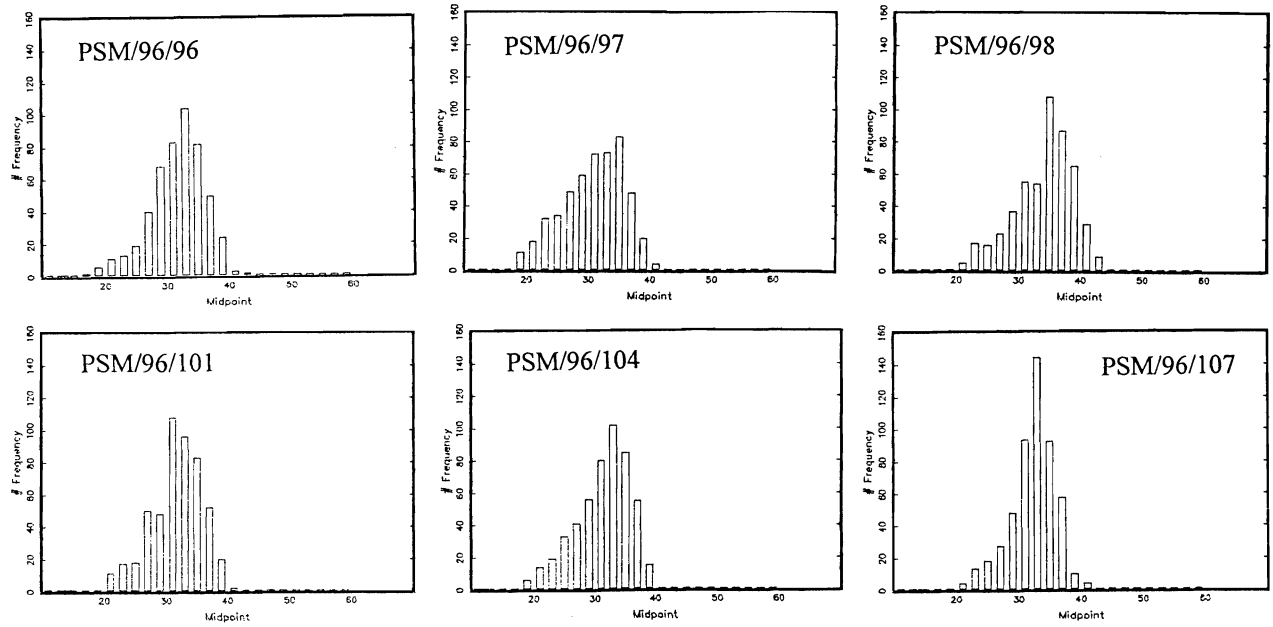


Figure 2b - Distributions of Rice Parameter 3 for pure non-Basmati samples

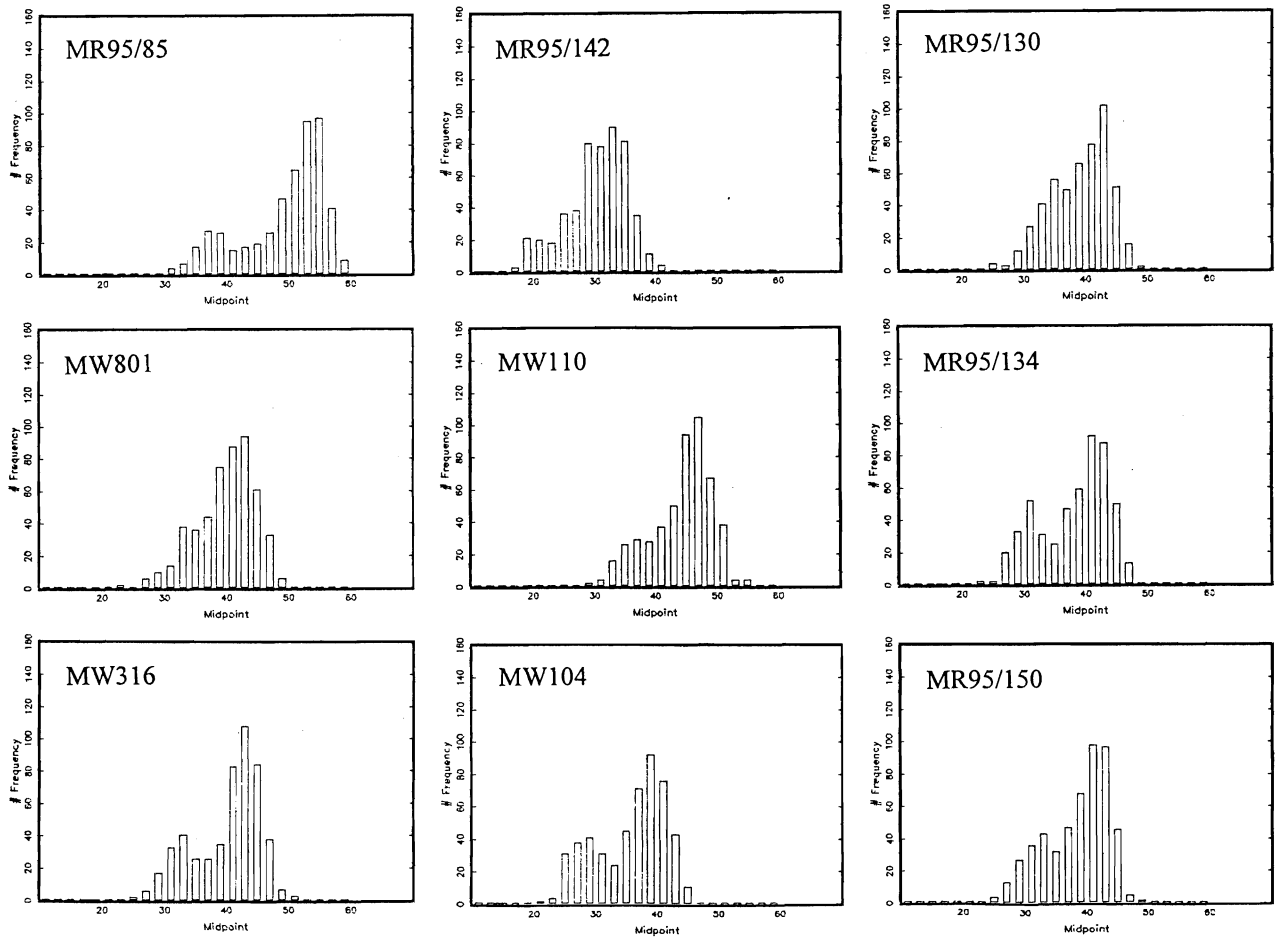


Figure 2b - continued

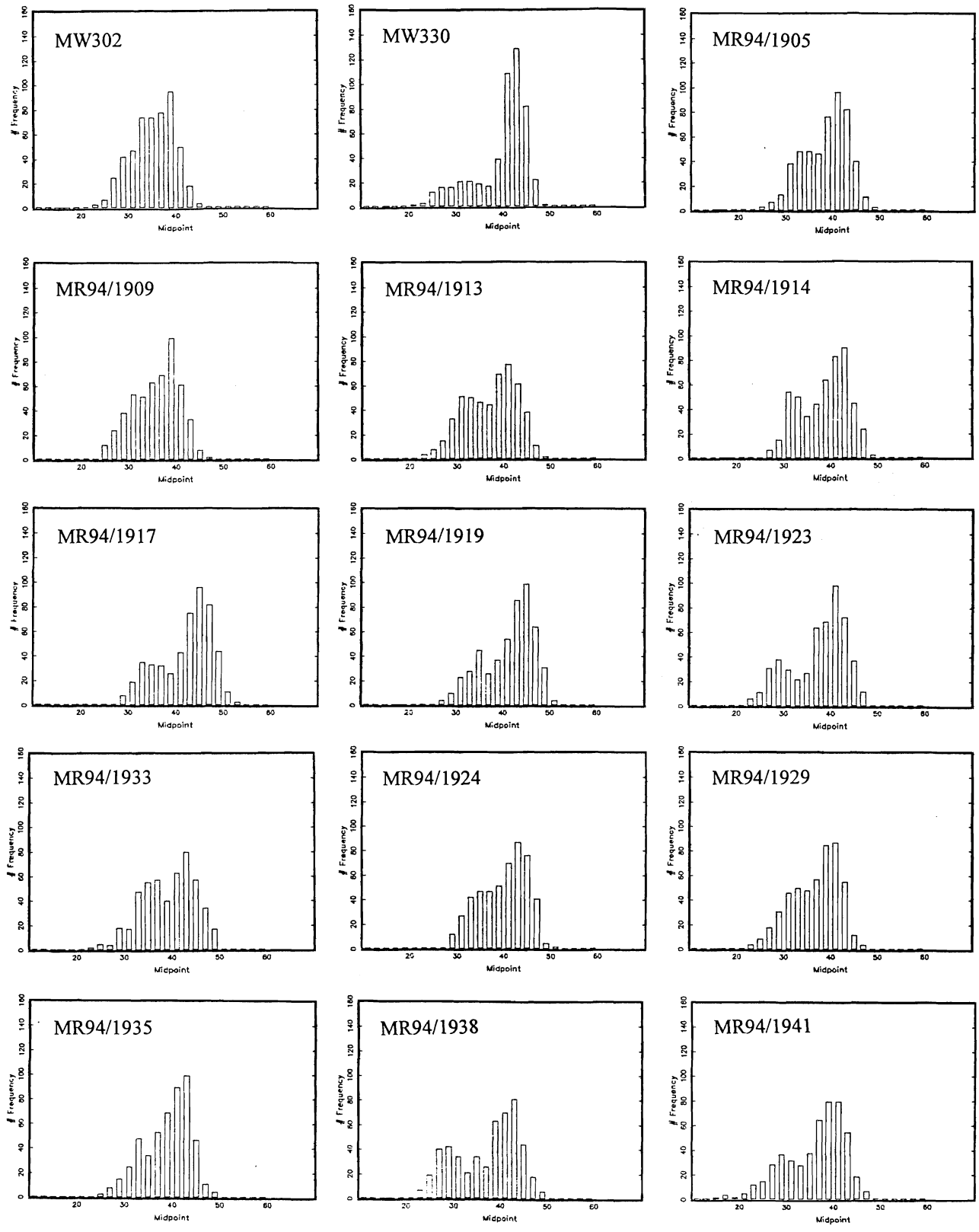
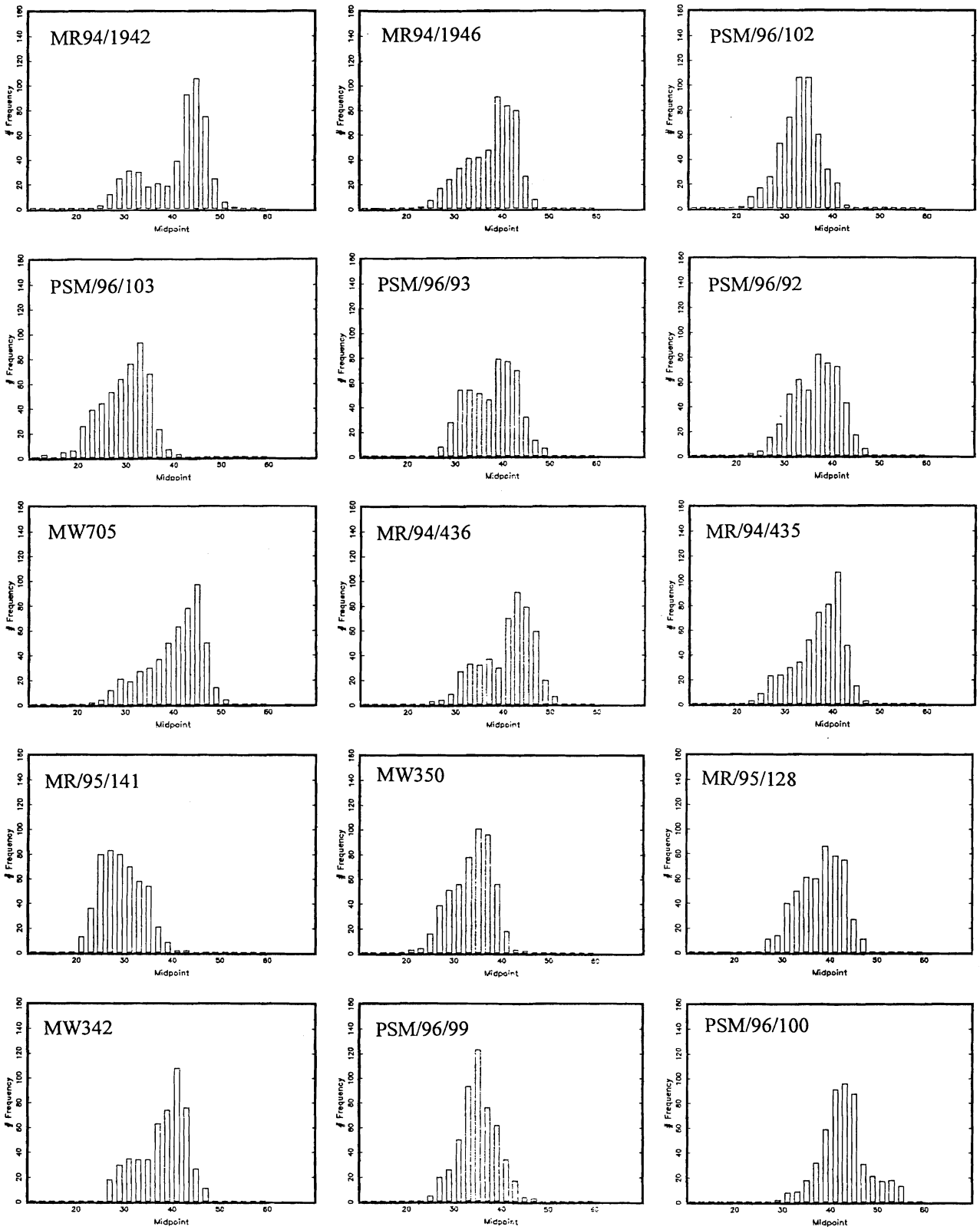


Figure 2b - continued



provide information on the characteristics of each population for use in statistical tests for the authenticity of unknown samples. The sample size of 500 grains per sample represents an improvement on previous studies in which only 46 grains were tested in most cases, and allows the distributions to be characterised more precisely. As had been demonstrated in previous studies, the variance of the RP3 distributions was similar for Basmati and non-Basmati samples, but the former typically had a lower median. The median RP3 value for each pure sample is given later, in Tables 5a and 5b. In previous work, RP3 distributions of pure samples were approximated as normal distributions for the purpose of fitting to distributions measured for test samples. On the basis of the more comprehensive data collected in this project, it is apparent that this assumption is invalid and that many of the distributions are skewed, and in several cases, are bimodal. For both Basmati and non-Basmati samples, the skew is generally negative and for bimodal distributions, the stronger component occurs at higher values of RP3 than the weaker one. When this phenomenon was first observed, it was thought that it might be caused by variations in the orientation in which rice grains were placed on the lightbox for analysis. As described previously (Whitworth *et al.*, 1996), the procedure specifies that grains should be placed with their symmetry plane parallel to the lightbox surface. If they are instead placed with their symmetry plane vertical, they form a narrower silhouette and produce lower RP3 values. Failure to follow the procedure correctly could therefore result in a bimodal distribution corresponding to grains in these two orientations, or could result in a skewed distribution caused by a range of grain orientations. When the bimodal distributions were first observed, this possibility was considered and particular attention was subsequently devoted to ensuring correct orientation of the grains. Even with this attention, skewed distributions were still obtained, and although variations in grain orientation may have have been partially responsible for some of the bimodal distributions, they could not account for the entire effect. It was concluded that the previous approximation of the RP3 distributions by normal distributions was inappropriate for accurate quantitative analysis. Although the median RP3 values are generally distinct for Basmati and non-Basmati samples, the major component of the RP3 distribution for Basmati samples coincides with the minor component of the bimodal distribution for non-Basmati samples, frustrating efforts at quantitative determination of the proportions of rice types in mixtures.

3.2.2 Mixtures

Measurements of RP3 were obtained for between 500 and 515 grains of each of the 36 mixed samples made. These are shown as histograms in Figures 3a, b and c. Each figure shows the results for mixtures using a single Basmati sample and each column of histograms represents addition of a different adulterant to this sample. The level of addition increases down each column. It can be seen in most cases that as the level of adulteration increases, the distribution of RP3 results develops a shoulder at the high RP3 end due to the measurements contributed by the adulterant grains. In some cases, particularly for high levels of adulteration, the distribution becomes noticeably bimodal. This alteration of the distribution provides the basis for statistical tests designed to determine the presence or level of adulteration, although the variably skewed and bimodal nature of the RP3 distributions for pure samples limits the extent to which such tests could be made quantitative.

3.3 Statistical analysis of results

3.3.1 Previously devised pass / fail criterion

In a previous study (Whitworth *et al.*, 1996), a simple pass/fail criterion was devised to indicate the presence of non-Basmati rice in a sample based on a set of RP3 measurements. A sample was considered inconsistent with pure Basmati rice if:

$$Q(50) > 36.5 \quad \text{or} \quad T = Q(90)-Q(50) > 8$$

where Q(50) and Q(90) are the 50th percentile (i.e. the median) and the 90th percentile of the RP3 distribution respectively. The Q(50) criterion is intended to identify when a major component of the sample is inconsistent with Basmati samples and the T criterion is designed to identify smaller quantities of non-Basmati rice within a Basmati sample. Results of this analysis for the samples analysed in this study are shown in Table 5 as Q(50) and T values, the extent by which these exceed the cut off values of 36.5 and 8 respectively, and the conclusion as to whether the results satisfy the criterion (A) for consistency with Basmati rice. These results are also shown as a plot of T against Q(50) in Figure 4.

Figure 3a - Distributions of Rice Parameter 3 for mixed samples

Base Basmati Sample: PSM/96/107

Adulterant Sample:	PSM/96/100	MW801	MR/94/435
5% admixture	CMS/97/533	CMS/97/633	CMS/97/637
10% admixture	CMS/97/534	CMS/97/634	CMS/97/638
20% admixture	CMS/97/535	CMS/97/635	CMS/97/639
30% admixture	CMS/97/632	CMS/97/636	CMS/97/640

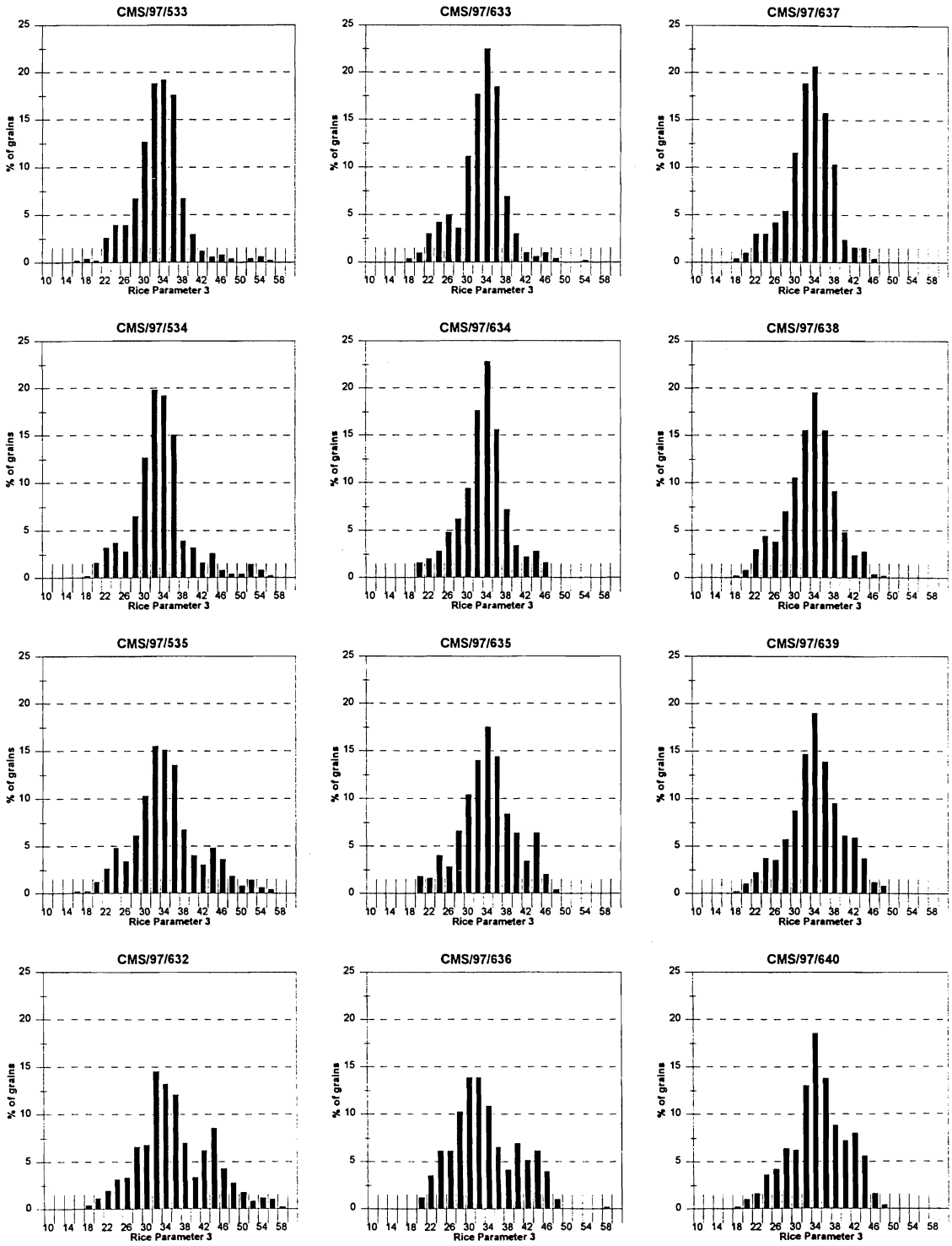


Figure 3b - Distributions of Rice Parameter 3 for mixed samples

Base Basmati Sample: PSM/96/97

Adulterant Sample:	PSM/96/100	MW801	MR/94/435
5% admixture	CMS/97/641	CMS/97/645	CMS/97/649
10% admixture	CMS/97/642	CMS/97/646	CMS/97/650
20% admixture	CMS/97/643	CMS/97/647	CMS/97/651
30% admixture	CMS/97/644	CMS/97/648	CMS/97/652

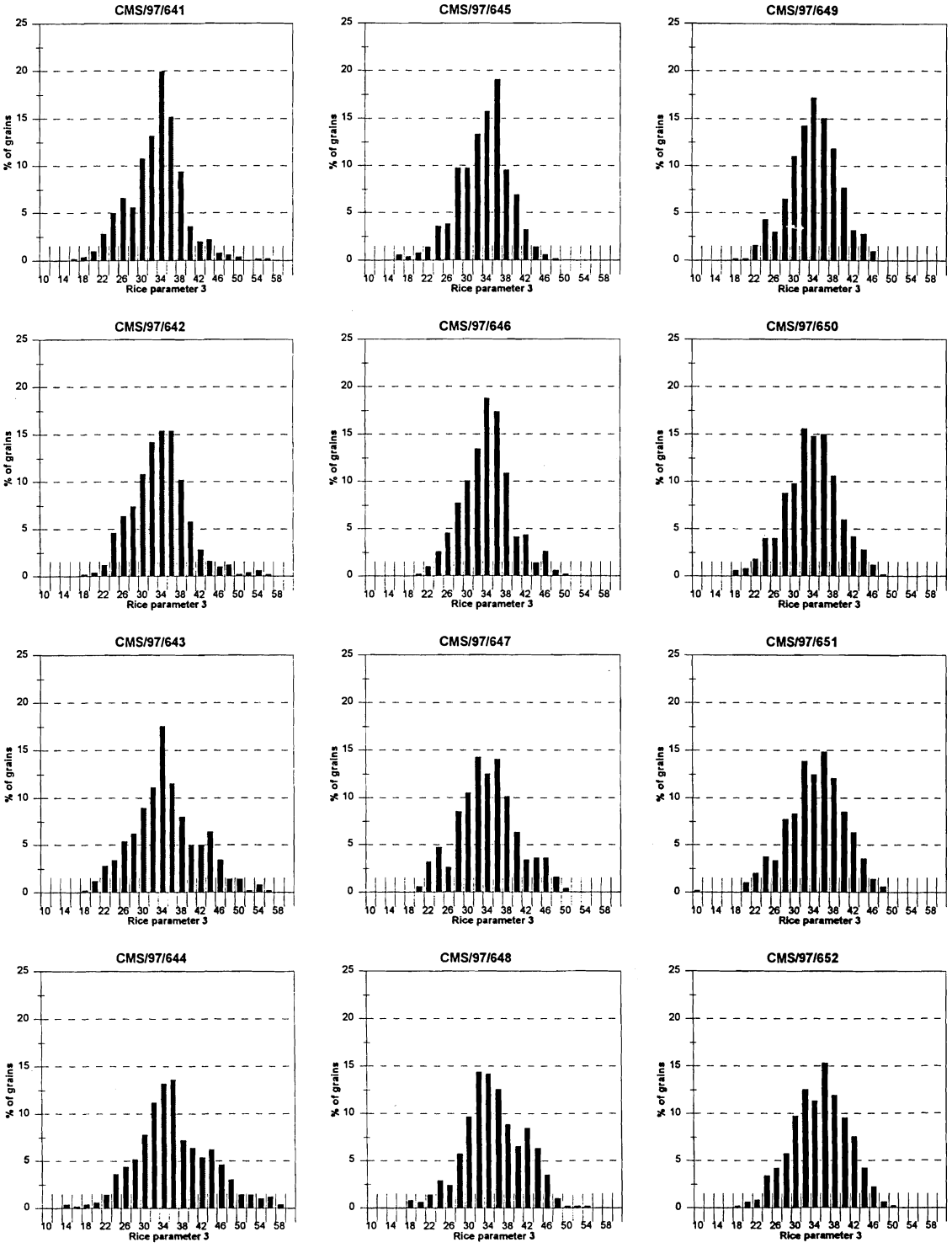


Figure 3c - Distributions of Rice Parameter 3 for mixed samples

Base Basmati Sample: PSM/96/104

Adulterant Sample:	PSM/96/100	MW801	MR/94/435
5% admixture	CMS/97/653	CMS/97/657	CMS/97/661
10% admixture	CMS/97/654	CMS/97/658	CMS/97/662
20% admixture	CMS/97/655	CMS/97/659	CMS/97/663
30% admixture	CMS/97/656	CMS/97/660	CMS/97/664

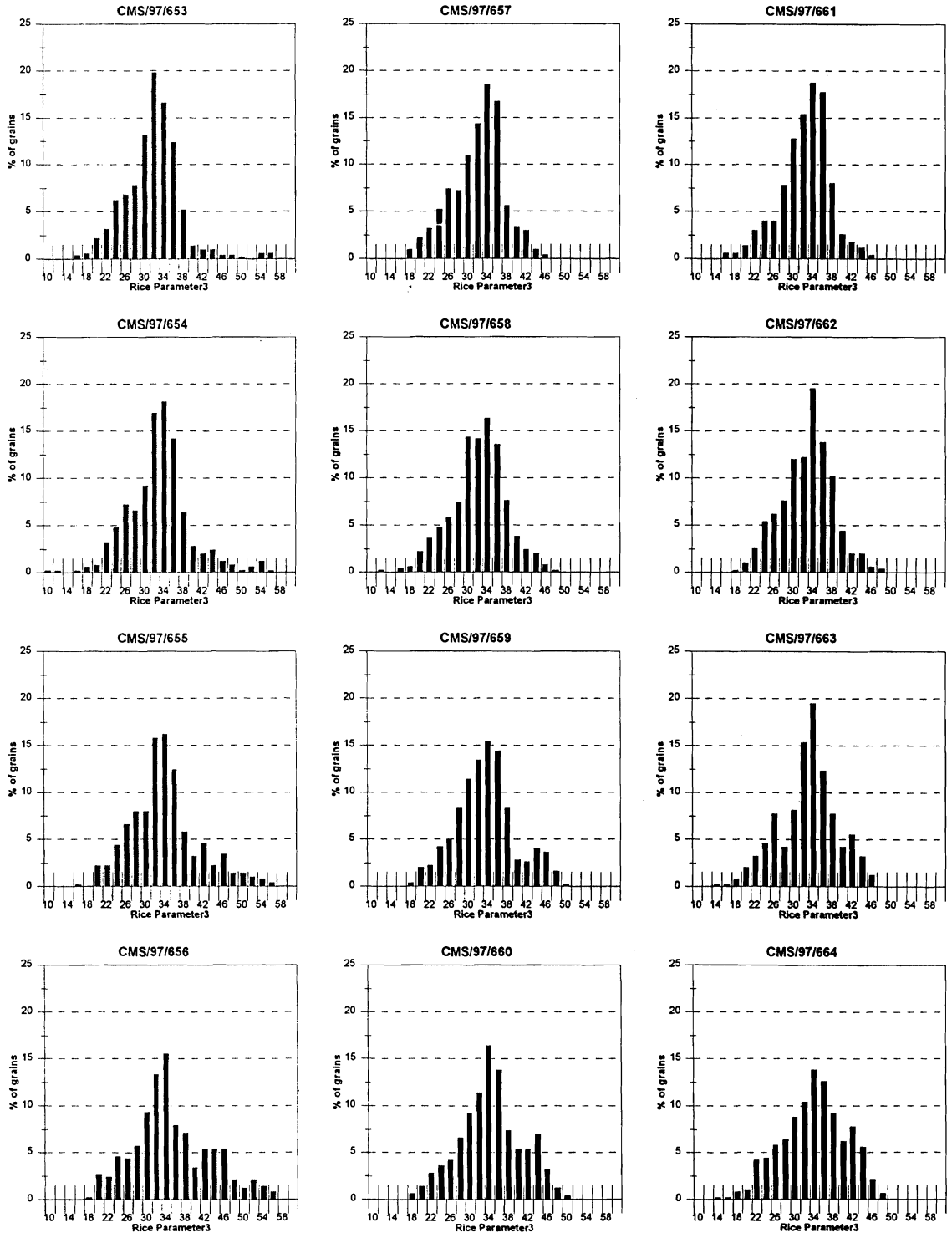


Table 5a - Results of a pass/fail test for consistency with Basmati.

Pure Basmati samples

Criterion A: Consistent with Basmati if Q(50)-36.5 ≤ 0 and T-8 ≤ 0

Criterion B: Consistent with Basmati if Q(50)-36.5 ≤ 0 and T-6.5 ≤ 0

Sample	Q(50)	T	Q(50) -36.5	T-8	T-6.5	Consistent with Basmati?		Sample	Q(50)	T	Q(50) -36.5	T-8	T-6.5	Consistent with Basmati?	
						A	B							A	B
MR94/432	30.58	5.02	-5.92	-2.98	-1.48	✓	✓	PSM/97/28	30.92	4.79	-5.58	-3.21	-1.71	✓	✓
MW337	31.15	4.61	-5.35	-3.39	-1.89	✓	✓	MR/94/433	31.91	4.71	-4.59	-3.29	-1.79	✓	✓
MR94/434	32.13	4.39	-4.37	-3.61	-2.11	✓	✓	PSM/96/94	30.02	5.24	-6.48	-2.76	-1.26	✓	✓
MW339	31.07	4.19	-5.43	-3.81	-2.31	✓	✓	PSM/96/95	31.48	5.31	-5.02	-2.69	-1.19	✓	✓
MW335	29.92	4.53	-6.58	-3.47	-1.97	✓	✓	PSM/96/96	32.33	4.74	-4.17	-3.26	-1.76	✓	✓
MW338	29.20	5.26	-6.58	-2.74	-1.24	✓	✓	**PSM/96/97**	31.44	5.16	-5.06	-2.84	-1.34	✓	✓
MW336	30.35	4.12	-6.15	-3.88	-2.38	✓	✓	PSM/96/98	34.98	4.63	-1.52	-3.37	-1.87	✓	✓
MR94/431	32.68	4.60	-3.82	-3.40	-1.90	✓	✓	PSM/96/101	32.02	4.94	-4.48	-3.06	-1.56	✓	✓
MR95/143	27.38	5.98	-9.12	-2.02	-0.52	✓	✓	**PSM/96/104**	32.08	4.50	-4.42	-3.50	-2.00	✓	✓
MR95/140	32.28	4.37	-4.22	-3.63	-2.13	✓	✓	**PSM/96/107**	32.69	3.79	-3.81	-4.21	-2.71	✓	✓
MW324	31.43	4.54	-5.07	-3.46	-1.96	✓	✓								

Samples identified in bold are those which were also used to make mixtures

Table 5b - Results of a test for consistency with Basmati.

Pure non-Basmati samples

Criterion A: Consistent with Basmati if Q(50)-36.5 ≤ 0 and T-8 ≤ 0

Criterion B: Consistent with Basmati if Q(50)-36.5 ≤ 0 and T-6.5 ≤ 0

Sample	Q(50)	T	Q(50) -36.5	T-8	T-6.5	Consistent with Basmati?		Sample	Q(50)	T	Q(50) -36.5	T-8	T-6.5	Consistent with Basmati?	
						A	B							A	B
MR95/85	51.64	4.36	15.14	-3.64	-2.14	×	×	MR94/1929	37.69	4.86	1.19	-3.14	-1.64	×	×
MR95/142	31.23	4.73	-5.27	-3.27	-1.77	✓	✓	MR94/1935	39.99	4.39	3.49	-3.61	-2.11	×	×
MR95/130	39.92	4.32	3.42	-3.68	-2.18	×	×	MR94/1938	38.97	5.67	2.47	-2.33	-0.83	×	×
MW801	40.71	4.61	4.21	-3.39	-1.89	×	×	MR94/1941	37.59	5.18	1.09	-2.82	-1.32	×	×
MW110	45.57	4.25	9.07	-3.75	-2.25	×	×	MR94/1942	43.17	4.17	6.67	-3.83	-2.33	×	×
MR95/134	39.48	4.90	2.98	-3.10	-1.60	×	×	MR94/1946	38.99	4.47	2.49	-3.53	-2.03	×	×
MW316	41.69	4.10	5.19	-3.90	-2.40	×	×	PSM/96/102	33.36	4.86	-3.14	-3.14	-1.64	✓	✓
MW104	37.18	4.99	0.68	-3.01	-1.51	×	×	PSM/96/103	30.46	4.63	-6.04	-3.37	-1.87	✓	✓
MR95/150	39.72	4.34	3.22	-3.66	-2.16	×	×	PSM/96/93	38.36	5.64	1.86	-2.36	-0.86	×	×
MW302	35.75	4.90	-0.75	-3.10	-1.60	✓	✓	PSM/96/92	36.99	5.91	0.49	-2.09	-0.59	×	×
MW330	41.72	3.15	5.22	-4.85	-3.35	×	×	MW705	41.61	4.85	5.11	-3.15	-1.65	×	×
MR94/1905	39.52	4.64	3.02	-3.36	-1.86	×	×	MR/94/436	42.16	5.02	5.66	-2.98	-1.48	×	×
MR94/1909	36.68	5.12	0.18	-2.88	-1.38	×	×	**MR/94/435**	38.13	4.53	1.63	-3.47	-1.97	×	×
MR94/1913	38.24	5.78	1.74	-2.22	-0.72	×	×	MR/95/141	29.08	6.32	-7.42	-1.68	-0.18	✓	✓
MR94/1914	39.73	5.03	3.23	-2.97	-1.47	×	×	MW350	34.24	4.60	-2.26	-3.40	-1.90	✓	✓
MR94/1917	43.66	4.55	7.16	-3.45	-1.95	×	×	MR/95/128	38.48	5.06	1.98	-2.94	-1.44	×	×
MR94/1919	42.58	4.78	6.08	-3.22	-1.72	×	×	MW342	39.12	4.50	2.62	-3.50	-2.00	×	×
MR94/1923	38.92	4.89	2.42	-3.11	-1.61	×	×	PSM/96/99	34.99	5.48	-1.51	-2.52	-1.02	✓	✓
MR94/1933	40.07	6.09	3.57	-1.91	-0.41	×	×	**PSM/96/100**	42.69	7.16	6.19	-0.84	0.66	×	×
MR94/1924	40.98	4.96	4.48	-3.04	-1.54	×	×								

Samples identified in bold are those which were also used to make mixtures

Table 5c - Results of a pass/fail test for consistency with Basmati.

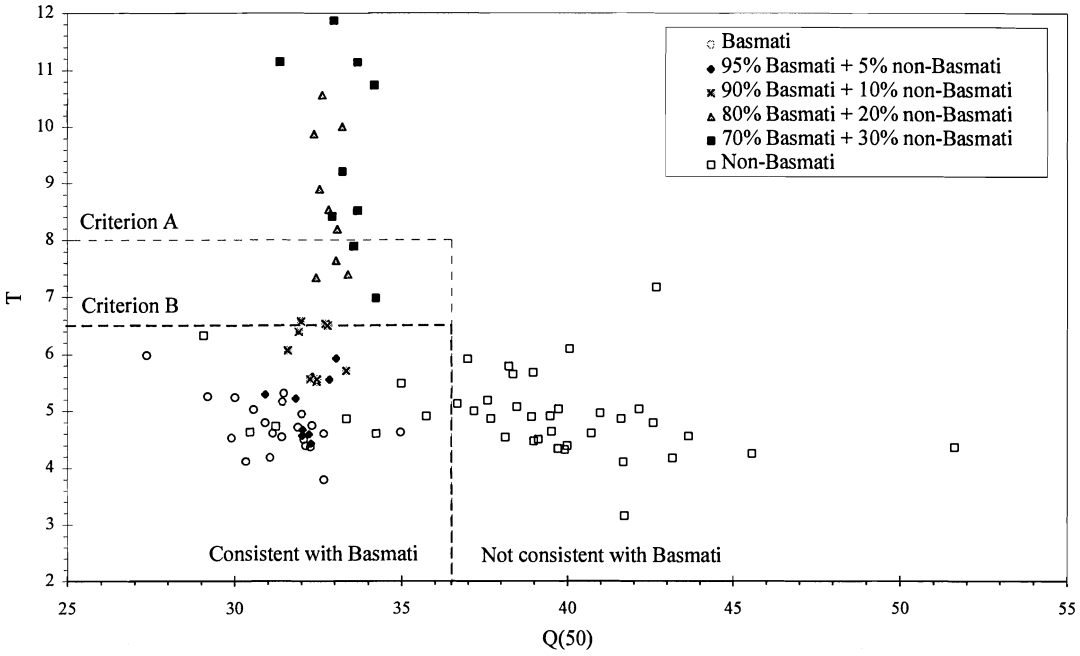
Mixed Samples

Criterion A: Consistent with Basmati if Q(50)-36.5 ≤ 0 and T-8 ≤ 0

Criterion B: Consistent with Basmati if Q(50)-36.5 ≤ 0 and T-6.5 ≤ 0

Adult -erant	% added	Basmati sample																			
		PSM/96/107							PSM/96/97						PSM/96/104						
		Q50	T	Q(50)-36.5	T-8	T-6.5	Basm. character?		Q50	T	Q(50)-36.5	T-8	T-6.5	Basm. character?		Q50	T	Q(50)-36.5	T-8	T-6.5	Basm. character?	
							A	B						A	B						A	B
PSM/96/100	5	32.04	4.67	-4.46	-3.33	-1.83	✓	✓	32.33	5.60	-4.17	-2.40	-0.90	✓	✓	30.92	5.29	-5.58	-2.71	-1.21	✓	✓
	10	31.93	6.39	-4.57	-1.61	-0.11	✓	✓	32.73	6.53	-3.77	-1.47	0.03	✓	×	32.00	6.57	-4.50	-1.43	0.07	✓	×
	20	32.64	10.56	-3.86	2.56	4.06	×	×	33.24	10.00	-3.26	2.00	3.50	×	×	32.39	9.87	-4.11	1.87	3.37	×	×
	30	33.70	11.13	-2.80	3.13	4.63	×	×	34.21	10.73	-2.29	2.73	4.23	×	×	32.99	11.86	-3.51	3.86	5.36	×	×
MW 801	5	32.29	4.43	-4.21	-3.57	-2.07	✓	✓	32.84	5.65	-3.66	-2.35	-0.85	✓	✓	31.84	5.22	-4.66	-2.78	-1.28	✓	✓
	10	32.46	5.51	-4.04	-2.49	-0.99	✓	✓	33.35	5.71	-3.15	-2.29	-0.79	✓	✓	31.61	6.07	-4.89	-1.93	-0.43	✓	✓
	20	33.09	8.19	-3.41	0.19	1.69	×	×	32.83	8.53	-3.67	0.53	2.03	×	×	32.56	8.98	-3.94	0.98	2.48	×	×
	30	31.37	11.14	-5.13	3.14	4.64	×	×	33.70	8.51	-2.80	0.51	2.01	×	×	33.25	9.20	-3.25	1.20	2.70	×	×
MR/94/435	5	32.23	4.59	-4.27	-3.41	-1.91	✓	✓	33.06	5.93	-3.44	-2.07	-0.57	✓	✓	32.03	4.57	-4.47	-3.43	-1.93	✓	✓
	10	32.47	5.56	-4.03	-2.44	-0.94	✓	✓	32.80	6.50	-3.70	-1.50	0.00	✓	✓	32.27	5.56	-4.23	-2.44	-0.94	✓	✓
	20	33.05	7.64	-3.45	-0.36	1.14	✓	×	33.41	7.40	-3.09	-0.60	0.90	✓	×	32.45	7.34	-4.05	-0.66	0.84	✓	×
	30	33.57	7.89	-2.93	-0.11	1.39	✓	×	34.24	6.97	-2.26	-1.03	0.47	✓	×	32.94	8.41	-3.56	0.41	1.91	×	×

Figure 4 - Discrimination of Basmati samples from other types and mixtures using a simple pass / fail statistical test



It can be seen that all the pure Basmati samples and 32 of the 39 pure non-Basmati samples tested were correctly classified. However, the test failed to identify 7 non-Basmati samples correctly. All but one of these (sample MW302, from Australia) were of Indian origin.

Of the mixed samples, addition of the non-Basmati samples PSM/96/100 (Indian white rice) and MW801 (Spanish Thaibonnet) was correctly identified at the 20% and 30% levels in all cases, but could not be detected at the 5% and 10% levels. This is consistent with the level of performance previously established by Whitworth *et al.*, 1996 and provides a useful validation of this work. The sample MR/94/435 (Indian, Sonasali Kh-93) has a lower median RP3 value than the other adulterants (38.13 as compared with 42.69 and 40.71 for the other adulterants) and was less readily distinguishable from Basmati. As such, its addition was only detected in one case of 30% addition in the mixtures tested. It is however, apparent that although for these samples, T only once exceeded the threshold of 8 units, it did increase steadily with the level of adulteration.

3.3.2 Revised pass / fail criterion

For the measurements made within this study, the highest value of T measured for any of the pure Basmati samples was 5.98 units. On this basis, it is possible to consider reducing the threshold to improve the sensitivity of the test. Caution should be exercised in lowering the threshold too far, since this would increase the likelihood of pure Basmati samples being incorrectly classified as adulterated. However, the values of T for the 21 pure Basmati samples have a distribution that is close to normal (see Figure 5a), with a mean of 4.76 and a standard deviation of 0.488. Since 99.9% of a normal distribution is less than (mean + 3.1 × s.d.), setting a limit at 4.76 + 3.1 × 0.488 = 6.27 would give an action limit with a risk of 1 in 1000 of false positives, assuming that the measured Basmati samples are representative, that the distribution of T is indeed normal and that the measurement procedure is stable. To provide a margin of safety to allow for such possibilities, a revised threshold value of 6.5 is considered appropriate.

Re-evaluating the threshold value for Q(50) in the same manner, this is also approximately normally distributed for the measured Basmati samples (Figure 5b) with a mean of 31.33 and a standard deviation of 1.524. An appropriate threshold value is therefore 31.33 + 3.1 × 1.524 = 36.05. Thus, the previously recommended value of 36.5 remains appropriate, and provides a slight additional margin for error.

Using the more sensitive criterion suggested for inconsistency with Basmati of

$$Q(50) > 36.5 \quad \text{or} \quad T = Q(90)-Q(50) > 6.5,$$

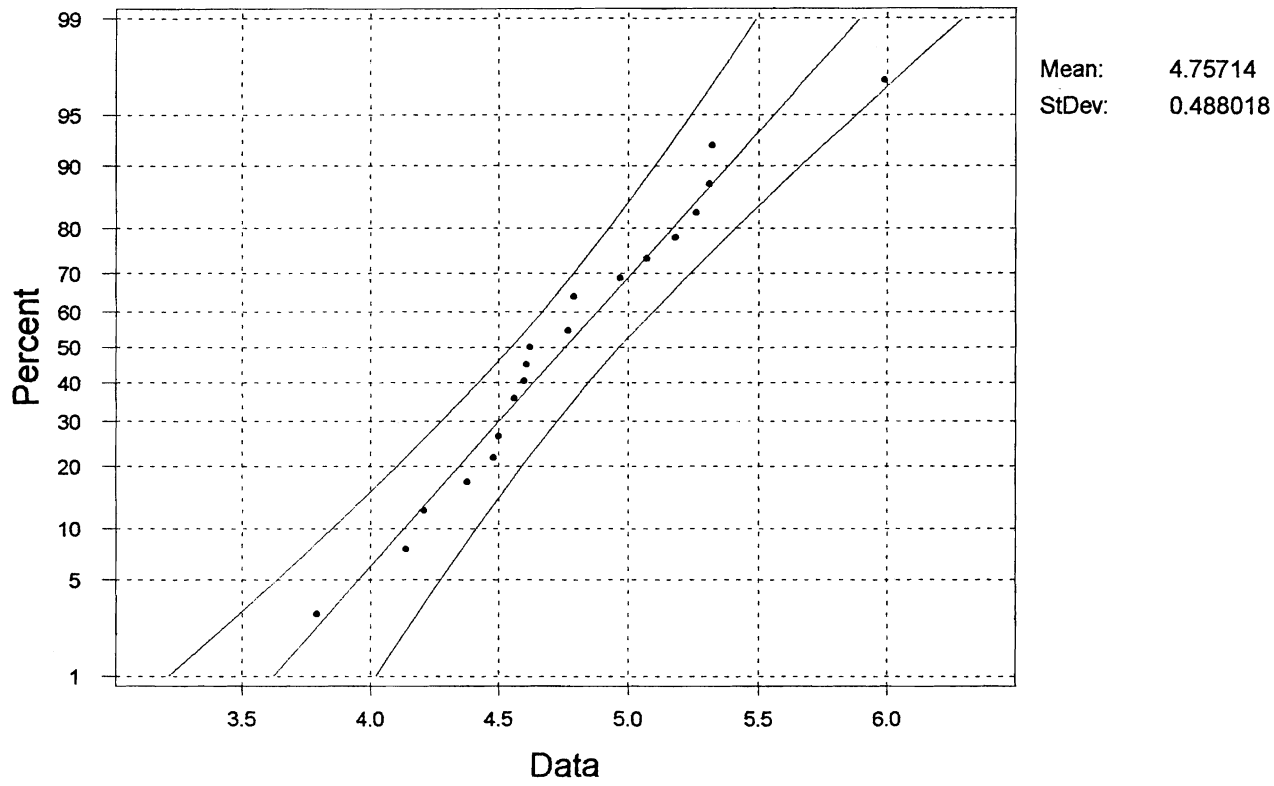
all of the mixtures would be detected at the 20% level, including those containing MR/94/435, and 2 of the 9 mixtures would be detected at the 10% level. The inverse criterion (i.e. for consistency with Basmati) has been denoted criterion B and the results for each sample, tested against this criterion, are given in Table 5.

3.3.3 Quantitative measurement by fitting of mixture distributions

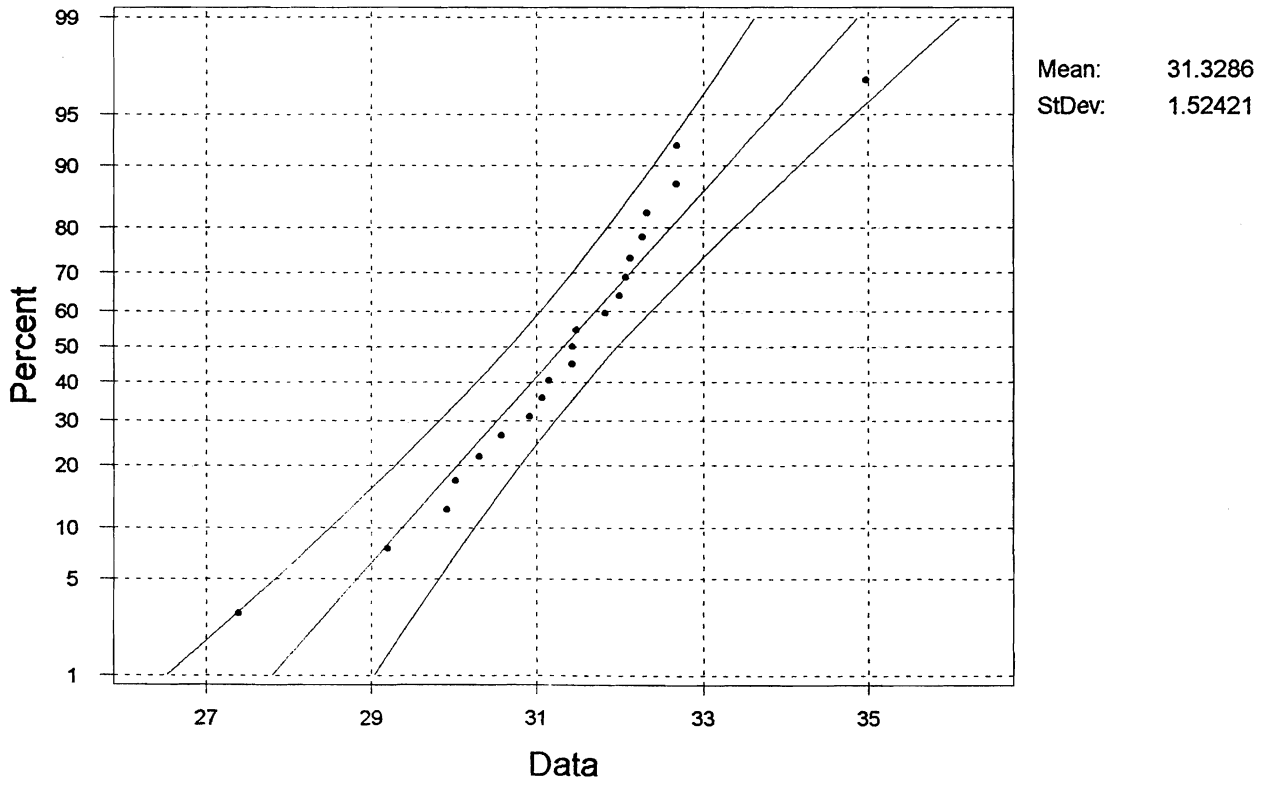
It was identified previously (Whitworth *et al.*, 1996) that a more sophisticated analysis of the RP3 distributions could be made by approximating the distributions of pure

Figure 5 - Normal probability plots of results for reference Basmati samples

(a) Normal probability plot for T



(b) Normal probability plot for Q(50)



Basmati and non-Basmati samples as normal distributions, and by fitting a sum of two such distributions to the measured data, subject to constraints on their means and standard deviations. By calculating the relative proportions of the two distributions providing the best fit to the measured data, a quantitative estimate of the level of adulteration could be made. The characteristics of the pure samples could also be estimated and confidence limits could be defined on each of the estimated quantities. In the previous study, reference data were available for only 46 grains of each reference sample and it was not considered prudent to extrapolate these data for analysis of test samples larger than 50 grains. The method showed promise, but at a sample size of only 50 grains, sampling errors were high.

The data collected in the current study have allowed the statistical fitting technique to be evaluated for samples of 500 grains each, which is considered to represent a suitable compromise between minimising sampling errors and the time required for analysis. As described in section 3.2.1, this has established that the RP3 values for pure samples are not normally distributed, and that it is inappropriate to approximate them in this way for the purposes of fitting to test sample data. Indeed, sufficient variation is seen in the distributions of pure samples, and in particular, for non-Basmati samples, that approximation by an alternative distribution, possibly characterised by more parameters would also be inappropriate for unambiguous estimation of the composition of a mixed sample.

4. Discussion

It has been established previously that measurements of a parameter, RP3, for cooked rice grains produce distributions which have lower mean and median values for Basmati rice samples than for most examples of other rice types tested. Using an earlier version of the measurement method, Whitworth *et al.* (1995) used a criterion of Mean RP3 > 37 to identify pure samples inconsistent with Basmati character. In further measurements using this method (Whitworth *et al.*, 1996), the maximum median RP3 value measured for a Basmati sample was 37.3, suggesting that a slightly greater median cut-off would be appropriate for detection of non-Basmati samples. In the same study, it was established that the revised measurement protocol now in use yielded RP3 values about

1.5 units lower than the old protocol, and a criterion of median RP3 > 36.5 was recommended. More extensive measurements of 500 grains each of 21 Basmati and 39 other samples have been made within the current study. These have confirmed that the criterion of median RP3 > 36.5 is appropriate for detection of a wide range of non-Basmati or heavily adulterated Basmati samples. All Basmati samples and 32 of the 39 non-Basmati samples tested were correctly classified by this criterion. Of the 7 non-Basmati samples not detected, 6 were of Indian origin. This suggests that adulteration of Basmati rice with native varieties at source is likely to be more difficult to detect with this method than adulteration with other rice types at other stages of the supply chain. However, despite the poorer performance for Indian varieties, some such samples were detected, and detection of further examples has also been demonstrated in previous studies.

Because of the above differences in median RP3 values, Basmati samples adulterated by addition of other samples also have RP3 distributions which differ from those of pure Basmati rice, in particular characterised by a broadening of the distribution at the high RP3 end, and in some cases by a bimodal character. In previous work (Whitworth *et al.*, 1996), it was established that this broadening could be measured as the difference, T, between the 50th and 90th percentiles of the distribution, and could be used as a simple indicator of the presence of admixture. Based on measurements of 50 grain reference samples using a previous version of the measurement method, a cut-off value of T>6 was initially identified as a suitable criterion. However, further measurements using the current measurement procedure indicated that it yielded pure sample distributions with a larger standard deviation, and it was determined that an increased cut-off of T>8 was likely to be more appropriate for data measured by this method. Tests of mixed samples demonstrated that this criterion was typically capable of detecting 20% admixture. Further tests of mixed samples in the current study have confirmed this performance. Two types of adulterant were detected at a level of 20% using this criterion, but the criterion was insufficient to detect a further adulterant reliably, even at 30% addition. The current study has provided the most complete set of reference data yet measured using the current measurement method, including data for about 500 grains of each of 21 Basmati and 39 other rice samples. Based on these measurements, it has been

possible to re-evaluate the T criterion more directly. It is apparent that the cut-off of T>8 was unnecessarily conservative, and that a more sensitive cut-off of T>6.5 is more appropriate. For the mixed samples measured in the previous study, this would improve the level of detection to 10% in many cases. For the mixed samples measured in the current study, all mixtures tested would be detected at the 20% level and some would be detected at the 10% level. Assuming that the test method is stable and that the Basmati samples measured are representative, the revised criterion would be expected to provide false identification of authentic Basmati samples as adulterated on only about 0.1% of occasions, although these assumptions should be noted carefully when interpreting results for test samples.

Although the work discussed confirms the performance of the Q(50) and T criteria previously devised and has shown that it is appropriate to modify these criteria to improve the sensitivity of the test, a further aim of this work was to establish whether RP3 distributions could also be analysed to provide a quantitative measurement of the level of adulteration of Basmati samples. In a previous study (Whitworth *et al.*, 1996), it was suggested that it might be possible to achieve this by fitting measured data with a weighted sum of two distributions typical of the populations of reference Basmati and non-Basmati samples. This was demonstrated by approximating the reference sample distributions by normal distributions. However, data were only available for 46 grains per reference sample, and it was unclear whether the assumption that their RP3 values were normally distributed, was valid.

In this study, 500 grains each have been measured for reference samples. It has become clear from these data that the RP3 measurements are not normally distributed for each sample, and that the distributions are in many cases skewed, and in some cases, bimodal. In particular, although the main component of the RP3 distribution of most non-Basmati samples is distinct from that of Basmati samples, many of them have a lower tail which significantly overlaps the RP3 distributions typical of Basmati. The shape and prominence of this lower tail varies considerably between different potential adulterants. Therefore, unless the characteristics of a suspected adulterant were known

in advance, it would not be possible to determine its concentration in a mixed sample with any useful degree of accuracy.

5. Conclusions

In previous studies, a test for Basmati rice authenticity was developed based on comparison of test results with those for reference samples. A significant difficulty in developing and evaluating such tests has been the limited availability of reference samples of known authenticity. Within this project, 30 new rice samples have been collected of which 13 are of Basmati type and many of the others are also of Indian origin. These represent a useful addition to the samples previously collected.

In this study, the amount of reference data available for rice samples measured with the current image analysis procedure has been increased by measurement of 500 grains each of 21 Basmati samples and 39 other samples. 36 mixed samples were also analysed and used to evaluate the performance of the test.

The new data revealed that the RP3 measurements for pure reference samples were not normally distributed, and that a previously suggested method for quantitative estimation of adulteration based on fitting such distributions, could not be applied. The distributions were skewed and bimodal to varying degrees, suggesting that an accurate quantitative analysis based on RP3 measurements is unlikely to be achievable.

Although quantitative estimation of adulteration was not possible, identification of the presence or absence of adulteration was more successful. The performance of a previously established analysis criterion was verified, and a more sensitive criterion was identified. Using the criterion, $Q(50)>36.5$ or $T>6.5$, all pure Basmati samples and most pure non-Basmati samples were correctly classified, although a significant number of Indian non-Basmati samples could not be discriminated from Basmati. For mixtures, the new criterion detected adulteration at levels of 10 to 20%.

5.1 Recommendations for use of the test

It is considered that the test described is suitable for use by rice traders as a screening tool to identify Basmati rice consignments which may have been adulterated with other rice types, and which would be worthy of further investigation. The test is typically sensitive to levels of adulteration of 10 to 20% and above. However, the results given in this report demonstrate that the test is not capable of detecting all types of adulteration, and it is therefore recommended that it should not be relied on as a sole method for ensuring authenticity, but should be used in combination with other methods such as audit trails. Further, although none of the authentic Basmati samples used in this study were misclassified by the test, it cannot be guaranteed that authentic samples, or samples containing small, acceptable levels of other rice types may not occasionally be misidentified as adulterated.

Because of the possibility of misclassifying authentic samples as described above, the test would be unsuitable as the basis for prosecutions. However, it would be suitable for survey purposes to provide an indication of the incidence of other rice types in Basmati samples, although the incidence of certain types of rice to which the test is insensitive, may be underestimated. Although such surveys would provide an indication of the presence of adulteration in samples, the work conducted in this project suggests that it would not be possible to determine the level of such adulteration quantitatively. In particular, it would not be possible to distinguish reliably between small, acceptable levels of impurities and levels sufficient to constitute adulteration, although discriminant parameters exceeding threshold values by a large margin are more likely to signify high levels of adulteration.

6. Acknowledgements

The authors would like to thank the following:

- Mr. David Marks and Miss Emma Gatherer for taking the measurements of the rice samples.
- Mrs. Barbara Smith for assistance with preparing mixed samples.
- Mrs. Carolyn Middleton, Mrs. Susan Slatter and Miss Sarah Hoare for assistance with cataloguing of the samples collected.
- The Natural Resources Institute and Mr. Bob Holben of Vogan & Co Ltd, Fullbourn, Cambridge for milling paddy rice samples.
- Mr. John Humphries of Whitworth Bros. Ltd, the Grain and Feed Trade Association (Gafta), IRRI, Phillipines, the Directorate of Rice Research, India, and Ricetec Inc. for supplying samples of rice.

7. Publications arising from this work

Whitworth, M. and Millar, S. "Further Evaluation of Image Analysis Method for Basmati Rice", CCFRA Research Summary Sheet No. 1997/22.

8. References

Osborne, B. G., Whitworth, M. B., Greenwell, P., Mertens, B., Thompson, M., Fearn, T. "The Authentication of Basmati Rice Using Near Infrared Transmittance Spectroscopy" Final Report on a project for MAFF, March 1993.

Scanlon , M. G., Pritchard, P. E., Osborne, B. G. "Determination of the Authenticity of Rice Varieties" Report on Project N1662 for MAFF, 1991.

Whitworth, M. B., Greenwell, P., Fearn, T., Osborne, B. G. "Physical Techniques for Establishing the Authenticity of Rice" Final Report on MAFF project ANO 634, June 1995. (Reproduced as CCFRA R&D Report No. 16)

Whitworth, M. B., Hall, M. N., Fearn, T., Seitz, R. R. L., Franciosi, N. "Authenticity Testing of Rice" Final Report on MAFF project No. 2A073 (AN1123), April 1996. (Reproduced as CCFRA R&D Report No. 27)