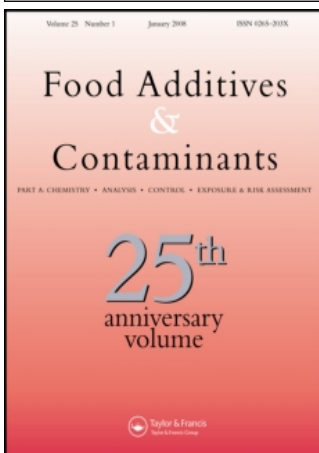


This article was downloaded by:[Central Science Laboratory]
On: 16 July 2008
Access Details: [subscription number 791350051]
Publisher: Taylor & Francis
Informa Ltd Registered in England and Wales Registered Number: 1072954
Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Food Additives & Contaminants Part A - Chemistry, Analysis, Control, Exposure & Risk Assessment

Publication details, including instructions for authors and subscription information:
<http://www.informaworld.com/smpp/title~content=t713599661>

Comparison of calibration methods for the quantification of Basmati and non-Basmati rice using microsatellite analysis

A. Colyer^a; R. Macarthur^a; J. Lloyd^a; H. Hird^a
^a Central Science Laboratory, York, UK

First Published on: 13 June 2008

To cite this Article: Colyer, A., Macarthur, R., Lloyd, J. and Hird, H. (2008)
'Comparison of calibration methods for the quantification of Basmati and
non-Basmati rice using microsatellite analysis', Food Additives & Contaminants,

To link to this article: DOI: 10.1080/02652030802040141
URL: <http://dx.doi.org/10.1080/02652030802040141>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article maybe used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

Comparison of calibration methods for the quantification of Basmati and non-Basmati rice using microsatellite analysis

A. Colyer*, R. Macarthur, J. Lloyd and H. Hird

Central Science Laboratory, Sand Hutton, York, UK

(Received 9 August 2007; final version received 1 March 2008)

The accurate quantification of non-Basmati rice in Basmati rice is central to the successful prosecution of adulteration, where non-Basmati rice has been substituted for Basmati rice. The current method and three alternatives of constructing calibration curves for the measurement of non-Basmati rice in Basmati rice using microsatellite analysis were investigated. The methods compared involved power regression, linear regression (with and without \log_{10} transformation) and hyperbolic regression of the ratio of Basmati to non-Basmati peak areas. Assessments were made using error uncertainty, standard error at the agreed limit of adulteration, and 95% confidence intervals for five example data sets. The linear regression of the ratio of peak areas to the ratio of content proportions was found to give the most precise calibration and thus enhanced quantification of the level of adulteration of Basmati rice with non-Basmati rice.

Keywords: microsatellite; Basmati; DNA; calibration

Introduction

The consumption of Basmati rice has gained in popularity in recent years, particularly in the UK, which consumes 70% of the Basmati rice imported into the European Union (approximately 155,000 tonnes/year). Its popularity stems from its unique characteristics, which include a popcorn-like fragrance and light, fluffy grains on cooking. All rice varieties are closely related and only skilled technicians are able to use a variety of objective measures to distinguish varieties apart. However, varietal identification using DNA-based analysis has been reported in the literature. The range of DNA-based techniques includes random amplified polymorphic DNA (Ko et al. 1994; Ohtsubo et al. 1997) amplified fragment length polymorphisms (Mackill et al. 1996) and microsatellite analysis (Bligh et al. 1999; Archak et al. 2007; Vemireddy et al. 2007). Of these techniques, microsatellite analysis appears to be the most successful and was recently used by the UK Food Standards Agency (FSA) to survey the authenticity of Basmati rice on sale in the country. The survey found that 17% of retail Basmati rice samples contained over 20% non-Basmati rice (FSA 2004). The adventitious mixing of Basmati with non-Basmati rice can occur during the production of Basmati rice and hence up to 7% non-Basmati rice may be present legally in 'Basmati rice' (agreed level of admixture in the UK Code of Practice). However, the presence of more than 20% non-Basmati in some

Basmati rice samples indicates wilful substitution, which if not declared, would be illegal under food labelling rules. The accurate quantification of the level of non-Basmati rice is therefore of paramount importance for successful prosecution. However, microsatellite analysis is a multi-step process, the accuracy of which is affected by, and not limited to, the selection of the rice varieties as appropriate calibrants, the microsatellites used for analysis, the methodology used for DNA extraction, polymerase chain reaction (PCR) and capillary electrophoresis, the efficiency of the software used for data collection and the methodology used for subsequent data analysis. We have focused on improving the methodology used to calculate the level of adulteration with non-Basmati rice. We report a comparison of the current method (construction of a standard curve with a best-fit parabolic line) and possible alternative methods of constructing calibration curves for the measurement of non-Basmati rice in Basmati rice using microsatellite analysis, to enable a more accurate measurement. Calibration methods were compared using a range of measures of uncertainty. The calibration method that was found to give the most precise calibration was the linear regression of the ratio of Basmati peak area to non-Basmati peak area against the ratio of the proportion of Basmati rice to the proportion of non-Basmati rice, which when used together with the measurement of uncertainty, should provide enhanced quantification of

*Corresponding author. Email: a.colyer@cs.l.gov.uk

the level of adulteration of Basmati rice with non-Basmati rice.

Materials and methods

Sample DNA extraction and purification

Calibrants were prepared by gravimetrically mixing milled samples of Taraori (Basmati rice type) and Sherbati (long-grain rice type). These samples were further ground to a fine powder, in a food processor, before DNA extraction using the Nucleon® Phytospure kit (Amersham Pharmacia Biotech Little Chalfont, UK), with the following minor modifications: the cooled cell lysate was centrifuged at 4500g for 15 min before transferring 0.5 ml supernatant to a fresh microcentrifuge tube and at the end of the procedure the DNA pellet was resuspended in TE buffer which had been previously heated to 85°C.

PCR assay conditions

Each PCR reaction (25 µl), contained 1× AmpliTaq Gold® reaction buffer, 0.625 Units AmpliTaq Gold® DNA Polymerase (Applied Biosystems, Warrington, UK) and 0.5 µM of each deoxynucleotide triphosphate (Sigma, Poole, UK) and 5 pmols of sense (CTTAAATGGGCCACATGCG, 5' labelled with hexachlorofluorescein) and 5 pmols of anti-sense (TGCAAGAATCTGACCCGG) oligonucleotide

primers (Sigma) for rice microsatellite marker RM222. DNA template was diluted 1:4 DNA:water and 5 µl added per reaction. Reactions were assembled in Axygen thin-walled PCR tubes and run on ABI 7900 PCR machine with the following thermal cycling protocol: 94°C for 10 min followed by 45 cycles of 94°C for 1 min, 55°C for 1 min and 72°C for 1 min with a hold for 60 min at 60°C.

Fluorescence gel electrophoresis

Amplicons were electrophoresed under denaturing conditions on an ABI377 (Applied Biosystems). Each reaction was run with a set of *N,N,N',N'*-tetramethyl-6-carboxyrhodamine (TAMRA) labelled size standards (Applied Biosystems). The results were analysed using GeneScan analysis software (Applied Biosystems) and the total area, under each peak, for both alleles, at each calibration point, was calculated (Table 1).

Calibration methods

Four calibration methods were compared for their ability to estimate accurate measurements of non-Basmati in Basmati rice. The error uncertainty, standard error at the agreed limit of adulteration and 95% confidence intervals were determined. The methods of calibration investigated were as follows.

Table 1. Total peak area under Basmati and non-Basmati peaks for rice microsatellite marker RM222, from calibration data 1.

| Percentage of non-Basmati | Ratio of proportions ^a | Basmati peak area | Non-Basmati peak area | Ratio of peak areas ^b |
|---------------------------|-----------------------------------|-------------------|-----------------------|----------------------------------|
| 90 | 0.11 | 8573 | 107 530 | 0.080 |
| 90 | 0.11 | 5338 | 77 420 | 0.069 |
| 80 | 0.25 | 25 523 | 117 400 | 0.217 |
| 80 | 0.25 | 16 418 | 76 782 | 0.214 |
| 80 | 0.25 | 10 755 | 49 969 | 0.215 |
| 60 | 0.67 | 45 955 | 62 983 | 0.730 |
| 60 | 0.67 | 35 775 | 48 708 | 0.734 |
| 60 | 0.67 | 36 784 | 49 601 | 0.742 |
| 40 | 1.50 | 108 668 | 71 095 | 1.528 |
| 40 | 1.50 | 44 688 | 26 491 | 1.687 |
| 40 | 1.50 | 31 717 | 18 572 | 1.708 |
| 30 | 2.33 | 45 393 | 17 985 | 2.524 |
| 30 | 2.33 | 30 896 | 11 838 | 2.610 |
| 30 | 2.33 | 79 124 | 31 434 | 2.517 |
| 20 | 4.00 | 40 269 | 5137 | 7.839 |
| 20 | 4.00 | 35 717 | 4785 | 7.464 |
| 10 | 9.00 | 58 007 | 5382 | 10.778 |
| 10 | 9.00 | 37 721 | 3891 | 9.694 |
| 10 | 9.00 | 32 834 | 3865 | 8.495 |
| 7 | 13.29 | 48 012 | 3125 | 15.364 |
| 7 | 13.29 | 43 202 | 2590 | 16.680 |
| 7 | 13.29 | 43 736 | 3061 | 14.288 |

Notes: ^aRatio of proportions = proportion of Basmati rice in sample/proportion of non-Basmati rice in sample, e.g. 90% non-Basmati implies the ratio of proportions is 0.1/0.9 = 0.11.

^bRatio of peak areas = Basmati peak area/non-Basmati peak area.

Power regression curve

A frequently used method of constructing calibration curves, where a power regression curve is fitted to the ratio of Basmati peak area to non-Basmati peak area against the proportion of non-Basmati rice, is shown in Equation (1):

$$r = ap^b \quad (1)$$

where r is the ratio of Basmati peak area to non-Basmati peak area; p is the proportion of non-Basmati rice; and a and b are constants estimated during model fitting.

Linear regression of the ratio of peak areas onto the ratio of proportions

An alternative option is to fit a linear regression to the ratio of Basmati peak area to non-Basmati peak area onto the ratio of the proportion of Basmati to the proportion of non-Basmati, as in Equation (2):

$$r = cm + d \quad (2)$$

where m is the ratio of the proportion of Basmati rice to the proportion of non-Basmati rice, i.e. $m = (1 - p)/p$; and c and d are constants estimated during model fitting.

Linear regression of the log of the ratio of peak areas onto the log of the ratio of proportions

This method fits a linear regression to the log (base10) of the ratio of Basmati peak area to non-Basmati peak area against the log of the ratio of the proportion of Basmati to the proportion of non-Basmati, as in Equation (3):

$$\log_{10}(r) = e \log_{10}(m) + f \quad (3)$$

where e and f are constants estimated during model fitting.

Hyperbolic regression curve

This alternative regresses the ratio of Basmati peak area to the non-Basmati peak area against the proportion of non-Basmati in the form of a hyperbolic line, as in Equation (4). This is a generalized form of Equation (2) that allows for the presence of an additional constant (Bligh 2000):

$$r = g + \frac{h}{1 + kp} \quad (4)$$

where g , h and k are constants estimated during model fitting.

Statistical methods

Statistical analyses of five example calibration curves were performed by applying aforementioned regression

techniques to investigate which calibration method would provide the most accurate results. The uncertainty associated with each calibration method was estimated by performing residual analyses (the residual is the difference between the observed response and its fitted model value) and by calculation of standard errors at the agreed limit of 7% non-Basmati rice in Basmati rice, by a method of inverse calibration (Brown 1993, p. 22). The residual analyses consisted of comparing the residual standard deviation and a graphical inspection to test modelling assumptions. In addition, 95% confidence intervals were calculated and the adjusted R^2 statistic (the proportion of total variability in the response that is accounted for by the model, adjusted for degrees of freedom) was recorded. Appropriate transformations were performed for Equation (3) to enable an appropriate comparison of errors and confidence widths to the other calibration methods. Without the transformation, measures may falsely appear smaller, as they are on the log(ratio of peak areas) scale rather than the ratio of peak areas scale. All analyses were performed using GenStat® 8.1 statistical software (VSN International, Hemel Hempstead, UK; <http://www.vsn.co.uk/products/genstat/>).

Results and discussion

Microsatellite analysis has now become the 'gold standard' method for the identification and quantification of long grain adulterant rice in Basmati rice following the survey conducted by the UK FSA in 2004 and the uptake of the methodology by analytical laboratories worldwide (Archak et al. 2007; Vemireddy et al. 2007). Rice varieties are closely related and no one marker can definitively identify a variety of rice, therefore, a panel of microsatellite markers are used to build a profile of the rice varieties present. Analysis of samples is usually a two-step process, where the rice varieties of rice present is determined using up to twelve microsatellite markers, followed by quantitative analysis, using either a single or occasionally, two microsatellite markers, appropriate for the constituents of the sample. The choice of microsatellite markers for both the identification and quantification stages are a matter of choice, although some, for example RM1, RM55 and RM44, are used by many laboratories (FSA 2004; Archak et al. 2007). The work presented in this study uses the microsatellite marker RM222 for illustrative purposes, since this marker easily distinguishes between the two rice varieties used to construct the calibration curve: Taraori Basmati and Sherbati long grain. However, the approach outlined in this paper has also been shown to be applicable for quantitation using the microsatellite markers M16, RM44, RM171 and RM201 and should in fact be

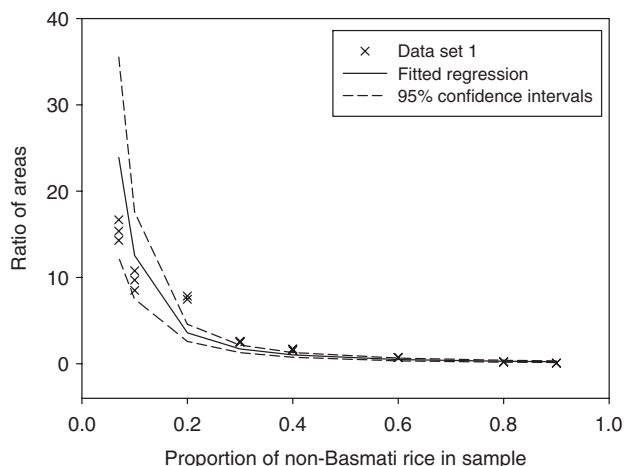


Figure 1. Calibration of data set 1 using Equation (1) – fitted power regression of the ratio of peak areas against the proportion of non-Basmati rice in the sample with 95% confidence intervals (where the ratio of peak areas = Basmati peak area/non-Basmati peak area).

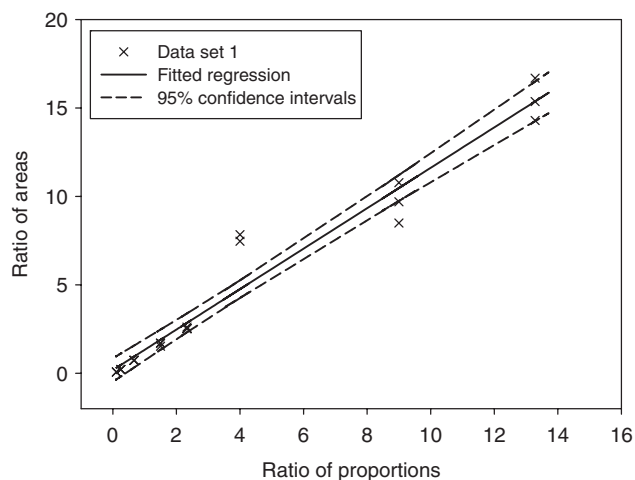


Figure 2. Calibration of data set 1 using Equation (2) – fitted linear regression of the ratio of peak areas against the ratio of proportions 95% confidence intervals (where the ratio of peak areas = Basmati peak area/non-Basmati peak area; and the ratio of proportions = proportion of Basmati rice/proportion of non-Basmati rice in sample).

appropriate for any microsatellite marker optimized for quantification.

The data for the current study was generated by the analysis of multiple identical calibration curves, constructed from samples taken from single batches of Taraori Basmati and Sherbati long-grain rice. In all, five calibration curves were prepared and taken through the complete analysis, under repeatability conditions to provide independent, but linked data sets, enabling statistical analysis.

For illustration, one of the five calibration data sets used to compare the calibration methods is shown

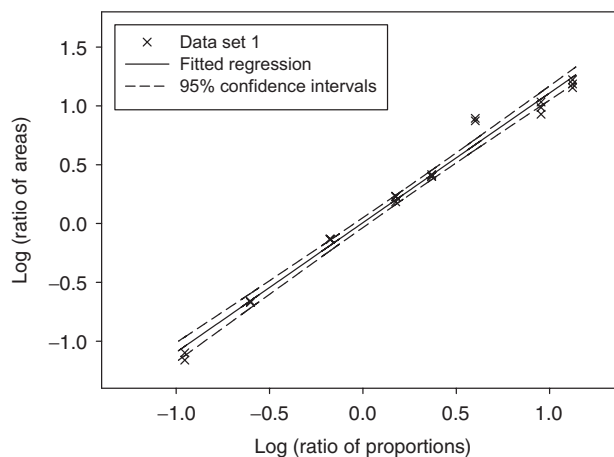


Figure 3. Calibration of data set 1 using Equation (3) – fitted linear regression of the log (ratio of peak areas) against the log (ratio of proportions) 95% confidence intervals (where the ratio of peak areas = Basmati peak area/non-Basmati peak area; and the ratio of proportions = proportion of Basmati rice/proportion of non-Basmati rice in sample).

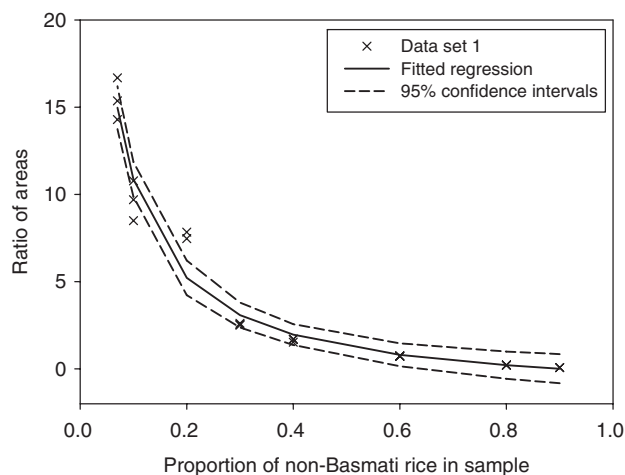


Figure 4. Calibration of data set 1 using Equation (4) – fitted hyperbolic model of the ratio of peak areas against the proportion of non-Basmati rice in the sample (solid) 95% confidence intervals (dashed) (where the ratio of peak areas = Basmati peak area/non-Basmati peak area).

in Table 1. From this, the calculation of the ratio of peak areas and the ratio of proportions was determined.

Calibration models, in Equations (1) to (4), applied to this data set are illustrated in Figures 1–4. The changes in uncertainty between models are demonstrated by 95% confidence intervals for mean values. The linear regression of the ratio of peak areas onto the ratio of proportions (Figure 2 and Equation 2) has the tightest intervals on the ratio of peak areas scale.

A visual check may lead the reader to believe that Equation (3) (Figure 3) has a lesser scatter of points or tighter confidence intervals than the other equations.

Table 2. Maximum width of confidence intervals on the ratio of peak areas scale (after the transformation for Equation 3 to enable comparability).

| Calibration | Model | | | |
|-------------|--------------|--------------|--------------|--------------|
| | Equation (1) | Equation (2) | Equation (3) | Equation (4) |
| 1 | 23.25 | 2.24 | 5.33 | 2.46 |
| 2 | 37.28 | 4.51 | 7.20 | 4.72 |
| 3 | 29.64 | 1.86 | 6.48 | 1.49 |
| 4 | 42.99 | 3.76 | 11.26 | 4.48 |
| 5 | 10.89 | 2.32 | 3.34 | 1.31 |

Table 3. Adjusted R^2 values for calibration models.

| Calibration | Model | | | |
|-------------|--------------|--------------|--------------|--------------|
| | Equation (1) | Equation (2) | Equation (3) | Equation (4) |
| 1 | 89.1 | 95.8 | 98.6 | 95.9 |
| 2 | 89.7 | 87.2 | 98.5 | 89.4 |
| 3 | 86.6 | 96.7 | 98.3 | 98.3 |
| 4 | 89.2 | 93.6 | 97.6 | 93.5 |
| 5 | 91.0 | 91.5 | 97.7 | 98.2 |

Table 4. Residual standard deviation for calibration models (after the transformation for Equation 3 to enable comparability).

| Calibration | Model | | | |
|-------------|--------------|--------------|--------------|--------------|
| | Equation (1) | Equation (2) | Equation (3) | Equation (4) |
| 1 | 3.529 | 1.018 | 1.468 | 1.052 |
| 2 | 6.033 | 1.080 | 2.457 | 1.068 |
| 3 | 4.629 | 1.048 | 1.839 | 1.029 |
| 4 | 6.106 | 1.022 | 2.297 | 1.048 |
| 5 | 1.113 | 1.054 | 1.291 | 1.043 |

However, the plots are on different scales (log(ratio of peak areas) for Figure 3 compared with ratio of peak areas for Figures 1, 2 and 4) and so the spread can not be crudely compared in this way. An investigation of the maximum confidence interval widths for each calibration model (Table 2), after appropriate transformations for Equation (3), show Equation (1) can have over 15 times the confidence width of Equations (2) and (4). Equation (3) can have up to four times that of Equations (2) and (4). Therefore, Equations (2) and (4) give more accurate estimation of the ratio of peak areas.

Graphical residual checks highlighted a non-random structure for Equation (1) due to the poor fit of the model. Alternative model residuals were much improved with a random scatter of points around zero. This may be expected as the relationship between the ratio of peak areas against the ratio of proportions (Equation 2) should be linear if the relative peak area for each type of rice is constant for each rice across the

calibration range, or if relative peak areas are equal for each type of rice. In addition, the log of ratio of the peak areas against the log of ratio of proportions (Equation 3) may be desirable as measurements based on PCR are cyclic in nature and therefore follow a log-normal distribution (assuming all other sources of error are minimal). Also, due to the reference concentration values often chosen, the log transformation can lead to a more evenly distributed set of calibration points.

Adjusted R^2 values for each calibration data set for Equations (2), (3) and (4) are mostly greater than 90%, where as the majority of adjusted R^2 values of Equation (1) are under 90% (Table 3). Hence, Equation (1) explains less of the variation in the data sets than the alternatives and so has a poorer fit.

Residual standard deviation values for each calibration data set for Equations (2) and (4) are all less than 1.1 (Table 4). Equations (1) and (3) can have over five and two times the residual standard deviation of

Table 5. Standard error of the calibration models at the agreed limit of 7% non-Basmati rice in Basmati rice (after the transformation of Equation 3 to enable comparability).

| Calibration | Model | | | |
|-------------|--------------|--------------|--------------|--------------|
| | Equation (1) | Equation (2) | Equation (3) | Equation (4) |
| 1 | 0.0090 | 0.0023 | 0.0043 | 0.0032 |
| 2 | 0.0088 | 0.0043 | 0.0046 | 0.0077 |
| 3 | 0.0098 | 0.0020 | 0.0047 | 0.0025 |
| 4 | 0.0095 | 0.0030 | 0.0064 | 0.0046 |
| 5 | 0.0078 | 0.0032 | 0.0055 | 0.0006 |
| Maximum | 0.0098 | 0.0043 | 0.0064 | 0.0077 |

Equations (2) and (4), respectively. Thus, Equations (2) and (4) provide better fits to the calibration data.

Standard error values at the agreed limit of 7% non-Basmati rice in Basmati rice show Equation (1) to be consistently larger than the alternatives (Table 5). Equation (2) has the smallest standard error for the first four calibrations and has the smallest maximum standard error overall. Hence, Equation (2) performs more accurately at the limit of 7% non-Basmati rice.

An observation concerning the available data is the modest amount of information for the very low or very high percentage non-Basmati rice, where most assessments and predictions will be made. Therefore, it is advised that more reference materials at very high and very low concentrations are used for calibration.

In addition, when constructing calibration curves it is advised that the modelling checks that have been carried out during this investigation, such as graphical exploration of residuals, are continually performed. This enables monitoring of any underlying structure in the data which is not at first evident but which could lead to a poor fitting and so unreliable calibration curve being used for prediction of non-Basmati rice content.

Conclusion

In conclusion, after investigation into current methods for constructing calibration curves for quantitative microsatellite analysis, it is recommended that a linear regression of the ratio of Basmati peak area to non-Basmati peak area onto the ratio of the proportion of Basmati to the proportion of non-Basmati (Equation 2) is used in future calibrations.

It was found that Equations (2) and (4) gave more accurate assessments when comparing error uncertainty and structure, standard errors at the agreed limit of 7% non-Basmati rice and 95% confidence intervals, when compared with Equations (1) and (3). Overall, for these data sets, Equation (2) was found to give the most precise calibration, as it reported the lowest standard error at the agreed limit of 7% non-Basmati on average. For future data sets of this kind, Equations

(2) and (4) may give very similar performances. However, due to ease of implementation in a non-statistical package, Equation (2) is favoured.

Using this calibration method, and the measurement uncertainty derived from it, will provide improved quantification of the level of adulteration of Basmati rice with non-Basmati rice. Furthermore, this method can be utilized for other measurements taken from an electropherogram where the aim is to compare peak areas.

References

- Archak S, Lakshminarayananreddy V, Nagaraju J. 2007. High-throughput multiplex microsatellite marker assay for detection and quantitation of adulteration in Basmati rice (*Oryza sativa*). *Electrophoresis*. 28:2396–2405.
- Bligh HFJ. 2000. Detection of adulteration of Basmati rice with non-premium long-grain rice. *Int J Food Sci Tech*. 35:257–265.
- Bligh HFJ, Blackhall NW, Edwards KJ, McClung AM. 1999. Cultivar identification of brown and white milled rice using AFLPs and SSLPs. *Crop Sci*. 39:1715–1721.
- Brown PJ. 1993. Measurement, regression and calibration. Oxford: Oxford University Press.
- Food Standards Agency (FSA). 2004. Basmati rice survey. March. Food Surveillance Information Sheet No. 47/04. London: FSA.
- Ko HL, Cowan DC, Henry RJ, Graham GC, Blakeney AB, Lewin LG. 1994. Random amplified polymorphic DNA analysis of Australian rice (*Oryza sativa* L.) varieties. *Euphytica*. 80:179–189.
- Mackill DJ, Zhang Z, Redona ED, Colowit PM. 1996. Level of polymorphism and genetic mapping of AFLP markers in rice. *Genome*. 39:969–977.
- Ohtsubo K, Fujii T, Hashino Y, Toyoshima H, Okadome H, Nakamura S, Kawasaki S. 1997. Identification of domestic rice cultivars by RAPD method using milled rice grains as samples. *Nippon Shokuhin Kagaku Kogaku Kaishi*. 44:386–390.
- Vemireddy LR, Archak S, Nagaraju J. 2007. Capillary electrophoresis is essential for microsatellite marker based detection and quantification of adulteration of Basmati rice (*Oryza sativa*). *J Agric Food Chem*. 55: 8112–8117.