E

BMT/6/2
ORIGINAL: English
DATE: January 31, 2000

UPOV

INTERNATIONAL UNION FOR THE PROTECTION OF NEW VARIETIES OF PLANTS

GENEVA

WORKING GROUP ON BIOCHEMICAL AND MOLECULAR TECHNIQUES AND DNA-PROFILING IN PARTICULAR

Sixth Session
Angers, France, March 1 to 3, 2000

MOST SIMILAR VARIETY COMPARISONS IN CHRYSANTHEMUMS

prepared by experts from the United Kingdom and the United States of America

MOST SIMILAR VARIETY COMPARISONS IN CHRYSANTHEMUMS.

J. R. Law, J. C. Reeves, J. Jackson, P. Donini, D. Matthews, J. S. C. Smith* and R. J. Cooke.

NIAB, Cambridge CB3 0LE, UK; *Pioneer Hi-Bred International, Johnson Iowa 50131, USA.

1. Introduction

There is much interest in approaches that reduce the workload and associated costs of DUS testing, for instance the elimination of unnecessary comparisons by grouping existing and candidate varieties prior to more formal testing. We have previously reported that in maize, the use of various types of DNA profiling (molecular markers) to group related varieties for this purpose has advantages over the morphological characters normally used in DUS testing (BMT/5/3, 1998, Law *et al.*, 1999). We have now extended this approach to consider varieties of chrysanthemums. Two types of molecular markers are compared:

(i) *5'-Anchored Inter-SSR PCR (ISSR)* (Zietkiewicz *et al*., 1994), a technique which is useful for identifying the presence of targeted repeated elements in DNA and evaluating their distribution. ISSR analysis is based on the fact that if an oligonucleotide comprising a repetitive microsatellite-like region and a redundant 5' anchor is used as the sole primer in a PCR, then it will amplify genomic sequences flanked by closely spaced, inversely orientated microsatellite sequences. ISSR has been utilised for variety identification in various species, including chrysanthemums (Woolf *et al.*, 1995);

(ii) *Inverse Sequence-Tagged Repeat Profiling (ISTR),* a PCR-based profiling method developed by Rohde (1996). ISTR primers are specific for the long terminal repeats of *gypsy-/copia*-like retrotransposons, which are high copy number elements universally distributed among eukaryotes (Wessler *et al*., 1995). The primers face outwards from the elements and amplify regions between randomly inserted and nested elements. Rohde (1996) has demonstrated that polymorphic profiles can be generated by this method from a wide range of eukaryotic species, although its application for the analysis of plant varieties has not been widely reported.

In this paper we use these two methods and morphological descriptors to measure the distances between a set of chrysanthemum varieties and to select the most similar variety. Some of the implications for DUS testing are discussed.

2. Materials and Methods

2.1 Plant material and DNA extraction

A set of 39 varieties of chrysanthemum (*Dendranthema grandiflora* Tzvelev) was selected for analysis from a much larger collection of varieties. These comprised a number of unrelated varieties (designated A, B, C, D, F, G, H, I, J, L, M, N, O, P) along with sports and families of sports of some of these (designated A1, A2, etc. for sports derived from A, and A4a, A4b, etc. for a family of sports derived from variety A4). The varieties were grown under controlled conditions in the greenhouse at NIAB. Leaf material was collected from each of the 39

varieties and DNA extracted using a modified version of the method described by Fulton *et al.* (1995).

2.2 DNA profiling

ISSR analysis was carried out using two different primers - PCT12 and PCT18 - with PCR conditions and semi-automated analysis of the PCR products using a Li-Cor Gene ReadIR 4200 DNA Analyser as described elsewhere (Jackson *et al.*, 1999). ISTR analysis was performed using two pairs of the primers described by Rohde (1996) –ISTR3 + ISTR2 (hereafter referred to as ISTR32) and ISTR3 +ISTR3 (ISTR33). PCR was performed with an initial denaturation step of 3 min at 95°C, followed by 40 cycles of [30 sec at 95°C, 30 sec at 45°C, 2 min at 72°C] and 10 min extension at 72°C. Reactions were carried out in 25ml aliquots comprising 25ng genomic DNA, 200mM each dNTP, 2.5mM MgCl2, 75mM Tris-HCl, pH 8.8, $(NH_4)_2SO_4$, 0.01 % (w/v) Tween-20, 2.5pmoles each of forward and reverse primers and 1 unit of *Taq* polymerase (BioGene). For analysis of PCR products on the Li-Cor, one IR-labelled primer was included in the reaction.

2.3 Statistical Analyses

Morphological data were derived from records held at NIAB. Matrices were compiled for each data set, comprising the absence/presence of bands at specific places on a gel for ISSR and ISTR, or the character state (as defined in the UPOV Guidelines) for the morphological data. Similarity matrices were constructed using the Jaccard method (ISSR and ISTR) or City Block method (morphology), as described in BMT/5/3 and Law *et al.* (1999). Comparisons of the similarity matrices and determination of the most similar variety were carried out as previously (BMT/5/3, Law *et al.*, 1999).

3. Results

3.1 Discrimination between varieties

Both of the molecular methods, as well as the morphological characters, were very efficient in distinguishing between the 39 varieties - Table 1 shows the discrimination rates (pair-wise separation coefficients) achieved at various degrees of stringency (i.e. different numbers of bands or character states needed for distinctness, see Law *et al.*, 1998). For the morphological characters, it was necessary with this set of varieties to exceed a stringency of 12 morphological character state differences before the discrimination rate fell below 100%. It must be remembered however that these characters have largely been chosen especially for their value for DUS testing, which is clearly not the case with the molecular markers. The larger number of ISSR bands in general provided higher varietal discrimination rates (see also Law *et al.*, 1998). However, the rates for the combined ISTR data were comparable. The overall average PIC (polymorphism information content) values for the individual components of the DNA profiles were higher for ISTR than for ISSR ($PIC_{ISTR32} = 0.36$, $PIC_{ISTR33} = 0.46$, $PIC_{PCT12} = 0.19$, $PIC_{PCT18} = 0.18$). Again this may in part be a feature of the larger number of ISSR bands, which increases the potential number of less informative bands.

Comparisons of the entire similarity matrices were made by calculation of pair-wise correlations. The whole similarity matrix correlations for morphology *vs* ISSR and morphology *vs* ISTR were 0.152 and 0.469 respectively, whilst for ISSR *vs* ISTR the correlation was 0.445. These values may reflect the parts and amount of the genome that are sampled by the different techniques.

A summary of the distribution of all pair-wise varietal distance (1 – similarity) estimates for the three approaches is shown in Figure 1. As was the case for maize, the magnitude of the distances as measured by morphology was much less than those as measured using molecular markers.

3.2 Most similar variety comparisons

As has been previously reported for maize (Law et al 1999), in order to evaluate the relative usefulness of morphology and molecular markers for defining associations between varieties, the 'most similar variety' was calculated for each of the varieties using each method in turn (Table 2). For example, for target variety number 1 (A), the most similar variety by morphology was number 30 (H1b, with a distance (1-similarity coefficient) value of 0.151), by ISTR analysis was number 5 (A2a, distance 0.111) and by ISSR was number 17 (A7a, distance 0.535). Cases where the most similar variety falls into the same family group of sports (and hence might be expected to be truly most similar by pedigree) are underlined. Table 2 also shows the range of 'minimum distances' for each method. The range is smallest, and the mean value lowest, for the morphological characters, which is also the case in maize (Law *et al.*, 1999) and confirms the distribution of distances (Figure 1). Agreement between the methods was as follows:

Methods	**No. of cases of agreement in most similar variety**	**Variety numbers**
Morphology & ISTR	6	21, 25, 26, 30, 31, 34
Morphology & ISSR	6	14, 15, 18, 33, 34, 39
ISTR & ISSR	9	2, 5, 7, 8, 19, 23, 27, 28, 34

These data only consider the actual most similar variety and a possible criticism is that 'near misses' in terms of minimum distance are not taken into account. The lack of good quality pedigree data precludes the approach taken previously with maize. However, two additional approaches were undertaken to examine this. Firstly, using the whole data matrix, the percentage correct identification of the most similar variety within the same family group of sports (e.g. variety A4a most similar to another member of the A4 family) was calculated for each method. The percentages were: for morphology - 43.6%, for ISSR – 33.3% and for ISTR 35.9%. The second approach, considering only those varieties for which there were two or more family members, cf. Table 2, was to compare the methods in turn for their ability to identify each of the varieties in a family group as being within the top 10, 20 or 30% of ranked minimum distances. This is based on the assumption that members of a family group will be more closely related to one another than to other varieties. The data (Table 3) showed that if the top 30% were considered, then the performance of ISTR was almost as good as that of morphology, with 48% and 53% inclusions respectively. The rate for ISSR was 32%. To check if the performance of the morphology was being driven by a set of characters that relate to flower colour (which might be of particular importance with sports), nine such characteristics were removed and the analyses re-computed. The within family identification rate fell slightly, to 50% (for the top 30% criterion).

4. Discussion

We have previously shown that in maize, morphology is a relatively poor way of assessing relationships between varieties as determined by pedigree. Molecular methods of various kinds were much more able to identify a minimum set of close varieties likely to contain the truly 'most similar' variety, taken as that defined by pedigree. The range of distances between varieties as assessed by morphological characters was also very much less than the values obtained from DNA profiling (BMT/5/3, Law *et al.*, 1999). In the current work, a similar conclusion can be reached with regard to the values of the distance measurements (Figure 1) and the pair-wise 'minimum distances' (Table 2). However, as pedigree data of comparable quality for are not available the chrysanthemum varieties, direct comparisons with pedigree are impossible, although it is reasonable to assume that families of sports are more closely related to each other than to other varieties. Given that, the morphological characters seem to perform well as predictors of relationship in chrysanthemums (Tables 2 and 3). However, ISTR appears to be almost as efficient and there is a reasonably good correlation (0.469) between the ISTR and morphology similarity matrices. Hence it could be argued that in terms of selecting the most similar varieties for grouping prior to more formal DUS testing, ISTR would be a useful approach. It provides data that correlate well with morphology (and relatedness), but in a shorter time and probably at lower cost (without the need to grow fully mature plants). Whilst ISSR also shares some of these advantages, it suffers somewhat in that the profiles generated are more susceptible to analytical conditions (i.e. less robust), the gels are more difficult to score and the correlations with morphology are worse, both in terms of overall similarity and in recognising family groups.

Comparison of these results with those previously obtained with maize (BMT/5/3, Law *et al.*, 1999) confirms that it would be wise to adopt a crop by crop, and method by method, approach when considering the questions of relatedness and most similar variety comparisons. It is unlikely that it will be possible to propose a single best analytical method, applicable to all species. As well as being of interest to DUS testing *per se*, and for potential grouping purposes, these results are significant for other situations which require knowledge of the associations between varieties, e.g. assessments of minimum distance and establishing criteria for the definition of essential derivation. Hence it is clear that DNA profiling methods have much to offer for determining varietal associations and for variety grouping and related topics.

ACKNOWLEDGEMENTS

This work was largely funded by the Ministry of Agriculture, Fisheries and Food (MAFF), to whom we are grateful.

REFERENCES

Fulton, P.J., Chungwongse, J. and Tanksley, S.S., 1995. Microprep protocol for extraction of DNA from tomato and other herbaceous plants. Plant Mol. Biol. Rep 13: 207-209.

Law, J.R.,. Donini, P, Koebner, R.M.D, Reeves, J.C. and Cooke, R.J., 1998. DNA profiling and plant variety registration. 3: The statistical assessment of distinctness in wheat using amplified fragment length polymorphisms. Euphytica 102: 335-342.

Law, J.R., Cooke, R.J., Reeves, J.C., Donini, P and Smith, J.S.C., 1999. Most similar variety comparisons as a grouping tool. Plant Varieties & Seeds (in press).

Jackson, J., Matthews, D. and Cooper, W., 1999. A modified inter-SSR PCR protocol for use in conjunction with the Li-Cor Gene ReadIR 4200 DNA Analyser. Biotechniques (in press).

Rohde, W. ,1996. Inverse sequence tagged repeat (ISTR) analysis a novel and universal PCR based technique for genome analysis in the plant and animal kingdom. J. Genet. And Breeding 50: 249-251

Wessler, S. R., Bureau, T. E. and White, S. E., 1995. LTR-retrotransposons and mites: important players in the evolution of plant genomes. Curr. Opinion in Genet. Dev. 5: 814-821.

Wolff, K., Zietkiewicz, E. and Hofstra, H., 1995. Identification of chrysanthemum cultivars and stability of DNA fingerprint patterns. Theor. Appl. Genet. 91: 439-447.

Zietkiewicz, E., Rafalski, A. and Labuda, D., 1994. Genome fingerprinting by simple sequence repeat (SSR)-anchored polymerase chain reaction amplification. Genomics 20: 176-183.

~~~~~

Table 1 – Discrimination rates for molecular methods and morphology with various criteria used for distinctness.

Method	No of polymorphic bands/characters	Discrimination rate (%) at different criteria (1, 2 or 3 bands/character states)		
		1	2	3
ISSR – PCT12	88	100	100	100
ISSR – PCT 18	53	100	100	94.9
ISSR – PCT 12 + PCT 18	141	100	100	100
ISTR – 32	16	59.0	10.3	2.6
ISTR – 33	32	94.9	74.4	56.4
ISTR – 32 + 33	48	100	100	76.9
Morphology	58	100	100	100
~~~~~

Table 2 – Most similar variety to target variety (and distance (1-similarity coefficient) values).

Target variety	Morphology			ISTR			ISSR		
1 A	30	0.151	H1b	5	0.111	A2a	17	0.535	A7a
2 A1	3	0.153	A1a	5	0.105	A2a	5	0.543	A2a
3 A1a	2	0.153	A1	1	0.294	A	6	0.595	A2a1
4 A2	11	0.140	A4d	15	0.095	A6a	5	0.500	A2a
5 A2a	15	0.098	A6d	4	0.100	A2	4	0.500	A2
6 A2a1	5	0.148	A2A1	1	0.333	A	3	0.595	A1a
7 A3a1	14	0.106	A6	5	0.150	A2a	5	0.550	A2a
8 A4a	9	0.082	A4b	15	0.143	A6a	15	0.531	A6a
9 A4b	16	0.068	A7	7	0.190	A3a1	14	0.667	A6
10 A4c	16	0.057	A7	8	0.158	A4a	15	0.571	A6a
11 A4d	12	0.049	A4f	2	0.444	A1	13	0.627	A4g
12 A4f	11	0.049	A4d	4	0.136	A2	14	0.595	A6
13 A4g	11	0.099	A4d	12	0.174	A4f	7	0.591	A3a1
14 A6	15	0.051	A6a	4	0.261	A2	15	0.471	A6a
15 A6a	14	0.051	A6	4	0.095	A2	14	0.471	A6
16 A7	10	0.057	A4c	5	0.368	A2a	11	0.689	A4d
17 A7a	12	0.089	A4f	4	0.136	A2	15	0.500	A6a
18 D	19	0.194	D1	2	0.350	A1	19	0.486	D1
19 D1	6	0.167	A2a1	18	0.368	D	18	0.486	D
20 C	37	0.177	N	21	0.150	C1	31	0.607	H1a
21 C1	20	0.257	C	20	0.150	C	22	0.656	B
22 B	26	0.165	B4	24	0.125	B2	23	0.550	B1b
23 B1a	26	0.128	B4	22	0.188	B	22	0.550	B
24 B2	26	0.129	B4	25	0.118	B3	7	0.581	A3A1
25 B3	24	0.225	B2	24	0.118	B2	22	0.564	B
26 B4	23	0.128	B1a	23	0.278	B1a	22	0.619	B
27 F1	12	0.180	A4f	28	0.417	F2	28	0.618	F2
28 F2	12	0.169	A4f	27	0.417	F1	27	0.618	F1
29 G1	31	0.113	H1a	5	0.304	A2a	27	0.625	F1
30 H1b	31	0.074	H1a	31	0.267	H1a	27	0.655	F1
31 H1a	30	0.074	H1b	30	0.267	H1b	19	0.593	D1
32 I1	39	0.232	P1	18	0.389	D	22	0.590	B
33 J1	34	0.096	J1a	22	0.667	B	34	0.474	J1a
34 J1a	33	0.096	J1	33	0.692	J1	33	0.474	J1
35 L1	6	0.193	A2a1	29	0.500	G1	23	0.617	B1a
36 M1	30	0.229	H1b	5	0.263	A2a	1	0.691	A
37 N	20	0.177	C	39	0.636	P1	38	0.595	O1
38 O1	29	0.226	G1	3	0.412	A1a	37	0.595	N
39 P1	32	0.232	I1	18	0.471	D	32	0.692	I1
Minimum MD*		0.049			0.095			0.471	
Maximum MD		0.257			0.667			0.692	
Mean MD		0.135			0.278			0.576	
Median MD		0.129			0.263			0.591	

*MD = minimum distance

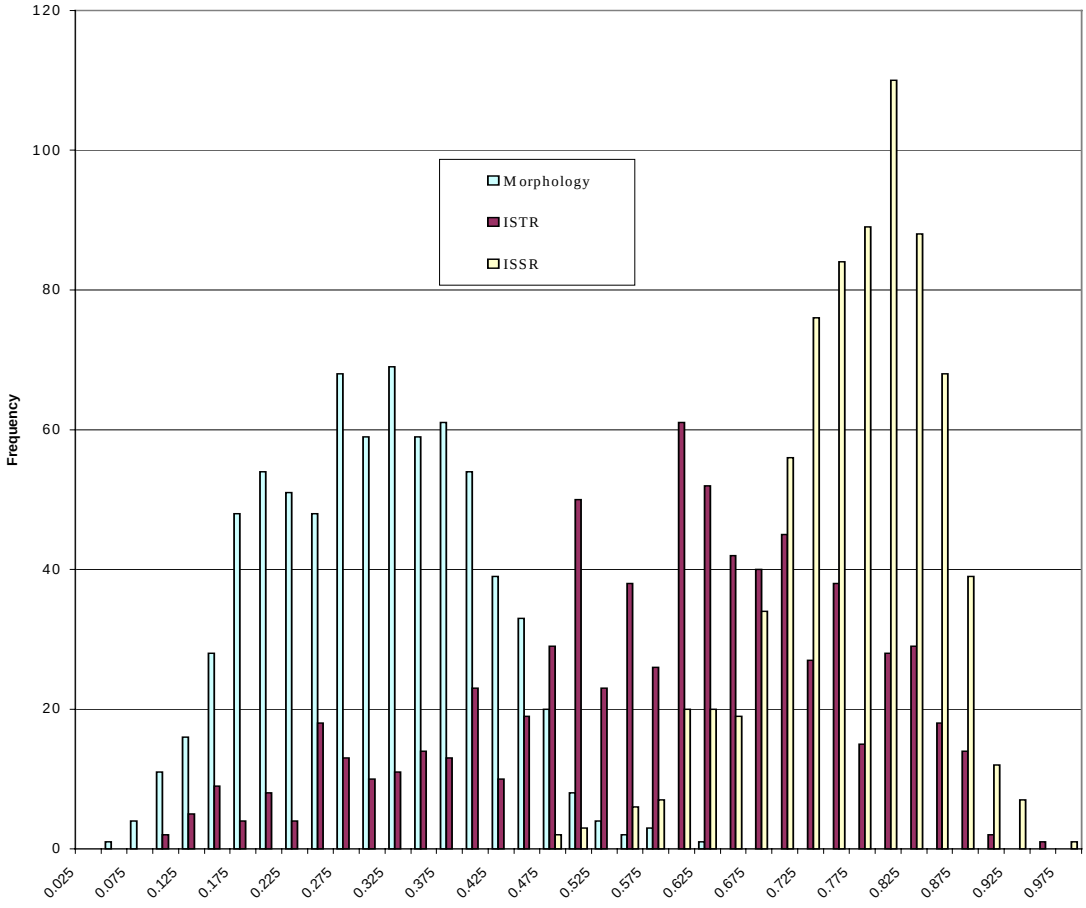
Table 3 – Estimates of within family relatedness for molecular methods and morphology – comparison of the ability of a method to identify each of the varieties in a family group as being within the top x% (where x = 10, 20 or 30) of ranked minimum distances.

	10%	20%	30%
Morphology	33	51	53
ISSR	20	27	32
ISTR	29	40	48
Morphology (2)*	34	45	50

* - flower colour characteristics removed

~~~~

Figure 1 – Summary of pair-wise distance estimates for three different approaches.



[End of document]
~~~~