E

BMT/7/16
ORIGINAL: English
DATE: November 21, 2001

UPOV

INTERNATIONAL UNION FOR THE PROTECTION OF NEW VARIETIES OF PLANTS

GENEVA

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

Seventh Session
Hanover, Germany, November 21 to 23, 2001

EU DEMONSTRATION PROJECT

prepared by experts from the Netherlands

Standardisation of STMS databases for tomato and wheat

Ben Vosman and Gerard Bredemeijer

Department of Biodiversity and Identity
Plant Research International
PO Box 16, 6700 AA Wageningen, The Netherlands

Introduction

A demonstration project entitled 'Molecular markers for variety testing' was carried out within the European Union Biotechnology program. This multi-national project aimed to demonstrate the technical viability of the STMS approach for variety identification and discrimination in two important European crop species: tomato (*Lycopersicon esculentum*) and wheat (*Triticum aestivum*). Initially, a set of approximately 30-40 STMS primer pairs was selected for each crop. These STMS markers had previously shown to reveal polymorphisms within small collections of wheat varieties (Plaschke *et al*. 1995; Röder *et al*. 1995 and 1998) and tomato varieties (Smulders *et al*. 1997; Bredemeijer *et al*. 1998; Areshchenkova and Ganal 1999). Each lab involved used for the detection of STMS polymorphism the system it had already available. This was considered to be a more realistic situation than using the same equipment in all the laboratories involved, as many of these apparatus might not be available anymore in future. The detection systems differed with respect to their mode of action, the costs involved in setting up the detection system and the type of laboratory facilities needed. Automated DNA sequencers were used by PRI (ALF express), IPK (ALF and ALFexpress) and Nunhems (ABI). NIAB used IR-labelled primers in combination with a LI-COR DNA Analyser 4200 whilst Agrogene used a combined fluoro-phosphoimager in conjunction with fluorescently labelled primers. After standardisation of the methodology and selection of 20 primer pairs for each crop, databases containing the molecular description of the most common varieties of these crops grown during the last 10 years in Europe were constructed and tested.

Results and discussion

Standardisation

For each crop species a minimum of 20 primer pairs had to be selected from existing collections of PRI and IPK in a way that the microsatellite technologies used in different laboratories are able to identify alleles in the same way. As several systems for the detection of STMS polymorphism were used it was important to standardise both the methodology (including sampling of material, DNA extraction and estimation, PCR conditions) and interpretation of the results. Standardised protocols for DNA extraction and PCR amplification were agreed upon between the partners (Vosman et al. 2000 and 2001). Two standard sets, of 8 varieties each, were analysed with 30-40 microsatellites and the best primer pairs were selected (Vosman et al. 2000 and 2001). The criteria for choosing a marker included scorability of the patterns, reproducibility of scoring between laboratories, distribution of the markers throughout the genome (map position) and the level of polymorphism detected between varieties.Then an inter-laboratory ring test was conducted to test the methodology and interpretation of the results. All participants analysed the same collection of 16 varieties of both species with the selected markers. In spite of the use of different technologies allele scoring by the partners resulted mostly in the same allele classification of varieties (Vosman et al. 2000).

The discrepancies found between duplicate samples were analysed by rescoring the peak and banding patterns. The number of discrepancies was strongly dependent on the marker (Vosman et al. 2000). If necessary, experiments were repeated in both laboratories and DNA was exchanged between the respective partners. This revealed that the discrepancies were caused by:

Methodological problems (differences in resolution capacity of individual gel systems, echo bands, extra base addition, and thresholds for allelic peaks, missing data, no PCR product, bad quality of DNA isolation, overloading and underloading of gels).

Data entry errors (mis-typing, mis-scoring).

Heterogeneity of samples (residual or hybrid heterozygosity).

The methodological problems could to a large extent be overcome by using gels with a high resolution capacity and by the use of PIG-tailed primers, which circumvents the problem of extra base addition as all fragments get the addition (Brownstein et al. 1996).

Comparing data sets for tomato was relatively difficult due to the occurrence of heterozygotes in hybrids, often in combination with differences in the amount of amplification product for the two alleles. The latter phenomenon may also be due to heterogeneity of the seed samples used or to differential amplification of the alleles. To distinguish between these two possibilities DNA from individual plants have been tested in some cases if the partners scored differently (see examples in Vosman et al. 2000, Fig. 2). In wheat one of the main reasons for the

occurrence of discrepancies between data from different laboratories or data for duplicate varieties seemed to be the occurrence of internal heterogeneity, defined as the identification of more than one allele for a given marker in a single variety.

Database construction

For the construction of the STMS databases of wheat and tomato 500 varieties were collected for each crop from which material was provided by the breeders. These varieties were analysed in duplicate (at 2 laboratories) using the final selection of 20 markers. Each lab analysed a bulked sample of six individual seeds. For a correct allele-recognition, the alleles defined by the initial 22 varieties that were analysed in the standardisation experiments, were used as reference alleles.
Table 1 summarises the results of the analyses that are described in detail in Bredemeijer et al. (2002) and (Röder et al. (2002). As can be seen from this table with the selected set of microsatellites it was possible to discriminate between most of the varieties. In a pairwise comparison more than 99.9% of the varieties can be distinguished. The varieties that did not show a unique banding pattern were mostly obtained from the same breeding company. Also in several cases it turned out that these varieties were clearly related by descent. As the selection of microsatellite markers was not aimed at reaching the maximum discriminative power it is very likely that the addition of a few extra markers or substituting some of the less informative markers for others will improve the discriminative power (Bredemeijer et al. 2002).

Table 1: Discriminative power of wheat and tomato microsatellites

	Wheat	Tomato
Varieties tested	554	521
Potentially different	502	508
Unique patterns	468 (93%)	468 (92%)
Number of pairs	15	18
Number of triplets	0	0
Groups of 4	1	1

For most varieties identical scores were obtained in both labs. Nevertheless, a relatively large number of both tomato and wheat varieties showed discrepancies between the two labs for one or more microsatellite markers . This was to a very large extent due to heterogeneity within the seed samples (Bredemeijer et al. 2002; Röder et al. 2002). A more detailed study involving 36 individual seeds from varieties that are homogenous in the bulked seed sample analysis showed that even in these varieties low levels of heterogeneity can be detected with just a small set (6 to 9) microsatellite markers (for tomato see Vosman and Bredemeijer 2001).

Conclusions

From the data presented above, it can be concluded that STMS databases can be made and used in a reliable way. For the production of the data the use of automated sequencers is recommended. With the set of markers used more than 99.9% of all pairs of varieties can be distinguished. A problem is the high level of heterogeneity within the varieties and some level needs to be accepted.

Acknowledgements

The authors acknowledge the financial support of the European Commission DGXII Biotechnology grant BIO4-CT97-2377.
We also thank the breeding companies for providing the tomato and wheat seeds.

References

Areshchenkova, T. and Ganal, M.W., 1999. Long tomato microsatellites are predominantly associated with centromeric regions. Genome 42: 536-544.

Bredemeijer, G.M.M., Arens, P., Wouters, D., Visser, D. and Vosman, B., 1998. The use of semi-automated fluorescent microsatellite analysis for tomato cultivar identification. Theor.Appl. Genet. 97: 584-590.

Bredemeijer, G.M.M., Cooke, R.J., Ganal, M.W., Peeters, R., Isaac, P., Noordijk, Y., Rendell, S., Jackson, J., Röder, M.S., Wendehake, K., Dijcks, M., Amelaine, M., Wickaert, R., Bertrand, L. and Vosman, B., 2002. Construction and testing of a microsatellite database containing more than 500 tomato varieties. Submitted.

Broun, P. and Tanksley, S.D., 1996. Characterization and genetic mapping of simple repeat sequences in the tomato genome. Mol. Gen. Genet. 250: 39-49.

Brownstein, M. J., Carpten, J. D. and Smith, J.R., 1996. Modulation of non-templated nucleotide addition by *Taq* DNA polymerase: primer modifications that facilitate genotyping. BioTechniques 20: 1004-1010.

Plaschke, J., Ganal, M. W. and Röder, M. S., 1995. Detection of genetic diversity in closely related bread wheat using microsatellite markers. Theor. Appl. Genet. 91: 1001-1007.

Röder, M. S., Plaschke, J., König, S. U., Börner, A., Sorrells, M. E., Tanksley, S. D. and Ganal, M. W., 1995. Abundance, variability and chromosomal location of microsatellites in wheat. Mol. Gen. Genet. 246: 327-333.

Röder, M.S., Korzun, V., Wendehake, K., Plaschke, J., Tixier, M.H., Leroy, P. and Ganal, M.W., 1998. A microsatellite map of the wheat genome. Genetics 149: 2007-2023.

Röder, M.S., Wendehake, K., Korzun, V., Bredemeijer, G., Visser, D., Isaac, P., Rendell, S., Jackson, J., Cooke, R.J., Vosman, B. and Ganal, M., 2002. Construction and analysis of a microsatellite-based database of european wheat varieties. Submitted

Smulders, M. J. M., Bredemeijer, G., Rus-Kortekaas, W., Arens, P. and Vosman, B., 1997. Use of short microsatellites from database sequences to generate polymorphisms among *Lycopersicon esculentum* cultivars and accessions of other *Lycopersicon* species. Theor. Appl. Genet. 94: 264-272.

Vosman, B., Cooke, R., Ganal, M., Peeters, R., Isaac, P., Röder, M., Jackson, J., Rendell, S., Dijcks, M., Kleyn, Y., Visser, D., Wendehake, K., Areshchenkova, T., Korzun, V., Amelaine, M., Wickaert, V. and Bredemeijer, G., 2000. Standardisation of molecular marker systems for variety testing. Document for UPOV working group on Biochemical and Molecular techniques and DNA profiling in particular (BMT/6/12).

Vosman, B., Cooke, R., Ganal, M., Peeters, R., Isaac, P. and Bredemeijer, G., 2001. Standardization and application of microsatellite markers for variety identification in tomato and wheat. Acta Hort. 546: 307-316.

Vosman, B. and Bredemeijer, G., 2001. Construction of the STMS database for tomato. Document for UPOV working group on Biochemical and Molecular techniques and DNA profiling in particular (BMT-TWV Tomato/1/1).

[End of document]