E

BMT/7/10
ORIGINAL: English
DATE: November 6, 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

PRE-SCREENING OF SUGAR BEET VARIETIES USING MICROSATELLITE MARKERS

prepared by experts from Belgium

PRE-SCREENING OF SUGAR BEET VARIETIES USING MICROSATELLITE MARKERS

I. Everaert[1], D. Esselink[2], J. De Riek[1] & B. Vosman[2]
[1] DvP-CLO, Melle, Belgium
[2] Plant Research International, Wageningen, the Netherlands

Introduction

The aim of this research is to develop a well defined set of molecular markers which allow a preliminary characterisation of new sugar beet varieties submitted for PBR during the winter before sowing. This can be an aid in the planning of the next trials and may also give a first indication for distinctness. Three different molecular techniques form part of this pre-screening test: Amplified Fragment Length Polymorphisms (AFLP), Cleaved Amplified Polymorphic Site (CAPS) and Sequenced Tagged Microsatellite Sites (STMS). With 3 AFLP primer combinations, 13 CAPS markers and about 20 STMS markers a set of 40 varieties will be characterised to develop a feasible pre-screening procedure. Data for 8 microsatellite markers on a set of 20 varieties (30 plants per variety) are presented here.

Materials and Methods

Plant material

Twenty sugar beet varieties were included. Seeds of the year 2000 seed delivery were obtained from the KBIVB-Tienen (Belgium). The same seed lot was used as in the official variety trials (Table 1). Thirty individual plants per variety were analysed (in total 600 individual plants). Plants were sown on wetted filter paper. The plant meristem was brought into tissue culture. After two multiplication rounds, plant material was harvested and the culture was stored at 4°C in a cooled growing chamber. This can be further used as reference collection.

Table 1: List of the varieties tested

Variety	Seed company	Type of trial in 2000	Ploidy level
Ariana	KWS	C	3n
Atlantis	Van der Have	C	3n
Aurelia	KWS	C	3n
Claudia	KWS	C	3n
Fortis	Hilleshög	C	2n
Opus	Dieckmann	C	3n
Princesse	Delitzsch	C	3n
Ravel	Kühn	C	3n
Sylvester	Van der Have	C	3n
Winner	Kühn	C	3n
A 8106	Agrosem	R2	3n
DS 3014	Danisco	R2	3n
H 66377	Vanderhave	R2	3n
HM 5432	Hilleshög	R2	3n
KWS 8123	KWS	R2	2n
H 66411	Van der Have	R1	3n
HI 0032	Hilleshög	R1	2n
KWS 9226	KWS	R1	3n
MK 9907	Kühn	R1	3n
S 1901	SES	R1	3n

* R1 = first year of registration trials; R2 =second year of registration trials; Cn = n^{th} year of trial for a listed variety

DNA isolation

At harvest, approximately 1 g fresh weight of leaf material was immediately immersed in liquid nitrogen and subsequently lyophilised during 48 h. The dry material was vacuum-packed for storage at -20°C until DNA extraction. Stored material was ground using a Culatti mechanical mill. The DNA isolation protocol was based on the CTAB method by Doyle and Doyle (1987). To 25 mg lyophilised ground tissue, 1 ml CTAB extraction buffer (100 mM Tris-HCl pH 8, containing 2% CTAB, 20 mM EDTA, 1.4 M NaCl, 0.5 mM $Na_2S_2O_5$, 0.4 % β-mercaptoethanol and 1% PVP MW 40000) and RNase (10 U) was added. Samples were incubated for 40 min. at 65°C. Afterwards, samples were homogenised with 1 ml chloroform/isoamylalcohol (24/1) and centrifuged during 15 min. at 10000g. The supernatant was transferred to a fresh tube and the DNA precipitated with 1 ml of ice-cold (-20°C) isopropanol. After centrifugation (5000g; 15 min.), the pellet was washed with EtOH (76%) - 0.2 M NaOAc, dried and dissolved in water. DNA concentration and quality was constantly checked compared to a standard series of lambda-DNA on a 1.5% TAE buffered agarose gel after electrophoresis.

Microsatellite analysis

Microsatellite containing loci were sequenced from an enriched library made by Plant Research International. In the flanking regions of the microsatellite locus primers suitable for direct PCR amplification were designed using Primerselect 5.0 (DNAstar inc.). Amplifications were performed using a Perkin-Elmer 9600. The reaction conditions were first optimised for annealing temperature, template and primer concentration and number of PCR cycles. According to corresponding reaction conditions 3 multiplex sets, each containing 3 microsatellite loci labelled with different fluorescent dyes (FAM, HEX, NED) were composed. The amplified fragments were separated on an ABI Prism 377 DNA Sequencer using 36 cm gels (4.25% denaturing polyacrylamide (4.25% acrylamide/bisacrylamide 19/1, 6 M urea in 1 X TBE)). GS-500 ROX labelled size standard (Perkin Elmer) was loaded in each lane in order to facilitate the automatic analysis of the gel and the sizing of the fragments. During a run on the ABI 377, the fluorescent signal in each lane is being recorded continuously. Genescan 2.1 was used to estimate detection time, signal peak height and surface for each fragment. Sizing of the fragments was performed by the Genescan software module by interpolation to the internal lane standard according to the Local Southern algorithm (as recommended by the manufacturer).The Genotyper software was used for allele calling and scoring of the PCR products Alleles were scored as present or absent (no complete co-dominant scoring).

Results and Discussion

Characteristics of the set of STMS loci

Up to now, only the data from 9 microsatellite loci on a set of 20 varieties are completely analysed. One STMS locus appeared to be not polymorphic. Some properties of the 8 loci that were polymorphic in the set of tested varieties are listed in Table 2.

Table 2: Characteristics of the used set of STMS loci

Primer	Type of repeat	Number of alleles	Null alleles detected in dataset	Rare alleles (f < 0.01)	PIC*
Bvv 17	Tri	3	Yes	0	0,66
Bvv 37	Di	34	Yes	29	0,89
Bvv 43	Di	10	Yes	8	0,63
Bvv 48	Di	4	No	2	0,56
Bvv 51	Di	6	No	2	0,73
Bvv 53	Di	10	No	6	0,77
Bvv 60	Tri	5	Yes	3	0,52
Bvv 61	Di	19	Yes	17	0,67

*PIC: Polymorphic Information Content = $1-\Sigma(f_i)^2$ with f_i the frequency of the i^{th} allele

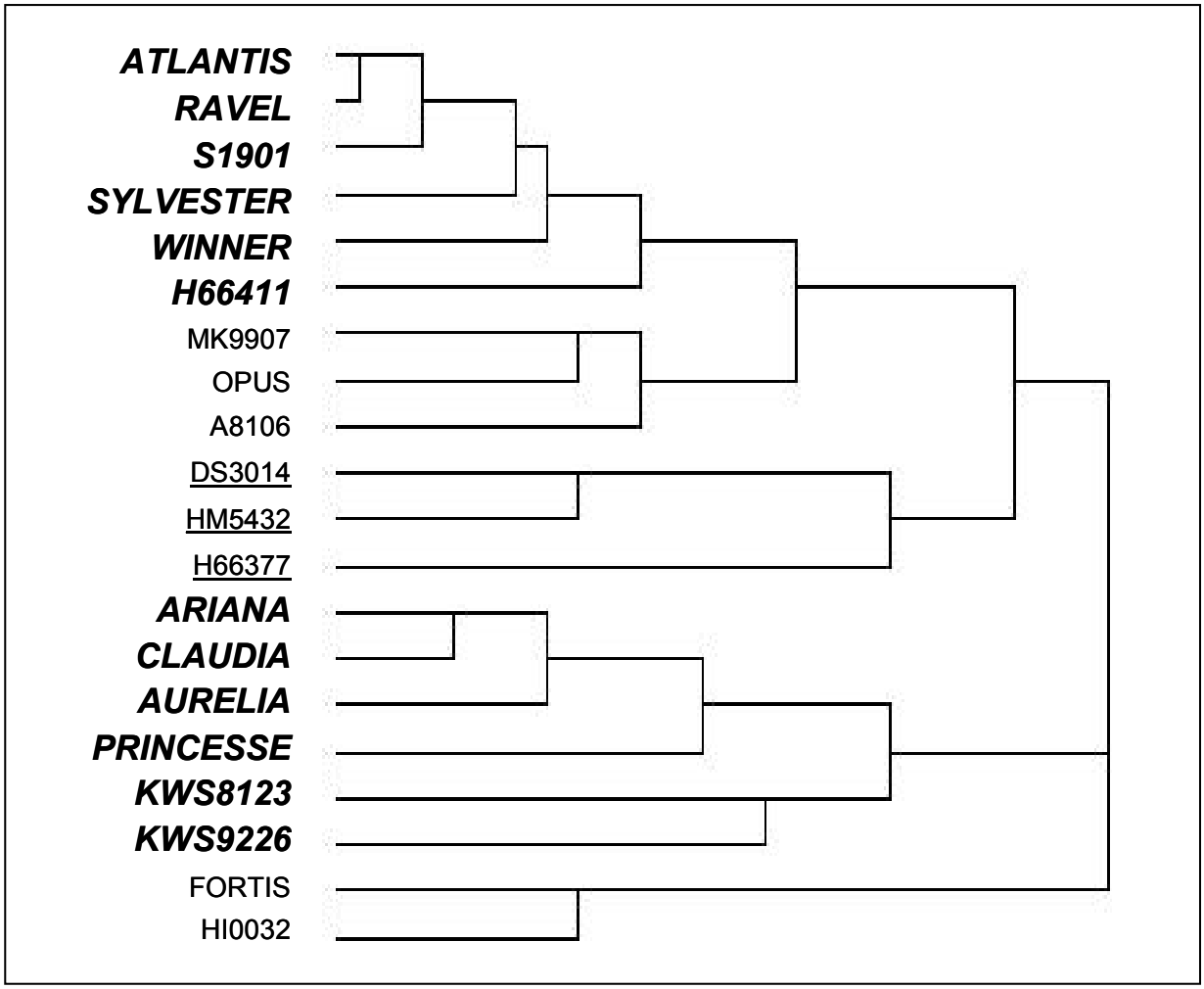
As there are 6 dinucleotide repeat type of markers in the STMS set, which are the most polymorphic, the number of alleles detected can be very high (e.g. Bvv 37 generated 34 alleles). However, often rare alleles are found. However, the Polymorphic Information Content of the used markers is in general satisfactory. For a marker locus having many alleles that are equally distributed, the PIC value will approach to 1 (the highest informative loci). Both a low number of alleles and an unequal distribution will lower the PIC value. Nevertheless, with this limited set of 8 STMS loci in total 91 alleles are detected; 24 of these have an allele frequency above 0.01.
As a first entrance to the data, the number of alleles detected per variety is listed (Table 3). Because of the high number of rare alleles, the number of alleles with an allele frequency (within each variety) below 0.1 are also indicated. Both if one looks to the total number of alleles present or to the useful number of alleles (the total minus the number of rare alleles) the diploid varieties KWS 8123, Fortis and Hi 0032 show the lowest number of alleles. Among the triploid varieties, a ranking to allele numbers is little discriminative. Alleles that are specific to one or a small number of varieties could be detected, however, this were most often rare alleles (within the total dataset and also within the varieties). An identification solely based on the presence and absence of variety specific alleles was not possible with this set of microsatellites used so far.

Table 3: The number of alleles detected per variety (30 individual plants analysed) with indication of the number of rare alleles (f < 0.1)

ras	Bvv 17	Bvv 37	Bvv 43	Bvv 48	Bvv 51	Bvv 53	Bvv 60	Bvv 61
A 8106	3(0)	13(9)	7(4)	3(0)	4(2)	5(1)	2(1)	5(2)
Ariana	3(0)	7(3)	6(1)	3(1)	4(1)	4(2)	5(3)	7(4)
Atlantis	3(0)	13(9)	5(3)	3(1)	5(1)	6(1)	2(1)	5(2)
Aurelia	2(0)	11(7)	5(2)	2(0)	5(2)	3(0)	4(1)	3(1)
Claudia	3(0)	12(9)	5(1)	3(1)	5(1)	7(4)	2(1)	8(5)
DS 3014	3(1)	6(3)	2(0)	3(0)	4(1)	3(1)	2(0)	6(3)
Fortis	3(1)	6(4)	3(0)	4(2)	4(1)	3(1)	3(1)	7(5)
H 66377	3(0)	11(7)	8(4)	3(1)	4(2)	4(1)	2(0)	6(3)
H 66411	3(0)	12(8)	4(3)	3(0)	5(0)	6(2)	2(0)	3(1)
HI 0032	3(1)	7(5)	4(2)	2(0)	3(0)	5(3)	2(1)	6(2)
HM 5432	3(0)	6(2)	5(2)	3(0)	4(1)	3(0)	4(2)	3(1)
KWS 8123	2(1)	4(2)	2(0)	2(0)	3(1)	2(0)	2(0)	3(1)
KWS 9226	2(1)	9(7)	5(1)	3(1)	3(0)	4(1)	4(2)	4(1)
MK 9907	3(1)	9(6)	6(2)	3(1)	5(2)	5(2)	4(2)	5(2)
Opus	3(0)	12(7)	6(4)	3(1)	4(1)	6(0)	3(1)	8(5)
Princesse	3(0)	10(7)	5(2)	2(0)	5(1)	4(0)	2(0)	6(3)
Ravel	3(0)	13(10)	3(2)	2(0)	5(1)	5(1)	2(0)	4(2)
S 1901	3(0)	14(10)	5(3)	2(0)	4(1)	6(1)	3(1)	4(1)
Sylvester	3(0)	11(8)	5(4)	3(1)	4(0)	4(0)	2(1)	7(3)
Winner	3(0)	12(9)	2(0)	2(0)	4(0)	5(1)	2(0)	4(2)

Classifications based on ordinations of the tested varieties

A first ordination was derived from the average presence or absence of alleles in each of the individual plants for each variety. For each allele, the total number of presences in each accession (30 plants per variety) was counted. Based on these values, Euclidean distances were calculated between the varieties and a dendrogram was constructed using the UPGMA-algorithm (Figure 1).
Five clusters could be distinguished: 1.) a group containing “Atlantis”, “Ravel”, “S1901”, “Sylvester”, “Winner” and “H66411”; 2.) linked to the first cluster “MK9907”, “Opus” and “A8106”; 3.) an intermediate group with “DS3014”, “HM5432” and “H66377”; 4.) a group containing “Ariana”, “Claudia”, “Aurelia”, “Princesse”, “KWS8123” and “KWS9226”; and 5.) a last cluster with “Fortis” and “HI0032”. From this ordination it is clear that varieties belonging to the same or related breeding programs tend to cluster together, indicating the genetic relatedness of these varieties. From some varieties studied earlier with AFLP (De Riek *et al.*, 2001) the grouping obtained here corresponds to the one based on AFLP markers.
Classifications based on the individual plant data were able to reveal larger clusters of related



plants, merely belonging to the same variety; however, plants also tended to be grouped across varieties (data not shown).

Figure 1: Dendrogram based on Euclidean distances between the total number of presences of each microsatellite allele (UPGMA-clustering)

Assignment tests based on individual plant data

Estimates for the within variety genetic variation can also directly be assessed from the pair wise resemblance data for individual plants (Jaccard similarity coefficients were used on the presence/absence data for each microsatellite allele). The distribution for the assignment of individual genotypes to a certain variety is given in Table 4. In the top panel the assignment of the 10 most similar partners is given, below the assignment of the 100 most similar ones. Table 4 must be read horizontally: e.g. for all individual plants analyzed from 'A8106', the 10 most similar partners were 66 times tracked back to 'A8106' itself, 2 times to 'Ariana', 15 times to 'Atlantis', 6 times to 'Aurelia' and so on.

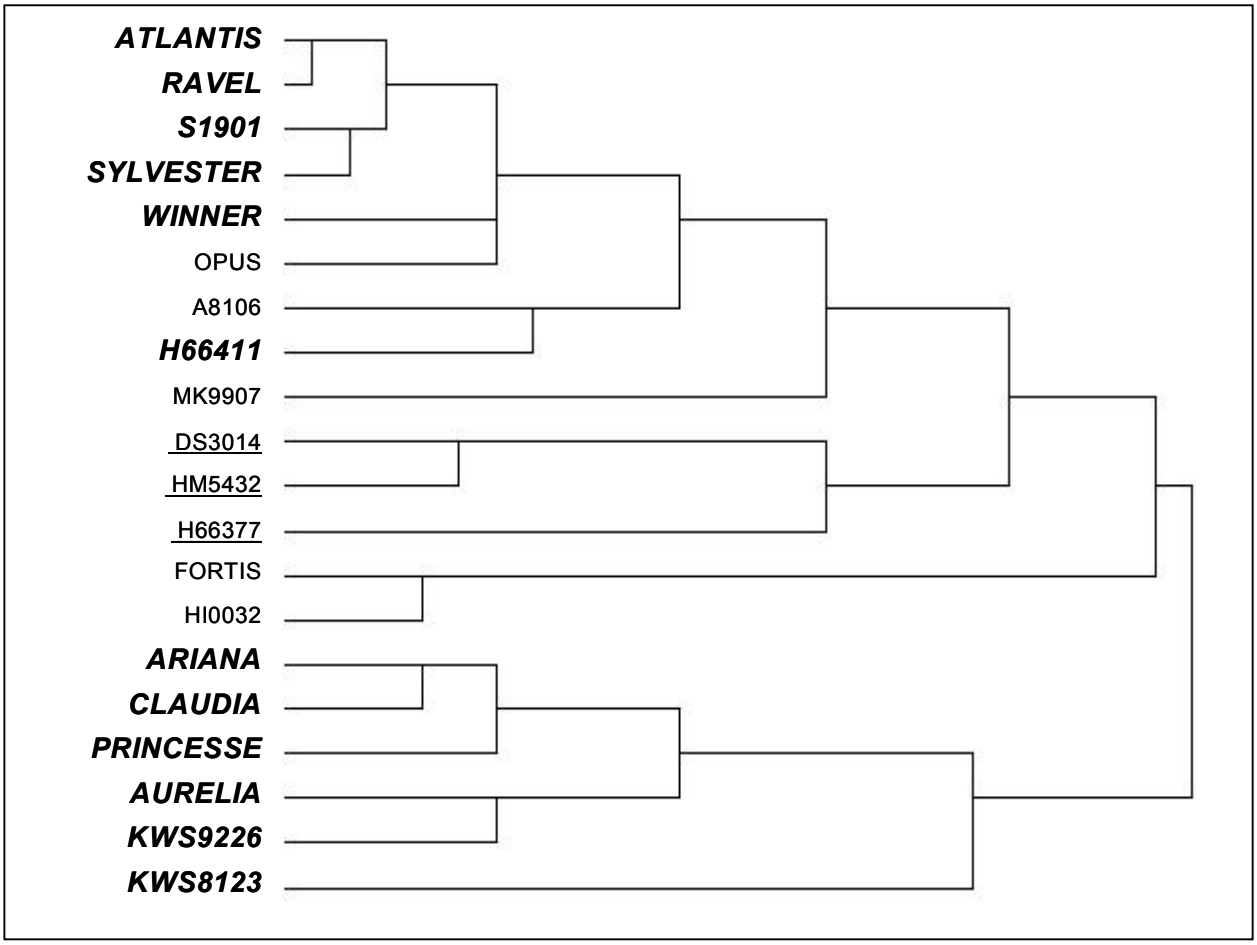


Figure 2: Dendrogram based on the top 100 assignment (Table 4). Euclidean distances between the assignment values to each variety (UPGMA-clustering)

The assignment tests reveal differences between varieties in the degree they are cross attributing to each other. "DS3014" and "KWS8123" can be taken as examples of varieties that are very good distinguishable from the others although they might refer to a common gene pool (e.g. "KWS8123 clearly refers to the KWS breeding program). On the other hand varieties like "Opus", "Atlantis" and "Ravel" refer to many other varieties, although clearly related to the Vander Have breeding gene pool. If this is an indication for a broader genetic background of these varieties can not be confirmed from this type of data.

To demonstrate that the assignment tests reveal in another way the same genetic structure as the previous analysis based on allele frequencies, the tabular data from the Top 100 assignment were converted into a dendrogram (Figure 2). In this ordination more or less the same structure of grouping is retained as in Figure 1.

Table 4: Assignment test showing the distribution of the 10 and 100 most similar genotypes among the different varieties

Top 10	A8106	Ariana	Atlantis	Aurelia	Claudia	DS3014	Fortis	H66377	H66411	HI0032	HM5432	KWS8123	KWS9226	MK9907	Opus	Princesse	Ravel	S1901	Sylvester	Winner
A8106	**66**	2	15	6	3	18	14	20	15	12	26	12	5	25	15	2	16	6	8	14
Ariana	1	**60**	2	51	36	12	6	1	10	3	40	9	40	3	2	13	1	1	7	2
Atlantis	13		**39**	6	9	6	7	5	16	1	3	6	6	15	30		38	27	29	44
Aurelia	7	41	1	**90**	17	11	5		11	2	19	21	49	5	1	10	2	2	2	4
Claudia	5	32	5	27	**78**	12	5	1	6	7	34	12	20	3	4	36	5	2	2	4
DS3014	3	1		1		**172**	2	8	7	9	69	1			6	1	1	3	9	7
Fortis	4	1		2	6	21	**131**	2	5	82	19	8	5		2		2		1	7
H66377	13	1	7	2	1	38	5	**112**	3	6	16	6		8	26		8	4	6	38
H66411	12	8	15	10	4	21	13	3	**65**	6	45	12	17	1	7	2	21	13	11	13
HI0032	8	1	2	2	5	13	53	3	2	**180**	3	15	1				1	3	2	2
HM5432	6	7		3	2	72	10	12	8	1	**153**	3		1	5	1	4		4	8
KWS8123	1	13	1	8	3	4	9		3	13	3	**214**	22			4				
KWS9226	4	35	3	42	10		7		9	1	2	38	**114**	1	1	17	2	1	1	2
MK9907	31	4	17	12	3	6	1	7	8		6	10	1	**137**	22	3	13	1	12	6
Opus	14	2	38	1	6	22	2	14	10		12			37	**43**	3	30	4	19	23
Princesse		25	3	19	49	9		1	12		21	11	23	1	5	**99**	5	1	8	8
Ravel	17		28	1	6	10	4	12	27		16	2	2	12	23	3	**45**	38	29	25
S1901	5	1	26	5	2	17	3	6	30	6	19	1	1	1	7	1	44	**57**	36	32
Sylvester	10		24		2	36	3	6	10	6	17	3	2	9	15	2	36	23	**76**	20
Winner	5	1	48	1	1	20	9	22	12	9	20			9	13		23	22	17	**68**

Top 100	A8106	Ariana	Atlantis	Aurelia	Claudia	DS3014	Fortis	H66377	H66411	HI0032	HM5432	KWS8123	KWS9226	MK9907	Opus	Princesse	Ravel	S1901	Sylvester	Winner
A8106	**332**	59	129	106	59	214	112	187	195	100	269	132	81	236	145	34	168	94	152	196
Ariana	89	**391**	49	332	300	181	66	69	160	69	333	133	240	49	43	250	68	53	68	57
Atlantis	138	37	**284**	85	94	120	68	119	199	60	128	76	77	136	212	47	324	262	232	302
Aurelia	104	280	32	**437**	237	126	84	55	201	76	244	293	296	110	42	170	63	34	54	62
Claudia	71	260	72	256	**306**	192	87	86	174	83	250	152	190	99	76	245	106	89	114	92
DS3014	121	63	81	65	104	**666**	96	239	161	168	528	40	27	38	84	44	108	75	148	144
Fortis	121	49	73	98	93	262	**455**	98	196	455	308	191	96	27	44	34	86	89	67	156
H66377	198	42	121	53	66	392	58	**490**	142	57	322	86	36	88	163	30	151	106	104	295
H66411	156	107	133	154	101	204	122	98	**300**	109	291	146	112	85	96	84	203	169	167	162
HI0032	100	60	58	116	84	274	369	72	166	**609**	278	254	98	16	27	17	60	86	101	151
HM5432	150	135	54	112	121	505	152	150	221	122	**619**	47	35	42	80	60	73	63	113	146
KWS8123	137	140	36	344	153	63	164	95	181	248	94	**662**	348	99	16	94	53	20	21	30
KWS9226	110	242	50	359	217	60	110	58	185	64	107	382	**462**	49	19	201	76	36	41	72
MK9907	300	63	169	160	133	161	38	119	141	39	177	129	69	**466**	234	48	199	59	171	125
Opus	160	32	242	69	116	170	38	154	147	33	144	43	27	229	**266**	71	226	169	242	222
Princesse	52	272	60	243	318	166	37	57	191	22	254	117	232	63	84	**458**	77	79	98	120
Ravel	152	43	249	64	97	169	52	131	236	38	168	40	63	124	172	46	**326**	279	279	272
S1901	103	46	229	50	108	196	43	132	237	77	187	28	43	49	155	75	296	**330**	306	310
Sylvester	133	45	198	59	101	251	31	95	202	94	180	38	46	105	189	53	282	285	**377**	236
Winner	146	22	234	48	69	215	105	223	197	110	242	34	31	75	160	47	226	233	197	**386**

Conclusions

Although still preliminary, because of the partial results, it can already be concluded from the presented data that microsatellite analysis offers comparable results for the analysis of sugar beet varieties as was obtained before with AFLP analysis (De Riek *et al.*, 2001) presented before.
Nevertheless, as can already be concluded from a subset of samples that were analysed a second time, the reproducibility of the PCR analysis and scoring is higher than for AFLP samples. This facilitates the construction of databases over different testing years.
The data set is to be enlarged with another set of 9 to 10 microsatellites; the total number of varieties to be analysed is 40. This data will then be compared with results from AFLP and CAPS analysis (taken from Schneider *et al.*, 1999).

Acknowledgements

The authors wish to thank the KBIVB-IRBAB for the good co-operation and Laurence Desmet and Veerle Cools for their efforts in doing DNA preparations and microsatellite analysis. The Belgian Ministry of Small Enterprises, Traders and Agriculture DG4 funded this research.

References:

De Riek J., Calsyn E., Everaert I., Van Bockstaele E. & De Loose M. (2001) AFLP based alternatives for the assessment of Distinctness, Uniformity and Stability of sugar beet varieties, Theor. Appl. Genet. (in press)

Schneider, K., Borchardt, D.C., Schäfer-Pergl, R., Nagl, N., Glass, C., Jeppsson, A., Gebhardt, C. & Salamini, F. (1999) PCR-based cloning and segreation analysis of functional gene homologues in *Beta vulgaris*. Mol. Gen. Genet. 262: 515-524.

[End of document]