*Diversity* **2010**, *2*, 439-449; doi:10.3390/d2040439

OPEN ACCESS

*diversity*

**ISSN 1424-2818**

www.mdpi.com/journal/diversity

*Article*

# Genetic Variability of Macedonian Tobacco Varieties Determined by Microsatellite Marker Analysis

**Katarina Davalieva [1], Ivana Maleva [1], Kiril Filiposki [2], Ognen Spiroski [1] and Georgi D. Efremov [1,*]**

[1] Research Centre for Genetic Engineering and Biotechnology, Macedonian Academy of Sciences and Arts, Krste Misirkov 2, POB 428, 1000 Skopje, Republic of Macedonia; E-Mails: katarina@manu.edu.mk (K.D.); maleva_i@yahoo.com (I.M.); ognens@gmail.com (O.S)

[2] University "St Kliment Ohridski", Bitola, Tobacco Institute, Kicevski pat bb, 7500 Prilep, Republic of Macedonia; E-Mail: tobacco_institute_prilep@yahoo.com

* Author to whom correspondence should be addressed: E-Mail: gde@manu.edu.mk; Tel.: +389-2-3235-411; Fax: +389-2-3115-434.

*Received: 20 January 2010; in revised form: 4 February 2010 / Accepted: 2 March 2010 / Published: 24 March 2010*

---

**Abstract:** Tobacco (*Nicotiana tabacum* L.) is an important agricultural crop plant for the economy of many countries. Assessment of the genetic diversity of cultivated tobacco varieties is of importance for long-term tobacco improvement. Microsatellite markers are currently the marker system of choice for genetic analysis of allopolyploid plants. In this study, we evaluated the use of 30 microsatellite markers for identification of 10 varieties of cultivated tobacco in the Republic of Macedonia. We found 24 of the microsatellite markers to be polymorphic and sufficient for identification of these varieties. Cluster analysis showed that Macedonian tobacco varieties are classifiable into three distinct groups.

**Keywords:** *Nicotiana tabacum* L; PCR; microsatellite markers; tobacco varieties; genetic diversity

---

## 1. Introduction

Tobacco (*Nicotiana tabacum* L.) is the most important non-food crop in the world's agriculture and is cultivated in more than 100 countries on approximately 4.2 million hectares of crop land (http://www.fao.org/docrep/006/y4956e/y4956e00.htm). Tobacco is an amphidiploid and one of the 64 species in the genus *Nicotiana* that carries 48 chromosomes in its somatic cells. Tobacco is considered to be an interspecific hybrid between *Nicotiana sylvestris* (n = 12) and *Nicotiana tomentosiformis* (n = 12) [1,2]. Despite its high economic importance, tobacco is also considered to be one of the most important model systems in plant biotechnology and highly promising for the production of commercially important substances such as medical drugs and vaccines [3,4].

Tobacco is an important commercial crop in the Republic of Macedonia. Cultivated varieties in Macedonia are highly polymorphic with a wide range of morphological types, and are mainly used in cigarette manufacturing. There are three main groups of cultivated tobacco in the Republic of Macedonia: oriental, semioriental and broad-leaf [5]. Oriental varieties are small leaf, highly aromatic tobacco varieties that are divided into three sub-groups: Prilep, Yaka and Djebel. Prilep varieties are derived from the local variety Djumaj-bale from Upper Djumaja (Republic of Bulgaria) and were introduced in Macedonia in the 30 s of the last century. Yaka varieties originate from Xanthian Yaka (Xanthy, Greece), which was derived from Cuban large-leaf local variety. They were introduced in Macedonia in the first half of the last century. Djebel varieties descend from cultivated Xanthian Yaka in the Djebel region, in the Republic of Bulgaria. Semioriental varieties are represented by Otlia varieties, descending from local Greek varieties, and are grown in the northern parts of the country, in much smaller quantities that other tobacco varieties. Broad-leaf varieties are represented by Virginia and Burley varieties that descend from America and were brought to Macedonia at the end of 1960.

The determination of the genetic diversity in tobacco has a high potential use for variety identification and in breeding programs. Prior to the availability of DNA markers, variety identification of tobacco was based only on morphological markers. Restriction fragment length polymorphism (RFLP) markers were among the first molecular markers applied in tobacco research and were used to study the function of particular cloned genes [6]. PCR based random amplified polymorphic DNA (RAPD) markers have been applied to identify tobacco chromosome regions carrying disease resistance genes [7,8] and for variety identification purposes [9-12]. Amplified fragment length polymorphism (AFLP) markers were also used to distinguish among related varieties of cultivated tobacco [13,10]. The use of microsatellite markers in identification of tobacco varieties was first reported by Bindler *et al.* [14]. Microsatellites, also known as simple sequence repeats (SSRs), have several advantages over the other molecular markers, because of their high variability, linkage to traits and linkage maps to specific chromosomes. Bindler *et al.* [14] reported 637 functional microsatellite markers in tobacco, of which 282 were highly polymorphic and useful for identification of closely related varieties of tobacco.

Until now, only limited information has been available on the relationship between morphological variability and genetic diversity in cultured tobacco. However, attempts are made to examine the degree of relatedness among tobacco cultivars at the genetic level. RAPD and AFLP markers analysis was used to assess the polymorphism, similarities and relationships among flue-cure tobacco cultivars in China [15]. Genetic diversity and genotyping of Indian FCV and burley tobacco cultivars was also

investigated using RAPD technique [12]. AFLP analysis was used to determine genetic variation in 54 varieties of cultivated tobacco in India [16]. Molecular diversity of cultivated tobacco from USA, as well as its changes over decades of cultivar development, was determined by microsatellite markers analysis [17]. Comparative studies on the genetic relatedness of cultivated tobacco varieties from USA, Central and South America were also done by microsatellite markers analysis [18].

Although approximately 25 tobacco varieties have been developed and cultivated in the last century in the Republic of Macedonia, their classification until now was based only on morphological, biological and technological characteristics such as selection based on pest resistance, morphological traits and leaf quality parameters. In this study, we have evaluated the use of 30 microsatellite markers [14] for the identification of 10 varieties of oriental and broad-leaf cultivated tobacco that are currently grown in different tobacco growing areas in the Republic of Macedonia.

## 2. Results and Discussion

The application of the molecular marker technology to the cultivated tobacco, especially microsatellite markers in the last few years, has provided fast and reliable methods for variety identification. By our experience, the use of microsatellite markers in variety identification is highly reproducible and is not affected by the quality of DNA or type of material used for DNA extraction (green leafs, dry leafs or seeds). On the other hand, identification of cultivated tobacco varieties using techniques such as AFLP and RAPD are reported to be highly affected by the quality of DNA and not able to provide a clear distinction of the individual groups [10].

The size of amplified PCR products (ranging from 149−345 bp) using 30 SSR primer pairs is shown in Table 2. Among microsatellite markers used, 24 were polymorphic (PT20021, PT20165, PT20172, PT20176, PT20242, PT20388, PT20445, PT30005, PT30087, PT30096, PT30138, PT30144, PT30150, PT30164, PT30274, PT30378, PT30480, PT30021, PT30077, PT30392, PT40021, PT40035, PT30417, PT40015), which represents 80% of the selected markers.

A total of 90 alleles for 30 SSR loci were detected among 10 tobacco varieties (Table 1). The number of alleles per locus ranged from one (PT1193, PT30160, PT30188, PT30321, PT30255 and PT30463) to six (PT20172, PT30021 and PT40015) with an average of three alleles per locus (Table 1). The highest polymorphic information content (PIC) showed dinucleotide and trinucleotide markers, while markers with repeats consisting of more than three nucleotides were monomorphic.

An allele that was observed once in the 10 genotypes was considered rare. A total of 13 rare alleles were observed at nine of the 30 SSR loci (Table 1). These rare alleles were present in only two tobacco genotypes (Virginia MV-1 and Burley B-2/93), which can be attributed to the low number of broad leaf varieties used in this study.

On average, 70% of the varieties shared a common allele. There was considerable range in highest allele frequency, as illustrated by the fact that at PT20242 locus, the highest frequency allele was present in only 30% of the genotypes, while PT1193, PT30160, PT30188, PT30231, PT30255 and PT30463 loci were monomorphic.

**Table 1.** Data on the number of alleles, number of rare alleles, allele size range, highest frequency allele and polymorphic information content (PIC) found among 10 tobacco varieties for 30 microsatellite markers.

| | | | | Highest frequency allele | | |
|---|---|---|---|---|---|---|
| **Marker** | **No. of alleles** | **No. of rare alleles** | **Size range (bp)** | **Size (bp)** | **Frequency (%)** | **PIC Value** |
| PT 1193 | 1 | 0 | 228 | 228 | 100 | 0.00 |
| PT 20021 | 2 | 0 | 342–345 | 342 | 90 | 0.18 |
| PT 20165 | 2 | 1 | 202–205 | 202 | 95 | 0.10 |
| PT 20172 | 6 | 2 | 196–244 | 196 | 45 | 0.71 |
| PT 20176 | 2 | 0 | 255–261 | 255, 261 | 50 | 0.50 |
| PT 20242 | 5 | 1 | 195–216 | 210, 213 | 30 | 0.76 |
| PT 20388 | 3 | 0 | 182–191 | 191 | 80 | 0.34 |
| PT 20445 | 3 | 0 | 196–271 | 202, 271 | 40 | 0.64 |
| PT 30005 | 4 | 2 | 227–254 | 251 | 75 | 0.41 |
| PT 30021 | 6 | 2 | 222–251 | 241 | 40 | 0.73 |
| PT 30077 | 3 | 1 | 205–213 | 205 | 65 | 0.49 |
| PT 30087 | 2 | 0 | 177–180 | 180 | 90 | 0.18 |
| PT 30096 | 4 | 0 | 229–241 | 238 | 40 | 0.70 |
| PT 30138 | 3 | 1 | 216–222 | 216 | 85 | 0.27 |
| PT 30144 | 4 | 0 | 262–286 | 262 | 70 | 0.48 |
| PT 30150 | 2 | 0 | 186–189 | 189 | 65 | 0.46 |
| PT 30160 | 1 | 0 | 181 | 181 | 100 | 0.00 |
| PT 30164 | 2 | 0 | 152–155 | 152 | 80 | 0.32 |
| PT 30188 | 1 | 0 | 155 | 155 | 100 | 0.00 |
| PT 30231 | 1 | 0 | 180 | 180 | 100 | 0.00 |
| PT 30255 | 1 | 0 | 225 | 225 | 100 | 0.00 |
| PT 30274 | 5 | 2 | 193–208 | 202 | 45 | 0.67 |
| PT 30378 | 3 | 0 | 219–228 | 219 | 60 | 0.56 |
| PT 30392 | 5 | 1 | 253–271 | 253 | 35 | 0.70 |
| PT 30417 | 2 | 0 | 182–191 | 182 | 90 | 0.18 |
| PT 30463 | 1 | 0 | 149 | 149 | 100 | 0.00 |
| PT 30480 | 4 | 0 | 153–171 | 159 | 60 | 0.58 |
| PT 40015 | 6 | 0 | 156–172 | 168 | 40 | 0.76 |
| PT 40021 | 4 | 0 | 151–182 | 154 | 50 | 0.64 |
| PT 40035 | 2 | 0 | 196–199 | 199 | 80 | 0.32 |
| Total | 90 | 13 | | | | |
| Mean | 3 | 0.43 | | | 70.00 | 0.39 |

**Table 2.** Size (bp) of PCR fragments from 30 microsatellite regions in 10 studied tobacco varieties.

| No. | Marker | Djebel | P 12-2/1 | P 23 | Jb-125/3 | Jk-48 | P-80pt | P-66-9/7 | NS-72 | B-2/93 | V MV1 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | PT1193 | 228 | 228 | 228 | 228 | 228 | 228 | 228 | 228 | 228 | 228 |
| 2 | PT20021 | 345 | 342 | 342 | 342 | 342 | 342 | 342 | 342 | 342 | 342 |
| 3 | PT20165 | 202 | 202 | 202 | 202 | 202 | 202 | 202 | 202 | 202/205 | 202 |
| 4 | PT20172 | 220 | 196 | 196 | 226 | 220 | 196 | 196 | 196/220 | 190/223 | 244 |
| 5 | PT20176 | 255 | 261 | 261 | 261 | 261 | 255 | 255 | 261 | 255 | 255 |
| 6 | PT20242 | 210 | 213 | 213 | 216 | 216 | 213 | 210 | 210 | 195/204 | 195 |
| 7 | PT20388 | 191 | 191 | 191 | 191 | 191 | 191 | 191 | 191 | 185 | 182 |
| 8 | PT20445 | 271 | 202 | 202 | 271 | 271 | 202 | 202 | 271 | 196 | 196 |
| 9 | PT30005 | 254 | 251 | 251 | 251 | 251 | 251 | 251 | 251 | 227/254 | 230/251 |
| 10 | PT30021 | 231 | 241 | 241 | 231 | 231 | 241 | 243 | 241 | 225/251 | 222 |
| 11 | PT30077 | 205 | 205 | 205 | 205/207 | 205/207 | 205 | 205 | 207 | 205/213 | 207 |
| 12 | PT30087 | 180 | 180 | 180 | 180 | 180 | 180 | 180 | 180 | 180 | 177 |
| 13 | PT30096 | 232 | 238 | 238 | 232 | 232 | 238 | 241 | 238 | 229 | 229 |
| 14 | PT30138 | 216 | 216 | 216 | 216 | 216 | 216 | 216 | 216 | 216/222 | 219/222 |
| 15 | PT30144 | 262 | 262 | 262 | 262 | 262 | 266 | 262 | 262 | 286 | 270 |
| 16 | PT30150 | 186 | 189 | 189 | 186 | 186 | 189 | 189 | 189 | 186/189 | 189 |
| 17 | PT30160 | 181 | 181 | 181 | 181 | 181 | 181 | 181 | 181 | 181 | 181 |
| 18 | PT30164 | 152 | 152 | 152 | 152 | 152 | 155 | 155 | 152 | 152 | 152 |
| 19 | PT30188 | 155 | 155 | 155 | 155 | 155 | 155 | 155 | 155 | 155 | 155 |
| 20 | PT30231 | 180 | 180 | 180 | 180 | 180 | 180 | 180 | 180 | 180 | 180 |
| 21 | PT30255 | 225 | 225 | 225 | 225 | 225 | 225 | 225 | 225 | 225 | 225 |
| 22 | PT30274 | 208 | 202 | 202 | 208 | 208 | 202 | 202 | 202/208 | 196/199 | 193 |
| 23 | PT30378 | 219 | 219 | 219 | 228 | 228 | 219 | 222 | 222 | 219 | 219 |
| 24 | PT30392 | 271 | 256 | 256 | 253 | 253 | 256 | 256 | 253 | 253/259 | 265 |
| 25 | PT30417 | 182 | 182 | 182 | 182 | 182 | 182 | 182 | 182 | 182 | 191 |
| 26 | PT30463 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 |
| 27 | PT30480 | 171 | 153 | 159 | 159 | 159 | 153 | 162 | 159 | 159 | 159 |
| 28 | PT40015 | 158 | 168 | 168 | 156 | 156 | 168 | 168 | 170 | 172 | 160 |
| 29 | PT40021 | 154 | 151 | 151 | 154 | 154 | 151 | 154 | 154 | 182 | 179 |
| 30 | PT40035 | 199 | 199 | 199 | 199 | 199 | 199 | 199 | 199 | 196 | 196 |

The level of polymorphism of the studied microsatellite loci was evaluated by calculating PIC values for each of the 30 SSR loci (Table 1). PIC values varied widely among loci and ranged from 0.0 (PT1193, PT30160, PT30188, PT30231, PT30255 and PT30463) to 0.76 (PT20242 and PT40015), with an average value of 0.39 per locus.

We identified several microsatellite markers (PT20172, PT20242, PT20388, PT20445, PT30005, PT30021, PT30096, PT30138, PT30144, PT30274, PT30392, PT40015 and PT40021) that could readily distinguish the broad-leaf varieties from oriental varieties (Table 2). Even more specifically, 14 markers (PT20021, PT20172, PT20176, PT20242, PT20445, PT30005, PT30021, PT30096, PT30150, PT30274, PT30392, PT30480, PT40015 and PT40021) can distinguish between Djebel and Prilep varieties. In addition, we identified one marker (PT30480) that distinguishes between very closely related Prilep varieties (P 12-2/1 and P 23). These SSRs can be used to prepare a DNA fingerprint database of the existing and newly developed tobacco varieties in the Republic of Macedonia or in the region, which can be used for varietal registration and protection of plant breeders' rights.

The genetic distances among the 10 tobacco varieties were calculated with two different software (see Experimental Section) and the obtained data was indicated as similarity coefficients according to Edwards [19], Nei [20,21] and Rogers [22]. Genetic similarity coefficients obtained with different programs and methods showed the same genetic distances, resulting in the same phylogenetic tree (Figure 1). Here we present only the genetic distance coefficients according to Nei (Table 3).

The genetic distance coefficient within the studied tobacco varieties ranged from the very low value of 0.03 for the combination of Jk-48 and Jb-125/3 to a high value of 1.08 for the combination of Virginia MB1 and Djebel 1, indicating that a wide range of genetic diversity was present in the selected varieties (Table 3). A cluster analysis for the selected 10 tobacco varieties showed that there are three distinct clusters (Figure 1). The first cluster is represented by broad-leaf varieties—Virginia MB1 and Burley-2/93. The second cluster is represented by Yaka and Djebel varieties (Djebel 1, Yaka Jb-125/3, Yaka Jk-48, NS-72), while the third cluster is represented by Prilep varieties (Prilep P12-2/1, Prilep
P-23, Prilep P66-9/7, Prilep P-80pt). These three clusters of tobacco are clearly genetically distinct.

**Figure 1.** Cluster analysis of the 10 tobacco varieties based on the data of 24 microsatellite markers, using the genetic distance of Nei [20,21].

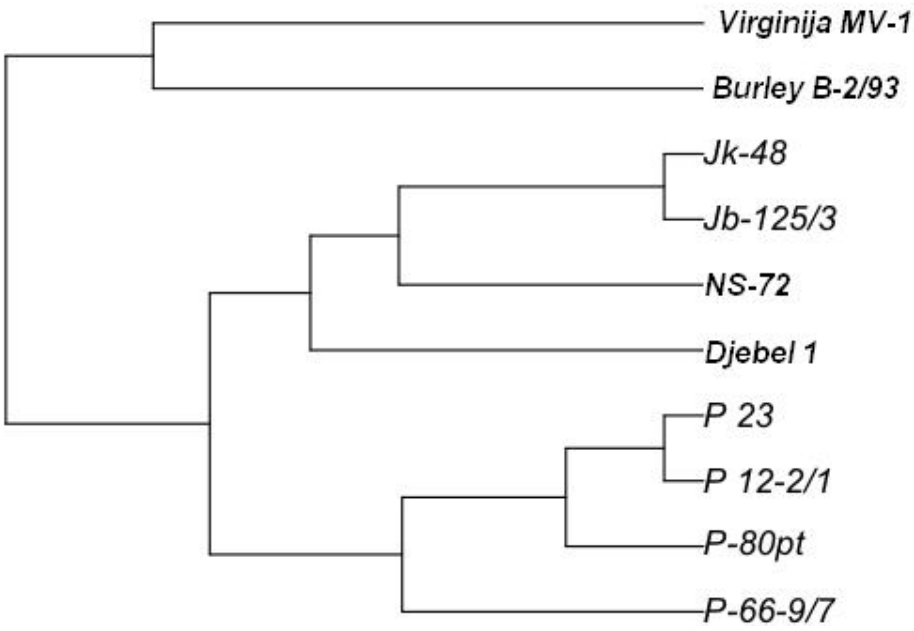

**Table 3.** Comparison of the similarity of the 10 studied tobacco varieties based on the data for 24 polymorphic microsatellite markers, indicated as Nei genetic distance [20,21].

| | **B-2/93** | **Djebel** | **NS- 72** | **Jb 125/3** | **Jk 48** | **P 12-2/1** | **P 23** | **P 66-9/7** | **P 80pt** | **V MB1** |
|---|---|---|---|---|---|---|---|---|---|---|
| **B-2/93** | 0.00 | | | | | | | | | |
| **Djebel** | 0.71 | 0.00 | | | | | | | | |
| **HC 72** | 0.73 | 0.49 | 0.00 | | | | | | | |
| **Jb 125/3** | 0.72 | 0.37 | 0.28 | 0.00 | | | | | | |
| **Jk 48** | 0.72 | 0.32 | 0.26 | 0.03 | 0.00 | | | | | |
| **P 12-2/1** | 0.75 | 0.63 | 0.34 | 0.53 | 0.53 | 0.00 | | | | |
| **P 23** | 0.67 | 0.63 | 0.29 | 0.48 | 0.48 | 0.03 | 0.00 | | | |
| **P 66-9/7** | 0.83 | 0.57 | 0.39 | 0.59 | 0.59 | 0.31 | 0.31 | 0.00 | | |
| **P 80pt** | 0.75 | 0.69 | 0.49 | 0.72 | 0.72 | 0.11 | 0.14 | 0.27 | 0.00 | |
| **V MB1** | 0.54 | 1.08 | 0.84 | 0.98 | 0.98 | 0.94 | 0.86 | 1.03 | 0.94 | 0.00 |

The lowest similarity exists between the broad-leaf varieties (Virginia MV-1 and Burley B-2/93) and the oriental tobacco varieties, where similarity coefficient ranged from 0.71 to 1.08. Within the oriental cluster, there are two different subclusters: the subcluster of Prilep varieties and the subcluster of Djebel and Yaka varieties. Djebel and Yaka varieties have the higher similarity to each other (coefficient of 0.32–0.49) compared with the lowest similarity between Djebel and Prilep varieties (genetic distance coefficient from 0.57 to 0.69). This can be explained by the fact that present cultivated Djebel and Yaka varieties in the Republic of Macedonia originated from the same variety (Xanthian Yaka), but from different geographical origin and therefore are more similar to each other than with Prilep varieties, which descended from variety Djumaj-bale [5].

The position of Prilep P12-2/1 in our philogenetic tree regarding to other oriental (Djebel and Yaka) and broad—leaf varieties (Burley B-2/93) is comparable to its position in the philogenetic tree of Bindler *et al.* [14]. However, further studies including more oriental tobacco varieties from the broader region, as well as varieties from other parts in the world, will help in providing a clearer picture on the genetic relatedness of oriental varieties and their position in the genetic tree of cultivated tobacco in the world.

## 3. Experimental Section

### *3.1. Plant Material and DNA Isolation*

Green leaves and seeds from 10 tobacco varieties (Djebel 1, Prilep P12-2/1, Prilep P-23, Prilep P66-9/7, Prilep P-80pt, Yaka Jb-125/3, Yaka Jk-48, NS-72, Virginia MV-1 and Burley B-2/93) were provided by the Tobacco Institute, Prilep, R Macedonia. For each tobacco variety, 2 DNA samples were isolated: one from seeds and one from leaves. Plant material was homogenized in analytical mill Cole-Parmer 4301-02 (Cole-Parmer Instrument Co., USA). DNA was isolated from 100

mg homogenized material with DNeasy Plant Mini Kit (QIAGEN GmbH, Hamburg, Germany) according to the procedure. DNA was eluted in 100 μL of elution buffer (10 mM Tris-HCl, 0.5 mM EDTA pH 9.0) and the quality was assessed by 1% TBE agarose gel electrophoresis. The concentration of isolated DNA was determined spectrophotometrically and the stock DNA was diluted to make a working solution of 50ng/μL for PCR analysis.

### *3.2. Molecular Markers*

A total of 30 SSR markers, out of 282 reported by Bindler *et al.* [14], were used in this study. The markers were chosen on the basis of their chromosomal location and type of repeat (di-, tri- or more repeat units). The name of the markers, type of repeat and approximate size of the amplified fragment [14] is given in Table 4. Forward primers were labeled with different ABI dyes (FAM, HEX or TET) for fragment analysis on an Applied Biosystems 310 Genetic Analyzer.

**Table 4.** Microsatellite markers used in the study.

| Name | Repeat | Expected size (bp) |
|---|---|---|
| PT 1193 | AC/AT | 328 |
| PT 20021 | AAC | 345 |
| PT 20165 | CT/CTT | 208 |
| PT 20172 | CTT | 203 |
| PT 20176 | CTT | 258 |
| PT 20242 | AGG | 200 |
| PT 20388 | AAG | 185 |
| PT 20445 | AAG | 193 |
| PT 30005 | TAA | 230 |
| PT 30021 | TA | 224 |
| PT 30077 | CA | 210 |
| PT 30087 | CAA | 177 |
| PT 30096 | GAA | 234 |
| PT 30138 | GAA | 222 |
| PT 30144 | TA | 266 |
| PT 30150 | TA | 229 |
| PT 30160 | TAAAAA | 186 |
| PT 30164 | CAG | 153 |
| PT 30188 | GA | 159 |
| PT 30231 | CAT | 184 |
| PT 30255 | GAAA | 228 |
| PT 30274 | GGA | 213 |
| PT 30378 | CGA | 222 |

**Table 4.** *Cont.*

| | | |
|---|---|---|
| PT 30392 | TAA | 270 |
| PT 30417 | CAA | 193 |
| PT 30463 | TAAA | 169 |
| PT 30480 | CAG | 163 |
| PT 40015 | GA | 170 |
| PT 40021 | CAG | 153 |
| PT 40035 | CAT | 199 |

### *3.3. PCR Conditions and Allele Detection*

PCR was performed in 25 μL final volume containing 1xPCR buffer (10 mM Tris-HCl pH 8.3, 50 mM KCl, 1.5 mM $MgCl_2$, 0.001% (w/v) gelatin), 0.2 mM dNTPs, 10 pmol of each forward and reverse primers, 1U of AmpliTaq Gold (Applied Biosystems) and 50 ng of template DNA. For each variety, two PCR assays were carried out with DNA isolated from seeds and leaves, respectively. The PCR program consisted of initial denaturation at 94 $^{o}$C for 10 min, followed by 35 amplification cycles of denaturation at 94 $^{o}$C for 10 s, primer annealing at 55 $^{o}$C for 45 s and DNA extension at 72 $^{o}$C for 1 min, followed by final extension step at 72 $^{o}$C for 20 min. Incubation was carried out in a model 2720 thermal cycler (Applied Biosystems). Samples for capillary electrophoresis contained 0.5 μL PCR product, 0.5 μL internal size standard and 20 μL formamide. Alleles were detected using the GeneScan software package of Applied Biosystems.

### *3.4. Diversity Analysis*

The polymorphic information content (PIC) value [23] was calculated for each microsatellite locus as follows:

$$PIC_i = 1 - \sum_{j=1}^{n} P_{ij}^{2} \quad \text{in}$$

Where $PIC_i$ is the polymorphic information content of a marker i; $P_{ij}$ is the frequency of the $j^{th}$ allele for $i^{th}$ marker, and the summation extends over n alleles.

The genetic clustering of the 10 tobacco varieties was done using the "R" software [24]. The genetic distances were calculated with the "ade4" package [25,26], according to Edwards [19], Nei [20,21] and Rogers [22], while the philogenetic tree was constructed with the "ape" package [27]. In order to confirm these results, the genetic clustering was done also with GenoDrive software [28], and the genetic distances were calculated using the methods of Nei [20,21] and Rogers [22].

## 4. Conclusions

Microsatellite marker analysis provides high resolution and can be used for fast and accurate variety identification. Assessment of genetic diversity among varieties can be used in evaluation of purity and

stability of genotypes entering into a breeding and seed multiplication program. The 24 microsatellite markers used in this study allow identification of all 10 cultivated tobacco varieties in the Republic of Macedonia. The obtained data could also be used for reducing duplicate collection of germplasm, and therefore can be of great use in tobacco breeding programs in the Republic of Macedonia.

This study also provided additional information about *N. tabacum* population genetics and expanded the knowledge about the genetic variation within diverse gene pools of cultivated tobacco.

## Acknowledgements

This work was supported in part by the FP7 Grant 229458 from the European Commission and by the Funds for science of the Macedonian Academy of Sciences and Arts (No.09-40/12), Skopje, Republic of Macedonia.

## References and Notes

1. Kenton, A.; Parokonny, A.S.; Gleba, Y.Y.; Bennett, M.D. Characterization of the *Nicotiana tabacum* L. genome by molecular cytogenetics. *Mol. Gen. Genet.* **1993**, *240*, 159-169.
2. Lim, K.Y.; Matyasek, R.; Kovarik, A.; Leitch, A.R. Genome evolution in allotetraploid *Nicotiana*. *Biol. J. Linn. Soc.* **2004**, *82*, 599-606.
3. Gadani, F.; Ayers, D.; Hempfling, W. Tobacco: a tool for plant genetic engineering research and molecular farming. Part I. *Agro. Food Ind. Hi. Tech.* **1995**, *6*, 19-24.
4. Brandle, J.; Bai, D. Biotechnology: uses and applications in tobacco improvement. In *Tobacco: Production, Chemistry and Technology*; Davis, N., Ed.; Wiley-Blackwell: Oxford, UK, 1999; pp. 49-65.
5. Korubin-Aleksoska, A. *Tobacco Varieties from Tobacco Institute-Prilep*; Trajkovski, F., Ed.; Alfa print: Skopje, Macedonia, 2004.
6. Martz, F.; Maury, S.; Pincon, G.; Legrand, M. cDNA cloning, substrate specificity and expression study of tobacco caffeoyl-Coa 3-O-methyltransferase, a lignin biosynthetic enzyme. *Annu. Rev. Plant. Physiol.* **1998**, *36*, 427-437.
7. Bai D.; Reeleder, R.; Brandle, J.E. Identification of two RAPD markers tightly linked with the *Nicotiana debneyi* gene for resistance to black root rot of tobacco. *Theor. Appl. Genet.* **1995**, *91*, 1184-1189.
8. Bai, D. Three novel *Nicotiana debneyi* specific repetitive DNA elements derived from a RAPD marker. *Genome* **1999**, *42*, 104-109.
9. Del Piano L.; Abet, M.; Sorrentino, C.; Acanfora, F.; Cozzolino, E.; Di Muro, A. Genetic variability in *Nicotiana tabacum* and *Nicotiana* species as revealed by RAPD markers: Development of the RAPD procedure. *Contrib. Tob. Res.* **2000**, *19*, 1-15.
10. Rossi, L.; Bindler, G.; Pijnenburg, H.; Isaac, P.G.; Giraud-Henry, I.; Mahe, M.; Orvain, C.; Gadani F. Potential of molecular marker analysis for variety identification in processed tobacco. *Plant Var. Seeds* **2001**, *14*, 89-101.
11. Tezuka, T.; Onosato, K.; Hijishita, S.; Marubashi W. Development of Q-chromosome-specific DNA markers in tobacco and their use for identification of a tobacco monosomic line. *Plant Cell Physiol.* **2004**, *45*, 1863-1869.

12. Sarla, K.; Rao, R.V.S. Genetic diversity in Indian FCV and burley tobacco cultivars. *J. Genet.* **2008**, *87*, 159-163.
13. Ren, N.; Timko, M.P. AFLP analysis of genetic polymorphism and evolutionary relationships among cultivated and wild *Nicotiana* species. *Genome* **2001**, *44*, 559-571.
14. Bindler, G.; van der Hoeven, R.; Gunduz, I.; Plieske, J.; Ganal, M.; Rossi, L.; Gadani, F.; Donini, P. A microsatellite marker based linkage map of tobacco. *Theor. Appl. Genet.* **2007**, *114*, 341-349.
15. Zhang, H.Y.; Liu, X.Z.; He, C.S.; Yang, Y.M. Genetic Diversity among flue-cured tobacco cultivars on the basis of RAPD and AFLP markers. *Braz. Arch. Biol. Technol.* **2008**, *51*, 1097-1101.
16. Siva-Raju, K.; Madhav, M.S.; Sharma, R.K.; Murthy, T.G.K.; Mohapatra, T. Genetic polymorphism of Indian tobacco types as revealed by amplified fragment length polymorphism. *Curr. Sci.* **2008**, *94*, 633-639.
17. Moon, H.S.; Nicholson, J.S.; Heineman, A.; Lion, K.; van der Hoeven, R.; Hayes, A.J; Lewis, R.S. Changes in genetic diversity of U.S. flue-cured tobacco germplasm over seven decades of cultivar development. *Crop Sci.* **2009**, *49*, 498-508.
18. Moon, H.S.; Nifong, J.M.; Nicholson, J.S.; Heineman, A.; Lion, K.; van der Hoeven, R.; Hayes, A.J.; Lewis, R.S. Microsatellite-based analysis of tobacco (*Nicotiana tabacum* L.) genetic resources. *Crop Sci.* **2009**, *49*, 2149-2159.
19. Edwards, A.W. Distances between populations on the basis of gene frequencies. *Biometrics* **1971**, *27*, 873-81.
20. Nei, M. Genetic distances between populations. *Amer. Naturalist* **1972**, *106*, 283-292.
21. Nei M. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* **1978**, *23*, 341-369.
22. Rogers, J.S. Measures of genetic similarity and genetic distances. *Stud. Gen.* **1972**, *7213*, 145-153.
23. Anderson, J.A.; Churchill, G.A.; Autrique, J.E.; Tanksley, S.D.; Sorrellis, M.E. Optimizing parental selection for genetic linkage maps. *Genome* **1993**, *36*, 181-186.
24. The R Development Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2009.
25. Chessel, D.; Dufour, A.B.; Thioulouse, J. The ade4 package-I- One-table methods. *R. News* **2004**, *4*, 5-10.
26. Dray, S.; Dufour, A.B.; Chessel, D. The ade4 package-II: Two-table and K-table methods. *R. News* **2007**, *7*, 47-52.
27. Paradis, E.; Claude, J.; Strimmer, K. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* **2004**, *20*, 289-290.
28. Meirmans, P.G.; Van Tienderen, P.H. GENOTYPE and GENODIVE: two programs for the analysis of genetic diversity of asexual organisms. *Mol. Ecol. Notes* **2004**, *4*, 792-794.