Agricultural University of Plovdiv
Faculty of Agronomy
Department of Plant Physiology, Biochemistry and Genetics

Bojin Maximov Bojinov

Use of molecular markers in genetic diversity studies and for DNA profiling

ABSTRACT
of dissertation work
for awarding a scientific degree
"doctor of science"

Higher Education Area ***4.0. Natural sciences, mathematics and computer science;***
Professional field: ***4.3. Biological Sciences*;**
Scientific specialty: "***Genetics"***

P L O V D I V

2 0 2 5

The dissertation contains 254 pages in A4 format, including 17 photos, 33 tables and 63 figures. The bibliography includes 280 titles, of which 279 are in Latin and 1 in Cyrillic.

The dissertation was discussed and proposed for defense before a Scientific Jury at a meeting of the Department of Plant Physiology, Biochemistry and Genetics at the Faculty of Agronomy of the Agricultural University - Plovdiv, held on …… (minutes No. …… of ……. 2025).

The official defense will be held on ……2025 from ……. h. at …….. at the “Faculty of Agronomy".

Scientific jury:

1. Acad. Atanas Atanasov – reviewer
2. Prof. Diana Svetleva – reviewer
3. Associate Professor Samir Naimov – reviewer
4. Corresponding Member Ivan Atanasov – participation with an opinion
5. Prof. Anelia Yancheva, PhD – participation with a statement
6. Prof. Dr. Andon Vassilev – participation with a statement
7. Assoc. Prof. Dr. Lyubka Koleva-Vulkova – participation with a statement

Acknowledgements

The author of this work would like to express his gratitude for the support provided by the International Atomic Energy Agency (Agreement No. 12814/2004) and the Bulgarian Science Fund (Agreements VU-AN-205/2008 and DFNI B01/16/2012). The views expressed in this document are solely the responsibility of the author and do not necessarily reflect the views of the IAEA and the National Science Fund on the issues discussed.

Furthermore, the author would like to express his gratitude to Jean-Marc Lacape, Marc Giband, Maxim Bozhinov, Zhivko Danailov, Daniela Ganeva, Atanas Pavlov and Mariyana Nakova for their valuable input with ideas and discussions. The author would also like to express his gratitude for the contributions of Elena Todorovska, Bilyana Ivanova, Katya Dimitrova, Silvia Vasileva, Milena Kostova, Atanas Vuchkov, Todorka Srebcheva, as well as Mihail Angelov, Tsvetan Venovski, Dilyana Ivanova, Sasho Georgiev, Ivan Ivanov, Zapryana Gancheva, Nevena Petrova, Stefani Dimitrova and Maria Tashuteva.

The author also wishes to express his gratitude for the timely HVI service provided by Cotton Research Station in Nazilli, Turkey.

Contents

Introduction

Since the 1990s, a number of types of DNA-based molecular markers have been available, the potential for practical application of which has been repeatedly discussed by both developers and a whole galaxy of researchers who have characterized their applicability in various organisms - wild and cultivated plants and animals, as well as various microorganisms. Their ability to facilitate both conventional selection and efforts aimed at accelerated development of new forms of organisms that better respond to rapidly changing environmental conditions and market requirements has been identified and well documented. As a result, a number of marker systems have been used for years to accelerate the selection process, guarantee the genetic identity of breeding materials and protect their intellectual property in large breeding companies. Due to the high initial cost of application, the use of these marker systems by breeding teams relying on public funding in our country was limited until recently. With the development of DNA sequencing and amplification technologies, the cost of their use has decreased exponentially and these tools are now easily accessible both financially and technologically. However, their use in Bulgarian state institutes and universities engaged in breeding tasks remains very limited. The same applies to the institutions responsible for the registration and control of seed and propagation material (IACS) and animal reproductive material (IASRJ). The root of the problem lies in the lack of standardization of the systems and protocols used between the two groups of institutions, due to the different requirements for the functioning of each of them.

The author's many years of experience in using a number of marker systems in a wide range of organisms allows him to formulate a hypothesis for solving the problem thus identified, the defense of which is the subject of the present dissertation. He summarizes various applications of several dominant (i.e. AFLP, ISSR, CAPS) and co-dominant (RFLP, SSR) marker systems. This includes the use of both types of markers in analyzing the genetic diversity of different sources of germplasm, along with the creation and proposal of approaches for DNA genotyping of several crops, microorganisms and two animal breeds. The arguments for choosing one or the other tool are presented in the context of the relevant studies, as carried out by the author.

This work summarizes the author's diverse efforts to make molecular tools understandable and usable for the Bulgarian breeding community in its main profiles - plant breeding and animal breeding.

ISSR markers as a toolkit, equally convenient for research and preservation of genetic diversity, in selection and breeding activities and in state variety

testing (in the case of RHS testing of candidate varieties and breeds), is argued.

Aim and objectives

The aim of this work is to establish a molecular marker system and to propose it for reliable and reproducible identification of genetic diversity, both in plant, microbial and animal forms. At the same time, the identified marker system should be easily scalable, so as to allow use both for selection purposes and for control purposes in variety testing and seed production/breeding of local breeds.

To achieve the set aim, the following tasks have been formulated for implementation:

1. To conduct a comparative test of a set of marker systems for identifying genetic diversity in plant species (cotton, tomatoes, tobacco, barley, etc.), eukaryotic phytopathogens (*Phytophthora* and *Fusarium* fungi) and animals (goats).
2. To establish the possibilities for applying the studied marker systems in order to discover loci for quantitative traits of specific importance for local populations.
3. To characterize the possibilities for applying the selected markers in variety testing, variety maintenance and conservation of the existing biodiversity of local plant forms and animal breeds.

Additionally, 4 subtasks have been formulated:

- To establish the possibilities for transferring markers between different genetic maps (previously published and developed within the framework of this study) in some local plant and animal varieties;
- To determine molecular markers associated with qualitative characteristics in characterizing decaying intraspecific populations;
- To determine loci for quantitative traits in different genotypes with practical applicability for Bulgarian breeding;
- To specify the ability of the proposed marker system to identify genetic diversity at different levels - intravarietal/ intrabreed, intervarietal/ interstrain /intraspecific, interspecific.

Materials and methods

Plant material

Cotton

The mapping population for the AFLP and SSR marker mapping was composed of 75 BC1 plants and originated from the cross ‘Guazuncho 2’ × ‘VH8-4602’. ‘Guazuncho 2’ is a modern inbred *G. hirsutum* variety developed at

INTA, Argentina. It was selected for its overall good production performance, evaluated in different tropical countries (B. Hau, personal communication). 'VH8-4602' originates from Antigua and is derived from a cross of two forms *of G. barbadense,* MSI and V135. MSI refers to *G. barbadense varieties* cultivated in Barbados in the 1980s and 1990s, while V135 is a selection from the island of Saint Vincent of more ancient origin (Pauly, 1991). The initial cross used five different female Guazuncho 2 × male VH8 plants; then the F_{1s} were backcrossed with four different Guazuncho 2 plants used as mothers.
The Bulgarian variety Chirpan 603 was selected to study the effects of mutagenic treatments on fiber quality characteristics, as it is a standard variety covering more than 80% of the growing area from 2005 to 2016. The approach involved inducing mutations by gamma irradiation of the standard genotype, followed by selfing the progeny for several generations before quantitative trait loci (QTL) analysis. The induction of mutations was aimed at improving fiber quality and technological performance. The mutant segregating population that resulted from this treatment was used for the purpose of constructing a genetic map of quantitative trait loci with a significant effect under local conditions.

Tobacco

The study included four commercial varieties of the two most common types of tobacco in Bulgaria (Oriental and Burley). To assess the existing intravarietal diversity, eight plants of each were used - three Oriental varieties (Plovdiv 7, Nevrokop 1146 and Krumovgrad 90) and one type of Burley (Burley 1317), which are widely distributed in our country.

Tomatoes

Twenty-four tomato genotypes, provided by the Institute of Plant Physiology and Genetics (IPPG, former Institute of Genetics "Prof. D. Kostov") and representing three main groups, were used to study the applicability of the dominant ISSR marker system:
1. Wild species (8 genotypes) - *Lycopersicum hirsutum* IVT 66087, *Lycopersicum glandolosum* IVT 48090, *Lycopersicum minutum* 911, *Lycopersicum pimpinellifolium* LA 0722, *Lycopersicum esculentum* var. *cerasiforme* LA 1226, *Lycopersicum parviflorum* LA 2072, *Lycopersicum f. glabratum* LA 2864 and *Lycopersicum peruvianum* LA 2172;
2. Determinant (sp) varieties and breeding lines (10 genotypes) - 533/06 (Trapezitsa), 540/06 (Bela), 541/06, 560/06, 561/06, 564/06, 573/06, 618/06, 632/06 and 636/06

3. Indeterminate (sp+) selection lines (6 genotypes) - 707/06, 708/06, 710/06, 792/06, 798/06 and 803/06.
All determinant and indeterminate varieties and breeding lines belong to the species *Lycopersicum esculentum* Mill. (*Solanum lycopersicum)*.
Wild species are maintained by the IFRG / IZK "Maritsa" and are used in breeding programs as donors of economically important characteristics such as resistance to biotic and abiotic stress. In addition, some of these species have proven to be valuable sources of qualities related to improving fruit quality - dry matter content and antioxidant components (ascorbic acid, lycopene, beta-carotene, etc.).
To study the applicability of dominant and co-dominant markers, plant material from 8 Bulgarian samples was used: the Plovdivska Karotina variety (with a genetic background of *Solanum chillense)*, the IZK Alya variety (cherry type – with a genetic background of *Solanum pimpinellifolium)* and 6 tomato breeding lines (L21β, L53β, L1140, L1116, L975, L984) from the Institute of Vegetable Crops (IZK "Maritsa"), Plovdiv, Bulgaria. Each genotype was represented by 7 individual plants.
For the purposes of the study with SCAR markers, 20 tomato lines from the breeding program of the Maritsa Plant Breeding Institute were used, which were tested for resistance to ToMV. pathotypes 0, 1 and 2. Lines known to carry the relevant resistance genes were used as controls, and the variety "Ideal", known to have no resistance genes, was used as a negative control.
The plants were grown in terrines, on a peat-perlite mixture (2.5:1) in a phytostat chamber at a temperature of 22-25°C and a photoperiod of 14/10 h.

Paulownia

Several genotypes available at the Agricultural University - Plovdiv at the time of the experiment were tested. They were selected as:
- *P. tomentosa* is described as frost-resistant (down to -27°C);
- *P. fortunei* – as cold-resistant (up to -0°C), but extremely fast-growing and with high ecological plasticity; and
- *P. elongata* – frost resistant (up to -16°C).
The following lines/hybrids were analyzed: *P. elongata, P. tomentosa x P. fortunei, P. elongata x P. elongata.*

Barley

Twenty-four breeding lines and cultivars of *Hordeum vulgare* ssp. *distichum* L. were included in the study according to preliminary information about their genetic distance. Genotypes 82105326, 89105100, 96105023, 96105024, 96105027, 96105046, 96105050, 99105020, 99105030 and I- Da

/102 were provided by the National Gene Bank (IPGR) – Sadovo. Promising breeding lines 508, 511 and 622 are the result of the breeding work of the Department of Genetics and Breeding of the Agricultural University – Plovdiv. The varieties Alexis, Beta Ketsoras, Emon, Kamenitsa, Karavela, Kaskadyor, Obzor and the newly created varieties Neda and Gorast were also included as reference material. Orpheus and Kifi were provided by the Institute of Agriculture, Karnobat.

Pepper

For the purpose of the study, local populations and cultivars belonging to the species *Capsicum annuum, Capsicum frutescens* and *Capsicum chinense were used.*

The specimens are maintained on private farms and are used in breeding programs, as donors of genes for antioxidant compounds (capsaicin, capsanthin) and other fruit characteristics, for seed production and as ornamental plants. Some of these lines also serve as valuable sources of traits of agronomic importance.

Wheat

The object of the study were 11 wheat genotypes, planted in a collection nursery in the Selection Garden of the Department of Genetics and Breeding at the Agricultural University of Plovdiv. Some of the samples were collected from all over Bulgaria, and others were selected in the former Department of Genetics and Breeding. The genotypes used are Avenue, Sofru, Teres, Enola, Solehio, Helix, Predel, Saya, Reyadur *and* Pibrak.

Corn

The object of study was 15 maize genotypes, representing the main hybrids grown in the country. Plants of each hybrid were grown in the greenhouse of the Department of Plant Breeding and Breeding at the Agricultural University-Plovdiv to obtain the necessary leaf mass.

The genotypes used are DKC 4709, LG 31.325 Limagold RS, KWS Giro, KWS Hypolito, SOUFFLET ISULEA, SOUFFLET PIANELLO, DKC 4611, DKC 5092, DKC 5075, P 9241, P 9889, P 9415, P9610, P 8834, P 9363.

Eukaryotic plant pathogens

Phytophthora fungi

The study was performed with 22 isolates of *Phytophthora* sp., taken from the root collar area of infected cherry, apple and pear trees, as well as two isolates from strawberries and redwood.

Fusarium

The study was performed with 8 isolates of *Fusarium* sp., collected from the Plovdiv region. The species identification of the isolates was performed in the Department of Phytopathology of the University of Plovdiv.

Animal organisms

Goats

Two local goat breeds with a distinctive exterior, stably transmitted to the offspring – Kalofer long-haired goat and Bulgarian long-haired goat (Sedefchev et al., 2011; Vuchkov and Dimov, 2011; Vuchkov et al., 2011) were selected for study. The subpopulation of the Kalofer goat in the separate region of Southwestern Bulgaria was also studied. In Blagoevgrad region, there were 5 herds of Bulgarian long-haired goats located in the following villages – Kresna; Karnalovo village of Drenovo; Petrich; Harsovo municipality, Sandanski; Ploshki village, Sandanski.

A total of 45 female and 15 male specimens (mature) of the Bulgarian long-horned goat were studied.

In the same way, 45 females and 15 males were selected from 5 herds of Kalofer long-haired goats in the regions of Bansko, Simitli, Vlahi, Kresna and Dolna Gradeshnitsa.

DNA isolation

Cotton

Cotton DNA was extracted from young fully developed leaves. For AFLP analysis, three grams of sample, including BC1 population, parents, F_1 and diploid species, were frozen and ground. DNA was then isolated by the methyl-alkyltrimethylammonium bromide (MATAB) method as described in (Risterucci et al., 2000) and further purified by ultracentrifugation to sedimentation equilibrium in a cesium chloride–ethidium bromide gradient.

For ISSR and SSR analyses, 150 mg of each sample, including parents, original variety (in case of mutation population), F_1/M_1, F_2/M_2 and subsequent generations were frozen in liquid nitrogen and ground in a mortar. DNA was then isolated using the PhytoPure DNA extraction kit (Amersham-Pharmacia). The manufacturer's protocol was followed with minor modifications.

Tomatoes

Tomato DNA samples were extracted from frozen leaf tissue (250–300 mg) of field-grown plants using a standard CTAB procedure according to (Murray and Thompson, 1980) with a modified extraction buffer containing 2% CTAB, 4% PvP, NaCl, Tris-HCl (pH 8.0) and EDTA-Na2 (pH 8.0). The addition of 8M

LiCl to the extraction buffer was a useful step to remove the large amount of RNA in the initial extraction step. After several (3–4) extractions with chloroform -isoamyl alcohol, the DNA was precipitated with cold isopropanol. The DNA pellet was washed with 76% EtOH containing 10 mM ammonium acetate for 30–45 min at 4 ° C to eliminate traces of polysaccharides. The dried DNA was dissolved in 1 x TE buffer and the remaining RNA content was removed by treatment with RNaseA for 45 minutes at 370°C. After extraction with chloroform -isoamyl alcohol, the DNA was precipitated with 5M NaCl and 2 volumes of cold 96% ethanol to a final concentration of 0.2M. The DNA was further washed with cold 70% ethanol, dried and dissolved in a small volume of 1 x TE buffer. The isolated DNA was of high quality (A260/280 = 1.7÷2.0), which allowed for trouble-free amplification of microsatellite loci.

Tobacco

Tobacco DNA was extracted from the last fully developed leaf of greenhouse-grown plants. Approximately 200 mg of leaf samples were frozen in liquid nitrogen before being ground in a mortar. DNA from each plant was extracted using a standard DNA extraction kit (Nucleon PhytoPure, Amersham Biosciences UK Ltd., Buckinghamshire, HP7 9NA England) with minor modifications to the manufacturer's protocol.

Paulownia

DNA was extracted from the last fully developed leaf of plants grown in a greenhouse. Approximately 200 mg of leaf samples were frozen in liquid nitrogen before being ground in a mortar. DNA from each plant was extracted using a standard DNA extraction kit (Nucleon PhytoPure, Amersham Biosciences UK Ltd., Buckinghamshire, HP7 9NA England) with minor modifications to the manufacturer's protocol.

Barley

DNA from barley specimens was isolated from 40 mg of frozen leaf tissue. The samples were ground in a mortar in liquid nitrogen to a fine powder. The “EZNA® Plant DNA Mini Kit“ (Omega Biotek, Inc., Georgia, USA) was used for DNA purification as recommended by the manufacturer, using the rapid extraction protocol.

Pepper

Only plants that fully met the UPOV phenotypic descriptions for the respective form or variety were used for DNA extraction. The homogeneity of these genotypes has not been tested using molecular markers to date.

DNA was isolated from the last young, fully developed leaf of the selected and marked plants for the study, grown outdoors, in containers, under identical conditions.
Three hundred milligrams of leaf material from each sample was ground after freezing in liquid nitrogen to a fine light green powder. The method chosen for DNA isolation was the EZNA Plant DNA Kit from Omega Bio-Tek.

Wheat

DNA was isolated from 60 mg of frozen leaf tissue. The samples were ground in a mortar in liquid nitrogen to a fine powder. The "EZNA® Plant DNA Mini Kit" (Omega Biotek, Inc., Georgia, USA) was used for DNA purification as recommended by the manufacturer, using the rapid extraction protocol.

Corn

DNA was isolated from 60 mg of frozen leaf tissue. The samples were ground in a mortar in liquid nitrogen to a fine powder. The "EZNA® Plant DNA Mini Kit" (Omega Biotek, Inc., Georgia, USA) was used for DNA purification as recommended by the manufacturer, using the rapid extraction protocol.

Goats

The choice of starting material for genomic DNA extraction from goats was made to implement non-invasive sample collection and avoid any stress to the animals.
Genomic DNA isolation from hair follicles of Kalofer Longhair and Bulgarian Vitoroga Longhair was performed with the innuPREP DNA Kit (IST Innuscreen GmbH, Germany). The manufacturer's protocol was followed and the steps were optimized for the conditions in our laboratory.
Visualization of the isolated genomic DNA from all samples was performed after electrophoretic separation of the products in a 1% agarose gel. Five µl of the final solution of extracted DNA was applied to each slot and staining was performed with ethidium bromide.

Marker analysis

RFLP analysis

A set of 210 RFLP probes developed by A. Paterson (Plant Genome Mapping Laboratory, University of Georgia, Athens, Georgia) was tested. These probes were obtained from various Gossypium spp. genomic and cDNA libraries described in (Reinisch et al., 1994). The probes were obtained by PCR amplification using consensus M13 reverse and forward primers. In addition, 4 probes corresponding to candidate fiber elongation genes that were found to be polymorphic were mapped. They were obtained by PCR amplification

of genomic DNA of cotton (*G. hirsutum* 'Coker 310') using primers based on published sequences. Finally, of the 80 *Arabidopsis thaliana* ESTs (prefixed At) probes developed within the European EuDicotMap project, 14 showed clear hybridization with cotton, and 11 were mapped. The relationships of the mapped probe numbers ' At ' to the reference probe nomenclature are given in Table 1.

Table 1. Correspondence of locus names with clone references and EMBL sequences of 11"At"loci obtained from partially sequenced cDNAs of Arabidopsis thaliana.

Locus	Reference	Reference sequence in EMBL	Chromosome
At03	34B8T7	T04245	c14
At09	82G7T7	T20487	c14
At16	109G22T7	T41886	c23
At17	113C1T7	T42378	c18
At18	114I16T7	T43072	c3
At24	129H11T7	T45078	c20
At30	146A20T7	T75767	c26
At47	193B3T7	H75998	D08
At51	221F10T7	N38650	c18
At55	247P2T7	AA597935	c16
At60	E3H3T7	AA042198	c9bot

DNA samples (7 µg) of both parents were digested with five different enzymes, EcoRI, EcoRV, BamHI, HindIII and XbaI, using 6U of enzyme per reaction. DNA fragments were separated on 0.8% w/v Tris -acetate-EDTA (TAE) agarose gels for 17 h at 20 V. Southern blots of the parents were prepared for screening for polymorphism and band segregation. After selecting the best probe-enzyme combinations from the parental forms, the samples were then hybridized to the corresponding BC1 plants, which also included integrated DNA from both parents, the F_1 and the three diploid species. Blotting, probe labeling, Southern hybridization and autoradiography were performed as described in (Lanaud et al., 1995), except that the prehybridization and hybridization steps were performed at 65°C for 4 h and overnight, respectively.

AFLP analysis

AFLP analysis was performed using the Life Technology AFLP™ Analysis System I (Gibco BRL, Gathersburg, MD) using a two-step amplification as described by (Vos et al., 1995). Genomic DNA samples were digested using 2.5 U EcoRI and MseI. Preliminary PCR amplification was performed using +1 primers (EcoRI+A, MseI+C). A subsequent selective amplification reaction using

+3-primers was performed with [γ-^{33}P]ATP end-labeled EcoRI primers. The protocol was performed as recommended by the supplier, with minor modifications, as detailed in (Risterucci et al., 2000). Eight EcoRI and eight MseI primers were used, making a total set of 64 EcoRI- MseI combinations (Table 2).

Table 2. Matrix of combinations of selective nucleotides used for the MseI and EcoRI primers.

	***Mse* I primer**	**M1** CAA	**M2** CAC	**M3** CAG	**M4** CAT	**M5** CTA	**M6** CTC	**M7** CTG	**M8** CTT
Selective nucleotides									
EcoRI primer									
E1	AAC	x	x	x	x	x	x	x	x
E2	AAG	x	x	x	x	x	x	x	x
E3	ACA	x	x	x	x	x	x	x	x
E4	ACT	x	x	x	x	x	x	x	x
E5	ACC	x	x	x	x	x	x	x	x
E6	ACG	x	x	x	x	x	x	x	x
E7	AGC	x	x	x	x	x	x	x	x
E8	AGG	x	x	x	x	x	x	x	x

SSR analysis

Cotton

The SSR markers used for integrated map construction were developed at Brookhaven National Laboratory (BNL prefix). All of these BNL microsatellites were obtained from *G. hirsutum* 'Acala Maxxa' genomic libraries enriched for GA and CT inserts according to (Ostrander et al., 1992). The cloned sequences used for primer design are available at https://www.cottongen.org/.

The conditions for SSR analysis were described in (Risterucci et al., 2000) using a [γ-^{33}P]ATP end-labeled "forward" primer and a 55°C melting temperature. The SSR primers were initially tested on both parents for their ability to produce polymorphic products. Of the 216 SSRs tested, 188 were both correctly amplified and polymorphic.

The selection of SSR primers for use in the detection of QTL for fiber quality in the Bulgarian genetic background was made after comparing several published genetic maps of cotton. At the time of the selection, several maps were already available (Guo et al., 2006; Jia et al., 2016; Mei et al., 2004; Nguyen et al., 2004; Park et al., 2005; Song et al., 2005; Zhang et al., 2002). Our selection was based on the available information from the map of (Zhang et al., 2005) who tested many SSR markers developed at Brookhaven National Laboratory (BNL) on an intraspecific cross that included the widely used

reference variety TM-1. From this map, six SSRs located on or closely linked to QTLs for fiber quality were selected. Additional SSR markers were selected from the map of (Song et al., 2005) who worked on the BC1 population from another intraspecific cross (TM-1 x Hai 7124). The selection was reduced to 15 SSRs after performing a comparative analysis of these two maps with a reference publication for QTL analysis (Lacape and Nguyen, 2005). SSR primer sequences were downloaded from the CMD database (https://www.cottongen.org/) and oligonucleotides synthesized by Microsynth Plc. (Switzerland).

Tomatoes

Published primers for tomato microsatellite markers (http://solgenomics.net;http://marker.kazusa.or.jp/Tomato), designated as locus-specific primers (LP), were synthesized with generic non-complementary nucleotide sequences tagF 5'ACGACGTTGTAAAA3' and tagR 5'CATTAAGTTCCCATTA3', respectively, at their 5' ends, as described in (Hayden et al., 2008). Aliquots of the primers (50 pM /µl) were prepared by mixing equimolar amounts of forward and reverse primers in miliQ H_2O and were designated as core primer sets for each locus. In addition, two generic tag primers were synthesized, namely tagF ' and tagR ' with the same sequences 5'ACGACGTTGTAAAA3' and 5'CATTAAGTTCCCATTA3'. The primer tagF ' (5'ACGACGTTGTAAAA3') was labeled at its 5' end with one of the following fluorescent dyes: FAM, ATTO565, ATTO550 and YELLOW YAKIMA (Applied Biosystems), allowing direct detection of alleles on an automated capillary sequencer (ABI3730, Applied Biosystems). All primers were synthesized by Microsynth AG (Switzerland).

All uniplex PCR reactions (PCR) were performed according to (Hayden et al., 2008) in a 6 µl reaction mixture containing 20–25 ng genomic DNA, 2x MyTaq HS mix (Bioline), 75 nM each of dye-labeled tagF ' and unlabeled tagR ' primers and an appropriate concentration (20, 30 or 60 nM) of unlabeled locus-specific primer (LPS), depending on the marker groups used.

PCR was performed on a Veriti96 Thermal Cycler (Applied Biosystems) using the following PCR conditions, depending on the melting temperature of the locus-specific primers. They were performed according to (Hayden et al., 2008) and (Tsonev et al., 2013) with some minor modifications.

Electrophoresis and visualization of tomato SSRs were performed on an ABI3730 DNA analyzer (Applied Biosystems). A standardized multiple pooling procedure was used to prepare SSR products for electrophoresis.

The pooling of PCR products with different dyes at 1:1:1:1 was intended to account for differences in the relative fluorescence of each fluorophore. In

cases where the intensity of the SSR bands specific for a given SSR was 2-3 times higher, a change in the ratio was performed so that the intensity of the bands labeled with different fluorophores was relatively equalized.

CAPS analysis

The sequences located in the Tm-2 locus (tm-2, Tm-2 and Tm-22) were published by (Lanfermeijer et al., 2003) in the GenBank database at NCBI under the numbers AF536199, AF536200 and AF536201, corresponding to alleles tm-2, Tm-2 and Tm-22, respectively.
The Tm-2 locus in tomato was genotyped using a single CAPS marker (Table 3) described by (Shi et al., 2011).

Table 3. Primer sequences used to perform CAPS analysis.

Marker	Primer	Sequence
CAPS	Tm2RS-f3	TGGAGGGGAATATTTGTGGA
	Tm2RS-r3	ACTTCAGACAACCCATTCGG

Expected size of the resulting fragment – 703 bp.

Restriction of the PCR product for the CAPS marker

5 µl of each PCR product was subjected to restriction with three enzymes – BoxI, KspAI and Alw21I (Fast Digest, Fermentas) according to the manufacturer's instructions in a volume of 15 µl reaction mixture. The enzymes used were isoschizomers of those indicated by (Shi et al., 2011) for PshAI, HpaI and BsiHKAI, respectively.

Visualization of the resulting DNA fragments on an agarose gel

10 µl of restriction products or 5 µl of PCR products were loaded onto a 1% agarose gel. Electrophoresis was performed for 60-75 min at 5 V/cm. The fragments were stained with DNA Stain G (Serva) and visualized with a transilluminator at a wavelength of 312 nm.

ISSR analyses

Cotton

A total of 14 primers were tested on the cotton samples studied. After eliminating primers that did not produce polymorphisms in the initial test, ISSR analysis was continued using a set of 10 primers, the sequences of which are presented in Table 4. The presence or absence of each ISSR fragment was scored in a binary system (1 = present, 0 = absent).

Table 4. Sequences of ISSR primers used to perform PCR reactions with cotton DNA.

Primer	DNA sequence	Length (bp)
ISSR 1	(CA) 8AA+GG	**20**
ISSR 2	(CA) 8AA+GC+T	**21**
ISSR 3	(GA) 8C+TC	**19**
ISSR 4	(AG) 8C+TC	**19**
ISSR 5	(AC) 8C+TA	**19**
ISSR 6	(AC) 8C+TG	**19**
ISSR 7	(AG) 8C+TG	**19**
ISSR 8	(AC) 8C+TT	**19**
ISSR 9	(AG) 8C	**17**
ISSR 10	(GA) 8T	**17**

Tobacco

A total of 4 primers were tested on tobacco specimens with their sequences presented in Table 5. The presence or absence of each ISSR fragment was scored as a binary unit symbol (1 = present, 0 = absent).

Table 5. Sequences of ISSR primers used to perform PCR reactions with tobacco DNA

Primer	DNA sequence	Length (bp)
ISSR 1	(CA) 8AA+GG	**20**
ISSR 2	(CA) 8AA+GC+T	**21**
ISSR 6	(AC) 8C+TG	**19**
ISSR 7	(AG) 8C+TG	**19**

Tomatoes

ISSR analysis was performed by testing several primers. The selection of the final set of 10 primers Table 6 is based on the screening of the initially selected primers against the available data on the frequency of tandem repeats in the tomato genome.

Table 6. Sequences of ISSR primers used to perform PCR reactions with tomato DNA.

Primer	DNA sequence	Length (bp)
ISSR 1	(CA) 8AA+GG	**20**
ISSR 2	(CA) 8AA+GC+T	**21**
ISSR 3	(GA) 8C+TC	**19**
ISSR 4	(AG) 8C+TC	**19**
ISSR 5	(AC) 8C+TA	**19**
ISSR 6	(AC) 8C+TG	**19**
ISSR 7	(AG) 8C+TG	**19**
ISSR 8	(AC) 8C+TT	**19**
ISSR 9	(AG) 8C	**17**
ISSR 10	(GA) 8T	**17**

Barley

To estimate the relative genetic distance between barley genotypes, a set of 9 ISSR primers was used (Vasileva, 2018), with their characteristics listed in Table 7.

Table 7. List of ISSR primers used in barley experiments.

Primer	DNA sequence	Length (bp)
ISSR_4	AGA GAG AGA GAG AGA GCT C	19
ISSR_6	ACA CAC ACA CAC ACA CCT G	19
ISSR_7	AGA GAG AGA GAG AGA GCT G	19
ISSR_8	ACA CAC ACA CAC ACA CCT T	19
ISSR_10	GAG AGA GAG AGA GAG AT	17
ISSR_5	RYG ACA GAC AGA CA	14
ISSR_6a	CTC TCT CTC TCT CTC TRG	18
ISSR_7a	CTC TCT CTC TCT CTC TRC	18
ISSR_11	GAG AGA GAG AGA GAG AA	17

PCR reactions were performed in a 25 µl volume with the following cycling regime: denaturation at 94°C for 3 min, 40 cycles of 94°C – 1 min, AT – 45 s, 72°C – 45 s, followed by a final extension at 72°C – 4 min, where AT is the melting temperature for each primer, calculated according to (Kochieva et al., 2002).

The analysis of the obtained DNA fragments was performed in 2% w/v agarose gels prepared according to Maniatis (Sambrook et al., 1987). Electrophoretic separation was performed for 1 hour at 150 V in an electrophoretic buffer containing: 10.78 g/l TRIS, 5.50 g/l boric acid and 0.74 g/l EDTA.$Na_2.2H_2O$.

Pepper

The primers used to perform the ISSR analysis (Table 8) were selected from a set of primers that showed high levels of reproducibility and potential for identification of polymorphisms in previous studies in the department.

PCR reactions were performed in the following reaction mixture: PCR buffer – 2.5 µl, dNTPs – 1.5 µl, ISSR primer – 1.5 µl, Taq polymerase – 0.12 µl, DNA - 1 µl, dH2O – 18.38 µl (total volume 25 µl). First, 1 µl of DNA was poured into tubes, then 24 µl of the final reaction mixture was taken and added to each sample.

The cycle used for the PCR reaction begins with denaturation at 94°C for 3 minutes, followed by 40 cycles of 94°C for 1 minute, Tx – 45 seconds, 72°C – 45 seconds, followed by a final extension at 72°C for 4 minutes (Tx – melting temperature calculated by Kochieva and al., 2002).

Table 8. Characteristics of the primers used to conduct ISSR analysis in pepper.

Primer	DNA sequence	Length (sun)	Temperature hybridization (°C)
E5	$(AC)_8$C+TA	19	55
E8	$(AC)_8$C+TT	19	55

Wheat

ISSR analysis was performed using primers (Gemmill and Grierson, 2021) that were available in the Molecular Biology Laboratory of the Department of Biological Sciences at the beginning of the study (Table 9).

Table 9. Sequences of ISSR primers used to conduct PCR analysis in wheat.

Primer	DNA sequence	Length (bp)
ISSR 7	AG(8)CTG	19
Volume 14	AG(8)YT	18
Volume 11	GA(8)YC	18
Volume 3	CA (7)RG	1 6

PCR reactions were performed in a volume of 25 µl.

Paulownia

A total of 4 primers were tested on tobacco specimens with their sequences presented in Table 10. The presence or absence of each ISSR fragment is scored as a binary unit symbol (1 = present, 0 = absent).

Table 10. Sequences of ISSR primers used to perform PCR reactions with paulownia DNA

Primer	DNA sequence	Length (bp)
ISSR 1	(CA)8AA+GG	20
ISSR 2	(CA)8AA+GC+T	21
ISSR 6	(AC)8C+TG	19
ISSR 7	(AG)8C+TG	19

Corn

A total of 4 primers were used to analyze the hybrids from this crop, which were tested both on mixed DNA (from all grown plants of each hybrid) and on each individual plant, in order to analyze the applicability of the system for DNA profiling of hybrids and to establish their homogeneity. The characteristics of the ISSR primers used are presented in Table 11.

Table 11. Sequences of ISSR primers used for analysis of maize hybrids.

Primer	DNA sequence	Length (bp)
ISSR 7	AG(8)CTG	19
Volume 14	AG(8)YT	18
Volume 11	GA(8)YC	18
Volume 3	CA(7)RG	16

Fusarium and *Phytophthora*

For the study of the two phytopathogenic organisms, 10 primers currently available in the Molecular Biology Laboratory were used. Their sequences are presented in Table 12.

Table 12. Sequences of ISSR primers used for analysis of phytopathogenic organisms.

Name of the primer	DNA sequence	Length (sun)
ISSR 1 (E1)	(CA)8AA+GG	**20**
ISSR 2 (E2)	(CA)8A A+GC+T	**21**
ISSR 3 (E3)	(GA)8C+TC	**19**
ISSR 4 (E4)	(AG)8C+TC	**19**
ISSR 5 (E5)	(AC)8C+TA	**19**
ISSR 6 (E6)	(AC)8C+TG	**19**
ISSR 7 (E7)	(AG)8C+TG	**19**
ISSR 8 (E8)	(AC)8C+TT	**19**
ISSR 9 (E9)	(AG)8C	**17**
ISSR 10 (E10)	(GA)8T	**17**

Goats

ISSR analyses were performed by PCR reactions in a QB-96 Thermal Cycler (Quanta Biotech, London, UK). Total polymorphism levels were estimated from generated multilocus anonymous dominant markers. The sequences of ISSR primers used to perform PCR reactions are listed in Table 13.

Table 13. Sequences of ISSR primers used to perform PCR reactions with goat DNA.

Primer	DNA sequence	Length (bp)
E7	AG (8)CTG	19
Volume 14	AG (8)YT	18
Volume 11	GA (8)YC	18
Volume 5	CT (8)RG	18

Linkage analysis

The cotton map was constructed using MapMaker software. version 3.0 (Lander et al., 1987). All pairs of linked markers were first identified using the "group" command (LOD>5, recombination frequency (RF) = 0.30). The "order" command was used a minimum of three times to establish the best frame order of the markers within the groups. The map was constructed using the Kosambi mapping function (Kosambi, 1943) with the "error correction" command included. The latter command takes into account the probability for each locus that its allelic configuration with respect to neighboring markers arises in part from typing errors. Significant corrections in the total length of the map may result from missing (2% in our case) or erroneous data

(Lincoln and Lander, 1992), as previously observed in barley by (Castiglioni et al., 1998). Additional information available from similar loci placements on other maps or from putative homologous duplications was used to test additional linkages of unrelated groups or individual markers to framework groups using the "try" and "compare" commands. Finally, aligned marker sequences were confirmed using the "ripple" command.
Linkage clusters were plotted using Visual Basic for Excel. The localization and visualization of clusters were assessed by scanning the linkage clusters using a 10 cM interval on the linkage clusters and by merging regions with more than 8 loci.

Statistical analysis and software used

Amplified fragments for AFLP, SSR and ISSR analyses were scored for presence (1) or absence (0) of the corresponding bands at the studied loci.

1. Sample allele frequencies in SSR studies are calculated as pu = nu / (2n), where n is the number of individuals.
2. The genetic diversity index is calculated for each primer and each pattern frequency using formula (1):

(1) $H = 1 - \Sigma pi^2$

where H is the index of genetic diversity and p_i is the frequency of occurrence of the corresponding marker (Nei, 1973).

3. The polymorphism information content (PIC) (Botstein et al., 1980) of a marker corresponds to its ability to detect polymorphisms among individuals of the corresponding population, and the higher this capacity, the greater its value. It is one of the indicators of marker quality in genetic research. PIC values for co-dominant markers such as SSRs range from 0 (monomorphic) to 1 (very highly informative, with several alleles of equal frequency).

The information content of polymorphisms for each SSR (marker polymorphism) is calculated using the formula PIC = 1 - Σpi2, where pi is the frequency of the i-th allele for each SSR marker locus.
For dominant markers (such as ISSRs), the calculation of the information content of polymorphisms can be simplified to:

(2) $PIC = 2f(1-f)$,

where ƒ is the frequency of observed bands in the gel; and 1−ƒ is the frequency of missing bands (De Riek et al., 2001).
For dominant marker systems, we use a binary matrix of presence and absence of bands resulting from the polymorphism revealed by the marker. If we consider the frequency of bands present as *p* and the frequency of absent

bands as *q,* we can arrive at the following general equation for estimating the information content of polymorphisms (PIC):

(3) $PIC = 1 - (p^2 + q^2)$

This equation considers polymorphism based on the frequencies of homozygous individuals for Hardy-Weinberg equilibrium, unlike ***(1),*** which considers the frequencies of heterozygous individuals. This is a simplified equation for quickly calculating the PIC (Serrote et al., 2020), whose theoretical distribution of values is presented in Figure 1.

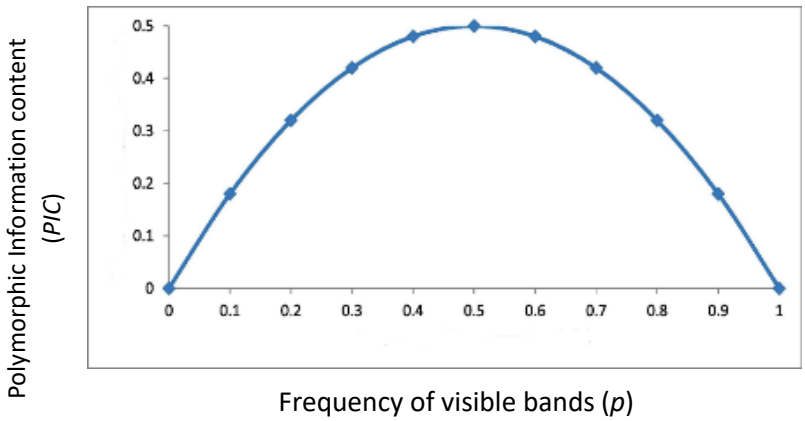

Figure 1. Graphical representation of PIC values as a function of the frequencies of present and absent bands in a gel with PCR-generated fragments.

As can be seen from the figure, high PIC values are obtained when the frequency of stripes (present or absent) is around 0.5. As we move left or right on the graph (when the difference between the frequencies of presence and absence of the stripe increases), PIC values decrease.

4. Nei distance matrix was used to construct the dendrogram with the unweighted group pairwise average method (UPGMA) using Power Marker 3.25 (Liu and Muse, 2005).
5. The molecular data collected during the present study were used to calculate relative genetic distances and create hierarchical clusters with the statistical package “SPSS for Windows”.
6. MapMaker 3.0 software (Lander et al., 1987) was used to construct the cotton genome reference map.
7. The free mapping software MapDisto (Lorieux, 2012) was used for QTL identification and mapping in a segregating population of cotton mutants. The procedure for calculating QTL in recombinant inbred lines (RILs) was used as the most suitable for characterizing a population of sister lines derived from mutational impact.

8. Disease severity index (DSI) was determined using a 0–4 rating scale according to the percentage of leaves affected by acropetal chlorosis, necrosis and wilting (where: 0 = healthy plant, 1 = 1–33% leaf area affected, 2 = 34–66% leaf area affected, 3 = 67–97% leaf area affected, 4 = 100% dead plants). The formula (4) was used to calculate DSI values:

(4) $$DSI = \frac{\Sigma N_i \times i}{4\Sigma N_i} \times 100,$$

where Ni is the number of plants with DSI score i.

The response of the plants was defined as resistant (R), moderately resistant (MR), or susceptible (S) according to the DSI values. The samples were defined as R if the calculated DSI was not higher than 25; MR if the DSI was in the range 25.1–45.0 and S if the DSI was 45.1 and higher (ElSharawy et al., 2015).

9. The following software programs were used for the CAPS analysis:
   - Primer BLAST hosted by NCBI at https://ncbi.nlm.nih.gov/tools/primer-blast/ (Ye et al., 2012)
   - Clustal Omega 1.2.1 for GNU/ Linux (Goujon et al., 2010; Sievers et al., 2011)
   - GENtle 1.9.4 for Microsoft Windows (Manske, 2006).

Results

Phenotypic variation

Cotton

After 5 generations of selfing and reproduction in the greenhouse and in the field, the lines were considered sufficiently homogeneous to determine and compare fiber quality properties to identify QTL.

Those found to carry multiple valuable traits at this stage can be used to initiate breeding as selection materials. Before proceeding in this direction, detailed phenotyping of the main technological characteristics and fiber was undertaken to identify individuals with potential for use in mechanized farming.

Phenotyping for technological traits revealed significant differences in the potential for their improvement within the study population (Table 14).

The results show that the variations generated by the mutagenic treatment are sufficient to identify lines with improved technological characteristics. The average boll weight also showed a wide range of variation, ranging from 3.56 to 7.33 g. Of the 132 progenies studied, 60 had a boll weight equal to or lower than the average, while 31 progenies had a significantly increased boll weight. The combination of increased boll weight with an increased number

of bolls was observed in 13 of the progenies. This kind of combination of useful characteristics has a large effect on the yield potential.

Table 14. Range of technological characteristics observed in the analyzed M6 plants.

Sign / Value	Bolls/plant No.	Boll weight g	Yield (3 bolls) kg / ha	Yield potential kg / ha	Plant height cm	Height of 1st fruiting branch cm	Length of 1st fruiting branch cm
Minimum	1	3.56	1708	1140	53	8	2.0
Maximum	11	7.33	5323	10270	107	30	15.0
Average	5	5.44	3516	5705	80	19	8.5
Untreated parent	6	5.21	2501	4915	82	10	7.0

To be well adapted to mechanical harvesting, cotton varieties should have uniform ripening combined with good earliness. Harvesting 3 bolls/plant results in yields of 2.2-2.5 t/ha (Table 14), which are satisfactory for farmers. Since the fields need to be prepared for the next crop, farmers are willing to forgo the difference to the full potential of the crop so that the combined income from the successive crops is maintained at optimal levels.

For mechanical picking to be successful, the aim is to pick the first 3 bolls, so that the earliest possible harvest can be achieved. As shown in Table 14, the main variety for the country has its first fruiting position 10 cm from the ground, which essentially makes it inaccessible for mechanical harvesting. The results of the mutagenic treatment show that this trait is influenced both positively and negatively in the population. The observed change in this trait in a negative direction was moderate (a decrease of only 2 cm). On the other hand, in some M6 plants a significant improvement was achieved – a very high position of the first fruiting branch (Figure 2, Table 14). Figure 2 shows that the increases in the height of the first fruiting position can result from different effects of the mutagenic treatment. Selecting progeny with a combination of desirable characteristics for mechanical harvesting in our case means selecting those that exhibit simultaneously a high position of the first fruiting branches, a large number of mature bolls at the end of the growing season, and a high average pod weight.

Furthermore, it is desirable that the first fruiting branch is not only in a high position but also grows upwards from this position so that the first mature boll is optimally positioned for mechanical harvesting. Not only must the fruiting branch be in the right position, but its length must be kept under control, as branches that are excessively long bend towards the ground under the weight of the developing bolls. Such bending can result in the fruit being positioned even lower than the position in which the branch itself developed.

Therefore, strong branches of 7-10 cm in length are most desirable for this position (Figure 2, Table 14).
In addition to agronomic characteristics, a successful new variety must possess good fiber quality characteristics. Although they do not directly affect the price a farmer receives for the crop, these characteristics significantly affect the efficiency of cotton ginning operations and are thus a major factor in making decisions about the sources of raw materials for cotton ginning factories. Therefore, fiber quality was also a major indicator tracked in our study. Phenotyping in the fiber properties population revealed significant differences in fiber quality (Table 15).

Table 15. Range of quality characteristics observed in the marked M6 plants.

	micronaire	strength	length	uniformity	elongation
minimum	3	23.8	23.3	79.3	5.9
Maximum	6.2	34.9	30.1	86.4	15.2
Average	4.3	28.8	26.8	83.2	8.9
Chirpan 603 untreated	4.9	27.6	26.6	83.5	9.6

The results show that the generated polymorphism is sufficient to identify loci associated with fiber quality traits. The overall mean value for micronaire, for example, was reduced by 0.6 units, with a lowest observed value of 3.0. In contrast, the overall mean value for fiber strength was increased (by 1.2 g/tex). The offspring with the lowest value also had the lowest micronaire, so the variation in the trait in this case can be attributed to incomplete fiber development.

Figure 2. Phenotypic variation observed in the M6 generation. Leftmost is the untreated parent, which clearly shows the combination of good earliness and yield potential with an undesirable position of the first fruiting branch.

At the other extreme, there were four progenies with fiber strength varying very little (34.2-34.9 g/tex). The micronaire of these progenies also varied little (4.5-4.8), while fiber length was more variable (24.8-27.3 mm). Interestingly, the elongation of the progenies with the longest fibers (27.2-27.3) within this group was close to that of the parental genotype and varied very little (7.7-8.1), indicating that in these progenies the only affected characteristic was fiber strength.

In line with prior expectations, fiber length was affected both positively and negatively with extremes evenly spaced from the mean. Three of the four progeny with the shortest fibers had high to very high micronaire (4.7-6.2) and equal or higher fiber strength than the parental genotype (27.2-29.2g/tex), indicating that these progenies had reached full maturity and their fiber length was actually negatively affected by the mutagenic treatment. At the other end of the range, the picture was not so clear. The progeny with the longest fiber (30.1 mm) had strength (26.8 g/tex), elongation (10.2) and uniformity (83.3%), essentially equal to those of the parent, while its micronaire was significantly reduced (to 3.9). The next three progenies with the longest fiber (29.0-29.3 mm) had significantly increased tenacity (28.1-31.4 g/tex), while the micronaire was moderately to significantly reduced (3.8-4.5).

Of the other fiber characteristics studied, elongation was of greatest interest as a characteristic of significant importance for yarn spinning. Regarding elongation, there was one progeny exhibiting an unusual value. As shown in Figure 2, in one of the progenies this trait had a value of 15.2, thus deviating positively by more than 5 units from the parental genotype and by more than 6 from the population mean. The same progeny had very good fiber length (28.0 mm), strength (28.5 g/tex), uniformity (82.6) and micronaire (4.3), making it a very good genotype for future development as a candidate variety carrying multiple positive fiber characteristics.

The results of clustering the genotypes based on the full set of information on their qualitative and agronomic characteristics are presented in Figure 3. The distribution pattern in the dendrogram would be difficult to explain based only on the analysis of the combinations of traits in each progeny. For example, the top cluster of 10 progenies combines plants with medium to tall plant height (75-95 cm), short to very short fibers (23-27 mm) and low yield potential (2-4 t/ ha). The progenies at the other end of the dendrogram also have short fibers, low yield potential and vary according to plant height. Neither group has any other truly distinctive characteristics. Therefore, explaining the division into the two groups would be difficult, if not impossible, without the PCA results.

First of all, PCA shows that the different traits studied have moderate to low weight in determining the overall distribution in the dendrogram. As Table 16 clearly shows, the variation explained by the loadings of the first two components is less than 50%. When the residual variation is so high, the clustering of genotypes is difficult to identify through empirical observations and only with the help of the breeder's experience.
The true utility of PCA becomes apparent when the data in Table 17 is carefully analyzed. As the table shows, the characteristics with the highest loadings on the first principal component are boll number, cotton seed weight, and yield potential. However, the total variation explained by this component is about 31%.
Therefore, an analysis of the loadings of the traits on the second principal component is necessary, in which the weight of the bolls and the yield (determined on the basis of 3 bolls/plant) appear as the traits with the highest loading on the second component. The fact that the first two components explain less than 50% of the total variation requires attention to the traits that have the highest loading on the third principal component, which are the plant height and the height of the 1st fruiting branch. The inclusion of the third component leads to an explanation of more than 60% of the total variation.

Table 16. Component loadings and percentage of total variation explained by principal component analysis of genotypic diversity in the M6 population.

Explained total variation						
Component	Original own values			Extracted sums of squares loads		
	Total	% of deviation	Cumulative %	Total	% of deviation	Cumulative %
1	3,447	31,333	31,333	3,447	31,333	31,333
2	1,876	17,055	48,388	1,876	17,055	48,388
3	1494	13,581	61,969	1494	13,581	61,969
4	1,428	12,982	74,951	1,428	12,982	74,951
5	1,155	10,496	85,446	1,155	10,496	85,446
6	,939	8,533	93,980			
7	,474	4,310	98,289			
8	,163	1,482	99,771			
9	,025	,229	100,000			
Extraction method: Principal component analysis.						

Since we are working with a segregating population originating from a mutagenic treatment, low component loadings should be expected because the effect of the treatment is not directional and therefore the traits are affected in both a positive and a negative direction.

The observation that boll number, boll weight, total yield potential, seed weight, 3-boll yield/plant, plant height, and 1st fruiting branch height are the traits with the highest loadings suggests that offspring combining simultaneous improvements in most (if not all) of these traits can be expected in the population. Furthermore, they should group into one (or two closely related) clusters on the dendrogram.

In fact, the progeny clustering analysis (Dimova and Bojinov, 2001) revealed that the last group of the left largest cluster unites progenies with high yield potential (7-10 t/ ha), medium to high pod weight (5-7 g), fast initial development (22-23 days from 1st true leaf to flowering) and medium to high height of 1st fruiting branch. In combination with good fiber length (28-31 mm) this group represents the most promising set of progenies in our experiment. Within this group, progenies M_5_2_1 and M_5_2_3 have micronaires of 4.07 and 4.43, fiber strength of 32.9 and 37.7 and uniformity of 89.7 and 89.1. The combination of these characteristics makes these two progenies with the highest potential to develop into new varieties.

Table 17. Relative weight of features by extracted component from PCA.

Principal components matrix					
Signs	Main components				
	1	2	3	4	5
Number of bolls	0.965	0.073	0.028	0.192	0.069
Seed weight	0.886	0.333	0.091	0.279	0.114
Boll weight	-0.546	0.791	0.170	0.191	0.024
Yield 3 b./plant	-0.546	0.791	0.170	0.191	0.024
Yield potential	0.886	0.333	0.091	0.279	0.114
Plant height	0.092	-0.364	0.685	-0.070	0.379
1st fruiting branch height	-0.343	-0.119	0.542	0.017	0.605
1st fruiting branch length	0.077	-0.140	0.274	0.273	-0.358
Cotyledon to 1st true leaf	-0.022	0.01	-0.682	0.041	0.593
Cotyledon to flowering	-0.385	-0.257	-0.309	0.746	0.256
1st true leaf to flowering	-0.265	-0.408	0.140	0.723	-0.259

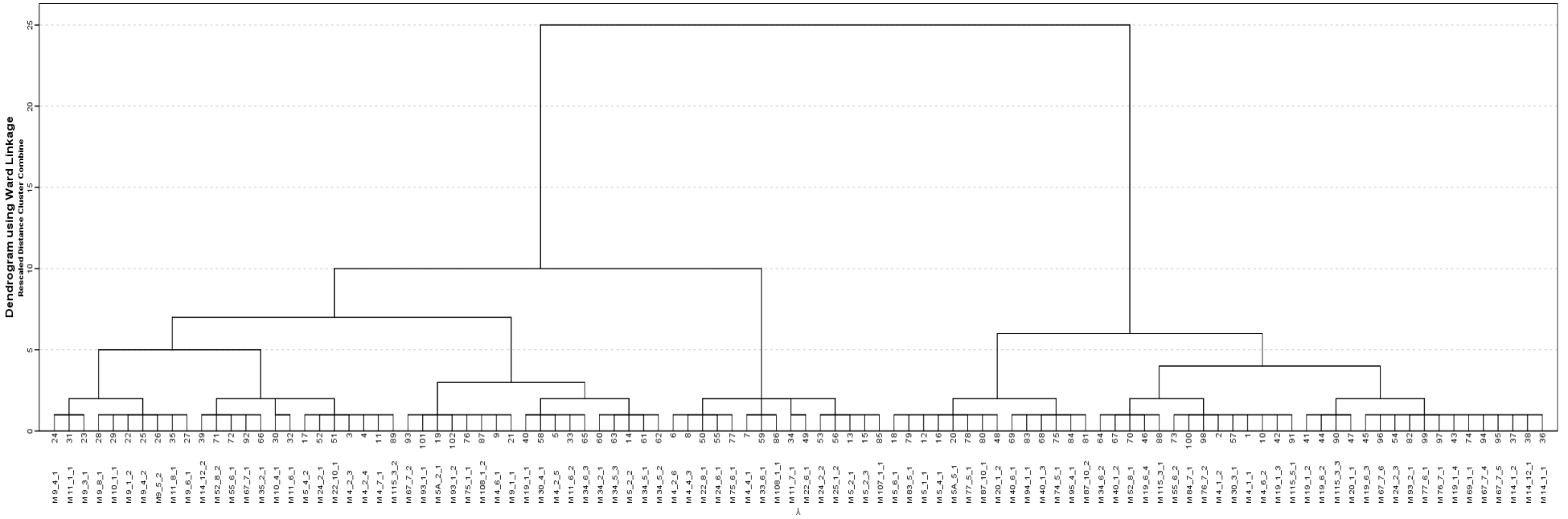

Figure 3. Grouping of genotypes based on both technological and quality characteristics of the fiber.

Tomato

We used 8 tomato cultivars and lines with large variations in fruit characteristics, including size (large, medium, small), shape (round, flat-round, oval), color (yellow-orange, red-orange, red, red-violet, pink, purple-black), etc., as well as growth habit (determinant and indeterminate). However, high morphological variability is not always reflected at the molecular level (Wang et al., 2006) and detailed genetic studies at the molecular level are very necessary when planning crosses between elite germplasm samples.

The studied tomato varieties and lines are characterized by diversity in the main chemical indicators (Table 18).

Table 18. Chemical indicators of fruits of tomato varieties and lines.

Variety / line	Dry matter	Ascorbic acid	Titratable organic acids	Total pigments	Lycopene	Beta-carotene
	Re (%)	(mg %)	(%)	(mg %)	(mg %)	(mg %)
Plovdivska carotena	**5.8** ± 0.5	**54.14** ± 5.53	**0.35** ± 0.02	**9.79** ± 1.99	**4.17** ± 1.04	**4.76** ± 0.82
L21β	**5.3** ± 0.8	**30.57** ± 4.87	**0.38** ± 0.05	**4.79** ± 0.99	**-***	**4.46** ± 0.92
L1116	**4.9** ± 1.0	**33.55** ± 4.91	**0.35** ± 0.08	**6.85** ± 1.22	**6.15** ± 1.36	**0.28** ± 0.19
L1140	**7.0** ± 0.3	**40.80** ± 3.33	**0.54** ± 0.05	**5.21** ± 0.49	**2.80** ± 0.45	**2.12** ± 0.62
IZK Alya	**9.7** ± 0.7	**60.09** ± 6.91	**0.66** ± 0.05	**9.81** ± 1.01	**9.23** ± 1.18	-
L984	**5.1** ± 0.2	**25.06** ± 1.38	**0.38** ± 0.05	**9.03** ± 0.89	**8.64** ± 1.16	-
L975	**5.4** ± 0.5	**19.23** ± 2.69	**0.37** ± 0.04	**7.48** ± 0.52	**7.08** ± 0.69	-
L53β	**5.4** ± 0.34	**27.59** ± 2.81	**0.37** ± 0.03	**3.68** ± 0.39	**0.65** ± 0.10	**2.78** ± 0.34

* - the quantity is below the sensitivity threshold of the method

Goats

Based on the study conducted in the populations of the two autochthonous goat breeds (Kalofer Long-Haired Goat and Bulgarian Vitorog Long-Haired Goat) raised in Southwestern Bulgaria, several basic colors can be differentiated.

Black pigmented coat. The coat is uniformly, deep black, without any lightening areas on the body. It occurs in both autochthonous breeds. Using the classification of goat colors of (Adalsteinsson et al., 1994), we have a typical eumelanin pigmentation. With age, in completely black specimens, gray fibers are also observed along the dorsal line, especially pronounced in the croup area.

Red -brown pigmented coat. The color of this type of pigmentation varies from dark chocolate brown to light ochre. Characteristic for this color is that

the mucous membranes of the lips, nose and eyelids are a deep brick brown. There is no black pigment whatsoever. It occurs in both autochthonous breeds, it should be noted that the kids are born reddish-brown and subsequently the hair on the body may lighten.

Silver grey pigmented coat. The *coat* consists of evenly distributed white and black hairs, giving a uniform silver grey hue to this colour type. It can vary from lighter to darker depending on the proportion of white and black hairs. Lightening is observed around the eyes towards the muzzle, the inner part of the ears and the back of the limbs. It is widely found in both autochthonous breeds. It corresponds to a pigmentation type determined by the Agouti locus, and is designated as "striped grey" according to (Adalsteinsson et al., 1994).

"Barza" type pigmentation. An extremely specific coloration that the local population defines with the name "barza" and is a combination of zonal light coloration of the front half of the body, with the back being dark – black, forming a "dark/black cloak" on the croup and loin. Due to the long hair cover in the two local breeds, the black "cloak" can descend almost to the ground in the back of the body. The fur on the lower part of the legs is dark to black. On the front part of the head there are very specific longitudinal black stripes – one above the eye, the other below it. Some variation is observed in the coloring of the light area of the body in this coloration. It can be from almost cream-white to brick-brown.

Black "garish" pigmented hair coat. A characteristic form of pigmentation of the hair coat in goats, in which on a black background of the body there are specifically localized spots of a lighter brownish color. The specific localizations of the light spots are on the face of the head, the inner side of the ears, the lower part of the limbs, the abdomen and the tail. The lighter spots can vary in color from a rich brick brown to almost white. This color is mainly found in the Bulgarian Vitorog long-haired goat. It is extremely rarely observed in the population of the Kalofer long-haired goat, and is very often perceived as a sign of crossbreeding. Some authors, depending on the size and location of the distribution of the light spots on the body, define this color as "black and tan", determined by the Agouti locus.

Intermediate forms of hair pigmentation. Genetic interactions under the influence of the Agouti locus determine an exceptional diversity in pigmentation, combining different elements of the basic colors. Therefore, there are a number of forms of hair pigmentation of autochthonous goats, combining elements of the basic color types, and their determination is significantly difficult. This diversity of intermediate forms is expressed to a significant extent in the population of the Bulgarian long-haired goat.

All of the above-mentioned variations in pigmentation make phenotypic characterization of individuals very complex and in many cases unreliable. Therefore, the development of more reliable ways to establish the genotypic affiliation of individuals is of particular importance for such local animal breeds.

DNA extraction

Cotton

Immediately before flowering, 10 typical plants were selected and DNA was extracted from them using the standard set of chemicals from Amersham-Pharmacia (PhytoPure DNA extraction kit). The extraction results are presented in Figure 4. As can be seen from the figure, the DNA obtained during the extraction is of similar quality (lack of degraded fragments) and in approximately the same amounts as most extractions. The exceptions are No. 1, 2, 6 and 10, in which either the amount of DNA obtained is smaller (No. 6 and 10) or practically absent (No. 1 and 2). From such plants, it was necessary to conduct additional DNA extraction in order to obtain a sufficient quantity and quality of the extracted DNA, so that subsequent analyses would be possible with DNA from all marked plants.

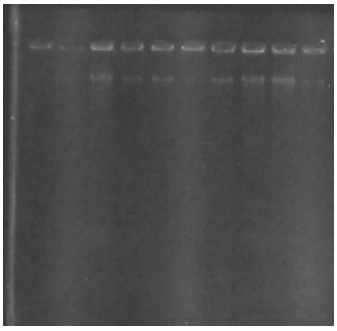

Figure 4. Results of the extraction of total genomic DNA from young leaves of the Chirpan 603 variety

Tomatoes

The DNA extraction procedure resulted in the production of high-quality genomic DNA from both cultivated and wild tomato genotypes. In most cases, between 300 and 500 µg of genomic DNA was isolated. When the quantity or quality of the DNA obtained was considered unsatisfactory, the extraction was repeated so that these criteria could be met.

Tobacco

Initially, plants were selected and DNA was isolated and purified from them using the standard set of chemicals from Amersham- Pharmacia (PhytoPure DNA extraction kit).

Barley

Total genomic DNA was extracted from 24 samples included in the study. In most cases, the quality of the DNA obtained was good enough. When the quantity or quality of DNA was considered unsatisfactory, the extraction was repeated.

Pepper

from all samples using the standard Omega Bio-Tek chemical set. The DNA extracted was of similar quality (no degraded fragments) and in approximately equal amounts (in most cases between 300 and 500 µg genomic DNA). When the quantity or quality of the DNA obtained was considered unsatisfactory, the extraction was repeated to meet the specified parameters.

Paulownia

At the beginning of the growing season, plants were selected and DNA was isolated and purified from them using the standard set of chemicals from Amersham- Pharmacia (PhytoPure DNA extraction kit).

Corn

DNA from the genotypes used was isolated and purified using the standard BioOne Plant DNA extraction mini kit. The DNA obtained during extraction was of similar quality and in approximately equal amounts from each extraction.

Wheat

Initially, a selection of forms was carried out and during the vegetation period, DNA was isolated and purified from them using the standard set of chemicals Plant DNA Preparation Kit. The results presented in Figure 5 show that the DNA obtained during the extraction was of appropriate quality

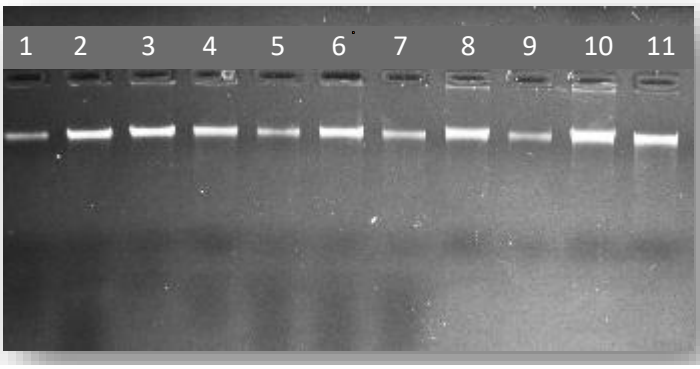

Figure 5. Results of the extraction of total genomic DNA from young leaves of 11 of the studied varieties. (1-Predel, 2-Avenue, 3-Reyadur, 4-Saya, 5-Teres, 6-Solechio, 7-Enola, 8-Pibrac, 9-Sofru, 10-Helix, 11-Flavor).

(absence of degraded fragments) and in approximately equal amounts from each extraction.

Goats

For ISSR analysis, genomic DNA from both goat genotypes was isolated with the standard innuPREP DNA kit (AnaliticaJena), according to the supplier's recommendations. Isolation of sufficient quantity of genomic DNA with good quality was achieved when 15 hair follicles were used (Figure 6), which allowed for successful ISSR analyses later. This resulted in a total of 300 to 500 µg of genomic DNA. When the quantity or quality of the obtained DNA was considered insufficient, the DNA extraction procedure was repeated until adequate results were obtained.

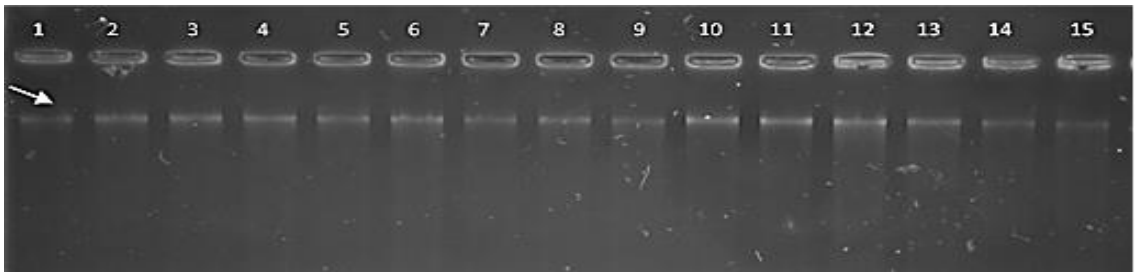

Figure 6. Results of genomic DNA extraction from hair follicles with the in-nuPREP DNA Kit (arrow points to the position of the genomic DNA fragments).

Levels of polymorphism of molecular markers and sample profiles

Co-dominant markers

Cotton

Approximately 90% of the RFLP probes and SSR primers revealed polymorphic bands for at least one locus between the two parents, “Guazuncho 2“ and “VH8”. Among the 377 polymorphic RFLP and SSR markers for which banding could be clearly inferred, 306 (81%) were duplicated, confirming the high evolutionary rate of duplication in the tetraploid cotton genome. Two-thirds of these duplications were either polymorphic for both loci (four bands in F1) or polymorphic at one locus and monomorphic at the second (three bands in F1). Sixty-three of the polymorphic RFLPs and SSRs were inferred to be of the single-locus type (one single and distinct band in each parent).
A total of 15 SSR primer pairs were screened for polymorphism between lines to conduct molecular marker analysis in the mutant population. Of these, 5 did not generate any polymorphisms, 6 generated 2 polymorphic bands and 4 generated more than 2 polymorphic bands. A total of 27 polymorphisms were identified in the studied population, most of which had either single or very few (2-6) bands.

Tomatoes

A total of 165 publicly available tomato microsatellite markers were used to assess genetic diversity in a set of 8 Bulgarian inbred lines and cultivars. Among them, 100 were genomic and EST-SSR markers from the SOL Genomics Network (http://solgenomics.net); 32 were TGS and TES markers from the Kazusa Tomato Genomics Database (https://marker.kazusa.or.jp/app / crop.php?crop = tomato); 16 were SLMs developed from anchored BAC clones on chromosomes 6 and 12 (Geethanjali et al., 2011) ; 16 were TMS and EST markers developed by (Areshchenkova and Ganal, 2002) and one SSR - LEMDDna (Smulders et al., 1997). The markers were chosen to cover the 12 chromosomes with a minimum of 7 SSRs per chromosome and to be located close to already published QTL and genes responsible for phenotypic variations in quality and other morpho-physiological traits.

This study used standardized PCR conditions at two different melting temperatures (50°C and 63°C), allowing for further amplification of markers that could be further developed for multiplex amplification.

Standardized PCR conditions were achieved by performing an initial optimization step for each locus-specific primer concentration (20, 30, 50, 60, and 80 nM). Adjusting the concentration of each locus-specific primer allows PCR specificity and yield to be controlled. These conditions also help to prevent nonspecific binding of locus-specific primers during the first few PCR cycles. Our test showed that primers for all genomic SOL SSRs, TES, TGS, TMS, EST, and LEMDDNA markers gave clear amplification products at a concentration of 20–30 nM, while SLM primers showed the best amplification at a concentration of 60 nM.

In our study, only five out of 165 primers (3.03%) failed to amplify the expected PCR fragments. Seventy-nine (49.38%) markers amplified monomorphic bands, while the remaining 81 (50.62%) markers generated polymorphic ones.

In this study, using the novel fluorescence-based SSR genotyping protocol developed by (Hayden et al., 2008) and (Tsonev et al., 2013), we reported a lack of amplification at only a few SSR loci compared to other studies. For example, (Benor et al., 2008) reported that 31.7% of the 60 SSRs used failed to amplify the expected PCR products, while (El-Awady et al., 2012) reported a total of 10% SSRs that were unable to produce amplicons.

A total of 299 alleles were identified in 160 SSR loci in the present study. The average PIC for all 160 SSR markers was 0.196 with values ranging from 0.00 for markers generating monomorphic stripes to 0.786 for marker SLM6-7. The number of alleles per locus ranged from 1 to 6 with an average of 1,869 alleles per locus (Table 19).

Table 19. Number of amplified microsatellite loci, total and average number of alleles, gene diversity (GD), observed heterozygosity (Ho), and polymorphic information content (PIC).

	Number of amplified loci	Number of alleles	GD	Ho	PIC
Total	**160**	**299**	35,553	6,625	31,439
Average		**1,869**	**0.222**	**0.041**	**0.196**

These results indicate that allelic variation in the studied tomato varieties and lines is limited.

In our study, the majority of polymorphic SSR loci (55.55%) generated two and three (29.63%) alleles. AT and TA were the most common repeat types (22.22% and 17.28%, respectively), followed by CGG, TTC and AAG, each with 3.7%. Contrary to the reports of (He et al., 2003) and (Benor et al., 2008), which found no relationship between PIC and the number of nucleotides per repeat, our study is in correlation with those of (Blair et al., 1999) and (Jones et al., 2001). Both authors reported that the level of polymorphism in trinucleotide repeats was lower than that in dinucleotide repeats in rice and ryegrass.

The SSR marker groups used here show different levels of polymorphism in the 8 Bulgarian tomato genotypes studied (Table 20).

Table 20. Total number of polymorphic loci detected with different marker groups.

Marker groups	Number of microsatellite loci	Number of polymorphic loci	Polymorphic loci (%)
TMS and EST	15	10	**66.66%**
SSR	99	50	**50.51%**
SLM	15	15	**100.00%**
TGS and TES	30	5	**16.66%**
LEMDDNA	1	1	**100.00%**
Total	**160**	**81**	**50.62%**

The highest level of polymorphism (100%) was generated with SLM markers developed from anchored BAC clones on chromosomes 6 and 12 (Geethanjali et al., 2011). All 15 amplified SLM loci were polymorphic. Among the 99 genomic and EST-SSR markers used from the SOL Genomics Database (http://solgenomics.net), 50.51% were polymorphic, while TMS and EST SSR markers of (Areshchenkova and Ganal, 2002) showed 66.66% polymorphism. The lowest level of polymorphism (16.66%) was observed in the TGS and TES loci (https://marker.kazusa.or.jp/app/crop.php?crop=tomato). The latest markers are expected to be more suitable for distinguishing elite tomato germplasm from its wild relatives or local species. Since some of the markers used here are located close to QTL for qualitative traits and genes for some

morpho -physiological characteristics (Mazzucato et al., 2008), their further use in association studies is highly recommended.
Of the 160 SSR markers used, the most polymorphic are the SLM markers originating from overlapping BAC clones involving parts of chromosomes 6 and 12, followed by TMS (Areshchenkova and Ganal, 2002) and SSR (http://solgenomics.net), and the least polymorphic are the TGS and TES markers (http://marker.kazusa.or.jp/Tomato). The latter show a higher level of polymorphism between wild relatives and cultivated tomato than within the species Lycopersicum esculentum (Shirasawa et al., 2010) and could be used to demonstrate introgressions in the case of crosses between them. In the present study with this group of markers, polymorphism was observed in only 5 loci out of a total of 30 amplified (16.6% in total), 3 of which are known to be localized adjacent to genes encoding *Glucan endo-1 3-beta-glucosidase 3, Alkaline ceramidase 3-like* and *Exocytosol complex component 7* (TES1276, TGS1606 and TES1711), and for the other two (TGS914 and TGS959) – to/in genes encoding as yet unknown proteins.
Polymorphism between the studied genotypes with the group of SSR markers (total 99 amplified) from the American database (http://solgenomics.net) was identified in 50 loci.
The group of SLM markers is extremely interesting (Geethanjali et al., 2010; Geethanjali et al., 2011). Of the 15 loci analyzed, all show polymorphism in the studied genotypes.

Dominant markers

Cotton

After obtaining genomic DNA with the necessary characteristics, the work continued with the search for hereditary determinants for which a close association with the observed phenotypic variation of the target traits (earliness, fiber quality and technological traits) could be established. Due to the lack of prior information on the potential localization of the sequences that determine the observed phenotypic variation, as a first step, a screening of the genomes with polymorphic anonymous markers was undertaken to cover them to the maximum extent. The 64 AFLP primer combinations yielded a total of 1294 polymorphic bands (i.e. specific for “Guazuncho 2“ or “VH8"), of which 726 were specific for “VH8“ and, therefore, informative for a backcross population (Table 21). Considering an average of 91 bands per primer pair, the polymorphism revealed by AFLP between the two parents G. hirsutum and G. barbadense was 22%. Between 2 and 27 (mean 11.3) informative AFLP markers were counted per primer pair (Table 21). There was no correlation between the total number of bands and the number of polymorphic

ones. A positive correlation was observed between the number of polymorphisms and the A/T content of the +3/+3 EcoRI and MseI selective nucleotides, confirming the results of (Keim et al., 1997) and (Young et al., 1999). The mean number of polymorphic AFLPs was 13.8, 11.9 and 7.8, corresponding to 0.66, 0.5 and 0.33% A/T content, respectively. Of the 726 informative AFLPs, 576 could be reliably confirmed in the BC1 population and therefore used for mapping analysis.

Table 21. Results from using AFLP primer combinations.

		Mse I primer	M1	M2	M3	M4	M5	M6	M7	M8		
	Selective nucleotides		CAA	CAC	CAG	CAT	CTA	CTC	CTG	CTT	Total	Average
EcoRI primer												
E1	AAC	Polym.	11	14	8	11	19	16	9	11	99	12.4
		Total	142	98	105	136	121	113	91	132		117
E2	AAG	Polym.	9	9	13	15	17	27	16	13	119	14.9
		Total	157	110	105	148	142	128	89	156		129
E3	ACA	Polym.	16	13	8	15	24	16	18	16	126	15.8
		Total	122	100	72	129	112	89	78	115		102
E4	ACT	Polym.	10	15	7	14	12	10	11	8	87	10.9
		Total	106	81	96	92	84	76	58	91		86
E5	ACC	Polym.	7	9	11	10	12	11	7	4	71	8.9
		Total	99	71	53	93	92	69	67	83		78
E6	ACG	Polym.	3	6	5	11	20	16	7	10	78	9.8
		Total	59	88	69	49	59	60	71	65		65
E7	AGC	Polym.	13	10	2	13	15	10	4	13	80	10
		Total	82	68	67	91	69	58	77	91		75
E8	AGG	Polym.	16	5	6	3	8	11	5	12	66	8.3
		Total	103	65	60	89	76	56	59	71		72
	Total	Polym.	85	81	60	92	127	117	77	87	726	
	Average	Polym.	10.6	10.1	7.5	11.5	15.9	14.6	9.6	10.9		11.3
		Total	109	85	78	103	94	81	74	101		91

The AFLP marker system was also used to analyze the diversity in the original Bulgarian variety Chirpan 603 and its M1 mutant generation.

The analysis in this group of individuals was focused on the search for hereditary determinants for which a close association could be established with the observed phenotypic variation of the target traits. Due to the lack of prior information on the potential effective sequences, a first stage was undertaken to screen the genome with polymorphic anonymous markers to cover it to the maximum extent.

For reasons described elsewhere (Bojinov & Lacape, 2003), AFLP analysis was initially chosen to identify DNA polymorphisms between the two genotypes. Four EcoRI/ MseI primer pairs were used, which demonstrated low levels of polymorphism between the genotypes. This result is in full agreement with the low levels of polymorphism observed in a population of over 160

genotypes from the CIRAD gene bank, studied with all possible 64 primer combinations of the Life Technology AFLP™ analysis system I (our unpublished studies). In order to determine informative polymorphisms, the results of fiber quality measurements in segregating populations are needed.
In this first phase of the study, M1 plants from irradiated seeds of the original Bulgarian genotype were also included. Although a number of changes were observed in terms of phenotypic characteristics (partial albinism, modified sequence of flower formation, altered plant habit), plants from the M1 generation had slightly higher levels of DNA polymorphism compared to the original forms. Our assumption is that this somewhat unexpected result is due to two main factors – the selection of a relatively small number of plants for DNA extraction (8 from each parental genotype and each of the M1 populations) and the need to carry out the DNA extractions before the result of determining the fiber qualities was known. However, the extraction of DNA from a larger number of plants from the M1 populations and their testing with a larger number of primer pairs was not justified for the reasons described above.
The second marker system used was ISSR. With its help, the studied genotypes were characterized, the results of which (Figure 7) showed that the source material used was sufficiently homogeneous for the purposes of the present study. At the same time, the source variety Chirpan 603 can be easily differentiated from other samples using the selected marker system. The figure presents the results of the comparative characterization of two genotypes – Chirpan 603 (lanes 1-8) and 2D-18 (lanes 9-16). As can be seen, both genotypes are composed of highly homogeneous groups of individuals, while both groups are well distinguishable in terms of their DNA profiles.
When using this marker system to characterize a mutant segregating population (M6), good discrimination of individuals was achieved with only a limited set of primers. As a result of using only 4 primers, a total of 21 polymorphic loci were identified, which allowed obtaining unique profiles for the majority of individuals. The results are visually presented in Figure 8 by a dendrogram of the distribution of individuals in terms of their relative genetic relatedness.

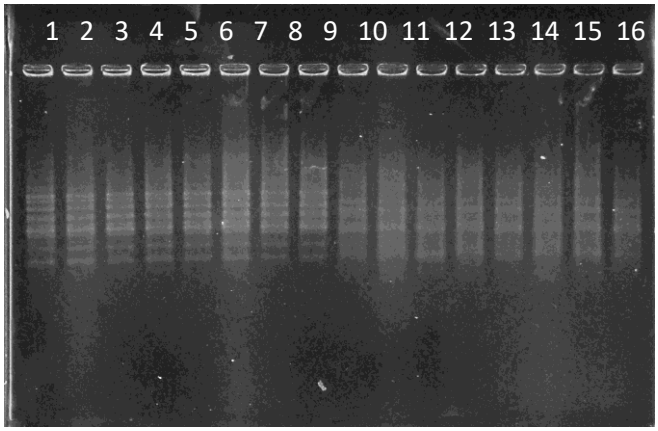

Figure 7. Results of the comparative study of two genotypes using one ISSR primer. Lanes 1-8 – variety Chirpan 603, lanes 9-16 – breeding line 2D-18.

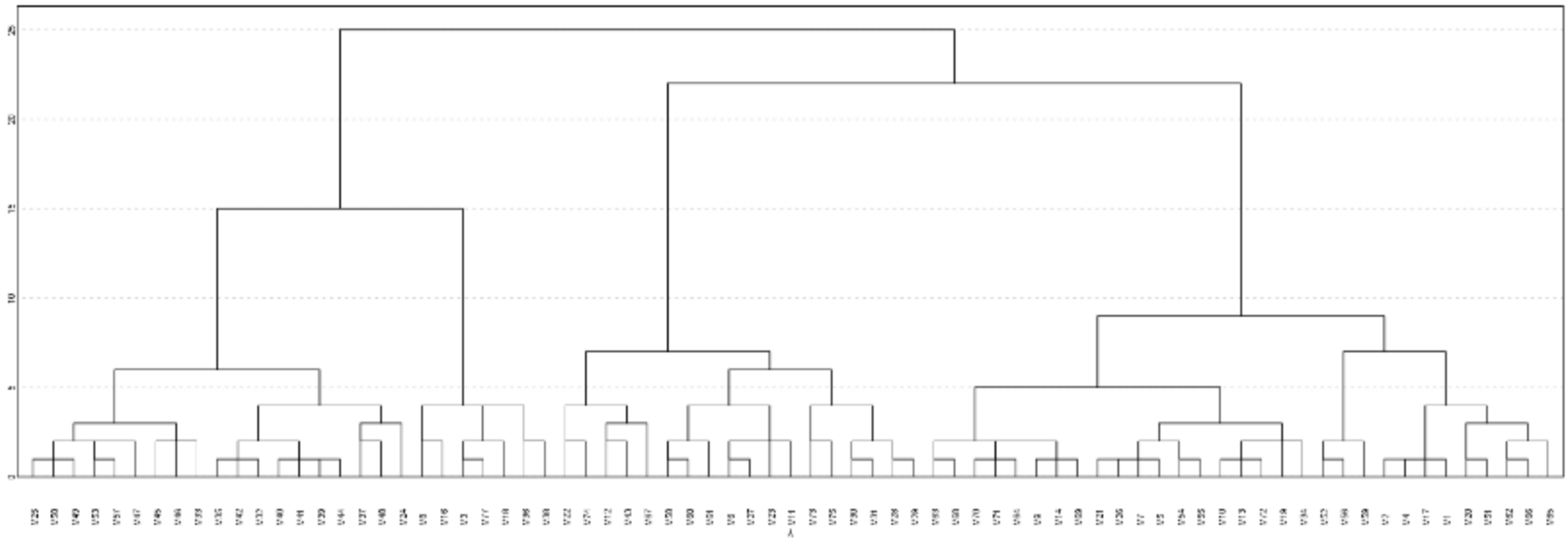

Figure 8. Grouping of individuals from the M6 population based on genetic diversity, established with ISSR markers.

Tomato

CAPS marker

After PCR with the CAPS marker, a product corresponding in size to the expected 703 bp was obtained from all plants (Shi et al., 2011). The fragment was subjected to restriction with each of the endonucleases – BoxI, KspAI, Alw21I. In all analyzed plants from lines 709, 720, 868, 1042 and one from line 1048 (Table 22), two fragments (458 and 245 bp) were obtained only after restriction with the enzyme KspAI. The initial PCR product of 703 bp remained intact after treatment with BoxI and Alw21I (Figure 10). The restriction profile of the positive control (line 975) was similar (Figure 9). This result indicates that the lines are homozygous for the Tm-22 allele (Table 23). Restriction analysis of the PCR product in all plants of line 981 and five plants of line 1048 (Table 22) showed two fragments with sizes of 538 and 165 bp after restriction with BoxI, and 458 and 245 bp after restriction with KspAI (Figure 10). The obtained restriction profiles with the CAPS marker correspond to a heterozygous genotype (*Tm-2 ^{2}/tm-2) (Table 23). The presence of the Tm-2* allele was not detected in any of the tested breeding lines.

Table 22. Summary table of identified resistance genes in six tomato breeding lines.

	Lines					
Genotype	**709**	**720**	**868**	**981**	**1042**	**1048**
tm-2/Tm-2 ^{2}	0	0	0	3	0	5
Tm-2 ^{2}/Tm-2 ^{2}	2	2	3	0	4	1

Table 23. Restriction profiles of a 703 bp PCR fragment with BoxI, KspAI and Alw21I for identification of alleles at the Tm-2 locus.

Allele	Box I	KspAI	Alw21I
tm-2	538 + 165 bp	703 bp	703 bp
Tm-2	703 bp	458 + 245 bp	358 + 345 bp
Tm-2 ^{2}	703 bp	458 + 245 bp	703 bp

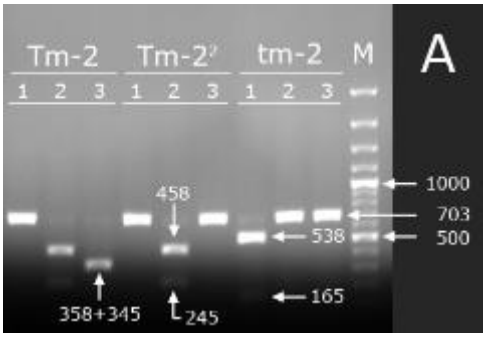

Figure 9. Restriction analysis of PCR product with CAPS marker from control plants homozygous for Tm-2, Tm-22 and tm-2. 1 – BoxI; 2 – KspAI; 3 – Alw21I; M – DNA marker, GeneRuler™ 100 bp Plus DNA Ladder (Thermo Scientific).

To confirm and complement our results, *an in-silico* analysis of the primers for the CAPS marker was performed against the sequences of the three alleles at the Tm-2 locus (AF536199 - *tm-2,* AF536200 - *Tm-2* and AF536201 - *Tm-2* [2)], published in the NCBI database. Primer BLAST was used for this purpose. (Ye et al., 2012). As a result, it was found that the resulting virtual product corresponds in size to the one obtained by us (703 bp).

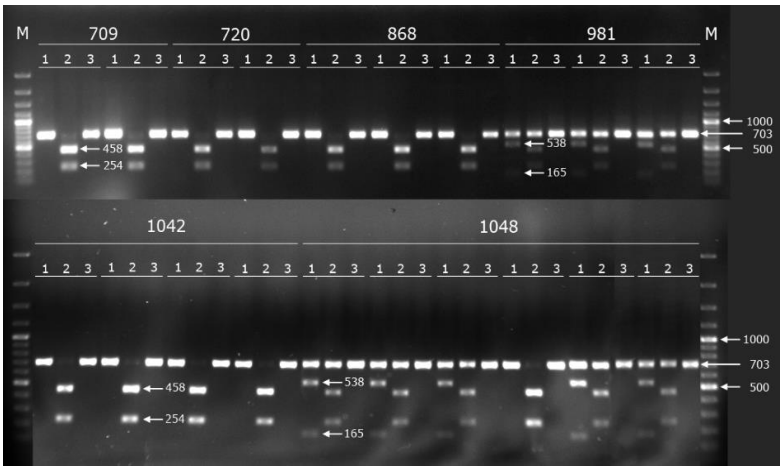

Figure 10. Restriction analysis of PCR product with CAPS marker from resistant plants of breeding lines: 709 and 720 – with 2 plants each; 868 and 981 – with 3 plants each; 1042 – with 4 plants; 1048 – with 6 plants. 1-BoxI; 2-KspAI; 3- Alw21I; M – DNA marker, GeneRuler™ 100 bp Plus DNA Ladder (Thermo Scientific). The size of the specific products is indicated in bp.

Subsequently, the virtual sequences were aligned using the software program Clustal Omega. (Goujon et al., 2010; Sievers et al., 2011) and the presence of restriction sites for the enzymes listed in Table 23 was confirmed. There are 39 differences between the three fragments, with 36 of them being between the recessive and dominant alleles only.

The polymorphism between the three sequences at 359 bp serves as a restriction site for the Alw21I enzyme, used to characterize the *Tm-2 allele.* A

polymorphism at 459 bp is found between *tm-2* and the two dominant alleles (*Tm-2* and *Tm-2 2*), which serves as a restriction site for the KspAI enzyme, characterizing the dominant alleles. A restriction site is also present due to the presence of a SNP in *tm-2,* which serves to characterize this allele using PshAI.

ISSR markers

For reasons discussed above, our preferred initial screening system is based on tandem repeat polymorphisms. ISSR markers are relatively simple and inexpensive to develop, while providing multilocus screening in a single reaction and good reproducibility between different laboratories. As a first step, we screened individual plants from each of the genotypes (Figure 11) to test the capacity of the selected marker system to reveal a sufficient number of polymorphisms within the species *Solanum lycopersicum.*

The results of this initial screening demonstrated the capacity of the selected marker system to differentiate genotypes within the genus. Figure 11 presents the results obtained with 16 of the genotypes. Two groups can be clearly identified, with the wild species group being significantly more heterogeneous. Even in this initial screening, each wild species could be identified by its own profile, easily distinguishable from other wild species. Differences within the group of breeding lines and cultivars are much less pronounced, with some of them being completely indistinguishable with respect to their profiles obtained with primer ISSR 2 (Figure 11).

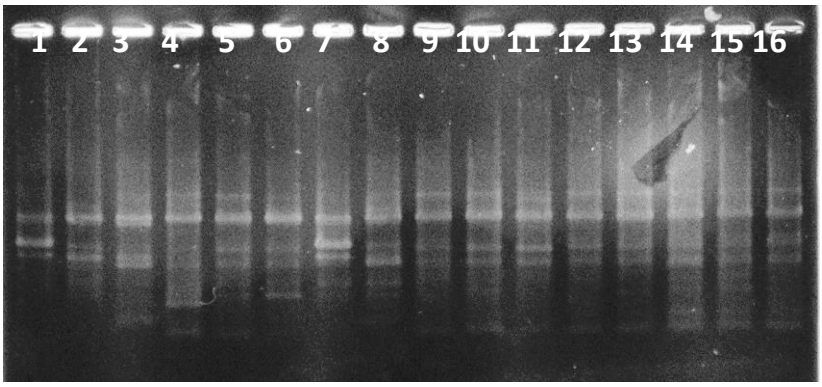

Figure 11. Results of testing wild and cultivated species with ISSR primer 2. Genotypes are arrayed on the gel as follows: 1. Lycopersicum hirsutum IVT 66087, 2. Lycopersicum glandolosum IVT 48090, 3. Lycopersicum minutum 911, 4. Lycopersicum pimpinellifolium LA 0722, 5. Lycopersicum esculentum var. cerasiforme LA 1226, 6. Lyco-persicum parviflorum LA 2072, 7. Lycopersicum f. glabratum LA 2864, 8. Lycopersicum peruvianum LA 2172, 9. cv. Table, 10. cv. Bella, 11. 541/06, 12. 560/06, 13. 561/06, 14. 564/06, 15. 573/06 and 16. 618/06.

The above results were obtained using genomic DNA extracted from a single plant for each genotype. In order to verify the usability of the individual

bands as specific for each genotype, pooled DNA samples should preferably be used. It should be borne in mind, however, that the use of pooled DNA samples masks the levels of inherent heterogeneity within each genotype and thus may compromise the reproducibility of the results if heterogeneous material with incomplete representation is included in the sample. The chances of such heterogeneity occurring within each sample are not negligible, as discussed by (Cooke et al., 2003) and (Bredemeijer et al., 2002). Without prior DNA data for the genotypes used in our study, this possibility had to be assessed before pooling samples from selected plants.
The expansion of the initial screening towards increasing the number of individuals per genotype showed the capacity of the selected marker system to differentiate even individual plants within the samples. Figure 12 presents the results obtained with 7 plants from one of the genotypes.

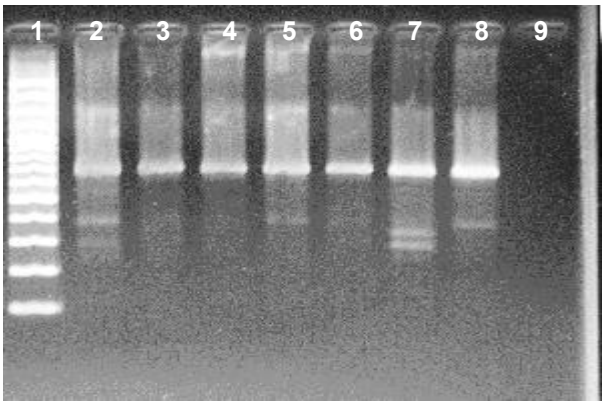

Figure 12. Revealing heterogeneity in products obtained with primer ISSR 2. Lane 1 – standard size DNA with bands of 500 and 1000 bp appearing as strong signals. Bands 2-8 – products of PCR reactions with individual plants of line 216. Lane 9 – control.

Similarly, the assessment of the existing heterogeneity in each of the cultivated genotypes was performed by testing 7 plants with all ISSR primers and looking for the appearance of unusual bands (Figure 14, Figure 13). Interestingly, none of the genotypes showed complete identity of the profiles of individual plants (as would be expected from a strictly self-pollinating species).

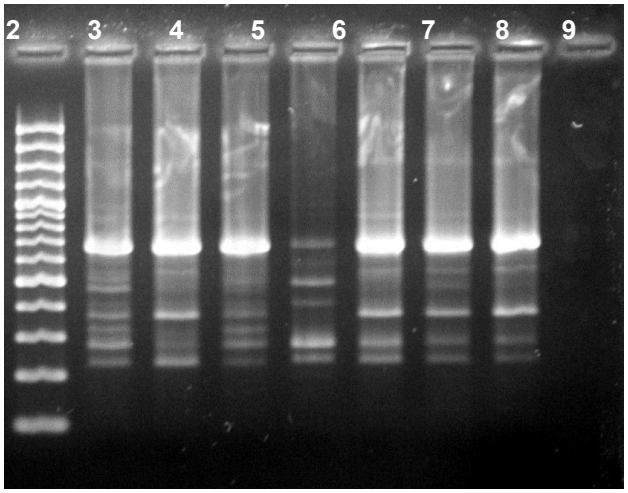

Figure 13. Detection of heterogeneity using ISSR primer 1. Lane 1 – standard size DNA with bands of 500 and 1000 bp appearing as strong signals. Lanes 2-8 – PCR products from individual plants of line 1140. Run 9 – control.

In all genotypes, at least one plant showed differences in the resulting profiles. As shown in Figure 13, plant No. 4 (Lane 5) had a band of about 380 bp, which was present in all other plants of the same variety. On the other hand, a band of 410 bp appeared as a relatively strong signal in this plant, which is unusual for other representatives of the same genotype.

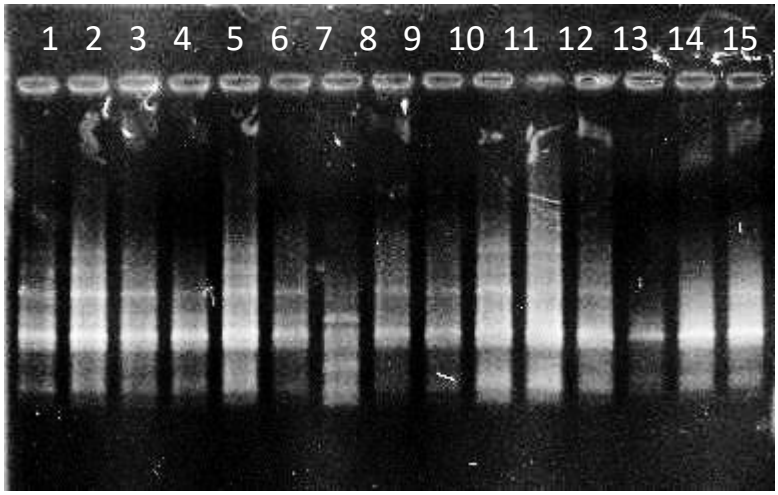

Figure 14. Detection of heterogeneity within cultivated genotypes with primer ISSR 7. Runs 1-12 – PCR products from individual plants of cv. Trapezitsa. Runs 13-15 – DNA PCR products from individual plants of cv. Bella.

Wheat

Similar to the analyses conducted on tomatoes and cotton, the results of the wheat samples demonstrated the ability of the ISSR system to distinguish samples of the same species. In this crop, we tested 5 ISSR primers, which produced a total of 55 polymorphic bands. Of these, the smallest number (5 polymorphic bands) was produced by primer T11, and the largest (11 polymorphic bands) by primer T14. At the same time, the markers thus obtained demonstrated significantly different levels of PIC (0.165 -0.496; average 0.335).

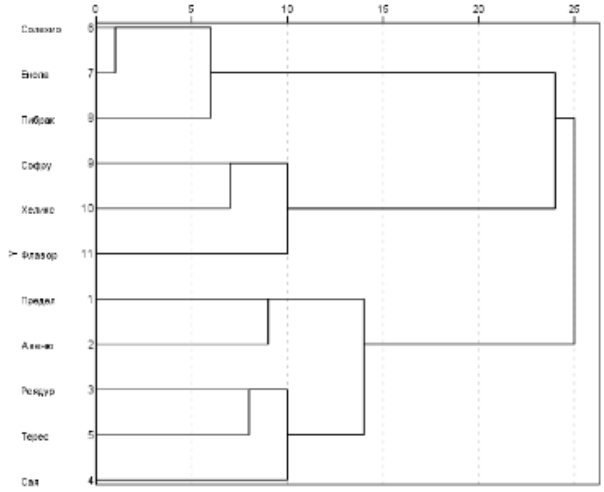

Figure 15. Grouping of wheat samples as a result of analysis with 5 ISSR primers.

The distribution of genotypes on the dendrogram (Figure 15) shows that, using the genetic profiles thus obtained, the studied set of samples is divided into three relatively equal-sized clusters. The first includes Solechio, Enola and Pibrac, the second – Sofru, Helix and Flavor, and the largest – Predel, Avenue, Readur, Teres and Saya.

Barley

ISSR analysis performed in barley revealed a total of 83 DNA fragments, of which 45 (54.22%) were polymorphic (Table 24). Amplification with primers ISSR_4, ISSR_6, ISSR_7, ISSR_8, ISSR_5 and ISSR_6a led to the identification of a number of polymorphisms among the simple repeat sequences. Figure 16 clearly shows a significant number of polymorphic fragments.

A total of 45 polymorphic bands were identified, the combinations of which could create unique banding patterns (DNA profiles) for each individual genotype (Figure 16).

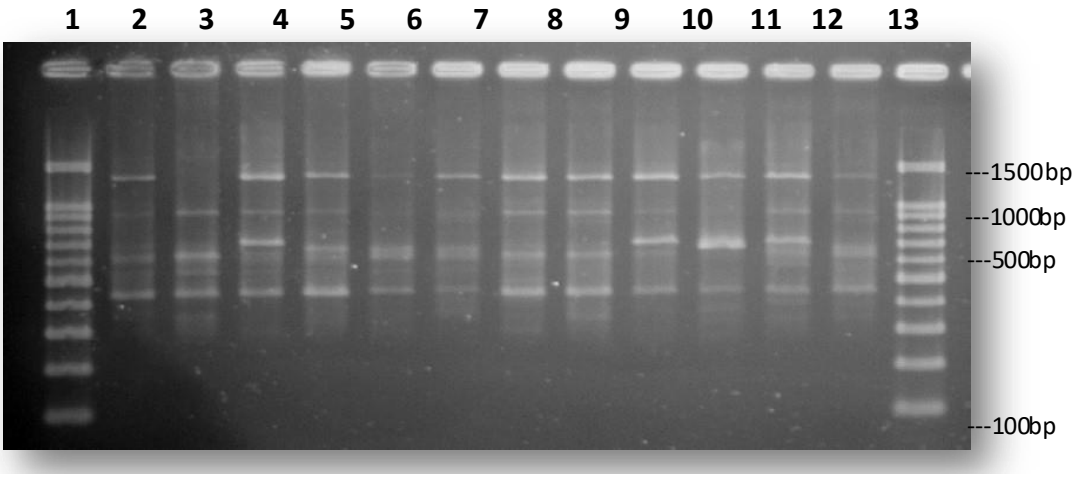

Figure 16. Amplification with primer ISSR_4. Lanes: 1) Marker, 2) I-DA/102, 3) Alexis, 4) Beta Ketsoras, 5) Emon, 6) Kamenitsa, 7) Caravela, 8) Kaskadyor, 9) Obzor, 10) Neda, 11) Gorast, 12) Orpheus, 13) Kifi, 14) Marker.

Analysis of the identified marker polymorphisms (Table 24) reveals that they vary considerably. While markers such as ISSR_10 do not show polymorphism in the studied population, 10 polymorphisms could be identified in the reactions performed with primer ISSR_5.

The example presented in Figure 16 demonstrates a significant number of polymorphic fragments obtained from performing PCR reactions with primer ISSR_4. This, together with other similar results obtained with other polymorphic markers, allowed the reliable differentiation of closely related barley genotypes, which was part of the objectives of the present study.

Table 24. ISSR polymorphisms revealed using the tested primers.

Primer	DNAsequence (3'-5')	Total number fragments	Number of polymorphic fragments
ISSR_4	(AG) 8C+TC	11	8
ISSR_6	(AC) 8C+TG	9	5
ISSR_7	(AG) 8C+TG	13	7
ISSR_8	(AC) 8C+TT	11	6
ISSR_10	(GA) 8T	4	-
ISSR_5	RY (GACA) 3	10	10
ISSR_6a	(CT) 8+RG	9	9
ISSR_7a	(CT) 8+RC	7	-
ISSR_11	(GA) 8A	9	-
	Total:	83.00	45.00
	Average:	9.22	7.50
			Polymorphisms: 54.22%

Using established polymorphic markers, a grouping of the studied genotypes was subsequently performed, which can be used to increase the efficiency of selection in this crop.

Pepper

Individual plants from each sample were examined to confirm the ability of the selected marker system to detect a sufficient number of polymorphisms in each of them. After an initial screening of the primers, it was found that some of them did not lead to the production of amplified fragments. This is most likely due to the absence or small number of the corresponding microsatellite sequence, flanked by the corresponding selective nucleotides, in the genome of the tested samples.

The reactions with the selected primers resulted in the production of one or more polymorphic fragments per sample. Some samples produced fewer polymorphic bands that were not sufficient for the identification of each plant from the respective sample using only one primer (Figure 17). Such results were expected due to the low levels of genetic diversity in a self-pollinating plant such as pepper.

Determining whether individual fragments are specific to each genotype (sample) requires the use of pooled samples. On the other hand, the use of pooled samples would hide the level of heterogeneity in the genotypes, which would reduce the reproducibility of the results if the sample contains a heterogeneous group of individuals. Since the probability of such heterogeneity within each genotype is not negligible, as shown by the studies of (Cooke et al., 2003) and (Bredemeijer et al., 2002), and we did not have prior information about the genotypes we analyzed in this study, we resorted to comparing the profiles of only individual plants from each sample.

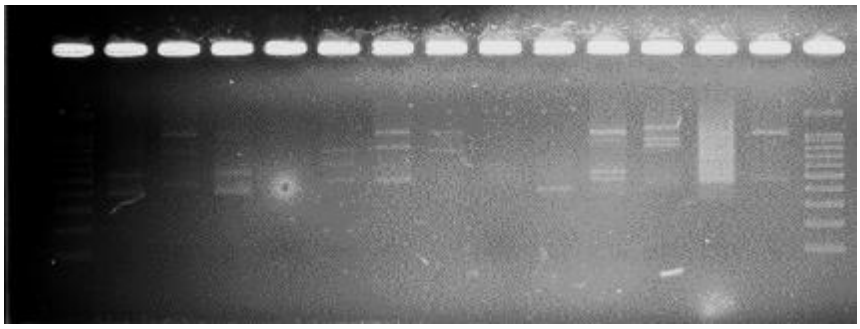

Figure 17. Results of conducting ISSR reactions with primer E8 on the thirteen samples used in the study.

Tobacco

Screening of 4 ISSR primers revealed a variable number of bands generated by different primers. The total number of bands varied between 13 for ISSR-

E7 and 4 for primer ISSR-E2. Similarly, the number of polymorphic bands was also variable. No correlation was observed between the total number of bands and the number of polymorphic bands.
Further testing revealed that plants with different DNA profiles could be found within individual cultivars. Even more interestingly, the use of primer E6 revealed polymorphisms in the Burley 1317 cultivar (indicating that two of the plants were distinct from the others), while producing essentially identical profiles for the four varieties (Figure 18 and Figure 19). The highly similar profiles for these two varieties from different production types obtained using primer E6 indicate that the DNA fingerprinting technique may pose a particular challenge in this crop.

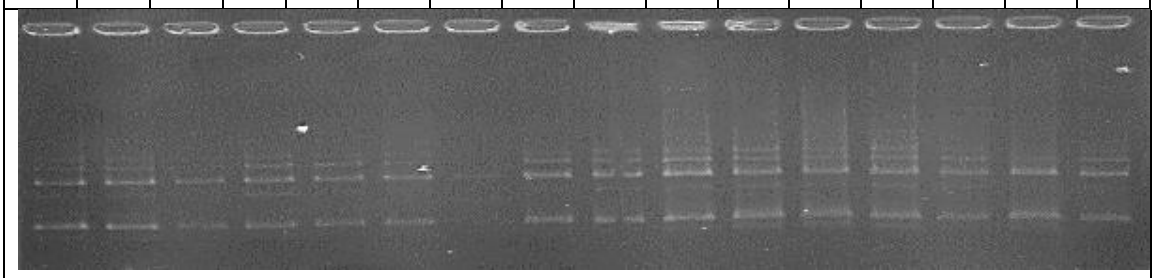

Figure 18. Polymorphisms observed when testing varieties Burley 1317 and Nevrokop 1146 with primer ISSR-E6. Each variety is represented by eight plants.

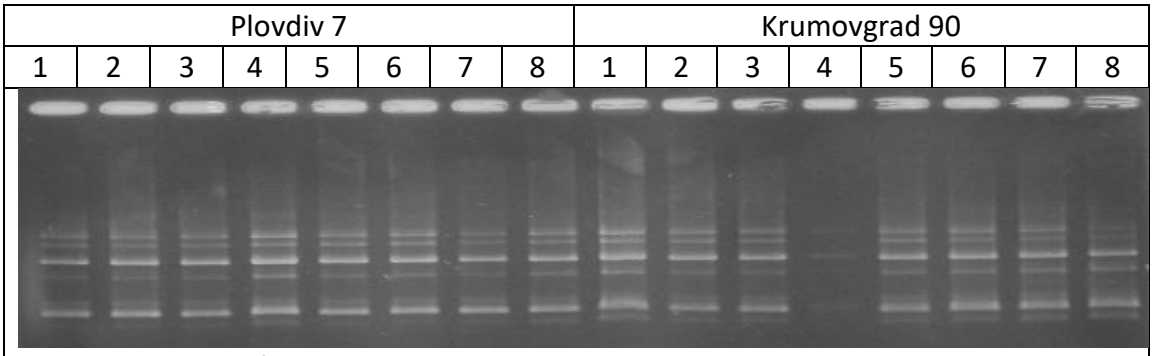

Figure 19. Testing of varieties Plovdiv 7 and Krumovgrad 90 with primer ISSR-E6. Each variety is represented by eight plants.

Corn

The conducted studies of the selected group of hybrids with a set of ISSR markers showed that in this crop too the system allows for effective differentiation of individual genotypes. The use of this type of markers allows for easy differentiation of individual hybrids with only a small set (1-2) of reactions. Moreover, as a result of the use of some of the primers, plants with different profiles within individual hybrids were identified.

And although in general the maize hybrids showed high levels of homogeneity (with rare exceptions of single deviant plants), in one of them a particularly significant diversity was found among the group of 7 studied individuals (Figure 20).

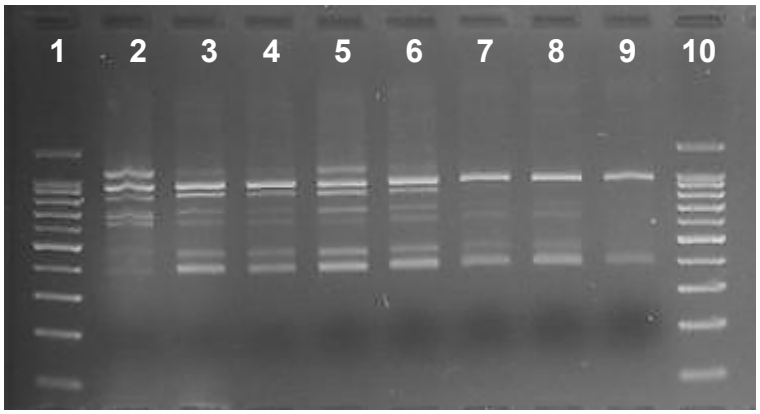

Figure 20. Genetic diversity detected within the most heterogeneous maize hybrid. Lanes 1 and 10 – DNA standard. Lanes 3-9 – DNA profiles of 7 plants from one hybrid. Lane 2 – plant from another hybrid (positive control).

Phytophthora

As a result of testing all 10 primers, it was found that primers E6, E7 and E10 failed to produce polymorphic fragments when conducting reactions with the 22 isolates of Phytophtora sp. tested. After conducting PCR reactions with these primers, gel electrophoresis did not show satisfactory visual results and therefore they cannot be commented on.

When analyzing the electrophoretic data using primer E2, it was found that runs 1 and 2, which have the same genotype, give different profiles. This difference can be explained by the poor primer performance when using DNA from the f2 genotype. Considering the above information, we noted the similarity between f1 and f4 (with common bands of 200 bp, 300 bp, 400 bp, 900 bp and 1100 bp). This similarity was not expected, since the two isolates are made from pathogens found in completely different hosts. What they have in common is the region. Atypical similarities were also found between runs 6 and 8, which contain *Phytophthora genomes* parasitizing different hosts.

Similar results were obtained with the other primers, which we will not comment on here for the sake of brevity.

Fusarium

The results of the initial screening of the primer set used demonstrated the potential of the selected ISSR marker system to differentiate genotypes within the genus.

When examining the genotypes with primer E2 (Figure 21), it was found that the similarity between runs 2, 3, 4 and 6 observed when using the previous

primer was not observed in the reaction using primer E2. Performing the reaction with this primer showed that runs 6, 7 and 8 have an identical distribution of fragments with three clearly distinguishable bands at 900bp, 1100bp and 1200bp.
Although testing the studied *Fusarium* strains with primer E3 yielded between 5 and 12 clearly distinguishable fragments, the use of this primer was not able to produce a unique profile for each of the genotypes studied.
The use of primer E3 showed that runs 7 and 9 had identical fragment distribution with four clearly distinguishable bands at 300bp, 500bp, 1000bp, and 1200bp. On the other hand, preliminary data on the origin of these isolates was that the two runs did not contain the same material.

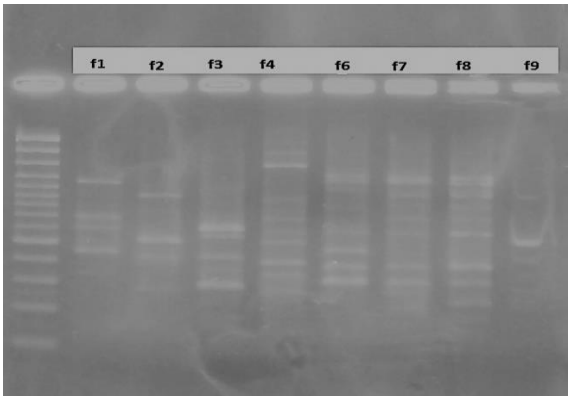

Figure 21. Results of testing the studied Fusarium strains with primer E2.

Goats

The initial screening with pre-selected ISSR markers (Figure 22) aimed to verify the capacity of the selected marker system to reveal sufficient levels of polymorphism within the two breeds. As shown in the figure, the use of the selected ISSR primers produced a number of polymorphic bands between individual animals. Therefore, they were used to screen both populations in an attempt to reveal genetic diversity both within and between the two local goat breeds.
As expected, the use of different ISSR primers resulted in the detection of different numbers of polymorphic bands in the two breeds. As a result of screening several primers and optimizing PCR conditions, 2 ISSR primers were identified that produced informative polymorphisms in these local breeds.

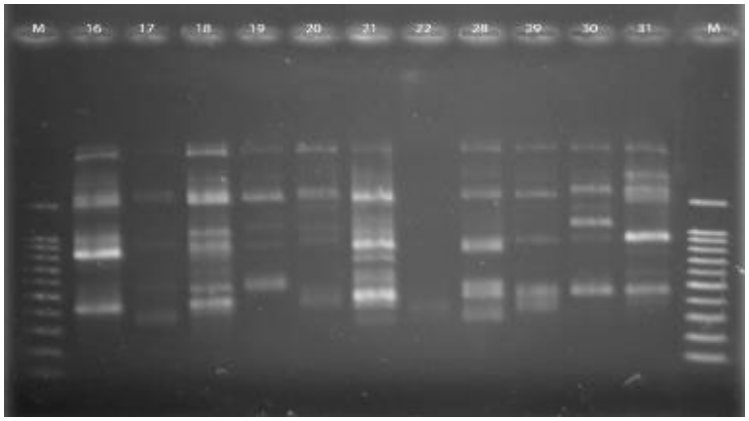

Figure 22. Polymorphisms in Kalofer Longhair breed revealed with primer E7. Line M – 1kb DNA standard. Numbered lanes – PCR products from different animals.

Using dominant markers for QTL mapping

The complex nature of inheritance and the requirement to combine several different (and often negatively correlated) traits make breeding varieties with good technological performance, carrying multiple improved fiber characteristics, a difficult task that every cotton breeder strives to solve. Marker-assisted selection (MAS) has been a favored tool in recent years, as it greatly facilitates the selection of genotypes for crossing in each generation. As more and more alleles with a positive effect on interesting traits are identified and molecularly tagged, breeders are increasingly using this toolkit in their real-world programs.

One of the key steps in identifying suitable markers for use in MAS is linking markers to the traits of interest. Several such markers were identified in the present study. Furthermore, mapping them to a genetic map provided further insight into what might actually lead to modification of the traits of interest.

For the analysis of molecular markers associated with qualitative and quantitative traits not mapped during the development of the unified map, a total of 3 SSR and 13 ISSR primer pairs were screened for their ability to reveal polymorphisms.

The ISSR primers were able to produce a large number of markers. Screening the population with the 13 primers resulted in large differences in the production of polymorphic bands, where 3 of them did not reveal any polymorphisms, while the remaining 10 gave between 2 and 8 polymorphic bands. In total, the ISSR primers produced 41 polymorphic bands in a population of 126 individuals. Of these, only 35 could be mapped, as some had only one occurrence in the studied population (Table 25).

Table 25. Linkage groups of ISSR markers.

Linkage groups from all loci													
Group	**Number of loci**	**ISSR marker no.**											
Unlinked:	*11*	2	6	7	9	22	24	25	26	31	33	34	
group 1:	*12*	1	4	5	10	14	16	17	20	21	27	30	32
group 2:	*10*	3	8	13	15	18	19	23	28	29	35		
group 3:	*2*	11	12										

* LOD min : 3.00; r max : 0.10; cM max : 10.1

Of the 27 polymorphic bands identified in the present study, only one (designated CM14-300) had a significant effect on fiber length. This locus was associated with LG1 and had a statistically significant effect, explaining 6.6% of the total variation in the trait. Fiber strength was significantly influenced by two different loci – one (CM8-100) on LG3, explaining 4.7% of the variability in the trait, and the other on LG4 (CM7-95), explaining 7.7%. Uniformity, on the other hand, was influenced by two different regions – one included markers CM1-290 and CM3-100 (spanning 7.2 cM on LG1 and explaining 18% of the total variation), while the other contained a single marker (CM7-100) responsible for 9.8% of the variation and located on LG3. Elongation is influenced by a single region on LG3 that explains 14.3% of the variation in the trait. The region includes markers CM13-95, CM14-95 and CM15-95 and spans 17.8 cM of this linkage group.

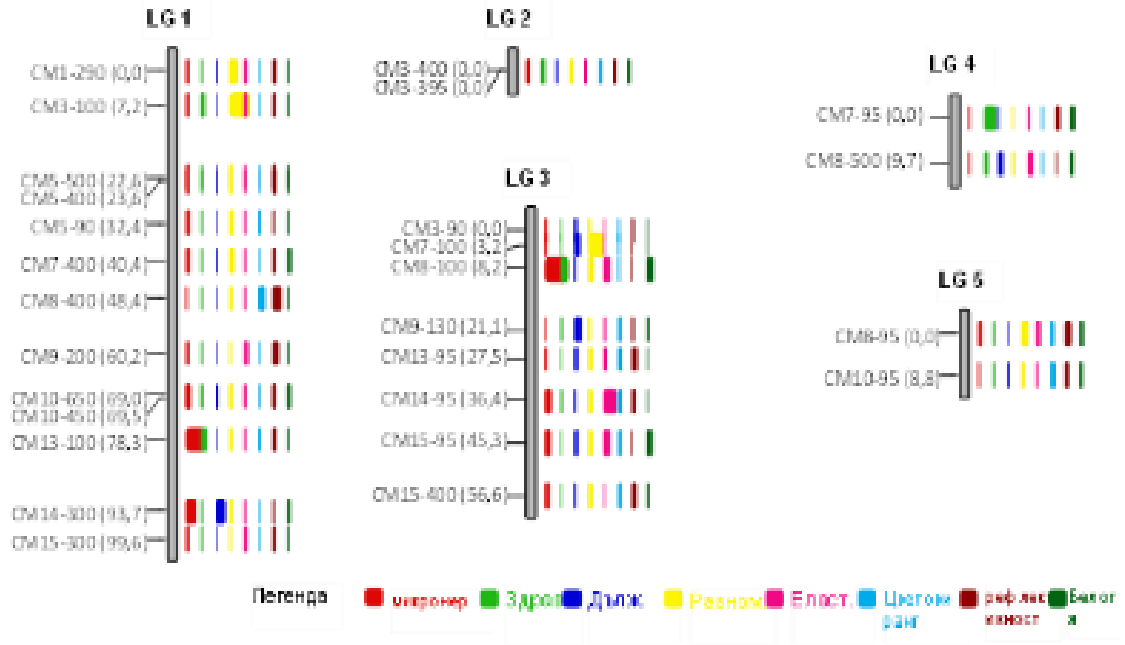

Figure 23. Combined genetic map for all studied traits in the mutant segregating population. The width of the colored bars represents the percentage of variation in the trait explained by the presence of a marker at the corresponding locus.

Although providing some insight into the effects of different loci on the studied fiber traits, association mapping for single traits could not provide evidence to support the correlation (whether positive or negative) between different traits that has been observed in many classical selection studies. Combining maps for all traits (Figure 23) provided such evidence, thus supporting at least some of the previously well-known correlations. In our study, the upper part of LG3 contained a region spanning 5 cM (defined by markers CM7-100 and CM8-100) that had a statistically significant effect on fiber strength, micronaire and elongation. This finding is consistent with the observations that micronaire, as a measure of both fiber fineness and fiber maturity, is affected by the accumulation of cellulose fibrils (which improves strength but reduces elongation). Therefore, the mutagenic treatment applied at the beginning of the present study may have affected some fundamental processes of cellulose accumulation, which are determined by the locus(es) marked by the aforementioned markers. Further studies are needed to establish which of the many possible mechanisms (related to either cellulose synthesis or deposition) are affected, but the simultaneous modification of all three traits strongly supports this hypothesis.

Discussion

SSR markers

All of the above shows that the applicability of microsatellite markers can be sought mainly in the direction of their use for breeding purposes. Obviously, this type of markers are best applicable in the context of the specific populations for which they were developed. This means that the transferability between populations developed on a significantly different genetic basis would be very low and, accordingly, their use should be limited mainly to the specific context of the breeding program for which they were originally developed.

The use of microsatellite markers in breeding programs leads to accelerated selection of specific alleles, without, however, monitoring the effect of this selection pressure on other alleles in parallel. This means that in such cases, one can expect the absence of selection pressure on alleles that have been identified as having an effect in a different genetic context. The expected return to the normal distribution of these alleles in such cases may make them more applicable for the purposes of RHC testing. For this purpose, an additional check of the normality of their distribution should be performed (which by definition also includes a check for the elimination of some of the alleles, i.e. for those that have passed into a monomorphic state).

CAPS markers

The use of molecular markers for genotyping the Tm-2 locus in tomatoes began in the early 1990s, after it was mapped by (Tanksley et al., 1992). Based on this information, a number of authors described molecular markers associated with the locus, but despite the positive results obtained from them, when analyzing other gene plasma, it was found that the application of most of them was unsuccessful (Panthee et al., 2013). According to the same authors, this may be due to the location of the markers in the region around the Tm-2 locus, and in the event of possible recombination, the link between the gene and the marker may be lost.

Therefore, for the purposes of our study, markers were selected, prepared for the Tm-2 locus by (Shi et al., 2011) based on sequences with numbers AF536199, AF536200 and AF536201, published in NCBI by (Lanfermeijer et al., 2003). The results obtained from the application of the CAPS marker to identify the Tm-2 and Tm-22 alleles correspond to those indicated by (Shi et al., 2011). These results were also confirmed by *in-silico* analysis of the fragment enclosed between the primers for the CAPS marker, i.e. the locations of the restriction sites of the used endonucleases were confirmed. The products obtained after the restriction were distinct and the determination of the genotype of the plants was easy, which is important for the routine application of the marker (Figure 9 and Figure 10). The results also showed compliance with the biological and serological tests carried out, which is expected taking into account the fact that the amplified fragment is located in the resistance genes themselves. However, we consider the practical applicability of this type of markers to be relatively limited due to the requirement to conduct PCR and restriction reactions sequentially. This simultaneously increases the requirements for the qualification of the personnel and complicates and increases the cost of conducting the analyses.

ISSR markers

Cotton

The application of the ISSR marker system to a disintegrating mutant population (M6) showed that it is able to identify a sufficient set of differences even in such a group of closely related individuals (Figure 8). When using this marker system to characterize a mutant segregating population (M6), good differentiation of individuals was achieved with only a limited set of primers. As mentioned above, as a result of using only 4 primers, a total of 21 polymorphic loci were identified, which made it possible to obtain unique profiles for the majority of individuals. Considering that the studied samples are of highly closely related origin (induced mutagenesis in a purebred variety, the

homogeneity of which was checked before subjecting it to mutagenic effects), we believe that the results presented here provide a clear example of the high efficiency of ISSR DNA profiling for selection purposes.
On the other hand, within the framework of the present study, we also demonstrated the potential for using the ISSR marker system for the differentiation of varieties and breeding lines (Figure 7), which would be useful in PCR testing. As shown, the system allows for the simultaneous differentiation of variety and line, along with the verification of the homogeneity of individuals of both genotypes.

Tomatoes

In the analysis of tomato cultivars and lines, differences in the DNA profiles obtained were observed between plants belonging to the same cultivar/line and which appeared identical in terms of their phenotypic characteristics. One possible explanation for the observed discrepancy between phenotypic homogeneity and identified genotypic heterogeneity is that the heterozygous loci are not linked to the traits for which the genotypes were selected during the breeding process. On the other hand, (Bredemeijer et al., 2002) suggested that such results could be explained by the presence of residual heterogeneity mainly in non-coding sequences and/or in quantitative trait loci (QTL) with very small phenotypic effect. Eliminating such residual heterogeneity would be impossible with classical breeding approaches. It is questionable whether attempting to eliminate such heterogeneity using molecular markers is of any practical benefit. This is due not so much to the lack of phenotypic effects from it, as these polymorphisms may be associated with other characteristics not currently accounted for in the breeding program. A main criterion for the practical applicability of efforts to eliminate such residual heterogeneity should be to verify whether it has a significant effect on other traits – either the biochemical composition of plants or their individual parts, or effects on architectonics or other properties of the crop that could be of practical interest from the point of view of growing technology.
The demonstrated capacity of ISSRs to identify the presence of such heterogeneity in the plant materials studied provides striking evidence in support of the careful selection of the marker system. A marker system that is meaningful to be introduced into breeding programs and formal PCR testing must be reliable, efficient, rapid and easy to apply. In cases where a plant deviates from the group of 7 of the same genotype, removal of that plant can easily lead to elimination of the heterogeneity from the selected genotype. However, in cases where more than one plant appears with a different profile, the selection of a representative set of bands for a particular genotype can

be difficult, especially when the plants appear to be divided into two more or less equal groups. In such cases, the question arises as to what should be considered a “significant deviation“ and whether a representative profile can be created for this particular genotype so that it can be distinguished from other genotypes in PCR testing.

An additional aspect that the demonstrated capacity of ISSR markers to identify the presence of such heterogeneity in the studied material presents is the possibility of a properly selected marker system to serve both the purposes of PCR testing and the needs of the selection process. As already commented, the requirements for the respective marker systems in the two directions are significantly different (UPOV, 2021). Above, we demonstrated that the ISSR marker system we chose can be introduced into breeding programs because it is reliable, effective, fast and easy to apply. For example, the application of ISSR in our study demonstrates applicability for increasing the efficiency of the selection process in the selection of parents and F1 individuals. In the case of tomatoes, the use of the ISSR marker system led to the identification of F1 plant No. 3 (Figure 24, lane 5) as a non-hybrid individual. The conclusion that this is more likely an individual originating from selfing of the P1 parent is based on the presence in its profile of a 510 bp band that is present in the P1 parent and not present in the P2 parent. Furthermore, the same individual lacks the 590, 800 and 1350 bp bands that are present in P2 but not in P1. Our results showed that DNA profiling of the F1 hybrid plants, allowing identification of the non-hybrid individual and confirmation of the hybrid nature of the other individuals, can be achieved with the use of a single ISSR primer (ISSR 3 in the case presented in Figure 24).

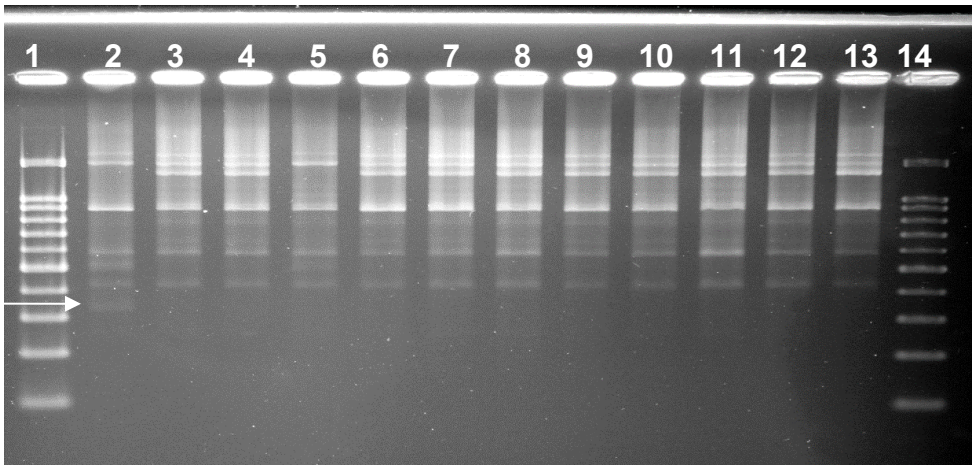

Figure 24. Results of testing parents and F1 hybrid tomato plants with primer ISSR 3. Genotypes are arranged on the gel as follows: Lane 2 – Parent 1 (PK), Lanes 3-12 – F1 hybrid plants. Lane 13 – Parent 2 (L 1116). Lanes 1 and 14 – standard-sized DNA. Arrows indicate bands that are polymorphic between parents.

The ISSR marker system is well suited for use in breeding programs because it requires a small amount of initial genomic DNA, uses a simple application protocol, and gives reliable results. (Penner et al., 1993; Rajput et al., 2006). The results of our study showed the capacity of the selected marker system to differentiate not only the parental lines but also some of the individual plants within the parental lines. Furthermore, the application of the ISSR system allowed the identification of deviant plants from the F1 hybrid population, which deviation was later confirmed phenotypically.

Barley

Combining data from different types of analyses was able to provide a significantly more complete and accurate picture of the genetic distance between the genotypes studied. This in turn has the potential to significantly facilitate the selection of parental forms from this working collection for future breeding programs.

Tobacco

As a first step towards the practical application of this technique in Bulgarian tobaccos, we presented here a study of intra- and intervarietal diversity in local forms, with an emphasis on practical applicability.

When developing a DNA “fingerprinting” technique in genetic diversity research, one consideration that must be taken into account is determining levels of within-variety polymorphism such as that observed in the present study.

This is important not only for the reliability of variety identification, but also for determining the significance of the differences in terms of distinctness (attribution of the genotype tested to an existing variety or granting the status of a new variety). Extensive experimentation should be carried out by each variety testing authority before DNA fingerprinting is fully adopted as a standard method of use. Assuming that (for practical reasons) no more than 4 ISSR primer pairs will be used in such distinctness and homogeneity testing, this will allow keeping the costs of determining polymorphism within the proposed new sample low in each government authority. The reliability and reproducibility of microsatellite markers (SSR) and amplified fragment length polymorphisms (AFLP) make these two marker systems preferred in many different plant species (Chen et al., 2015; Chen et al., 2000; Faccioli et al., 1999; Garg et al., 2001). Our previous comparative study on the potential of SSR and AFLP for DNA fingerprinting of cotton (Bojinov and Lacape, 2003) outlined the respective advantages of these two types of markers depending on the size of the population, its diversity and the availability of the necessary

equipment and licenses. As presented in the present study, AFLP would currently be the technique of choice when DNA fingerprinting of large populations is required. In all other cases, however, our study shows that ISSR is the recommended system to use.
Our data show that a limited set of ISSR markers would be sufficient for variety identification and adequate grouping of tobacco lines and cultivars according to the purpose of the respective study. Furthermore, in the future, analysis of additional tobacco lines using an expanded set of ISSR markers will be able to provide a much more detailed picture of the genetic relatedness of tobacco lines within and between the individual germplasm groups used in Bulgarian breeding programs.

Wheat

In this culture, we studied 5 ISSR primers, which produced a total of 55 polymorphic bands. Of these, the smallest number (5 polymorphic bands) produced primer T11, and the largest (11 polymorphic bands) – primer T14. At the same time, the markers thus obtained demonstrated significantly different levels of PIC (0.165-0.496; average 0.335). As was commented in the Material and methods section, low PIC values are obtained in cases where we have only a single present/absent band in the profile of one of the individuals of the studied population. Such bands can be highly informative regarding specific genetic determinants and thus be useful for differentiating the genotype from the others in the group. Such markers usually represent a specific selection interest, therefore they are mainly used in this direction. On the other hand, markers with high PIC (highly polymorphic markers in the studied population) are useful in DNA profiling of genotypes for the needs of PCR testing.

Corn

The use of the ISSR marker system allowed the differentiation of the set of maize hybrids used in our study. The use of only one primer is able to lead to the identification of a significant number of polymorphisms, which is a necessary condition for increasing the efficiency of the system. Despite the limited number of primers that were tested, the successful obtaining of unique profiles for a number of hybrids, as well as the demonstrated ability of the system to identify deviant individuals within the hybrids, indicate its applicability for the needs of PCR testing. In the case of this crop, no studies were conducted to associate ISSR markers with qualitative and/or quantitative traits, since the AU lacks the necessary capacity such as laboratory equipment and technological complex for phenotyping.

Pepper

Until the present study, no application of MAS tools with practical applicability for Bulgarian breeding programs has been demonstrated, as those developed abroad are based on foreign germplasm, and studies in the country are in their early stages. The only known study (Tsonev et al., 2017) aims to group 19 pepper (*Capsicum) cultivars annuum* L.) from the collection of the Maritsa Institute of Vegetable Crops, Plovdiv, Bulgaria in clusters according to their relative genetic distances, estimated by technological characters and 9 di- and tri-nucleotide ISSR markers, and to assess the relationship between them. Although this represents an extremely important first step in the selection of parental forms for hybridization, there is no data available that the study was continued with the analysis of F1 hybrids and/or decaying populations from initiated crosses.

The genotypic profiling of a series of samples conducted by us using the two ISSR primers (E5 and E8) was used to cluster them according to their relative genetic distance.

The only known study of Bulgarian pepper samples (Tsonev et al., 2017) found that with the 7 ISSR primers they used, they failed to obtain unique profiles for each sample from the collection of 19 varieties. Moreover, an entire group of 5 samples turned out to be completely indistinguishable. All this demonstrates the importance of the correct selection of ISSR primers for DNA profiling of the Bulgarian gene pool in this crop.

Phytophthora

As a result of the analysis, a set of primers based on the ISSR marker system was identified and have the potential to establish the presence of sufficiently informative polymorphisms between the tested *Phytophthora genotypes.* sp. The most interesting results were obtained using primer E2, where the largest number of polymorphisms in genotypes within the genus were identified. Performing a PCR reaction using the remaining primers was able to identify sufficient levels of polymorphism.

Polymorphisms, besides genotypes from different phenotypic groups, were also identified in pathogens with similar or identical phenotypes. Such is the example with primer E2, where the most atypical profiles were observed. Our results showed that phenotypically identical genotypes gave atypical and different profiles (f1- f2 and f15-f16), and others, which are phenotypically different, were grouped as identical (f1 - f4; f6 – f8; f9 – f10 and f20 – f21 – f22.) In subsequent studies, it would be interesting to establish the reasons for these manifestations.

The presented results show that using ISSR markers can be used to develop a system for the identification of genotypes of *Phytophthora.* This demonstrates the potential of the proposed method for revealing differences even within the same phenotypic groups and analyzing genetic diversity in this phytopathogenic genus.

Fusarium

As a summary of the results obtained in this work, we can note that the use of the ISSR marker system for the identification of *Fusarium genotypes* can be done using at least 2 primers. The primer pairs that are able to produce unique profiles for all studied genotypes are primer E1 and primer E3, primer E2 and primer E3, primer E3 and primer E4, primer E3 and primer E6.
The presence of such a large number of combinations indicates that the selected system has sufficiently large potential for the identification of a significant number of polymorphisms to justify the search for new primers that would enable the identification of all genotypes with a single reaction.

Goats

The initial screening with two ISSR primers was able to demonstrate the capacity of the chosen system to identify informative polymorphisms in both breeds of long-haired goats. When using only two ISSR primers, the differentiation of individuals from the two breeds was incomplete, as in both breeds this was illustrated by the presence of groups of practically genetically indistinguishable individuals. Although expected for such breeds of relatively recent origin, these results did not meet the requirement for adequate differentiation of each of the animals studied.
After preliminary screening of several more ISSR primers for the produced levels of polymorphism, a total of 4 ISSR primers were included in the study, which was sufficient to obtain a grouping of individuals that well corresponds to their presumed kinship. The use of a set of 4 ISSR primers led to obtaining completely satisfactory results in terms of grouping of individuals. Thus, the application of the ISSR marker system demonstrates the necessary potential for the detection of polymorphisms that produce informative bands, with the help of which to achieve a satisfactory differentiation of animals according to the degree of their kinship. This is an important step towards the development of an effective system for genotyping animals of both breeds in order to expand and supplement our knowledge about the degree of their kinship and differentiation.

Applicability of ISSR markers in PCR testing

With the examples presented in this study, based on organisms from different groups (plant and animal), we consider the proposal for the use of the ISSR marker system as standard in RHS testing in our country to be justified. Examples and analyses of the system's compliance with the UPOV set of requirements for its application were presented in detail. (UPOV, 2021). Below we will comment on the compliance of the specific parameters demonstrated of the system with the UPOV set of criteria.

According to the UPOV document, important considerations for choosing DNA profiling methods that generate high-quality molecular data are:

1. *Reproducibility of data acquisition within and between laboratories and detection platforms (different types of equipment).*

As we have shown with the presented literature data and with our data and analyses, obtaining data for ISSR profiling of samples does not depend on the equipment used. The 3 different types of PCR apparatuses used by us (2 gradient and 1 regular) were repeatedly comparatively tested, demonstrating high reproducibility of the results between repetitions. At the same time, a number of researchers have also demonstrated their high reproducibility between different laboratories, which fully satisfies this first criterion.

2. *Repeatability over time.*

The studies presented in this paper have been conducted for more than 15 years, with the comparability of the resulting profiles being confirmed by the periodic use of reference DNA samples isolated back in time. This type of verification has been mainly carried out with cotton and tomato samples, but sufficient information is also available for barley, wheat and tobacco.

3. *Discriminatory capacity*

In the framework of this study, a number of examples of the high (if necessary) discriminatory capacity of the ISSR system were presented. Examples of discrimination between wild and cultivated species, between endemic varieties and breeds, between modern varieties and lines, between mutant lines originating from a highly homogeneous line, within breeding lines, as well as between parents and their F1 generation were presented. In this way, it was demonstrated that the use of this system can achieve the necessary level of differentiation depending on the aims of the respective study - be it breeding or for RHS testing.

4. *Time and labor intensity*

Due to its multilocus nature, the ISSR system produces a significant number of polymorphic fragments within a single reaction. Our research has shown that in the context of Bulgarian breeding materials, the correct selection of primers can lead to the production of 10 or more polymorphic fragments in

a single reaction, which makes the time and labor required to obtain sufficient information quite short. Our personal experience shows that from the receipt of plant/animal samples in the laboratory to the receipt of the results, a period of 1 to a maximum of 2 days is required (with a number of samples – up to 20). Naturally, with an increase in the number of samples, the required time increases, but this is mainly at the expense of the time for DNA extraction, while the actual conduct and documentation of the ISSR reactions becomes significantly faster. Moreover, the equipment used allows simultaneous analysis of up to 96 samples.

5. *Stability of performance over time and conditions (sensitivity to subtle changes in protocol or conditions)*

As already mentioned above, here we demonstrated the high stability of the ISSR system over time. Moreover, since a number of PhD and graduate students participated in the research, it is obvious that their inevitable subtle changes in the protocol and conditions do not affect the final result.

6. *Flexibility of the method, possibility of varying the number of samples and/or the number of markers*

Within the framework of the variations presented in this work regarding the organisms used, the number of analyzed samples, and the number of polymorphic fragments studied in the different plant and animal forms, both the indifference of the ISSR marker system to the origin of the studied samples and the possibilities for rapid and effective increase in the number of studied samples and obtained markers according to the needs and aims of the respective analysis were demonstrated.

7. *The interpretation of the data obtained is independent of the equipment*

Since ISSR reactions produce clear bands with binary appearance (presence/absence), the interpretation of the obtained results is quite easy and unambiguous. As demonstrated by the results presented above, the identification of polymorphic bands can be easily done even by people with minimal knowledge and specialized training. The results presented by us were obtained using 3 different PCR apparatuses, several different extraction kits, different combinations of power supply units and electrophoresis baths of varying sizes. Despite all these variations in the equipment used, the data obtained allow easy comparability and, accordingly, inclusion in summary analyses.

8. *Database Sustainability*

This criterion has low applicability to the ISSR marker system, since obtaining this type of marker can be done without the prior existence of databases. In turn, the obtained results can be aggregated into databases with an

extremely simplified structure, which in turn makes them resistant to degradation even with repeated and multi-user use/editing/filling.

9. *Accessibility of the methodology*

The presented methodology is extremely easy to implement and requires a very minimal set of molecular biology equipment (PCR apparatus, electrophoresis bath, centrifuge, transilluminator) and skills from the staff. This is also confirmed by the participation of dozens of graduate students who have passed through the AU laboratory and been trained in its use in the development and defense of their thesis based on the diverse applications of ISSR markers.

10. *Independence from a specific machine, specific chemistry, specific supplier, specific partners or products*

The fulfillment of this criterion was already somewhat demonstrated in the analysis of the fulfillment of criteria 1 and 2. We will recall that we used 3 different types of PCR apparatus, several different DNA extraction kits, various sets of chemicals for PCR reactions, etc., whereby the results obtained were completely comparable, regardless of who the suppliers of the relevant products were;

11. *Suitability for automation*

Due to their dominant nature and significant amounts of the products obtained, ISSR reactions produce clear bands that are both easy to read by operators. At the same time, this makes them easy to automate, since in automated result reporting systems it is imperative to set a threshold value for the strength of the received signal. When for the operators themselves the determination of each signal has only two options (presence/absence) and, when the presence is easy to establish, the determination of these threshold values becomes easy. At the same time, the necessary periodic verification/validation of the automatic collection of results is alleviated.

11. *Suitability for multiplexing*

By their very nature, ISSR reactions are not particularly suitable for multiplexing. This is mainly due to their multilocus nature, where the use of only one primer results in a large number of products (from 4 to 18, observed in our experiments). On the other hand, the very fact of obtaining such a high number of products is a result of amplifying a significant number of loci simultaneously, which is the main purpose of multiplexing reactions when using monolocus markers.

12. *Cost-effective (costs, number of samples and number of markers are balanced)*

As demonstrated in the significant number of crops and in the two autochthonous animal breeds used in the present study, obtaining a sufficient

amount of information (depending on the objectives of the analysis) when using the ISSR marker system can be done with the help of a very small number of reactions. Since the system uses the PCR principle, the amount of DNA required is very small and can be obtained from 50-80 mg of plant tissue or a few hair bulbs. This makes the DNA extraction itself relatively cheap (in contrast to RFLP analysis), which leads to a lower overall cost of the ISSR analysis. Somewhat similar to AFLP analysis, when using the ISSR marker system, the inclusion of only one primer (one primer pair in AFLP) is often sufficient to obtain a sufficiently good characterization of the genetic diversity in the studied set of genotypes. Unlike AFLP analyses, however, the ISSR reactions do not require double nested PCR reactions, which again makes the cost of the ISSR analysis lower. Even in cases where more than one reaction is required, their number rarely exceeds 3-4, which is common for conducting SSR analyses. Very often, this is the reason for looking for multiplexing options in order to reduce the final number of reactions performed. This, in addition to reducing the consumables used, also significantly reduces the time required to perform the analyses, which is the main reason why the requirement for multiplexing is considered essential for this type of marker system.

All of the above, together with the low requirements for the equipment and the qualification of the personnel who can perform this range of analyses, makes the ISSR marker system significantly more efficient than the SSR, AFLP and RFLP markers used for comparison in the present study. This efficiency is with regard to both the PCR testing procedures and their use for selection purposes, which once again emphasizes the flexibility of the system with regard to its application in different research contexts.

Unlike ISSR markers, the results obtained from the tested locus-specific markers confirm the thesis expressed by other authors that a significant part of the published markers are unsuitable for immediate use in diverse germplasm. Additional steps are often required for their successful application, partly due to reasons such as loss of complementary sites of the primers due to point mutations, interruption of the connection between them and the searched gene due to recombination or differences in PCR conditions. Although tomato is a typically self-pollinating species, some residual heterogeneity is observed even in PCR-stabilized lines (Bredemeijer et al., 2002; Cooke et al., 2003). The probability of point mutations that do not disrupt gene function is not negligible, and if such a mutation occurs in a region of the gene related to the marker used, its applicability may not be complete, even if it is created according to the sequence of the corresponding gene.

Conclusions

The present study presents the differences in the application of several types of molecular markers. In particular, it demonstrates the ability of ISSR as a tool for identifying varieties in a number of crops and two autochthonous goat breeds. Since the ISSR system has been proposed for a long time but has not been sufficiently studied in the Bulgarian context, here the capacity of a set of primers to produce polymorphism in the main local crops and breeds for the country is assessed. However, further fine-tuning of the technique will be necessary before it can be applied to its full capacity – in PCR testing, protection of breeders' rights, sorting of breeders' collections and gene banks from redundant records, development of genotype “identity cards”, etc. The main conclusions in the presented study can be summarized as:

1. Until this study, there are no MAS tools available with demonstrated practical applicability for Bulgarian selection and breeding programs, as those developed outside the country are based on foreign germplasm, and studies in the country are in their infancy. Using germplasm that is most widely available for the respective target crop/breed in the country, this study provides knowledge and tools necessary to maintain the competitiveness of Bulgarian selection programs.
2. Our study confirms the applicability of the ISSR marker system in plant breeding, as we demonstrated how the use of a single primer can lead to early confirmation of the hybrid nature of F1 individuals, as well as the detection of those of non-hybrid origin. This could be of great benefit in crops such as tomato, where F1 populations are usually established by transplanting, and such early detection of deviant individuals can save expensive experimental space and reduce the need for skilled labor in the early stages of the breeding program. The effects are multiplied for cases where deviant characteristics are inherited as recessive traits and thus their earliest identification by phenotyping is only possible in the F2 generation.
3. Using some of the most important commercial varieties in the country in crops such as cotton, tomatoes, barley, maize and wheat, this study attempts to provide the knowledge and tools necessary to maintain the competitiveness of Bulgarian breeding programs in the institutes of the SSA.
4. The application of the ISSR marker system not only reveals sufficient polymorphisms between different plant and animal genotypes but is also able to identify outliers in sample collections and among F1 individuals. This demonstrates the ability of the proposed technique to

reveal differences even within supposedly highly homogeneous genotypes such as self-pollinating varieties of the species *Solanum lycopersicum.*

5. The results presented in this paper demonstrate that ISSR markers can effectively differentiate genotypes even within phenotypically highly homogeneous populations of plant, microbial, and animal organisms. This is most clearly demonstrated in the cases of cotton, tomatoes, and goats.
6. ISSR markers are capable of not only identifying diversity within the genotypes studied but also carry the potential to effectively identify individuals with different metabolic profiles within different varieties/breeds.
7. ISSRs were used effectively to differentiate different species and hybrids in Paulownia sp. Their effectiveness in the genus has been demonstrated by distinguishing *P. elongata* and *P. elongata x P. elongata* hybrid.
8. The use of ISSR molecular markers is effective for the differentiation of isolates within genera *Fusarium* and *Phytophthora.*
9. Despite the relatively small size and long history of inbreeding in the two studied indigenous goat breeds, significant diversity was revealed, as determined by molecular markers. The molecular analyses performed revealed that the ISSR marker system can identify significant genetic diversity among the indigenous breeds.
10. The ease of use and high reliability of the system make it a very good candidate for practical application in plant and animal breeding programs.
11. AFLP markers have a high capacity to discriminate individuals in breeding collections, including those working with the limited diversity of local samples available in the country. Therefore, they were effectively used to saturate the cotton genome map, where SSR and RFLP provided the initial mapping framework.
12. Based on their high reproducibility and polymorphic content, SSR markers were applied to study cotton and tomato collections and to identify loci associated with key quality traits of these crops.
13. The CAPS marker (with primers Tm2RS-f3 and Tm2RS-r3) can be successfully applied to identify the allelic status of tm-2, Tm-2 and Tm-2 2 in tomato breeding lines. Its application, however, is limited to a specific genetic context and is further limited by the requirement for higher levels of qualification from personnel in the procedure itself for its establishment.

Contributions

Theoretical contributions

1. Based on a number of parameters (polymorphic information content, uniformity of distribution along the length of the genomes, reproducibility, reliability, etc.), a molecular marker system has been identified that has the capacity to meet the requirements of various areas of use.
2. Methods have been developed for the use of the ISSR marker system in a number of crop plants, two eukaryotic phytopathogens, and two autochthonous animal breeds.
3. Within the framework of this work, for the first time in our known literature, the ability of the ISSR marker system to identify intraspecific diversity in phytophthora fungi has been demonstrated, which opens up new opportunities for the development of more effective systems for detection and identification of this group of phytopathogens.
4. The ability of a set of dominant (i.e. AFLP, ISSR, CAPS) and co-dominant (RFLP, SSR) molecular marker systems to detect genetic diversity in a range of plant and animal organisms of local origin has been characterized.
5. Through the use of comparative analysis, the advantages and disadvantages of each system have been identified depending on the planned direction of use - for selection, breeding or for the purposes of registering and maintaining the biodiversity of the relevant organisms.

Applied contributions

1. A molecular marker system (ISSR) has been identified, which has a wide range of applications in breeding programs of local forms of plants and animals, as well as for the detection and identification of phytopathogenic organisms.
2. The applicability of the ISSR marker system for analyzing genetic diversity in a large set of plant, microbial, and animal organisms has been demonstrated, thus proving its universality in terms of use in eukaryotes in such a context.
3. By performing comparative analyses, the effectiveness of the application of the ISSR marker system in the selection programs of a number of crops of national importance and in two autochthonous goat breeds has been proven.
4. The potential of the ISSR marker system for effective use for both breeding purposes and for the purposes of state variety testing,

approbation and seed control (in plants) and for the independent certification and breeding of local animal breeds has been demonstrated.

5. The compliance of the ISSR marker system with the system of criteria used by UPOV to determine the applicability of various marker systems for the purposes of PCR testing has been characterized in detail, thus proving its applicability in the IACS and IASR systems.
6. New marker/trait associations have been obtained for a number of economically valuable traits in local cotton and tomato varieties.
7. A significant number of molecular markers have been mapped in cotton, significantly enriching the existing genetic map with loci that have significant effects in local agro-ecological conditions.
8. A significant number of loci for quantitative traits have been identified in local samples, and those with pleiotropic effects on a number of qualitative and quantitative traits have also been identified.
9. The genetic diversity in two autochthonous breeds (Bulgarian Vitoroga Longhair and Kalofer Longhair) of goats at the level of hereditary material has been characterized, and recommendations have been given for the future breeding of both breeds, so as to reduce the risks of inbreeding, potentially leading to inbreeding depression.