

---

**MARCATORI MOLECOLARI DEL  
DNA PER LA TRACCIABILITA'  
NELLA FILIERA AGROALIMENTARE  
DEL POMODORO**

**Martina Caramante**

Dottorato in Scienze Biotechnologiche – XXII ciclo  
Indirizzo Biotecnologie per le Produzioni Vegetali  
Università di Napoli Federico II





---

Dottorato in Scienze Biotechologiche – XXII ciclo  
Indirizzo Biotecnologie per le Produzioni Vegetali  
Università di Napoli Federico II



---

**MARCATORI MOLECOLARI DEL  
DNA PER LA TRACCIABILITA'  
NELLA FILIERA AGROALIMENTARE  
DEL POMODORO**

---

**Martina Caramante**

Dottoranda: Martina Caramante

Relatore: Prof.ssa Rosa Rao

Coordinatore: Prof. Ettore Benedetti



---

*Ai miei nonni. . . . alla mia famiglia*



---

## INDICE

|  |         |
|--|---------|
| <b>RIASSUNTO</b>   | pag. 1  |
| <b>SUMMARY</b>   | pag. 2  |
| <b>1. INTRODUZIONE</b>   | pag. 7  |
| 1.1 Il pomodoro: classificazione botanica  | pag. 7  |
| 1.2 Origine della pianta di pomodoro   | pag. 7  |
| 1.3 L'importanza economica   | pag. 8  |
| 1.4 La variabilità ed il miglioramento genetico  | pag. 9  |
| 1.4.1 La produzione di ibridi  | pag. 10 |
| 1.5 Il 'San Marzano': una cultivar campana di pregio   | pag. 11 |
| 1.6 Marcatori genetici   | pag. 12 |
| 1.6.1 Marcatori morfologici e biochimici   | pag. 12 |
| 1.6.2 Marcatori molecolari del DNA   | pag. 13 |
| 1.6.2.1 VNTR (Variable Number of Tandem Repeats)   | pag. 13 |
| 1.6.2.2 SSR (Simple Sequence Repeat)   | pag. 14 |
| 1.6.2.3 SNP (Single Nucleotide Polymorphism)   | pag. 16 |
| 1.6.2.4 DNA barcoding  | pag. 17 |
| 1.7 Sequenziamento del genoma  | pag. 18 |
| 1.8 La tracciabilità nella filiera agroalimentare del pomodoro                                     | pag. 19 |
| 1.9 Scopo della tesi   | pag. 21 |
| <b>2. MATERIALI E METODI</b>   | pag. 22 |
| 2.1 Materiale vegetale   | pag. 22 |
| 2.2 Estrazione DNA genomico  | pag. 25 |
| 2.3 Quantificazione del DNA genomico estratto  | pag. 25 |
| 2.4 DNA <i>fingerprint</i> mediante il minisatellite (GATA) <sub>4</sub>                           | pag. 26 |
| 2.4.1 Southern blot  | pag. 26 |
| 2.4.2 Marcatura $\gamma^{32}\text{PdATP}$ 5' terminale e ibridazione con sonda (GATA) <sub>4</sub> | pag. 26 |
| 2.5 Analisi mediante SSR   | pag. 27 |
| 2.5.1 Condizioni di amplificazione PCR per l'analisi di loci SSR                                   | pag. 27 |
| 2.5.2 Elettroforesi capillare degli amplificati SSR mediante ABI PRISM 3100-AVANT                  | pag. 28 |
| 2.6 Analisi CAPS   | pag. 29 |
| 2.6.1 Condizioni di amplificazioni di loci EST-SNP   | pag. 29 |
| 2.6.2 Sequenziamento ed analisi dei prodotti PCR per la validazione degli SNP                      | pag. 30 |
| 2.7 Analisi di marcatori in/del  | pag. 30 |
| 2.8 Analisi molecolari mediante SNP  | pag. 31 |

---

|  |         |
|--|---------|
| <b>2.8.1 Condizioni di amplificazione PCR dei frammenti contenenti gli SNP e Estensione del Primer Allele-Specifica (ASPE)</b> | pag. 31 |
| <b>2.8.2 Ibridazione alle microsferre FlexMAP</b>  | pag. 33 |
| <b>2.8.3 Analisi a citometria a flusso mediante Luminex 200</b>  | pag. 34 |
| <b>2.9 Analisi dei dati molecolari</b>   | pag. 34 |
| <b>2.9.1 Calcolo delle relazioni genetiche</b>   | pag. 34 |
| <b>2.9.2 Analisi statistica della struttura delle popolazioni</b>  | pag. 34 |
| <b>2.10 Analisi morfologica</b>  | pag. 35 |
| <b>3. RISULTATI</b>  | pag. 37 |
| <b>3.1 Valorizzazione e protezione di ecotipi di pomodoro campani</b>  | pag. 37 |
| <b>3.1.1 Analisi molecolare mediante minisatellite (GATA)<sub>4</sub></b>  | pag. 37 |
| <b>3.1.2 Caratterizzazione fenotipica</b>  | pag. 39 |
| <b>3.1.3 Analisi molecolare mediante SSR</b>   | pag. 41 |
| <b>3.1.3.1 Amplificazione PCR ed elettroforesi capillare degli ecotipi di pomodoro campani</b>                                 | pag. 41 |
| <b>3.1.3.2 Analisi dei dati SSR per la caratterizzazione e discriminazione degli ecotipi campani</b>                           | pag. 43 |
| <b>3.1.4 Analisi molecolare mediante SNP</b>   | pag. 45 |
| <b>3.1.5 Confronto tra marcatori molecolari nell'analisi degli ecotipi di pomodoro</b>   | pag. 49 |
| <b>3.1.6 Discriminazione di ecotipi di pomodoro 'San Marzano'</b>  | pag. 50 |
| <b>3.1.6.1 Analisi morfologica</b>   | pag. 50 |
| <b>3.1.6.2 Analisi mediante il minisatellite (GATA)<sub>4</sub></b>  | pag. 50 |
| <b>3.1.6.3 Analisi SSR</b>   | pag. 52 |
| <b>3.1.6.4 Analisi CAPS</b>  | pag. 53 |
| <b>3.1.6.5 Analisi SNP</b>   | pag. 55 |
| <b>3.2 Identificazione e discriminazione varietale</b>   | pag. 57 |
| <b>3.2.1 Caratterizzazione varietale mediante SSR</b>  | pag. 57 |
| <b>3.2.2 Identificazione di SNP per la discriminazione varietale</b>   | pag. 60 |
| <b>3.2.3. Analisi mediante in/del</b>  | pag. 62 |
| <b>3.2.4 Correlazione tra i marcatori molecolari</b>   | pag. 64 |
| <b>3.3 Confronto tra marcatori molecolari SSR, SNP, in/del in accessioni di pomodoro coltivato</b>                             | pag. 64 |
| <b>3.4 Relazioni genetiche tra pomodoro coltivato e specie selvatiche mediante analisi SSR</b>                                 | pag. 65 |
| <b>3.5 Analisi della struttura delle popolazioni</b>   | pag. 66 |
| <b>3.6 Tracciabilità della filiera agro-alimentare del pomodoro</b>  | pag. 69 |
| <b>3.6.1 Sviluppo di un protocollo di estrazione e amplificazione di DNA genomico da pomodoro trasformato</b>                  | pag. 69 |

---

|   |          |
|---|----------|
| <b>3.6.2 Utilizzo degli SSR selezionati per la tracciabilità molecolare dei prodotti di filiera</b>   | pag. 70  |
| <b>3.6.3 Validazione sperimentale degli EST-SNP mediante conversione in CAPS e sequenziamento</b>   | pag. 72  |
| <b>4. DISCUSSIONI</b>   | pag. 76  |
| <b>4.1 Valorizzazione e protezione di ecotipi di pomodoro campani</b>   | pag. 77  |
| <b>4.2 Polimorfismo nei diversi marcatori</b>   | pag. 78  |
| <b>4.3 Il pomodoro coltivato in relazione alle specie selvatiche</b>  | pag. 79  |
| <b>4.4 Tracciabilità genetica dei prodotti di filiera</b>   | pag. 80  |
| <b>5. CONCLUSIONI</b>   | pag. 82  |
| <b>6. BIBLIOGRAFIA</b>  | pag. 84  |
| <br>  |          |
| <b>Appendice</b>  | pag. 96  |
| <b>Periodo di formazione all'estero</b>   | pag. 96  |
| <b>Pubblicazioni</b>  | pag. 97  |
| <b>Discrimination of 'San Marzano' accessions: A comparison of minisatellites, CAPS and SSR markers in relation to morphological traits</b> | pag. 98  |
| <b>Innovazioni genetiche per l'identificazione e la protezione di prodotti tipici italiani</b>  | pag. 103 |
| <b>VNTR, SSR, and morphological characterization of tomato accessions spread in Campania region</b>   | pag. 110 |
| <b>A comparison of minisatellite and SSR markers in relation to morphological traits for discrimination of 'San Marzano' accession</b>      | pag. 111 |
| <b>Diversità genetica di accessioni di pomodoro campane</b>   | pag. 112 |

---

## RIASSUNTO

L'applicazione delle tecniche dei marcatori molecolari per la tracciabilità nella filiera agro-alimentare rappresenta una nuova frontiera nell'ambito della tutela di prodotti di alta qualità. La tracciabilità, divenuta dal gennaio 2005 obbligatoria per tutte le aziende agro-alimentari, è definita come "la capacità di rintracciare e seguire un alimento, un mangime, un animale produttore di alimenti o una sostanza attraverso tutti gli stadi della produzione e della distribuzione" (Reg. CE 178/2002). Al fine di rispettare questa normativa, deve essere applicato un continuo controllo di tutte le fasi della catena produttiva, dalla raccolta del materiale in campo, alla valutazione della materia prima in ingresso, all'analisi del prodotto finale. A tale scopo si è indagata la capacità dei marcatori molecolari del DNA, SSR, SNP ed In/Del, per il controllo della filiera agro-alimentare del pomodoro, al fine di offrire una metodica sicura e rapida per la tracciabilità, tramite l'identificazione varietale del materiale fresco e la tracciabilità del prodotto finito. I dati generati hanno portato all'individuazione di un ridotto numero di loci SSR che ha permesso la discriminazione di tutte le varietà ibride da industria e degli ecotipi di pomodoro campani in esame; sono state individuate le condizioni ottimali di estrazione e di amplificazione di DNA da matrice di pomodoro processata di tre differenti tipologie (polpa, pelato, pomodorino inscatolato) ed è stato dimostrato che esiste un limite nelle dimensioni dei frammenti di DNA amplificabile quando proveniente da matrici alimentari complesse e che, nonostante l'elevato grado di degradazione dovuto ai processi produttivi, il DNA possiede frammenti lunghi abbastanza da permettere l'amplificazione di prodotti di peso molecolare superiore alle 150 bp. I profili allelici ottenuti mediante l'utilizzo dei loci SSR opportunamente selezionati sono stati, quindi, confrontati in tre diversi segmenti della filiera agroalimentare del pomodoro (foglia, bacca, prodotto finito) e sulla base dei risultati ottenuti è stata dimostrata la fattibilità dell'impiego dei marcatori SSR per gli studi di tracciabilità di filiera e, in alcuni casi, per l'identificazione della materia prima con la quale il materiale processato monovarietale è stato prodotto. Un ulteriore approccio adoperato per la caratterizzazione delle cultivar e degli ecotipi di pomodoro in esame, è stato l'utilizzo di marcatori SNP ed In/Del, che nelle varietà commerciali hanno mostrato di essere altamente polimorfici con elevato valore informativo. Negli ecotipi, invece, la maggior parte dei loci SNP ed In/Del è risultata essere monomorfa ma il ridotto numero di loci polimorfici riscontrati ha comunque permesso la discriminazione della maggior parte delle accessioni in esame. L'analisi della struttura delle popolazioni dei genotipi analizzati mediante marcatori SNP mostra chiaramente che gli ecotipi campani sono geneticamente distinti dal restante materiale in analisi. Sulla base dei risultati ottenuti e dal confronto delle principali caratteristiche dei marcatori utilizzati per l'analisi delle varietà commerciali e degli ecotipi di pomodoro campani, gli SSR hanno mostrato possedere un potere discriminante più elevato.

## SUMMARY

During last decades the use of local varieties was strongly reduced due to introduction of modern cultivars characterized by higher yield and bred for different traits of agronomic value. However, these cultivars did not always have the quality aspects that were found in older, traditional lines. Performance also depended on the know-how of traditional cultivation. Nowadays the practise of intensive agriculture selects only a small number of species and varieties. This has a consequent reduction of the diversity in agro-ecosystems and risks the loss of important alleles characterizing genetic materials adapted to specific environments.

The tomato variety 'San Marzano' is a symbol of the Italian food-industry and represents, arguably, its most significant horticultural product. However, during the past few decades, the use of this variety has suffered a decline due to its genetic deficiencies and the subsequent diffusion of more productive tomato cultivars with similar fruit shapes (Monti et al., 2004). Modern hybrids possess morphological and quantitative traits that undoubtedly make them more suitable for fresh market or food processing purposes, but they are genetically different from the original variety.

The creation of quality marks of the European Union proved to be a successful system to protect typical products through the Denomination of Origins (PDO- Protected Denomination of Origin and PGI- Protected Geographical Indication). However, the protection of quality needs efficient instruments to discriminate DOP or IGP varieties in the field and to trace them along the agro-food chain. It is essential to have adequate tools to discriminate the original variety from genetic material of different origins.

Over last decade consumers' demand for high quality products have made it necessary to develop a system to certify the genetic origin of material from the field to the table. Traceability has become one of the most important issues in the food chain. The concept of traceability is recognized within the European Union under the regulation ((EC) No. 178/2002) in which traceability is defined as the ability to trace and follow food, feed and ingredients through all stages of production, processing and distribution.

DNA molecular markers are a useful tool for variety identification since they are based on plant genotypes and not much influenced by environmental variation. They can be highly polymorphic and amenable to high throughput systems like microarrays. In addition, DNA is much more resistant to industrial processing procedure, such as exposure to high temperature and low pH, than other molecules, such as proteins or secondary metabolites (Martinez et al., 2003). In a specie such as tomato, despite its low level of genetic diversity (Miller and Tanksley, 1990) due to its self-pollination mating system, DNA molecular markers, if opportunely selected and validated can be highly informative and a powerful tool for variety discrimination. The ideal marker system is highly polymorphic, codominant, accurate, reproducible, high-throughput and low cost (both in terms of capital investment and cost per assay).

Genotyping of crop plants is often carried out with simple sequence repeats (SSRs). Analyses with these markers are high reproducibility, multiallelic and co-dominant. Moreover SSRs are commonly use for mapping studies, variety identification and marker-assisted selection. SSR markers have the advantages of species specificity, being highly informative and only require a short template for amplification in PCR procedure. In tomato a large number of SSR loci have been characterized (He et al., 2003; Areshchenkova 2000; Smulders et al., 1997). He et al. (2003) reported a study

where 500 DNA sequences of tomato genome were investigated for SSRs. Of the 158 pairs of SSR primers screened against a set of 19 different tomato cultivars, 129 pairs produced the expected DNA fragments in their PCR products, and 65 of them were polymorphic with the polymorphism information content (PIC) ranging from 0.09 to 0.67. Smulders et al. (1997) showed that, despite the vast majority of the microsatellites loci have bi- or tri- allelic repeats, they were able to discriminate among a high number of tomato cultivars. However, practically, there could be some problems in sizing SSR alleles due to PCR and electrophoresis artifacts; null alleles arising from mutations in the primer region and size homoplasmy, whereby alleles of the same size may not necessarily be identical in sequence (Jones et al., 2007).

The generation of sequence information for many crop species offers new opportunities in DNA marker studies. SNPs (Single Nucleotide Polymorphisms) are currently considered the new generation of DNA markers. as they be extremely numerous and spread all over the genome. In addition, SNPs are suitable to high throughput procedures thus allowing the contemporary evaluation of a very large number of DNA sequences. They are usually bi-allelic and although less polymorphic than SSRs this limitation can be compensated by their abundance.

In a previous study (Yang et al., 2004), the GenBank EST database was mined for SNPs which resulted in 101 putative SNPs, 83 of which were confirmed empirically. Among these, 53% proved useful within cultivated germplasm. A similar study identified 764 putative SNPs from a similar dataset from which only 27% (28/103) of the putative SNPs tested could be validated experimentally (Labate and Baldo, 2005). The paucity of SNPs within the ESTs can be attributed to low rate of polymorphism within gene coding regions of tomato and a limited sampling of genotypes in the databases. Extrapolation of these experimentally defined SNP rates to the gene and genome level projects a rate of 1 SNP every 4000–8500 bp of coding sequence.

An additional significant advantage of SNP markers is that results across time and labs can be readily compared without the need of extensive checks that is required for comparative sizing of SSR alleles. Similarly, SNPs collected with different SNP chemistries and platforms can be readily compared, a practice not currently feasible with SSRs, where any differences in platform or protocol can effect allele designation. SNPs therefore not only offer lower cost, higher volume, repeatability and accuracy, but they also offer increased opportunities to develop germplasm databases.

The present study documents the comparative efficacy of different molecular markers (minisatellite, SSRs and SNPs) in the discrimination of a representation of tomato accessions spread in the Campania region and their relationship to morphological traits. A previous study indicated that DNA fingerprinting based on the (GATA)<sub>4</sub> minisatellite showed an high discrimination power (Rao et al., 2006). On the basis of this knowledge, an homogeneity analysis was carried out on the ecotypes' population using the minisatellite. The data revealed an high level of polymorphism with 87.10% of polymorphic bands in the analyzed germplasm. The results indicated the presence of ten homogeneous and six heterogeneous ecotypes that showed the presence of two or four different genotypes in the same ecotype.

In addition, the power of SSR and SNP fingerprinting carried out only on the homogeneous ecotypes selected was investigated. As expected for non-hybrid tomato cultivars, the SSR analysis showed that all the genotypes are homozygous. Of the 14 SSRs analyzed, 13 have been shown to be polymorphic among the accessions studied. The mean number of alleles obtained was 2.8 and the highest

PIC was 0.76, with a mean PIC value of 0.46 considering all the 14 SSR loci. The data indicated that the SSR markers revealed high levels of polymorphism allowing the discrimination of all the ecotypes analyzed, but 'San Marzano Cilindrico' and 'San Marzano Murano' accessions that share also the same (GATA)<sub>4</sub> profile, a fingerprint that can be considered the only 'San Marzano' specific one (Rao et al., 2006). It has been found that three SSR amplifications allowed the same level of discrimination. The most informative SSRs had different utilities: some permitted the identification of accessions featuring a specific allele, others distinguished between types of cultivars and others even differentiated forms of the same type. The results indicate that, despite the limited genetic diversity of tomato (Miller and Tanksley, 1990), DNA fingerprints based on SSRs allowed for discrimination not only of different cultivars (He et al., 2003; Smulders et al., 1997), but also of morphologically similar types.

A low level of polymorphism has been found using SNP markers. Only 12 SNP loci out of 71 showed polymorphism among the local cultivars, with a mean PIC value of 0.315 when considering only the polymorphic loci. It was possible to identify five SNPs that could be representative of the level of discrimination of all the SNPs analyzed and that can distinguish the same accessions. The SNP markers used were not able to generate a unique profile of all the accessions. Two ecotypes, closely related on the basis of SSRs and minisatellite analysis, shared the same allelic profile at all the SNP loci.

In addition, the high number of SNP markers analyzed allowed the analysis of the population structure. The accessions spread in the Campania region were compared to a collection of accessions geographically different and with different purposes. The results showed that the ecotypes cultivated in the Campania region clustered separately, indicating that these accessions are genotypically different from all the other material.

In conclusion, both SSRs and the (GATA)<sub>4</sub> minisatellite, most likely due to their multiallelic nature, were useful in depicting genetic relationships consistent with the hierarchical pattern generated by the morphological data, yet SSR analysis possesses various advantages over (GATA)<sub>4</sub>-based discrimination. The former is faster and more suitable for high-throughput analysis, as well as traceability in the agro-food chain (Melchiade et al., 2007; Schlotterer, 2004), as all current DNA-based methods for determining the authenticity of food rely on PCR-based molecular markers (Martinez et al., 2003).

### **Comparison between the molecular markers**

The main features of the three molecular markers analyzed (14 SSRs, 71 SNPs, 6 In/Dels) have been compared between the two populations of cultivated tomato examined, varieties and ecotypes.

SSR markers showed high levels of polymorphism in the examined germplasm. The PIC value and the percentage of polymorphic loci for the SSR markers were found to be high in both varieties (mainly hybrids) and ecotypes. The PIC showed values in the range of 0.441 (varieties) to 0.447 (ecotypes) with an average of 0.444 across the germplasm analyzed. The percentage of polymorphic loci was 92.86%, the same value in both the populations. These results indicated that the discrimination power of SSR markers selected was not influenced by the genetic variability of the examined population. A set of only five SSR markers was found to be able to give a unique profile of all the accessions examined.

SNP markers are mainly bi-allelic and therefore a maximum PIC value of 0.50 can be expected for a given locus. In some case it has almost reached this threshold detecting a PIC value of 0.49. However, SNP markers showed an intermediate level of polymorphism. A higher level of diversity was observed among the varieties as compared to ecotypes. The PIC values calculated for these markers were in the range of 0.25 (varieties) to 0.059 (ecotypes) with an average of 0.154 across the germplasm collection examined and the percentage of polymorphic loci was 75.76% for the varieties and 21.21% for the ecotypes. However, if one takes into consideration only the polymorphic loci to calculate the PIC, this value has been found higher overall among the ecotypes (0.277) where the majority of the SNP markers were monomorphic. This set of SNP markers was not able to distinguish all the accessions, neither among the varieties nor the ecotypes.

Similar results were found using In/Del markers. All the In/Del markers out of six analyzed, were found to be polymorphic among the varieties, but not able to allow the discrimination of every cultivar. Only one In/Del marker out six was found to be polymorphic among the ecotypes.

In conclusion, the data indicated that SSR markers showed the highest level of polymorphism and the highest discrimination power among the germplasm collection analyzed.

### **DNA molecular markers for the traceability in the tomato agro-food chain**

Molecular markers techniques for the traceability in the agro-food chain represent a new frontier to certify high quality food products. The aim of this work was to give a robust and fast tool that allows tomato food to be traced through every step of the production chain back to its origin.

DNA extraction from difficult food matrices is a complex procedure because of the low amount of DNA present and the high degree of degradation due to the processing procedure used for food production. Among the food preservation techniques, canning is carried out according to two main conditions: high temperature and acid pH that are causes of fragmentation of high molecular weight DNA, which can affect its capacity to be amplified.

This study analyzes DNA from different processed foods, chosen in order to assure the presence of a representative sample for each typology (peeled tomato, paste tomato, cherry tomato). In order to obtain high concentration and high quality DNA to be amplified, genomic DNA from processed foods was extracted using different amounts of starting material and following the procedure of different DNA extraction kits (e.g. Genelute Plant Genomic DNA Kit, SIGMA) and the standard CTAB-based DNA extraction protocol. It was not possible to determinate the amount of DNA extracted by agarose gel electrophoresis, because of its low concentration and the high degree of degradation, so it was directly used as template DNA for SSR amplification.

The second step of this research was to select SSR loci able to discriminate all the cultivars used in this project and able to generate amplified fragments when DNA from processed food is used as template. Fourteen tomato SSR loci were selected for their high polymorphism, their easily scored patterns and the ability to generate small size PCR products to amplify degraded DNA. All of them were used to genotype a set of seventeen different tomato cultivars in order to obtain an allele profile of all the genotypes using tomato leaf DNA. Among those microsatellites, three were found to be able to generate a unique profile for each cultivar allowing

their identification and discrimination. All the 14 SSRs were used to analyze DNA extracted from tomato fruits and processed foods. The best results in terms of fragment amplification when DNA from processed foods is used as a template were obtained using dried tissues and the kits procedures for the extraction. The kits allowed to have specific amplifications visible on the gel and peaks detectable by capillary electrophoresis. It was necessary to modify the PCR conditions to obtain a good amplification and PCR products of expected size. Only with a small amount of template DNA it was possible to get amplicons. It could be explained by reduced inhibition of polymerase activity due to the lower presence of inhibitory compounds that are normally present in industrial products (Tengel et al., 2002; Breton et al., 2004; De la Torre et al., 2002). Nine out of the fourteen SSR primer pairs used were able to produce the expected DNA products in a range between 94 -164 bp, while the other five failed to amplify due to the higher molecular sizes of expected fragments. This result demonstrates that there is a limit in amplicon size when the template comes from processed food due to its high degree of degradation. However, the concentration of DNA did not appear to be limiting; rather, successful amplification likely depends on the ability of the different kits to free DNA from the PCR inhibitors present in the samples and on amplicon size.

SSR allelic profiles obtained for tomato leaves, fruits and processed foods were compared. As expected for hybrid tomato cultivars, two peaks were obtained for the major part of analyzed loci, demonstrating that all the genotypes are heterozygous. Of the nine amplified primer pairs, seven were polymorphic and two were monomorphic among the tomato processed foods tested. The results show that the allelic profiles obtained from three of the tomato fruits and processed foods examined ('Tomito', 'Minidor' and 'Cirio 3') matched with their respective allelic profiles obtained from leaf DNA. On the other hand, for 'Talent' there was not a correlation between the allelic profiles obtained from leaf and fruit DNA compared to DNA isolated from processed food. Instead, it matched to the allelic profile of genotype 'Ercole'. This could be explained by the substitution of varieties with the same shape and purpose as 'Talent' at some point in the production chain.

In conclusion, the data shows that it is possible to trace the entire chain of tomato food production and to identify the variety from which tomato mono-varietal processed food has been made. In addition, a protocol has been developed to extract DNA from tomato processed food and it has been shown that DNA concentration are sufficient for consistent amplification if DNA was well purified from other compounds that could inhibit the polymerase activity. Furthermore, it has been found that although DNA in processed food is highly degraded, fragments up to 150 bp can be successfully amplified using microsatellites. Finally, it is possible to conclude that SSRs are a suitable tool to identify tomato fresh materials and to trace them through all the stages of the agro-food chain, even though the PCR product size could be a limiting factor to analyze some SSR loci.

## 1. INTRODUZIONE

### 1.1 Il pomodoro: classificazione botanica

Il pomodoro appartiene alla ampia e diversificata famiglia delle *Solanaceae*, che consiste di 96 generi e più di 2800 specie (Nee et al., 1991). Quella delle *Solanaceae* è una delle più importanti famiglie delle angiosperme per valore economico e comprende molte delle più comuni piante coltivate tra cui la patata, il pomodoro, il peperone, la melanzana, la petunia ed il tabacco. I membri di questa famiglia sono estremamente differenti in termini di *habitat* e morfologia (Knaap et al., 2004). Il pomodoro coltivato, genere *Lycopersicon*, è tra i più piccoli dei generi delle *Solanaceae*. È stato originariamente denominato *Solanum lycopersicum* da Linnaeus (Linnaeus 1753). Nel 1754, Miller separò il pomodoro e designò il genere *Lycopersicon* e la specie *esculentum* per il pomodoro coltivato (Miller 1754). Il genere *Lycopersicon* è stato inizialmente distinto dal genere *Solanum* a causa delle differenti caratteristiche delle antere e delle foglie. Mentre *Lycopersicon* ha le antere che deiscono lateralmente e foglie che si presentano principalmente pinnate, *Solanum* ha antere che deiscono dalle estremità terminali e foglie che tendono ad essere semplici. Le relazioni filogenetiche tra i due generi sono state oggetto per lungo tempo di grandi dibattiti; molti dei ricercatori riconoscevano *Lycopersicon* come un genere distinto dagli altri, mentre altri suggerivano di accorparlo con *Solanum*. Più recentemente, basandosi sulle informazioni derivate dai dati molecolari e morfologici, è stata suggerita una nuova classificazione tassonomica del pomodoro e la riadozione della nomenclatura *Solanum lycopersicum* per il pomodoro coltivato (Knapp et al., 2004; Peralta e Spooner 2001; Olmstead et al., 1999; Spooner et al., 1993; Terrell et al., 1983).

### 1.2 Origine della pianta di pomodoro

L'origine della pianta del pomodoro sembra essere l'America del Sud, in particolare quella larga regione che si estende dall'Ecuador al Perù e dal Cile settentrionale alle isole Galapagos, dove tuttora vive come pianta selvatica. La coltivazione della pianta del pomodoro era già diffusa in età precolombiana come pianta ornamentale in quanto veniva considerata velenosa. Da qui in seguito si diffuse anche in America Centrale e a farla conoscere in Europa furono gli Spagnoli nel XVI secolo, ma ancora esclusivamente a scopo ornamentale.

Il pomodoro originariamente aveva un portamento strisciante e perennante, ma grazie al miglioramento genetico ed alle condizioni climatiche degli areali dove viene coltivato è diventato una pianta eretta ed annuale. In Europa, soltanto in Italia e successivamente in Spagna, il pomodoro fu utilizzato per fini alimentari, soprattutto presso le popolazioni rurali. Probabilmente in Italia furono avviati i primi tentativi di coltivazione che poi si è diffusa principalmente nei paesi del Mediterraneo (Soressi 1969; Esquinas-Alcazar e Nuez, 1995; Garcia-Martinez et al., 2006). Si ritiene, però, che il pomodoro sia stato reintrodotta in Nord America nel XVIII secolo (Rick 1976). In Nord America lo standard del frutto di pomodoro è in realtà la forma globulare, liscia e solida; le altre forme di pomodoro furono coltivate come *heirloom* o varietà da giardino (Bailey et al., 1960; Noble 1994). Di contro, in Italia e negli altri paesi europei, le forme di pomodoro appiattita, periforme, a cuore, allungata, *cherry* e costoluta furono subito apprezzate e coltivate. Molti di questi tipi sono stati utilizzati nei programmi di miglioramento genetico del pomodoro ed alcuni sono stati

tramandati, nel corso degli anni, come varietà locali presenti ancora oggi sui mercati locali.

### 1.3 L'importanza economica

Il pomodoro coltivato *S. lycopersicum*, è una coltura di rilevante valore economico mondiale. È tra le tre colture ortive maggiormente consumate al mondo con una produzione sul mercato fresco di quasi 130 milioni di tonnellate (FAOSTAT 2007; <http://faostat.fao.org>). Nonostante sia una pianta tropicale, il pomodoro è coltivato in ogni angolo del mondo. È coltivato nelle serre nei luoghi dove le basse e rigide temperature sono un fattore limitante. I paesi che rappresentano i maggiori produttori di pomodoro, in ordine discendente, includono Cina, USA, Turchia, India, Egitto ed Italia (<http://faostat.fao.org>). In Italia sono 38.000 le aziende agricole impegnate nella coltivazione del pomodoro (14.400 nella produzione del pomodoro da trasformazione e 23.500 nella produzione di quello da mensa). Degli oltre 550.000 ettari coltivati in Italia ad ortaggi, 68.658 sono quelli che nel 2008 sono stati coltivati per produrre pomodoro da avviare alla trasformazione, 23.000 quelli che invece sono stati coltivati per produrre pomodoro da mensa per il consumo fresco. La produzione è pari a 4,9 milioni di tonnellate di pomodoro per la trasformazione (43,7% nel Nord Italia, 7% nel centro e 49,3% nel Sud Italia). La produzione è concentrata per il 34% in Puglia e il 28% in Emilia Romagna. La Lombardia è la terza regione per superficie coltivata a pomodoro e la Campania la quarta. Dal 2003 al 2007, l'Italia ha prodotto mediamente 6,5 milioni di tonnellate di pomodoro (pomodoro da industria e pomodoro fresco), pari al 38% dell'intera produzione dell'Unione Europea a 27 Paesi membri. Di questi, 5,2 milioni di tonnellate sono destinati alla trasformazione (il 53% del volume trasformato dall'intera Unione Europea) e 1,3 milioni di tonnellate di pomodoro per il consumo fresco. In questo ultimo segmento di mercato l'Italia gioca un ruolo meno importante, rappresentando soltanto il 19% dell'intera offerta europea (fonte: AgroalimentareNews).

Inoltre, il pomodoro è stato un eccellente sistema modello per gli studi di genetica vegetale, in ricerche di base ed applicate. Questo è dovuto a differenti ragioni (Rick e Yoder, 1988), tra cui la possibilità di coltivazione in differenti ambienti, il ciclo vitale breve, l'elevato grado di autoimpollinazione ed omozigosi, il grande potenziale riproduttivo, la facilità di controllarne l'ibridazione, la presenza di un corredo diploide con un genoma di piccole dimensioni (950 Mbp) (Peterson et al., 1998; Arumuganathan e Earle, 1991) e la disponibilità di risorse genetiche e genomiche (<http://tgrc.ucdavis.edu>; <http://www.sng.cornell.edu>). Recentemente, la disponibilità di librerie di inserti genomici di elevato peso molecolare, YAC (Bonnema et al., 1996) e BAC (Budiman et al., 2000; Hamilton et al., 1999), ha facilitato il clonaggio *map-based* e posizionale. Pomodori transgenici sono facilmente prodotti con metodi di trasformazione che utilizzano *Agrobacterium tumefaciens* (Patil et al., 2002; McCormick et al., 1986). Il pomodoro è stata la prima coltura per cui, negli USA, varietà geneticamente ingegnerizzate sono state commercializzate (Breuning e Lyons, 2000) e per cui un gene di resistenza è stato clonato (Martin et al., 1993).

## 1.4 La variabilità ed il miglioramento genetico

Il pomodoro coltivato possiede una limitata variabilità genetica, conseguenza diretta dell'elevato livello di autoimpollinazione e dei differenti eventi di selezione naturale ed artificiale che sono avvenuti durante la domesticazione e la diffusione delle moderne cultivar (Rick 1976). I primi pomodori, introdotti in Europa dagli Spagnoli, fornirono l'intera base genetica delle moderne cultivar e conseguentemente le attuali varietà presenti sul mercato nazionale ed internazionale, sono altamente simili le una alle altre. E' stato stimato che all'interno del genere *S. lycopersicum* e' presente solo il 5% della variabilità genetica totale presente in *Solanum* (Miller e Tanksley, 1990; Rick e Fobes, 1975). Le relative specie selvatiche, comunque, sono una ricca fonte di geni e caratteristiche idonee per il miglioramento genetico.

Per questo motivo, durante gli ultimi 70 anni, le specie selvatiche di pomodoro sono state utilizzate in programmi di *breeding* per migliorare il pomodoro coltivato (Rick 1982; Rick 1979; Rick 1973). Il pomodoro coltivato rappresenta, infatti, il tipico esempio di specie vegetale che ha significativamente beneficiato delle introgressioni da germoplasma esogeno. Numerose sono le accessioni selvatiche identificate che posseggono caratteristiche agronomiche di rilevanza come l'elevata qualità del frutto e resistenza a stress abiotici. E' stato stimato che piu' di 62.800 accessioni di pomodoro coltivato e selvatico sono conservate in diverse banche di germoplasma nel mondo, includendo quelle presenti presso il "United States Department of Agriculture" (USDA) ed il "Plant Genetic Resources Unit" (PGRU) a Geneva (NY) negli USA, il "CM Rick Tomato Genetics Resource Center" (TGRC) dell'Universita' della California, Davis (CA) USA e presso l' "Asia Vegetable Research and Development Center" (AVRDC) in Cina. Altre collezioni di germoplasma sono anche conservate in Olanda (IVT), Russia (VIR), Giappone (NIAS), Perù (DHUNA) e Cuba (INIFAT) (Kalloo 1991).

Il *breeding* di nuove cultivar di pomodoro con migliori caratteristiche acquisite e' cominciato piu' di 200 anni fa in Europa, principalmente in Italia (Stevens e Rick, 1986; Tigchelaar 1986). Ad oggi, le cultivar destinate al mercato fresco ed alla trasformazione industriale sono abbastanza distinte, come risultato delle differenti richieste e necessita' di qualità in funzione della destinazione d'uso. L'obiettivo principale del miglioramento sul pomodoro, per entrambi gli scopi, e' stato l'incremento della produttività. Altre caratteristiche comuni includono la presenza di resistenze a differenti malattie, l'adattabilità a differenti condizioni climatiche e numerose altre caratteristiche riguardanti la qualità del frutto come la precocità di maturazione, la resistenza alla rottura, la consistenza. Caratteristiche specifiche richieste per le cultivar destinate al processamento riguardano la qualità del frutto includendo la compattezza, il colore, il pH, l'acidità, la consistenza. Caratteristiche specifiche, invece, richieste per le cultivar destinate al mercato fresco sono l'uniformità della dimensione dei frutti, la forma ed il colore, l'assenza di anomalie, il gusto e l'aroma (Stevens e Rick, 1986; Tigchelaar 1986).

Uno dei principali limiti della coltivazione di pomodoro è il danno causato da agenti patogeni, quali virus, batteri e funghi. Il loro controllo coinvolge principalmente tre strategie, che sono le tecniche di allevamento, l'applicazione di prodotti chimici, l'uso di varietà resistenti. L'uso di varietà resistenti e' forse lo strumento principale per ridurre i danni causati da agenti patogeni con un impatto minimo sull'ecosistema.

Finora, l'allevamento di pomodoro ha portato a varietà con resistenza ad oltre una decina di agenti patogeni che comprendono virus, batteri e funghi. Per questa specie, la maggior parte delle resistenze utilizzate è monogenica e dominante.

Attualmente esistono diverse varietà coltivate di pomodoro che presentano geni di resistenza contro almeno tre o quattro agenti patogeni (es. varietà VFN con resistenza a *Verticillium*, *Fusarium* e Nematodi). Con l'incremento della produzione di ibridi F<sub>1</sub> sarà possibile utilizzare varietà con quattro a sei resistenze acquisite combinate.

Sin dal loro iniziale sviluppo nel 1980, i marcatori molecolari sono ampiamente utilizzati come strumento principale per l'allevamento di molti raccolti, tra cui il pomodoro. Per la semplicità della base genetica e a causa delle difficoltà riscontrate nell'effettuare la selezione basata sugli aspetti fenotipici, un grande sforzo è stato realizzato per trovare marcatori molecolari legati ai geni resistenza alle malattie. L'utilizzo di marcatori molecolari e le tecniche di MAS (Marker Assisted Selection) hanno facilitato l'identificazione, il mappaggio ed il trasferimento di numerosi geni di resistenza a stress biotici ed abiotici e di loci che controllano caratteri quantitativi (QTL) in pomodoro. Fino ad oggi, più di 40 geni (tra cui molti singoli geni e QTL) che conferiscono resistenza a tutte le principali categorie di agenti patogeni delle piante sono stati mappati in pomodoro (mappa molecolare) e/o clonati da altre *Solanaceae* (Bai et al., 2003; Chunwongse et al., 2002; Parrella et al., 2002; Grube et al., 2000). Indipendentemente dal tipo di tecnologia utilizzata per la selezione, i marcatori strettamente legati ai geni di resistenza possono essere di grande aiuto in programmi di resistenza alle malattie, consentendo di seguire il gene sotto selezione attraverso le generazioni in assenza della verifica (o con una verifica limitata) dell'espressione fenotipica del gene di resistenza (Tanksley et al., 1989). Inoltre, marcatori legati alla resistenza alle malattie possono essere utilizzati anche per piramidizzare nello stesso genotipo geni di resistenza. Infine, diversi geni di resistenza del pomodoro, tra i quali *Cf-2*, *Cf-4*, *Cf-5*, *Cf-9*, *Pto*, *Mi*, *I2*, e *Sw-5*, sono stati identificati grazie alla presenza di marcatori associati. Tra i più rilevanti programmi pubblici di *breeding* vanno citati quelli della University of Florida (Scott et al., 2009), della North Carolina State University (Levis et al., 2005), dell'Ohio State University (Sim et al., 2009), della Pennsylvania State University (Ashrafi et al., 2009) e della Cornell University (Orsi e Tanksley, 2009).

Sebbene il pomodoro rappresenti una delle specie vegetali in cui la selezione assistita ha fornito dei risultati tangibili (Yang e Francis, 2005; Tanksley e Nelson, 1996) l'imminente completamento del sequenziamento delle regioni eucromatiche del genoma fornirà una serie di informazioni genomiche molto più complete, utili sia per lo sviluppo di nuovi marcatori molecolari adatti alla selezione assistita che l'identificazione di geni di resistenza contro i diversi stress.

#### 1.4.1 La produzione di ibridi

Per lungo tempo il miglioramento genetico sul pomodoro era principalmente finalizzato allo sviluppo di cultivar *inbred* ed alla loro commercializzazione. Dal 1970 maggiore enfasi è stata data alla produzione di ibridi F<sub>1</sub>. In molti paesi produttori di pomodoro (USA, Giappone, Europa) la produzione è soprattutto basata sull'utilizzo di varietà ibride. L'uso di ibridi in pomodoro non è soltanto dovuto al beneficio dell'eterosi in sé, ma a fattori legati principalmente alla protezione degli investimenti che i *breeders* hanno effettuato sulla ricerca per combinare un insieme di attributi come le resistenze a malattie e la produzione di cultivar che presentano geni di interesse in condizioni di eterozigosi (Scott e Angell, 1998; Georgiev 1991). Al giorno d'oggi, quasi tutte le varietà commerciali di pomodoro destinate sia al mercato fresco che al processamento industriale sono ibridi.

Come in campo industriale è possibile brevettare una novità scientifica o tecnologica, facendo valere i propri diritti economici su di essa, anche in campo vegetale è prevista la richiesta dei diritti di tutela per le nuove varietà vegetali.

L'UPOV (International Union For The Protection Of New Varieties Of Plants) è una organizzazione intergovernativa fondata nel 1961 ed entrata in vigore nel 1968 per la protezione delle nuove varietà di piante. Scopo dell' UPOV è quello di promuovere un efficiente sistema di protezione sui ritrovati vegetali ed assicurare che i membri dell'Unione riconoscano i risultati raggiunti dai costitutori vegetali, concedendogli un diritto di proprietà intellettuale. Inoltre, assiste i paesi membri (Italia inclusa) nel processo di implementazione nella propria legislazione nazionale. I *breeders* che operano in Italia hanno due possibilità alternative per tutelare le proprie novità vegetali: a) tutela solo per l'Italia, con la protezione nazionale per le nuove varietà vegetali, attraverso il Decreto Legislativo 10 febbraio 2005, n.30; b) tutela per tutto il territorio della Comunità Europea, con il regime comunitario di privativa per i ritrovati vegetali (Reg. (CE) 2100/94).

La diffusione di varietà ibride più produttive, omogenee e migliorate per vari caratteri agronomici importanti ha, però, portato negli ultimi decenni al declino dell'utilizzo di cultivar locali caratterizzate da qualità e tipicità. Le caratteristiche di qualità di queste varietà sono dovute anche al *know-how* dei sistemi di coltivazione tradizionale. L'agricoltura intensiva utilizza un limitato numero di varietà, con conseguente diminuzione della diversità negli agro-ecosistemi e rischio di perdita di alleli importanti presenti in materiali genetici adattati a specifiche condizioni ambientali. Effettuare la caratterizzazione molecolare risulta essere uno strumento utile per proteggere e valorizzare ecotipi locali che potenzialmente aspirano ad avere marchi di qualità (Denominazione d'Origine Protetta- DOP, Indicazione Geografica Protetta - IGP), come già accaduto per il pomodoro 'San Marzano'.

### 1.5 Il 'San Marzano': una cultivar campana di pregio

Secondo la leggenda tradizionale il primo seme di pomodoro è giunto in Italia intorno al 1770, come dono del Regno del Perù al Regno di Napoli, e sarebbe stato piantato nella zona che corrisponde al comune di San Marzano, dando quindi origine a tale varietà. Sebbene l'origine del 'San Marzano' sia ignota ed ancora non definita, è probabile che tale cultivar sia derivata dall'ibridazione o mutazione di pre-esistenti popolazioni locali di pomodoro a frutto allungato. Tali materiali sono, poi, stati selezionati e migliorati dagli agricoltori locali, fino alla definizione di vere e proprie varietà, frutto dell'attività di miglioramento genetico.

Il pomodoro 'San Marzano' è ampiamente conosciuto e commercializzato nel mondo per le sue qualità organolettiche ed è un classico esempio di varietà locale con un valore aggiunto (Porretta 1995). Nonostante questo frutto di pomodoro abbia assunto un rilevante apprezzamento dal punto di vista gastronomico anche per la produzione di processato "pelato", esso possiede caratteristiche agronomiche indesiderate come l'assenza di resistenze a patogeni ed una bassa produttività rispetto alle moderne cultivar ibride con bacche simili (Monti et al., 2004) che sono, però, geneticamente molto differenti dalla varietà originale. Purtroppo, numerosi prodotti che derivano da cultivar ibride di origine e *background* genetico incerti, etichettate come 'San Marzano', sono commercialmente disponibili sul mercato fresco ed industriale.

Dal 1999 la legittima e tradizionale produzione di 'San Marzano' è protetta dalla certificazione del marchio di tutela DOP (Denominazione di Origine Protetta). La creazione di marchi di tutela da parte dell'Unione Europea si è rivelata un utile

strumento per proteggere la tipicità dei prodotti alimentari e rivitalizzare alcuni comparti produttivi che includono produzioni tradizionali di pregio. La Denominazione di Origine Protetta (DOP) estende la tutela del marchio nazionale DOC (Denominazione di Origine Controllata) a tutto il territorio europeo. Il marchio si riferisce ad un prodotto originario di un paese e di una regione caratterizzato da una qualità essenzialmente o esclusivamente dovuta all'areale geografico, inclusi i fattori naturali e umani che lo caratterizzano.

L'Ente Regionale di Sviluppo Agricolo in Campania (E.R.S.A.C.) ha promosso uno studio per assicurare la conformità genetica e morfologica delle piante 'San Marzano' coltivate nella area DOP e per identificare marcatori molecolari capaci non solo di discriminare nell'ambito di differenti genotipi, ma anche potenzialmente utili per identificare inequivocabilmente i prodotti 'San Marzano' nella filiera agroalimentare.

## 1.6 Marcatori genetici

### 1.6.1 Marcatori morfologici e biochimici

Per definizione, ogni carattere espresso in forme multiple può essere considerato ed utilizzato come un marcatore genetico. Fin dal 1900, i *breeders* conoscono la potenzialità di utilizzare i marcatori genetici come criterio di selezione. L'uso di marcatori morfologici è, però, spesso associato a difficoltà inerenti l'espressione dominante, le interazioni epistatiche, gli effetti pleiotropici, la penetranza incompleta. In pomodoro ci sono più di 1300 geni che influenzano le caratteristiche morfologiche, fisiologiche e che sono coinvolti nella resistenza a malattie (Kalloo, 1991), di cui meno di 400 sono stati mappati (Chetelat 2002; Tanksley 1993; Mutschler 1987).

L'IPGRI (International Plant Genetic Resources Institute) ha pubblicato la lista di descrittori morfologici del pomodoro (IPGRI 1996, Descriptors for Tomato, *Lycopersicon* spp.), che include descrittori con diverso potenziale informativo. "Descriptors for Tomato (*Lycopersicon* spp.)" è una revisione di una lista di descrittori che è apparsa nel 1981 in "Genetic Resources of Tomatoes and Wild Relatives" (IBPGR 1981). L'UPOV ha fatto propri questi descrittori (UPOV 2001) che sono attualmente disponibili ed utilizzabili come strumento per l'identificazione e la protezione varietale dagli enti preposti a tale attività (ENSE, Ente Nazionale delle Sementi Elette).

Per essere idonee alla protezione, le varietà devono rispondere ad un set di caratteri e requisiti stabiliti (principalmente morfologici) utilizzati nell'esame DUS (Distinguibilità dalle varietà già esistenti, Uniformità e Stabilità) (Ardley e Hoptroff, 1996). L'esame prevede la documentazione e la registrazione di una serie di rilevanti caratteri morfologici (es. colore, forma, dimensione) e fisiologici (es. tolleranza a malattie) (UPOV 2001) sulle varietà candidate in confronto con varietà simili già esistenti (varietà di riferimento) per un periodo di almeno 2 anni. Questo, insieme al fatto che la discriminazione tra accessioni strettamente correlate è spesso possibile solo durante gli ultimi stadi del ciclo vitale della pianta, come nel caso del pomodoro, ha incanalato l'interesse verso l'utilizzo di marcatori molecolari del DNA a complemento e/o validazione dell'esame DUS (De Riek et al., 2001; Tommasini et al., 2003).

La seconda generazione di marcatori genetici, gli isozimi, divenne popolare durante il 1970 e 1980 (Tanksley 1987). Nonostante presentassero numerosi vantaggi, gli

isozimi sono presenti in numero limitato e spesso non risultano essere polimorfici tra genotipi relativamente vicini (Foolad et al., 1993; Tanksley e Orton, 1983).

### 1.6.2 Marcatori molecolari del DNA

Con l'avvento delle tecnologie basate sui marcatori del DNA nel 1980 (Botstein et al., 1980) ed agli inizi del 1990, molte limitazioni associate ai marcatori morfologici ed agli isozimi sono state superate ed il mappaggio genetico e' entrato in una nuova era con la promessa di un significativo aumento dell'efficienza della ricerca associata alla genetica vegetale ed al *breeding*. Un marcatore molecolare del DNA e' tipicamente una piccola regione di DNA che mostra una sequenza polimorfica tra diversi individui all'interno della stessa specie o tra specie diverse.

Negli ultimi decenni, sono stati sviluppati differenti tipi di marcatori molecolari tra cui RFLP (Restriction Fragment Length Polymorphism) (Smulders et al., 1997; Sandbrink et al., 1995; Van Ooijen et al., 1994), RAPD (Randomly Amplified Polymorphic DNA) (Grandillo e Tanksley, 1996; Stevens et al., 1995), AFLP (Amplified Fragment Length polymorphism) (Vos et al., 1995), VNTR o minisatellite (Variable Number of Tandem Repeat) (Andreakis et al., 2004), SSR o microsatelliti (Simple Sequence Repeat) (He et al., 2003; Smulders et al., 1997), CAPS (Cleaved Amplified Polymorphic Sequence) (Yang et al., 2004; Bai et al., 2004), EST (Expressed Sequence Tag) (Yamamoto et al., 2005; Labate e Baldo, 2005), COS (Conserved Ortholog Set) (Labate et al., 2009; Van Deynze et al., 2007; Fulton et al., 2002), SNP (Single Nucleotide Polymorphism) (Labate e Baldo, 2005) e In/Del (Insertion Deletion) (Yang et al., 2004).

L'elevato numero di marcatori (piu' di 15000) che sono stati sviluppati per gli studi sul pomodoro utilizzando differenti metodi, sono consultabili in una banca dati disponibile sul sito del Sol Genomics Network (SGN, [www.solgenomics.net](http://www.solgenomics.net)) (Shibata 2005).

E' da considerare, pero', che uno dei maggiori problemi dei marcatori in pomodoro e' che molti mostrano bassi livelli di polimorfismo intra ed inter specifico (Ruitz et al., 2005; Saliba-Colombani et al., 2000; Foolad et al., 1993; Williams e St. Clair, 1993; Miller e Tanksley, 1990). Questa limitazione restringe il loro utilizzo in programmi di *breeding* finalizzati all'ottenimento di nuove varietà. A tal proposito, l'immissione dei dati del progetto di sequenziamento del genoma del pomodoro in una banca dati pubblica rappresenta un utile mezzo di informazione per lo sviluppo di marcatori ad alta risoluzione, dando inizio alla scoperta *in silico* di marcatori.

#### 1.6.2.1 VNTR (Variable Number of Tandem Repeat)

Il genoma del pomodoro, cosi' come quello di altri organismi vegetali, include molte sequenze di DNA ripetute. Esse si originano probabilmente a causa di cicli di duplicazione e ricombinazione ineguale e/o conversione genica. Tra le diverse classi di sequenze altamente e mediamente ripetute in un genoma, di interesse ai fini del *fingerprint* molecolare, vi sono quelle caratterizzate da una ripetizione di lunghezza inclusa tra 10 e 60 nucleotidi (Jeffreys et al., 1985). Tali sequenze sono ripetute solitamente in tandem e vengono definite "minisatellite". Sebbene la classificazione in letteratura non e' universale, i minisatelliti sono considerati appartenenti al gruppo dei VNTR (Tausz 1989).

I minisatelliti mostrano polimorfismo di lunghezza derivante dalle differenze nel numero di ripetizioni e per questo sono da considerarsi dei marcatori multiallelici, sebbene le variazioni delle sequenze fiancheggianti i siti *target* della sonda possono

dare origine a polimorfismi tra individui. L' ibridazione del DNA genomico con una sequenza ripetuta che rappresenta il *core* produce un *pattern* di ibridazione che solitamente e' altamente specifico per un individuo od una varieta'. Nonostante il loro elevato polimorfismo e l'enorme successo del loro utilizzo in medicina legale e per i test di paternità (Aronson 2005), i minisatelliti non sono stati ampiamente utilizzati per la genetica di popolazione e la mappatura del genoma. Il *pattern* di ibridazione complesso rende praticamente impossibile l'assegnazione degli alleli presenti in un dato locus. Inoltre, l'impiego dei minisatellite per studi di mappaggio ed associazione e' limitato dal fatto che la distribuzione di queste sequenze nel genoma non e' uniforme. Sebbene sia possibile sviluppare minisatelliti che si riferiscono ad un singolo locus (Armour 1990), questa tecnica per la sua complessità, per il contemporaneo sviluppo di marcatori basati sulla PCR e per la necessita' di DNA di alto peso molecolare e' stata impiegata solo per applicazioni in cui il poter discriminante rappresenta l'esigenza principale.

Per quanto riguarda i VNTR, la letteratura è abbastanza concorde nell'indicare nella sequenza (GATA)<sub>4</sub> il motivo che in pomodoro ha il più elevato potere discriminante (Rao et al., 2006; Kaemmer et al., 1995; Vosman et al., 1992). Il motivo (GATA)<sub>4</sub>, come atteso, non e' pero' uniformemente distribuito all'interno del genoma di pomodoro (Arens et al., 1995; Areshchenkova e Ganal, 1999). Altri motivi che hanno mostrato un elevato potere discriminante sono rappresentati da (GGAT)<sub>4</sub>, (CCTA)<sub>4</sub>, (CA)<sub>8</sub> (Kaemmer et al., 1995). Vosman e Arens (1997) hanno mostrato che i motivi (GATA)<sub>4</sub> e (GACA)<sub>4</sub> sono spesso associati con un altro elemento ripetitivo, che e' stato chiamato U30. La funzione e l'origine di tali sequenze non e' pero' completamente chiarita, sebbene esse possano contribuire alla generazione del polimorfismo. La localizzazione cromosomica di diverse sequenze ripetute di pomodoro e' stata recentemente effettuata con esperimenti di FISH (Fluorescent In Situ Hybridization). Il (GATA)<sub>4</sub> e' principalmente localizzato sull' eterocromatina pericentromerica (circa il 65% del genoma del pomodoro) ed in parte anche nella regione NOR (Chang et al., 2008). Questi dati hanno rappresentato una conferma degli esperimenti di mappatura genetica eseguiti da Arens et al. (1995) and Broun e Tanksley (1996). Nove frammenti contenenti sequenze ripetute (GATA) sono, infatti, stati mappati all'interno di cluster adiacenti ai centromeri (Broun e Tanksley, 1996; Arens et al., 1995).

### 1.6.2.2 SSR (Simple Sequence Repeat)

Un'altra categoria di sequenze ripetute e' rappresentata dai microsatelliti. Conosciuti anche come semplice sequenza di ripetizioni (SSRs) o brevi ripetizioni in tandem (STRs), essi rappresentano sequenze specifiche di DNA costituito da unità ripetute in tandem di lunghezza generalmente compresa tra 1 e 6 nucleotidi (Hancock 1999; Jeffreys et al., 1985). Queste sequenze sono abbondanti nei genomi procarioti ed eucarioti, sia nelle regioni codificanti che non codificanti.

I microsatelliti sono classificati in base alla ripetizione del *core* come: a) perfetti (*core* ininterrotto di ripetizioni, ad esempio (GTG)<sub>15</sub>); b) imperfetti (*core* interrotto con sostituzioni di base, ad esempio (GTG)<sub>7</sub>CTCTG(GTG)<sub>8</sub>); c) composti (*core* costituito da due o più sequenze di unità di ripetizione differente, ad esempio (GTG)<sub>8</sub>(AT)<sub>16</sub>). I microsatelliti si ereditano in modo mendeliano e sono codominanti.

Nelle piante coltivate, è stata documentata la variazione di ripetizioni di-nucleotidiche (AC)<sub>n</sub> e (GA)<sub>n</sub> (Gupta e Varshney, 2000) e di ripetizioni tri-nucleotidiche, le più frequenti delle quali sono (AAG)<sub>n</sub> e (AAT)<sub>n</sub> (Gupta et al., 1996). In uno studio su 54

specie Wang et al. (1994) hanno osservato che l'ordine di abbondanza delle sequenze dei microsatelliti è:  $(AT)_n > (A)_n > (AG)_n > (AAT)_n > (AAC)_n > (AGC)_n > (AAG)_n > (AATT)_n > (AAAT)_n > (AC)_n$ .

Si ritiene che il polimorfismo della lunghezza delle sequenze ripetute del *core* di un microsatellite si origini *in vivo* in seguito allo slittamento (*slippage*) della DNA polimerasi durante la replicazione e successivi *crossing over* ineguali tra cromosomi omologhi durante la meiosi (Levinson e Gutman, 1987). Conoscendo le sequenze fiancheggianti i microsatelliti, è dunque possibile sintetizzare primer per l'amplificazione selettiva di singoli loci. Le differenze di lunghezza degli alleli di un locus SSR sono difficili da risolvere su gel di agarosio. Pertanto, è necessario ricorrere a gel di poliacrilammide in combinazione con marcatura radioattiva o colorazione con nitrato di argento (Scrimshaw, 1992). L'uso di primer fluorescenti in combinazione con l'utilizzo di analizzatori di frammenti di DNA ha dimostrato di essere una soluzione tecnicamente più efficiente rispetto ai metodi tradizionali (Schwengel et al., 1994).

I vantaggi dei microsatelliti sono diversi e comprendono il loro elevato polimorfismo, l'elevata ripetibilità rispetto ad altri marcatori, la possibilità di automazione e di *multiplexing* (Schlotterer 2004; Mitchelson et al., 1997). Per tutti questi motivi non sorprende che gli SSR hanno trovato applicazioni in moltissimi campi con un notevole impatto nell'ambito del miglioramento genetico del pomodoro, così come di altre specie. In particolare sono stati utilizzati per il *fingerprinting* varietale, la costruzione di mappe genetiche, negli studi di associazione con geni di importanza agraria (*gene tagging*, analisi QTL) e in programmi di selezione assistita dai marcatori (MAS).

Per quanto riguarda il pomodoro, riportiamo alcuni dati relativi ad analisi bioinformatiche del "Solanaceae Genomics Resource" al TIGR (Tabella 1.1). Il criterio adottato per la definizione di microsatellite in una sequenza di DNA si basa sul numero minimo di ripetizioni in funzione della lunghezza del motivo ripetuto.

Tabella 1.1: Analisi bioinformatica degli SSR in pomodoro. In totale sono disponibili 4532 SSR, che sono stati classificati in base al numero di unità che sono ripetute nel *core*.

| Dimensioni della unità ripetuta ( <i>core</i> ) in bp | Numero minimo di ripetizioni considerate | Numero di SSR individuati in pomodoro |
|---|--|---------------------------------------|
| 1   | 10                                       | 2626                                  |
| 2   | 6  | 556                                   |
| 3   | 5  | 1119                                  |
| 4   | 5  | 29                                    |
| 5   | 5  | 3                                     |
| 6   | 5  | 19                                    |

Sono attualmente disponibili 31.838 sequenze di pomodoro che rappresentano poco meno di 22 milioni di basi. In base ai criteri citati, sono stati identificati 4.532 SSR in 3.601 sequenze (ci sono quindi 578 sequenze che contengono più di 1 SSR, separato da almeno 100 bp). Inoltre, la tabella 1.1 mostra la distribuzione degli SSR in funzione delle ripetizioni del *core*. Non tutti questi SSR sono stati validati sperimentalmente. È stato riportato che da 500 sequenze di DNA di pomodoro è stato possibile utilizzarne solo 129 (28%) per il disegno di primer SSR. Ciò era dovuto al fatto che la maggioranza, pari al 72%, delle sequenze di DNA aveva la ripetizione semplice sia all'estremità 5' che 3', alcune di esse risultavano duplicate o

ridondanti ed altre avevano delle corte sequenze fiancheggianti ricche in A/T (He et al., 2003).

Inoltre, la variabilità genetica del pomodoro è ridotta ed inegualmente distribuita, con rare isole di polimorfismo che originano da introgressioni (Van der Beek et al., 1992). Infatti, il numero di medio di alleli per locus SSR polimorfici (65 di 129) e' stato di 2,7 (He et al., 2003), un valore simile ad altri riportati in letteratura per il pomodoro (3 alleli per locus riportati da Smulders et al., 1997). Non necessariamente un basso livello di alleli si associa ad un basso potere discriminante. E' stato, infatti, dimostrato che sebbene la maggioranza dei loci SSR siano bi-allelici (49%) o tri-allelici (34%), essi possono comunque discriminare un numero elevato di cultivar. La possibilità di discriminare cultivar di pomodoro e' probabilmente legata ai materiali impiegati. In uno studio relativo a numerosissime varietà di pomodoro, e' riportato che non è stato possibile raggiungere la discriminazione di tutte le 508 diverse varietà in esame utilizzando un set di 20 SSR (Bredemeijer et al., 2002).

Gli SSR in pomodoro sono un esempio di marcatori molecolari che possono essere sviluppati da sequenze esistenti. Smulder et al. (1997), dallo screening delle banche dati EMBL e GenBank, hanno identificato 36 coppie di primer che identificano polimorfismi all'interno o in prossimità di regioni codificanti. In studi recenti, 2.627 SSR sono derivati dall'analisi di 26.363 sequenze EST e 57.222 sequenze di cDNA disponibili al MiBASE (<http://www.accelrys.com/products/gcg>). Inoltre, piu' di 250.000 EST derivate da librerie di cDNA sono attualmente catalogate e disponibili sul sito del SGN. Tutte queste sequenze sono una potenziale fonte di nuovi marcatori, SSR e SNP, utili per l'analisi del genoma del pomodoro.

### 1.6.2.3 SNP (Single Nucleotide Polymorphism)

Gli SNP sono la seconda classe di marcatori molecolari identificabili da sequenze disponibili. Il rilevamento del polimorfismo SNP e' particolarmente interessante a causa delle limitazioni degli SSR dovute alla difficoltà di determinare accuratamente, ed in modo riproducibile nelle diverse piattaforme, le dimensioni alleliche; alla presenza di artefatti durante l'amplificazione PCR e all'omoplasia degli alleli (Jones et al., 2007). Gli SNP sono biallelici e, quindi, generalmente meno informativi degli SSR. E' necessario impiegarne un numero elevato per ottenere una risoluzione ed un polimorfismo comparabile a quello degli SSR. La variazione genetica causata dall'alternanza del singolo nucleotide permette un facile confronto tra i risultati collezionati da diversi laboratori utilizzando diverse chimiche e tecnologie, senza la necessità di estensivi controlli.

In molte specie vegetali, gli SNP sono presenti con una frequenza sufficiente da offrire una metodica alternativa per il mappaggio genetico e la MAS. In mais, la frequenza del polimorfismo riscontrata in linee *inbred* e' di 1 SNP ogni 31 bp nelle regioni non codificanti, e 1 SNP ogni 124 bp nelle regioni codificanti (Ching et al., 2002). Nella soia, studi effettuati sulla diversità genetica in 22 diversi genotipi hanno rilevato la presenza di 1,64 SNP ogni 1000 bp nelle regioni codificanti e 4,85 SNP ogni 1000 bp nelle regioni non codificanti (Zhu et al., 2003). Una ricerca effettuata sull'orzo da Kanazin et al. (2002) riporta una frequenza di 1 mutazione ogni 189 bp. In studi effettuati sul pomodoro (Yang et al., 2004) e' stata consultata la banca dati di EST in GenBank per la ricerca di SNP. I risultati sono stati che di 101 SNP candidati in 44 geni considerati, l'83% (43 SNP) e' stato confermato sperimentalmente. Simili studi (Labate e Baldo, 2005) hanno predetto la presenza di 2.527 SNP in 764 geni. Il risequenziamento di 53 prodotti PCR ha confermato la presenza del solo 27% (62

SNP) degli SNP candidati. Da questi risultati si può dedurre che la frequenza di SNP in pomodoro è di 1 ogni 4000-8500 bp nelle regioni codificanti (Van Deynze et al., 2007). L'insufficienza di SNP in sequenze EST potrebbe essere attribuita al basso livello di polimorfismo nelle regioni codificanti e ad un limitato campionamento di genotipi nella banca dati.

La variabilità genetica è dimostrata essere, invece, più elevata nelle regioni non tradotte (UTR) (Ching et al., 2002) e negli introni, nei quali la diversità nucleotidica è più tollerata in confronto a quella nelle regioni codificanti (Van Deynze et al., 2007). La maggior parte degli SNP (35/62) riportati da Labate e Baldo (2005) non erano predetti dalle sequenze EST, ma sono stati identificati proprio negli introni amplificati. Questo implica che l'analisi degli introni potrebbe essere un altro valido approccio per gli studi della diversità esistente in pomodoro. Un modesto numero di SNP intra-specifici validati è disponibile consultando le banche dati del Tomato Mapping Resource Database ([www.tomatomap.net](http://www.tomatomap.net)) (Francis et al., 2006), del Micro-Tom (MiBase, [www.kazusa.or.jp/jsol/microtom/index.html](http://www.kazusa.or.jp/jsol/microtom/index.html)) (Yamamoto et al., 2005) e del dbSNP del National Center for Biotechnology Information (NCBI) (Labate e Baldo, 2005).

Al pari delle altre specie vegetali, il numero, la diversità e la disponibilità di marcatori molecolari del DNA in pomodoro è in continuo aumento e molti più sforzi sono dedicati all'identificazione di marcatori SNP (Baldo et al., 2007; Ganai et al., 2007; [www.tomatomap.net](http://www.tomatomap.net)). Recentemente, *array* basati sull'utilizzo di oligonucleotidi sono stati impiegati per l'identificazione di polimorfismi di sequenza del DNA in differenti specie vegetali, pomodoro incluso (Sim et al., 2009). Da questo ne deriva che sviluppo di tecnologie ad alta processività, ripetibilità ed accuratezza basate sull'utilizzo dei marcatori molecolari del DNA rappresentano un valido strumento per la caratterizzazione molecolare del germoplasma vegetale a costi sempre più ridotti.

#### 1.6.2.4 DNA barcoding

Sulla scia di numerosi lavori scientifici (Kelly et al., 2007; Kerr et al., 2007; Smith et al., 2006; Hebert et al., 2003b), diversi enti stanno promuovendo ambiziosi progetti con l'obiettivo di associare ad ogni specie vivente una o poche sequenze di DNA in grado di identificarla univocamente. Questo è il principio su cui si basa una nuova metodica che prevede l'integrazione dell'approccio tradizionale all'identificazione degli organismi con il supporto molecolare del DNA barcoding ([www.barcoding.si.edu](http://www.barcoding.si.edu); Hebert et al., 2003a). L'approccio integrato, morfologico e molecolare, permetterebbe la generazione di un *data set* di riferimento che possa essere utilizzato come sistema di identificazione anche per un "non esperto" e che permetterebbe analisi standardizzate su larga scala.

Data la natura pionieristica di questo aspetto della ricerca, dovranno essere per prima cosa identificati marcatori molecolari idonei a consentire la discriminazione a livello di specie all'interno dei taxa, e la variabilità genetica intra ed interspecifica dovrà essere testata sia su specie strettamente correlate dal punto di vista filogenetico che su diverse popolazioni all'interno dell'areale distributivo di specie *target*. Il DNA barcoding è, infatti, basato sull'assunto che una breve sequenza genica standard possa permettere l'attribuzione specifica di campioni non identificati, dato che la variabilità genetica interspecifica eccede quella intraspecifica.

Il potere risolutivo dell'approccio DNA barcoding nell'attribuzione specifica di campioni non identificati è stato confermato in numerose recenti ricerche (Kelly et al., 2007; Kerr et al., 2007; Smith et al., 2006). Ci sono, comunque, problemi, che

limitano al momento un utilizzo generalizzato del DNA barcoding: a) un sistema di identificazione basato sul DNA può funzionare solo se tutte le specie esistenti hanno le loro sequenze depositate in una banca dati. Un *data base* incompleto, infatti, potrà solo permettere agli utilizzatori di determinare se una data sequenza è diversa da quelle depositate. Tale risultato non permetterà, comunque, di identificare una specie, nè automaticamente implicherà che l'individuo in questione sia membro di una nuova specie (Dayrat 2005). b) perché l'approccio molecolare sia efficace, deve essere possibile discriminare tra la variabilità intraspecifica e interspecifica della sequenza utilizzata come "codice a barre".

Di recente, è stata sottolineata l'utilità del coinvolgimento di esperti tassonomi nella ricerca, come premessa indispensabile per individuare e testare il potere discriminatorio delle sequenze su base filogenetica (Kelly et al., 2007).

Il DNA barcoding possiede la potenzialità di essere impiegato sia per ricerche di base in sistematica biologica, sia come sistema di diagnosi e di tracciabilità agro-alimentare, sia per la ricostruzione della filogenesi.

## 1.7 Sequenziamento del genoma

Il genoma nucleare di pomodoro comprende 12 cromosomi e approssimativamente 950 Mbp di DNA, contenente il 59% di regioni non codificanti, il 28% di sequenze codificanti, l'11% di trasposoni e il 2% di sequenze organellari (Wang et al., 2005). Approssimativamente il 77% del DNA cromosomico è compreso in regioni eterocromatiche centromeriche, che sono prive di geni (Wang et al., 2005; Wang et al., 2006). Il genoma del pomodoro codifica per circa 35000 geni, la maggior parte dei quali è presente nelle regioni eucromatiche distali dei cromosomi (van der Hoeven et al., 2002; Peterson et al., 1998; Khush e Rick, 1968) con una densità di 6.7 kb/gene, simile a quella di *Arabidopsis* e riso (Wang et al., 2005).

Ad oggi, i 12 cromosomi di pomodoro sono in via di sequenziamento da parte di un Consorzio internazionale composto da 10 paesi ai quali sono stati assegnati i diversi cromosomi: Cina (3), Francia (7), India (5), Italia (12), Giappone (8), Korea (2), Olanda (6), Spagna (9), Regno Unito (4) e USA (1,10,11) ([http://www.sgn.cornell.edu/help/about/tomato\\_sequencing.pl](http://www.sgn.cornell.edu/help/about/tomato_sequencing.pl)). Questo progetto è parte di una iniziativa conosciuta come "International Solanaceae genome Project (SOL): System Approach to Discovery and Adaptation".

Il 23 ottobre 2009 Keygene ha consegnato al Consorzio la mappa fisica del genoma del pomodoro costruita utilizzando la tecnologia "Whole Genome Profiling" (WGP) ([www.keygene.com](http://www.keygene.com)). È stata coperta l'intera sequenza di 950 Mb che costituiscono il genoma del pomodoro mediante assemblaggio dei cloni BAC sotto condizioni di elevata stringenza. Questa mappa sarà utilizzata come scheletro per produrre una bozza della sequenza del genoma del pomodoro di più elevata qualità entro la fine del 2009.

Iniziato nel 2003, il progetto di sequenziamento ha ambiziosi obiettivi di ricerca da raggiungere nei successivi 10 anni, inclusi gli studi di genomica evolutiva e funzionale della famiglia delle *Solanaceae* (Mueller et al., 2005).

## 1.8 La tracciabilità nella filiera agroalimentare del pomodoro

Il successo del pomodoro è da sempre legato alla sua trasformazione industriale. Furono proprio gli imprenditori che esportavano ortaggi e frutta all'estero a fornire i capitali per lo sviluppo delle prime industrie di trasformazione in Italia, intuendo la possibilità di ampliare il numero di mercati raggiungibili sorpassando il problema della deperibilità dei prodotti freschi. Un impulso notevole alla coltivazione del pomodoro si è, infatti, registrato in Italia agli inizi del secolo scorso, in collegamento con lo sviluppo tecnologico che ha permesso la nascita dell'industria conserviera e di trasformazione.

Inizialmente, il pomodoro veniva semplicemente inscatolato e sterilizzato, secondo quella che veniva definita "appertizzazione". Successivamente, principalmente grazie alla apertura di alcuni stabilimenti nel Mezzogiorno (specialmente nell'agro-nocerino sarnese), si diffuse la "pelatura" dei pomodori con bacca lunga, tipicamente coltivati in tali aree. Lo sviluppo di questo settore industriale causò la conversione di vaste aree agricole alla coltivazione del pomodoro, condizionando in modo positivo gli aspetti socio-economici delle popolazioni locali. Indipendentemente dalla qualità del prodotto, è stata la trasformazione industriale a rendere il pomodoro 'San Marzano' conosciuto a livello mondiale (Porretta 1995).

Sviluppare strumenti che possano essere utili per garantire l'identità dei materiali nella filiera del pomodoro è, quindi, una attività che ha ricadute in diversi settori.

La tracciabilità dei prodotti delle filiere agroalimentari sta assumendo una rilevanza sempre maggiore per l'economia Europea, soprattutto per la pressione che i consumatori esercitano per conoscere in maniera più approfondita, non solo il valore nutritivo dei cibi, ma anche la provenienza e l'autenticità delle materie prime impiegate per il loro ottenimento. La tracciabilità permette di garantire il consumatore da possibili frodi e tutelare scelte alimentari individuali. Inoltre, la possibilità di verificare con sistemi oggettivi e scientifici l'origine di prodotti accresce il valore della certificazione di qualità (come ad esempio i marchi IGP, DOP), favorendo lo sviluppo di aree ad economia marginale attraverso la valorizzazione di prodotti tipici e di nicchia e fornendo incentivi alla conservazione di ecotipi locali preservandone la biodiversità. Il consumatore è sempre più esigente e vuole verificare la qualità e la sicurezza dei cibi, soprattutto alla luce delle ultime emergenze zootecniche riguardanti la BSE (Bovine Spongiform Encephalopathy) e la diossina nei mangimi zootecnici in Belgio, scoprendo quanto siano cambiate le tecniche di produzione degli alimenti e diversificate le strade che li conducono alle nostre tavole.

Oggi, con lo sviluppo dei sistemi organizzativi e gestionali è possibile tornare a dare completa visibilità alle filiere produttive. Lo strumento per assicurare la piena visibilità della filiera è la tracciabilità degli alimenti, cioè l'identificazione documentata dei materiali in ingresso e di tutte le operazioni che concorrono alla formazione del prodotto finito venduto al consumatore. Si tratta di un sistema che permette al consumatore di conoscere, nel più breve tempo e nel miglior modo possibile, la provenienza ed i metodi di produzione, di trasferimento e commercializzazione dei prodotti alimentari. La tracciabilità è indicata anche come strumento fondamentale nel Libro Bianco sulla sicurezza alimentare redatto dalla Commissione dell'Unione Europea e rappresenta un elemento di base delle normative europee sulla responsabilità per danni da prodotto "difettoso". Lo scopo principale è di proteggere la salute del consumatore e assegnare livelli di responsabilità primaria a coloro che sono coinvolti nella produzione agricola ed agroindustriale. Verifica e controllo di questa politica richiedono la capacità di rintracciare i materiali grezzi in ogni fase e ad

ogni passaggio della filiera (Regolamento CE 178/2002).

La tracciabilità è indispensabile per qualsiasi politica di qualità: dalle produzioni a denominazione d'origine protetta o garantita (DOP, IGP) ai prodotti tradizionali. Rappresenta la chiave della qualità per prodotti normali ottenuti in modo conforme alle leggi, in quanto vengono caratterizzati dalla peculiarità di avere una storia identificata. Da una tracciabilità fatta principalmente di documentazione (HACCP) si sta passando sempre più all'utilizzo di nuove tecnologie molecolari per la tracciabilità delle materie prime in filiera agroalimentare.

Il Reg. CE 178/2002 sulla sicurezza alimentare definisce la rintracciabilità come: *"la possibilità di ricostruire e seguire il percorso di un alimento, di un mangime, di un animale destinato alla produzione alimentare o di una sostanza destinata o atta ad entrare a far parte di un alimento o di un mangime attraverso tutte le fasi della produzione, della trasformazione e della distribuzione"*. Affinché il percorso produttivo di un alimento sia rintracciabile, è necessario che venga tracciato. Nella pratica, tracciare una filiera agroalimentare, significa raccogliere i dati che si generano lungo il percorso dal campo alla tavola, ogni volta che si completa una fase produttiva in qualsiasi punto della filiera: settore vivaistico o sementiero, azienda agricola, impresa di prima lavorazione, impresa di trasformazione, distribuzione e consumatore finale. Con il termine di tracciabilità si intende, quindi, il processo informativo che segue il prodotto da monte a valle della filiera produttiva; mentre per rintracciabilità si intende il processo inverso che permette di risalire da valle a monte le informazioni distribuite lungo la filiera. Dal 1 gennaio 2005 sono entrati in vigore la Direttiva Europea 2000/13/CE e il Regolamento Europeo 178/2002 secondo cui tutte le aziende produttrici e di commercializzazione devono adottare le idonee procedure, espresse nella citata direttiva, per poter individuare la provenienza di tutti gli elementi utilizzati in un determinato processo produttivo agro-alimentare per garantirne la trasparenza. Ciò significa che le singole aziende devono provvedere a raccogliere i dati lungo la filiera nelle varie fasi di lavorazione, organizzandoli in modo da mettere a disposizione delle autorità competenti le informazioni inerenti tutto il percorso produttivo. Un documento di grande importanza per il settore agroalimentare è la norma internazionale UNI EN ISO 22005 dal titolo "Rintracciabilità nelle filiere agroalimentari - Principi generali e requisiti di base per la progettazione di sistemi e attuazione". Pubblicata nel gennaio 2008, la norma fornisce i principi e specifica i requisiti di base per progettare ed attuare un sistema di rintracciabilità nel settore agroalimentare, risulta essere pertanto, uno strumento flessibile, utilizzabile da tutti i soggetti della filiera agroalimentare ai fini di determinare l'origine di un prodotto (o un ingrediente), facilitarne la verifica e contribuire alla ricerca delle cause di non conformità. I benefici derivanti dall'adozione di un sistema di rintracciabilità sono molteplici e riguardano sia i consumatori, sia le imprese, delle quali può aumentare la competitività, facilitandone l'inserimento in percorsi di certificazione e riconoscimento. Pertanto, oggi è possibile affermare che il futuro del sistema agroalimentare dipenderà, in gran parte, dal rispetto di due specifiche normative europee, una relativa alla tracciabilità di filiera, ormai divenuta obbligatoria per tutto il settore agro-alimentare; l'altra relativa all'assenza di sostanze chimiche pericolose per la salute umana, che possono residuare nei prodotti agro-alimentari.

L'identificazione diretta a livello del DNA offre uno strumento potente per controllare ed autenticare i sistemi di identificazione tradizionali poiché il codice genetico è inalterabile, presente in ogni organo della pianta ed unico per ciascun individuo. Il recente sviluppo della biologia molecolare e delle tecnologie per l'analisi diretta del DNA ha reso i sistemi per l'identificazione genetica varietale di semplice

applicazione, relativamente poco costosi, sensibili ed adatti all'automazione. Attualmente, le classi di marcatori molecolari più utilizzate sono quelle basate sulla reazione a catena della polimerasi (PCR) (Martinez et al., 2003), dai microsatelliti (Melchiade et al., 2007; Schlotterer, 2004) alle tecnologie ad alta efficienza che vedono lo sviluppo di *microarray* e l'utilizzo di SNP (Sim et al., 2009), la nuova generazione di marcatori che sta sostituendo le altre classi di marcatori.

## **1.9 Scopo della tesi**

Le applicazioni della genetica molecolare avanzata trovano utilizzo nella costituzione varietale (miglioramento genetico e caratterizzazione varietale), nella produzione di sementi (controllo purezza varietale e genetica della produzione sementiera), nella valutazione della novità varietale (esame DUS per il registro e/o un titolo di tutela), nella difesa della tutela varietale (identificazione delle varietà). L'applicazione delle tecniche dei marcatori molecolari per la tracciabilità nella filiera agro-alimentare rappresenta una nuova frontiera nell'ambito della tutela di prodotti di alta qualità. Il presente studio riguarda l'utilizzo dei marcatori molecolari del DNA per il controllo della filiera agro-alimentare del pomodoro, al fine di offrire una metodica sicura e rapida per la tracciabilità, tramite l'identificazione varietale del materiale fresco e la tracciabilità del prodotto finito. Il lavoro riporta la caratterizzazione di ecotipi locali campani di pomodoro che aspirano ad avere il marchio DOP ed IGP, con diversi marcatori molecolari (VNTR, SSR, SNP, In/Del); l'utilizzo di marcatori SNP ed In/Del e l'individuazione di un ridotto numero di SSR per la discriminazione varietale; lo sviluppo ed ottimizzazione di protocolli di estrazione di DNA amplificabile da matrici alimentari complesse; l'utilizzo degli SSR selezionati per la tracciabilità molecolare dei prodotti di filiera.

## 2. MATERIALI E METODI

### 2.1 Materiale vegetale

Il materiale vegetale analizzato comprende una collezione italiana di pomodoro coltivato, una collezione di accessioni appartenenti alla “Core Collection” dell'Ohio State University ([www.tomatomap.net](http://www.tomatomap.net)) e specie selvatiche di pomodoro. Il germoplasma di pomodoro (*S. lycopersicum*) italiano e le specie selvatiche analizzate sono elencati in tabella 2.1 dove sono riportati i codici identificativi relativi ad ogni genotipo utilizzati per le rappresentazioni grafiche dei risultati. La collezione italiana comprende un totale di 64 accessioni, che includono ecotipi campani forniti dalla ditta sementiera “La Semiorto Sementi” e varietà ibride commerciali di rilevante valore economico sul mercato italiano e mondiale. Gli ecotipi campani sono stati selezionati principalmente in base alle loro caratteristiche morfologiche, in modo tale da avere una buona rappresentazione della variabilità fenotipica di pomodoro della regione Campania. Le varietà commerciali sono state selezionate sulla base della forma del frutto e della destinazione d'uso e risultano essere rappresentative del mercato italiano 2007-2008-2009. Delle diciassette varietà ibride commerciali, sette sono state selezionate per analizzare DNA estratto da bacca e rispettivo prodotto trasformato (Tabella 2.1). La selezione è stata effettuata in modo da avere una rappresentanza di prodotto finito per ogni tipologia (pelato, passata, pomodorino inscatolato).

La “Core Collection” in esame è costituita da 96 accessioni di differente provenienza e tipologia (Tabella 2.2) e fa parte della collezione di germoplasma di pomodoro del laboratorio del prof. David Francis dell'Ohio State University. Su questa collezione è stata costruita una banca dati che comprende 300 marcatori tra cui SSR e SNP che è disponibile sul sito “Tomato Mapping Resource Database” ([www.tomatomap.net](http://www.tomatomap.net)) (Francis et al., 2006) e che rientra nel progetto “SolCAP” (Solanaceae Coordinated Agricultural Project).

Tabella 2.1: Nome delle accessioni della collezione di germoplasma di pomodoro coltivato italiano e delle specie selvatiche di pomodoro, codice identificativo, sistema di riproduzione, presenza di materiale ibrido, tipologia del frutto, destinazione d'uso, ente di riferimento e localita' di provenienza.

|  | Nome accessione                       | Codice           | Sistema di riproduzione | Ibrido                  | Tipologia del frutto | Destinazione d'uso           | Ente di riferimento/Localita'  |
|--|---------------------------------------|------------------|-------------------------|-------------------------|----------------------|------------------------------|--------------------------------|
| Varieta' ( <i>Solanum lycopersicum</i> ) | Auspicio                              | Auspicio         | Autogama                | F1                      | Cilindrico           | Pelato                       | "Annalisa" Lodato S.p.A/Italia |
|  | Cirio3*                               | CI3              | Autogama                | F1                      | Cilindrico           | Pelato                       | "Annalisa" Lodato S.p.A/Italia |
|  | Diaz                                  | Diaz             | Autogama                | F1                      | Cilindrico           | Pelato                       | "Annalisa" Lodato S.p.A/Italia |
|  | Discovery                             | Discovery        | Autogama                | F1                      | Cilindrico           | Pelato                       | "Annalisa" Lodato S.p.A/Italia |
|  | Ercole*                               | E                | Autogama                | F1                      | Cilindrico           | Pelato                       | "Annalisa" Lodato S.p.A/Italia |
|  | Genius                                | Genius           | Autogama                | F1                      | Cilindrico           | Pelato                       | "Annalisa" Lodato S.p.A/Italia |
|  | Jet                                   | Jet              | Autogama                | F1                      | Tondo-Prismatico     | Passata                      | "Annalisa" Lodato S.p.A/Italia |
|  | Leader*                               | L                | Autogama                | F1                      | Tondo-Prismatico     | Passata                      | "Annalisa" Lodato S.p.A/Italia |
|  | Minidor*                              | MI               | Autogama                | F1                      | Tondo Piccolo        | Pomodorino inscatolato       | "Annalisa" Lodato S.p.A/Italia |
|  | Player                                | Player           | Autogama                | F1                      | Cilindrico-Piriforme | Pelato                       | "Annalisa" Lodato S.p.A/Italia |
|  | Podium                                | Podium           | Autogama                | F1                      | Tondo                | Passata                      | "Annalisa" Lodato S.p.A/Italia |
|  | PS1296*                               | PS               | Autogama                | F1                      | Tondo                | Passata                      | "Annalisa" Lodato S.p.A/Italia |
|  | PS513                                 | PS513            | Autogama                | F1                      | Cilindrico           | Pelato                       | "Annalisa" Lodato S.p.A/Italia |
|  | Scipio                                | Scipio           | Autogama                | F1                      | Cilindrico           | Pelato                       | "Annalisa" Lodato S.p.A/Italia |
|  | Suerte                                | Suerte           | Autogama                | F1                      | Tondo-Quadrato       | Passata                      | "Annalisa" Lodato S.p.A/Italia |
|  | Talent*                               | TA               | Autogama                | F1                      | Cilindrico           | Pelato                       | "Annalisa" Lodato S.p.A/Italia |
|  | Tomito*                               | TO               | Autogama                | F1                      | Tondo Piccolo        | Pomodorino inscatolato       | "Annalisa" Lodato S.p.A/Italia |
|  | M82                                   | M82              | Autogama                | no                      | Tondo Grande         | Passata                      | DISSPAPA**/Italia              |
|  | Red Setter                            | Red Setter       | Autogama                | no                      | Tondo Grande         | Passata                      | DISSPAPA**/Italia              |
|  | UC82                                  | UC82             | Autogama                | no                      | Cilindrico           | Passata e Concentrato        | DISSPAPA**/Italia              |
| Ecotipo ( <i>Solanum lycopersicum</i> )  | Cuor di Bue Indeterminato             | CB               | Autogama                | no                      | Tondo Grande         | Mensa                        | "La Semiorto Sementi"/Italia   |
|  | Cerise Determinato                    | CE               | Autogama                | no                      | Tondo Piccolo        | Mensa e Inscatolato          | "La Semiorto Sementi"/Italia   |
|  | Corbarino Indeterminato               | CI               | Autogama                | no                      | Tondo Piccolo        | Mensa e Inscatolato          | "La Semiorto Sementi"/Italia   |
|  | Galatino Determinato                  | GD               | Autogama                | no                      | Tondo Piccolo        | Mensa e Inscatolato          | "La Semiorto Sementi"/Italia   |
|  | Principe Borghese Indeterminato       | PB               | Autogama                | no                      | Tondo Piccolo        | Mensa                        | "La Semiorto Sementi"/Italia   |
|  | Principe Borghese Determinato         | PBD              | Autogama                | no                      | Tondo Piccolo        | Mensa                        | "La Semiorto Sementi"/Italia   |
|  | Sorrento Globoso Rosato Indeterminato | SGRI             | Autogama                | no                      | Tondo Globoso Grande | Mensa                        | "La Semiorto Sementi"/Italia   |
|  | San Marzano Cilindrico                | SMC              | Autogama                | no                      | Cilindrico allungato | Mensa e Pelato               | "La Semiorto Sementi"/Italia   |
|  | San Marzano Morini                    | SMMO             | Autogama                | no                      | Cilindrico allungato | Mensa e Pelato               | "La Semiorto Sementi"/Italia   |
|  | San Marzano Murano                    | SMMU             | Autogama                | no                      | Cilindrico allungato | Mensa e Pelato               | "La Semiorto Sementi"/Italia   |
|  | Sorrento Semiorto                     | SorrentoSemiorto | Autogama                | no                      | Tondo Grande         | Mensa                        | "La Semiorto Sementi"/Italia   |
|  | Sorrento Tondo Liscio Rosato          | STLR             | Autogama                | no                      | Tondo Liscio Grande  | Mensa                        | "La Semiorto Sementi"/Italia   |
|  | Tondino Determinato                   | TD               | Autogama                | no                      | Tondo                | Mensa                        | "La Semiorto Sementi"/Italia   |
|  | Tondino Indeterminato                 | TI               | Autogama                | no                      | Tondo                | Mensa                        | "La Semiorto Sementi"/Italia   |
|  | Tondo Liscio Indeterminato            | TLI              | Autogama                | no                      | Tondo                | Mensa                        | "La Semiorto Sementi"/Italia   |
|  | Tondo Marmande Indeterminato          | TM               | Autogama                | no                      | Tondo Grande         | Mensa                        | "La Semiorto Sementi"/Italia   |
|  | Ventura Determinato                   | VD               | Autogama                | no                      | Piriforme            | Mensa                        | "La Semiorto Sementi"/Italia   |
|  | Sorrento Gragnano                     | SorrentoGragnano | Autogama                | no                      | Tondo Grande         | Mensa                        | DISSPAPA**/Italia              |
|  | Sorrento P-98 61                      | Sorrento61       | Autogama                | no                      | Tondo Grande         | Mensa                        | DISSPAPA**/Italia              |
|  | Sorrento P-98 62                      | Sorrento62       | Autogama                | no                      | Tondo Grande         | Mensa                        | DISSPAPA**/Italia              |
|  | Sorrento P-98 65                      | Sorrento65       | Autogama                | no                      | Tondo Grande         | Mensa                        | DISSPAPA**/Italia              |
|  | Agro/Nocerino 2                       | AN2              | Autogama                | no                      | Tondo Piccolo        | Mensa                        | Agro-Nocerino/Italia           |
|  | Agro/Nocerino 3                       | AN3              | Autogama                | no                      | Tondo Piccolo        | Mensa                        | Agro-Nocerino/Italia           |
|  | Agro/Nocerino 4                       | AN4              | Autogama                | no                      | Tondo Piccolo        | Mensa                        | Agro-Nocerino/Italia           |
|  | Agro/Nocerino 5                       | AN5              | Autogama                | no                      | Tondo Piccolo        | Mensa                        | Agro-Nocerino/Italia           |
|  | Agro/Nocerino 6                       | AN6              | Autogama                | no                      | Tondo Piccolo        | Mensa                        | Agro-Nocerino/Italia           |
|  | Agro/Nocerino 7                       | AN7              | Autogama                | no                      | Tondo Piccolo        | Mensa                        | Agro-Nocerino/Italia           |
|  | Agro/Nocerino 8                       | AN8              | Autogama                | no                      | Tondo Piccolo        | Mensa                        | Agro-Nocerino/Italia           |
|  | Agro/Nocerino 9                       | AN9              | Autogama                | no                      | Tondo Piccolo        | Mensa                        | Agro-Nocerino/Italia           |
|  | Agro/Nocerino 13                      | AN13             | Autogama                | no                      | Tondo Piccolo        | Mensa                        | Agro-Nocerino/Italia           |
|  | Agro/Nocerino 14                      | AN14             | Autogama                | no                      | Tondo Piccolo        | Mensa                        | Agro-Nocerino/Italia           |
|  | Agro/Nocerino 16                      | AN16             | Autogama                | no                      | Tondo Piccolo        | Mensa                        | Agro-Nocerino/Italia           |
|  | Agro/Nocerino 17                      | AN17             | Autogama                | no                      | Tondo Piccolo        | Mensa                        | Agro-Nocerino/Italia           |
|  | Agro/Nocerino 18                      | AN18             | Autogama                | no                      | Tondo Piccolo        | Mensa                        | Agro-Nocerino/Italia           |
|  | Agro/Nocerino 19                      | AN19             | Autogama                | no                      | Tondo Piccolo        | Mensa                        | Agro-Nocerino/Italia           |
|  | Agro/Nocerino 21                      | AN21             | Autogama                | no                      | Tondo Piccolo        | Mensa                        | Agro-Nocerino/Italia           |
|  | Agro/Nocerino 22                      | AN22             | Autogama                | no                      | Tondo Piccolo        | Mensa                        | Agro-Nocerino/Italia           |
|  | Area Vesuviana 14                     | AreaVes14        | Autogama                | no                      | Tondo Piccolo        | Mensa                        | Agro-Nocerino/Italia           |
|  | Area Vesuviana 31                     | AreaVes31        | Autogama                | no                      | Tondo Piccolo        | Mensa                        | Agro-Nocerino/Italia           |
|  | Corbara 1.6                           | CO1.6            | Autogama                | no                      | Tondo Piccolo        | Mensa                        | Agro-Nocerino/Italia           |
|  | Corbara 2.7                           | CO2.7            | Autogama                | no                      | Tondo Piccolo        | Mensa                        | Agro-Nocerino/Italia           |
|  | B.P.                                  | B.P.             | Autogama                | no                      | Tondo Piccolo        | Mensa                        | Montanaso Lombardo/Italia      |
| Datterino                                | Datterino                             | Autogama         | no                      | Tondo Piccolo           | Mensa                | Montanaso Lombardo/Italia    |                                |
| Pisanello                                | PSN                                   | Autogama         | no                      | Tondo Appiattito Grande | Mensa                | Monsampolo del Tronto/Italia |                                |
| Specie selvatica                         | S. Arcanum LA1708                     | ARC              | Allogama-SI             |                         |                      |                              | Peru'                          |
|  | S. Chmielewskii LA1327                | CHW              | Facoltativa-SC          |                         |                      |                              | Peru'                          |
|  | S. Habrochaites LA1777                | HBC              | Allogama-SI             |                         |                      |                              | Peru'                          |
|  | S. Neorickii LA2133                   | NRK              | Autogama-SC             |                         |                      |                              | Ecuador                        |
|  | S. Pennellii LA0716                   | CHW              | Facoltativa-SC          |                         |                      |                              | Peru'                          |

\*Varieta' ibride commerciali selezionate per le analisi di tracciabilita' della filiera agroalimentare del pomodoro.

Le sigle rappresentano i codici identificativi riportati nel testo, nelle tabelle e nei grafici ed i numeri sono indicativi della pianta analizzata.

\*\*DISSPAPA: Dipartimento del Suolo, della Pianta, dell'Ambiente e delle Produzioni Animali, Facolta' di Agraria, Portici (NA), Italia.

Tabella 2.2: Nome delle accessioni appartenenti alla "Core Collection", tipologia o destinazione d'uso, specie di pomodoro di appartenenza, ente donatore o risorsa, localita' di provenienza.

| Accessione       | Tipologia/Destinazione | Specie   | Donatore/Risorsa                          | Provenienza |
|------------------|------------------------|--|---|-------------|
| BananaLegs       | Heirloom               | <i>S. lycopersicum</i>                         | Ohio State University                     | USA (OH)    |
| Beauty           | Heirloom               | <i>S. lycopersicum</i>                         | Ohio State University                     | USA (OH)    |
| BlackPlum        | Heirloom               | <i>S. lycopersicum</i>                         | Ohio State University                     | Russia      |
| Campbell28       | Mercato fresco         | <i>S. lycopersicum</i>                         | Pubblica                                  | USA (NJ)    |
| Campbell31       | Inscatolato            | <i>S. lycopersicum</i>                         | Pubblica                                  | USA (NJ)    |
| Campbell34       | Inscatolato            | <i>S. lycopersicum</i>                         | Pubblica                                  | USA (NJ)    |
| Campbell1327     | Inscatolato            | <i>S. lycopersicum</i>                         | Pubblica                                  | USA (NJ)    |
| E3259            | Inscatolato            | <i>S. lycopersicum</i>                         | Ohio State University                     | USA (OH)    |
| E6203            | Inscatolato            | <i>S. lycopersicum</i>                         | Ohio State University                     | USA (CA)    |
| Favorite         | Heirloom               | <i>S. lycopersicum</i>                         | Ohio State University                     | USA (OH)    |
| Fla. 7060        | Mercato fresco         | <i>S. lycopersicum</i>                         | University of Florida                     | USA (FL)    |
| Fla. 7547        | Mercato fresco         | <i>S. lycopersicum</i>                         | University of Florida                     | USA (FL)    |
| Fla. 7600        | Mercato fresco         | <i>S. lycopersicum</i>                         | University of Florida                     | USA (FL)    |
| Fla. 7770        | Mercato fresco         | <i>S. lycopersicum</i>                         | University of Florida                     | USA (FL)    |
| Fla. 7771        | Mercato fresco         | <i>S. lycopersicum</i>                         | University of Florida                     | USA (FL)    |
| Fla. 7775        | Mercato fresco         | <i>S. lycopersicum</i>                         | University of Florida                     | USA (FL)    |
| Fla. 8233        | Mercato fresco         | <i>S. lycopersicum</i>                         | University of Florida                     | USA (FL)    |
| Fla. 8249        | Mercato fresco         | <i>S. lycopersicum</i>                         | University of Florida                     | USA (FL)    |
| Flora-Dade       | Mercato fresco         | <i>S. lycopersicum</i>                         | Pubblica                                  | USA (FL)    |
| Globe            | Heirloom               | <i>S. lycopersicum</i>                         | Ohio State University                     | USA (OH)    |
| GoldBall         | Heirloom               | <i>S. lycopersicum</i>                         | Ohio State University                     | USA (OH)    |
| Grushovka        | Heirloom               | <i>S. lycopersicum</i>                         | Ohio State University                     | Russia      |
| H722             | Inscatolato            | <i>S. lycopersicum</i>                         | Pubblica, Tomato Genetics Resource Center | USA (OH)    |
| H1706            | Inscatolato            | <i>S. lycopersicum</i>                         | Pubblica, Tomato Genetics Resource Center | USA (OH)    |
| Ha7981           | Unimproved             | <i>S. lycopersicum</i>                         | Pubblica                                  | USA         |
| Ha7998           | Unimproved             | <i>S. lycopersicum</i>                         | Tomato Genetics Resource Center           | USA         |
| Heinz 1370       | Land Race              | <i>S. lycopersicum</i>                         | United States Department of Agriculture   |             |
| HowardGerman     | Heirloom               | <i>S. lycopersicum</i>                         | Ohio State University                     |             |
| Hunt100          | Inscatolato            | <i>S. lycopersicum</i>                         | Pubblica, Tomato Genetics Resource Center | USA         |
| JersyDevil       | Heirloom               | <i>S. lycopersicum</i>                         |   |             |
| LA 410           | Land Race              | <i>S. lycopersicum</i>                         | United States Department of Agriculture   |             |
| LA0407           | Selvatica              | <i>S. habrochaites</i>                         | Ohio State University                     | Ecuador     |
| LA0716           | Selvatica              | <i>S. pennellii</i>                            | Ohio State University                     | Peru        |
| LA1208           | Selvatica              | <i>S. lycopersicum</i> var. <i>cerasiforme</i> | Oregon State University                   | Colombia    |
| LA1269           | Selvatica              | <i>S. pimpinellifolium</i>                     | Ohio State University                     | Peru        |
| LA1464           | Selvatica              | <i>S. lycopersicum</i> var. <i>cerasiforme</i> | Oregon State University                   | Honduras    |
| LA1547           | Selvatica              | <i>S. pimpinellifolium</i>                     | Ohio State University                     | Ecuador     |
| LA1589           | Selvatica              | <i>S. pimpinellifolium</i>                     | Ohio State University                     | Peru        |
| LA2184           | Selvatica              | <i>S. pimpinellifolium</i>                     | Ohio State University                     | Peru        |
| LA2312           | Selvatica              | <i>S. lycopersicum</i> var. <i>cerasiforme</i> | Oregon State University                   | Peru        |
| LA2378           | Mercato fresco         | <i>S. lycopersicum</i>                         | Oregon State University                   |             |
| LA2393           | Selvatica              | <i>S. lycopersicum</i> var. <i>cerasiforme</i> | Oregon State University                   | Costa Rica  |
| M82              | Inscatolato            | <i>S. lycopersicum</i>                         | Pubblica, Tomato Genetics Resource Center | USA (CA)    |
| Marglobe         | Heirloom               | <i>S. lycopersicum</i>                         |   | USA         |
| Moneymaker       | Heirloom               | <i>S. lycopersicum</i>                         | Ohio State University                     | Olanda      |
| NC23E-2          | Mercato fresco         | <i>S. lycopersicum</i>                         | North Carolina State University           | USA (NC)    |
| NC265-1          | Mercato fresco         | <i>S. lycopersicum</i>                         | North Carolina State University           | USA (NC)    |
| NC84173          | Mercato fresco         | <i>S. lycopersicum</i>                         | North Carolina State University           | USA (NC)    |
| NC99471-3        | Mercato fresco         | <i>S. lycopersicum</i>                         | North Carolina State University           | USA (NC)    |
| NCEBR-1          | Mercato fresco         | <i>S. lycopersicum</i>                         | North Carolina State University           | USA (NC)    |
| NCEBR-2          | Mercato fresco         | <i>S. lycopersicum</i>                         | North Carolina State University           | USA (NC)    |
| OH7814           | Inscatolato            | <i>S. lycopersicum</i>                         | Ohio State University                     | USA (OH)    |
| OH7983           | Inscatolato            | <i>S. lycopersicum</i>                         | Ohio State University                     | USA (OH)    |
| OH8243           | Inscatolato            | <i>S. lycopersicum</i>                         | Ohio State University                     | USA (OH)    |
| OH8245           | Inscatolato            | <i>S. lycopersicum</i>                         | Ohio State University                     | USA (OH)    |
| OH8446           | Inscatolato            | <i>S. lycopersicum</i>                         | Ohio State University                     | USA (OH)    |
| OH8556           | Inscatolato            | <i>S. lycopersicum</i>                         | Ohio State University                     | USA (OH)    |
| OH86120          | Inscatolato            | <i>S. lycopersicum</i>                         | Ohio State University                     | USA (OH)    |
| OH87160          | Inscatolato            | <i>S. lycopersicum</i>                         | Ohio State University                     | USA (OH)    |
| OH88119          | Inscatolato            | <i>S. lycopersicum</i>                         | Ohio State University                     | USA (OH)    |
| OH9241           | Inscatolato            | <i>S. lycopersicum</i>                         | Ohio State University                     | USA (OH)    |
| OH9242           | Inscatolato            | <i>S. lycopersicum</i>                         | Ohio State University                     | USA (OH)    |
| OH981049         | Inscatolato            | <i>S. lycopersicum</i>                         | Ohio State University                     | USA (OH)    |
| OH981057         | Inscatolato            | <i>S. lycopersicum</i>                         | Ohio State University                     | USA (OH)    |
| OH981067         | Inscatolato            | <i>S. lycopersicum</i>                         | Ohio State University                     | USA (OH)    |
| OH981136         | Inscatolato            | <i>S. lycopersicum</i>                         | Ohio State University                     | USA (CA)    |
| OH987034         | Inscatolato            | <i>S. lycopersicum</i>                         | Ohio State University                     | USA (OH)    |
| Ohio11           | Mercato fresco         | <i>S. lycopersicum</i>                         | Ohio State University                     | USA (OH)    |
| OhioMR13         | Mercato fresco         | <i>S. lycopersicum</i>                         | Ohio State University                     | USA (OH)    |
| Opalka           | Heirloom               | <i>S. lycopersicum</i>                         | Ohio State University                     | Polonia     |
| OR11             | Mercato fresco         | <i>S. lycopersicum</i>                         | Oregon State University                   | USA (OR)    |
| Oxheart          | Heirloom               | <i>S. lycopersicum</i>                         | Ohio State University                     | USA (OH)    |
| PARAGON          | Heirloom               | <i>S. lycopersicum</i>                         | Ohio State University                     | USA (OH)    |
| PI114490         | Selvatica              | <i>S. lycopersicum</i> var. <i>cerasiforme</i> | Ohio State University                     |             |
| PI128216         | Selvatica              | <i>S. pimpinellifolium</i>                     | Ohio State University                     | Bolivia     |
| PI365914         | Selvatica              | <i>S. pimpinellifolium</i>                     | United States Department of Agriculture   | Ecuador     |
| PrincipeBorghese | Heirloom               | <i>S. lycopersicum</i>                         | Ohio State University                     | Italia      |
| R001             | Inscatolato            | <i>S. lycopersicum</i>                         | University of Guelph                      | CANADA      |
| R9812            | Inscatolato            | <i>S. lycopersicum</i>                         | University of Guelph                      | CANADA      |
| R996             | Inscatolato            | <i>S. lycopersicum</i>                         | University of Guelph                      | CANADA      |
| R997             | Inscatolato            | <i>S. lycopersicum</i>                         | University of Guelph                      | CANADA      |
| RioGrande        | Mercato fresco         | <i>S. lycopersicum</i>                         | Tomato Genetics Resource Center           |             |
| RomaVF           | Heirloom               | <i>S. lycopersicum</i>                         | Ohio State University                     |             |
| Rutgers          | Heirloom               | <i>S. lycopersicum</i>                         | Ohio State University                     | USA (NJ)    |
| SanMarzano       | Heirloom               | <i>S. lycopersicum</i>                         | Ohio State University                     | Italia      |
| Sausage          | Heirloom               | <i>S. lycopersicum</i>                         | Ohio State University                     | USA (OH)    |
| Simcoe102        | Inscatolato            | <i>S. lycopersicum</i>                         | Ohio State University                     | USA (OH)    |
| Spitz            | Heirloom               | <i>S. lycopersicum</i>                         | Ohio State University                     | USA (OH)    |
| Sun1642          | Inscatolato            | <i>S. lycopersicum</i>                         | Pubblica                                  | USA (CA)    |
| T-5              | Mercato fresco         | <i>S. lycopersicum</i>                         | Tomato Genetics Resource Center           | USA (CA)    |
| Tropic           | Mercato fresco         | <i>S. lycopersicum</i>                         | Pubblica                                  |             |
| Turrialba        | Land Race              | <i>S. lycopersicum</i>                         | United States Department of Agriculture   |             |
| Unilever265      | Inscatolato            | <i>S. lycopersicum</i>                         | Pubblica                                  | USA (CA)    |
| Walter           | Mercato fresco         | <i>S. lycopersicum</i>                         |   |             |
| Yellow Pear      | Heirloom               | <i>S. lycopersicum</i>                         | Ohio State University                     |             |
| Yellow Stuffer   | Heirloom               | <i>S. lycopersicum</i>                         | Ohio State University                     |             |

## 2.2 Estrazione DNA genomico

L'estrazione di DNA per analisi (GATA)<sub>4</sub> e' stata effettuata nel seguente modo: mortaio e pestello sono stati raffreddati in azoto liquido e 3-4 g di giovani foglie sono stati triturati fino ad essere ridotti allo stato di polvere molto fine e mantenuti sempre in azoto liquido. Successivamente il triturato e' stato trasferito in tubi Falcon da 50 ml e sono stati aggiunti 15 ml di soluzione di estrazione 0,1 M Tris-HCl (pH 8,0), 0,5 M NaCl, 0,05 M EDTA, 2 ml di 10% SDS e 10,5 µl di β-mercaptoetanol. I campioni sono stati agitati vigorosamente ed incubati a 65°C per 20 min. Dopo aggiunta di 5 ml di 5 M potassio acetato, i campioni sono stati lasciati in ghiaccio per 15 min e successivamente centrifugati per 10 min a 7000 rpm in una "Eppendorf Centrifuge 5810 R" utilizzando un rotore angolare fisso (F-34-6-38). Il supernatante e' stato poi filtrato su carta da filtro trasferendolo in un nuovo tubo Falcon da 50 ml. Ad ogni campione sono stati aggiunti 10 ml di isopropanolo a 4°C ed il tutto centrifugato per 30 min a 8000 rpm. Rimosso il supernatante e lasciato asciugare il pellet a temperatura ambiente, 400 µl di 1X TE (0,01 M Tris-HCl, 0,01 mM EDTA; pH 8.0) (Sambrook *et al.*, 2001), sono stati aggiunti per risospendere il precipitato. E' stato effettuato un trattamento aggiungendo 10 µl di 10mg/ml RNasi con successiva incubazione a 37°C per 15 min. Sono stati aggiunti 400 µl di soluzione contenente 2% CTAB (p/v), 2 M NaCl, 0,2 M Tris-HCl (pH 7,0), 50 mM EDTA ed effettuata un'incubazione di 15 min a 65°C. I campioni sono stati trasferiti in Eppendorf da 2 ml, sono stati aggiunti 800 µl cloroformio: ottanolo (24:1), miscelati per inversione a temperatura ambiente e successivamente centrifugati per 5 min a 14000 rpm in una centrifuga "Eppendorf Centrifuge 5810 R" con un rotore ad angolo fisso F-45-30-11. E' stata poi delicatamente prelevata la fase superiore di ogni campione e trasferita in Eppendorf da 2 ml, a cui sono stati aggiunti 800 µl di isopropanolo con successiva centrifugazione di 5 min a 14000 rpm. Allontanato il supernatante, il precipitato e' stato lavato con 500 µl di etanolo al 70%. I campioni sono stati, poi, nuovamente centrifugati per 5 min a 13000 rpm. Il precipitato è stato asciugato sotto vuoto e, quindi, risospeso in 300 µl di 1X TE precedentemente riscaldato a 65°C.

Per i campioni di pomodoro analizzati mediante SSR e SNP, il DNA genomico e' stato estratto utilizzando 100 mg di giovani foglie congelate oppure 250 mg di frutti di pomodoro congelati seguendo la procedura prevista dal "Genelute Plant Genomic DNA Kit" (G2N70-SIGMA), previa triturazione del tessuto vegetale in azoto liquido. Le foglie provengono da singola pianta per ognuno dei genotipi analizzati, mentre le bacche delle varietà selezionate per gli studi di tracciabilità di filiera, fornite dalla ditta di trasformazione "Annalisa", sono state raccolte da piante diverse e collezionate insieme. Il DNA dai prodotti processati e' stato estratto utilizzando 250 mg di materiale precedentemente liofilizzato. Sono stati impiegati il metodo basato sul CTAB precedentemente descritto ed il protocollo previsto dal "Genelute Plant Genomic DNA Kit" (G2N70-SIGMA).

## 2.3 Quantificazione del DNA genomico estratto

Il DNA genomico estratto è stato quantificato mediante elettroforesi orizzontale in gel di agarosio all'1% (p/v) in tampone 1X TAE (Sambrook *et al.*, 2001). La corsa elettroforetica è stata effettuata per circa 15 min ad un voltaggio di 100 V. Il DNA è stato visualizzato per esposizione ai raggi UV su di un trans-illuminatore (Bio-Rad UV Transilluminator 2000) e poi fotografato mediante l'uso della macchina "Kodak Gel Logic 100". Per determinare la quantità di DNA estratto è stata paragonata l'intensità

di ciascuna banda con quella relativa al DNA genomico del fago  $\lambda$  (Promega) caricato in tre differenti quantita' (50, 100 e 200ng).

## 2.4 DNA *fingerprint* mediante il minisatellite (GATA)<sub>4</sub>

### 2.4.1 Southern blot

Sono state analizzate quattro piante per ogni ecotipo. Di ognuno dei campioni 5  $\mu\text{g}$  di DNA genomico sono stati digeriti con enzima di restrizione *TaqI* (Promega). La reazione di digestione e' stata allestita aggiungendo al DNA genomico, 40 U di *TaqI*, 1/10 del volume totale di 10 $\mu\text{g}/\mu\text{l}$  BSA e 1X buffer in un volume finale di reazione di 450  $\mu\text{l}$ . La reazione e' stata incubata a 65°C per 16 ore. E' seguita una precipitazione con 1/10 del volume totale di 3 M sodio acetato (pH 5,2) e 3 volumi di 100% etanolo, successiva incubazione a -80°C per 5 min, centrifugazione a 14000 rpm per 15 min a 4°C, eliminazione del supernatante, lavaggio del precipitato mediante 70% etanolo e centrifugazione per 2 min a 14000 rpm. Il precipitato e' stato lasciato asciugare e risospeso in 40  $\mu\text{l}$  di 1X TE. Ai 40  $\mu\text{l}$  di ogni campione e' stato aggiunto 1/10 del volume di "gel-loading buffer" (Sambrook et al., 2001) e sono stati, quindi, caricati su gel di agarosio (0,8 -1,0%) contenente 1X TAE buffer. Il gel e' stato corso a 25-30 V per 18-24 ore, in modo da separare tutti i frammenti di digestione. Successivamente il gel e' stato immerso in una soluzione di denaturazione (1,5 M NaCl, 0,5 M NaOH) per 30 min a temperatura ambiente, in lenta agitazione. Allontanata la soluzione di denaturazione, il gel e' stato incubato nella soluzione di neutralizzazione (1 M Tris (pH 7,4), 1,5 M NaCl) per 45 min a temperatura ambiente, in lenta agitazione. E' stato allestito il Southern blot ed utilizzata la soluzione 20X SSC (Sambrook et al., 2001). Il trasferimento per capillarita' del DNA dal gel alla membrana di nylon (Hybond N, Amersham) e' avvenuto *over/night*. Disassemblato il blot, il DNA e' stato fissato sulla membrana mediante esposizione agli UV (150 mJoule) utilizzando un apparecchio "Bio-Rad GS Gene Linker".

### 2.4.2 Marcatura $\gamma^{32}\text{PdATP}$ 5' terminale e ibridazione con sonda (GATA)<sub>4</sub>

La membrana e' stata preibridata per almeno 1 h nel buffer di ibridazione contenente 0,5 M Na<sub>2</sub>PO<sub>4</sub> (pH 7,2), 7% SDS, 1 mM EDTA, precedentemente riscaldato a 35°C. La marcatura 5' terminale della sonda (GATA)<sub>4</sub> e' stata allestita in un volume finale di 50  $\mu\text{l}$  utilizzando 1 X buffer kinase (Promega), 25 pmol di oligo (GATA)<sub>4</sub>, 0,2 U di T4 polinucleotide kinase (Promega) e 5  $\mu\text{l}$  di  $\gamma^{32}\text{PdATP}$  (Amersham). La reazione e' stata incubata per 1 h a 37°C e successivamente purificata mediante colonne Microspin G-25 (GE Healthcare) secondo le istruzioni fornite dal produttore. La membrana e' stata posta in 25 ml della soluzione di ibridazione dove sono stati aggiunti 50  $\mu\text{l}$  di sonda marcata purificata e lasciata in incubazione alla temperatura di 35°C *over/night*. La membrana e' stata, poi, lavata con una soluzione di lavaggio (3 X SSC, 0,1 % SDS) mediante due successivi lavaggi di 2 min ed asciugata su un foglio di carta "Whatman". La membrana e' stata posta nella cassetta per impressionare la lastra fotografica (Kodak X-Omat) e riposta a -80°C con uno schermo per l'amplificazione del segnale (Sambrook et al., 2001). Ventotto bande sono state utilizzate per analisi di variabilita' genetica ed ogni frammento e' stato assunto essere rappresentativo di un singolo locus. Solo bande riproducibili sono state prese in considerazione per presenza (1) oppure assenza (0) e monitorate

mediante analisi comparativa del profilo di ibridazione ottenuto per quattro piante differenti per ogni ecotipo.

## 2.5 Analisi mediante SSR

### 2.5.1 Condizioni di amplificazione PCR per l'analisi di loci SSR

Sono stati selezionati dalla letteratura 14 loci SSR (He et al., 2003; Smulders et al., 1997) rappresentativi di varie classi di ripetizione ed analizzati mediante amplificazione PCR (Tabella 2.3). Le reazioni di amplificazione del DNA genomico estratto da foglia e bacca sono state condotte in un volume finale di 25  $\mu$ l contenente 20 ng di DNA genomico, 1X PCR buffer (Promega), 0,1 mM dNTP, 0,2  $\mu$ M di primer forward marcato all'estremità 5' con un fluoroforo (Tabella 2.3), 0,2  $\mu$ M di primer reverse non marcato, 0,5 U di *Taq* polimerasi (Promega). Le reazioni di amplificazione sono state eseguite in "Eppendorf Mastercycler Gradient thermocycler" il cui programma è stato impostato con le seguenti temperature e tempi di reazione: un primo step di denaturazione del DNA a 94°C per 3 min; 35 cicli comprendenti la denaturazione a 94°C per 45 sec, l'*annealing* per 45 sec alla specifica  $T_a$  (Tabella 2.3) e l'elongazione a 72°C per 1 min 30 sec; uno step finale di elongazione a 72°C per 5 min.

Il DNA estratto da prodotto processato e' stato amplificato con le seguenti modificazioni al ciclo di reazione precedentemente descritto: il volume di DNA stampo utilizzato e' stato di 1  $\mu$ l del totale estratto e le condizioni di amplificazione sono state: un primo step di denaturazione del DNA a 94°C per 5 min; 40 cicli comprendenti la denaturazione a 94°C per 50 sec, l'*annealing* per 50 sec alla specifica  $T_a$  (Tabella 2.3) e l'elongazione a 72°C per 4 min; uno step finale di elongazione a 72°C per 9 min. I prodotti PCR sono stati separati mediante elettroforesi orizzontale in gel di agarosio per verificare la presenza dei prodotti di amplificazione e la corrispondenza delle loro dimensioni con i pesi molecolari attesi.

Tabella 2.3: Caratteristiche dei loci SSR utilizzati: nome del locus SSR; tipo di ripetizione del core; dimensione attesa dei prodotti in bp; numero di alleli attesi; indice di diversità (D); sequenza 5'-3' dei primer forward e reverse; tipo di marcatura (solo al 5' del primer forward); temperatura di *annealing*.

| SSR         | Tipo di Ripetizione                      | Dimensione dei Prodotti (bp) | Num. di Alleli | Indice di Diversità (D) | Sequenza dei Primers (5'-3')<br>Forward e Reverse       | Marcatura<br>(fw-5') | Ta °C |
|-------------|--|------------------------------|----------------|-------------------------|---|----------------------|-------|
| LE 20592*   | (TAT) <sub>15-1</sub> (TGT) <sub>4</sub> | 158-167                      | 4              | 0,58                    | CTGTTTACTTCAAGAAGGCTG<br>ACTTTAACTTTATTATTGCCACG        | 6-FAM                | 55    |
| LE 21085*   | (TA) <sub>2</sub> (TAT) <sub>9-1</sub>   | 98-113                       | 4              | 0,36                    | CATTTTATCATTATTTGTGTCTTG<br>ACAAAAAAGGTGACGATACA        | 5HEX                 | 50    |
| LE at002**  | (AT) <sub>9</sub>                        | 203                          | 2              | 0,5                     | ACTGCATTTTCAGGTACATACTCTC<br>ATAAACTCGTAGACCATACCCTC    | 6-FAM                | 57    |
| LE aat002** | (AAT) <sub>12</sub>                      | 106                          | 3              | 0,55                    | GCGAAGAAGATGAGTCTAGAGCATAG<br>CTCTCTCCCATGAGTTCTCCTCTTC | 6-FAM                | 55    |
| LE aat007** | (AAT) <sub>12</sub>                      | 100                          | 3              | 0,52                    | CAACAGCATAGTGGAGGAGG<br>TACATTCTCTCTCTCCCATGAG          | VIC                  | 55    |
| LE caa001** | (CAA) <sub>7</sub>                       | 105                          | 2              | 0,33                    | AGAAGGCGTGAGAGGCAAC<br>CTTAGCACTTGATGTTGATTGG           | 6-FAM                | 52    |
| LE ct001**  | (CT) <sub>12</sub>                       | 111                          | 3              | 0,35                    | TCCAATTCAGTAAGGACCCCTC<br>CCGAAAACCTTTGCTACAGAGTAGA     | NED                  | 52    |
| LE ct001**  | (CTT) <sub>9</sub>                       | 101                          | 2              | 0,39                    | CCTCTTTCACCTCTTTACAATTTCC<br>CACTGGTCATTAGTCTACAGCC     | PET                  | 57    |
| LE EF1Aa*   | (TA) <sub>8</sub> (ATA) <sub>9</sub>     | 198-213                      | 7              | 0,67                    | AAATAATTAGCTTGCCAATTG<br>CTGAAAGCAGCAACAGTATTT          | 5HEX                 | 50    |
| LE ga003**  | (GA) <sub>20</sub>                       | 241                          | 4              | 0,58                    | TTCGGTTTATTCTGCCAACC<br>GCCTGTAGGATTTTCGCCTA            | VIC                  | 52    |
| LELE25*     | (TA) <sub>11</sub>                       | 211-217                      | 4              | 0,36                    | TTCTCCGTATGAGTGAGT<br>CTCTATTACTTATTATTATCG             | 6-FAM                | 50    |
| LE ta003**  | (TA) <sub>9</sub>                        | 111                          | 4              | 0,43                    | GCTCTGCTTACAAATGATACCTCC<br>CAATGCTGGGACAGAAGATTTAATG   | VIC                  | 52    |
| LE ta015**  | (TA) <sub>15</sub>                       | 107                          | 2              | 0,49                    | ATATGCATGGACAAATCTTGAGGG<br>CTCGGCATCAAATTAATGATCAG     | PET                  | 55    |
| LE tat002** | (TAT) <sub>12</sub>                      | 196                          | 3              | 0,58                    | ACGCTTGGCTGCCTCGGA<br>AACTTTATTATTGCCACGTAGTCATGA       | NED                  | 55    |

\* Smulders et al., 1997

\*\* He et al., 2003

## 2.5.2 Elettroforesi capillare degli amplificati SSR mediante ABI PRISM 3100-AVANT

Per la discriminazione allelica gli amplificati fluorescenti sono stati analizzati mediante il sistema di elettroforesi capillare automatizzato ABI PRISM 3100- AVANT (Applied Biosystems), con capillari in silice fusa di 36 cm di lunghezza riempiti con il polimero POP4 (Applied Biosystems). Tutti i frammenti passano attraverso una finestra di lettura che converte l'informazione fluorescente in informazione elettronica che a sua volta è trasferita al computer per essere processata attraverso i programmi "ABI PRISM GeneScan" e "Genotyper 2.5" (v. 3.7) (Applied Biosystems). Le fluorescenze usate sono quelle riportate in tabella 2.3. Per ogni campione sono state effettuate opportune diluizioni delle reazioni di amplificazione e ad 1 µl di ogni diluizione sono stati aggiunti 0,25 µl di standard GeneScan 500 Liz (Applied Biosystems) per le marcature FAM, VIC, NED e PET, 0,25 µl di standard GeneScan 400 HD Rox (Applied Biosystems) per la marcatura HEX e 9,25 µl di formammide. I

campioni sono stati denaturati a 95°C per 5 min e immediatamente posti in ghiaccio per altri 2 min. Le condizioni di corsa elettroforetica sono state 15kV alla temperatura di 60°C per una durata di 45 min. L'altezza del segnale e la dimensione allelica del picco vengono calcolate mediante il software "ABI PRISM Genotyper 2.5" (v. 3.7) (Applied Biosystems).

## 2.6 Analisi CAPS

### 2.6.1 Condizioni di amplificazioni di loci EST-SNP

Sono stati selezionati dalla letteratura 21 EST-SNP di *S. lycopersicum* (Yang et al., 2004) e l'approccio adoperato è stato quello dell'analisi CAPS (Cleaved Amplified Polymorphic Sequence). Le condizioni di amplificazione PCR sono state le stesse di quelle adoperate per amplificazione dei loci SSR, precedentemente descritta nel paragrafo 2.5.1, utilizzando i primer riportati in tabella 2.4.

Tabella 2.4: Loci EST-SNP selezionati per analisi CAPS e loro principali caratteristiche.

| SNP           | Sequenza dei Primer (5'-3')<br>Forward e Reverse | Dimensione frammenti attesa<br>(bp) | Enzima di restrizione | EST di riferimento |     |           |    |
|---------------|--|-------------------------------------|-----------------------|--------------------|-----|-----------|----|
| LEOH2         | CTTGAAGATGGCCGAACACT<br>CTGGTCTGGGGGAATACCTT     | 201                                 | <i>BsaWI</i>          | EST478594          | G   | EST253241 | A* |
| LEOH8.4       | CCACTGATCAATGTGGTGA<br>CAACCACAAATGGCTCCTAAA     | 208                                 | <i>HaeIII</i>         | EST252213          | C*  | EST307657 | T  |
| LEOH11.1      | TATGTTGCTGCCAGACTCA<br>ACATCATGACCAACCATTCA      | 189                                 | <i>HhaI</i>           | EST328328          | T   | EST307804 | C* |
| LEOH12.2      | CCAGATGGGAGATGGGTCTA<br>CAGCAGTAACACCAGGAGCA     | 181                                 | <i>MseI</i>           | EST549757          | T*  | EST308207 | C  |
| LEOH16.2-16.3 | TCGACGCTGCACAGAAATAC<br>TTCTCCTCCTTATCTCCTTCA    | 206                                 | <i>BsaWI</i>          | EST301659          | C*  | EST285093 | T  |
| LEOH20.1      | CAGACCTAACAAGACAGGCAA<br>ATCAGGCATGACCATGGAAG    | 246                                 | <i>HaeIII</i>         | EST327354          | A   | EST284995 | G* |
| LEOH23.1      | GAGAGAAAAAGGGCACAAAGG<br>ACCGACAAACGCATAGATCA    | 191                                 | <i>MspI</i>           | EST546919          | del | EST256088 | G* |
| LEOH25.1      | GGAGGAAATAGGGTTCTAGGG<br>AATGGCCTGGCTAATCTGTG    | 325                                 | <i>HincII</i>         | EST511738          | C*  | EST285647 | G  |
| LEOH28        | CGCGGAGTTCTGTTAGCTTC<br>GCCGACGAATTACGAACATC     | 223                                 | <i>TaqI</i>           | EST273353          | A*  | EST253877 | T  |
| LEOH31.3      | TTGCAATGGCTTCTCCTC<br>ACTTGTCCGTTTCTCGCTTG       | 238                                 | <i>MspI</i>           | EST583372          | G*  | EST308897 | A  |
| LEOH31.4      | TTGCAATGGCTTCTCCTC<br>ACTTGTCCGTTTCTCGCTTG       | 238                                 | <i>MseI</i>           | EST583372          | G*  | EST308897 | A  |
| LEOH32.1      | TGGTGTGGATCCTGCTGTTA<br>TGGAATCACACCAAAACGA      | 293                                 | <i>HaeIII</i>         | EST358606          | C*  | EST256921 | T  |
| LEOH35.5      | CATCAGCCTCGCTCTTCT<br>CAAACCTGCAAGCCATTGAA       | 480                                 | <i>TaqI</i>           | EST549543          | G*  | EST280079 | A  |
| LEOH 31.6     | TGTTGATGTCTGGTCCATTCT<br>CCCTGCCCAAACATCTAAAA    | 197                                 | <i>TaqI</i>           | EST583372          | C*  | EST308897 | T  |
| LEOH 31.7     | TGTTGATGTCTGGTCCATTCT<br>CCCTGCCCAAACATCTAAAA    | 197                                 | <i>MseI</i>           | EST583372          | A   | EST308897 | T* |
| LEOH 31.10    | TGTTGATGTCTGGTCCATTCT<br>CCCTGCCCAAACATCTAAAA    | 197                                 | <i>AluI</i>           | EST583372          | G*  | EST308897 | T  |
| LEOH39        | AGAGAGTGGTCAAGTTAG<br>CATAGGCACAGTAATGAGAT       | 191                                 | <i>EcoRII</i>         | EST552218          | C*  | EST280326 | T  |
| LEOH 63       | CCCCTCGTGTAGGTGTCACT<br>ATCCTCCGATCATCAGCAGT     | 199                                 | <i>TaqI</i>           | BT013861           | C*  | -         | -  |
| CosOH73       | CTTCCCGACAAGCACAAAA<br>CGAATGCTCTGTACCATTTC      | 189                                 | <i>AluI</i>           | -                  | -   | -         | -  |

Note: l'asterisco indica la base del sito di taglio riconosciuto dallo specifico enzima di restrizione.

I frammenti di amplificazione sono stati digeriti mediante gli specifici enzimi di restrizione riportati in tabella 2.4. Le reazioni di digestione sono state allestite in un volume totale di 20 µl utilizzando 5 µl di prodotto di amplificazione, 1 X buffer, 4 U dello specifico enzima di restrizione ed 1/10 del volume finale di BSA. Le reazioni sono state, quindi, incubate alla rispettiva temperatura ottimale e per il tempo richiesto indicati per ogni enzima. I frammenti sono stati, poi, separati mediante elettroforesi orizzontale in gel di agarosio al 3% (Sambrook et al., 1989) ed i loro rispettivi pesi molecolari sono stati calcolati sulla base di una curva standard

costruita sulle ultime dieci bande (da 2000 a 100 bp) del “1 Kb plus DNA ladder” (Invitrogen) mediante il “Molecular Imaging Software v. 4.0.3” (Kodak).

## 2.6.2 Sequenziamento ed analisi dei prodotti PCR per la validazione degli SNP

L'identità degli amplificati e' stata effettuata tramite sequenziamento e tale verifica è avvenuta per almeno due genotipi per ogni SNP. La reazione di sequenza è stata effettuata utilizzando il “Big Dye RR terminator v1.1” per frammenti fino a 150 bp e “Big Dye RR terminator v3.1” (Applied Biosystems) per frammenti compresi tra 500 e 700 bp. A 30 ng di prodotto di amplificazione sono stati aggiunti 3,2 pmoli del primer e il buffer 1X RR (Applied Biosystems), secondo quanto riportato dal produttore, in una reazione dal volume finale di 10 µl. Per le reazioni di sequenziamento del DNA i primer utilizzati sono gli stessi di quelli impiegati per le amplificazioni PCR del locus di interesse, in reazioni allestite separatamente, l'una che utilizza il forward e l'altra il reverse. La reazione di sequenza ha incluso una fase iniziale di 96°C per 5 min seguita da 25 cicli di amplificazione costituiti da uno step a 96°C per 10 sec, uno step alla relativa temperatura di *annealing* per 5 sec e una fase di polimerizzazione a 60°C per 4 min. I prodotti sono stati purificati mediante precipitazione in 2,5 volumi di 95% etanolo e 1/10 del volume dei campioni di 3 M sodio acetato (pH 5,2), incubazione a -80°C per 10 min, centrifugazione per 2 min a 14000 rpm, eliminazione del supernatante, lavaggio del precipitato con 70% etanolo e solubilizzazione in 10 µl di formammide Hi-Di (Applied Biosystems). I campioni sono stati denaturati a 96°C per 5 min e successivo rapido abbassamento della temperatura a 4°C e, quindi, analizzati al sequenziatore (AB 3130; Applied Biosystems). Le sequenze ottenute sono state analizzate mediante il programma “Chromas Lite 2.01” (Technelysium Pty Ltd) e per ognuna delle sequenze relative ad un locus SNP sono stati generati contigs per assemblaggio delle sequenze ottenute con il forward ed il reverse mediante il programma “CodonCode Aligner v3.0.1” ([www.codoncode.com](http://www.codoncode.com)). Le sequenze sono state, poi, allineate e confrontate con quelle corrispondenti EST-SNP selezionate dalla banca dati mediante il programma “ClustalW 2.0.11” (Thompson et al., 1994).

## 2.7 Analisi di marcatori in/del

Sono stati selezionati dalla letteratura 6 marcatori in/del (Tabella 2.5) tra quelli disponibili sul sito [www.tomatomap.net](http://www.tomatomap.net) (Francis et al., 2006) ed utilizzati per la genotipizzazione del materiale in esame.

Tabella 2.5: Loci in/del selezionati, cromosoma di appartenenza, metodo di rilevazione e dimensioni dei prodotti.

| Locus in/del | Cromosoma | Metodo di rilevazione <sup>a</sup> | Dimensione prodotti |
|--------------|-----------|------------------------------------|---------------------|
| CT10042      | 3         | 2%                                 | 126/150bp           |
| CT10184      | 4         | 4%                                 | 149/161bp           |
| CT10373      | 5         | 2%                                 | 192/270bp           |
| CT10737      | 11        | 4%                                 | 182/195bp           |
| CT20210      | 5         | 4%                                 | 176/194bp           |
| CT20268      | 1         | 4%                                 | 253/268bp           |

<sup>a</sup> percentuale di agarosio utilizzato per le analisi in elettroforesi orizzontale.

Le condizioni di amplificazione PCR sono le stesse di quelle precedentemente descritte nel paragrafo 2.5.1. La separazione dei prodotti di amplificazione e' avvenuta mediante elettroforesi orizzontale in gel alle percentuali di agarosio indicate in tabella 2.5 ed i loro rispettivi pesi molecolari sono stati calcolati sulla base di una curva standard costruita sulle ultime dieci bande (da 2000 a 100 bp) del "1 Kb plus DNA ladder" (Invitrogen) mediante il "Molecular Imaging Software v. 4.0.3 (Kodak).

## **2.8 Analisi molecolari mediante SNP**

### **2.8.1 Condizioni di amplificazione PCR dei frammenti contenenti gli SNP e Estensione del Primer Allele-Specifica (ASPE)**

Sono stati selezionati 71 SNP tra quelli disponibili sul sito "Tomato Mapping Resource Database" ([www.tomatomap.net](http://www.tomatomap.net)) (Francis et al., 2006). Ogni SNP e' localizzato in un amplicone di dimensioni molecolari comprese nel range da 300 a 1000 bp (Tabella 2.6, pagina successiva) e sono stati selezionati sulla base di quelli che erano gia' stati utilizzati per la genotipizzazione della "Core Collection" dell'Ohio State University (Francis et al., 2006).

---

Tabella 2.6 Loci SNP analizzati; nucleotidi associati allo SNP; sequenze dei primer forward; temperatura *melting* dei primer forward in gradi Celsius; sequenze dei primer reverse; temperatura *melting* dei primer reverse in gradi centigradi, peso molecolare dei prodotti di amplificazione in bp; numero indicativo della microsfera 1; sequenza del primer 1 per ASPE; numero indicativo della microsfera 2; sequenza del primer 2 per ASPE; direzione: forward o reverse.

| Locus   | SNP | Primer fw              | TmF  | Primer rv               | TmR  | PM   | m.1 | ASPE primer 1            | m.2 | ASPE primer 2            | D. |
|---------|-----|------------------------|------|-------------------------|------|------|-----|--------------------------|-----|--------------------------|----|
| CT10004 | AG  | CACGACATTAGCTGTTGGGTT  | 60,0 | TGGAGCTGGAATCAGTGTG     | 59,8 | 820  | 33  | TTCCAGGTCACCTTGAATGT     | 34  | TTCCAGGTCACCTTGAATGT     | R  |
| CT10018 | AG  | CCAATTCCTCTCTGTGATGC   | 60,6 | CTTCATACCACCTTGTCCA     | 59,8 | 1383 | 74  | CCTTTGTCTCCTTTGAAGAT     | 75  | CCTTTGTCTCCTTTGAAGAT     | R  |
| CT10024 | G   | GCTCATGTCAATCTTCTCTGG  | 58,9 | GAGCTGAAGACCTTCTCGC     | 59,3 | 307  | 72  | TGGCGTATGTGACATCTAG      | 73  | TGGCGTATGTGACATCTAG      | F  |
| CT10043 | AT  | CTGTGGAAGCTACTGCTAGGC  | 59,3 | CCCTCTATGCTACCAATCAA    | 59,1 | 869  | 35  | GAACCTTAAAGTCGTGTCACA    | 36  | GAACCTTAAAGTCGTGTCACA    | F  |
| CT10050 | G   | AAGTGGTGTGATGTAATGGTC  | 59,7 | TTCATGGCTGCATGTACGG     | 59,7 | 964  | 35  | CCCACCCATGGGATTACA       | 36  | CCCACCCATGGGATTACA       | F  |
| CT10075 | C   | GGGCTCTAGTTTACCCAGC    | 60,7 | TTCGACAGAGAAGCAAGGA     | 60,4 | 753  | 17  | GCTAAGTTGCTTGGTCC        | 18  | GCTAAGTTGCTTGGTCC        | F  |
| CT10102 | AT  | TTTGTCCGCTCATCTTCCA    | 60,3 | AGCAACGCAACCAATTCTG     | 59,9 | 420  | 17  | GCTTCGCATTGAGTCATACA     | 18  | GCTTCGCATTGAGTCATACA     | R  |
| CT10120 | AG  | AGCCGACTTTCTATTGCTGG   | 59,5 | GGAGATCACATTCTCCTCCCT   | 59,5 | 1242 | 42  | GCAAAATTTATCTGGAGGGTACA  | 43  | GCAAAATTTATCTGGAGGGTACA  | F  |
| CT10136 | AG  | GTCCTCGCCCTATCTGCTTTG  | 60,0 | GCTTCTCAAACCTCTCCCTT    | 59,0 | 1261 | 42  | GTTAGGCCATCCTTTACCC      | 43  | GTTAGGCCATCCTTTACCC      | F  |
| CT10138 | C   | ATGGTGCCTCTCAAGGTTCT   | 60,1 | CAGCGACATACCCCTCCAAA    | 59,7 | 1277 | 42  | GGTTGATTTTCAGCCCTTC      | 43  | GGTTGATTTTCAGCCCTTC      | F  |
| CT10153 | G   | GTCGAATCAAAGCCTCTCTC   | 58,9 | CAAGGTTTTCACACGTTTCG    | 60,2 | 1422 | 44  | AACTCCCAAGTCCAATAAGG     | 45  | AACTCCCAAGTCCAATAAGG     | R  |
| CT10162 | AG  | AGCTCCCTCAACAGGGTTG    | 59,7 | CAATAACCGTCTCCCTCTT     | 58,0 | 1242 | 46  | CCCTGGCTAGAAATCGAT       | 47  | CCCTGGCTAGAAATCGAT       | R  |
| CT10184 | C   | GATGGAGATGAACCGGAGA    | 60,2 | CACAATGCTTCAATCCAAAG    | 60,1 | 1022 | 51  | GTTTCGATTGAACCCGTAC      | 52  | GTTTCGATTGAACCCGTAC      | F  |
| CT10187 | AG  | GTGAATAATGGCGGTGAG     | 59,9 | AGATTATGCCCTCTCCGATGT   | 59,9 | 1120 | 53  | TGTACAAAGAGGACCAAT       | 54  | TGTACAAAGAGGACCAAT       | R  |
| CT10190 | AG  | GAACAATCAAGACGGGAAAC   | 59,7 | TCAAGGAAACCAAGTCCATTC   | 59,9 | 894  | 21  | TAATTTCTCTTCGAGCCA       | 24  | TAATTTCTCTTCGAGCCA       | F  |
| CT10192 | AG  | CATCACAATAACCAAGCTGC   | 60,5 | TCTTCAGGATCTCAGCTTTCC   | 60,0 | 1666 | 41  | AGTTAAAGTAGTACCCTT       | 11  | AGTTAAAGTAGTACCCTT       | R  |
| CT10205 | AG  | GGGAGCTTCTTTGTTGATG    | 59,7 | AAACGGCAACTGGACAA       | 60,1 | 962  | 44  | TAGCAGCATCTACACGCC       | 45  | TAGCAGCATCTACACGCC       | F  |
| CT10206 | AG  | GGTAATAATGGCGGTGAG     | 60,0 | GATGGCAGGTGATTCTTGATG   | 60,5 | 1125 | 42  | GATACCTGAGTTACCAAGCT     | 43  | GATACCTGAGTTACCAAGCT     | R  |
| CT10215 | C   | TCCCATACTGCGGTTACT     | 59,6 | GAAACTGGCATCCACATGAA    | 59,5 | 1219 | 61  | CTTGGGTGGGATAACTG        | 62  | CTTGGGTGGGATAACTG        | R  |
| CT10238 | AG  | GGCGTAGCGGTTGCTCT      | 61,2 | TCATACCAATCCCAACCAAGC   | 60,7 | 975  | 63  | ACCATAGGCTCTTAGGCTAT     | 64  | ACCATAGGCTCTTAGGCTAT     | R  |
| CT10242 | AT  | AGGAGTTGGACGTGTGAAC    | 60,0 | CCAGCAGTGATAGCGTTGG     | 60,4 | 879  | 72  | TCGAGCTAACAAAACACAGT     | 73  | TCGAGCTAACAAAACACAGT     | R  |
| CT10322 | AC  | ACAGCAAGGAGCGTTTCAAG   | 60,1 | GGCCCAATAATGCAATGTG     | 61,1 | 1254 | 76  | AGGAAAAAGCAACCCAGAA      | 77  | AGGAAAAAGCAACCCAGAA      | F  |
| CT10346 | AT  | TTCTCTCTCCACCAAAAG     | 59,3 | TCCCTTCTCTCCATTTACAG    | 60,2 | 1298 | 21  | AAACGTACACTACAGAAACAAT   | 24  | AAACGTACACTACAGAAACAAT   | R  |
| CT10347 | C   | GGGATGCAACTGCTTGTGT    | 59,7 | GCAAGATCCCGATCTCTAAGT   | 58,9 | 1441 | 25  | GCCAGCCAATAAGAGCCA       | 26  | GCCAGCCAATAAGAGCCA       | R  |
| CT10351 | AT  | CTGTGCTGCTAGTCTACTGCT  | 60,0 | AACTAGGGCAACCTTCCCA     | 59,6 | 1135 | 27  | AAACCCAAACCCCAACAGA      | 28  | AAACCCAAACCCCAACAGA      | F  |
| CT10373 | C   | CTCGCCAGTCTGTTATGTC    | 60,7 | TGCTCAAGTCGAAATCCATC    | 60,1 | 1579 | 19  | GCTACAGGGGTGGAATCC       | 20  | GCTACAGGGGTGGAATCC       | F  |
| CT10375 | AT  | TTGCTGATATTTGGGCTCT    | 59,7 | GTCAAATAATGACCGAAGGC    | 59,7 | 1586 | 51  | GGTGATATCCCATTTGTGT      | 52  | GGTGATATCCCATTTGTGT      | R  |
| CT10377 | AG  | CATGGATTGACCTTCATCG    | 58,9 | TCAAAGTGGCGTCTAGTCTTC   | 59,4 | 1460 | 35  | TTGTTGTGCTTTCAGCTT       | 36  | TTGTTGTGCTTTCAGCTT       | R  |
| CT10385 | AG  | AGCTTTGTTAGTTGGGCTG    | 59,4 | TGCTAATCCAAACCCACTGA    | 59,1 | 929  | 76  | CTATTTCCGGTCTGCTCATCA    | 77  | CTATTTCCGGTCTGCTCATCA    | F  |
| CT10386 | G   | CACCTGCAACCACTCCCTCTG  | 59,5 | CTCCTTAGGTGGCGCTTC      | 60,0 | 1461 | 21  | GAATGGATGCTCTGTATGAGG    | 24  | GAATGGATGCTCTGTATGAGG    | F  |
| CT10390 | AG  | GCCCAATAGTTCACAGCTC    | 61,0 | TGTCGTTGGTATTAAGCGCA    | 60,3 | 1729 | 37  | AGCCAAATTCATCTCACCA      | 38  | AGCCAAATTCATCTCACCA      | F  |
| CT10396 | AG  | GCGATTGGTAGAAGCAGATG   | 60,6 | GCAAGTTGGGAACCTGACTTT   | 59,5 | 1461 | 25  | TCTGTACGATATGCTTGTCT     | 26  | TCTGTACGATATGCTTGTCT     | R  |
| CT10401 | AT  | GGGAGACTCATCTGATTGT    | 59,6 | ATTTGTTCTCCCAATGCTG     | 59,9 | 1090 | 46  | ATGAGATATGAGCCCAACTA     | 47  | ATGAGATATGAGCCCAACTA     | R  |
| CT10432 | AC  | GAAACAGGCAAGCAAGGTG    | 60,5 | CACAGGTCACAGCAATGCTC    | 60,2 | 691  | 27  | CCTTCACTCCGCGAAGAT       | 28  | CCTTCACTCCGCGAAGAT       | R  |
| CT10437 | AT  | ACTTAGAGCTTCAGCCACCG   | 59,6 | ACTTCTTCTCTGCTGCTCT     | 60,5 | 826  | 29  | GACCCAAAGCAACTTTGGA      | 32  | GACCCAAAGCAACTTTGGA      | F  |
| CT10450 | C   | CGAATCATGTCATCTCTC     | 60,0 | GCAAGAAAACCCCACTTCA     | 60,4 | 476  | 12  | GAAGAAAGGTGATAGAGGCG     | 13  | GAAGAAAGGTGATAGAGGCG     | R  |
| CT10469 | A   | TGTCGATTTCTGTTTCACT    | 58,1 | CTAACTACCACCTCTTAAACCAC | 58,8 | 984  | 14  | CATTCCTTGAAGAACGATAA     | 6   | CATTCCTTGAAGAACGATAA     | F  |
| CT10471 | C   | CATGCTGCAACTCTCCATCT   | 60,7 | GCCCTGACACTGCTGGTT      | 59,8 | 1404 | 44  | CACATGGAGTGTATGTCATG     | 45  | CACATGGAGTGTATGTCATG     | R  |
| CT10485 | AC  | AGCACTCGGTGAGCTGTTAG   | 59,7 | ATCCTCATCTCCAGTGCAT     | 59,5 | 1100 | 53  | GTGGCTGAAAGCGGTATCA      | 54  | GTGGCTGAAAGCGGTATCA      | F  |
| CT10494 | AG  | GGCCTAATTTGCTTCAAGAC   | 60,0 | GGCTCATGTCGACTGATTG     | 59,0 | 1740 | 42  | AGAAGTGGAGGTGTAAGAACT    | 43  | AGAAGTGGAGGTGTAAGAACT    | R  |
| CT10526 | C   | GAACTGTCTACCACTTGAAGGG | 58,3 | AGGTCGGAGTTTCCAGAAGA    | 58,9 | 706  | 46  | TAGCTCAGGTAAGCTAGAAC     | 47  | TAGCTCAGGTAAGCTAGAAC     | F  |
| CT10533 | AC  | TCGTGCGAAGCACTGTTACT   | 59,7 | GTCCCTCCAGTAATAACTGCGAC | 57,9 | 1640 | 46  | TGTGATCGGAAGCATAGA       | 47  | TGTGATCGGAAGCATAGA       | F  |
| CT10535 | AC  | CTAGCAAGTGGCAGCATGG    | 60,4 | CCTCCGCAACAGGATACAAT    | 60,1 | 1009 | 51  | CACATAGGAGAGGTAACCT      | 52  | CACATAGGAGAGGTAACCT      | R  |
| CT10554 | AG  | AGAGAAGCATTGGCCCTT     | 59,8 | AGCAGAACGAAAGATCGC      | 59,7 | 769  | 55  | ACAGCAATAAGTTTAGGAAAGA   | 56  | ACAGCAATAAGTTTAGGAAAGA   | F  |
| CT10556 | AG  | CACAAGTGGTCTGAGGAAAT   | 60,2 | TGAGGGCGAATGAAGAAAT     | 59,8 | 659  | 61  | CCTTCTCTCTCGAAGAAC       | 62  | CCTTCTCTCTCGAAGAAC       | F  |
| CT10591 | AT  | CTGATCTCATTGCCCTTAAACC | 60,0 | ACTCCAGTATGACCGCTCCA    | 59,7 | 447  | 55  | CAGCTAGTCTTCAAGTCT       | 56  | CAGCTAGTCTTCAAGTCT       | R  |
| CT10639 | AG  | GGTCTCACAATGTGCTTACTG  | 57,4 | TTCGCAATCAGTTTAGGGTC    | 57,8 | 1609 | 63  | AGAATTTTGACGAGGCTGTA     | 64  | AGAATTTTGACGAGGCTGTA     | F  |
| CT10678 | G   | GAACGGAGACATTGAAGTGC   | 58,3 | AACTGGATCTTCTGCCAACA    | 58,3 | 1454 | 65  | TATGGTTGAAACCCAGGCTTG    | 66  | TATGGTTGAAACCCAGGCTTG    | F  |
| CT10715 | C   | CATTGGGACTGCTTGAC      | 58,7 | GAGACAATCCCGAATCCA      | 57,5 | 1038 | 53  | AGTACTGTAACAGCCOAG       | 54  | AGTACTGTAACAGCCOAG       | R  |
| CT10736 | C   | TGTAATGCAATGGCTCATC    | 59,7 | AGACCAGATCCAGAGGGA      | 59,6 | 667  | 76  | TGTTTACAGAAAGTGGGGC      | 77  | TGTTTACAGAAAGTGGGGC      | F  |
| CT10768 | C   | TTTGGCTCCAGAACGAGG     | 59,9 | GGACGATAATGATGATCTG     | 59,2 | 585  | 17  | CCAGATGATTGGAAGTTGC      | 18  | CCAGATGATTGGAAGTTGC      | F  |
| CT10778 | AG  | TTTGGTCTCAGGCTATTGG    | 60,1 | CAGACGAATCTACCGCTTC     | 59,8 | 486  | 74  | TTTGAATCGAGGTAACACCA     | 75  | TTTGAATCGAGGTAACACCA     | F  |
| CT10781 | G   | CTGCAACTGGGAATGAGGTC   | 60,7 | ACACCAGACTAGACCGCTTGA   | 59,9 | 1193 | 21  | CTTACTCAGAGGCCOACC       | 24  | CTTACTCAGAGGCCOACC       | R  |
| CT10790 | AG  | CCCTCCGATAGCAACATCAC   | 60,5 | TGCACCCTTCTGGGCTCA      | 59,3 | 643  | 55  | CTCCAGATATATITCAGCAAT    | 56  | CTCCAGATATATITCAGCAAT    | R  |
| CT10793 | C   | CTTGCCACTCCCAATGACGG   | 60,8 | CTAGCTGTTCCGATCATCG     | 60,6 | 277  | 25  | TCCCTACTGTTGAGGAAAC      | 26  | TCCCTACTGTTGAGGAAAC      | F  |
| CT10801 | C   | CAGTTTCTCAGTGTGGAGC    | 59,8 | TCTCTTTCAGACATCTCGCT    | 60,0 | 655  | 27  | AAOCTCATATCGGGTCTATC     | 28  | AAOCTCATATCGGGTCTATC     | F  |
| CT10807 | AT  | CACGACATTAAGCTGTTGGTT  | 60,3 | TGGAGCTGGAATCAGTGTG     | 60,4 | 936  | 35  | TTGGTCTCAATTTCAAAGCTA    | 36  | TTGGTCTCAATTTCAAAGCTA    | R  |
| CT10809 | AG  | ACCAATGTTGCAAAGGGTGT   | 57,7 | CATCAGACGTTCAATCATCG    | 57,6 | 897  | 29  | GAGCAAGAGGTGAGAAAA       | 32  | GAGCAAGAGGTGAGAAAA       | F  |
| CT10909 | AG  | AAGGATTGTGCGCTTACAT    | 59,6 | AGCATCTAATGGAGCTGGTTC   | 58,4 | 662  | 72  | GGAAACATTATTACAAGTCTCAGT | 73  | GGAAACATTATTACAAGTCTCAGT | R  |
| CT10915 | C   | GGCAAGGAAAGAAATCCCAT   | 60,3 | GAGTTCCACCATGAGACAGGC   | 59,7 | 895  | 25  | ACAACCTTTCCGAGACTGAC     | 26  | ACAACCTTTCCGAGACTGAC     | F  |
| CT10923 | C   | CACACAACCTGATGATTGGC   | 60,0 | GTGACATGCTCCATCACGAC    | 60,1 | 1520 | 42  | CAGAGACATGATGATGATC      | 43  | CAGAGACATGATGATGATC      | F  |
| CT10963 | AG  | TTGCCAAGTAATGACAGAGA   | 58,3 | GAAGATACGCGAAAACAATGG   | 58,2 | 581  | 29  | AACTGAAAAAGAGAAAAAGAAA   | 32  | AACTGAAAAAGAGAAAAAGAAA   | F  |
| CT20017 | C   | GTTGAAAGAGCCCAAGTTTGG  | 59,9 | CACCTCCAAGCTGTGCTTTC    | 61,0 | 1051 | 63  | AGGCATGCTCATAGAGAG       | 64  | AGGCATGCTCATAGAGAG       | R  |
| CT20023 | AT  | AGAGCTTACTGCGGCTGAAC   | 59,8 | GACCTCGTGAAGAGCATCC     | 59,8 | 625  | 65  | GACTGCTTCTTCTAAGCA       | 66  | GACTGCTTCTTCTAAGCA       | R  |
| CT20028 | AT  | CGTATTGAACCTTTCAAGGAG  | 59,8 | ATTGACGAGGGTACTGGCAT    | 59,4 | 1775 | 72  | CGAAAACCTACACCTTGCA      | 73  | CGAAAACCTACACCTTGCA      | R  |
| CT20051 | AG  | TGCAACCAACCAAGTATG     | 60,1 | AATGCTGCTTCCAGCTCT      | 60,1 | 883  | 41  | CACAAGTATTGGTTTATTGGTCA  | 41  | CACAAGTATTGGTTTATTGGTCA  | F  |
| CT20059 | AC  | CCAATTCCTCTGCTGATCG    | 59,1 | CTTCCATTACCACCTTGTCCA   | 59,7 | 1575 | 44  | GAACAGTATCTGTAGCAA       | 45  | GAACAGTATCTGTAGCAA       | F  |
| CT20097 | AT  | TCTGCAATGACCACTGGAAG   | 59,8 | GGATGATCTTTGGAACCAG     | 59,7 | 459  | 65  | CAGCTAGTCTTCAAGTCA       | 66  | CAGCTAGTCTTCAAGTCA       | F  |
| CT20127 | AG  | TGAATCAGCAGAAAGGAGGC   | 59,7 | GTTGCCAAGGTCATGAAAG     | 60,5 | 369  | 61  | GCTCAGTCTATGAATTGCT      | 62  | GCTCAGTCTATGAATTGCT      | R  |
| CT20159 | G   | CGTGCTCTGGAGAGAAACC    | 60,0 | CAAGATGTGGGAGTGAAGTGG   | 59,8 | 1444 | 25  | TGCAACCTTAGCTAGTGC       | 26  | TGCAACCTTAGCTAGTGC       | R  |
| CT20181 | AG  | TGCAATGCACTTATGCTTGG   | 59,8 | TTCAGGTGCAAGTGAAGCTCA   | 60,2 | 1150 | 12  | CGATATTGGTGGCATTGA       | 13  | CGATATTGGTGGCATTGA       | F  |

Lo SNP target e' stato pre-amplificato e tutte le amplificazioni PCR sono state allestite in un volume finale di 15  $\mu$ l contenente 3 ng di DNA genomico, 1X PCR buffer (Promega), 0,05 mM dNTP, 0,1  $\mu$ M di primer forward e 0,1  $\mu$ M di primer reverse, 0,5 U di *Taq* polimerasi (Promega). Le reazioni sono state eseguite in un termociclatore "DNA Engine Tetrad 2 Peltier Thermal Cycler" (Bio-Rad) con le seguenti condizioni: un ciclo iniziale di denaturazione a 94°C per 2 min; due cicli comprendenti ognuno uno step a 94°C per 30 sec, uno a 62°C per 30 sec e uno a 72°C per 1 min e 30 sec; sei cicli comprendenti ognuno uno step a 94°C per 30 sec, una temperatura di *annealing* in un range tra 61°C e 56°C per 30 sec che decresce ad ognuno dei sei cicli di 1°C, uno step a 72°C per 1 min e 30 sec; 31 cicli comprendenti ognuno uno step a 94°C per 30 sec, uno a 55°C per 30 sec e uno a 72°C per 1 min e 30 sec; ed un ciclo finale di estensione a 72°C per 7 min. Verificata la presenza dei prodotti di amplificazione attesi mediante elettroforesi in gel di agarosio, 4  $\mu$ l di ogni prodotto PCR per ogni genotipo sono stati raggruppati e precipitati mediante aggiunta di due volumi di soluzione di precipitazione contenente 90% etanolo e 2 M di 10% sodio acetato (pH 7,0), centrifugazione per 30 min a 14000 rpm, eliminazione del supernatante, lavaggio del precipitato mediante 70% etanolo, centrifugazione per 10 min a 14000 rpm e solubilizzazione del *pellet* mediante aggiunta di 8  $\mu$ l di H<sub>2</sub>O. Per la reazione di estensione del primer allele-specifica (ASPE), 4  $\mu$ l di tutti i prodotti PCR precipitati sono stati aggiunti a 10  $\mu$ l di una reazione contenente 1X ASPE buffer, 1,25 mM MgCl<sub>2</sub>, 100 mM dNTP (-dCTP) mix (Invitrogen), 5  $\mu$ M biotina-dCTP, 25 nM *tag*-ASPE primer mix e 0,038 U *Tsp* DNA polimerasi (Invitrogen). I primer per ASPE sono stati disegnati in modo da avere una sequenza *tag* di 21 nt all'estremita' 5' e includere lo SNP all'estremita' 3'. Ogni primer per ASPE e' stato sintetizzato per ognuna delle due varianti alleliche di quel dato SNP ed ogni sonda differisce nella scelta del *tag* e del nucleotide polimorfico presente all'estremita' 3'. Le condizioni di amplificazione per la reazione ASPE sono state le seguenti: un primo step di denaturazione del DNA a 96°C per 2 min; 30 cicli comprendenti la denaturazione a 94°C per 30 sec, l'*annealing* a 55°C per 1 min e l'elongazione a 74°C per 2 min; uno step finale a 4°C.

### 2.8.2 Ibridazione alle microsferi FlexMAP

Le microsferi sono internamente marcate con due fluorocromi diversi e pre-acoppiate con delle sonde anti-*tag*. La reazione di ibridazione e' stata allestita in un volume finale di 50  $\mu$ l contenente 40  $\mu$ l di microsferi ( $2,5 \times 10^5$  microsferi/ml) in una soluzione di ibridazione 2X (0,4 M NaCl, 0,2 M Tris (pH 8,0), 0,16% Triton X-100) e 10  $\mu$ l di prodotti di reazione ASPE. Dopo uno passaggio di denaturazione a 96°C per 1 min 30 sec, e' previsto un successivo step di ibridazione a 37°C per 1 h. Dopo centrifugazione per 5 min a 2250 g, il precipitato e' stato risospeso mediante aggiunta di 40  $\mu$ l di soluzione di ibridazione 1X e centrifugazione per 2 min a 2000 rpm. Prima dell'analisi mediante apparecchio "Luminex 200", alla reazione e' stata aggiunta streptavidina per una concentrazione di 4  $\mu$ g/ml e la reazione e' stata incubata a 37°C per 30 min.

### 2.8.3 Analisi a citometria a flusso mediante Luminex 200

La fluorescenza delle microsfere e' stata rilevata mediante tecnologia "Luminex's xMAP" utilizzando un apparecchio "Luminex 200" equipaggiato con un "Luminex Analyzer". La fluorescenza sulla superficie delle microsfere e' stata rilevata e convertita in un valore di intensita' di fluorescenza media (MFI). Mediante il programma "MasterPlexGT" e' stato possibile elaborare i dati e ottenere i valori relativi di fluorescenza, la percentuale di eterozigoti ottenuta, il numero delle microsfere che ha effettivamente ibridato ed il nucleotide associato allo SNP.

## 2.9 Analisi dei dati molecolari

### 2.9.1 Calcolo delle relazioni genetiche

I dati derivanti dai marcatori molecolari sono stati sottoposti ad analisi genetica e sono stati calcolati il numero effettivo di alleli,  $N_e = 1/\sum p_i^2$ ; l'eterozigosita' osservata,  $H_o = N_o$  di  $H_e/N$  dove  $N_o$  di  $H_e$  rappresenta il numero di eterozigoti ed  $N$  si riferisce al numero di campioni; il contenuto polimorfico informativo (equivalente all'eterozigosita' attesa,  $H_e$ )  $PIC = H_e - \sum p_i^2$ , dove  $p_i$  e' la frequenza dell'allele  $i^{th}$  al dato locus (Hartl e Clark, 1997). Le matrici delle distanze genetiche sono state utilizzate per le rappresentazioni grafiche. I dendrogrammi sono stati costruiti utilizzando l'algoritmo del metodo Neighbor-Joining (NJ) (Saitou e Nei, 1987) e del metodo UPGMA (Unweighted Pair Group Method with Arithmetic mean).

La correlazione statica tra le matrici e' stata valutata con il test di Mantel (Mantel, 1967) ed il coefficiente di correlazione lineare di Pearson ( $r$ ) e' stato usato per la relazione tra le variabili genetiche, la cui significativita' e' stata stabilita in seguito a 999 permutazioni casuali dei dati delle matrici. L'Analisi delle Coordinate Principali (PCA), eseguita dopo standardizzazione dei dati sulla matrice della covarianza ottenuta dalle similarita' genetiche, e' stata effettuata per mostrare la relazione genetica tra le accessioni in esame. Tutte le analisi sono state eseguite mediante i programmi GeneAlex 6 (Genetic Analysis in Excel) (Peakall et al., 2006), NTSYSpc v. 2.1 (Numerical Taxonomy System) (Rohlf, 1998) e MEGA v. 3.1 (Molecular Evolutionary Genetics Analysis) (Tamura et al., 2007).

### 2.9.2 Analisi statistica della struttura delle popolazioni

Possibili strutture di popolazione sono state investigate mediante l'utilizzo del *software* "Structure" v. 2.1 (Pritchard et al., 2000). Per marcatori co-dominanti e' stato applicato il metodo Bayesiano implementato nel *software* "Structure". L'algoritmo definisce il numero di popolazioni di riferimento, ciascuna caratterizzata da un insieme di frequenze alleliche e poi assegna ogni individuo in maniera probabilistica. Il risultato che si ottiene e' la probabilita' a posteriori che un individuo "i" appartenga alla popolazione K (sconosciuta) dato un genotipo X e stimate le frequenze alleliche di tutte le popolazioni. La probabilita' a priori, cioe' la probabilita' che un individuo appartenga ad una qualsiasi delle K popolazioni di riferimento, e' definita come  $1/K$ . Il *software* consente l'analisi dei dati provenienti da marcatori dominanti utilizzando il "*no admixture model*" (ogni individuo e' ipotizzato appartenere in maniera discreta ad una determinata popolazione) e permette di scegliere il numero di simulazioni (*burn-in period*) con il quale minimizzare gli effetti dati dalla configurazione casuale di partenza, e il numero di *routine* MCMC (Markov Chain Monte Carlo) per il calcolo

delle probabilità a posteriori  $\ln Pr(X/K)$ . Il modello assume la presenza dell'equilibrio Hardy-Weinberg e di Linkage Equilibrium tra i marcatori impiegati. Un set di 77 marcatori, comprendenti in/del e SNP (Tabelle 2.5 e 2.6, rispettivamente) e' stato utilizzato per l'analisi. La maggior parte dei parametri utilizzati rientrano nei valori di *default* stabiliti dal manuale del programma. Con il modello "no admixture" utilizzando un *burning period* di 50000 ed una *routine* MCMC di 500000 ripetizioni sono state effettuate 20 analisi per ogni valore di K (da 2 a 16) in esame. Il numero di popolazioni esistenti (K) e' stato identificato usando il massimo valore di  $\ln Pr(X/K)$  calcolato nel seguente modo: alla media del logaritmo della probabilita' dei dati ad ogni step di MCMC e' sottratta meta' della loro varianza (Pritchard et al., 2000).

## 2.10 Analisi morfologica

La caratterizzazione fenotipica e' stata condotta su sedici ecotipi forniti dalla ditta sementiera "La Semiorto Sementi" su almeno sei piante per ecotipo. Quindici caratteri morfometrici quantitativi e venti qualitativi sono stati analizzati in due anni consecutivi e si riferiscono a diversi stadi di sviluppo della pianta. Le tabelle 2.7 e 2.8 riportano i caratteri morfologici esaminati in funzione dello stadio e dell'organo considerato per l'analisi discriminante. Gli ecotipi sono stati sottoposti ad analisi di gruppo. Il coefficiente della distanza Euclidea e' stato utilizzato per il calcolo della matrice di dissimilarita' per i caratteri quantitativi ed il coefficiente di *Jaccard* per ottenere la matrice di similarita' per i caratteri qualitativi, indipendentemente. La matrice dei caratteri qualitativi e' stata convertita in una matrice di dissimilarita' moltiplicando per (-1,0) (Romesburg, 1984). Entrambe le matrici sono state standardizzate ed utilizzate, indipendentemente, per l'analisi cluster mediante il metodo UPGMA (Unweighted Pair-Group Method with Arithmetic Averages). I valori delle matrici cofenetiche e delle distanze genetiche (Sneath e Sokal, 1973) dei cluster UPGMA basati sui dati morfologici quantitativi e qualitativi sono stati anche utilizzati per calcolare il coefficiente di correlazione ( $r$ ) tra le matrici con 999 permutazioni (Mantel, 1967). Queste analisi sono state condotte mediante NTSYS-PC v. 2.1 (Rohlf, 1998).

Tabella 2.7: Lista dei caratteri morfologici quantitativi analizzati sugli ecotipi in esame e loro stadio o organo considerato.

| Carattere quantitativo | Stadio/organo considerato |
|------------------------|---------------------------|
| Altezza fusto          | plantula                  |
| Lunghezza foglia       | pre-trapianto             |
| Altezza pianta         | pre-trapianto             |
| Lunghezza foglia       | pianta                    |
| Larghezza foglia       | pianta                    |
| Lunghezza picciolo     | foglia                    |
| Altezza I palco        | fiore                     |
| Lunghezza internodo    | fiore                     |
| Lunghezza picciolo     | frutto                    |
| Numero frutti/palco    | frutto                    |
| Numero frutti/pianta   | frutto                    |
| Lunghezza frutto       | frutto                    |
| Diametro frutto        | frutto                    |
| Peso frutto            | frutto                    |
| Altezza pianta         | frutto                    |

Tabella 2.8: Lista dei caratteri morfologici qualitativi analizzati sugli ecotipi in esame e loro stadio o organo considerato.

| Carattere qualitativo   | Stadio/organo considerato |
|-------------------------|---------------------------|
| Vigore                  | plantula                  |
| Vigore pianta           | pre-trapianto             |
| Colorazione antocianina | pianta-fusto              |
| Rami avventizi laterali | pianta-fusto              |
| Rami avventizi basali   | pianta-fusto              |
| Colore                  | foglia                    |
| Bollosità               | foglia                    |
| Scanalatura picciolo    | foglia                    |
| Colore nervature        | foglia                    |
| Portamento              | foglia                    |
| Precocità               | fiore                     |
| Frutti al I palco       | frutto                    |
| Copertura fogliare      | frutto                    |
| Bicchiere               | frutto                    |
| Punta o goccia          | frutto                    |
| Attacco pistillare      | frutto                    |
| Colore                  | frutto                    |
| Spalla verde            | frutto                    |
| Costolatura             | frutto                    |
| Colore                  | frutto                    |

### 3. RISULTATI

#### 3.1 Valorizzazione e protezione di ecotipi di pomodoro campani

##### 3.1.1 Analisi molecolare mediante minisatellite (GATA)<sub>4</sub>

Gli ecotipi campani sono stati sottoposti a caratterizzazione molecolare mediante una analisi di omogeneità utilizzando il minisatellite (GATA)<sub>4</sub>. È stato analizzato il DNA di quattro piante selezionate per ognuno dei sedici ecotipi in esame forniti dalla ditta sementiera "La Semiorto Sementi". La figura 3.1 riporta un esempio di *fingerprint* molecolare basato sull'utilizzo della sonda (GATA)<sub>4</sub> per tre degli ecotipi analizzati.

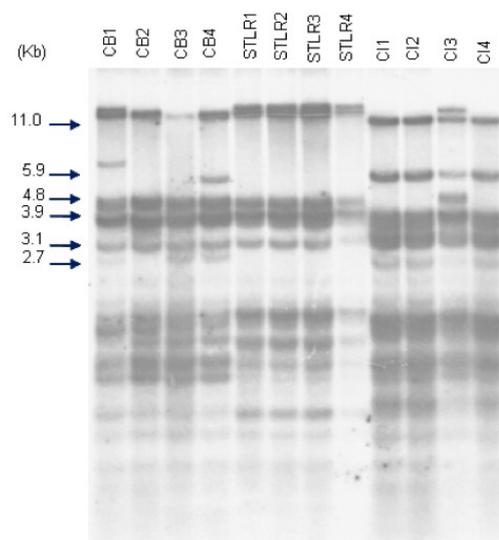


Figura 3.1: Esempio di eterogeneità ed uniformità di tre ecotipi tramite profilo (GATA)<sub>4</sub>. I nomi dei genotipi sono riportati con le rispettive sigle identificative (Tabella 2.1). Il numero indica la pianta analizzata.

La figura mostra eterogeneità per l'ecotipo 'Cuore di Bue' (CB) (quattro genotipi diversi) e per l'ecotipo 'Corbarino' (CI) (due diversi genotipi); un esempio di omogeneità è, invece, riportato per l'ecotipo 'Sorrento Tondo Liscio Rosato' (STLR). In particolare, l'analisi di diverse piante dell'ecotipo 'Cuore di Bue' mostra che la prima banda al di sopra di 11 Kb è presente solo nel genotipo CB1, così come la banda al di sopra di 5,9 Kb; la banda al di sotto di 5,9 Kb, invece, è presente solo nel genotipo CB4. Un'ulteriore differenza nel *pattern* di ibridazione è la presenza di una banda di circa 2,7 Kb presente solo nei genotipi CB3 e CB4. L'eterogeneità riscontrata nell'ecotipo 'Corbarino' è dovuta alla presenza di tre bande rilevate solo nel genotipo CI3 (la prima ad un peso molecolare maggiore di 11 Kb e le altre due a 4,8 Kb) e all'assenza della banda a circa 2,7 Kb nel genotipo CI3, presente invece nei genotipi CI1, C2 e CI3.

L'analisi dei sedici ecotipi ha rilevato una percentuale di bande polimorfiche di 87,10% ed un contenuto polimorfico informativo (PIC) di 0,224.

La matrice della distanza genetica è stata usata per ottenere una rappresentazione grafica delle relazioni genetiche (Figura 3.2).

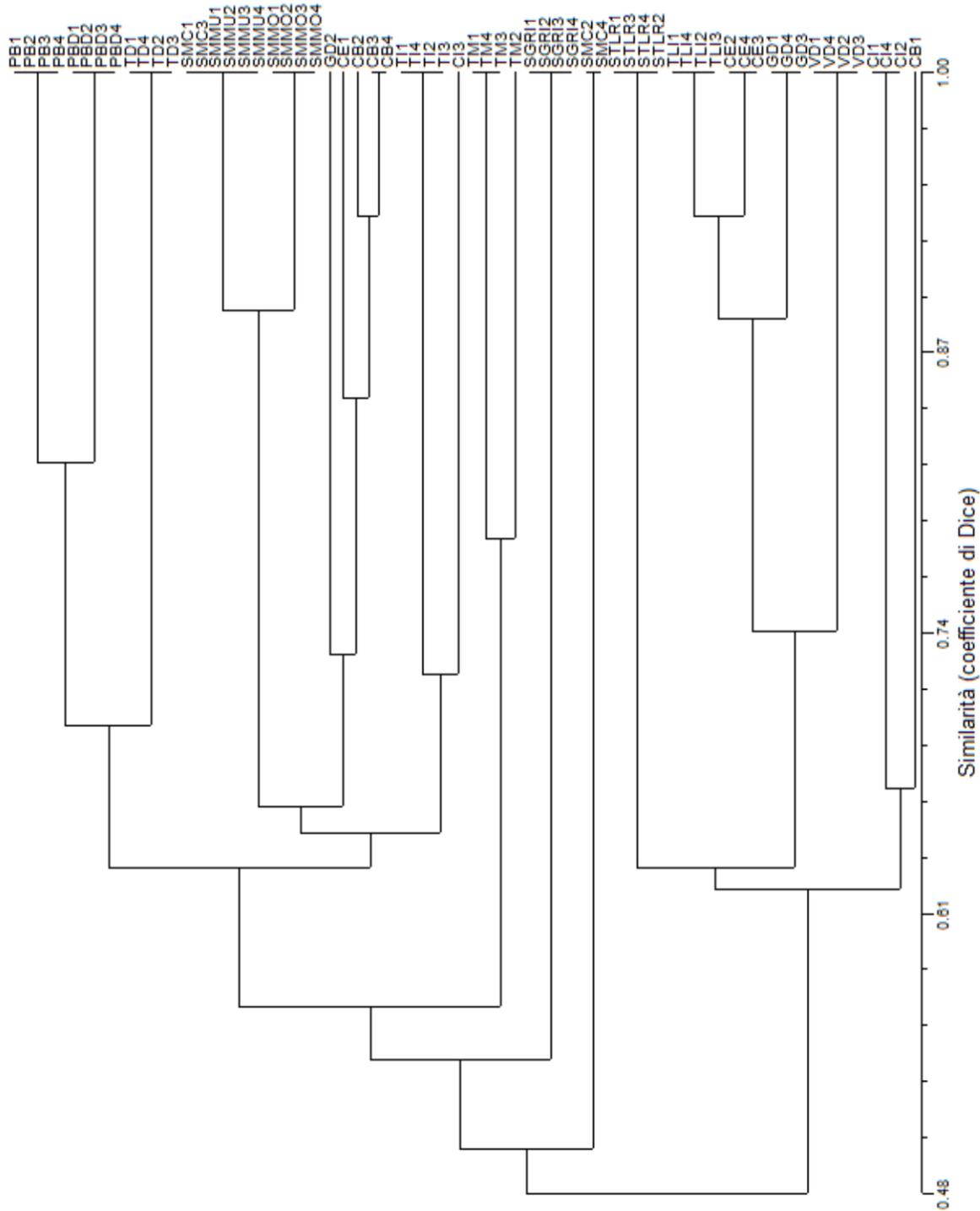


Figura 3.2: Dendrogramma generato mediante l'analisi cluster UPGMA utilizzando il programma NTSYS (Numerical Taxonomy System) sulla base dei dati (GATA)<sub>4</sub> degli ecotipi forniti da "La Semiorto Sementi". Le sigle sono identificative degli ecotipi riportati in tabella 2.1 ed il numero indica la pianta analizzata.

La figura mostra il grado di similarita' che indica che per la maggior parte degli ecotipi analizzati le repliche rappresentano lo stesso genotipo. Contaminazioni, invece, sono state riscontrate nell'ambito dell'ecotipo 'San Marzano Cilindrico' (SMC) che mostra la presenza di due genotipi diversi, uno dei quali risulta essere indistinguibile dal 'San Marzano Murano' (SMMU) con cui condivide lo stesso profilo di ibridazione. La presenza di due diversi genotipi e' stata riscontrata anche negli ecotipi 'Galatino Determinato' (GD), 'Cerise Determinato' (CE), 'Corbarino Indeterminato' (CI) e 'Tondo Marmande Indeterminato' (TM). Nell'ambito dei primi tre ecotipi, il contaminante, cosi' identificato perche' presente con una percentuale piu' bassa (25%) rispetto all'altro genotipo, risulta raggruppare separatamente dal genotipo principale; mentre per l'ecotipo 'Tondo Marmande Indeterminato' i due genotipi mostrano una elevata similarita'. L'ecotipo 'Cuor di Bue Indeterminato' (CB), invece, mostra la presenza di quattro diversi genotipi, tre dei quali raggruppano insieme risultando essere maggiormente correlati. I dati sono riassunti in tabella 3.1, dove si evince che dei 16 ecotipi, 10 mostrano un profilo (GATA)<sub>4</sub> omogeneo, nei rimanenti sei si evidenzia la presenza di 2-4 genotipi diversi nell'ambito dello stesso ecotipo.

Tabella 3.1: Uniformità ed eterogeneità degli ecotipi di pomodoro campano analizzati.

| Numero di ecotipi | Genotipi presenti |
|-------------------|-------------------|
| 10                | 1                 |
| 5                 | 2                 |
| 1                 | 4                 |

I dati hanno, quindi, indicato la presenza di un totale di 24 genotipi. Sono stati selezionati gli ecotipi risultati omogenei per le successive analisi morfologiche e genetiche.

### 3.1.2 Caratterizzazione fenotipica

I dieci ecotipi risultati omogenei da analisi (GATA)<sub>4</sub> sono stati sottoposti ad analisi morfologiche. La caratterizzazione fenotipica è stata condotta su almeno sei piante per ecotipo prendendo in esame quindici parametri morfologici quantitativi e venti qualitativi descritti in tabella 1.7 e 1.8, rispettivamente. Le figure 3.3 (A) e (B) mostrano l'analisi cluster UPGMA dei dieci ecotipi in esame. Si può osservare che, dal punto di vista morfologico, esiste una sufficiente variabilità da discriminare tutti gli ecotipi.

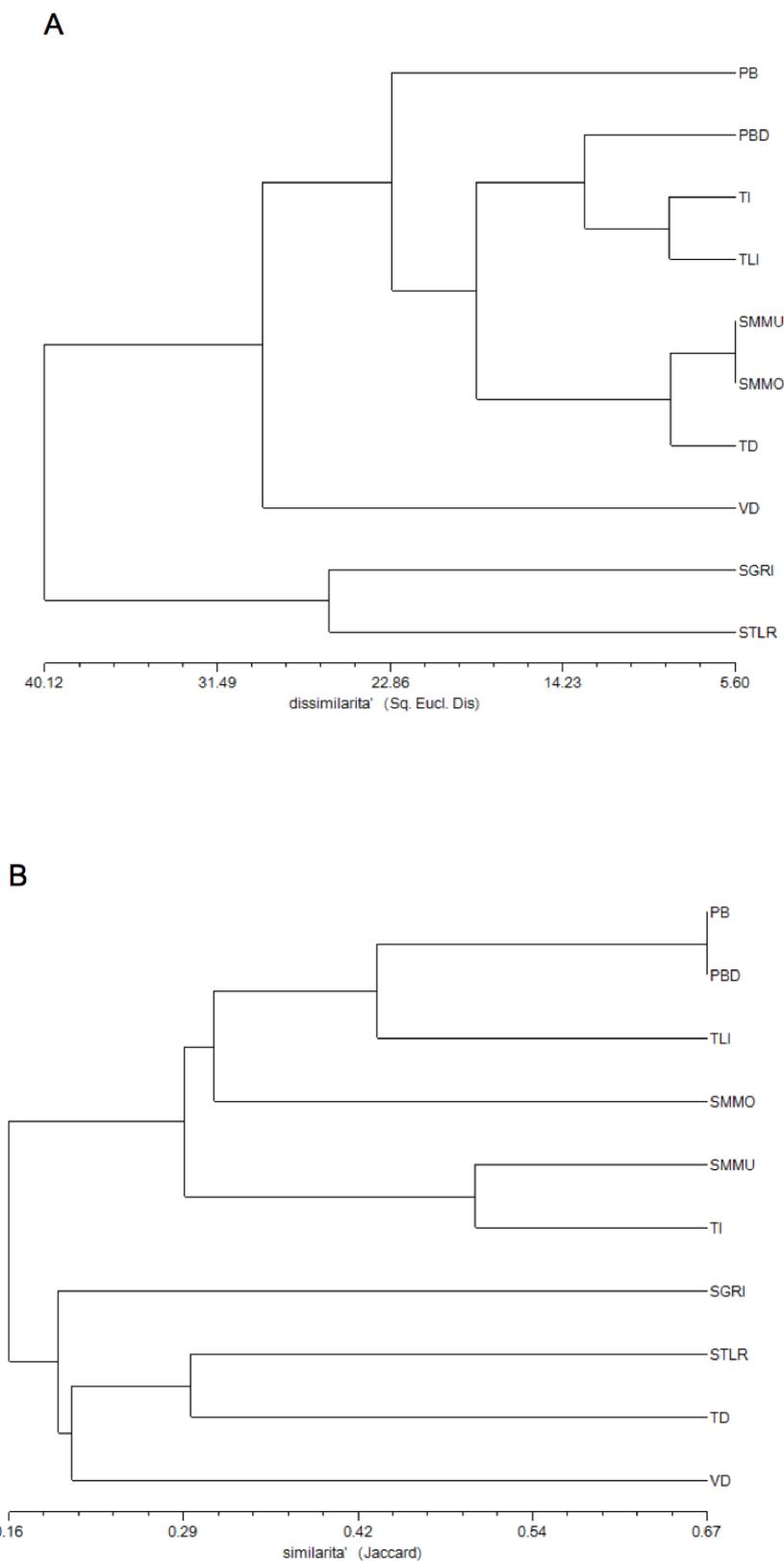


Figura 3.3: Dendrogramma generato mediante l'analisi cluster UPGMA utilizzando il programma NTSYS (Numerical Taxonomy System) basato sulle distanze genetiche calcolate utilizzando dati morfologici quantitativi (A) e qualitativi (B). Le distanze sono state ottenute utilizzando la distanza Euclidea per i caratteri quantitativi e il coefficiente *Jaccard* per quelli qualitativi.

Per verificare l'eventuale esistenza di una correlazione tra i dati morfologici qualitativi e quantitativi, e' stato condotto il test di Mantel e calcolato l'indice di Pearson ( $r$ ), la cui significativita' e' stata stabilita in seguito a 999 permutazioni casuali dei dati delle matrici. La correlazione tra le distanze genetiche basate sui dati quantitativi e qualitativi non e' risultata essere significativa, mostrando un  $r = 0,16$  e una significativita' statistica  $p = 0,20$ . La correlazione tra le matrici cofenetiche utilizzate per la costruzione dei dendrogrammi e' di  $r = 0,35$  con una significativita' statistica di  $p = 0,048$ .

### 3.1.3 Analisi molecolare mediante SSR

#### 3.1.3.1 Amplificazione PCR ed elettroforesi capillare degli ecotipi di pomodoro campani

Una caratterizzazione molecolare mediante microsatelliti SSR e' stata effettuata utilizzando quattordici coppie di primer (Tabella 2.3). Il DNA dei dieci ecotipi omogenei è stato amplificato con ognuno dei loci in esame che hanno generato specifici *pattern* di amplificazione nel range di peso molecolare atteso. In figura 3.4 e' riportato un esempio di visualizzazione su gel di agarosio delle amplificazioni ottenute per il locus L<sub>Eta</sub>015, per il quale la dimensione degli alleli attesa e' di 107 bp.

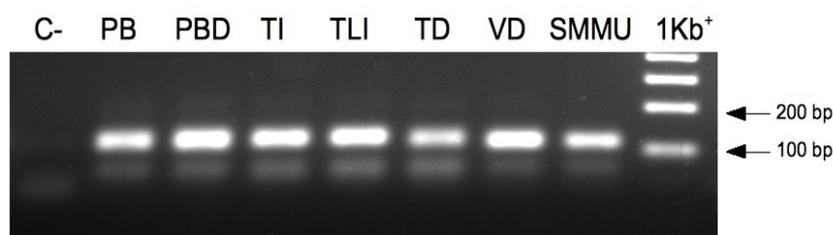


Figura 3.4: Gel al 2% di agarosio dei prodotti di amplificazione del locus SSR L<sub>Eta</sub>015 del DNA genomico estratto da foglie di alcuni degli ecotipi di pomodoro in analisi. I codici sono identificativi degli ecotipi come riportato in tabella 2.1; C-: controllo negativo, mix priva di template; 1Kb plus ladder (Invitrogen).

Gli amplificati con le quattordici coppie di primer SSR dei dieci campioni di DNA sono stati sottoposti ad elettroforesi capillare utilizzando un analizzatore di frammenti automatico ABI PRISM 3100-AVANT. La figura 3.5 illustra esempi di elettroferogrammi dei frammenti di DNA amplificati con il locus L<sub>Eta</sub>015, che ha una ripetizione dinucleotidica. E' possibile notare che il tracciato di ciascun microsatellite è costituito da più picchi di diversa altezza e lunghezza. Il picco più alto corrisponde all'allele del microsatellite analizzato mentre i picchi che lo affiancano costituiscono le "stutter bands" ( $n+1$ ,  $n-1$ ,  $n-2$ ,  $n-3$ , dove  $n$  rappresenta le dimensioni del picco più alto) (Applied Biosystems, 2004). Queste ultime sono sempre presenti nel profilo SSR e si formano, durante l'amplificazione, in seguito allo slittamento della *Taq* polimerasi sulla sequenza ripetuta del microsatellite (Delmotte et al., 2001).

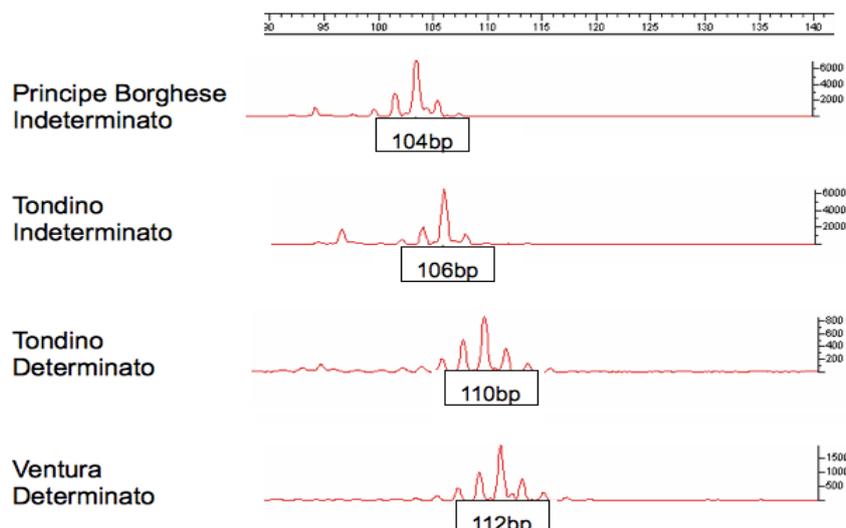


Figura 3.5: Esempi di profili SSR di quattro dei dieci ecotipi analizzati. Gli elettroferogrammi sono stati ottenuti mediante elettroforesi capillare degli amplificati rilevati al locus LEta015. Per ogni elettroferogramma la barra superiore indica le dimensioni degli alleli mentre la barra verticale di ogni profilo indica l'intensità dei picchi (RFU). Sono riportate le dimensioni degli alleli in paia di basi, arrotondate considerando la ripetizione del *core* del locus in esame.

Come atteso per ecotipi di pomodoro non ibridi, e' presente un solo allele per locus ad indicare che tutti i campioni in esame sono omozigoti. La tabella 3.2 riporta i profili allelici riscontrati dopo amplificazione ed elettroforesi capillare del DNA estratto dal materiale vegetale per ognuno dei quattordici loci SSR analizzati.

Tabella 3.2: Dimensioni degli alleli (bp) dei dieci ecotipi di pomodoro analizzati per ognuno dei quattordici loci SSR in esame.

| SSR      | Ecotipi |     |     |     |     |     |      |      |      |      |     |
|----------|---------|-----|-----|-----|-----|-----|------|------|------|------|-----|
|          | PB      | PBD | TI  | TLI | TD  | VD  | STLR | SGRI | SMMU | SMMO |     |
| LE20592  | 164     | 167 | 164 | 164 | 164 | 164 | 164  | 164  | 164  | 167  | 167 |
|          | 164     | 167 | 164 | 164 | 164 | 164 | 164  | 164  | 164  | 167  | 167 |
| LELE25   | 220     | 220 | 218 | 220 | 220 | 220 | 220  | 220  | 220  | 220  | 220 |
|          | 220     | 220 | 218 | 220 | 220 | 220 | 220  | 220  | 220  | 220  | 220 |
| LE21085  | 102     | 102 | 102 | 102 | 102 | 102 | 102  | 102  | 102  | 102  | 102 |
|          | 102     | 102 | 102 | 102 | 102 | 102 | 102  | 102  | 102  | 102  | 102 |
| LEEF1Aa  | 205     | 201 | 195 | 203 | 205 | 207 | 203  | 201  | 203  | 203  | 203 |
|          | 205     | 201 | 195 | 203 | 205 | 207 | 203  | 201  | 203  | 203  | 203 |
| LEta003  | 106     | 100 | 104 | 104 | 102 | 106 | 104  | 104  | 104  | 106  | 104 |
|          | 106     | 100 | 104 | 104 | 102 | 106 | 104  | 104  | 104  | 106  | 104 |
| LEat002  | 197     | 197 | 197 | 201 | 201 | 201 | 201  | 197  | 201  | 201  | 201 |
|          | 197     | 197 | 197 | 201 | 201 | 201 | 201  | 197  | 201  | 201  | 201 |
| LEga003  | 233     | 233 | 233 | 233 | 233 | 233 | 233  | 231  | 233  | 233  | 233 |
|          | 233     | 233 | 233 | 233 | 233 | 233 | 233  | 231  | 233  | 233  | 233 |
| LEtat002 | 195     | 198 | 195 | 195 | 195 | 195 | 195  | 195  | 195  | 198  | 198 |
|          | 195     | 198 | 195 | 195 | 195 | 195 | 195  | 195  | 195  | 198  | 198 |
| LEaat002 | 101     | 101 | 101 | 104 | 104 | 101 | 104  | 104  | 101  | 101  | 101 |
|          | 101     | 101 | 101 | 104 | 104 | 101 | 104  | 104  | 101  | 101  | 101 |
| LEcaa001 | 100     | 100 | 100 | 100 | 100 | 100 | 100  | 100  | 100  | 100  | 100 |
|          | 100     | 100 | 100 | 100 | 100 | 100 | 100  | 100  | 100  | 100  | 100 |
| LEaat007 | 96      | 96  | 96  | 99  | 99  | 96  | 99   | 99   | 96   | 96   | 96  |
|          | 96      | 96  | 96  | 99  | 99  | 96  | 99   | 99   | 96   | 96   | 96  |
| LEct001  | 105     | 101 | 105 | 105 | 105 | 97  | 101  | 97   | 105  | 105  | 105 |
|          | 105     | 101 | 105 | 105 | 105 | 97  | 101  | 97   | 105  | 105  | 105 |
| LEctt001 | 94      | 94  | 94  | 94  | 91  | 94  | 94   | 91   | 94   | 94   | 94  |
|          | 94      | 94  | 94  | 94  | 91  | 94  | 94   | 91   | 94   | 94   | 94  |
| LEta015  | 104     | 108 | 106 | 108 | 110 | 112 | 108  | 108  | 116  | 112  | 112 |
|          | 104     | 108 | 106 | 108 | 110 | 112 | 108  | 108  | 116  | 116  | 112 |

### 3.1.3.2 Analisi dei dati SSR per la caratterizzazione e discriminazione degli ecotipi campani

I valori relativi alle dimensioni alleliche calcolate per ogni singolo campione amplificato con i quattordici loci SSR sono stati utilizzati per il calcolo dei principali indici utili a stimare la variabilità genetica e la diversità allelica. La tabella 3.3 riporta i valori ottenuti per i diversi indici genetici che caratterizzano la popolazione presa in considerazione.

Tabella 3.3: Principali indici genetici e loro media ottenuti da analisi mediante SSR su ecotipi campani.

| SSR      | N  | Na           | Ne           | I            | Ho    | PIC          |
|----------|----|--------------|--------------|--------------|-------|--------------|
| LE20592  | 10 | 2,000        | 1,724        | 0,611        | 0,000 | 0,420        |
| LELE25   | 10 | 2,000        | 1,220        | 0,325        | 0,000 | 0,180        |
| LE21085  | 10 | 1,000        | 1,000        | 0,000        | 0,000 | 0,000        |
| LEEF1Aa  | 10 | 5,000        | 3,846        | 1,471        | 0,000 | 0,740        |
| LEta003  | 10 | 4,000        | 2,778        | 1,168        | 0,000 | 0,640        |
| LEat002  | 10 | 2,000        | 1,923        | 0,673        | 0,000 | 0,480        |
| LEga003  | 10 | 2,000        | 1,220        | 0,325        | 0,000 | 0,180        |
| LEtat002 | 10 | 2,000        | 1,724        | 0,611        | 0,000 | 0,420        |
| LEaat002 | 10 | 2,000        | 1,923        | 0,673        | 0,000 | 0,480        |
| LEcaa001 | 10 | 1,000        | 1,000        | 0,000        | 0,000 | 0,000        |
| LEaat007 | 10 | 2,000        | 1,923        | 0,673        | 0,000 | 0,480        |
| LEct001  | 10 | 3,000        | 2,273        | 0,950        | 0,000 | 0,560        |
| LEctt001 | 10 | 2,000        | 1,471        | 0,500        | 0,000 | 0,320        |
| LEta015  | 10 | <b>6,000</b> | <b>4,167</b> | <b>1,609</b> | 0,000 | <b>0,760</b> |
| media    | 10 | 2,818        | 2,204        | 0,787        | 0,000 | 0,460        |

Note: N, numero di ecotipi analizzati; Na, numero di alleli; Ne, numero effettivo di alleli; I, indice d'informazione. Ho, eterozigotità osservata; PIC, contenuto polimorfico informativo. I valori massimi per ogni indice sono riportati in grassetto.

E' stato riscontrato un elevato livello di polimorfismo per alcuni dei loci analizzati e la percentuale di loci polimorfici nella popolazione in esame e' risultata essere di 85,71%. Il valore Na identifica il numero di alleli ottenuto per ogni singolo locus SSR, in quanto tale è un numero naturale uguale o maggiore di 1. Un totale di 36 differenti alleli sono stati identificati, con una media di 2,818 alleli per locus. Due dei loci in esame (LE21085 e LEcaa001) sono risultati essere monomorfici nella popolazione considerata. Il locus LEta015 ha mostrato il numero di alleli più alto con un valore di Na pari a 6. In riferimento ai valori di Ne, esso rappresenta il numero effettivo di alleli e può assumere valori variabili da 1 a n alleli. Il locus con il valore maggiore di Ne (4,167) è risultato essere il locus LEta015, mentre i loci LE21085 e LEcaa001 hanno presentato il valore minimo (1). Il valore medio di Ne per tutti e quattordici i loci è stato 2,204. I valori osservati per l'indice I, il quale rappresenta l'indice d'informazione, possono essere usati per stimare la diversità allelica e genetica. Come riportato nella tabella, il valore maggiore di I (1,609) è stato riscontrato per il locus LEta015 mentre il valore minore (0) è identificativo dei loci LE21085 e LEcaa001. Ho è l'eterozigotità osservata, che risulta essere zero per tutti i loci e le accessioni analizzate. In riferimento al contenuto polimorfico informativo (PIC) esso rappresenta un indice della diversità genetica. Il valore massimo e' stato riscontrato per il locus LEta015 (0,76) che risulta, quindi, essere il maggiormente discriminante.

Il valore minimo e' stato ottenuto per i loci monomorfici (0). Il valore PIC medio considerando tutti i loci in esame è risultato 0,46, indicando un elevato potere discriminante del set di loci SSR utilizzati.

I suddetti profili unici sono stati utilizzati per identificare il minor numero di loci utilizzabili per discriminare i dieci ecotipi analizzati. In figura 3.6 viene proposta una chiave di identificazione grafica in cui i dieci ecotipi vengono discriminati utilizzando un set minimo di tre loci SSR capace di assegnare un profilo unico identificativo di ogni ecotipo.

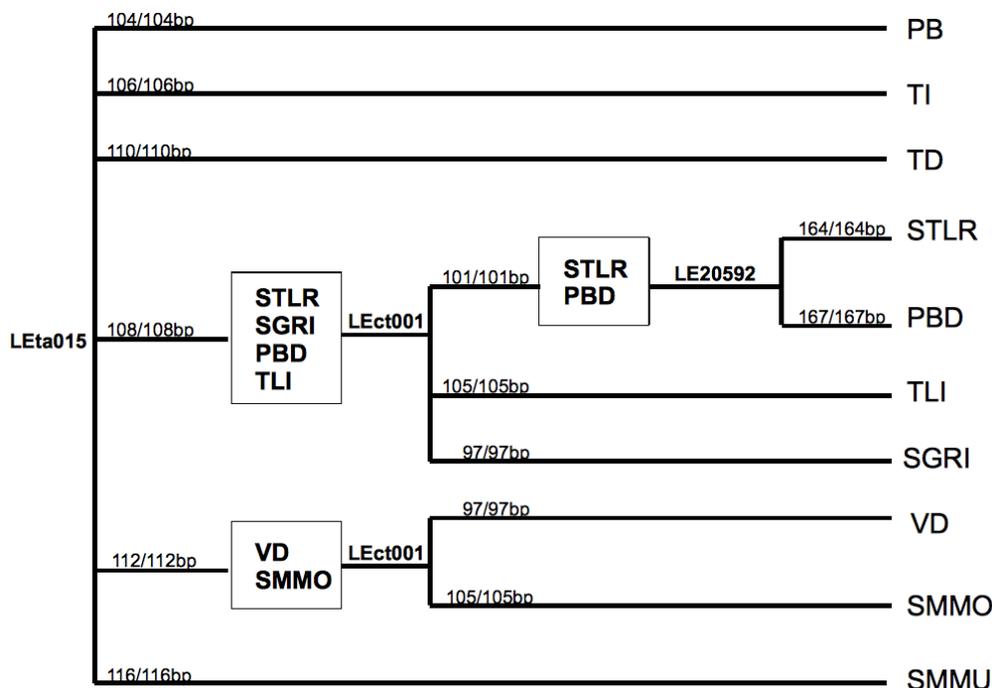


Figura 3.6: Chiave di identificazione grafica per discriminare i dieci ecotipi di pomodoro campano utilizzando i profili SSR ottenuti con un set minimo di tre coppie di primer (LEta015, LEct001, LE20592).

Il locus LETA015 e' risultato essere il maggiormente informativo e capace di discriminare la maggior parte degli ecotipi in esame. Risultano, pero', avere lo stesso profilo allelico omozigote al locus LETA015, gli ecotipi STLR, SGRI, PBD e TLI con alleli di 108bp e gli ecotipi VD e SMMO che mostrano un profilo omozigote con alleli di 112bp. Il locus LEct001 e' risultato capace per discriminare tra di loro tutti gli ecotipi non discriminabili dal locus LETA015, eccetto che STLR e PBD per i quali e' risultato necessario utilizzare i gia' citati due loci in combinazione con un terzo rappresentato da LE20592.

Le distanze genetiche di Nei basate sui dati SSR sono state utilizzate per costruire l' albero mostrato in figura 3.7.

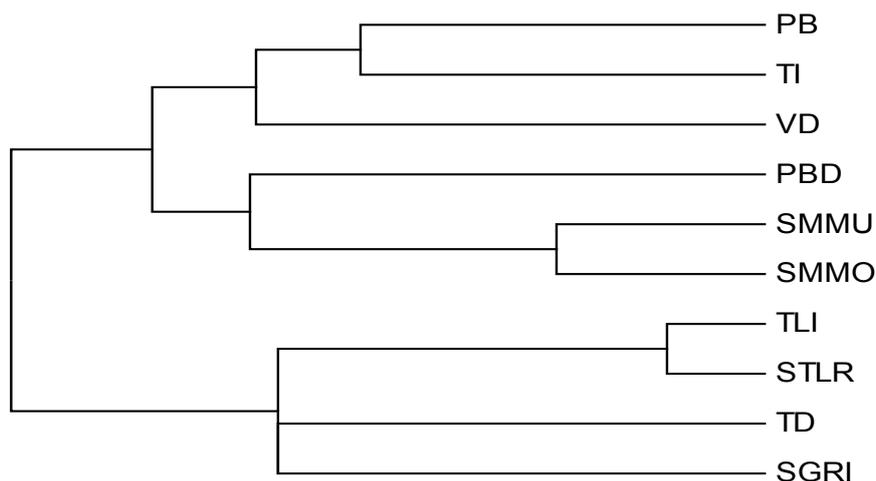


Figura 3.7: Dendrogramma ottenuto utilizzando l'algoritmo Neighbor-Joining basato sulle distanze genetiche calcolate sui dati SSR. I codici sono identificativi degli ecotipi riportati in tabella 2.1.

Dall'analisi del cluster si evidenzia come tutti gli ecotipi sono nettamente distinti ed inequivocabilmente individuano genotipi differenti. Questo risultato è paragonabile a quello ottenuto mediante l'analisi del minisatellite  $(GATA)_4$  sugli stessi ecotipi di pomodoro. L'analisi SSR è, però, più breve ed è automatizzata e può essere utilizzata anche per l'analisi del DNA estratto da matrici complesse.

### 3.1.4 Analisi molecolare mediante SNP

È stato investigato il potere discriminante di 71 loci SNP selezionati dalla letteratura (Francis et al., 2006). I dieci ecotipi di pomodoro campani omogenei da analisi  $(GATA)_4$  sono stati amplificati mediante PCR con le singole coppie di primers e successiva reazione ASPE mediante gli specifici primer allele-specifici per ognuno dei loci SNP biallelici analizzati (Tabella 2.6). L'analisi mediante tecnologia "Luminex 200" ha rilevato la presenza del segnale di solo uno dei due alleli dello SNP per ogni ecotipo in esame, indicando che tutti gli ecotipi sono omozigoti. L'elaborazione e l'analisi dei dati ottenuti ha rilevato che dei 71 loci SNP analizzati, 12 sono polimorfici mentre la restante parte è risultata essere monomorfica nella popolazione in esame mostrando la presenza dello stesso allele in tutti gli ecotipi. In tabella 3.4 sono riportati i nucleotidi associati ad ognuno degli SNP analizzati sui dieci ecotipi di pomodoro. Dei loci SNP polimorfici, il 75% delle mutazioni puntiformi è rappresentato da transizioni (pirimidina-pirimidina), di cui più del 77% è rappresentato da sostituzioni C/T. I dati sono stati sottoposti ad analisi per il calcolo di alcuni dei principali indici genetici, riportati in tabella 3.5.

Tabella 3.4: Nucleotide associato ad ognuno dei loci SNP analizzati sui dieci ecotipi di pomodoro in esame.

| SNP     | Genotipo |     |     |     |     |     |      |      |      |      |     |
|---------|----------|-----|-----|-----|-----|-----|------|------|------|------|-----|
|         | PB       | PBD | TI  | TLI | TD  | VD  | STLR | SGRI | SMMU | SMMO |     |
| CT10004 | C/C      | C/C | C/C | C/C | C/C | C/C | C/C  | C/C  | C/C  | C/C  | C/C |
| CT10018 | C/C      | C/C | C/C | C/C | C/C | T/T | C/C  | C/C  | C/C  | C/C  | C/C |
| CT10024 | G/G      | G/G | G/G | G/G | G/G | G/G | G/G  | G/G  | G/G  | G/G  | G/G |
| CT10043 | A/A      | A/A | A/A | A/A | A/A | A/A | A/A  | A/A  | A/A  | A/A  | A/A |
| CT10050 | C/C      | C/C | C/C | C/C | C/C | C/C | C/C  | C/C  | C/C  | C/C  | C/C |
| CT10075 | T/T      | T/T | T/T | C/C | T/T | C/C | C/C  | T/T  | T/T  | T/T  | T/T |
| CT10102 | A/A      | A/A | A/A | A/A | A/A | A/A | A/A  | A/A  | A/A  | A/A  | A/A |
| CT10120 | A/A      | A/A | A/A | A/A | A/A | A/A | G/G  | A/A  | G/G  | A/A  | A/A |
| CT10136 | T/T      | T/T | T/T | T/T | T/T | T/T | T/T  | T/T  | T/T  | T/T  | T/T |
| CT10138 | T/T      | T/T | T/T | T/T | T/T | T/T | T/T  | T/T  | T/T  | T/T  | T/T |
| CT10153 | G/G      | G/G | G/G | G/G | G/G | G/G | G/G  | G/G  | G/G  | G/G  | G/G |
| CT10162 | T/T      | T/T | T/T | C/C | T/T | T/T | T/T  | C/C  | T/T  | T/T  | T/T |
| CT10184 | T/T      | T/T | T/T | T/T | T/T | T/T | T/T  | -1   | T/T  | T/T  | T/T |
| CT10187 | C/C      | C/C | C/C | C/C | C/C | C/C | C/C  | C/C  | C/C  | C/C  | C/C |
| CT10190 | G/G      | G/G | G/G | G/G | G/G | G/G | G/G  | G/G  | G/G  | G/G  | G/G |
| CT10192 | C/C      | C/C | C/C | C/C | C/C | C/C | C/C  | C/C  | C/C  | C/C  | C/C |
| CT10205 | C/C      | C/C | C/C | C/C | C/C | C/C | C/C  | C/C  | C/C  | C/C  | C/C |
| CT10206 | C/C      | C/C | C/C | C/C | T/T | C/C | C/C  | C/C  | C/C  | C/C  | T/T |
| CT10215 | A/A      | A/A | G/G | G/G | G/G | A/A | G/G  | G/G  | G/G  | G/G  | G/G |
| CT10238 | C/C      | C/C | C/C | C/C | C/C | C/C | C/C  | C/C  | C/C  | C/C  | C/C |
| CT10242 | T/T      | T/T | T/T | T/T | T/T | T/T | T/T  | T/T  | T/T  | T/T  | T/T |
| CT10322 | C/C      | C/C | C/C | C/C | A/A | C/C | C/C  | C/C  | C/C  | C/C  | C/C |
| CT10346 | A/A      | A/A | A/A | A/A | A/A | A/A | A/A  | A/A  | A/A  | A/A  | A/A |
| CT10347 | G/G      | G/G | G/G | G/G | G/G | G/G | G/G  | G/G  | G/G  | G/G  | G/G |
| CT10351 | A/A      | A/A | A/A | A/A | A/A | A/A | A/A  | A/A  | A/A  | A/A  | A/A |
| CT10373 | C/C      | C/C | C/C | C/C | C/C | C/C | C/C  | C/C  | C/C  | C/C  | C/C |
| CT10375 | A/A      | A/A | A/A | A/A | A/A | A/A | A/A  | A/A  | A/A  | A/A  | A/A |
| CT10377 | C/C      | C/C | C/C | C/C | C/C | C/C | C/C  | C/C  | C/C  | C/C  | C/C |
| CT10385 | G/G      | G/G | G/G | G/G | G/G | G/G | G/G  | G/G  | G/G  | G/G  | G/G |
| CT10386 | G/G      | G/G | G/G | G/G | G/G | G/G | G/G  | G/G  | G/G  | G/G  | G/G |
| CT10390 | G/G      | G/G | G/G | G/G | G/G | G/G | G/G  | G/G  | G/G  | G/G  | G/G |
| CT10396 | C/C      | C/C | T/T | C/C | C/C | C/C | C/C  | C/C  | C/C  | C/C  | C/C |
| CT10401 | T/T      | A/A | T/T | T/T | T/T | T/T | T/T  | A/A  | A/A  | A/A  | A/A |
| CT10432 | G/G      | G/G | G/G | G/G | G/G | T/T | G/G  | G/G  | G/G  | G/G  | G/G |
| CT10437 | A/A      | A/A | A/A | A/A | A/A | A/A | A/A  | A/A  | A/A  | A/A  | A/A |
| CT10450 | A/A      | A/A | A/A | A/A | A/A | A/A | A/A  | A/A  | A/A  | A/A  | A/A |
| CT10469 | C/C      | C/C | C/C | C/C | C/C | C/C | C/C  | C/C  | C/C  | C/C  | C/C |
| CT10471 | G/G      | G/G | G/G | G/G | G/G | G/G | G/G  | G/G  | G/G  | G/G  | G/G |
| CT10485 | C/C      | C/C | C/C | C/C | C/C | C/C | C/C  | C/C  | C/C  | C/C  | C/C |
| CT10494 | T/T      | T/T | T/T | C/C | T/T | T/T | T/T  | C/C  | C/C  | T/T  | T/T |
| CT10526 | C/C      | C/C | C/C | C/C | C/C | C/C | C/C  | C/C  | C/C  | C/C  | C/C |
| CT10533 | A/A      | A/A | A/A | A/A | A/A | A/A | A/A  | A/A  | A/A  | A/A  | A/A |
| CT10535 | T/T      | T/T | T/T | T/T | T/T | T/T | T/T  | T/T  | T/T  | T/T  | T/T |
| CT10554 | A/A      | A/A | A/A | A/A | A/A | A/A | A/A  | A/A  | A/A  | A/A  | A/A |
| CT10556 | C/C      | C/C | C/C | C/C | C/C | C/C | C/C  | C/C  | C/C  | C/C  | C/C |
| CT10591 | T/T      | T/T | T/T | T/T | T/T | T/T | T/T  | T/T  | T/T  | T/T  | T/T |
| CT10639 | A/A      | A/A | A/A | A/A | A/A | A/A | A/A  | A/A  | A/A  | A/A  | A/A |
| CT10678 | G/G      | G/G | G/G | G/G | G/G | G/G | G/G  | G/G  | G/G  | G/G  | G/G |
| CT10715 | A/A      | A/A | A/A | A/A | A/A | A/A | A/A  | A/A  | A/A  | A/A  | A/A |
| CT10736 | T/T      | T/T | T/T | T/T | T/T | T/T | T/T  | T/T  | T/T  | T/T  | T/T |
| CT10768 | C/C      | C/C | C/C | C/C | C/C | C/C | C/C  | C/C  | C/C  | C/C  | C/C |
| CT10778 | G/G      | G/G | G/G | G/G | G/G | G/G | G/G  | G/G  | G/G  | G/G  | G/G |
| CT10781 | C/C      | C/C | C/C | C/C | C/C | C/C | C/C  | C/C  | C/C  | C/C  | C/C |
| CT10790 | T/T      | T/T | T/T | C/C | T/T | C/C | C/C  | C/C  | T/T  | T/T  | T/T |
| CT10793 | C/C      | C/C | C/C | C/C | C/C | C/C | C/C  | C/C  | C/C  | C/C  | C/C |
| CT10801 | C/C      | C/C | C/C | C/C | C/C | C/C | C/C  | C/C  | C/C  | C/C  | C/C |
| CT10807 | A/A      | A/A | A/A | A/A | A/A | A/A | A/A  | A/A  | A/A  | A/A  | A/A |
| CT10809 | G/G      | G/G | G/G | G/G | G/G | G/G | G/G  | G/G  | G/G  | G/G  | G/G |
| CT10909 | C/C      | C/C | C/C | C/C | C/C | C/C | C/C  | C/C  | C/C  | C/C  | C/C |
| CT10915 | C/C      | C/C | C/C | C/C | C/C | C/C | C/C  | C/C  | C/C  | C/C  | C/C |
| CT10923 | T/T      | T/T | T/T | T/T | T/T | T/T | T/T  | T/T  | T/T  | T/T  | T/T |
| CT10963 | G/G      | G/G | G/G | G/G | G/G | G/G | G/G  | G/G  | G/G  | G/G  | G/G |
| CT20017 | G/G      | G/G | G/G | G/G | G/G | G/G | G/G  | G/G  | G/G  | G/G  | G/G |
| CT20023 | T/T      | T/T | T/T | T/T | T/T | T/T | T/T  | T/T  | T/T  | T/T  | T/T |
| CT20028 | T/T      | T/T | T/T | T/T | T/T | T/T | T/T  | T/T  | T/T  | T/T  | T/T |
| CT20051 | A/A      | A/A | A/A | A/A | A/A | A/A | A/A  | A/A  | A/A  | A/A  | A/A |
| CT20059 | C/C      | C/C | C/C | C/C | C/C | C/C | C/C  | C/C  | C/C  | C/C  | C/C |
| CT20097 | T/T      | T/T | T/T | T/T | T/T | T/T | T/T  | T/T  | T/T  | T/T  | T/T |
| CT20127 | C/C      | C/C | C/C | C/C | C/C | C/C | C/C  | C/C  | C/C  | C/C  | C/C |
| CT20159 | A/A      | A/A | A/A | A/A | A/A | A/A | A/A  | A/A  | A/A  | A/A  | A/A |
| CT20181 | G/G      | G/G | G/G | G/G | G/G | G/G | G/G  | G/G  | G/G  | G/G  | G/G |

Note: le sigle sono identificative degli ecotipi riportati in tabella 2.1.

Tabella 3.5: Principali indici genetici e loro media ottenuti da analisi mediante SNP.

| SNP            | N         | Na           | Ne           | I            | Ho           | PIC          |
|----------------|-----------|--------------|--------------|--------------|--------------|--------------|
| CT10004        | 10        | 1,000        | 1,000        | 0,000        | 0,000        | 0,000        |
| <b>CT10018</b> | <b>10</b> | <b>2,000</b> | <b>1,220</b> | <b>0,325</b> | <b>0,000</b> | <b>0,180</b> |
| CT10024        | 10        | 1,000        | 1,000        | 0,000        | 0,000        | 0,000        |
| CT10043        | 10        | 1,000        | 1,000        | 0,000        | 0,000        | 0,000        |
| CT10050        | 10        | 1,000        | 1,000        | 0,000        | 0,000        | 0,000        |
| <b>CT10075</b> | <b>10</b> | <b>2,000</b> | <b>1,724</b> | <b>0,611</b> | <b>0,000</b> | <b>0,420</b> |
| CT10102        | 10        | 1,000        | 1,000        | 0,000        | 0,000        | 0,000        |
| <b>CT10120</b> | <b>10</b> | <b>2,000</b> | <b>1,471</b> | <b>0,500</b> | <b>0,000</b> | <b>0,320</b> |
| CT10136        | 10        | 1,000        | 1,000        | 0,000        | 0,000        | 0,000        |
| CT10138        | 10        | 1,000        | 1,000        | 0,000        | 0,000        | 0,000        |
| CT10153        | 10        | 1,000        | 1,000        | 0,000        | 0,000        | 0,000        |
| <b>CT10162</b> | <b>10</b> | <b>2,000</b> | <b>1,471</b> | <b>0,500</b> | <b>0,000</b> | <b>0,320</b> |
| CT10184        | 9         | 1,000        | 1,000        | 0,000        | 0,000        | 0,000        |
| CT10187        | 10        | 1,000        | 1,000        | 0,000        | 0,000        | 0,000        |
| CT10190        | 10        | 1,000        | 1,000        | 0,000        | 0,000        | 0,000        |
| CT10192        | 10        | 1,000        | 1,000        | 0,000        | 0,000        | 0,000        |
| CT10205        | 10        | 1,000        | 1,000        | 0,000        | 0,000        | 0,000        |
| <b>CT10206</b> | <b>10</b> | <b>2,000</b> | <b>1,471</b> | <b>0,500</b> | <b>0,000</b> | <b>0,320</b> |
| <b>CT10215</b> | <b>10</b> | <b>2,000</b> | <b>1,724</b> | <b>0,611</b> | <b>0,000</b> | <b>0,420</b> |
| CT10238        | 10        | 1,000        | 1,000        | 0,000        | 0,000        | 0,000        |
| CT10242        | 10        | 1,000        | 1,000        | 0,000        | 0,000        | 0,000        |
| <b>CT10322</b> | <b>10</b> | <b>2,000</b> | <b>1,220</b> | <b>0,325</b> | <b>0,000</b> | <b>0,180</b> |
| CT10346        | 10        | 1,000        | 1,000        | 0,000        | 0,000        | 0,000        |
| CT10347        | 10        | 1,000        | 1,000        | 0,000        | 0,000        | 0,000        |
| CT10351        | 10        | 1,000        | 1,000        | 0,000        | 0,000        | 0,000        |
| CT10373        | 10        | 1,000        | 1,000        | 0,000        | 0,000        | 0,000        |
| CT10375        | 10        | 1,000        | 1,000        | 0,000        | 0,000        | 0,000        |
| CT10377        | 10        | 1,000        | 1,000        | 0,000        | 0,000        | 0,000        |
| CT10385        | 10        | 1,000        | 1,000        | 0,000        | 0,000        | 0,000        |
| CT10386        | 10        | 1,000        | 1,000        | 0,000        | 0,000        | 0,000        |
| CT10390        | 10        | 1,000        | 1,000        | 0,000        | 0,000        | 0,000        |
| <b>CT10396</b> | <b>10</b> | <b>2,000</b> | <b>1,220</b> | <b>0,325</b> | <b>0,000</b> | <b>0,180</b> |
| <b>CT10401</b> | <b>10</b> | <b>2,000</b> | <b>1,923</b> | <b>0,673</b> | <b>0,000</b> | <b>0,480</b> |
| <b>CT10432</b> | <b>10</b> | <b>2,000</b> | <b>1,220</b> | <b>0,325</b> | <b>0,000</b> | <b>0,180</b> |
| CT10437        | 10        | 1,000        | 1,000        | 0,000        | 0,000        | 0,000        |
| CT10450        | 10        | 1,000        | 1,000        | 0,000        | 0,000        | 0,000        |
| CT10469        | 10        | 1,000        | 1,000        | 0,000        | 0,000        | 0,000        |
| CT10471        | 10        | 1,000        | 1,000        | 0,000        | 0,000        | 0,000        |
| CT10485        | 10        | 1,000        | 1,000        | 0,000        | 0,000        | 0,000        |
| <b>CT10494</b> | <b>10</b> | <b>2,000</b> | <b>1,724</b> | <b>0,611</b> | <b>0,000</b> | <b>0,420</b> |
| CT10526        | 10        | 1,000        | 1,000        | 0,000        | 0,000        | 0,000        |
| CT10533        | 10        | 1,000        | 1,000        | 0,000        | 0,000        | 0,000        |
| CT10535        | 10        | 1,000        | 1,000        | 0,000        | 0,000        | 0,000        |
| CT10554        | 10        | 1,000        | 1,000        | 0,000        | 0,000        | 0,000        |
| CT10556        | 10        | 1,000        | 1,000        | 0,000        | 0,000        | 0,000        |
| CT10591        | 10        | 1,000        | 1,000        | 0,000        | 0,000        | 0,000        |
| CT10639        | 10        | 1,000        | 1,000        | 0,000        | 0,000        | 0,000        |
| CT10678        | 10        | 1,000        | 1,000        | 0,000        | 0,000        | 0,000        |
| CT10715        | 10        | 1,000        | 1,000        | 0,000        | 0,000        | 0,000        |
| CT10736        | 10        | 1,000        | 1,000        | 0,000        | 0,000        | 0,000        |
| CT10768        | 10        | 1,000        | 1,000        | 0,000        | 0,000        | 0,000        |
| CT10778        | 10        | 1,000        | 1,000        | 0,000        | 0,000        | 0,000        |
| CT10781        | 10        | 1,000        | 1,000        | 0,000        | 0,000        | 0,000        |
| <b>CT10790</b> | <b>10</b> | <b>2,000</b> | <b>1,724</b> | <b>0,611</b> | <b>0,000</b> | <b>0,420</b> |
| CT10793        | 10        | 1,000        | 1,000        | 0,000        | 0,000        | 0,000        |
| CT10801        | 10        | 1,000        | 1,000        | 0,000        | 0,000        | 0,000        |
| CT10807        | 10        | 1,000        | 1,000        | 0,000        | 0,000        | 0,000        |
| CT10809        | 10        | 1,000        | 1,000        | 0,000        | 0,000        | 0,000        |
| CT10909        | 10        | 1,000        | 1,000        | 0,000        | 0,000        | 0,000        |
| CT10915        | 10        | 1,000        | 1,000        | 0,000        | 0,000        | 0,000        |
| CT10923        | 10        | 1,000        | 1,000        | 0,000        | 0,000        | 0,000        |
| CT10963        | 10        | 1,000        | 1,000        | 0,000        | 0,000        | 0,000        |
| CT20017        | 10        | 1,000        | 1,000        | 0,000        | 0,000        | 0,000        |
| CT20023        | 10        | 1,000        | 1,000        | 0,000        | 0,000        | 0,000        |
| CT20028        | 10        | 1,000        | 1,000        | 0,000        | 0,000        | 0,000        |
| CT20051        | 10        | 1,000        | 1,000        | 0,000        | 0,000        | 0,000        |
| CT20059        | 10        | 1,000        | 1,000        | 0,000        | 0,000        | 0,000        |
| CT20097        | 10        | 1,000        | 1,000        | 0,000        | 0,000        | 0,000        |
| CT20127        | 10        | 1,000        | 1,000        | 0,000        | 0,000        | 0,000        |
| CT20159        | 10        | 1,000        | 1,000        | 0,000        | 0,000        | 0,000        |
| CT20181        | 10        | 1,000        | 1,000        | 0,000        | 0,000        | 0,000        |
| media          | 9,99      | 1,17         | 1,09         | 0,08         | 0,000        | 0,05         |

Note: N, numero di ecotipi analizzati; Na, numero di alleli; Ne, numero effettivo di alleli; I, indice d'informazione. Ho, eterozigosità osservata; PIC, contenuto polimorfico informativo. I loci polimorfici ed i rispettivi indici genetici sono riportati in grassetto.

Come atteso per loci biallelici, il numero massimo di alleli per locus e' risultato essere due. Il numero totale degli alleli ottenuti e' 83, con una media  $N_a$  di 1,17. La percentuale di loci polimorfici, nell'ambito dei 71 analizzati, e' risultata essere del 16,90% ed il valore massimo di PIC vicino al valore soglia di 0,5 e' stato ottenuto per il locus CT10401 (0,480), che e' risultato essere il locus maggiormente discriminante. Gli altri 11 loci polimorfici e i rispettivi indici genetici calcolati sono evidenziati in grassetto. Il valore di PIC medio, considerando i soli loci SNP polimorfici sulla popolazione in esame, risulta essere di 0,315.

Per l'analisi cluster dei genotipi è stata calcolata la distanza genetica di Nei sulle frequenze alleliche sulla base della quale e' stato costruito un dendrogramma utilizzando l'algoritmo Neighbor-Joining (Figura 3.8). Dall'analisi del cluster si evidenzia come tutti gli ecotipi, fatta eccezione per TLI (Tondo Liscio Indeterminato) e STLR (Sorrento Tondo Liscio Rosato), vengono discriminate.

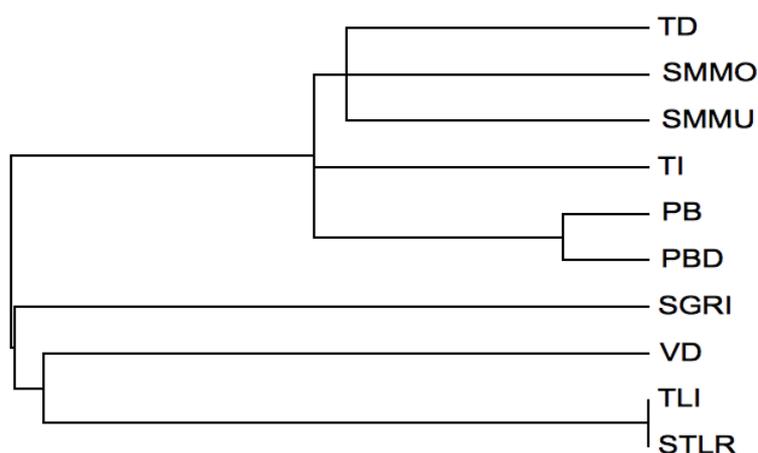


Figura 3.8: Dendrogramma ottenuto utilizzando l'algoritmo Neighbor-Joining basato sulle distanze genetiche calcolate sui dati SNP dei dieci ecotipi omogenei. I codici sono identificativi degli ecotipi riportati in tabella 2.1.

Le stesse relazioni rappresentate graficamente nella figura 3.8, sono mantenute utilizzando un set minimo di cinque loci SNP individuato tra i dodici loci polimorfici. In particolare, il locus CT10401 ha separato gli ecotipi in due gruppi, 'Sorrento Globoso Rosato Indeterminato' (SGRI), 'San Marzano Murano' (SMMU), 'San Marzano Morini' (SMMO), 'Principe Borghese Determinato' (PBD), da un lato e 'Principe Borghese Indeterminato' (PB), 'Ventura' (VD), 'Tondino Indeterminato' (TI), 'Sorrento Tondo Liscio Rosato' (STLR) dall'altro. La combinazione con il locus CT10215 ha permesso di discriminare SGRI, SMMU e SMMO da PBD, e PB, VD da TI, STLR, TD (Tondino Determinato), TLI (Tondo Liscio Indeterminato). Il terzo locus (CT10120) e' necessario per separare SGRI da SMMU e SMMO, e PB da VD; mentre mediante il locus CT10206 l'ecotipo SMMU risulta discriminato da SMMO, e TD e' separato da STLR, TI, TLI. Il locus CT10075 permette di discriminare TI da STLR e TLI; per i due ultimi ecotipi nessuno dei loci SNP analizzati e' stato utile per effettuare l'identificazione molecolare.

### 3.1.5 Confronto tra marcatori molecolari nell'analisi degli ecotipi di pomodoro

Il test della corrispondenza delle matrici di Mantel è stato utilizzato per comparare le matrici delle distanze genetiche ottenute dall'analisi dei dati dei minisatelliti, microsattelliti e SNP (Tabella 3.6). Tra i marcatori analizzati la correlazione esistente è bassa.

Tabella 3.6: Confronto a coppie tra i marcatori molecolari. Sono riportati i valori di correlazione genetica tra le matrici di distanza genetica.

|                     | Dati molecolari     |      |     |
|---------------------|---------------------|------|-----|
|                     | (GATA) <sub>4</sub> | SSR  | SNP |
| (GATA) <sub>4</sub> | 1                   |      |     |
| SSR                 | 0,15                | 1    |     |
| SNP                 | 0,22                | 0,01 | 1   |

Le analisi di gruppo basate sulle matrici delle distanze ottenute da ognuno dei tre marcatori molecolari ha evidenziato che tutti e tre i marcatori molecolari rilevano una elevata capacità di discriminazione degli ecotipi in esame. Il minisatellite ed i microsattelliti risultano essere capaci di identificare e inequivocabilmente discriminare tutti gli ecotipi analizzati, mentre il marcatore SNP è capace di discriminare tutte le accessioni fatta eccezione per gli ecotipi STLR (Sorrento Tondo Liscio Rosato) e TLI (Tondo Liscio Rosato) che presentano, comunque, un elevato valore di similarità per gli altri due marcatori utilizzati.

La tabella 3.7 riassume le principali caratteristiche riscontrate per i tre marcatori molecolari utilizzati.

Tabella 3.7: Confronto tra i tre marcatori molecolari utilizzati per analizzare gli ecotipi campani di pomodoro. Tipologia di marcatore, numero di loci analizzati, numero di alleli ottenuti, percentuale di loci polimorfici, contenuto polimorfico informativo (PIC) medio.

| Marker              | N. loci analizzati  | N. alleli individuati | % loci polimorfici | PIC medio                  |
|---------------------|---------------------|-----------------------|--------------------|----------------------------|
| (GATA) <sub>4</sub> | 31 (25 polimorfici) | 28                    | 80,65              | 0,22                       |
| SSR                 | 14 (12 polimorfici) | 36                    | 85,71              | 0,46 (0,53 <sup>a</sup> )  |
| SNP                 | 71 (12 polimorfici) | 83                    | 16,9               | 0,05 (0,315 <sup>a</sup> ) |

<sup>a</sup> valori medi di PIC considerando i soli loci polimorfici

I valori più elevati di loci polimorfici e di contenuto polimorfico informativo (PIC) sono stati riscontrati per il set di 14 loci SSR analizzati.

### 3.1.6 Discriminazione di ecotipi di pomodoro 'San Marzano'

Tra gli ecotipi di pomodoro campano riportati in tabella 2.1, ne sono stati selezionati tre rappresentativi di diverse tipologie di 'San Marzano', il 'San Marzano Murano', il 'San Marzano Cilindrico' ed il 'San Marzano Morini', in relazione ad un genotipo di pomodoro con frutti di forma globosa 'Sorrento Globoso Rosato Indeterminato' e, quindi, facilmente distinguibili dal 'San Marzano'. Sono state condotte analisi molecolari mediante marcatori (GATA)<sub>4</sub>, SSR e SNP in relazione alle caratteristiche morfologiche (Caramante et al., 2009).

#### 3.1.6.1 Analisi morfologica

Una analisi morfologica e' stata condotta per quantificare le differenze fenotipiche tra gli ecotipi in esame. La figura 3.9 mostra un dendrogramma UPGMA calcolato sui quindici caratteri morfologici quantitativi considerati per l'analisi.

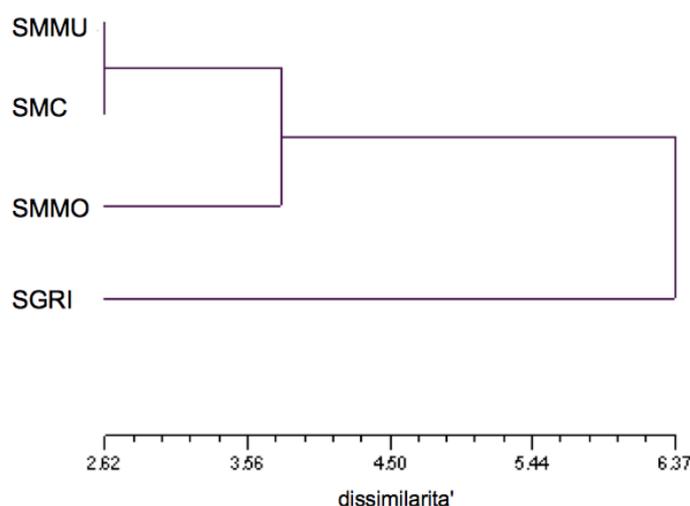


Figura 3.9: Dendrogramma generato mediante l'analisi cluster UPGMA basato sulla distanza Euclidea e calcolato sui dati di 15 caratteristiche morfologiche quantitative. I codici sono identificativi degli ecotipi riportati in tabella 2.1.

La tipologia 'San Marzano Murano' raggruppa con la tipologia 'San Marzano Cilindrico', mostrando un basso valore di dissimilarità. La tipologia 'San Marzano Morini' e', invece, risultata essere piu' distante dalle altre due. L'ecotipo 'Sorrento Globoso Rosato Indeterminato' e' risultato essere chiaramente separato dalle tipologie 'San Marzano', come atteso.

#### 3.1.6.2 Analisi mediante il minisatellite (GATA)<sub>4</sub>

E' stato effettuato il *fingerprint* molecolare delle tipologie 'San Marzano' in relazione all'ecotipo 'Sorrento Globoso Rosato Indeterminato' mediante il minisatellite (GATA)<sub>4</sub>. La figura 3.10 mostra il *pattern* di ibridazione dopo digestione mediante l'enzima *TaqI*, che ha rilevato la presenza di diversi frammenti distribuiti in un range di peso molecolare di 0,5-12,5 Kb. Il profilo dell'ecotipo 'San Marzano Murano' (SMMU) include frammenti le cui dimensioni sono state stimate essere di 11,0; 5,8; 4,6; 4,1; 3,3; 2,8 Kb, non presentando nessun frammento al di sopra di 2,8 Kb. Le sei bande ed il corrispondente *pattern* da essere generato sono descrittori molecolari della

cultivar 'San Marzano' (Rao et al., 2006). Lo stesso *pattern* e' presente nel *fingerprint* del DNA estratto dall'ecotipo 'San Marzano Cilindrico' (SMC), risultato non geneticamente uniforme. In figura 3.10B e' mostrato il *pattern* di ibridazione rappresentativo dei diversi genotipi presenti all'interno della tipologia 'San Marzano Cilindrico'.

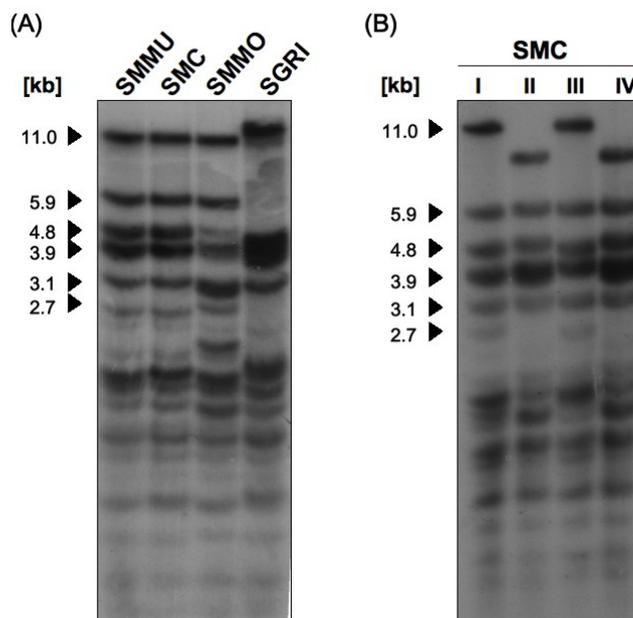


Figura 3.10: *Pattern* di ibridazione (GATA)<sub>4</sub> delle tipologie di 'San Marzano' (SMMU, SMC, SMMO) e dell'ecotipo 'Sorrento Globoso Rosato Indeterminato' (SGRI) analizzati (A) ed esempio di eterogeneità genetica dell'ecotipo 'San Marzano Cilindrico' (SMC) (B). SMC I e III mostrano possedere il profilo caratteristico del 'San Marzano', mentre SMC II e IV risultano essere i genotipi contaminanti. Sulla sinistra sono riportate le dimensioni molecolari stimate (Kb) delle sei bande specifiche del 'San Marzano'.

Il profilo genetico caratteristico del 'San Marzano' non e', invece, stato riscontrato negli ecotipi 'San Marzano Morini' (che mostra la presenza di un frammento al di sopra di 2,8 Kb e l'assenza del frammento a' 2,8 Kb) e 'Sorrento Globoso Rosato Indeterminato'. Quest'ultimo ha mostrato, il piu' differente *pattern* di ibridazione. Per una rappresentazione grafica delle distanze genetiche esistenti tra gli ecotipi analizzati, e' stato costruito un dendrogramma basato sulla matrice binaria ottenuta dai profili di ibridazione mostrati in figura 3.10A. Il dendrogramma in figura 3.11A ha confermato le relazioni osservate mediante analisi dei caratteri morfometrici. Per verificare l'eventuale esistenza di una correlazione statistica tra i dati genetici e morfologici, e' stato condotto il test di Mantel e calcolato l'indice di Pearson ( $r$ ), la cui significativita' e' stata stabilita in seguito a 999 permutazioni casuali dei dati delle matrici. Il confronto tra le matrici ha rilevato l'esistenza di una forte correlazione tra i dati molecolari e morfologici ( $r= 0,98$ ;  $p<0,05$ ).

### 3.1.6.3 Analisi SSR

E' stato valutato il potere discriminante relativo al set dei 14 microsatelliti selezionati (Tabella 2.3). In tabella 3.8 sono riportati i profili allelici ottenuti.

Tabella 3.8: Dimensioni degli alleli (bp) dei dieci ecotipi di pomodoro analizzati per ognuno dei quattordici loci SSR in esame.

| SSR      | Ecotipi |      |       |        |      |
|----------|---------|------|-------|--------|------|
|          | SMMU    | SMMO | SMC I | SMC II | SGRI |
| LE20592  | 167     | 167  | 167   | 164    | 164  |
|          | 167     | 167  | 167   | 164    | 164  |
| LELE25   | 220     | 220  | 220   | 220    | 220  |
|          | 220     | 220  | 220   | 220    | 220  |
| LE21085  | 102     | 102  | 102   | 102    | 102  |
|          | 102     | 102  | 102   | 102    | 102  |
| LEEF1Aa  | 203     | 203  | 203   | 203    | 201  |
|          | 203     | 203  | 203   | 203    | 201  |
| LEta003  | 106     | 104  | 106   | 104    | 104  |
|          | 106     | 104  | 106   | 104    | 104  |
| LEat002  | 201     | 201  | 201   | 201    | 197  |
|          | 201     | 201  | 201   | 201    | 197  |
| LEga003  | 233     | 233  | 233   | 233    | 231  |
|          | 233     | 233  | 233   | 233    | 231  |
| LEtat002 | 198     | 198  | 198   | 195    | 195  |
|          | 198     | 198  | 198   | 195    | 195  |
| LEaat002 | 101     | 101  | 101   | 101    | 104  |
|          | 101     | 101  | 101   | 101    | 104  |
| LEcaa001 | 100     | 100  | 100   | 100    | 100  |
|          | 100     | 100  | 100   | 100    | 100  |
| LEaat007 | 96      | 96   | 96    | 96     | 99   |
|          | 96      | 96   | 96    | 96     | 99   |
| LEct001  | 105     | 105  | 105   | 105    | 97   |
|          | 105     | 105  | 105   | 105    | 97   |
| LEctt001 | 94      | 94   | 94    | 94     | 91   |
|          | 94      | 94   | 94    | 94     | 91   |
| LEta015  | 116     | 112  | 116   | 112    | 108  |
|          | 116     | 112  | 116   | 112    | 108  |

Dei 14 loci analizzati, tre (LELE25, LE21085, Lecaa001) non hanno rivelato la presenza di alcun polimorfismo nell'ambito delle diverse tipologie 'San Marzano' e dell'ecotipo 'Sorrento Globoso Rosato Indeterminato' in esame. I dati derivanti dai restanti 11 loci SSR hanno indicato che il 'San Marzano Murano' (SMMU) ed il 'San Marzano Cilindrico' I (SMC I, il genotipo dell'ecotipo SMC che ha mostrato possedere il profilo specifico del 'San Marzano') non hanno mostrato avere differenze alleliche a tutti i loci analizzati. Le differenze alleliche tra SMC I e SMC II (quest'ultimo rappresentante il genotipo 'San Marzano Cilindrico', identificato come contaminante, che non mostra il profilo  $(GATA)_4$  specifico del 'San Marzano'), sono state riscontrate per quattro loci SSR. I risultati dei profili allelici hanno mostrato che solo due amplificazioni SSR possono discriminare l'ecotipo 'San Marzano Morini' (SMMO) dal 'San Marzano', mentre l'ecotipo 'Sorrento Globoso Rosato Indeterminato' (SGRI) e' risultato essere il piu' differente. Una classificazione gerarchica basata sulla matrice di similarita' dei dati SSR (Figura 3.11B) ha confermato le relazioni osservate

dall'analisi del minisatellite (GATA)<sub>4</sub>. La matrice di similarita' basata sui dati SSR non ha, pero', mostrato una correlazione significativa con i dati morfologici ( $r= 0,51$ ;  $p= 0,18$ ), dovuto probabilmente al fatto che la distanza genetica tra il gruppo delle tipologie 'San Marzano' e il 'Sorrento Globoso Rosato Indeterminato' e' maggiore quando calcolata sulla base dei dati SSR.

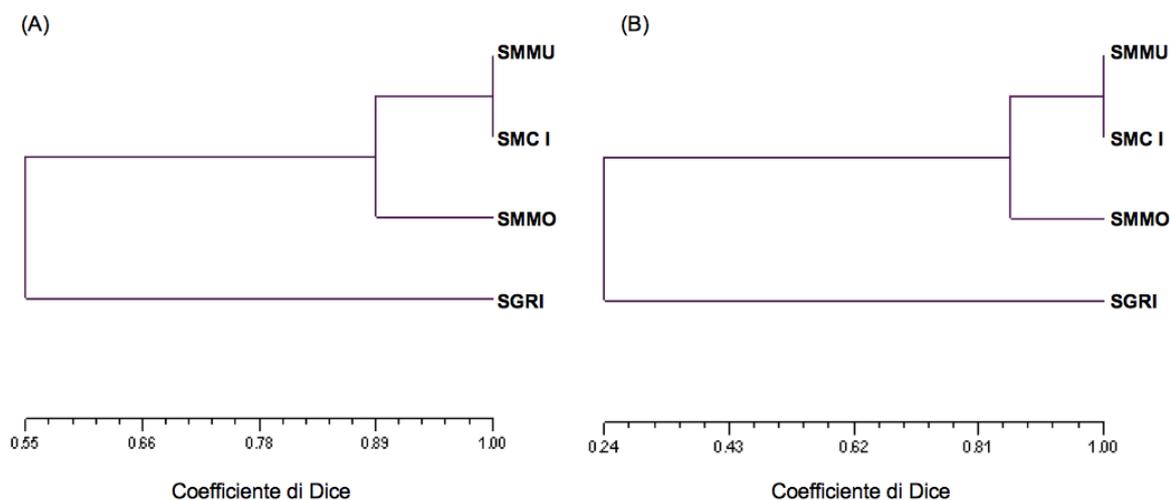


Figura 3.11: Dendrogrammi generati mediante l'analisi cluster UPGMA basati sulle matrici di similarita' delle distanze genetiche calcolate utilizzando i dati del minisatellite (A) e degli SSR (B).

#### 3.1.6.4 Analisi CAPS

Successivamente, 14 dei loci EST-SNP (Tabella 2.4), selezionati dalla letteratura (Yang et al., 2004) sulla base della elevata probabilita' di identificare polimorfismi a singolo nucleotide (53%) in *S. lycopersicum*, sono stati saggiati mediante analisi CAPS. L'amplificazione dei loci EST-SNP e' stata effettuata su DNA genomico, per cui alcuni prodotti di amplificazione sono risultati essere di dimensioni molecolari maggiori rispetto alle attese per la presenza di introni. In tabella 3.9 sono riportate le dimensioni molecolari dei prodotti di amplificazione e dei relativi frammenti di taglio ottenuti dopo digestione enzimatica.

Tabella 3.9: Conversione di EST-SNP in CAPS. Nome dello SNP; dimensione osservata dei prodotti PCR in bp; frammenti di restrizione in bp; nome dell'enzima di restrizione utilizzato per l'analisi CAPS.

| SNP       | Dimensione frammenti PCR (bp) | Frammenti di taglio (bp)     | Enzima di digestione |
|-----------|-------------------------------|------------------------------|----------------------|
| LEOH2     | 201                           | 109, 92                      | <i>BsaWI</i>         |
| LEOH8.4   | 208                           | 127, 81                      | <i>HaeIII</i>        |
| LEOH11.1  | 189                           | 148,41                       | <i>HhaI</i>          |
| LEOH12.2  | 181                           | 158,23                       | <i>MseI</i>          |
| LEOH16.2  | 206                           | 116; 90 <sup>a</sup>         | <i>BsaWI</i>         |
| LEOH 16.3 | 206                           | 152,54 <sup>a</sup>          | <i>BsaWI</i>         |
| LEOH20.1  | 246                           | 126; 120                     | <i>HaeIII</i>        |
| LEOH23.1  | 191                           | 127,64                       | <i>MspI</i>          |
| LEOH25.1  | 644                           | 471; 77; 52; 44 <sup>b</sup> | <i>HincII</i>        |
| LEOH28    | 223                           | 114; 109                     | <i>TaqI</i>          |
| LEOH31.3  | 388                           | 307; 81                      | <i>MspI</i>          |
| LEOH31.4  | 388                           | 217; 48; 87; 37 <sup>c</sup> | <i>MseI</i>          |
| LEOH32.1  | 521                           | 249; 242; 30 <sup>d</sup>    | <i>HaeIII</i>        |
| LEOH39    | 191                           | 162; 29                      | <i>EcoRII</i>        |

<sup>a</sup> la digestione simultanea di LEOH 16.2 e 16.3 genera 116; 54; 36

<sup>b</sup> l'assenza del sito polimorfico *HincII* genera 471; 128; 44

<sup>c</sup> l'assenza del sito polimorfico *MseI* genera 265; 87; 37

<sup>d</sup> l'assenza del sito polimorfico *HaeIII* genera 491; 30

La validazione dello SNP ed il calcolo della mappa di restrizione sono stati effettuati mediante sequenziamento degli amplificati e allineamento con le sequenze EST di riferimento. In tabella 3.10 sono riportati i risultati della analisi di restrizione, i cui profili di digestione non hanno rilevato nessuna differenza tra i cinque genotipi in esame.

Tabella 3.10: Profili SNP da analisi CAPS e validazione mediante sequenziamento degli ecotipi 'San Marzano' e 'Sorrento'.

| SNP       | SMMU | SMMO | SMC I | SMC II | SGRI |
|-----------|------|------|-------|--------|------|
| LEOH2     | +/+  | +/+  | +/+   | +/+    | +/+  |
| LEOH8.4   | -/-  | -/-  | -/-   | -/-    | -/-  |
| LEOH11.1  | -/-  | -/-  | -/-   | -/-    | -/-  |
| LEOH12.2  | +/+  | +/+  | +/+   | +/+    | +/+  |
| LEOH16.2  | +/+  | +/+  | +/+   | +/+    | +/+  |
| LEOH 16.3 | -/-  | -/-  | -/-   | -/-    | -/-  |
| LEOH20.1  | +/+  | +/+  | +/+   | +/+    | +/+  |
| LEOH23.1  | -/-  | -/-  | -/-   | -/-    | -/-  |
| LEOH25.1  | -/-  | -/-  | -/-   | -/-    | -/-  |
| LEOH28    | -/-  | -/-  | -/-   | -/-    | -/-  |
| LEOH31.3  | -/-  | -/-  | -/-   | -/-    | -/-  |
| LEOH31.4  | -/-  | -/-  | -/-   | -/-    | -/-  |
| LEOH32.1  | -/-  | -/-  | -/-   | -/-    | -/-  |
| LEOH39    | -/-  | -/-  | -/-   | -/-    | -/-  |

Sito di restrizione polimorfico: presenza (+); assenza (-).

### 3.1.6.5 Analisi SNP

E' stato, inoltre, valutato il potere discriminante di 71 loci SNP (Tabella 2.6). Come atteso per ecotipi di pomodoro non ibridi, l'analisi di tutti i loci SNP ha generato un singolo segnale indicando che tutti i genotipi in esame sono omozigoti. Dei sette loci polimorfici, quattro hanno permesso di discriminare SMCII (il genotipo contaminante dell'ecotipo SMC) dagli altri ecotipi analizzati; differenze alleliche per due loci sono state riscontrate tra l'ecotipo SGRI e le altre tipologie 'San Marzano', mentre l'ecotipo SMMO e' risultato essere discriminato per un solo locus di cui condivide lo stesso profilo allelico con SMCII. Gli ecotipi SMMU e SMCI hanno, invece, mostrato possedere lo stesso profilo allelico per tutti i loci analizzati (Tabella 3.11). Il calcolo dei principali indici genetici (Tabella 3.12) ha rilevato la presenza di un totale di 78 alleli ed un numero massimo di alleli per locus pari a due, come atteso per loci biallelici. Dei 71 loci SNP analizzati, sette hanno mostrato polimorfismo all'interno dei cinque ecotipi analizzati con una percentuale di loci polimorfici di 9,86%. Il locus che e' risultato essere il maggiormente discriminante e' stato CT10206 di cui e' stato stimato un PIC di 0,480.

Tabella 3.11: Nucleotide associato ad ognuno dei sette loci SNP polimorfici analizzati sui cinque ecotipi di pomodoro in esame. In grassetto sono riportati i profili che permettono la discriminazione tra gli ecotipi.

| SNP     | Ecotipo |            |      |            |            |
|---------|---------|------------|------|------------|------------|
|         | SMMU    | SMMO       | SMCI | SMCII      | SGRI       |
| CT10018 | C/C     | C/C        | C/C  | <b>T/T</b> | C/C        |
| CT10120 | A/A     | A/A        | A/A  | A/A        | <b>G/G</b> |
| CT10206 | C/C     | <b>T/T</b> | C/C  | <b>T/T</b> | C/C        |
| CT10215 | G/G     | G/G        | G/G  | <b>A/A</b> | G/G        |
| CT10401 | A/A     | A/A        | A/A  | <b>T/T</b> | A/A        |
| CT10494 | T/T     | T/T        | T/T  | T/T        | <b>C/C</b> |
| CT10790 | T/T     | T/T        | T/T  | <b>C/C</b> | T/T        |

Tabella 3.12: Principali indici genetici ottenuti da analisi mediante SNP sugli ecotipi 'San Marzano' e 'Sorrento Globoso Rosato Indeterminato'.

| SNP            | N | Na    | Ne           | I            | Ho    | PIC          |
|----------------|---|-------|--------------|--------------|-------|--------------|
| CT10004        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| <b>CT10018</b> | 5 | 2,000 | 1,471        | 0,500        | 0,000 | 0,320        |
| CT10024        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| CT10043        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| CT10050        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| CT10075        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| CT10102        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| <b>CT10120</b> | 5 | 2,000 | 1,471        | 0,500        | 0,000 | 0,320        |
| CT10136        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| CT10138        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| CT10153        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| CT10162        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| CT10184        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| CT10187        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| CT10190        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| CT10192        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| CT10205        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| <b>CT10206</b> | 5 | 2,000 | <b>1,923</b> | <b>0,673</b> | 0,000 | <b>0,480</b> |
| <b>CT10215</b> | 5 | 2,000 | 1,471        | 0,500        | 0,000 | 0,320        |
| CT10238        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| CT10242        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| CT10322        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| CT10346        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| CT10347        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| CT10351        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| CT10373        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| CT10375        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| CT10377        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| CT10385        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| CT10386        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| CT10390        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| CT10396        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| <b>CT10401</b> | 5 | 2,000 | 1,471        | 0,500        | 0,000 | 0,320        |
| CT10432        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| CT10437        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| CT10450        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| CT10469        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| CT10471        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| CT10485        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| <b>CT10494</b> | 5 | 2,000 | 1,471        | 0,500        | 0,000 | 0,320        |
| CT10526        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| CT10533        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| CT10535        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| CT10554        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| CT10556        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| CT10591        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| CT10639        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| CT10678        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| CT10715        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| CT10736        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| CT10768        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| CT10778        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| CT10781        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| <b>CT10790</b> | 5 | 2,000 | 1,471        | 0,500        | 0,000 | 0,320        |
| CT10793        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| CT10801        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| CT10807        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| CT10809        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| CT10909        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| CT10915        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| CT10923        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| CT10963        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| CT20017        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| CT20023        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| CT20028        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| CT20051        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| CT20059        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| CT20097        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| CT20127        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| CT20159        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| CT20181        | 5 | 1,000 | 1,000        | 0,000        | 0,000 | 0,000        |
| media          | 5 | 1,099 | 1,053        | 0,052        | 0,000 | 0,034        |

Note: N, numero di ecotipi analizzati; Na, numero di alleli; Ne, numero effettivo di alleli; I, indice d'informazione. Ho, eterozigosit  osservata; PIC, contenuto polimorfico informativo. I loci polimorfici ed i valori piu' elevati degli indici genetici sono riportati in grassetto.

L'analisi cluster riportata in figura 3.12 ha mostrato che, al pari del minisatellite e dei microsatelliti, i marcatori SNP sono risultati essere capaci di discriminare i diversi ecotipi tra di loro ed il contaminante all'interno dell'ecotipo SMC. Gli ecotipi SMMU e SMCI non risultano essere discriminabili dalle analisi morfologiche, ne' da analisi SNP e SSR e hanno mostrato possedere lo stesso profilo (GATA)<sub>4</sub> caratteristico ed identificativo del 'San Marzano'.

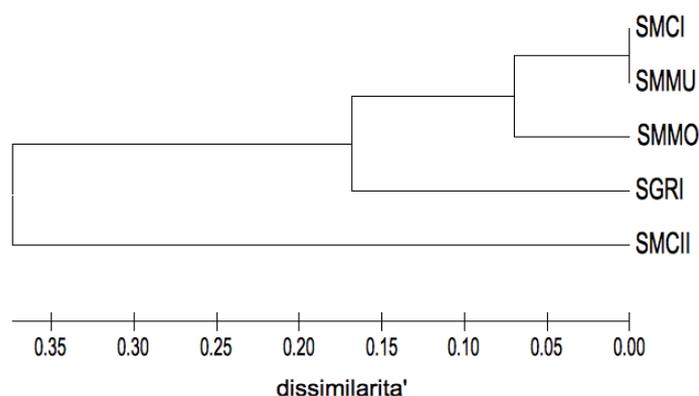


Figura 3.12: Dendrogramma generato mediante l'analisi cluster UPGMA basata sulla matrice di dissimilarità delle distanze genetiche calcolata utilizzando i dati SNP.

## 3.2 Identificazione e discriminazione varietale

### 3.2.1 Caratterizzazione varietale mediante SSR

I quattordici loci SSR selezionati (Tabella 2.3) sono stati utilizzati per l'identificazione e la discriminazione delle venti varietà commerciali di pomodoro in esame riportate in tabella 2.1. Tutti i loci analizzati hanno generato prodotti di amplificazione specifici delle rispettive dimensioni molecolari attese, su tutte le varietà. L'analisi mediante elettroforesi capillare dei frammenti di DNA amplificati ha rilevato che la maggior parte delle varietà ha mostrato la presenza di due alleli per gran parte dei loci esaminati, ad indicare la presenza di materiale ibrido ed eterozigote. In figura 3.13 sono riportati degli esempi di elettroferogrammi dove è possibile notare la presenza di genotipi omozigoti ed eterozigoti allo stesso locus analizzato.

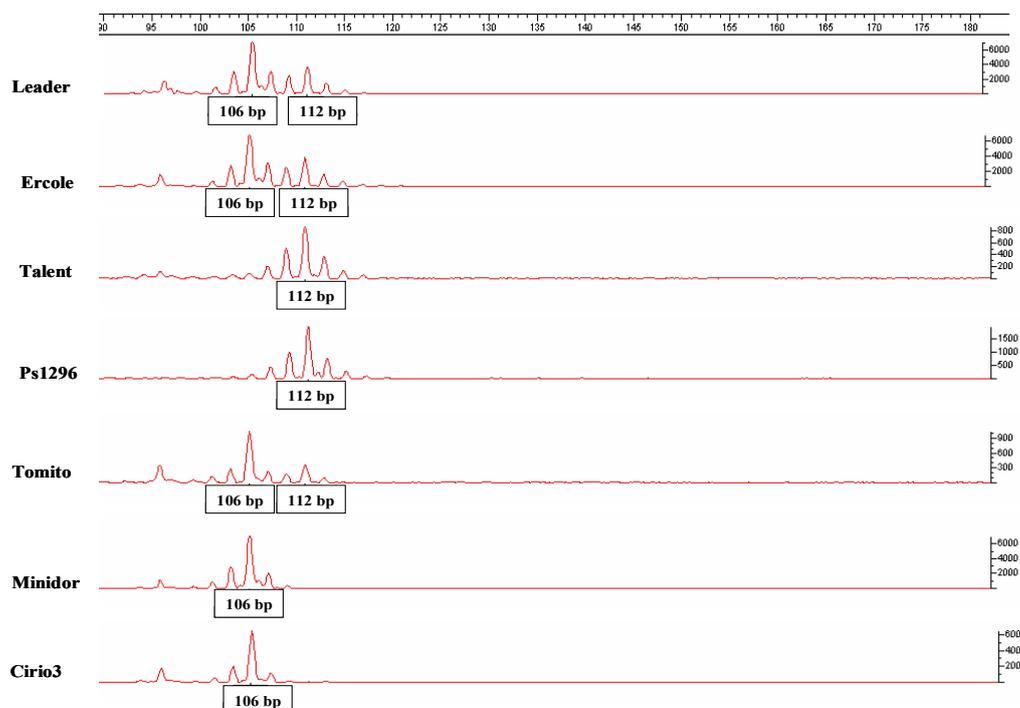


Figura 3.13: Esempi di elettroferogrammi del DNA genomico estratto da foglia di sette varietà ibride da industria amplificato con il locus SSR LEta015. Analisi dei dati effettuata mediante i software ABI PRISM Genotyper 2.5 (v. 3.7) e GeneScan (Applied Biosystems).

I dati generati dall'analisi dei quattordici loci SSR per ognuna delle venti varietà sono stati sottoposti ad analisi per il calcolo dei principali indici genetici che potessero portare alla rilevazione della variabilità genetica e della ricchezza allelica esistente. I risultati sono riportati in tabella 3.13.

Tabella 3.13: Principali indici genetici e loro media per ognuno dei 14 loci SSR analizzati sulle varietà.

| SSR      | N  | Na           | Ne           | I            | Ho           | PIC          |
|----------|----|--------------|--------------|--------------|--------------|--------------|
| LE20592  | 20 | 3,000        | 2,156        | 0,914        | 0,500        | 0,536        |
| LELE25   | 20 | 2,000        | 1,220        | 0,325        | 0,000        | 0,180        |
| LE21085  | 20 | 3,000        | 2,778        | 1,055        | 0,400        | 0,640        |
| LEEF1Aa  | 20 | <b>7,000</b> | <b>4,624</b> | <b>1,677</b> | 0,250        | <b>0,784</b> |
| LEta003  | 20 | 4,000        | 2,192        | 0,928        | 0,350        | 0,544        |
| LEat002  | 20 | 3,000        | 1,681        | 0,731        | 0,200        | 0,405        |
| LEga003  | 20 | 3,000        | 2,005        | 0,809        | <b>0,650</b> | 0,501        |
| LEtat002 | 20 | 3,000        | 2,156        | 0,914        | 0,500        | 0,536        |
| LEaat002 | 20 | 2,000        | 1,882        | 0,662        | 0,250        | 0,469        |
| LEcaa001 | 20 | 1,000        | 1,000        | 0,000        | 0,000        | 0,000        |
| LEaat007 | 20 | 2,000        | 1,882        | 0,662        | 0,250        | 0,469        |
| LEct001  | 20 | 2,000        | 1,105        | 0,199        | 0,000        | 0,095        |
| LEctt001 | 20 | 4,000        | 2,036        | 0,913        | 0,550        | 0,509        |
| LEta015  | 20 | 2,000        | 2,000        | 0,693        | 0,300        | 0,500        |
| media    | 20 | 2,929        | 2,051        | 0,749        | 0,300        | 0,441        |

Note: N, numero di varietà analizzate; Na, numero di alleli; Ne, numero effettivo di alleli; I, indice d'informazione. Ho, eterozigotità osservata; PIC, contenuto polimorfico informativo. I valori massimi sono riportati in grassetto.

Dei 14 loci SSR analizzati, tredici sono risultati essere polimorfici nella popolazione considerata, mostrando una percentuale di loci polimorfici di 92,86%. Il numero totale di alleli identificato e' stato di 41, con un valore massimo riscontrato per il locus LEEF1Aa (7). Il valore medio di Na e' di 2,929 alleli per locus analizzato. Il contenuto polimorfo informativo (PIC), che per marcatori multiallelici assume valori compresi tra 0 e 1, ha mostrato il valore massimo per il locus LEEF1Aa (0,784) che e' risultato possedere il maggior potere discriminante; il valore medio di PIC per tutti i loci in esame e' di 0,441 e assume un valore lievemente superiore se si considerano solo i loci polimorfici (0,474). Una rappresentazione grafica delle relazioni esistenti tra le varieta' in esame e del potere discriminante dei loci SSR selezionati e' mostrata in figura 3.14.

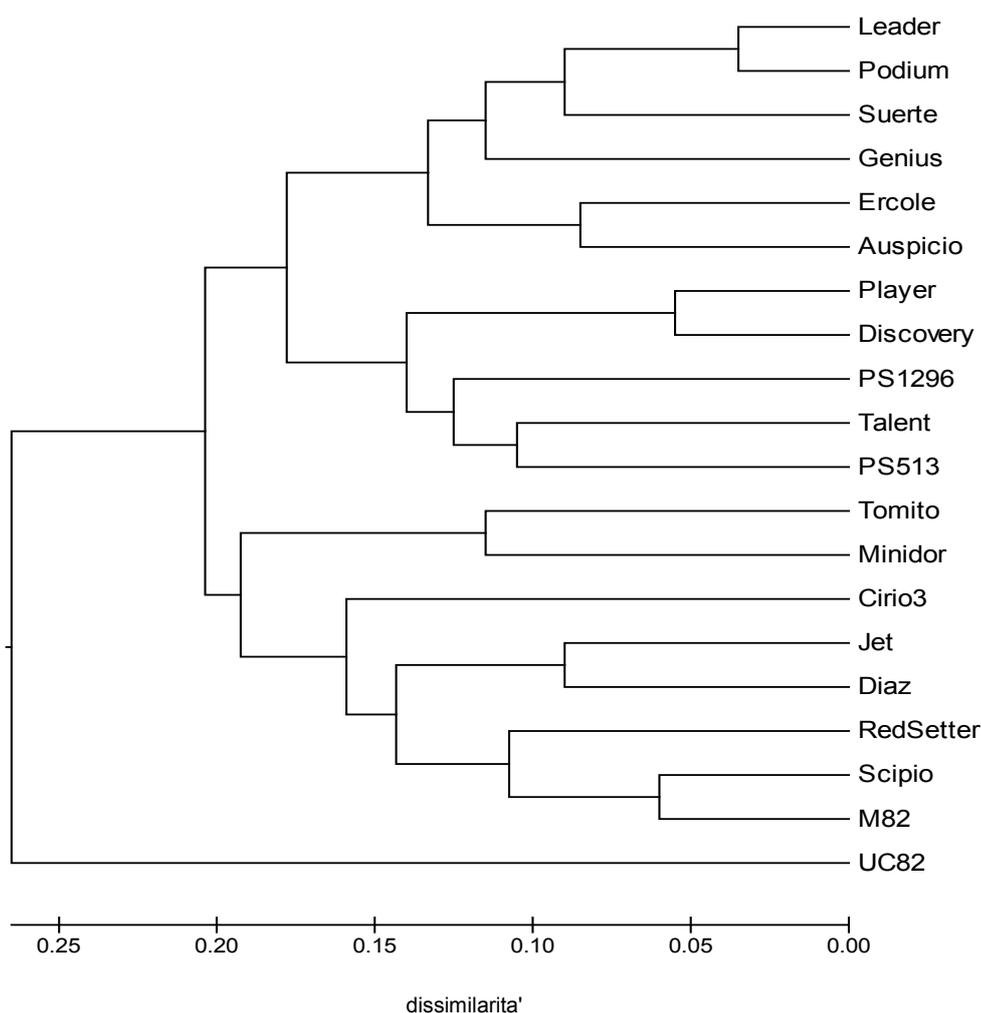


Figura 3.14: Dendrogramma ottenuto utilizzando l'algoritmo Neighbor-Joining basato sulle distanze genetiche calcolate sui dati di 14 loci SSR analizzati su 20 varieta' commerciali. L'analisi e' stata condotta con il *software* MEGA 3.1.

Come si evince dal dendrogramma, tutte le varieta' commerciali in esame sono state discriminate mediante i quattordici loci SSR analizzati e lo stesso livello di discriminazione e' mantenuto mediante l'utilizzo di un set minimo di tre loci SSR

(LE21085, LEEF1Aa, LEta015) di cui i profili allelici relativi ad ogni cultivar sono riportati in tabella 3.14.

Tabella 3.14: Dimensioni degli alleli (bp) delle venti varietà di pomodoro analizzate per i tre loci SSR selezionati.

|           | SSR     |     |         |     |         |     |
|-----------|---------|-----|---------|-----|---------|-----|
|           | LE21085 |     | LEEF1Aa |     | LEta015 |     |
| Leader    | 102     | 102 | 126     | 126 | 106     | 112 |
| Ercole    | 116     | 116 | 126     | 126 | 106     | 112 |
| Talent    | 102     | 116 | 201     | 201 | 112     | 112 |
| PS1296    | 102     | 102 | 205     | 205 | 112     | 112 |
| Tomito    | 102     | 116 | 203     | 205 | 106     | 112 |
| Minidor   | 102     | 116 | 201     | 201 | 106     | 106 |
| Cirio3    | 102     | 102 | 201     | 201 | 106     | 106 |
| PS513     | 102     | 102 | 203     | 209 | 106     | 112 |
| Auspicio  | 102     | 116 | 126     | 126 | 106     | 106 |
| Podium    | 102     | 102 | 126     | 126 | 112     | 112 |
| Suerte    | 100     | 100 | 126     | 126 | 106     | 112 |
| Player    | 100     | 116 | 203     | 205 | 112     | 112 |
| Scipio    | 100     | 116 | 203     | 203 | 106     | 106 |
| Jet       | 102     | 116 | 126     | 126 | 106     | 112 |
| Discovery | 102     | 116 | 203     | 203 | 112     | 112 |
| Genius    | 100     | 100 | 203     | 212 | 112     | 112 |
| Diaz      | 100     | 100 | 203     | 209 | 112     | 112 |
| M82       | 116     | 116 | 207     | 207 | 106     | 106 |
| RedSetter | 116     | 116 | 203     | 203 | 106     | 106 |
| UC82      | 116     | 116 | 205     | 205 | 106     | 106 |

La media del contenuto polimorfico informativo per il set di tre loci SSR individuato ha un valore di 0,64, più alto del PIC medio di tutti i loci polimorfici (0,471). Questo set minimo di tre loci SSR risulta, quindi, sufficiente per identificare e discriminare tutte le varietà in esame.

### 3.2.2 Identificazione di SNP per la discriminazione varietale

L'analisi molecolare mediante un ulteriore marcatore, rappresentato dagli SNP, è stata effettuata sulle venti varietà commerciali già caratterizzate mediante SSR. Sono stati utilizzati 66 dei loci SNP riportati in tabella 2.6. La maggior parte delle varietà ha mostrato la presenza di entrambe le varianti alleliche dello SNP nell'ambito dei loci selezionati, ad indicare la presenza di genotipi eterozigoti, come atteso per la presenza di ibridi. La percentuale di transizioni (purina-purina, pirimidina-pirimidina) è stata del 58%, di cui il 65,5% è rappresentato da sostituzioni C/T ed il 34% da sostituzioni A/G. Del 42% di transversioni (purina-pirimidina e viceversa), il 42,8% è rappresentato da sostituzioni A/T, il 23,8% da G/T e C/A ed il 9,6% da G/C. L'analisi degli indici genetici ha permesso il calcolo della eterozigosità osservata il cui valore medio è di 0,172 ed il valore massimo è stato riscontrato per il locus CT10469 (0,556). Il numero totale di alleli ottenuti è stato di 116 per tutti i loci analizzati ed il valore massimo di due è assunto per i loci polimorfici, come atteso per loci biallelici. Dei 66 loci SNP analizzati, 16 sono risultati essere monomorfi e 50 polimorfici nell'ambito delle cultivars in esame e la percentuale di loci polimorfici è stata di 75,76%. Il PIC medio è di 0,25. In tabella 3.15 sono riassunti i principali indici genetici riportati per i soli loci risultati essere polimorfici.

Tabella 3.15: Principali indici genetici e loro media ottenuti da analisi dei loci SNP sulle varietà in esame.

| SSR     | N    | Na    | Ne           | I            | Ho           | PIC          |
|---------|------|-------|--------------|--------------|--------------|--------------|
| CT10004 | 18   | 2,000 | 1,314        | 0,403        | 0,056        | 0,239        |
| CT10018 | 19   | 2,000 | 1,296        | 0,389        | 0,053        | 0,229        |
| CT10024 | 20   | 2,000 | 1,280        | 0,377        | 0,050        | 0,219        |
| CT10043 | 20   | 2,000 | 1,471        | 0,500        | 0,200        | 0,320        |
| CT10075 | 20   | 2,000 | 1,663        | 0,588        | 0,250        | 0,399        |
| CT10120 | 20   | 2,000 | 1,956        | 0,682        | 0,350        | 0,489        |
| CT10136 | 20   | 2,000 | 1,835        | 0,647        | 0,400        | 0,455        |
| CT10153 | 18   | 2,000 | 1,117        | 0,215        | 0,111        | 0,105        |
| CT10162 | 19   | 2,000 | 1,054        | 0,122        | 0,053        | 0,051        |
| CT10184 | 19   | 2,000 | 1,362        | 0,436        | 0,000        | 0,266        |
| CT10187 | 19   | 2,000 | 1,232        | 0,336        | 0,211        | 0,188        |
| CT10192 | 17   | 2,000 | 1,125        | 0,224        | 0,118        | 0,111        |
| CT10206 | 20   | 2,000 | 1,105        | 0,199        | 0,000        | 0,095        |
| CT10238 | 19   | 2,000 | 1,994        | 0,692        | 0,316        | 0,499        |
| CT10322 | 20   | 2,000 | 1,051        | 0,117        | 0,050        | 0,049        |
| CT10347 | 18   | 2,000 | 1,906        | 0,668        | 0,444        | 0,475        |
| CT10373 | 18   | 2,000 | 1,994        | 0,692        | 0,500        | 0,498        |
| CT10375 | 19   | 2,000 | 1,111        | 0,206        | 0,000        | 0,100        |
| CT10385 | 20   | 2,000 | 1,835        | 0,647        | 0,300        | 0,455        |
| CT10386 | 20   | 2,000 | 1,471        | 0,500        | 0,000        | 0,320        |
| CT10390 | 19   | 2,000 | 1,054        | 0,122        | 0,053        | 0,051        |
| CT10396 | 20   | 2,000 | 1,724        | 0,611        | 0,200        | 0,420        |
| CT10401 | 20   | 2,000 | 1,923        | 0,673        | 0,300        | 0,480        |
| CT10432 | 18   | 2,000 | <b>2,000</b> | <b>0,693</b> | 0,111        | <b>0,500</b> |
| CT10437 | 14   | 2,000 | 1,774        | 0,628        | 0,071        | 0,436        |
| CT10450 | 19   | 2,000 | 1,915        | 0,671        | 0,263        | 0,478        |
| CT10469 | 18   | 2,000 | 1,976        | 0,687        | <b>0,556</b> | 0,494        |
| CT10485 | 20   | 2,000 | 1,724        | 0,611        | 0,400        | 0,420        |
| CT10494 | 18   | 2,000 | 1,600        | 0,562        | 0,278        | 0,375        |
| CT10526 | 20   | 2,000 | 1,923        | 0,673        | 0,500        | 0,480        |
| CT10556 | 20   | 2,000 | 1,956        | 0,682        | 0,450        | 0,489        |
| CT10591 | 18   | 2,000 | 1,180        | 0,287        | 0,056        | 0,153        |
| CT10639 | 20   | 2,000 | 1,995        | 0,692        | 0,350        | 0,499        |
| CT10678 | 20   | 2,000 | 1,342        | 0,423        | 0,100        | 0,255        |
| CT10781 | 20   | 2,000 | <b>2,000</b> | <b>0,693</b> | 0,400        | <b>0,500</b> |
| CT10790 | 20   | 2,000 | 1,105        | 0,199        | 0,100        | 0,095        |
| CT10793 | 20   | 2,000 | 1,600        | 0,562        | 0,400        | 0,375        |
| CT10801 | 20   | 2,000 | 1,220        | 0,325        | 0,200        | 0,180        |
| CT10807 | 20   | 2,000 | 1,923        | 0,673        | 0,400        | 0,480        |
| CT10809 | 19   | 2,000 | 1,430        | 0,478        | 0,158        | 0,301        |
| CT10909 | 19   | 2,000 | 1,870        | 0,658        | 0,421        | 0,465        |
| CT10915 | 20   | 2,000 | <b>2,000</b> | <b>0,693</b> | 0,400        | <b>0,500</b> |
| CT10963 | 20   | 2,000 | 1,342        | 0,423        | 0,300        | 0,255        |
| CT20017 | 20   | 2,000 | 1,724        | 0,611        | 0,400        | 0,420        |
| CT20023 | 20   | 2,000 | <b>2,000</b> | <b>0,693</b> | 0,000        | <b>0,500</b> |
| CT20028 | 20   | 2,000 | 1,956        | 0,682        | 0,450        | 0,489        |
| CT20059 | 18   | 2,000 | 1,314        | 0,403        | 0,278        | 0,239        |
| CT20097 | 20   | 2,000 | 1,406        | 0,464        | 0,150        | 0,289        |
| CT20159 | 20   | 2,000 | 1,105        | 0,199        | 0,100        | 0,095        |
| CT20181 | 19   | 2,000 | 1,170        | 0,276        | 0,053        | 0,145        |
| media   | 19,4 | 1,76  | 1,43         | 0,37         | 0,17         | 0,25         |

Note: N, numero di varietà analizzate; Na, numero di alleli; Ne, numero effettivo di alleli; I, indice d'informazione. Ho, eterozigotità osservata; PIC, contenuto polimorfico informativo. Sono riportati i soli loci polimorfici ed i valori più elevati degli indici genetici sono riportati in grassetto. Valori di N inferiori a 20 sono dovuti alla presenza di dati mancanti.

Il dendrogramma in figura 3.15 generato dalla analisi dei dati SNP ha mostrato che tutte le varietà commerciali sono discriminate, fatta eccezione per le cultivars 'Red Setter' e 'UC82'.

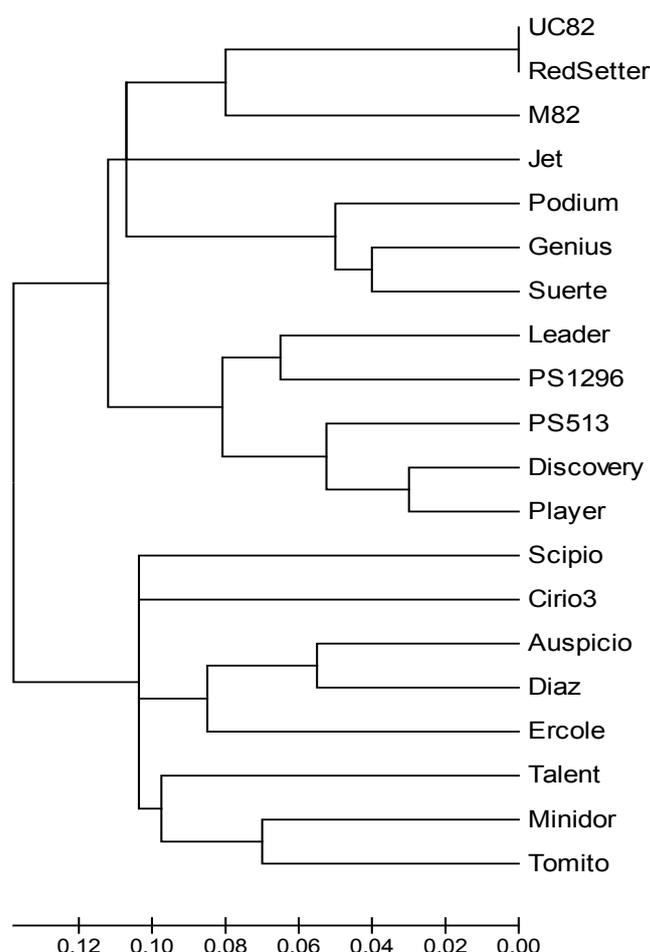


Figura 3.15: Dendrogramma ottenuto utilizzando l'algoritmo Neighbor-Joining basato sulle distanze genetiche calcolate sui dati di 66 loci SNP analizzati su 20 varietà. In scala è riportato il grado di dissimilarità.

### 3.2.3. Analisi mediante in/del

L'analisi delle venti varietà commerciali è proseguita valutando il potere discriminante di un'altra tipologia di marcatori molecolari, rappresentata da inserzioni e delezioni di frammenti di DNA in un range di peso molecolare compreso tra 12 e 82 bp. L'analisi di 6 in/del ha rilevato che tutti i loci sono risultati essere polimorfici (100%) con un numero di alleli totali di 12. I valori medi di  $H_o$  e PIC sono rispettivamente di 0,308 e 0,394 con valori massimi rilevati per il locus CT103731 (Tabella 3.16).

Tabella 3.16: Principali indici genetici e loro media ottenuti da analisi mediante 6 in/del sulle varietà.

| in/del   | N  | Na    | Ne           | I            | Ho           | PIC          |
|----------|----|-------|--------------|--------------|--------------|--------------|
| CT10042I | 20 | 2,000 | 1,782        | 0,631        | 0,350        | 0,439        |
| CT10184I | 20 | 2,000 | 1,724        | 0,611        | 0,300        | 0,420        |
| CT10373I | 20 | 2,000 | <b>1,980</b> | <b>0,688</b> | <b>0,600</b> | <b>0,495</b> |
| CT10737I | 20 | 2,000 | 1,406        | 0,464        | 0,150        | 0,289        |
| CT20210I | 20 | 2,000 | 1,471        | 0,500        | 0,300        | 0,320        |
| CT20268I | 20 | 2,000 | 1,663        | 0,588        | 0,150        | 0,399        |
| media    | 20 | 2,000 | 1,671        | 0,580        | 0,308        | 0,394        |

Note: N, numero di varietà analizzate; Na, numero di alleli; Ne, numero effettivo di alleli; I, indice d'informazione. Ho, eterozigotà osservata; PIC, contenuto polimorfico informativo. I valori più elevati degli indici genetici sono riportati in grassetto.

La rappresentazione grafica dei dati in figura 3.16 mostra che questo set di marcatori non è risultato capace di discriminare le cultivars 'UC82', 'Red Setter', 'M82' tra di loro, come anche le varietà 'Auspicio' e 'Talent' e le varietà 'Podium' e 'PS1296' risultano indistinguibili.

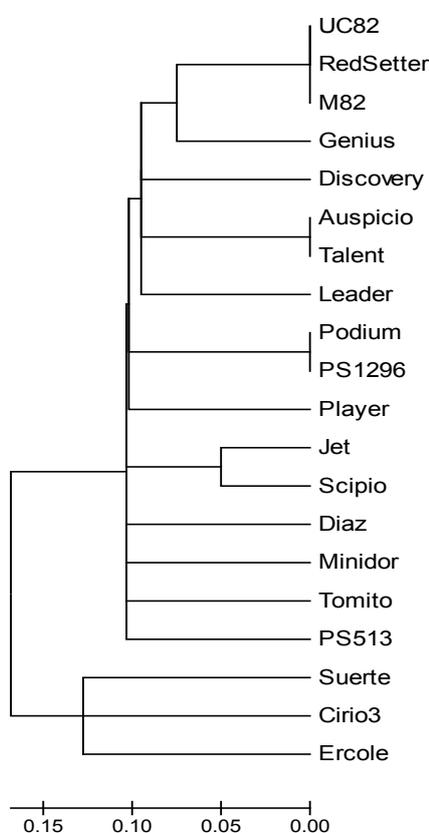


Figura 3.16: Dendrogramma ottenuto utilizzando l'algoritmo Neighbor-Joining basato sulle distanze genetiche calcolate sui dati di 6 loci in/del analizzati su 20 varietà. In scala è riportato il grado di dissimilarità.

### 3.2.4 Correlazione tra i marcatori molecolari

E' stata calcolata la correlazione esistente tra i marcatori molecolari (14 SSR, 66 SNP, 6 in/del) utilizzati per l'analisi delle venti varietà. I coefficienti di correlazione calcolati tra le matrici genetiche non hanno rilevato valori di significatività tra i marcatori in esame (Tabella 3.17).

Tabella 3.17: Confronto a coppie tra i marcatori molecolari. Sono riportati i valori di correlazione genetica tra le matrici di distanza genetica.

|        | Dati molecolari |      |        |
|--------|-----------------|------|--------|
|        | SSR             | SNP  | in/del |
| SSR    | 1               |      |        |
| SNP    | 0,01            | 1    |        |
| in/del | 0               | 0,47 | 1      |

### 3.3 Confronto tra marcatori molecolari SSR, SNP, in/del in accessioni di pomodoro coltivato

Le caratteristiche ed il potere discriminante di tre diversi marcatori molecolari sono stati confrontati su due diverse popolazioni di pomodoro coltivato. SSR, SNP e in/del sono stati utilizzati per l'analisi di 20 ecotipi e 20 varietà commerciali. I risultati derivanti dal confronto sono riportati in tabella 3.18.

Tabella 3.18: Confronto tra tre marcatori molecolari nell'analisi di due diverse popolazioni, ecotipi (20) e varietà commerciali (20). Tipologia di marcatore, numero di loci analizzati, numero di alleli ottenuti, percentuale di loci polimorfici, contenuto polimorfico informativo (PIC) medio.

| Marcatore | Varietà               |           |                    |                             | Ecotipi               |           |                    |                             | PIC medio tra popolazioni   |
|-----------|-----------------------|-----------|--------------------|-----------------------------|-----------------------|-----------|--------------------|-----------------------------|-----------------------------|
|           | N. loci               | N. alleli | % loci polimorfici | PIC medio                   | N. loci               | N. alleli | % loci polimorfici | PIC medio                   |                             |
| SSR       | 14 (13 <sup>a</sup> ) | 41        | 92,86              | 0,441 (0,474 <sup>b</sup> ) | 14 (13 <sup>a</sup> ) | 46        | 92,86              | 0,447 (0,481 <sup>b</sup> ) | 0,444 (0,477 <sup>b</sup> ) |
| SNP       | 66 (50 <sup>a</sup> ) | 116       | 75,76              | 0,250 (0,328 <sup>b</sup> ) | 66 (14 <sup>a</sup> ) | 80        | 21,21              | 0,059 (0,277 <sup>b</sup> ) | 0,154 (0,302 <sup>b</sup> ) |
| in/del    | 6 (6 <sup>a</sup> )   | 12        | 100                | 0,39                        | 6 (1 <sup>a</sup> )   | 7         | 16,67              | 0,080 (0,480 <sup>b</sup> ) | 0,237 (0,437 <sup>b</sup> ) |
| media     |                       |           |                    | 0,360 (0,397 <sup>b</sup> ) |                       |           |                    | 0,195 (0,412 <sup>b</sup> ) | 0,278 (0,405 <sup>b</sup> ) |

<sup>a</sup> numero di loci polimorfici

<sup>b</sup> PIC medio considerando i soli loci polimorfici

La percentuale di loci SSR polimorfici ed il valore di PIC risultano essere elevati sia nell'ambito delle varietà che degli ecotipi. Per quando riguarda i marcatori SNP ed in/del la percentuale di loci polimorfici ed il PIC assumono valori elevati solo nell'ambito delle varietà. In particolare, se per il calcolo del contenuto polimorfico informativo vengono considerati i soli loci polimorfici, questo indice genetico assume valori notevolmente maggiori soprattutto nell'ambito degli ecotipi per i quali la maggior percentuale di loci SNP ed in/del era risultata essere monomorfica. Considerando che il PIC per loci multiallelici, quali gli SSR, varia da 0 a 1, mentre per loci biallelici, quali gli SNP, varia da 0 a 0,5, entrambi i marcatori hanno mostrato possedere un contenuto polimorfico informativo medio elevato. Gli SSR sono, però,

risultati capaci di identificare e discriminare tutto il materiale in esame. Di contro, gli SNP falliscono nella discriminazione di alcuni genotipi sia nell'ambito delle varietà che degli ecotipi. Per quanto riguarda i marcatori in/del, pur mostrando la massima percentuale di loci polimorfici nell'ambito delle varietà, hanno mostrato possedere un basso livello di discriminazione per l'analisi del materiale in esame. I risultati ottenuti hanno rilevato che i marcatori SSR hanno i più alti livelli di polimorfismo e il più alto potere discriminante nell'ambito del germoplasma analizzato.

### 3.4 Relazioni genetiche tra pomodoro coltivato e specie selvatiche mediante analisi SSR

I venti ecotipi di pomodoro diffusi in Campania sono stati confrontati con le venti varietà commerciali in esame e cinque relativi selvatici del pomodoro (Tabella 2.1) sono stati inclusi come possibile riferimento esterno. Il polimorfismo del DNA è stato valutato in 14 loci SSR e sono stati confrontati i principali indici genetici tra le tre popolazioni in esame (Tabella 3.19).

Tabella 3.19: Principali indici genetici (media) dei genotipi analizzati.

|                   | Na    | Ne    | I     | Ho    | PIC   |
|-------------------|-------|-------|-------|-------|-------|
| Ecotipi           | 3,286 | 1,980 | 0,787 | 0,000 | 0,447 |
| Varietà           | 2,929 | 2,051 | 0,749 | 0,300 | 0,441 |
| Specie selvatiche | 4,857 | 4,235 | 1,380 | 0,271 | 0,674 |
| Tutti i genotipi  | 3,690 | 2,755 | 0,972 | 0,192 | 0,520 |

Note: Na, numero di alleli; Ne, numero effettivo di alleli; I, indice d'informazione. Ho, eterozigotità osservata; PIC, contenuto polimorfico informativo.

Gli indici indicano che nonostante le varietà commerciali siano nella maggior parte ibridi, gli ecotipi campani posseggono un numero medio di alleli di poco superiore a quello degli ibridi. Come atteso, la diversità genetica esistente tra le specie selvatiche di pomodoro è superiore a quella esistente tra le varietà della stessa specie. L'Analisi delle Coordinate Principali basata sulla conversione della matrice delle distanze genetiche in una matrice di covarianza (Figura 3.17), indica che gli ecotipi campani si raggruppano in una posizione che è chiaramente diversa da quella delle varietà commerciali, sottolineando che esistono dei fattori di natura genetica che distinguono tali materiali. Le varietà risultano, invece, essere maggiormente correlate alle specie selvatiche analizzate, che ne rappresentano la fonte di geni di resistenza a stress biotici ed abiotici.

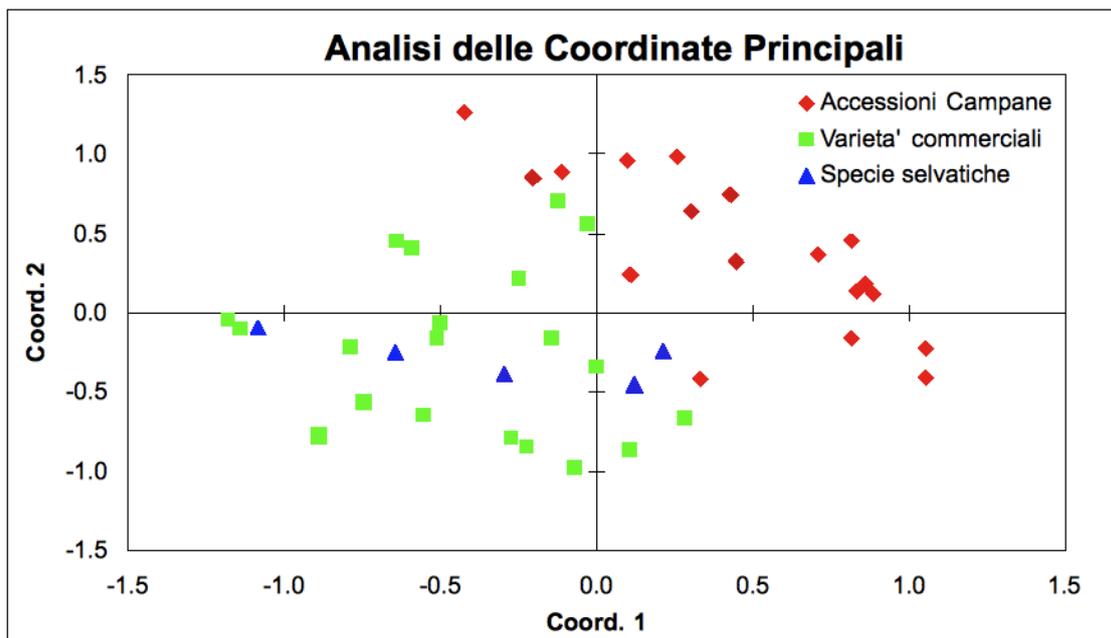


Figura 3.17: Analisi delle Coordinate Principali basata sui dati dei microsatelli. Il grafico mostra le relazioni esistenti tra le tre popolazioni di pomodoro in esame. Gli ecotipi sono raggruppati separatamente, mentre le specie selvatiche raggruppano nel *cluster* rappresentato dalle varietà.

### 3.5 Analisi della struttura delle popolazioni

Considerando i risultati dell'analisi SSR, si è voluto verificare se gli ecotipi campani rappresentino una popolazione distinta da altri materiali simili, quali varietà coltivate non ibride ed accessioni di origine geografica diversa. Si è pertanto esteso il numero dei materiali campani a 43 ecotipi di pomodoro (Tabella 2.1) che sono stati confrontati con la collezione di 96 accessioni di diversa tipologia e provenienza (Tabella 2.2), facenti parte della "Core Collection" del laboratorio del prof. David Francis dell'Ohio State University (USA) ([www.tomatomap.net](http://www.tomatomap.net); Francis et al., 2006). Per motivi legati principalmente al numero dei campioni, le analisi sono state condotte utilizzando 6 marcatori in/del e 71 loci SNP (Tabelle 2.5 e 2.6, rispettivamente). I risultati derivanti dall'analisi molecolare sono stati sottoposti ad analisi statistica al fine di identificare il numero minimo di popolazioni (K) capace di catturare la maggiore struttura dei dati. Per ogni valore di K testato (da 2 a 16) sono state effettuate 20 ripetizioni. La figura 3.18 mostra l'andamento della curva costruita sulla probabilità ( $\ln Pr(X/K)$ ) che ogni genotipo appartenga alla data popolazione per ognuna delle 20 analisi effettuate per ogni valore di K. Il coefficiente di correlazione  $r = 0,96$  indica la ripetibilità e la robustezza dei risultati ottenuti.

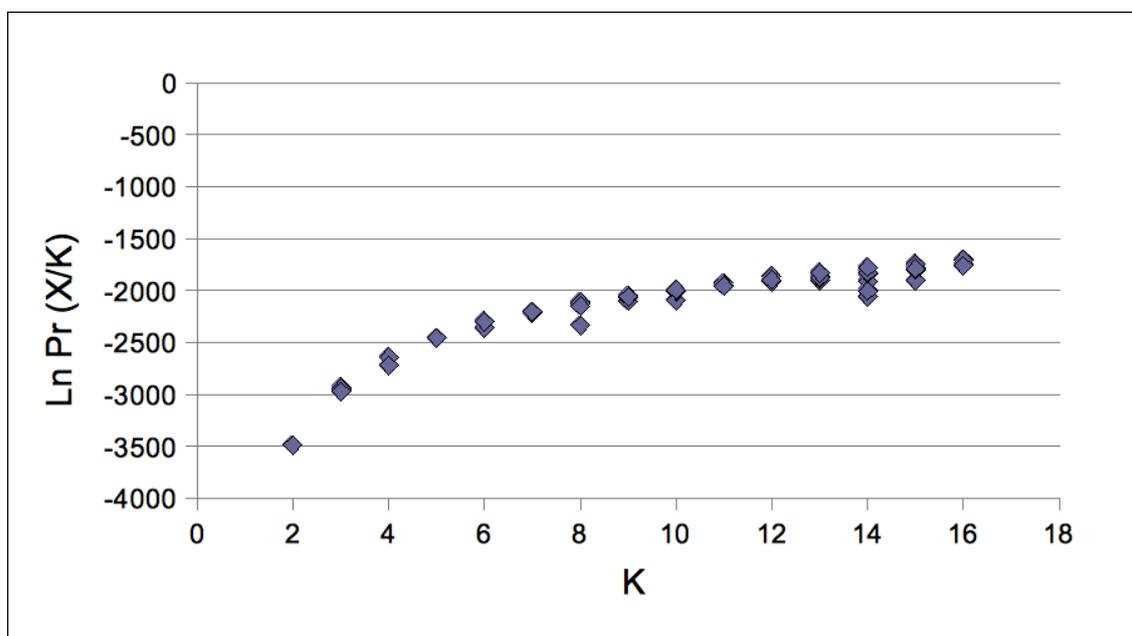


Figura 3.18: Grafico che riporta i valori di probabilita' dei dati ( $\ln \text{Pr} (X/K)$ ) ottenuti mediante analisi di 71 loci SNP e 6 in/del utilizzando il *software* "Structure" v2.1 in funzione dei raggruppamenti ipotizzati. Sono riportati i valori per ognuna delle 20 ripetizioni effettuate per ognuno dei valori di K testati (2 a 16) utilizzando un *burn-in period* di 50000 e una *routine* MCMC di 500000 ripetizioni.

La piu' elevata probabilita' a posteriori ( $\ln \text{Pr} (X/K)$ ) che ogni genotipo appartenga ad una data popolazione e' stato ottenuto per il valore di  $K=13$ , che rappresenta il numero minimo di popolazioni esistenti che raggruppa la maggiore struttura dei dati. La figura 3.19 riporta una rappresentazione grafica del numero di popolazioni riscontrate. Risulta evidente che il gruppo degli ecotipi campani (ad eccezione dei tipi 'Sorrento') rappresenta una popolazione indipendente e geneticamente diversa dal resto del materiale in esame.

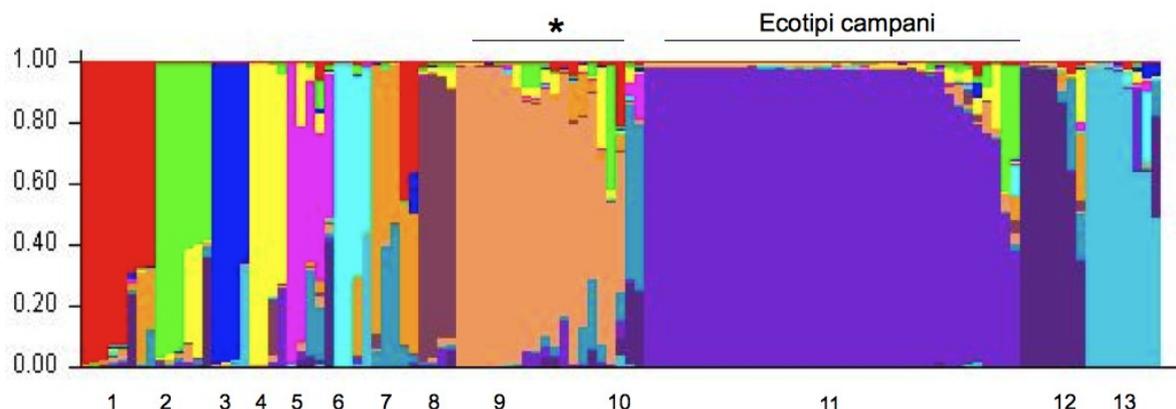


Figura 3.19: Classificazione degli individui effettuata mediante il programma "Structure". La distribuzione degli individui nei 13 differenti raggruppamenti e' indicata dai 13 diversi colori. Ogni individuo e' rappresentato da una linea verticale, che e' ripartita in diversi segmenti colorati la cui altezza e' funzione della probabilita' stimata di appartenenza del singolo individuo ad uno dei 13 raggruppamenti.

1: accessioni provenienti dalla Florida (USA) destinate al mercato fresco.

2, 3, 4: accessioni *heirloom* provenienti dagli USA.

5: accessioni provenienti dal North Carolina (USA) destinate al mercato fresco.

6: varieta' da industria non ibride destinate alla trasformazione.

7: accessioni provenienti dagli USA destinate al mercato fresco.

8: accessioni provenienti dall'Oregon (USA) destinate al mercato fresco.

9, 10: ecotipi della tipologia "Sorrento" campane e accessioni di forma globosa provenienti dagli USA destinate al mercato fresco.

11: ecotipi campani destinati al mercato fresco tra cui le tipologie 'Agro-Nocerino', 'San Marzano', 'Principe Borghese' ed *heirloom* 'San Marzano' e 'Principe Borghese' della "Core Collection".

12, 13: accessioni provenienti dall'Ohio (USA) destinate alla trasformazione.

Nota: l'asterisco e' indicativo di un raggruppamento che comprende i tipi 'Sorrento' campane ed altri genotipi dai frutti di forma globosa della "Core Collection".

Insieme a questo cluster rappresentativo degli ecotipi campani raggruppano anche i tipi 'Principe Borghese' e 'San Marzano' appartenenti alla "Core Collection" e di provenienza italiana. Le tipologie 'Sorrento', invece, facenti parte del germoplasma campano, raggruppano separatamente insieme ad altri genotipi dai frutti di forma globosa della "Core Collection". Tale raggruppamento e' mantenuto anche per valori di K inferiori a 13 (K 9 a 13). Gli altri raggruppamenti sembrano essere correlati alla tipologia di materiale, alla destinazione d'uso ed alla loro provenienza.

### 3.6 Tracciabilità della filiera agro-alimentare del pomodoro

#### 3.6.1 Sviluppo di un protocollo di estrazione e amplificazione di DNA genomico da pomodoro trasformato

Nell'ambito della tracciabilità della filiera del pomodoro sono state monitorate le fasi meno sicure in cui potrebbe essere compromessa la genuinità del prodotto ossia: "a monte" tramite la certificazione varietale del materiale in campo, "in produzione" mediante l'autenticazione delle bacche ed "a valle" con la tracciabilità del prodotto finito. Sono state selezionate sette varietà tra quelle analizzate e fornite dalla ditta di trasformazione "Annalisa" Lodato S.p.A. (Tabella 2.1) di cui il DNA genomico da foglia e bacca è stato estratto e quantificato. Tutte le amplificazioni PCR hanno generato frammenti specifici di peso molecolare atteso. Il DNA estratto dai prodotti trasformati è stato direttamente sottoposto ad amplificazione in quanto la quantità di DNA estraibile è risultata essere talmente esigua da impedirne la quantificazione tramite visualizzazione su gel di agarosio. Pertanto, si è proceduto con l'amplificazione mediante PCR per verificarne la presenza e l'analizzabilità.

Dal momento in cui il DNA è altamente processato e, quindi, presumibilmente degradato, sono state effettuate prove di amplificazione. Si sono impiegati quattro primer che riconoscono un gene a singola copia del genoma di pomodoro, l'Elongation Factor 1-alfa, le cui combinazioni generano frammenti di circa 680 bp, 400 bp e di circa 100 bp (Figura 3.20).

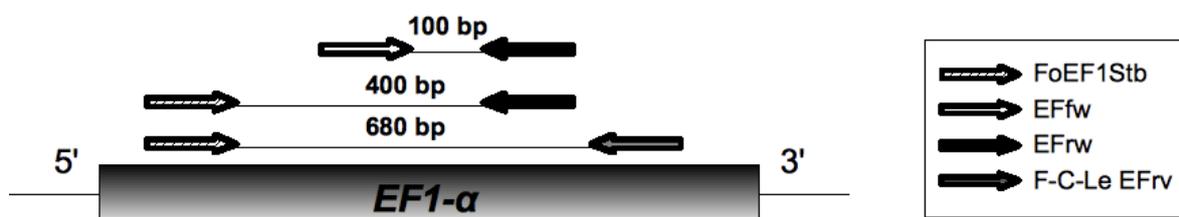


Figura 3.20: Schema di appaiamento che mostra le diverse combinazioni dei primer EF sullo stesso gene (EF1- $\alpha$ ). Le coppie di primer FoEF1Stb e F-C-Le EFrw generano frammenti di 680 bp; FoEF1Stb e EF rw generano frammenti di 400 bp; EF fw e EF rw generano frammenti di 100 bp.

L'analisi su gel di agarosio ed il confronto con gli amplificati derivanti dai rispettivi DNA estratti da foglia utilizzati come controlli positivi, rivelano l'assenza di amplificati delle dimensioni di 680 bp e 400 bp nei campioni di materiale processato, ma la presenza di una banda ad un peso molecolare di 100 bp (Figura 3.21).

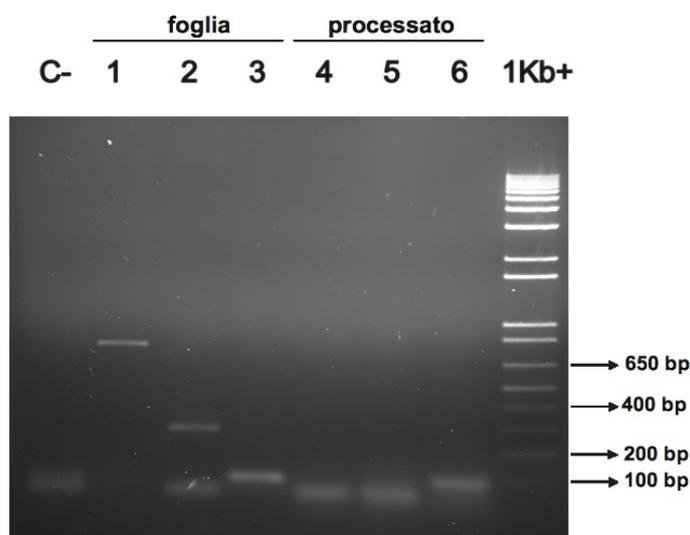


Figura 3.21: Elettroforesi in gel di agarosio delle PCR per la verifica dell'amplificabilità del DNA isolato da pelato. 1, 2, 3: DNA estratto da foglia amplificato con le coppie di primer FoEF1Stb e F-C-Le EFrw (che danno frammenti di 680 bp), FoEF1Stb e EF rw (che danno frammenti di 400 bp) e EF fw e EF rw (che danno frammenti di 100 bp), rispettivamente; 4, 5, 6: DNA estratto da pelato amplificato con le coppie di primer FoEF1Stb e F-C-Le EFrw (che danno frammenti di 680 bp), FoEF1Stb e EF rw (che danno frammenti di 400 bp) e EF fw e EF rw (che danno frammenti di 100 bp), rispettivamente; C-= Controllo negativo, 1Kb+= 1Kb plus ladder (Invitrogen).

Questo risultato conferma che esiste un limite nelle dimensioni del DNA amplificabile quando il DNA stampo proviene da bacche di pomodoro altamente processate nella filiera industriale. Le amplificazioni sono state eseguite variando la quantità di DNA utilizzato nella reazione di amplificazione, il numero dei cicli (da 35, per il DNA da foglia e da bacca, a 40 per il materiale processato) ed i tempi di *annealing* ed estensione al fine di individuare le condizioni ottimali di amplificazione che sono risultate essere diverse da quelle del DNA isolato da foglia e da bacca. Messe a punto le condizioni ottimali di amplificazione, si è proceduto con l'amplificazione mediante i loci selezionati.

### 3.6.2 Utilizzo degli SSR selezionati per la tracciabilità molecolare dei prodotti di filiera

Al fine di procedere con gli studi di tracciabilità, sia le bacche che i vari prodotti inscatolati sono stati forniti ed etichettati dalla ditta di trasformazione "Annalisa". Tali materiali fanno parte della stessa filiera di produzione delle piante analizzate. Il set di 14 loci SSR selezionato per l'identificazione e la discriminazione delle varietà commerciali è stato utilizzato per analizzare DNA estratto da bacca e prodotto processato fornito come monovarietale delle sette varietà in esame. I primer SSR selezionati generano frammenti di peso molecolare compreso in un *range* di 94-233 bp. L'amplificazione PCR del DNA estratto da bacca delle sette varietà da industria in esame ha generato frammenti delle dimensioni attese per tutti i loci analizzati. Per il DNA estratto da materiale processato, l'analisi su gel di agarosio mostra una buona amplificazione per nove dei 14 loci SSR, che hanno generato prodotti di amplificazione delle dimensioni attese in un *range* compreso tra 94-164 bp, mentre gli altri cinque loci hanno fallito nell'amplificazione probabilmente a causa del più elevato peso molecolare dei frammenti attesi. In figura 3.22 è mostrato un esempio di amplificazione del DNA estratto da prodotto processato delle sette varietà da

industria in esame al locus LEctt001 che genera frammenti di peso molecolare atteso di circa 100 bp.

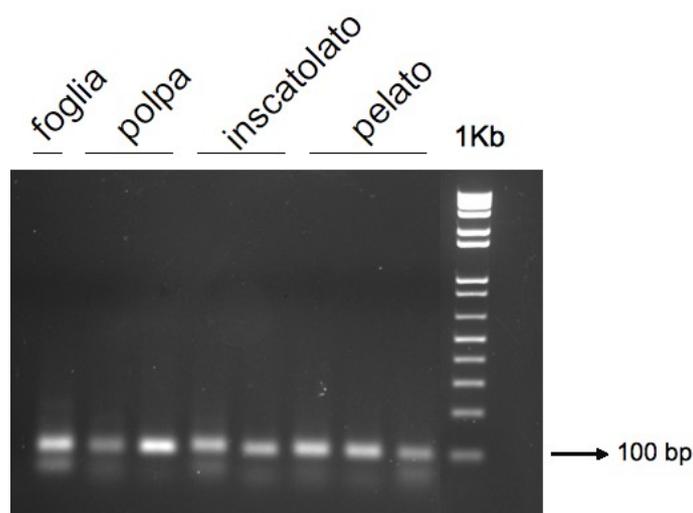


Figura 3.22: Elettroforesi in gel di agarosio dei prodotti di amplificazione, al locus SSR LEctt001, del DNA estratto da materiale processato di tre differenti tipologie (polpa, pomodorino inscatolato, pelato) e della foglia, utilizzata come controllo positivo; 1Kb= Ladder.

I prodotti di amplificazione ottenuti sono stati, pertanto, separati mediante elettroforesi capillare al fine di verificare la corrispondenza in termini di dimensioni alleliche degli amplificati del DNA estratto da prodotto processato con gli amplificati del DNA estratto da foglia e bacca della stessa varietà. In tabella 3.20 sono riportati i profili elettroforetici relativi ai nove loci SSR per la foglia, la bacca ed il prodotto trasformato di ognuna delle sette varietà in esame.

Tabella 3.20: Profili allelici da foglia, bacca e prodotto processato fornito come monovarietale delle sette varietà da industria generati mediante analisi di nove loci SSR. In grassetto sono riportate le varietà tra cui è stata riscontrata corrispondenza tra i profili allelici dei tre segmenti della filiera in esame.

| Varietà'       | SSR            |           |            |           |           |           |            |            |            |            |            |            |            |            |            |            |            |            |            |
|----------------|----------------|-----------|------------|-----------|-----------|-----------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
|                | LE ct001       |           | LE aat007  |           | LE ct001  |           | LE caa001  |            | LE ta003   |            | LE ta015   |            | LE 20592   |            | LE 21085   |            | LE aat002  |            |            |
| Foglia         | Leader         | 94        | 103        | 96        | 96        | 97        | 97         | 100        | 100        | 102        | 106        | 106        | 112        | 161        | 161        | 102        | 102        | 101        | 101        |
|                | <b>Ercole</b>  | <b>94</b> | <b>103</b> | <b>96</b> | <b>96</b> | <b>97</b> | <b>97</b>  | <b>100</b> | <b>100</b> | <b>106</b> | <b>106</b> | <b>106</b> | <b>112</b> | <b>161</b> | <b>164</b> | <b>116</b> | <b>116</b> | <b>101</b> | <b>101</b> |
|                | Talent         | 94        | 103        | 99        | 99        | 97        | 97         | 100        | 100        | 102        | 106        | 112        | 112        | 164        | 164        | 102        | 116        | 104        | 104        |
|                | PS 1296        | 94        | 94         | 96        | 99        | 97        | 97         | 100        | 100        | 102        | 106        | 112        | 112        | 161        | 164        | 102        | 102        | 101        | 104        |
|                | <b>Tomito</b>  | <b>97</b> | <b>100</b> | <b>96</b> | <b>99</b> | <b>97</b> | <b>97</b>  | <b>100</b> | <b>100</b> | <b>94</b>  | <b>106</b> | <b>106</b> | <b>112</b> | <b>161</b> | <b>164</b> | <b>102</b> | <b>116</b> | <b>101</b> | <b>104</b> |
|                | <b>Minidor</b> | <b>97</b> | <b>97</b>  | <b>99</b> | <b>99</b> | <b>97</b> | <b>97</b>  | <b>100</b> | <b>100</b> | <b>106</b> | <b>106</b> | <b>106</b> | <b>106</b> | <b>161</b> | <b>161</b> | <b>102</b> | <b>116</b> | <b>104</b> | <b>104</b> |
| <b>Cirio 3</b> | <b>94</b>      | <b>94</b> | <b>99</b>  | <b>99</b> | <b>97</b> | <b>97</b> | <b>100</b> | <b>100</b> | <b>102</b> | <b>102</b> | <b>106</b> | <b>106</b> | <b>161</b> | <b>161</b> | <b>102</b> | <b>102</b> | <b>104</b> | <b>104</b> |            |
| Bacca          | Leader         | 94        | 94         | 96        | 99        | 97        | 97         | 100        | 100        | 102        | 106        | 112        | 112        | 161        | 164        | 102        | 102        | 101        | 104        |
|                | <b>Ercole</b>  | <b>94</b> | <b>103</b> | <b>96</b> | <b>96</b> | <b>97</b> | <b>97</b>  | <b>100</b> | <b>100</b> | <b>106</b> | <b>106</b> | <b>112</b> | <b>161</b> | <b>164</b> | <b>116</b> | <b>116</b> | <b>101</b> | <b>101</b> |            |
|                | Talent         | 94        | 103        | 99        | 99        | 97        | 97         | 100        | 100        | 102        | 106        | 112        | 112        | 164        | 164        | 102        | 116        | 104        | 104        |
|                | PS 1296        | 94        | 94         | 96        | 96        | 97        | 97         | 100        | 100        | 106        | 106        | 112        | 112        | 161        | 164        | 102        | 102        | 101        | 101        |
|                | <b>Tomito</b>  | <b>97</b> | <b>100</b> | <b>96</b> | <b>99</b> | <b>97</b> | <b>97</b>  | <b>100</b> | <b>100</b> | <b>94</b>  | <b>106</b> | <b>106</b> | <b>112</b> | <b>161</b> | <b>164</b> | <b>102</b> | <b>116</b> | <b>101</b> | <b>104</b> |
|                | <b>Minidor</b> | <b>97</b> | <b>97</b>  | <b>99</b> | <b>99</b> | <b>97</b> | <b>97</b>  | <b>100</b> | <b>100</b> | <b>106</b> | <b>106</b> | <b>106</b> | <b>106</b> | <b>161</b> | <b>161</b> | <b>102</b> | <b>116</b> | <b>104</b> | <b>104</b> |
| <b>Cirio 3</b> | <b>94</b>      | <b>94</b> | <b>99</b>  | <b>99</b> | <b>97</b> | <b>97</b> | <b>100</b> | <b>100</b> | <b>102</b> | <b>102</b> | <b>106</b> | <b>106</b> | <b>161</b> | <b>161</b> | <b>102</b> | <b>102</b> | <b>104</b> | <b>104</b> |            |
| Processato     | Leader         | 94        | 103        | 96        | 99        | 97        | 97         | 100        | 100        | 102        | 106        | 106        | 112        | 161        | 164        | 102        | 102        | 101        | 104        |
|                | Ercole         | 94        | 94         | 96        | 99        | 97        | 97         | 100        | 100        | 102        | 104        | 106        | 112        | 164        | 167        | 102        | 102        | 101        | 104        |
|                | <b>Talent</b>  | <b>94</b> | <b>103</b> | <b>96</b> | <b>96</b> | <b>97</b> | <b>97</b>  | <b>100</b> | <b>100</b> | <b>106</b> | <b>106</b> | <b>106</b> | <b>112</b> | <b>161</b> | <b>164</b> | <b>116</b> | <b>116</b> | <b>101</b> | <b>101</b> |
|                | PS 1296        | 94        | 94         | 96        | 99        | 97        | 97         | 100        | 100        | 102        | 106        | 106        | 112        | 161        | 164        | 102        | 102        | 101        | 101        |
|                | <b>Tomito</b>  | <b>97</b> | <b>100</b> | <b>96</b> | <b>99</b> | <b>97</b> | <b>97</b>  | <b>100</b> | <b>100</b> | <b>94</b>  | <b>106</b> | <b>106</b> | <b>112</b> | <b>161</b> | <b>164</b> | <b>102</b> | <b>116</b> | <b>101</b> | <b>104</b> |
|                | <b>Minidor</b> | <b>97</b> | <b>97</b>  | <b>99</b> | <b>99</b> | <b>97</b> | <b>97</b>  | <b>100</b> | <b>100</b> | <b>106</b> | <b>106</b> | <b>106</b> | <b>106</b> | <b>161</b> | <b>161</b> | <b>102</b> | <b>116</b> | <b>104</b> | <b>104</b> |
| <b>Cirio 3</b> | <b>94</b>      | <b>94</b> | <b>99</b>  | <b>99</b> | <b>97</b> | <b>97</b> | <b>100</b> | <b>100</b> | <b>102</b> | <b>102</b> | <b>106</b> | <b>106</b> | <b>161</b> | <b>161</b> | <b>102</b> | <b>102</b> | <b>104</b> | <b>104</b> |            |

L'analisi mediante il sistema di elettroforesi capillare ha confermato la corrispondenza tra i profili allelici del DNA da foglia, bacca e rispettivo prodotto processato per le varietà 'Tomito', 'Minidor' e 'Cirio3'. In figura 3.23 sono riportati esempi di elettroferogrammi generati dall'analisi del locus LEta015 che mostrano la corrispondenza tra i profili allelici ottenuti dalle tre matrici (foglia, bacca e rispettivo prodotto processato monovarietale) della varietà 'Tomito'.

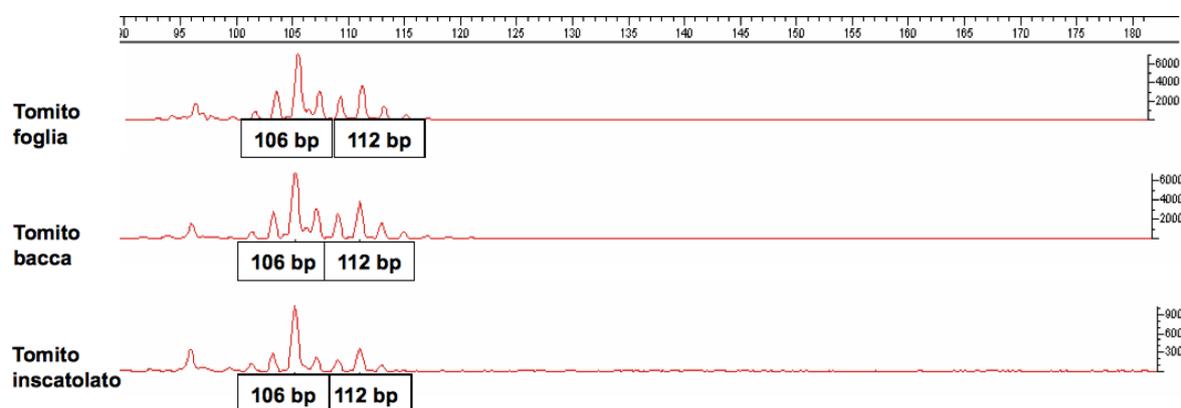


Figura 3.23: Esempi di elettroferogrammi del DNA genomico da foglia, bacca e prodotto trasformato monovarietale amplificato con il locus SSR LEta015 per la varietà ibrida da industria 'Tomito'. Analisi dei dati effettuata mediante i software "ABI PRISM Genotyper 2.5" (v. 3.7) e "GeneScan" (Applied Biosystems) usando come standard di peso molecolare il GeneScan 400 HD.

Per le altre varietà in esame non è stata riscontrata correlazione tra i profili allelici ottenuti per la foglia con quelli delle rispettive bacche e prodotto processato. In particolare, per la varietà 'Talent' non è stata riscontrata correlazione tra il profilo allelico ottenuto per la foglia e la bacca quando messo a confronto con quello ottenuto per il prodotto processato. Ci sono casi in cui si osserva una dimensione degli alleli minore ed in altri maggiore per alcuni dei loci analizzati. Il pelato etichettato come 'Talent' ha il profilo allelico della varietà 'Erocole'. 'Erocole' e 'Talent' sono varietà di pomodoro che condividono la stessa forma del frutto e destinazione d'uso. In questo caso è, quindi, ipotizzabile uno scambio tra le due varietà in uno dei passaggi della filiera di produzione. Per quanto riguarda la non corrispondenza nei tre segmenti della filiera in esame per le altre varietà è ipotizzabile, invece, un possibile mescolamento dei tipi in lavorazione.

### 3.6.3 Validazione sperimentale degli EST-SNP mediante conversione in CAPS e sequenziamento

Per lo sviluppo di sistemi innovativi e/o ad alta produttività per la tracciabilità di filiera particolare rilievo è dato ai marcatori molecolari che rilevano polimorfismi di una singola base del DNA. Per questo motivo, un ulteriore approccio seguito è stato la validazione di EST-SNP di *S. lycopersicum* selezionati dalla letteratura (Yang et al., 2004), che posseggono una elevata probabilità (53%) di identificare polimorfismi tra le varietà di pomodoro coltivato. I 20 EST-SNPs selezionati (Tabella 2.4) sono stati utilizzati per effettuare la genotipizzazione SNP del DNA genomico delle sette varietà da industria in esame. La selezione dei loci SNP è stata effettuata consultando una banca dati costruita sulla base di sequenze EST, per cui l'amplificazione del DNA genomico ha a volte generato frammenti di peso molecolare più alto rispetto a quello atteso per la presenza di introni (Tabella 3.21).

Tabella 3.21: EST-SNPs selezionati per l'analisi CAPS, tipo di sostituzione e presenza/assenza dell'introne.

| SNP        | Sostituzione | Introne  |
|------------|--------------|----------|
| LEOH 2     |              | Assente  |
| LEOH 8.4   | sinonima     | Assente  |
| LEOH 11.1  |              | Assente  |
| LEOH 12.2  |              | Assente  |
| LEOH 16.2  | non sinonima | Assente  |
| LEOH 16.3  | non sinonima | Assente  |
| LEOH 20.1  | non sinonima | Assente  |
| LEOH 23.1  | 5' UTR       | Assente  |
| LEOH 25.1  | sinonima     | Presente |
| LEOH 28    | non sinonima | Assente  |
| LEOH 31.3  | sinonima     | Presente |
| LEOH 31.4  | non sinonima | Presente |
| LEOH 31.6  | sinonima     | Presente |
| LEOH 31.7  | sinonima     | Presente |
| LEOH 31.10 | non sinonima | Presente |
| LEOH 32.1  | sinonima     | Presente |
| LEOH 35.5  | non sinonima | Presente |
| LEOH 39    |              | Assente  |
| LEOH 63    |              | Assente  |
| CosOH 73   |              | Assente  |

Dei 20 loci EST-SNP selezionati, 12 hanno generato frammenti delle dimensioni riportate in letteratura calcolate sulla base dell'amplificazione del cDNA ed otto hanno mostrato la presenza di introni generando frammenti di dimensioni molecolari piu' elevate. La figura 3.24 mostra alcuni esempi di digestione, e assenza di, nelle sette varietà commerciali.

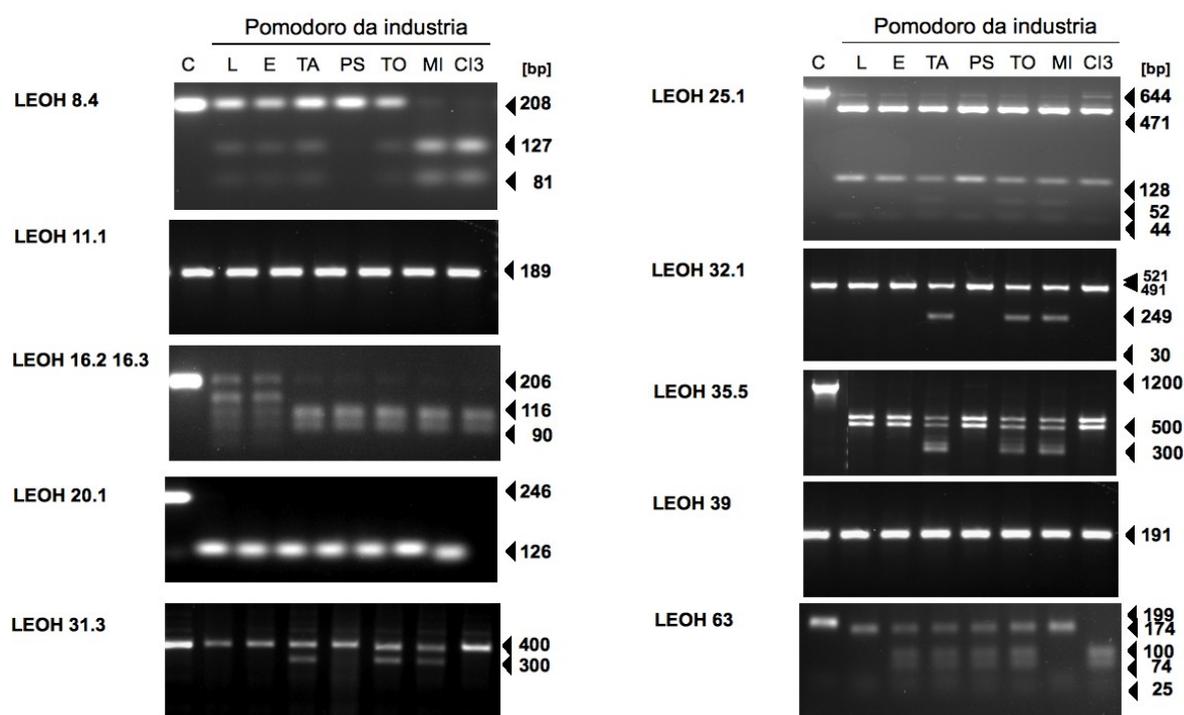


Figura 3.24: Esempi di analisi CAPS per dieci dei loci EST-SNP analizzati su sette varietà di pomodoro da industria. Sulla destra è riportata la dimensione (bp) dei frammenti prima e dopo la digestione. C= controllo, campione da PCR non digerito.

L'identità dello SNP ed il calcolo della mappa di restrizione sono stati verificati tramite sequenziamento dei prodotti di amplificazione del DNA genomico e successivo allineamento con le rispettive sequenze EST di riferimento. La mappa di restrizione ottenuta dall'analisi delle sequenze è riportata in tabella 3.22.

Tabella 3.22: Conversione di EST-SNP in CAPS. Nome dello SNP; dimensione osservata dei prodotti PCR in bp; frammenti di restrizione in bp; nome dell'enzima di restrizione utilizzato per l'analisi CAPS.

| SNP       | Dimensione frammenti PCR (bp) | Frammenti di taglio (bp)     | Enzima di digestione |
|-----------|-------------------------------|------------------------------|----------------------|
| LEOH2     | 201                           | 109, 92                      | <i>BsaWI</i>         |
| LEOH8.4   | 208                           | 127, 81                      | <i>HaeIII</i>        |
| LEOH11.1  | 189                           | 148,41                       | <i>HhaI</i>          |
| LEOH12.2  | 181                           | 158,23                       | <i>MseI</i>          |
| LEOH16.2  | 206                           | 116; 90 <sup>a</sup>         | <i>BsaWI</i>         |
| LEOH 16.3 | 206                           | 152,54 <sup>a</sup>          | <i>BsaWI</i>         |
| LEOH20.1  | 246                           | 126; 120                     | <i>HaeIII</i>        |
| LEOH23.1  | 191                           | 127,64                       | <i>MspI</i>          |
| LEOH25.1  | 644                           | 471; 77; 52; 44 <sup>b</sup> | <i>HincII</i>        |
| LEOH28    | 223                           | 114; 109                     | <i>TaqI</i>          |
| LEOH31.3  | 388                           | 307; 81                      | <i>MspI</i>          |
| LEOH31.4  | 388                           | 217; 48; 87; 37 <sup>c</sup> | <i>MseI</i>          |
| LEOH32.1  | 521                           | 249; 242; 30 <sup>d</sup>    | <i>HaeIII</i>        |
| LEOH35.5  | 1200                          | nd <sup>h</sup>              | <i>TaqI</i>          |
| LEOH39    | 191                           | 162; 29                      | <i>EcoRII</i>        |
| LEOH31.6  | 324                           | 227; 97                      | <i>TaqI</i>          |
| LEOH31.7  | 324                           | 218; 73; 33 <sup>e</sup>     | <i>MseI</i>          |
| LEOH31.10 | 324                           | 143; 152; 29 <sup>f</sup>    | <i>AluI</i>          |
| LEOH63    | 199                           | 100; 74; 25 <sup>g</sup>     | <i>TaqI</i>          |
| CosOH73   | 189                           | 100; 89                      | <i>AluI</i>          |

<sup>a</sup> la digestione simultanea di LEOH 16.2 e 16.3 genera 116; 54; 36

<sup>b</sup> l'assenza del sito polimorfico *HincII* genera 471; 128; 44

<sup>c</sup> l'assenza del sito polimorfico *MseI* genera 265; 87; 37

<sup>d</sup> l'assenza del sito polimorfico *HaeIII* genera 491; 30

<sup>e</sup> l'assenza del sito polimorfico *MseI* genera 251; 73

<sup>f</sup> l'assenza del sito polimorfico *AluI* genera 295; 29

<sup>g</sup> l'assenza del sito polimorfico *TaqI* genera 174; 25

<sup>h</sup> non disponibile: la bassa qualità della sequenza ottenuta è stata sufficiente per effettuare la validazione dello SNP ma non per il calcolo della mappa di restrizione

I dati hanno dimostrato che gli SNP sono stati utili per distinguere le sette varietà da industria (Tabella 3.23).

Tabella 3.23: Profili SNP da analisi CAPS e validazione mediante sequenziamento delle sette varietà da industria.

| SNP       | Leader | Ercole | Talent | PS1296 | Tomito | Minidor | Cirio3 |
|-----------|--------|--------|--------|--------|--------|---------|--------|
| LEOH2     | +/+    | +/+    | +/+    | +/+    | +/+    | +/+     | +/+    |
| LEOH8.4   | +/-    | +/-    | +/-    | -/-    | +/-    | +/+     | +/+    |
| LEOH11.1  | -/-    | -/-    | -/-    | -/-    | -/-    | -/-     | -/-    |
| LEOH12.2  | +/+    | +/+    | +/+    | +/+    | +/+    | +/+     | +/+    |
| LEOH16.2  | +/-    | +/-    | +/+    | +/+    | +/+    | +/+     | +/+    |
| LEOH 16.3 | +/-    | +/-    | +/+    | +/+    | +/+    | +/+     | +/+    |
| LEOH20.1  | +/+    | +/+    | +/+    | +/+    | +/+    | +/+     | +/+    |
| LEOH23.1  | +/-    | +/-    | -/-    | +/-    | +/-    | +/+     | +/+    |
| LEOH25.1  | -/-    | -/-    | +/-    | -/-    | +/-    | +/-     | -/-    |
| LEOH28    | -/-    | -/-    | -/-    | -/-    | -/-    | -/-     | -/-    |
| LEOH31.3  | -/-    | -/-    | +/-    | -/-    | +/-    | +/-     | -/-    |
| LEOH31.4  | -/-    | -/-    | +/-    | -/-    | +/-    | +/-     | -/-    |
| LEOH32.1  | -/-    | -/-    | +/-    | -/-    | +/-    | +/-     | -/-    |
| LEOH35.5  | -/-    | -/-    | +/-    | -/-    | +/-    | +/-     | -/-    |
| LEOH39    | -/-    | -/-    | -/-    | -/-    | -/-    | -/-     | -/-    |
| LEOH31.6  | -/-    | -/-    | +/-    | -/-    | +/-    | +/-     | -/-    |
| LEOH31.7  | +/+    | +/+    | +/-    | +/+    | +/-    | +/-     | +/+    |
| LEOH31.10 | -/-    | -/-    | +/-    | -/-    | +/-    | +/-     | -/-    |
| LEOH 63   | -/-    | +/-    | +/-    | +/-    | +/-    | -/-     | +/+    |
| CosOH 73  | +/+    | +/+    | +/+    | +/+    | +/+    | +/+     | +/+    |

Sito di restrizione polimorfico: presenza (+); assenza (-)

Nell'ambito di questo set di loci SNP analizzato sono stati individuati tre SNPs (loci LEOH8.4, LEOH31.3 ed LEOH 63) sufficienti a discriminare le sette varietà in esame. Con i dati del sequenziamento ottenuti, e' stato, quindi, possibile effettuare in collaborazione con l'Università degli Studi di Parma un *design* preliminare delle sonde a PNA (*Peptide Nucleic Acid Array*) da sintetizzare per la costruzione di un *array* contenente SNP per studi di tracciabilità nella filiera sementiera e di trasformazione.

#### 4. DISCUSSIONI

L'industria agro-alimentare italiana potrebbe trarre notevoli vantaggi dalla ampia disponibilit  e diversit  di specie orto-frutticole di pregio. Questi materiali genetici italiani, alle volte, non adeguatamente valorizzati, rappresentano una importante risorsa per incrementare la competitivita' del comparto agricolo italiano nei mercati internazionali. Inoltre, la crescente preferenza dei consumatori per alimenti salubri, tradizionali con legami con il territorio e sicuri, fornisce una ulteriore opportunit  all'agricoltura italiana per affermarsi anche sui mercati interni, tramite la valorizzazione dei prodotti tipici. Purtroppo, il fenomeno dell'agro-pirateria limita fortemente queste potenzialit . Uno studio del 2008 del Censis sulla tematica della contraffazione ha stimato che l'agricoltura italiana perde ogni anno oltre 3 miliardi di euro a causa dell'agro-pirateria sui mercati internazionali e che il mercato del falso dei prodotti alimentari in Italia, ha "fatturato" nel 2008 oltre 1,1 miliardi di euro.

L'obiettivo specifico di questo progetto e' stato lo sviluppo di innovazioni tecnologiche idonee per difendere le filiere agroalimentari del pomodoro. Infatti, negli ultimi anni l'aumento della richiesta di prodotti di alta qualit  ha canalizzato l'interesse scientifico verso l'ottenimento di metodiche sempre pi  sicure e rapide per garantire e certificare l'origine delle materie prime in ingresso ed in uscita delle filiere agro-alimentari. Spesso l'etichettatura non fornisce garanzie sufficienti in merito all'identit  genetica del prodotto, rendendo necessario lo sviluppo di tecnologie avanzate per identificare e/o autenticare i componenti degli alimenti trasformati, proteggendo cos  sia i consumatori che i produttori da sostituzioni illecite. La tracciabilit , divenuta dal gennaio 2005 obbligatoria per tutte le aziende agro-alimentari, obbliga le stesse al rispetto della normativa che necessita l'applicazione di un continuo controllo di tutte le fasi della catena produttiva, dalla raccolta del materiale in campo, alla valutazione della materia prima in ingresso, all'analisi del prodotto finale.

Da qui nasce l'esigenza di avvalersi di strumenti di valutazione sicuri, basati anche su un approccio inter-disciplinare tra vari settori, come, ad esempio, le scienze agro-alimentari e le biotecnologie vegetali. Le tecniche dei marcatori molecolari del DNA possono offrire una maggiore affidabilit  per l'identificazione dell'identit  genetica dei materiali in ingresso nella filiera produttiva. Inoltre, il DNA   pi  resistente alla degradazione durante il processamento dei cibi, come la cottura o la sterilizzazione, rispetto ad altri marcatori (Martinez et al., 2003), rendendo questa molecola particolarmente amena alle analisi dei prodotti in uscita dalla filiera.

Studi pionieri che riportano l'utilizzo di marcatori RFLP (Miller e Tanksley, 1990) e SNP (Nesbitt e Tanksley, 2002) hanno rilevato la presenza di una ridotta diversit  genetica nel pomodoro coltivato. Questo e' probabilmente una conseguenza della domesticazione e della sua stessa natura di autoimpollinazione. In questa specie, a causa della ridotta diversit  genetica (Miller e Tanksley, 1990), i marcatori molecolari del DNA, opportunamente selezionati e validati, possono essere particolarmente discriminativi ed informativi per l'identificazione di cultivar ed ecotipi, nonostante gli alti livelli di monomorfismo ed i bassi valori di contenuto polimorfico informativo (PIC) riportati in letteratura. Attualmente sono due le tipologie di marcatori molecolari del DNA utili per la tracciabilit  nella filiera agro-alimentare: gli SSR e gli SNP, perch  riescono a fornire informazioni anche con un DNA frammentato o danneggiato e ben si prestano a sistemi di automazione.

#### 4.1 Valorizzazione e protezione di ecotipi di pomodoro campani

La competitività, la sostenibilità e la tipicità dell'agricoltura italiana dipende anche dal continuo accesso a nuovo germoplasma, utile al *breeder* per selezionare o migliorare nuove varietà. Nessuna nazione al mondo è autosufficiente da questo punto di vista, ma le varietà tradizionali presenti in ogni regione rappresentano una risorsa troppo spesso inesplorata. Tale materiale possiede il vantaggio di essere già adattato ad un agro-sistema ed ameno a "*low-input practices*", e solitamente presenta caratteristiche qualitative del prodotto edule superiori e facilmente riconoscibili dai consumatori. In questa prospettiva, il pomodoro rappresenta un caso esemplare. La specie è caratterizzata da una ristretta diversità genetica ed il panorama varietale è dominato da ibridi che hanno progressivamente scalzato gli ecotipi locali.

La tracciabilità genetica nella filiera vivaistica è importante per la valorizzazione e protezione di ecotipi di pomodoro che hanno il potenziale per affermarsi come prodotti di qualità e che potenzialmente aspirano ad avere marchi di tutela (Denominazione d'Origine Protetta- DOP, Indicazione Geografica Protetta - IGP), come già accaduto per il 'San Marzano'. La protezione degli ecotipi è strettamente dipendente dagli strumenti che permettono la distinzione di genotipi contaminati e/o fenotipicamente simili che sono presenti anche nell'area di origine di coltivazione (Rao et al., 2006).

Il presente studio documenta l'efficienza di tre diverse tipologie di marcatori molecolari per la discriminazione di piante di pomodoro e la loro relazione con le caratteristiche morfologiche. L'analisi di omogeneità condotta sugli ecotipi campani mediante il minisatellite (GATA)<sub>4</sub> ha indicato sia la presenza di ecotipi che sono geneticamente uniformi che la presenza di ecotipi che sono una mistura di due o più genotipi. Come già riportato in letteratura (Rao et al., 2006), il minisatellite ha mostrato possedere un elevato potere discriminante ed è risultato essere altamente risolutivo ed informativo, mostrando la presenza dell'87,10% di bande polimorfiche nel materiale analizzato.

Sulla base di questi risultati, è stato possibile selezionare gli ecotipi uniformi per le successive analisi mediante marcatori SSR e SNP. I dati indicano che il *fingerprint* del DNA basato sull'analisi mediante marcatori molecolari SSR permette la discriminazione non solo di ecotipi differenti (He et al., 2003; Smulders et al., 1997), ma anche di tipi morfologicamente simili. Il numero medio di alleli osservato è stato di 2,8 ed il valore massimo di PIC riscontrato è stato di 0,76 con un valore medio di 0,46. Questi dati indicano che il set di loci selezionato è risultato essere particolarmente informativo e capace di discriminare gli ecotipi in esame. I loci SSR maggiormente informativi hanno mostrato differenti livelli di discriminazione: alcuni hanno permesso l'identificazione di ecotipi caratterizzati dalla presenza di specifici alleli, altri hanno permesso la distinzione tra i diversi ecotipi, altri hanno individuato la presenza di genotipi contaminanti all'interno dello stesso ecotipo. Un set minimo di tre loci SSR è stato individuato ed è risultato sufficiente per discriminare tutti gli ecotipi in esame, eccetto che i tipi 'San Marzano Cilindrico I' e 'San Marzano Murano' che condividono anche lo stesso profilo generato dall'analisi mediante il minisatellite, caratteristico ed identificativo del 'San Marzano' (Caramante et al., 2009).

Bassa è risultata, invece, la percentuale di loci SNP polimorfici. Solo dodici SNP su 71 hanno mostrato polimorfismo ed il valore di PIC medio, considerando i soli loci polimorfici, è risultato essere di 0,315. Cinque loci SNP individuati sono risultati

sufficienti a rappresentare il livello di discriminazione del totale degli SNP analizzati e capaci di discriminare gli stessi ecotipi discriminabili mediante analisi SSR e (GATA)<sub>4</sub>, eccetto che due ecotipi (STLR e TLI) che mostrano elevati livelli di similarità anche per gli altri marcatori utilizzati per l'analisi.

Di contro, entrambi SSR e minisatellite, data anche la loro natura multiallelica, sono stati capaci di identificare e discriminare tutti gli ecotipi di pomodoro in esame. I marcatori SSR posseggono, però, vari vantaggi sulla analisi basata sul minisatellite. L'analisi mediante SSR richiede tempi inferiori, e' estendibile a diverse tipologie di materiale data la limitazione del minisatellite di poter essere utilizzato solo per le analisi del materiale fresco. Nonostante il potere discriminante degli SSR in pomodoro sia più basso che in altre specie autogame (He et al., 2003; Smulders et al., 1997), nel presente studio e' stato dimostrato che alcuni marcatori SSR possono essere efficientemente utilizzati per la discriminazione intra-varietale, così come per la discriminazione di diversi tipi di 'San Marzano' sia a livello di genotipi molto simili che di genotipi contaminanti.

#### 4.2 Polimorfismo nei diversi marcatori

Il polimorfismo riscontrato nei diversi marcatori utilizzati per l'analisi delle due popolazioni in esame, ecotipi e varietà commerciali, e' stato valutato ed alcune delle loro principali caratteristiche sono state confrontate. Le caratteristiche dei marcatori SSR non sono risultate essere influenzate dal tipo di popolazione esaminata. Infatti, la percentuale di loci polimorfici e' risultata essere del 92,86% sia per gli ecotipi che per le varietà, mostrando la presenza di un solo locus polimorfico su 14 analizzati, ed il contenuto polimorfico informativo medio assume valori di 0,447 e 0,441, rispettivamente. I dati hanno, quindi, indicato che il potere discriminante dei microsattelliti non e' influenzato dalla diversità genetica della popolazione in esame. Questo dato può essere la conseguenza della selezione a monte di loci SSR particolarmente informativi.

I risultati ottenuti per i singoli SSR rispecchiano nella maggioranza dei casi le caratteristiche attese. In alcuni casi, però, il numero degli alleli ed il contenuto polimorfico informativo osservati sono risultati essere inferiori o superiori a quelli riportati in letteratura. Esempi sono il locus LELE25 che, in entrambe le popolazioni in esame, ha generato la presenza di un numero di alleli pari a 2 ed un PIC pari a 0,18 rispetto ai 4 alleli ed al PIC di 0,36 attesi ed il locus LEta015 per il quale, negli ecotipi campani, sono stati riscontrati un numero di alleli pari a 6 ed un PIC di 0,76 rispetto ai 2 alleli ed al PIC di 0,49 riportati in letteratura.

I marcatori SNP sono principalmente biallelici per cui il valore massimo di PIC atteso per un dato locus e' di 0,5. Nonostante i bassi livelli di polimorfismo SNP riscontrati nell'ambito degli ecotipi, in alcuni casi si e' quasi raggiunto il valore soglia rilevando un valore di PIC di 0,48 (CT10790). Il set di marcatori SNP non ha, però, mostrato possedere un livello di discriminazione tale da generare profili allelici unici ed identificativi di tutto il materiale in esame, sia nell'ambito degli ecotipi che delle varietà commerciali. Inoltre, l'efficacia dei marcatori SNP nel discriminare il materiale vegetale si riduce sensibilmente passando dalle varietà, altamente eterozigoti, agli ecotipi.

Risultato paragonabile e' stato riscontrato nell'analisi dello stesso materiale mediante marcatori in/del. In riferimento alle varietà, il 100% dei loci in/del e' risultato essere polimorfico, ma non capace di effettuarne la completa discriminazione. Per gli ecotipi, invece, un solo locus su sei e' risultato essere polimorfico. I risultati ottenuti hanno

rilevato che i marcatori SSR hanno i più alti livelli di polimorfismo e il più alto potere discriminante nell'ambito del germoplasma analizzato.

Il PIC medio per gli SSR (0,477), considerando per il calcolo sia i dati derivanti dagli ecotipi che quelli derivanti dalle varietà, risulta essere superiore al PIC medio (0,405) ottenuto dalla somma dei PIC medi di tutti i marcatori utilizzati (SSR, SNP, in/del) considerando per il calcolo di questo valore i soli loci polimorfici. La scelta e l'utilizzo delle diverse tipologie di marcatori molecolari per la stima della diversità genetica dipende da numerosi fattori inclusi i costi di genotipizzazione di larghe popolazioni ed i tempi di analisi. Dovrebbe essere considerato che lo sviluppo di SSR è costoso in termini di denaro e tempo e che i marcatori SSR utilizzati sono stati specificamente selezionati per la caratterizzazione di cultivar di pomodoro essendo specie-specifici. Il polimorfismo rilevabile mediante i marcatori SSR è sicuramente più elevato rispetto a quello degli SNP, dato che gli SSR sono marcatori multiallelici. Questa caratteristica associata con la loro natura co-dominante li rende i marcatori di elezione nell'ambito della genetica delle piante e gli studi di miglioramento genetico. L'abbondanza degli SNP potrebbe, però, altamente compensare la loro deficienza, in presenza di piattaforme genomiche che offrono un compromesso tra alta processività e costi ridotti. I dati di genotipizzazione ottenuti da analisi SNP possono essere valutati in un formato binario di presenza/assenza dell'allele tra i differenti genotipi e questo risulta essere molto conveniente per la costruzione di una banca dati genomica. Di contro, gli SSR necessitano un'accurata definizione delle dimensioni alleliche, che può variare tra differenti laboratori in funzione del sistema di analisi utilizzato e dell'esperienza dell'utente.

Considerando i vantaggi e gli svantaggi, gli SSR e gli SNP dovrebbero essere i marcatori di elezione per gli studi delle principali specie vegetali dal momento che sequenze EST e primer per tali tipologie di marcatori sono disponibili in quantità sufficienti ed il sequenziamento dei genomi è in corso.

#### 4.3 Il pomodoro coltivato in relazione alle specie selvatiche

Come atteso, specie selvatiche generalmente mostrano livelli di polimorfismo più elevati rispetto alle varietà coltivate di *S. lycopersicum*. Questo potrebbe anche essere dovuto alla presenza di alleli presenti nei genotipi selvatici, che sono stati persi durante la domesticazione. L'analisi mediante SSR effettuata sul pomodoro coltivato (ecotipi e varietà) e sulle cinque specie selvatiche in esame, ha rilevato la presenza di un numero medio di alleli osservati nelle specie selvatiche di 4,85, superiore a quello riscontrato negli ecotipi e nelle varietà (3,28 e 2,92, rispettivamente). Il valore di PIC medio risulta essere elevato nell'ambito di ognuna delle tre popolazioni in esame ed il valore medio totale per tutte e tre le popolazioni è di 0,52. Un set minimo di 5 loci SSR individuato è stato sufficiente per effettuare la discriminazione di tutto il materiale in esame (ecotipi, varietà, specie selvatiche) ed il PIC medio calcolato considerando l'insieme delle tre popolazioni è di 0,55, valore pari a quello riscontrato considerando tutti i loci in analisi.

Sulla base dei dati dei microsatelliti è stato effettuato il confronto tra 20 ecotipi campani e 20 varietà commerciali (*inbred* ed ibridi) tramite analisi gerarchiche e multivariata (PCA). Le varietà risultano raggruppare insieme alle specie selvatiche esaminate, ad indicare che esistono fattori genetici che accomunano questi materiali. Le specie selvatiche, infatti, sono la fonte più ricca di geni di resistenza a stress biotici del pomodoro e risorsa di alleli unici e favorevoli nei programmi di miglioramento genetico per la produzione di ibridi da industria.

L'analisi multivariata ha suggerito che gli ecotipi campani risultano essere una popolazione geneticamente distinta dalle comuni varietà coltivate di pomodoro. Tali ecotipi sono caratterizzati da una buona e probabilmente specifica ricchezza allelica data la presenza di 11 alleli ecotipo specifici. e devono essere considerati, quindi, un materiale da proteggere e migliorare che si differenzia significativamente dalle varietà commerciali, ma al tempo stesso anche una probabile risorsa di caratteri di qualità ed adattabilità a specifici ambienti.

Simili conclusioni sono state ottenute dall'analisi mediante marcatori SNP. L'elevato numero di marcatori SNP analizzato ha permesso l'analisi di una eventuale struttura di popolazione, che mostra chiaramente che gli ecotipi campani di pomodoro sono una popolazione geneticamente diversa dal restante materiale in analisi che comprende numerose tipologie di frutto a differente destinazione d'uso e di differente provenienza. I tipi 'Sorrento', invece, appartenenti alla collezione di germoplasma campano raggruppano separatamente insieme ad altri genotipi dal frutto di forma globosa provenienti da differenti aree geografiche.

#### **4.4 Tracciabilità genetica dei prodotti di filiera**

L'estrazione di DNA da matrici alimentari processate è una procedura complessa data la scarsa quantità di DNA e l'elevato grado di degradazione dovuti ai processi produttivi. Tra le tecniche di conservazione dei cibi, l'inscatolamento richiede due principali condizioni: elevate temperature e pH acidi, causa della frammentazione di molecole di DNA ad alto peso molecolare che ne influenza la capacità di essere amplificato mediante reazioni PCR.

In questo studio è stato analizzato DNA da tre differenti tipologie di prodotto processato putativo monovarietale (polpa, pelato, pomodorino inscatolato) allo scopo di avere una rappresentazione delle diverse tipologie di inscatolato di pomodoro. Al fine di ottenere DNA quantitativamente e qualitativamente migliore, sono stati saggiati diversi protocolli per l'estrazione di DNA da matrici complesse. Per tali estrazioni sono state utilizzate quantità variabili di materiale processato (tal quale o liofilizzato) e differenti protocolli di estrazione, il metodo tradizionale basato sul CTAB e il protocollo previsto dal "Genelute Plant Genomic DNA Kit" (SIGMA). Non è stato possibile determinare la quantità di DNA estratto a causa delle basse concentrazioni di DNA presente nell'inscatolato e l'elevato grado di degradazione, per cui è stato direttamente utilizzato come stampo per le amplificazioni mediante loci SSR.

Il set di loci SSR è stato selezionato sulla base di alcune caratteristiche quali la capacità di generare frammenti di amplificazione di piccole dimensioni, in previsione di utilizzare lo stesso set per la caratterizzazione del materiale fresco e del rispettivo prodotto processato. I 14 loci SSR sono stati utilizzati per la genotipizzazione del DNA estratto da foglia di tutte le varietà ibride in esame e sono risultati essere particolarmente informativi e capaci di effettuarne la completa discriminazione. La combinazione di pochi microsatelliti ha mantenuto lo stesso livello di discriminazione, generando un profilo allelico unico per ogni varietà analizzata.

Il set di loci SSR opportunamente selezionato genera frammenti di amplificazione di peso molecolare compreso tra 94 e 233 bp, per cui si è proceduto con il loro utilizzo su DNA estratto da bacca e processato di sette delle varietà selezionate per gli studi di tracciabilità. I migliori risultati in termini di specifici frammenti amplificati, quando DNA da prodotto processato è utilizzato come stampo, sono stati ottenuti quando per l'estrazione viene utilizzato materiale liofilizzato e seguito il protocollo previsto dal kit commerciale utilizzato. È stato, però, necessario individuare le condizioni ottimali

di amplificazione PCR per ottenere gli specifici frammenti attesi. Solo l'utilizzo di piccole quantita' di *templato* ha generato la presenza di ampliconi. Questo potrebbe essere spiegato dalla riduzione dell'inibizione dell'attivita' polimerasica dovuta alla bassa presenza di componenti inibenti che sono normalmente presenti nei prodotti industriali (Breton et al., 2004; Tengel et al., 2002; De la Torre et al., 2002).

Nove loci SSR su quattordici analizzati sono stati capaci di generare gli specifici alleli per tutte e tre le tipologie di processato in esame, in un range di peso molecolare compreso tra 94 e 164 bp. Gli altri cinque hanno fallito a causa, probabilmente, delle piu' elevate dimensioni degli alleli attese. Questi risultati indicano che il successo dell'amplificazione dipende principalmente dalla capacita' dei *kit* commerciali di liberare il DNA dagli inibitori della PCR presenti nei campioni e dalle dimensioni degli amplificati attesi.

I prodotti di amplificazione fluorescenti sono stati separati mediante elettroforesi capillare per la discriminazione allelica. Dei nove loci amplificati, sette sono risultati essere polimorfici e due monomorfici nell'ambito delle sette varieta' analizzate. I profili allelici ottenuti da foglia, bacca e processato sono stati confrontati. E' stata riscontrata corrispondenza tra i profili allelici, nei tre segmenti della filiera in esame, per le varieta' 'Tomito', 'Minidor' e 'Cirio3'. Questi risultati risultano essere incoraggianti e dimostrano che e' possibile tracciare l'intera catena di produzione del pomodoro. Tuttavia, non e' stata riscontrata corrispondenza negli stadi della filiera analizzati per le altre varieta'. In particolare, e' stata pero' riscontrata una corrispondenza, per tutti i nove loci SSR analizzati, tra i profili allelici ottenuti da foglia e bacca della varieta' 'Ercole' con i profili allelici ottenuti da processato della varieta' 'Talent'. Questo potrebbe essere spiegato come uno scambio tra varieta' che condividono la stessa forma del frutto e destinazione d'uso in qualche punto della catena di produzione. Per le altre varieta' di cui non e' stata riscontrata corrispondenza tra i profili allelici dei tre segmenti della filiera in esame, e' ipotizzabile la presenza di materiale non monovarietale all'interno dell'inscatolato.

## 5. CONCLUSIONI

L'applicazione delle moderne tecnologie genetico-molecolari rappresenta un importante contributo alla tutela del consumatore che, in molte parti del mondo, dispone di un reddito crescente ed è alla ricerca di cibi saporiti, tradizionali e genuini, e alla tutela del produttore attraverso la differenziazione del prodotto che può conseguire maggiore potere di mercato. Il *fingerprinting* del DNA basato su marcatori molecolari rappresenta un valido strumento per l'identificazione di cultivar ed accessioni, per la quantificazione della diversità genetica di collezioni di varietà ed ecotipi e per la tracciabilità genetica delle materie prime nelle filiere agro-alimentari fornendo un eccellente ausilio per la protezione di prodotti tipici italiani. Inoltre, gli agricoltori sono costantemente alla ricerca di opportunità originali ed inesplorate per creare nuovi sbocchi di mercato. Certamente tra queste, un ruolo chiave è occupato dalla tutela delle denominazioni dei prodotti alimentari mediante marchi di qualità, opportunamente difesi nell'ambito di consorzi ed associazioni di produttori. Le innovazioni di genetica e genomica vegetale forniscono, in tal senso, un contributo molto importante. E' stata generata una enorme risorsa di conoscenze che incoraggiano il *breeder* ad adottare strategie di miglioramento piu' razionali e rifinite. Considerevoli progressi sono stati effettuati nella realizzazione di infrastrutture per l'applicazione degli approcci di genomica. Nuove conoscenze e nuovi strumenti stanno cambiando lo scenario delle strategie utilizzate nella ricerca. Un importante fattore e' lo sviluppo di sistemi *high-throughput*, basati sull'amplificazione PCR ed in grado di identificare polimorfismi tra genotipi simili. L'aumento delle sequenze di DNA, in particolare EST, disponibili sta producendo utili informazioni per lo sviluppo di marcatori molecolari del DNA altamente informativi per studi di mappaggio, *fingerprinting* e MAS per il pomodoro coltivato. Di pari passo con i progressi nel sequenziamento del genoma del pomodoro e' atteso che piu' marcatori, in particolare SNP, saranno disponibili e che i progressi nelle tecnologie che utilizzano i marcatori molecolari avranno costi sempre piu' ridotti per il loro sviluppo ed uso. E' anche atteso che una maggiore enfasi sara' data allo sviluppo di marcatori funzionali. Numerose sono le risorse genomiche e genetiche gia' disponibili per specie vegetali di rilevante valore economico, biologico ed agronomico come il pomodoro. Esiste una continua necessita' di integrare discipline come la genomica strutturale, la trascrittomica, la proteomica e la metabolomica con la fisiologia vegetale ed il miglioramento genetico delle piante. Nell'era della post-genomica, approcci ad alta produttivita' in combinazione con l'automazione, stanno incrementando la quantita' di dati di sequenza a dominio pubblico ed il rafforzamento delle tecniche di bioinformatica contribuirà alla ricerca genomica per il miglioramento delle specie vegetali.

In questo studio, e' stato sviluppato un protocollo di estrazione di DNA da matrice alimentare complessa di pomodoro di tre tipologie ed e' stato dimostrato che le concentrazioni di DNA presenti sono sufficienti per ottenere amplificazioni consistenti se il DNA e' ben purificato dagli altri componenti che potrebbero inibire l'attivita' della polimerasi. E' stato dimostrato che nonostante il DNA presente nei prodotti processati sia altamente degradato, frammenti di dimensioni superiori alle 150 bp possono essere amplificati. I risultati hanno, quindi, dimostrato la fattibilita' dell'impiego dei loci SSR per gli studi di tracciabilita' della filiera agro-alimentare del pomodoro anche se le dimensioni dei prodotti PCR possono essere un fattore limitante per l'analisi di alcuni loci ed e' stato dimostrato che in alcuni casi e' possibile

risalire, tramite analisi SSR del DNA estratto da prodotto trasformato monovarietale, alle varietà utilizzate nella filiera produttiva.

Questo studio ha fornito una serie di solide informazioni per poter scegliere i marcatori molecolari più appropriati per gli studi di tracciabilità di filiera del pomodoro. Tali marcatori sono concordi nell'evidenziare la specificità degli ecotipi campani, giustificando azioni che ne garantiscono la loro protezione. Infine, il presente lavoro ha accertato la potenzialità delle analisi del DNA in substrati degradati ed ha messo in evidenza alcuni dei problemi di identificazione del materiale nei diversi segmenti della filiera agroalimentare.

Due aforismi celebri - *"Fa che il cibo sia la tua medicina e la medicina sia il tuo cibo"* di Ippocrate e *"L'uomo è ciò che mangia"* di L. Feuerbach - racchiudono l'essenza degli indirizzi più attuali ed innovativi nella ricerca scientifica legata alla nutrizione. Tutti gli alimenti che costituiscono la nostra dieta hanno origine da piante o animali. Lo studio e le applicazioni di nuove discipline scientifiche destinate a migliorare la qualità dei cibi ed il loro apporto salutare nei confronti dell'alimentazione umana comincia dalle materie prime, nei campi e negli allevamenti, e prosegue con le operazioni di trasformazione, conservazione e consumo. Un alimento pronto per essere portato alla bocca ha alle spalle, nella società occidentale, una storia che coinvolge molte professionalità, indirizzate a migliorare la produzione e la qualità dei singoli prodotti.

Dalla ricerca scientifica verso l'agroindustria: questa è la sfida che scienziati ed industriali possono affrontare congiuntamente in un futuro non lontano.

**BIBLIOGRAFIA**

Andreakis N., Giordano I., Pentangelo A., Fogliano V., Graziani G., Monti L. M., Rao R. (2004). DNA fingerprinting and quality traits of Corbarino cherry-like tomato landraces. *Journal of Agricultural and Food Chemistry*, vol. 52, pp. 3366-3371.

Ardley J., Hoptroff C. G. M. (1996). Protecting plant 'invention': the role of plant-variety rights and patents. *Trends in Biotechnology*, vol. 14, pp. 67-69.

Arens P., Odinet P., van Heusden A. W., Lindhout P., Vosman B. (1995). GATA and GACA repeats are not evenly distributed throughout the tomato genome. *Genome*, vol. 38, pp. 84-90.

Areshchenkova T., Ganai M. W. (1999). Long tomato microsatellites are predominantly associated with centromeric regions. *Genome*, vol. 42, pp. 536-544.

Armour A. L., Povey S., Jeremjah S., Jeffreys, A. (1990). Systematic cloning of human minisatellites from ordered array charomid libraries. *Genomics*, vol. 8, pp. 501-512.

Aronson J. D. (2005). DNA fingerprinting on trial: the dramatic early history of a new forensic technique. *Endeavour*, vol. 29, pp. 126-131.

Arumuganathan K., Earle E. D. (1991). Nuclear DNA content of some important plant species. *Plant Molecular Biology Reporter*, vol. 9, no. 3, pp. 208-218.

Ashrafi H., Kinkade M., Foolad M. R. (2009). A new genetic linkage map of tomato based on a *Solanum lycopersicum* x *S. pimpinellifolium* RIL population displaying location of candidate pathogen response genes. *Genome*, vol. 52, pp. 935-956.

Bai Y., Feng X., van der Hulst R., Lindhout P. (2004). A set of simple PCR markers converted from sequence specific RFLP markers on tomato chromosomes 9 to 12. *Molecular Breeding*, vol. 13, pp. 281-287.

Bai Y., Huang C. C., van der Hulst R., Meijer-Dekens F., Bonnema G., Lindhout P. (2003). QTLs for tomato powdery mildew resistance (*Oidium lycopersici*) in *Lycopersicon parviflorum* G1. 1601 co-localize with two qualitative powdery mildew resistance genes. *Molecular Plant-Microbe Interactions*, vol. 16, pp. 169-176.

Bailey L. H., Tracy W. W., Kyle E. J., Watts R. L. (1960). Tomato. In: Bailey LH (eds) The standard cyclopedia of horticulture. The Macmillan Company, New York, pp. 3353-3359.

Baldo A. M., Wan Y., Lamboy W. F., Simon C. J., Labate J. A., Sheffer S. M. (2007). SNP validation and genetic diversity in cultivated tomatoes and grapes. *International Plant and Animal Genomes XV Conference*, p. 171, San Diego (CA) USA.

Bonnema G., Hontelez J., Verkerk R. (1996). An improved method of partially digesting plant megabase DNA suitable for YAC cloning: application to the

construction of a 5.5 genome equivalent YAC library of tomato. *The Plant Journal*, vol. 9, no. 1, pp. 125–133.

Bredemeijer G., Cooke R., Ganai M., Peeters R., Isaac P., Noordijk Y., Rendell S., Jackson J., Roder M. S., Wendehake K., Dijcks M., Amelaine M., Wickaert V., Bertrand L., Vosman B. (2002). Construction and testing of a microsatellite database containing more than 500 tomato varieties. *Theoretical and Applied Genetics*, vol. 105, pp. 1019-1026.

Breton C., Claux D., Metton I., Skorski G., Berville A. (2004). Comparative study of methods for DNA preparation from olive oil samples to identify cultivar SSR alleles in commercial oil samples: possible forensic applications. *Journal of Agricultural and Food Chemistry*, vol. 52, pp. 531-537.

Breuning G., Lyons J. M. (2000). The case of the FLAVR SAVR tomato. *California Agriculture*, vol. 54, pp. 6–7.

Broun P., Tanksley S. D. (1996). Characterization and genetic mapping of simple repeat sequences in the tomato genome. *Molecular Genetics and Genomics*, vol. 250, pp. 39-49.

Budiman M. A., Mao L., Wood T. C., Wing R. A. (2000). A deep-coverage tomato BAC library and prospects toward development of an STC framework for genome sequencing. *Genome Research*, vol. 10, no. 1, pp. 129–136.

Caramante M., Rao R., Monti L. M., Corrado G. (2009). Discrimination of 'San Marzano' accessions: A comparison of minisatellite, CAPS and SSR markers in relation to morphological traits. *Scientia Horticulturae*, vol. 120, pp. 560-564.

Chang S.B., Yang T. J., Datema E., van Vugt J., Vosman B., Kuipers A., Meznikova M., Szinay D., Klein Lankhorst R., Jacobsen E., de Jong H. (2008). FISH mapping and molecular organization of the major repetitive sequences of tomato. *Chromosome Research*, vol. 16, pp. 919–933.

Chetelat R. T. (2002). Revised list of monogenic stocks. *Report of the Tomato Genetics Cooperative*, vol. 52, pp. 41–62.

Ching A., Caldwell K. S., Jung M., Dolan M., Smith O. S., Tingey S., Morgante M., Rafalski A. J. (2002). SNP frequency, haplotype structure and linkage disequilibrium in elite maize inbred lines. *BMC Genetics*, vol. 3, no. 1, p. 19.

Chunwongse J., Chunwongse C., Black L., Hanson P. (2002). Molecular mapping on the *Ph-3* gene for the late blight resistance. *The Journal of Horticultural Science and Biotechnology*, vol. 77, pp. 281-286.

Dayrat B. (2005). Towards integrative taxonomy. *Biological Journal of the Linnean Society*, vol. 85, pp. 407-415.

Decreto Legislativo 10 febbraio 2005, n. 30 (GU n. 052 Suppl.Ord. del 04/03/2005). Codice della proprietà industriale, a norma dell'articolo 15 della legge 12 dicembre 2002, n. 273.

De la Torre F., Bautista R., Canovas F. M., Claros G. (2004). Isolation of DNA from olive oil and oil sediments: application in oil fingerprinting. *Food, Agriculture and Environment*, vol. 2, pp. 84-89.

Delmotte F., Leterme N., Simon J. C. (2001). Microsatellite allele sizing: difference between automated capillary electrophoresis and manual technique. *BioTechniques*, vol. 31, pp. 810–818.

De Riek J., Calsyn E., Everaert I., Van Bockstaele E., De Loose M. (2001). AFLP based alternatives for the assessment of Distinctness, Uniformity and Stability of sugar beet varieties. *Theoretical and Applied Genetics*, vol. 103, pp. 1254—1265.

Direttiva 2000/13/CE del Parlamento Europeo e del Consiglio del 20 marzo 2000 relativa al ravvicinamento delle legislazioni degli Stati membri concernenti l'etichettatura e la presentazione dei prodotti alimentari, nonché la relativa pubblicità.

Esquinas-Alcazar J., Nuez F. (1995). Situacion taxonomica, domesticacion y diffusion del tomate. In: Nuez F (eds) *El cultivo del tomate*. Ed Mundi-Prensa, Madrid, Barcelona, Mexico, pp. 13–42.

Foolad M. R., Jones R. A., Rodriguez R. L. (1993). RAPD markers for constructing intraspecific tomato genetic maps. *Plant Cell Reports*, vol. 12, no. 5, pp. 293–297.

Francis D., Yang W., van der Knaap E., Hogenhout S., Deynze A. V., Darrigues A. (2006). DNA-microarray detection of single feature polymorphisms as a discovery tool for marker assisted selection within elite tomato populations. W299. *Plant and Animal Genome XIV*, San Diego (CA) USA.

Fulton T. M., van der Hoeven R., Eannetta N. T., Tanksley S. D. (2002). Identification, analysis, and utilization of conserved ortholog set markers for comparative genomics in higher plants. *The Plant Cell*, vol. 14, no. 7, pp. 1457–1467.

Ganal M. W., Durstewitz G., Kulosa D., Luerssen H., Polley A., Wolf M. (2007). Development of EST-derived SNP markers for plant breeding. *International Plant and Animal Genomes XV Conference*, p. 172, USDA, San Diego (CA) USA.

García-Martínez S., Andreani L., García-Gusano M., Geuna F., Ruiz J. J. (2006). Evaluation of amplified fragment length polymorphism and simple sequence repeats for tomato germplasm fingerprinting: utility for grouping closely related traditional cultivars. *Genome*, vol. 49, pp. 648–656.

Georgiev H. (1991). Heterosis in tomato breeding. Genetic Improvement of Tomato. Kalloo G. Ed. vol. 14 of *Monographs on Theoretical and Applied Genetics*, pp. 83–98, Springer, Berlin, Germany.

Grandillo S., Tanksley S. D. (1996). Genetic analysis of RFLPs, GATA microsatellites and RAPDs in a cross between *L. esculentum* and *L. pimpinellifolium*. *Theoretical and Applied Genetics*, vol. 92, pp. 957-965.

Grube R. C., Radwanski E. R., Jahn M. (2000). Comparative genetics of disease resistance within the *Solanaceae*. *Genetics*, vol. 155, pp. 873-887.

Gupta P. K., Balyan H. S., Sharma P. C., Ramesh B. (1996). Microsatellites in plants: a new class of molecular markers. *Current Science*, vol 70, pp. 45-54.

Gupta P. K., Varshney R. K. (2000) The development and use of microsatellite markers for genetic analysis and plant breeding with special emphasis on bread wheat. *Euphytica*, vol. 113, pp. 163-185

Hamilton C. M., Frary A., Xu Y., Tanksley S. D., Zhang H. B. (1999). Construction of tomato genomic DNA libraries in a binary-BAC (BIBAC) vector. *Plant Journal*, vol. 18, no. 2, pp. 223–229.

Hancock J. M. (1999). Microsatellites and other simple sequences: genomic context and mutational mechanisms. In: Goldstein D., Schlotterer C., editors. *Microsatellites - Evolution and Applications*. New York: Oxford University Press. pp. 1-9.

Hartl D. L., Clark A. G. (1997). Principles of population genetics. 3rd edition. Sunderland (MA): Sinauer Associates.

He C., Poysa V., Yu K. (2003). Development and characterization of simple sequence repeat (SSR) markers and their use in determining relationships among *Lycopersicon esculentum* cultivars. *Theoretical and Applied Genetics*, vol. 106, pp. 363–373.

Hebert P. D., Cywinska A., Ball S. L., deWaard J. R. (2003a). Biological identifications through DNA barcodes. *Proceedings Biological Sciences*, vol. 270, pp. 313-321.

Hebert P. D., Ratnasingham S., deWaard J. R. (2003b). Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings Biological Sciences*, vol.,270, pp. 96-99.

IBPGR (1981). Jose' T., Esquinas-Alcazar. Genetic Resources of Tomatoes and Wild Relatives- a Global Report. AGP: IBPGR/80/103/. *IBPGR Secretariat*, Rome, p. 65.

IPGRI (1996). Descriptors for Tomato (*Lycopersicon* spp.). ISBN 92-9043-294-2.

Jeffreys A. J., Wilson V., Thein S. L. (1985). Hypervariable 'minisatellite' regions in human DNA. *Nature*, vol. 314, no. 6006, pp. 67–73.

Jones E., Sullivan H., Bhatramakki D., Smith J. (2007). A comparison of simple sequence repeat and single nucleotide polymorphism marker technologies for the genotypic analysis of maize (*Zea mays* L.). *Theoretical and Applied Genetics*, vol. 115, pp. 361–371.

Kalloo G. (1991). Genetic Improvement of Tomato, *Springer*, Berlin, Germany.

Kaemmer D., Weising K., Beyermann B., Borner T., Eppelen J. T., Kahl G. (1995). Oligonucleotide fingerprinting of tomato DNA. *Plant Breeding*, vol. 114, pp. 12-17.

Kanazin V., Talbert H., See D., DeCamp P., Nevo E., Blake T. (2002). Discovery and assay of single-nucleotide polymorphisms in barley (*Hordeum vulgare*). *Plant Molecular Biology*, vol. 48, pp. 529-537.

Kelly R. P., Sarkar I. N., Eernisse D. J., Desalle R. (2007). DNA barcoding using chitons (genus *Mopalia*). *Molecular Ecology Notes*, vol. 7, pp. 177-183.

Kerr K. C. R., Stoeckle M. Y., Dove C. J., Weigt L. A., Francis C. M., Hebert P. D. N. (2007). Comprehensive DNA Barcode coverage of North American Birds. *Molecular Ecology Notes*, vol. 7, pp. 535-543.

Khush G. S., Rick C. M. (1968). Cytogenetic analysis of the tomato genome by means of induced deficiencies. *Chromosoma*, vol. 23, pp. 452-484.

Knapp S., Bohs L., Nee M., Spooner D. M. (2004). *Solanaceae*- a model for linking genomics with biodiversity. *Comparative and Functional Genomics*, vol. 5, no. 3, pp. 285-291.

Labate J. A., Baldo A. M. (2005). Tomato SNP discovery by EST mining and resequencing. *Molecular Breeding*, vol. 16, no. 4, pp. 343-349.

Labate J. A., Robertson L. D., Wu F., Tanksley S. D., Baldo A. M. (2009). EST, COSII, and arbitrary gene markers give similar estimates of nucleotide diversity in cultivated tomato (*Solanum lycopersicum* L.). *Theoretical and Applied Genetics*, vol. 118, pp. 1005-1014.

Levinson G., Gutman G. A. (1987). Slipped-strand mispairing: a major mechanism for DNA sequence evolution. *Molecular Biology and Evolution*, vol. 4, pp. 203-221.

Levis J. S., Thompson W. F., Csinos A. S., Stephenson M. G., Weissinger A. K. (2005). Transgenic attachment regions increase the efficiency and stability of RNA-mediated resistance to tomato spotted wilt virus in transgenic tobacco. *Transgenic Research*, vol. 14, pp. 193-206.

Libro Bianco sulla Sicurezza Alimentare, Bruxelles, 12.1.2000, COM (1999) 719 def.

Linnaeus C. (1753). *Species Plantarum*, Holmiae, Stockholm, Sweden, 1st edition.

Mantel N. (1967). The detection of disease clustering and a generalized regression approach. *Cancer Research*, vol. 27, pp. 209-220.

Martin G. B., Brommonschenkel S. H., Chunwongse J. (1993a). Map-based cloning of a protein kinase gene conferring disease resistance in tomato. *Science*, vol. 262, no. 5138, pp. 1432-1436.

Martin G. B., Vicente M. C., Tanksley S. D. (1993b). High resolution linkage analysis and physical characterization of the *Pto* bacterial resistance locus in tomato. *Molecular Plant Microbiology Interactions*, vol. 6, pp. 26-34.

- Martinez I., Aursand M., Erikson U., Singstad T. E., Veliyulin E., van der Zwaag C. (2003). Destructive and non destructive analytical techniques for authentication and composition analyses of foodstuffs. *Trends in Food Science & Technology*, vol. 14, pp. 489-498.
- McCormick S., Niedermeyer J., Fry J., Barnason A., Horsch R., Fraley R. (1986). Leaf disc transformation of cultivated tomato (*L. esculentum*) using *Agrobacterium tumefaciens*. *Plant Cell Reports*, vol. 5, no. 2, pp. 81–84.
- Melchiade D., Foroni I., Corrado G., Santangelo I., Rao R. (2007). Authentication of the 'Annurca' apple in agro-food chain by amplification of microsatellites loci. *Food Biotechnology*, vol. 21, pp. 33-43.
- Miller P. (1754). *The Gardeners Dictionary*, John and Francis Rivington, London, UK, 4th edition.
- Miller J. C., Tanksley S. D. (1990). RFLP analysis of phylogenetic relationships and genetic variation in the genus *Lycopersicon*. *Theoretical and Applied Genetics*, vol. 80, no. 4, pp. 437–448.
- Mitchelson K. R., Cheng J., Kricka L. J. (1997). The use of capillary electrophoresis for point-mutation screening. *Trends in Biotechnology*, vol. 15, pp. 448-458.
- Monti L. M., Santangelo E., Corrado G., Rao R., Soressi G. P., Scarascia Mugnozza G. T. (2004): Il 'San Marzano': Problematiche e prospettive in relazione alla sua salvaguardia e alla necessita' di interventi genetici. *Agroindustria*, vol. 3, pp. 161-170.
- Mueller L. A., Solow T. H., Taylor N., et al. (2005). The SOL genomics network. A comparative resource for *Solanaceae* biology and beyond. *Plant Physiology*, vol. 138, no. 3, pp. 1310–1317.
- Mutschler M. A., Tanksley S. D., Rick C. M. (1987). Linkage maps of the tomato (*Lycopersicon esculentum*). Report of the Tomato Genetics Cooperative, vol. 37, pp. 5–34.
- Nee M., Hawkes J. G., Lester R. N., Estrada N. (1991). *Solanaceae* III: Taxonomy, Chemistry, Evolution. *The Royal Botanic Garden, Kew, UK*.
- Nesbitt T. C., Tanksley S. D. (2002). Comparative sequencing in the genus *Lycopersicon*: implications for the evolution of fruit size in the domestication of cultivated tomatoes. *Genetics*, vol. 162, pp. 365–379.
- Noble D. (1994). Tantalizing tomatoes from the past. *Am Veget Grower*, vol. 42, pp. 44–47.
- Olmstead R. G., Sweere J. A., Spangler R. E., Bohs L., Palmer J. (1999). Phylogeny and provisional classification of the *Solanaceae* based on chloroplast DNA. *Solanaceae* IV, Nee M., Symon D. E., Lester R. N., Jessop J. P. Eds. pp. 111–117, *Royal Botanic Garden, Kew, UK*.

Orsi C. H., Tanksley S. D. (2009). Natural variation in an ABC transporter gene associated with seed size evolution in tomato species. *Plos Genetics*, vol. 5, e1000347.

Parrella G., Ruffel S., Moretti A., Morel C., Palloix A., Caranta C. (2002). Recessive resistance genes against potyviruses are localized in colinear genomic regions of the tomato (*Lycopersicon* spp.) and pepper (*Capsicum* spp.) genomes. *Theoretical and Applied Genetics*, vol., 105, pp. 855-861.

Patil R. S., Davery M. R., Power J. B., Cocking E. C. (2002). Effective protocols for *Agrobacterium*-mediated leaf disc transformation in tomato (*Lycopersicon esculentum* Mill.). *Indian Journal of Biotechnology*, vol. 1, no. 4, pp. 339–343.

Peakall R., Smouse P., (2006). GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* (electronic), vol 6, pp. 288-295.

Peralta I. E., Spooner D. M. (2001). Granule-bound starch synthesis (GBSSI) gene phylogeny of wild tomatoes (*Solanum* L. section *Lycopersicon* [Mill.] Wettst. subsection *Lycopersicon*). *American Journal of Botany*, vol. 88, no. 10, pp. 1888–1902.

Peterson D. G., Pearson W. R., Stack S. M. (1998). Characterization of the tomato (*Lycopersicon esculentum*) genome using in vitro and in situ DNA reassociation. *Genome*, vol. 41, no. 3, pp. 346–356.

Porretta S. (1995). Quality evaluation of canned whole tomatoes. *Food Science and Technology*, Int. 1, pp. 97-104.

Pritchard J. K., Stephens P., Donnelly P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, vol. 155, pp. 945–959.

Rao R., Corrado G., Bianchi M., Di Mauro A. (2006). (GATA)<sub>4</sub> DNA fingerprinting identifies morphologically characterized 'San Marzano' tomato plants. *Plant Breeding*, vol. 125, pp. 173-176.

Regolamento (CE) n. 2100/94 del Consiglio, del 27 luglio 1994, concernente la privativa comunitaria per ritrovati vegetali.

Regolamento (CE) n. 178/2002 del Parlamento europeo e del Consiglio, del 28 gennaio 2002, che stabilisce i principi e i requisiti generali della legislazione alimentare, istituisce l'Autorità europea per la sicurezza alimentare e fissa procedure nel campo della sicurezza alimentare.

Rick C. M. (1973). Potential genetic resources in tomato species: clues from observations in native habitats. *Genes, Enzymes and Populations*, Srb A. M., Ed., pp. 255–269, Plenum Press, New York (NY) USA.

- Rick C. M., Fobes J. F. (1975). Allozyme variation in the cultivated tomato and closely related species. *Bulletin of the Torrey Botanical Club*, vol. 102, no. 6, pp. 376–384.
- Rick C. M. (1976). Tomato, *Lycopersicon esculentum* (Solanaceae). *Evolution of Crop Plants*, Simmonds N. W., Ed., pp. 268–273, Longman, London, UK.
- Rick C. M. (1979). Potential improvement of tomatoes by controlled introgression of genes from wild speies. *Proceedings of the Conference on Broadening the Genetic Base of Crops*, pp. 167–173, Pudoc, Wageningen, The Netherlands.
- Rick C. M. (1982). The potential of exotic germplasm for tomato improvement. *Plant Improvement and Somatic Cell Genetics*, Vasil I. K., Scowcroft W. R., Frey K. J., Eds., pp. 1–28, Academic Press, New York (NY) USA.
- Rick C. M., Yoder J. I. (1988). Classical and molecular genetics of tomato: highlights and perspectives. *Annual Review of Genetics*, vol. 22, pp. 281–300.
- Rohlf M. (1998). NTSYS-pc. Numerical taxonomy and multivariate analysis system. Version 2. Department of Ecology and Evolution. State University of New York, Albany, N.Y.
- Romersburg H.C. (1984). Cluster Analysis for Research. Belmont (CA) USA. *Lifetime Learning Publications*.
- Ruiz J. J., Garcia-Martinez S., Pico' B., Gao M., Quiros C. F. (2005). Genetic variability and relationship of closely related Spanish traditional cultivars of tomato as detected by SRAP and SSR markers. *Journal of the American Society for Horticultural Science*, vol. 130, no. 1, pp. 88–94.
- Saitou N., Nei M. (1987). The Neighbor-Joining method: A new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, vol.4, pp. 406-425.
- Saliba-Colombani V., Causse M., Gervais L., Philouze J. (2000). Efficiency of RFLP, RAPD, and AFLP markers for the construction of an intraspecific map of the tomato genome. *Genome*, vol. 43, no. 1, pp. 29–40.
- Sambrook J., Fritsh E. F., Maniatis T. (1989). *Molecular cloning: A laboratory manual*, second edition. Cold Spring Harbor Laboratory Press, NY, USA.
- Sambrook J., Fritsh E. F., Maniatis T. (2001). *Molecular cloning: A laboratory manual*, third edition. Cold Spring Harbor Laboratory Press, NY, USA.
- Sandbrink J.M., Van Ooijen J.W., Purimahua C.C., Vrielink M., Verkerk R., Zabel P., Lindhout P. (1995). Localization of genes for bacterial canker resistance in *Lycopersicon peruvianum* using RFLPs. *Theoretical and Applied Genetics*, vol. 95, pp. 444-450.

- Schlotterer C. (2004). The evolution of molecular markers – just a matter of fashion? *Nature Reviews Genetics*, vol. 5, pp. 63-69.
- Schwengel D. A., Jedlicka A. E., Nanthakumar E. J., Weber J. L., Levitt R. C. (1994). Comparison of fluorescence-based semi-automated genotyping of multiple microsatellite loci with autoradiographic techniques. *Genomis*, vol. 22, pp. 46-54.
- Scott J. W., Angell F. F. (1998). Tomato. Hybrid Cultivar Development. Banga S. S., Banga S. K. Eds., pp. 451–475, Narosa, New Delhi, India.
- Scott J. W., Olson S. M., Bartz J. A. (2009). Trybeca hybrid tomato; Fla. 8124c and Fla. 8249 breeding lines. *HortScience*, vol. 44, pp. 471-473.
- Scrimshaw B. J. (1992). A simple nonradioactive procedure for visualization of (dC-dA)<sub>n</sub> dinucleotide repeat length polymorphisms. *Bio/Techniques*, vol. 13, p. 189.
- Shibata D. (2005). “Genome sequencing and functional genomics approaches in tomato,” *Journal of General Plant Pathology*, vol. 71, no. 1, pp. 1–7.
- Sim S. C., Robbins M. D., Chilcott C., Zhu T., Francis D. M. (2009). Oligonucleotide array discovery of polymorphisms in cultivated tomato (*Solanum lycopersicum* L.) reveals patterns of SNP variation associated with breeding. *BMC Genomics*, vol. 10, p. 466.
- Smith M. A., Woodley N. E., Janzen D. H., Hallwachs W., Hebert P. D. N. (2006). DNA barcodes reveal cryptic host-specificity within the presumed polyphagous members of a genus of parasitoid flies (Diptera: Tachinidae). *Proceeding of National Academy of Sciences of the United State of America*, vol. 103, pp. 3657-3662.
- Smulders M. J. M., Bredemeijer G., Rus-Kortekaas W., Arens P., Vosman B. (1997). Use of short microsatellites from database sequences to generate polymorphisms among *Lycopersicon esculentum* cultivars and accessions of other *Lycopersicon* species. *Theoretical and Applied Genetics*, vol. 97, pp. 264-272.
- Sneath P. H., Sokal R. R. (1973). Numerical taxonomy. The principle and practice of numerical classification. WH Freeman and Company, San Francisco (CA) USA.
- Soressi G. P. (1969). Il pomodoro. Ed Agricole, Bologna.
- Spooner D. M., Anderson G. J., Jansen R. K. (1993). Chloroplast DNA evidence for the interrelationships of tomatoes, potatoes, and pepinos (*Solanaceae*). *American Journal of Botany*, vol. 80, no. 6, pp. 676–688.
- Stevens M. A., Rick C. M. (1986). Genetics and breeding. The Tomato Crop: A Scientific Basis for Improvement, Atherton J. G., Rudich J., Eds., pp. 35–109, Chapman and Hall, New York (NY) USA.
- Stevens M. R., Lamb E. M., Rhoads D. D. (1995). Mapping the Sw-5 locus for tomato spotted wilt virus resistance in tomatoes using RAPD and RFLP analyses. *Theoretical and Applied Genetics*, vol. 90, pp. 451-456.

- Tamura K., Dudley J., Nei M., Kumar S. (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular Biology and Evolution*, vol. 24, no. 8, pp. 1596-1599.
- Tanksley S. D. (1993). Mapping polygenes. *Annual Review of Genetics*, vol. 27, pp. 205–233.
- Tanksley S. D. (1993). Linkage map of the tomato (*Lycopersicon esculentum*) (2N = 24). Genetic Maps: Locus Maps of Complex Genomes, S. J. O'Brian, Ed., pp. 6.3–6.15, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, USA.
- Tanksley S. D., Bernatzky R. (1987). Molecular markers for the nuclear genome of tomato. *Plant Biology*, vol. 4, Tomato Biotechnology, Nevins D. J., Jones R. A. Eds., pp. 37–44, Alan R. Liss, New York (NY) USA.
- Tanksley S. D., Nelson J. C. (1996). Advanced backcross QTL analysis: a method for the simultaneous discovery and transfer of valuable QTLs from unadapted germplasm into elite breeding lines. *Theoretical and Applied Genetics*, vol. 92, pp. 191–203.
- Tanksley S. D., Orton T. J. (1968). Isozymes. *Plant Genetics and Breeding*, Elsevier, Amsterdam, The Netherlands.
- Tanksley S. D., Young N. D., Paterson A. H., Bonierbale M. W. (1989). RFLP mapping in plant breeding: new tools for an oldscience. *BioTechnology*, vol. 7, pp. 257-264.
- Tauz D. (1989). Hypervariability of simple sequence as a general source for polymorphic DNA markers. *Nucleic Acids Research*, vol. 17, pp. 6463-6471.
- Tengel C., Schüssler P., Setzke E., Balles J., Sprenger-Haussels M. (2002). Detection of genetically modified soybean and maize in raw and processed foodstuffs. *Qiagen News*, vol. 2, pp. 16–18.
- Terrell E. E., Broome C. R., Reveal J. L. (1983). Proposal to conserve the name of the tomato as *Lycopersicon esculentum* P. Miller and reject the combination *Lycopersicon lycopersicum* (L.) Karsten (*Solanaceae*). *Taxon*, vol. 32, no. 2, pp. 310–314.
- Thompson J. D., Higgins D. G., Gibson T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, vol.22, no. 22, pp. 4673-80.
- Tigchelaar E. C. (1986). Tomato breeding. *Breeding for Vegetable Crops*. Bassett M. J., Ed., pp. 135–171, AVI, Westport, Conn, USA.
- Tommasini, L., Batley J., Arnold G. M., Cooke R. J., Donini P., Lee D., Law J. R., Lowe C., Moule C., Trick M., Edwards K. J. (2003). The development of multiplex simple sequence repeat (SSR) markers to complement Distinctness, Uniformity and

Stability testing of rape (*Brassica napus* L.) varieties. *Theoretical and Applied Genetics*, vol. 106, pp. 1091—1101.

UNI EN ISO 22005 (2008). "Rintracciabilità nelle filiere agroalimentari - Principi generali e requisiti di base per la progettazione di sistemi e attuazione".

UPOV (2001). Guidelines for the Conduct of Tests for Distinctness, Uniformity and Stability, vol. TG/44/10. UPOV, Geneva.

van der Hoeven R. S., Ronning C., Giovannoni J. J., Martin G., Tanksley S.D. (2002). Deductions about the number, organization and evolution of genes in the tomato genome based on analysis of a large expressed sequence tag collection and selective genomic sequencing. *The Plant Cell*, vol. 14, no. 7, pp.1441-1456.

Van Deynze A. E., Stoffel K., Buell R. C., Kozik A., Liu J., van der Knaap E., Francis D (2007). Diversity in conserved genes in tomato. *BMC Genomics*, vol. 81, p. 465.

Van Deynze A. E., Wilkins T. A., Stoffel K., Lee M., Stelly D., Kozik A. (2007). A Set of Informative Markers Designed Specifically for Breeding Cotton. *Plant and Animal Genome XV*, San Diego (CA) USA.

Van der Beek J. G., Verkerk R., Zabel P., Lindhout P. (1992). Mapping strategy for resistance genes in tomato based on RFLPs between cultivars: *Cf9* (resistance to *Cladosporium fulvum*) on chromosome 1. *Theoretical and Applied Genetics*, vol. 84, pp. 106-112.

Van Ooijen J.W., Sandbrink J.M., Vrieling M., Verkerk R., Zabel P., Lindhout P. (1994). An RFLP linkage map of *Lycopersicon peruvianum*. *Theoretical and Applied Genetics*, vol. 89, pp. 1007-1013.

Vos P., Hogers R., Bleeker M. (1995). AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research*, vol. 23, no. 21, pp. 4407–4414.

Vosman B., Arens P. (1997). Molecular characterization of GATA/GACA microsatellite repeats in tomato. *Genome*, vol. 40, pp. 25-33.

Vosman B., Arens P., Ruskortekaas W., Smulders M. J. M. (1992). Identification of highly polymorphic DNA regions in tomato. *Theoretical and Applied Genetics*, vol. 85, pp. 239-244.

Wang Y., van der Hoeven R. S., Nielsen R., Mueller L. A., Tanksley S. D. (2005). Characteristics of the tomato nuclear genome as determined by sequencing undermethylated *EcoRI* digested fragments. *Theoretical and Applied Genetics*, vol. 112, no. 1, pp. 72–84.

Wang Y., van der Hoeven R., Nielsen R. (2006). Characteristics of the tomato nuclear genome as determined by sequencing unmethylated DNA and euchromatic and pericentromeric BACs. *Plant and Animal Genome XIV Conference*, W176, p. 147, USDA, San Diego (CA) USA.

Wang Z., Weber J. L., Zhong G., Tanksley S. D. (1994). Survey of plant short tandem repeats. *Theoretical and Applied Genetics*, vol. 88, pp. 1-6.

Williams C. E., St. Clair D. A. (1993). Phenetic relationships and levels of variability detected by restriction fragment length polymorphism and random amplified polymorphic DNA analysis of cultivated and wild accessions of *Lycopersicon esculentum*. *Genome*, vol. 36, no. 3, pp. 619–630.

Yamamoto N., Tsugane T., Watanabe M., Yano K., Maeda F., Kuwata C., Toriki M., Ban Y., Nishimura S., Shibata D. (2005). Expressed sequence tags from the laboratory-grown miniature tomato (*Lycopersicon esculentum*) cultivar Micro-Tom and mining for single nucleotide polymorphisms and insertions/deletions in tomato cultivars. *Gene*, vol. 356, pp. 127–134.

Yang W., Bai X., Kabelka E., Eaton C., Kamoun S., van der Knaap E., Francis D. M. (2004). Discovery of single nucleotide polymorphisms in *Lycopersicon esculentum* by computer aided analysis of expressed sequence tags. *Molecular Breeding*, vol. 14, pp. 21-34.

Yang W., Francis D. M. (2005). Marker Assisted Selection for Combining Resistance to Bacterial Spot and Bacterial Speck in Tomato. *Journal of the American Society of Horticultural Science*, vol. 130, pp. 716-721.

Zhu Y. L., Song Q. J., Hyten D. L., Van Tassell C. P., Matukumalli L. K., Grimm D. R., Hyatt M., Fickus E. W., Young N. D., Cregan P. (2003). Single-Nucleotide Polymorphisms in Soybean. *Genetics*, vol. 163, pp. 1123-1134.

### **Periodo di formazione all'estero:**

Stage di 9 mesi (da settembre 2008 a dicembre 2008 e da aprile 2009 ad agosto 2009) presso il laboratorio del prof. David Francis, The Ohio State University, USA. Durante questo periodo e' stata effettuata la genotipizzazione del materiale oggetto di studio del progetto di dottorato mediante marcatori molecolari SNP ed In/Del, con particolare riguardo allo sviluppo di marcatori per gli studi di tracciabilita' genetica nella filiera agro-alimentare del pomodoro.

## Pubblicazioni

**Caramante M.**, Rao R., Monti L. M., Corrado G. (2009). Discrimination of 'San Marzano' accessions: A comparison of minisatellite, CAPS and SSR markers in relation to morphological traits. *Scientia Horticulturae*. 120:560-564.

Rao R., **Caramante M.**, Blanco A., Lanteri S., Lucchin M., Mazzucato A. (2009). Innovazioni genetiche per l'identificazione e la protezione di prodotti tipici italiani. *Italian Journal of Agronomy, Riv. Agron.* Vol. 4, No. 3 Suppl.

**Caramante M.**, Rao R., Monti L., Corrado G. VNTR, SSR and morphological characterization of tomato accessions spread in Campania. 52Mo convegno annuale SIGA- Societa' Italiana di Genetica Agraria, Padova 14/17 Settembre 2008.

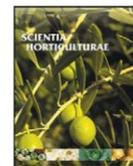
**Caramante M.**, Rao R., Corrado G. A comparison of minisatellite and SSR markers in relation to morphological traits for discrimination of 'San Marzano' accession. 52Mo convegno annuale SIGA- Societa' Italiana di Genetica Agraria, Padova 14/17 Settembre 2008.

Corrado G., Scarano D., **Caramante M.** Diversita' genetica di accessioni di pomodoro campane. Giornate Scientifiche, Facolta' di Agraria, Portici, Dicembre 2008.



Contents lists available at ScienceDirect

Scientia Horticulturae

journal homepage: [www.elsevier.com/locate/scihorti](http://www.elsevier.com/locate/scihorti)

Short communication

## Discrimination of 'San Marzano' accessions: A comparison of minisatellite, CAPS and SSR markers in relation to morphological traits

Martina Caramante, Rosa Rao, Luigi Maria Monti, Giandomenico Corrado\*

Dipartimento di Scienze del Suolo, della Pianta, dell'Ambiente e delle Produzioni Animali, Università degli Studi di Napoli 'Federico II', 80055 Portici, Napoli, Italy

## ARTICLE INFO

## Article history:

Received 1 September 2008

Received in revised form 24 November 2008

Accepted 9 December 2008

## Keywords:

Tomato

GATA

SNP

Microsatellite

Genetic distance

Morphological analysis

## ABSTRACT

'San Marzano' (SM) is one of the most widely known tomato (*Solanum lycopersicum* L.) cultivars, and is a classic example of a local variety with a premium value. Unfortunately, the original cultivated form is underrepresented in the Protected Denomination of Origin (PDO) area because of the incidence of contaminant and phenotypically similar genotypes. Our aim was to examine the ability of three DNA marker systems (minisatellite, cleaved amplified polymorphic sequence (CAPS) and simple sequence repeat (SSR)) to reveal the genetic diversity of tomato accessions that were, based on a morphological analysis, very similar. The data indicate that both minisatellites and SSRs can be used to genetically distinguish the analysed materials. Furthermore, these two marker systems depict relationships consistent with the hierarchical pattern obtained by the morphological data. As locally cultivated tomato accessions are often characterised by some degree of genetic variability, our results will be valuable in facilitating the purification, management and breeding of tomato germplasms. The differences between the marker systems employed are also discussed in relation to their usefulness in the agro-food chain.

© 2008 Elsevier B.V. All rights reserved.

### 1. Introduction

The tomato variety 'San Marzano' (SM) is a symbol of the Italian food-industry and represents, arguably, its most significant horticultural product. However, during the past few decades, the use of this variety has suffered a decline due to its genetic deficiencies and the subsequent diffusion of more productive tomato cultivars with similar fruit shapes (Monti et al., 2004). Modern hybrids possess morphological and quantitative traits that undoubtedly make them more suitable for fresh market or food-processing purposes, but they are genetically different from the original variety. Hence, they should be considered as such and labelled as 'San Marzano-type'. The approval of the EU Protected Denomination of Origin (PDO) certification for 'San Marzano' is a valid action to protect and value the 'San Marzano' trade name, but additional measures should be undertaken for its long-term preservation. In many instances, it is essential to have adequate tools to discriminate the original variety from genetic material of different origins.

A previous study indicated that DNA fingerprinting based on the (GATA)<sub>4</sub> minisatellite can identify San Marzano plants (Rao et al., 2006), and proved that there are two main problems for 'San Marzano' accessions in the PDO area. The first issue is the presence

of similar plants that possess many but not all the 'San Marzano' phenotypic attributes. The second is the variable occurrence in several accessions of contaminant genotypes among true 'San Marzano' plants. To solve these problems, it would be useful to be able to distinguish 'San Marzano' with a fast, cost-effective and accurate means of genetic identification.

Different molecular markers are available for detecting genetic variation. In tomato, high levels of polymorphism have been reported with the use of a (GATA)<sub>4</sub> oligonucleotide probe, which proved to be useful for distinguishing cultivars and accessions that are otherwise difficult to differentiate (Kaemmer et al., 1995; Rao et al., 2006). Cleaved amplified polymorphic sequence (CAPS) polymorphisms are differences in restriction fragment lengths caused by single nucleotide polymorphisms (SNPs) that modify restriction sites in PCR amplicons (Yang et al., 2004). CAPSs are therefore the less technically demanding approach for SNP analysis. Simple sequence repeats (SSRs) have the advantage of being highly informative (Bredemeijer et al., 1998; Cooke et al., 2003), and unlike SNPs, knowledge of the amplified sequences or further enzymatic manipulations are not needed (Schlotterer, 2004). However, SSRs require high resolution fluorescent-based capillary systems to fully exploit the information they can provide. Typically, different DNA marker systems have been evaluated for the detection of inter-varietal differences but the morphological similarity of the plant material analysed is rarely indicated. For these reasons, the aim of this work was to compare the efficacy of minisatellite, CAPS and SSR molecular markers for the distinction

\* Corresponding author. Tel.: +39 081 2539209; fax: +39 081 7753579.  
E-mail address: [giandomenico.corrado@unina.it](mailto:giandomenico.corrado@unina.it) (G. Corrado).

of the true 'San Marzano' cultivar from very closely related materials. This analysis was carried out in relation to data from a morphological analysis.

## 2. Materials and methods

### 2.1. Plant materials and morphological analysis

From among the different cultivated forms of 'San Marzano' tomato (*Solanum lycopersicum* L.) of the 'San Marzano' PDO area (Rao et al., 2006), we selected three accessions that are referred to in this study as SMAR, SMID and SMON. The first represents the true 'San Marzano' (Rao et al., 2006), the second is an example of a cultivated form containing a contaminant (i.e., a genotype, present at a notable level, that does not belong to 'San Marzano'), and the third is the tomato accession that looks very similar to 'San Marzano'. The 'Sorrento' (indicated in this study as RG), an indeterminate tomato with rounded fruits, was also included as an example of a locally cultivated variety with fruits that are easily distinguishable from 'San Marzano'. Furthermore, the efficiency of the CAPS was also evaluated on five commercial varieties that differ in their fruit shape, reported in brackets. The cultivars were 'Cirio 3' (elongated), 'Minidor' (cherry), 'PS 1296' (rounded), 'Talent' (elongated) and 'Tomito' (cherry).

Seeds were sown in March in pots with sterilised soil. Plants were grown in a greenhouse without supplemental lighting, and support was provided with stakes and ties as the plants grew. For each accession, 15 quantitative traits were analysed in two consecutive years at the two cotyledons stage (leaf length and plant height), before flowering (stem height) and when fully grown (leaf length and width, petiole length, height of the nodes with the first inflorescence, internode length). The fruit descriptors were pedicel length, number per vine and per plant, and length, width, height and weight at maturity. The dissimilarity coefficients between cultivars were computed using Euclidean distances after standardization of the data matrix. To calculate the standardized value  $Z_{ij}$  for any  $i$ th attribute and the  $j$ th object, we subtracted from the corresponding value  $X_{ij}$  in the data matrix, the mean of the values of the  $i$ th attribute, and divided the result by the standard deviation of the values of the  $i$ th attribute. The accessions were clustered by the unweighted pair-group method with arithmetic averages (UPGMA). A cophenetic value matrix (Sneath and Sokal, 1973) of the UPGMA clustering was also used to test for goodness of fit of the clustering to the similarity matrix on which it was based by computing the product-moment correlation ( $r$ ) with 1000 permutations (Mantel, 1967). These analyses were conducted using the NTSYS-PC v. 2.1 package (Rohlf, 1998).

### 2.2. DNA isolation and (GATA)<sub>4</sub> DNA fingerprints

Leaves were collected and immediately frozen in liquid nitrogen. Total DNA was isolated using a previously reported CTAB-based method and digested by *Taq* I (Andreakis et al., 2004). DNA fingerprinting using the (GATA)<sub>4</sub> minisatellite was also carried out as previously described (Rao et al., 2006; Vosman et al., 1992). Twenty-eight bands were used to test genetic diversity and each fragment was assumed to represent a single locus. Only reproducible polymorphic bands were scored for presence (1) or absence (0) and monitored by comparative analysis of the hybridisation pattern obtained from at least four different plants per accession.

### 2.3. SSR analysis

Fourteen tomato SSR loci, representing various repeat classes, were used for PCR amplification. The loci analysed were LE20592,

LE21085, LELE25, LEEF1Aa (Smulders et al., 1997), LEat002, LEga003, LEat002, LEat002, LEca001, LEat007, LEct001, LEct001, LEta015 and LEta003 (He et al., 2003). PCR amplification of DNA isolated from leaves was performed in a 25- $\mu$ l volume containing 20 ng genomic DNA, 1.5 mM MgCl<sub>2</sub>, 100  $\mu$ M deoxyribonucleotides, 0.2  $\mu$ M fluorescently labelled forward primer and unlabelled reverse primer, and 0.5 U Taq DNA polymerase (Promega) in 1 $\times$  PCR buffer. PCR reactions were performed in an Eppendorf Mastercycler Gradient thermocycler. The annealing temperature ( $T_a$ ) varied depending on the primer pair used, according to the literature (He et al., 2003; Smulders et al., 1997). The amplification conditions were one cycle of 94 °C for 3 min; 35 cycles of 94 °C for 45 s,  $T_a$  for 45 s and 72 °C for 90 s. After the final cycle, a 5-min step at 72 °C was added. Amplification products were separated by agarose gel-electrophoresis to verify the presence of amplified fragments. For allelic discrimination, the fluorescent fragments were resolved by capillary electrophoresis and detected in an ABI Prism 3100-Avant (Applied Biosystems) system using the POP4 polymer (Applied Biosystems). Signal peak height and allele sizes were calculated using the ABI Prism Genotyper 2.5 (v. 3.7) software (Applied Biosystems) on the basis of the GeneScan 500Liz molecular weight standard (Applied Biosystems).

### 2.4. CAPS analysis

For genotyping the tomato plants, fourteen *S. lycopersicum* EST-SNPs selected from the literature (Yang et al., 2004) were assayed as CAPS. Furthermore, we analysed the CAPS marker REX-1 that is tightly linked to the nematode resistance gene *Mi* (Williamson et al., 1994). The dominant allele is present in modern tomato lines and hybrids (VFN cultivars), and the analysis was carried out to exclude the presence of this contaminating material. PCR reactions were carried out as described for the SSR analysis using previously reported primers (Yang et al., 2004). The amplified fragments were digested with the restriction enzymes indicated in Table 1. Fragments were resolved on a 3% agarose gel (Sambrook et al., 1989) and their molecular weights were calculated on a standard curve based on the lower 10 bands (from 2000 to 100 bp) of the '1 kb Plus DNA ladder' (Invitrogen) with Molecular Imaging Software v. 4.0.3 (Kodak).

### 2.5. Molecular data analysis

Genetic similarities were evaluated by calculating the Dice coefficient (Dice, 1945), and samples were clustered by the UPGMA method. A cophenetic value matrix (Sneath and Sokal, 1973) of the UPGMA clustering was used to test the goodness of fit of the clustering to the similarity matrix on which it was based by computing the product-moment correlation ( $r$ ) with 1000 permutations (Mantel, 1967). The Mantel test was also used to test whether one set of relationships among a set of objects (i.e., distance matrices obtained using different data) is independent of another. These analyses were conducted using the NTSYS-PC v. 2.1 package (Rohlf, 1998).

## 3. Results

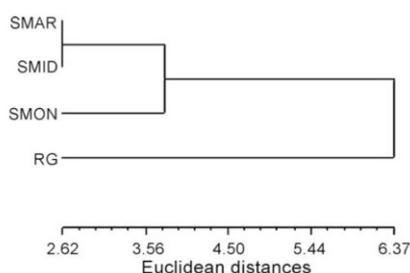
### 3.1. Morphological and minisatellite analysis

A morphological analysis was carried out to quantify phenotypic differences between the analysed accessions. Fig. 1 shows the UPGMA dendrogram based on Euclidean distances. While the SMAR clustered with the SMID accession, showing a very small dissimilarity, the SMON accession was more distantly related. The RG was clearly separated by the previous group, as expected. We

**Table 1**

SNPs selected for CAPS analysis and their main features.

| SNP       | Codon substitution | Restriction enzyme | PCR product (bp) | Cleaved fragments <sup>a</sup> (bp) |
|-----------|--------------------|--------------------|------------------|-------------------------------------|
| LeOH 2    |                    | BsaW I             | 201              | 109, 92                             |
| LeOH 8.4  | Synonymous         | Hae III            | 172              | 91, 81                              |
| LeOH 11.1 |                    | Hha I              | 189              | 148, 41                             |
| LeOH 12.2 |                    | Mse I              | 181              | 158, 23                             |
| LeOH 16.2 | Non-synonymous     | BsaW I             | 206              | 116, 90 <sup>b</sup>                |
| LeOH 16.3 |                    | BsaW I             | 206              | 152, 54 <sup>b</sup>                |
| LeOH 20.1 | Non-synonymous     | Hae III            | 246              | 126, 120                            |
| LeOH 23.1 | 5'-UTR             | Msp I              | 191              | 127, 64                             |
| LeOH 25.1 | Synonymous         | Hinc II            | 644              | 471, 77, 52, 44 <sup>c</sup>        |
| LeOH 28   |                    | Taq I              | 223              | 114, 109                            |
| LeOH 31.3 | Synonymous         | Msp I              | 388              | 307, 81                             |
| LeOH 31.4 |                    | Mse I              | 388              | 217, 87, 48, 37 <sup>d</sup>        |
| LeOH 32.1 | Synonymous         | Hae III            | 521              | 249, 242, 30 <sup>e</sup>           |
| LeOH 39   |                    | EcoR II            | 191              | 162, 29                             |
| REX-1     |                    | Taq I              | 650              | 500, 150                            |

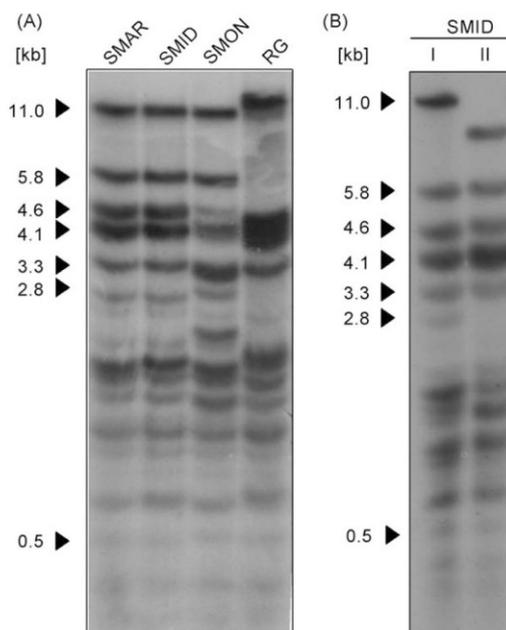
<sup>a</sup> Size of the digested PCR products in the presence of the polymorphic restriction site.<sup>b</sup> The simultaneous digestion of LEOH16.2 and LEOH16.3 gives rise to fragments of 116, 54 and 36 bp.<sup>c</sup> The absence of the polymorphic Hinc II site gives rise to fragments of 471, 128 and 44 bp.<sup>d</sup> The absence of the polymorphic Mse I site gives rise to fragments of 265, 87 and 37 bp.<sup>e</sup> The absence of the polymorphic Hae III site gives rise to fragments of 491 and 30 bp.**Fig. 1.** UPGMA dendrogram based on Euclidean distances calculated from 15 morphological traits.

then evaluated the genomic differences by  $(GATA)_4$  DNA fingerprinting. Fig. 2 shows the pattern of hybridisation after *Taq* I digestion, which revealed several fragments distributed over the 0.5–12.5-kb range. The profile of the SMAR accession included fragments with estimated sizes of 11.0, 5.8, 4.6, 4.1, 3.3 and 2.8 kb, with no other fragment above 2.8 kb, a fingerprint that can be considered the only 'San Marzano' specific one (Rao et al., 2006). The SM-pattern was also shared by the accession SMID. Nonetheless, minisatellite fingerprinting revealed that this accession is not genetically uniform, and Fig. 2B shows a representative hybridisation of the different genotypes that are present. The SM-pattern is not present in the SMON and RG accessions, with the latter showing the most different hybridisation pattern. To represent the genetic distances between samples, a UPGMA dendrogram was constructed on a similarity matrix based on a binary matrix of the pattern presented in Fig. 2A. The dendrogram (Fig. 3) confirmed the relationships observed with the morphological analysis. The comparison of the molecular and morphological similarity distances indicated a very strong association ( $r = 0.98$ ;  $p < 0.05$ ).

### 3.2. SSR analysis

We then investigated the power of SSR fingerprinting, and Table 2 shows the allelic profiles we obtained. As expected for non-hybrid tomato cultivars, we obtained a single peak for the analysed loci, indicating that all the genotypes are homozygous. Of the 14 SSR analysed, three did not reveal any polymorphism. The data indicate that SMAR and SMID I (the genotype of the SMID accession with the SM-pattern) do not show allelic differences at the

analysed loci. Moreover, allelic differences between the SMID I and the SMID II, the SMID genotype with a  $(GATA)_4$  profile different from SMAR, are present at four loci. Finally, we found that only two SSR amplifications could distinguish the SMON accession from 'San Marzano', while RG was observed to be the most different. UPGMA hierarchical classification based on SSR data (Fig. 3) confirmed the relationship observed from minisatellite analysis. However, the SSR-based similarity matrix was not significantly associated with the morphological data ( $r = 0.51$ ;  $p = 0.18$ ). This is mainly due to the fact that the genetic distance between the group of the 'San Marzano' accessions and RG is larger when calculated based on the SSR data.

**Fig. 2.**  $(GATA)_4$  DNA fingerprinting of tomato accessions analysed (A) and an example of the genetic heterogeneity of the SMID accession (B). SMID I is the genotype with the SM-profile, while SMID II is the contaminant genotype. The estimated molecular size (kb) of the six SM-specific bands is indicated on the left.

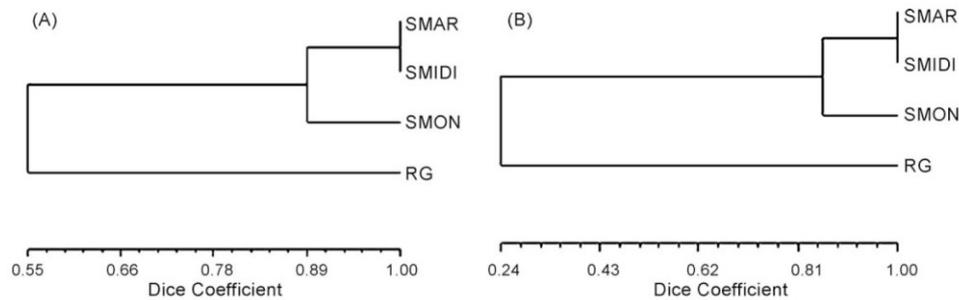


Fig. 3. UPGMA dendrograms based on molecular data obtained from minisatellite (A) and SSR markers (B). Genotypes are as in Fig. 2.

Table 2

Allelic profile at the 14 SSR loci in the 'San Marzano' accessions (see Section 2 for abbreviations). All genotypes were homozygous at all loci. Sizes are in bp.

| SSR      | Accession |      |        |         |     |
|----------|-----------|------|--------|---------|-----|
|          | SMAR      | SMON | SMID I | SMID II | RG  |
| LE20592  | 167       | 167  | 167    | 164     | 164 |
| LE21085  | 102       | 102  | 102    | 102     | 102 |
| LEEF1Aa  | 203       | 203  | 203    | 203     | 201 |
| LELE25   | 220       | 220  | 220    | 220     | 220 |
| LEaat002 | 102       | 102  | 102    | 102     | 105 |
| LEaat007 | 96        | 96   | 96     | 96      | 99  |
| LEat002  | 202       | 202  | 202    | 202     | 198 |
| LEcaa001 | 101       | 101  | 101    | 101     | 101 |
| LEct001  | 106       | 106  | 106    | 106     | 97  |
| LEctt001 | 93        | 93   | 93     | 93      | 91  |
| LEga003  | 233       | 233  | 233    | 233     | 231 |
| LEta003  | 106       | 104  | 106    | 104     | 104 |
| LEta015  | 116       | 112  | 116    | 112     | 108 |
| LEtat002 | 198       | 198  | 198    | 196     | 196 |

### 3.3. CAPS analysis

Subsequently, SNP datapoints were assayed by CAPS analysis firstly on the five genotypes identified by (GATA)<sub>4</sub> fingerprint (Fig. 2). Table 3 shows the result of the restriction analysis. The digestion profiles did not show any differences between genotypes. Furthermore, analysis of the *Mi* locus implied that contaminant genotypes are not correlated with modern VFN tomato cultivars (not shown). The 14 CAPSs employed for genotyping are based on SNPs that have a high probability (53%) of detecting polymorphisms in *S. lycopersicum* varieties (Yang et al., 2004). However, to exclude that the lack of polymorphism in the San Marzano types is due to a poor selection of CAPS markers, we tested their discriminating power on five commercial varieties with different fruit shape. The data indicated that the CAPS employed in this study can discriminate the commercial varieties. Furthermore, it is also possible to distinguish

the 'San Marzano' variety (SMAR) from the other cultivated varieties (Table 3). Overall, the data suggest that in tomato SNPs have a lower efficiency to detected intra-varietal differences in comparison to the other molecular markers tested.

### 4. Discussion

Assessing genetic diversity and performing unambiguous cultivar discrimination using molecular markers is important not only for tomato breeding but also for efficient management and conservation of plant genetic materials (Mazzucato et al., 2008; Ruiz et al., 2005). The preservation of the true 'San Marzano' accession is strictly dependent upon tools that allow the distinction of contaminant and phenotypically similar genotypes that are also present in the original area of cultivation (Rao et al., 2006). The genetic heterogeneity in local populations is probably the main cause of the difficulty in discriminating materials from modern hybrids when analyses are carried out based exclusively on compositional parameters (Loiudice et al., 1995). Furthermore, the presence of improperly labelled material can easily generate confusion and dissatisfaction among consumers who often have not had the opportunity to appreciate the distinctive features of the original 'San Marzano' tomato (Loiudice et al., 1995).

In recent years, a variety of molecular markers have been used to discriminate wild from cultivated tomato plants and to determine genetic distances (Grandillo and Tanksley, 1996; Ruiz et al., 2005; Smulders et al., 1997; Yang et al., 2004). Their usefulness depends on many factors, including the cost and simplicity of the assay. The present study documents the comparative efficacy of three marker types in the discrimination of tomato plants and their relationship to morphological traits. Our results indicate that, despite the limited genetic diversity of tomato (Miller and Tanksley, 1990), DNA fingerprints based on SSR allowed for discrimination not only of different cultivars (He et al., 2003; Smulders et al., 1997), but also of morphologically similar types. On the contrary, the CAPS assay, the most inexpensive

Table 3

SNP profiling as detected by a CAPS assay.

| Genotype | LeOH2 | LeOH8.4 | LeOH11.1 | LeOH12.2 | LeOH16.2 | LeOH16.3 | LeOH20.1 | LeOH23.1 | LeOH25.1 | LeOH28 | LeOH31.3 | LeOH31.4 | LeOH32.1 | LeOH39 |
|----------|-------|---------|----------|----------|----------|----------|----------|----------|----------|--------|----------|----------|----------|--------|
| SMAR     | +/+   | -/-     | -/-      | +/+      | +/+      | -/-      | +/+      | -/-      | -/-      | -/-    | -/-      | -/-      | -/-      | -/-    |
| SMON     | +/+   | -/-     | -/-      | +/+      | +/+      | -/-      | +/+      | -/-      | -/-      | -/-    | -/-      | -/-      | -/-      | -/-    |
| SMID I   | +/+   | -/-     | -/-      | +/+      | +/+      | -/-      | +/+      | -/-      | -/-      | -/-    | -/-      | -/-      | -/-      | -/-    |
| SMID II  | +/+   | -/-     | -/-      | +/+      | +/+      | -/-      | +/+      | -/-      | -/-      | -/-    | -/-      | -/-      | -/-      | -/-    |
| RG       | +/+   | -/-     | -/-      | +/+      | +/+      | -/-      | +/+      | -/-      | -/-      | -/-    | -/-      | -/-      | -/-      | -/-    |
| Cirio3   | +/+   | +/+     | -/-      | +/+      | +/+      | -/-      | +/+      | +/+      | -/-      | -/-    | -/-      | -/-      | -/-      | -/-    |
| Minidor  | +/+   | +/+     | -/-      | +/+      | +/+      | -/-      | +/+      | +/+      | +/+      | -/-    | +/+      | +/+      | +/+      | -/-    |
| PS1296   | +/+   | -/-     | -/-      | +/+      | +/+      | -/-      | +/+      | +/+      | -/-      | -/-    | -/-      | -/-      | -/-      | -/-    |
| Talent   | +/+   | +/+     | -/-      | +/+      | +/+      | -/-      | +/+      | +/+      | +/+      | -/-    | +/+      | +/+      | +/+      | -/-    |
| Tomito   | +/+   | +/+     | -/-      | +/+      | +/+      | -/-      | +/+      | +/+      | +/+      | -/-    | +/+      | +/+      | +/+      | -/-    |

Polymorphic restriction site: presence (+); absence (-).

method, did not reveal any level of intra-varietal polymorphism, although it is possible that a larger number of CAPSs may enable the distinction among closely related tomato accessions. Both SSRs and the (GATA)<sub>4</sub> minisatellite, most likely due to their multiallelic nature, were useful in depicting genetic relationships consistent with the hierarchical pattern generated by the morphological data, yet SSR analysis possesses various advantages over (GATA)<sub>4</sub>-based discrimination. The former is faster and more suitable for high-throughput analysis, as well as traceability in the agro-food chain (Melchiade et al., 2007; Schlotterer, 2004), as all current DNA-based methods for determining the authenticity of food rely on PCR-based molecular markers (Martinez et al., 2003).

In conclusion, the present study illustrates some advantages and disadvantages of three types of DNA molecular markers in tomato. Although the discriminating power of SSR in tomato is lower than in other autogamous species (He et al., 2003; Smulders et al., 1997), we demonstrated that some SSR markers can be efficiently used for intra-varietal distinction, as they discriminated between true 'San Marzano' and both very closely related genotypes and contaminant accessions.

#### Acknowledgment

We thank the "Semiorto Sementi" for sharing its germplasm collection.

#### References

- Andreakis, N., Giordano, I., Pentangelo, A., Fogliano, V., Graziani, G., Monti, L.M., Rao, R., 2004. DNA fingerprinting and quality traits of corbarino cherry-like tomato landraces. *Journal of Agricultural and Food Chemistry* 52, 3366–3371.
- Bredemeijer, G.M.M., Arens, P., Wouters, D., Visser, D., Vosman, B., 1998. The use of semi-automated fluorescent microsatellite analysis for tomato cultivar identification. *Theoretical and Applied Genetics* 97, 584–590.
- Cooke, R.J., Bredemeijer, G.M.M., Ganai, M.W., Peeters, R., Isaac, P., Rendell, S., Jackson, J., Roder, M.S., Korzun, V., Wendehake, K., Areshchenkova, T., Dijcks, M., Laborie, D., Bertrand, L., Vosman, B., 2003. Assessment of the uniformity of wheat and tomato varieties at DNA microsatellite loci. *Euphytica* 132, 331–341.
- Dice, L.R., 1945. Measures of the amount of ecologic association between species. *Ecology* 297–302.
- Grandillo, S., Tanksley, S.D., 1996. Genetic analysis of RFLPs, GATA, microsatellites and RAPDs in a cross between *L. esculentum* and *L. pimpinellifolium*. *Theoretical and Applied Genetics* 92, 957–965.
- He, C., Poysa, V., Yu, K., 2003. Development and characterization of simple sequence repeat (SSR) markers and their use in determining relationships among *Lycopersicon esculentum* cultivars. *Theoretical and Applied Genetics* 106, 363–373.
- Kaemmer, D., Weising, K., Beyermann, B., Borner, T., Epplen, J.T., Kahl, G., 1995. Oligonucleotide fingerprinting of tomato DNA. *Plant Breeding* 114, 12–17.
- Loiudice, R., Impembo, M., Laratta, B., Villari, G., Lovoi, A., Siviero, P., Castaldo, D., 1995. Composition of San Marzano tomato varieties. *Food Chemistry* 53, 81–89.
- Mantel, N.A., 1967. The detection of disease clustering and a generalized regression approach. *Cancer Research* 209–220.
- Martinez, I., Aursand, M., Erikson, U., Singstad, T.E., Veliyulin, E., van der Zwaag, C., 2003. Destructive and non-destructive analytical techniques for authentication and composition analyses of foodstuffs. *Trends in Food Science & Technology* 14, 489–498.
- Mazzucato, A., Papa, R., Bitocchi, E., Mosconi, P., Nanni, L., Negri, V., Picarella, M.E., Siligato, F., Soressi, G.P., Tiranti, B., Veronesi, F., 2008. Genetic diversity, structure and marker-trait associations in a collection of italian tomato (*Solanum lycopersicum* L.) landraces. *Theoretical and Applied Genetics* 116, 657–669.
- Melchiade, D., Foroni, I., Corrado, G., Santangelo, I., Rao, R., 2007. Authentication of the 'Annurca' apple in agro-food chain by amplification of microsatellite loci. *Food Biotechnology* 21, 33–43.
- Miller, J.C., Tanksley, S.D., 1990. RFLP analysis of phylogenetic-relationships and genetic-variation in the genus *Lycopersicon*. *Theoretical and Applied Genetics* 80, 437–448.
- Monti, L.M., Santangelo, E., Corrado, G., Rao, R., Soressi, G.P., Scarascia Mugnozza, G.T., 2004. Il 'San Marzano': Problematiche e prospettive in relazione alla sua salvaguardia e alla necessità di interventi genetici. *Agroindustria* 3, 161–170.
- Rao, R., Corrado, G., Bianchi, M., Di Mauro, A., 2006. (GATA)<sub>4</sub> DNA fingerprinting identifies morphologically characterized 'San Marzano' tomato plants. *Plant Breeding* 125, 173–176.
- Rohlf, F., 1998. Ntsyspc Version 2.0: User Guide. Applied Biostatistics.
- Ruiz, J.J., Garcia-Martinez, S., Pico, B., Gao, M.Q., Quiros, C.F., 2005. Genetic variability and relationship of closely related spanish traditional cultivars of tomato as detected by SRAP and SSR markers. *Journal of the American Society for Horticultural Science* 130, 88–94.
- Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. *Molecular Cloning, a Laboratory Manual*, 2nd ed. Cold Spring Harbor Laboratory Press.
- Schlotterer, C., 2004. The evolution of molecular markers—just a matter of fashion? *Nature Reviews Genetics* 5, 63–69.
- Smulders, M.J.M., Bredemeijer, G., Ruskortekaas, W., Arens, P., Vosman, B., 1997. Use of short microsatellites from database sequences to generate polymorphisms among *Lycopersicon esculentum* cultivars and accessions of other *Lycopersicon* species. *Theoretical and Applied Genetics* 94, 264–272.
- Sneath, P.H.A., Sokal, R.R., 1973. *Numerical Taxonomy*. Freeman, San Francisco.
- Vosman, B., Arens, P., Ruskortekaas, W., Smulders, M.J.M., 1992. Identification of highly polymorphic DNA regions in tomato. *Theoretical and Applied Genetics* 85, 239–244.
- Williamson, V.M., Ho, J.Y., Wu, F.F., Miller, N., Kaloshian, I., 1994. A PCR-based marker tightly linked to the nematode resistance gene, *Mi*, in tomato. *Theoretical and Applied Genetics* 87, 757–763.
- Yang, W.C., Bai, X.D., Kabelka, E., Eaton, C., Kamoun, S., van der Knaap, E., Francis, D., 2004. Discovery of single nucleotide polymorphisms in *Lycopersicon esculentum* by computer aided analysis of Expressed Sequence Tags. *Molecular Breeding* 14, 21–34.

*Ital. J. Agron. / Riv. Agron., 2009, 3 Suppl.:93-99*

## Innovazioni genetiche per l'identificazione e la protezione di prodotti tipici italiani

Rosa Rao<sup>\*1</sup>, Martina Caramante<sup>1</sup>, Antonio Blanco<sup>2</sup>, Sergio Lanteri<sup>3</sup>,  
Margherita Lucchin<sup>4</sup>, Andrea Mazzucato<sup>5</sup>

<sup>1</sup>Dipartimento di Scienze del Suolo, della Pianta, dell'Ambiente e delle Produzioni Animali,  
Università di Napoli 'Federico II', Via Università 100, 80055 Portici (NA)

<sup>2</sup>Dipartimento di Biologia e Chimica Agro-forestale ed Ambientale, Università di Bari,  
Via Amendola 165/A, 70126 Bari

<sup>3</sup>Dipartimento di Valorizzazione e Protezione delle Risorse Agroforestali – Settore di Genetica Agraria,  
Università di Torino, Via L. Da Vinci 44, 10095 Grugliasco (TO)

<sup>4</sup>Dipartimento di Agronomia Ambientale e Produzioni Vegetali, Università di Padova, Agripolis,  
Viale Università 16, 35020 Legnaro (PD)

<sup>5</sup>Dipartimento di Agrobiologia e Agrochimica, Università della Tuscia, Via San Camillo de Lellis, 01100 Viterbo

Società Italiana di Genetica Agraria

### Riassunto

Le produzioni agricole tradizionali italiane, caratterizzate da una ricca diversità genetica, possono incrementare la competitività del comparto agricolo mediante l'ottenimento di marchi di qualità da parte della Unione Europea. La protezione dei marchi di qualità richiede strumenti appropriati per l'accertamento dell'identità genetica dei materiali vegetali pregiati sia in campo che nelle rispettive filiere agro-alimentari. Il *fingerprinting* del DNA rappresenta un eccellente metodo di identificazione di cultivars ed accessioni di specie erbacee ed arboree e, quando basato su profili allelici di specifici marcatori quali gli SSR, può tracciare ed autenticare le materie prime nelle filiere agroalimentari. Vengono descritti esempi di caratterizzazione della diversità molecolare ed identificazione varietale di specie erbacee ed arboree mediante i profili SSR, AFLP e (GATA)<sub>n</sub>. Vengono, inoltre, descritti esempi di tracciabilità genetica dei materiali vegetali delle filiere olivo, melo, frumento duro e pomodoro.

*Parole chiave:* diversità molecolare, filiere agro-alimentari, identificazione varietale, SSRs, tracciabilità.

### Summary

#### NEW GENETIC TOOLS TO IDENTIFY AND PROTECT TYPICAL ITALIAN PRODUCTS

During last decades the use of local varieties was strongly reduced due to introduction of modern cultivars characterized by higher yield, and breed for different traits of agronomic value. However, these cultivars not always have the quality aspects that was found in old traditional and typical crops also depending from the know-how of traditional cultivation. Nowadays the practise of intensive agriculture select only a small number of species and varieties with a consequent reduction of the diversity in agro-ecosystems and risk of loss of important alleles characterizing genetic materials adapted to specific environments. The creation of quality marks of the European Union proved to be a successful system to protect typical products through the Denomination of Origins (PDO- Protected Denomination of Origin and PGI- Protected Geographical Indication). However, the protection of quality needs efficient instruments to discriminate DOP or IGP varieties in the field and to trace them along the agro-food chain. DNA fingerprinting represents an excellent system to discriminate herbaceous and tree species as well as to quantify the amount of genetic variability present in germplasm collections. The paper describes several examples in which AFLPs, SSRs and minisatellite markers were successfully used to identify tomato, artichoke, grape, apple and walnut varieties proving to be effective in discriminating also closely related genetic material. DNA fingerprinting based on SSR is also a powerful tool to trace and authenticate row plant materials in agro-food chains. The paper describes examples of varieties traceability in the food chains durum wheat, olive, apple and tomato pursued through the identification of SSR allelic profiles obtained from DNA isolated from complex highly processed food, such as bread, olive oil, apple pureè and nectar and peeled tomato.

*Key-words:* agro-food chains, molecular diversity, SSRs, traceability, variety identification.

\* Autore corrispondente: tel.: +39 081 2539204; fax: +39 081 2539481. Indirizzo e-mail: rao@unina.it

Rao R., Caramante M., Blanco A., Lanteri S., Lucchin M., Mazzucato A.

### Introduzione

Numerose sono le specie orto-frutticole di pregio coltivate in Italia ed esse rappresentano una importante risorsa dell'industria agro-alimentare. Le produzioni tradizionali italiane sono caratterizzate da una ricca diversità genetica talora di pregio e la competitività del comparto agricolo italiano può essere notevolmente incrementata dalla utilizzazione di queste risorse. Negli ultimi decenni, l'utilizzo di varietà locali ha sofferto un declino dovuto alla diffusione di cultivar più produttive, omogenee e migliorate per vari caratteri agronomici importanti, ma non sempre caratterizzate da pregi di qualità tipici delle varietà locali, determinati anche dal *know-how* tradizionale dei sistemi di coltivazione. L'agricoltura intensiva attualmente praticata utilizza un limitato numero di specie e varietà, con conseguente diminuzione della diversità negli agro-ecosistemi e rischio di perdita di alleli importanti presenti in materiali genetici adattati a specifiche condizioni ambientali. La creazione di marchi di qualità da parte dell'Unione Europea si è rivelata un utile strumento per proteggere la tipicità dei prodotti alimentari e rivitalizzare alcuni comparti produttivi che includono produzioni tradizionali di pregio. La Denominazione di origine protetta (DOP) estende la tutela del marchio nazionale DOC (Denominazione di origine controllata) a tutto il territorio europeo. Il marchio si riferisce ad un prodotto originario di un paese e di una regione caratterizzato da una qualità essenzialmente o esclusivamente dovuta all'areale geografico, inclusi i fattori naturali e umani che lo caratterizzano. Il marchio IGP (Indicazione geografica protetta) introduce nella tutela della qualità l'aspetto industriale della sua produzione, dando maggiore rilevanza alle tecniche di produzione rispetto al vincolo territoriale.

La tracciabilità, divenuta dal gennaio 2005 obbligatoria per tutte le aziende agro-alimentari, è definita come "la capacità di rintracciare e seguire un alimento, un mangime, un animale produttore di alimenti o una sostanza attraverso tutti gli stadi della produzione e della distribuzione" (Reg. CE 178/2002). Il perseguimento della tracciabilità così definita può essere facilitato dalla tracciabilità genetica dei materiali vegetali che sono utilizzati nelle filiere DOP o IGP. Da qui nasce l'esigenza di avvalersi di stru-

menti di valutazione sicuri, basati su approcci inter-disciplinari che includono le scienze agro-alimentari e le biotecnologie vegetali. L'analisi del DNA mediante i marcatori molecolari rappresenta un efficiente strumento di identificazione varietale che offre una maggiore affidabilità rispetto ai metodi tradizionali basati sulla valutazione di caratteri morfologici, soggetti all'influenza delle condizioni ambientali, ed utilizzabili solo sui materiali non processati. I marcatori molecolari, invece, non subiscono interferenze da parte dell'ambiente, forniscono una misura oggettiva della variabilità essendo direttamente correlati alla molecola di DNA, e sono praticamente illimitati oltre che rilevabili con sistemi automatizzati. Inoltre, il loro polimorfismo può spesso essere rilevato sul DNA estratto da matrici alimentari complesse consentendo l'identificazione genetica delle materie prime utilizzate nelle filiere agro-alimentari (Martinez et al., 2003).

Numerosi sono i marcatori utilizzati per la caratterizzazione della diversità molecolare di specie erbacee ed arboree, mentre per la tracciabilità genetica dei materiali vegetali nelle filiere agro-alimentari i marcatori che meglio si prestano allo scopo sono gli SSR (Sequenze semplici ripetute) e gli SNP (Polimorfismi a singolo nucleotide). Entrambe queste tipologie di marcatori consentono di rilevare i polimorfismi presenti anche a partire da un DNA templato frammentato o danneggiato come spesso accade al DNA delle matrici alimentari soggette ai processi produttivi industriali. Gli SSR, marcatori di elezione per i test di paternità e studi forensi, sono tra i più utili ed affidabili marcatori molecolari del DNA in quanto possiedono elevati livelli di polimorfismo, dovuto alle variazioni del numero delle ripetizioni, sono codominanti e dispersi nel genoma; hanno il vantaggio di essere specie-specifici, essendo quindi altamente informativi e richiedono un corto DNA-stampo per l'amplificazione, un requisito cruciale per l'analisi del DNA estratto da matrici alimentari. Gli SNP, polimorfismi a singolo nucleotide generati da mutazioni puntiformi, rappresentano la forma più comune di variazioni molecolari segreganti nelle popolazioni naturali. Uno dei principali vantaggi dei marcatori SNP è che essi consentono l'analisi delle variazioni di sequenza originatesi da sostituzioni di un singolo nucleotide. Lo svantaggio di que-

sti marcatori risiede nel fatto che spesso richiedono l'utilizzo di speciali sistemi di rilevazione.

**Applicazione di tecniche di analisi molecolare per la quantificazione della variabilità genetica e la discriminazione di varietà ed ecotipi**

Le applicazioni della genetica molecolare avanzata trovano utilizzo nella costituzione varietale (miglioramento genetico e caratterizzazione varietale), nella produzione di sementi (controllo purezza varietale e genetica della produzione sementiera), nella valutazione della novità varietale (esame DUS per il registro e/o un titolo di tutela), nella difesa della tutela varietale (identificazione delle varietà). Di seguito vengono descritti alcuni esempi di discriminazione varietale e tracciabilità in filiere agro-alimentari mediante l'uso di differenti tipologie di marcatori del DNA .

*AFLP e microsatelliti per la caratterizzazione di Cynara cardunculus L.*

La specie *Cynara cardunculus* L. include il carciofo (var. *scolymus* L.), il cardo coltivato (var. *altilis* D.C.) ed il cardo selvatico [var. *sylvestris* (Lamk) Fiori]. Quest'ultimo è considerato il progenitore di entrambe le forme coltivate, che sono il risultato dell'applicazione di criteri di selezione volti ad incrementare la dimensione e produzione di capolini (carciofo) o di steli carnosì (cardo coltivato). Il carciofo rappresenta un'importante componente dell'economia agricola del Sud Europa, ma la sua coltivazione è diffusa anche in nord Africa, sud America, USA e Cina; viceversa, il cardo coltivato ha una più limitata diffusione regionale. La caratterizzazione di *landraces* reperite in orti siciliani, mediante marcatori AFLP e microsatelliti, ha evidenziato che questi materiali rappresentano forme di transizione nel processo di domesticazione della specie (Mauro et al., 2009), presumibilmente avvenuto in Sicilia. Marcatori AFLP e microsatelliti sono stati applicati per la genotipizzazione di una collezione di germoplasma di carciofo rappresentativa del materiale coltivato a livello mondiale (Lanteri et al., 2004a) e di una di collezione di cardo coltivato che include le varietà più diffuse in Italia e Spagna (Portis et al., 2005a). Mediante tali marcatori, inoltre, è stata evidenziata l'ampia variabilità genetica

presente nei tipi varietali attualmente in coltivazione (Portis et al., 2005b) ed è stato effettuato il *fingerprinting* del DNA di materiali selezionati a seguito di un programma di selezione clonale (Lanteri et al., 2004b).

*Microsatelliti per la caratterizzazione del noce di Sorrento*

La produzione italiana di noci si basa in buona parte sulla cultivar "Sorrento" varietà italiana di noce pregiata, diffusa soprattutto nell'areale campano, considerato "la terra italiana del noce" in quanto particolarmente idoneo alla coltivazione per condizioni climatiche e pedologiche. Storicamente il noce in Campania è stato allevato con duplice attitudine: produzione del legno e dei frutti. Nel corso degli anni, spesso piante con buone caratteristiche del legno sono state innestate con genotipi caratterizzati da elevata produzione e qualità dei frutti. Ciò ha determinato la coesistenza di entità genetiche differenti ed una marcata disomogeneità dei frutti e delle rese ottenute da piante diverse, ascritte alla stessa cultivar. Questo aspetto ha contribuito al forte decremento della produzione di noci registrato a partire dalla metà del secolo scorso. Allo scopo di contribuire ad una più corretta classificazione delle piante di noce diffuse sul territorio campano, è stata monitorata la diversità genetica di 42 piante assegnate alla cultivar "Sorrento" mediante l'analisi di 12 loci SSR (Feroni et al., 2005, 2007). I profili SSR analizzati identificano due distinte popolazioni "Sorrento" distribuite in due differenti areali Campani, penisola Sorrentina e areale Casertano, indicando una probabile differente origine dei materiali in propagazione. L'analisi SSR ha, inoltre, escluso l'appartenenza di alcune piante alla cultivar "Sorrento" rivelandosi un efficiente strumento per la discriminazione varietale in noce.

*Caratterizzazione genetica di una collezione di vitigni dell'area Euganea*

I vitigni autoctoni dell'area Euganea sono stati caratterizzati con lo scopo di chiarire la presenza di omonimie e sinonimie e di stabilire eventuali relazioni di parentela tra essi. La caratterizzazione varietale è stata effettuata utilizzando 30 marcatori SSR che hanno permesso di costruire le relazioni filogenetiche esistenti tra i 19 vitigni in esame (Salmaso et al., 2008).

Rao R., Caramante M., Blanco A., Lanteri S., Lucchin M., Mazzucato A.

La definizione dei profili SSR ha integrato i dati ampelografici consentendo la discriminazione di tipologie affini presenti nella collezione di germoplasma del comprensorio. La maggior parte delle cultivars considerate è risultata costituire una risorsa indipendente di variabilità genetica, anche se alcune sinonimie non sono ancora definitivamente risolte.

*Marcatori molecolari per la caratterizzazione di ecotipi e varietà di S. lycopersicum*

La variabilità genetica delle specie vegetali è fortemente dipendente dal sistema di riproduzione della specie, dalla sua storia di domesticazione e dalle dimensioni della collezione disponibile. Nel pomodoro, sin dai primi studi effettuati mediante marcatori molecolari, fu chiaro che il livello di variabilità genetica era significativamente inferiore a quello riscontrato in altre specie ad auto-impollinazione e ciò è la conseguenza della stretta autogamia della specie oltre che della diffusa utilizzazione di materiali comuni in programmi di incrocio e selezione. Un marcatore dall'elevato potere discriminante è confermato essere il minisatellite (GATA)<sub>4</sub> (Rao et al., 2006; Andreakis et al., 2004), marcatore altamente informativo, che evidenzia un profilo genetico identificativo di ogni varietà o accessione analizzata ed individua contaminazioni genetiche presenti in una stessa accessione (Caramante et al., 2009). Tra le altre classi di marcatori molecolari, gli SSR hanno mostrato elevati livelli di polimorfismo nella specie coltivata *S. lycopersicum* (Bredemeijer et al., 2002). Questi marcatori hanno inoltre il vantaggio di avere tempi di analisi inferiori a quelli richiesti dal minisatellite e di essere estendibili anche all'analisi di matrici complesse. Allo scopo di effettuare il *fingerprinting* del DNA di ecotipi e varietà tradizionali campane moltiplicate da una ditta sementiera che opera in Campania e di varietà commerciali incluse nel disciplinare di produzione integrata pomodoro da industria, anno 2007, in uso presso una ditta di trasformazione campana, sono stati selezionati 14 loci SSR particolarmente informativi della diversità intra-specifica. La discriminazione allelica è stata effettuata mediante elettroforesi capillare automatizzata su ABI Prism 3100- AVANT (Applied Biosystem). I dati ottenuti hanno mostrato che gli SSR utilizzati identificano tutti gli ecotipi e le varietà all'a-

nalisi (Rao et al., manoscritto in preparazione). Risultati simili sono stati ottenuti con il marcatore (GATA)<sub>4</sub>. Il potere discriminante degli SSR si è rivelato simile a quello del (GATA)<sub>4</sub> anche per discriminare genotipi morfologicamente simili e per individuare genotipi contaminanti presenti in alcuni materiali sementieri (Caramante et al., 2009). Questi risultati indicano che, nonostante la limitata diversità genetica presente in pomodoro, il *fingerprinting* del DNA basato sugli SSR, discrimina non solo varietà differenti ma anche tipi morfologicamente simili e può quindi essere utilizzato per la protezione di materiali geneticamente pregiati.

I marcatori SSR, inoltre, offrono un importante contributo per la quantificazione della variabilità genetica in germoplasma di pomodoro. L'analisi SSR associata alla valutazione morfologica di varietà locali di pomodoro provenienti principalmente dall'Italia centrale e di varietà moderne, ha rivelato, nelle prime, una maggiore variabilità genetica rispetto a quella presente nelle cultivar moderne (Mazzucato et al., 2008).

**Applicazione di biotecnologie genetiche per la tracciabilità e la rintracciabilità di materie prime in filiere agro-alimentari**

*Tracciabilità della filiera pomodoro*

L'utilizzo dei marcatori molecolari del DNA per il controllo della filiera agro-alimentare del pomodoro ha la potenzialità di offrire una metodica sicura e rapida per l'accertamento dell'identità genetica dei materiali presenti nei vari segmenti della filiera agro-alimentare: dal prodotto fresco al prodotto finito. A tale scopo particolarmente idonei risultano quei loci SSR che, pur mantenendo un elevato potere discriminante, siano caratterizzati da alleli di dimensioni intorno a 100 bp. Infatti, prove di amplificazione di alleli SSR a partire da DNA estratto da lavorati di pomodoro (pelati, polpa, ecc.) hanno evidenziato buoni risultati solo nel caso di alleli di piccole dimensioni. Questa è la probabile conseguenza dell'elevato grado di degradazione del DNA che si estrae dal prodotto lavorato, danneggiato dai processi produttivi. I profili allelici dei loci SSR selezionati, ottenuti mediante elettroforesi capillare degli amplificati prodotti a partire da DNA estratto da foglia,

bacca e prodotto processato di una filiera mono-varietale di pomodoro pelato, hanno mostrato completa corrispondenza (Rao et al., manoscritto in preparazione) confermando che i marcatori SSR rappresentano un potente strumento per la tracciabilità genetica dei materiali nella filiera pomodoro.

Attualmente è in corso di sperimentazione l'uso di marcatori SNP per la produzione di un microarray che utilizza sonde PNA (Peptide Nucleic Acid) finalizzato alla tracciabilità nella filiera pomodoro (Marchelli et al., dati non pubblicati).

#### *Autenticazione dell'IGP "Melannurca Campana" nella filiera melicola*

L'"Annurca" è una tipologia di mela popolare e diffusa nel sud Italia che possiede caratteristiche organolettiche superiori e riconosciute. Per difendere la produzione locale di questo prodotto e per sostenerne il valore di mercato, la "Melannurca Campana" ha ricevuto il marchio IGP ed è stata definita come la "regina delle mele". Nel lavoro di Melchiade et al., 2007 è descritta la caratterizzazione SSR di 14 varietà locali campane e 3 accessioni di "Annurca" mediante il *fingerprinting* del DNA basato su 4 loci SSR. Il prodotto IGP è identificato dal profilo allelico di un singolo locus, mentre la combinazione dei profili allelici di due loci è sufficiente per discriminare tutte le varietà all'analisi. Lo stesso locus è stato utilizzato per l'autenticazione genetica delle mele IGP in prodotti finiti differenti ed altamente processati, il nettare e la purea. La corrispondenza dei profili allelici, evidenziata mediante elettroforesi capillare degli amplificati SSR, nel DNA estratto dai materiali dei differenti segmenti della filiera indica che è possibile tracciare geneticamente la presenza della "Melannurca Campana" nei prodotti della filiera melicola confermando il marcatore SSR essere uno strumento di elezione per l'autenticazione alimentare.

#### *Olio di Collina di Brindisi: un olio extravergine d'oliva DOP*

La Puglia è la regione italiana con la più elevata produzione di olio di oliva e parte di questa produzione è rappresentata da oli che per la loro tipicità hanno ottenuto marchi di qualità DOP a livello europeo. "Collina di Brindisi" è un olio extravergine di oliva salentino che ha

ottenuto questo marchio, e il requisito varietale del disciplinare di produzione prevede l'utilizzazione della cultivar "Ogliarola" (70% minimo) e di altre cultivar diffuse nell'area di produzione (MIPAF, DM 29/9/1998). Recentemente è stata individuata la presenza di tre differenti tipi di "Ogliarola" provenienti da tre differenti territori della regione pugliese (salentina, barese e garganica) ma, considerando l'area di produzione di quest'olio DOP, la presenza della cultivar "Ogliarola salentina" è fortemente prevalente. L'identificazione della cultivar utilizzata nella produzione dell'olio DOP è stata effettuata mediante l'uso di sette marcatori SSR caratterizzati da ampliconi di piccole dimensioni, comprese tra 80 e 205 bp e, pertanto, particolarmente idonei per analizzare DNA in parte degradato come quello estratto dall'olio. Il pattern elettroforetico del DNA dell'olio DOP "Collina di Brindisi" ha evidenziato il prevalere della varietà "Ogliarola salentina" presente in quantità superiore al 70% (Pasqualone et al., 2007a). I risultati ottenuti suggeriscono che il disciplinare di produzione dell'olio DOP dovrebbe fare riferimento alla cultivar "Ogliarola salentina" e non indicare la cultivar con il nome generico di "Ogliarola". Tale indicazione è rilevante ai fini della tracciabilità delle materie prime in filiere agro-alimentari dal momento che esistono omonimi della cultivar "Ogliarola" che dovrebbero essere meglio definiti.

#### *Rintracciabilità nella filiera del grano duro e rilevazione di frumento tenero nei prodotti finiti*

La presenza di grano tenero (*Triticum aestivum* L.) nelle semole di grano duro (*Triticum turgidum* L. var. *durum*) è sempre stata oggetto di particolare interesse nella preparazione di alcuni prodotti tipici, come il "Pane di Altamura" e il "Pane di Matera" che devono essere preparati utilizzando esclusivamente semola di grano duro (European Commission L. 181 del 19/3/2003; MIPAF, DM 28/09/2004). La tipicità di questi prodotti è legata all'esclusivo uso della semola derivata da determinate cultivar di grano duro coltivate in un ben definito e ristretto areale geografico, e solo con il rispetto di tali requisiti si può accedere ai marchi europei DOP e IGP.

L'identificazione di specifiche frazioni proteiche di grano tenero ha sempre incontrato non poche difficoltà a causa della significativa denaturazione delle proteine alle alte temperatu-

Rao R., Caramante M., Blanco A., Lanteri S., Lucchin M., Mazzucato A.

re di preparazione di questi prodotti tipici. L'analisi di semole e pane con marcatori molecolari che amplificano sequenze di DNA specifiche del genoma D ha permesso di rilevare la presenza di quantità di frumento tenero del 3-5% e persino del 2,5% con la tecnica RT-PCR. Il DNA estratto con kit commerciali è stato amplificato con marcatori SSR che hanno generato un polimorfismo idoneo per discriminare il grano tenero dal grano duro; in particolare un marcatore SSR ha evidenziato la presenza di un allele tipico del genoma D ed è risultato ideale per la messa a punto di un metodo routinario di identificazione del grano tenero (Pasqualone et al., 2007b). I risultati ottenuti rappresentano una importante applicazione diretta dell'uso degli SSR per la protezione di un marchio Europeo DOP come il "Pane di Altamura" che non deve contenere farina di grano tenero.

### Conclusioni

Il *fingerprinting* del DNA basato su marcatori molecolari rappresenta un valido strumento per l'identificazione di cultivar e accessioni di specie diverse, per la quantificazione della diversità genetica di collezioni di varietà ed ecotipi e per la tracciabilità genetica delle materie prime nelle filiere agro-alimentari fornendo un'eccellente ausilio per la protezione di prodotti tipici italiani.

Questa applicazione delle moderne tecnologie genetico-molecolari rappresenta, inoltre, un importante contributo alla tutela del consumatore che, in molte parti del mondo, dispone di un reddito crescente ed è alla ricerca di cibi saporiti, tradizionali e genuini, e alla tutela del produttore attraverso la differenziazione del prodotto che può conseguire maggiore potere di mercato. D'altra parte, gli agricoltori sono costantemente alla ricerca di opportunità originali ed inesplorate per creare nuovi sbocchi di mercato. Certamente tra queste, un ruolo chiave è occupato dalla tutela delle denominazioni dei prodotti alimentari mediante marchi di qualità, opportunamente difesi nell'ambito di consorzi ed associazioni di produttori. Le continue innovazioni tecnologiche che scaturiscono dalla ricerca in genetica e genomica vegetale forniscono un contributo molto importante che, in un futuro non lontano, potrà evolversi verso la

diretta identificazione, nei segmenti della filiera, dei geni che determinano la qualità di un prodotto, ad esempio di quei geni responsabili delle qualità organolettiche o della produzione di metaboliti benefici per l'uomo.

### Ringraziamenti

Lavoro parzialmente svolto nell'ambito del progetto finanziato dal MIUR - Laboratorio pubblico privato per la genomica del pomodoro, GenoPOM, DM17732.

### Bibliografia

- Andreakis N., Giordano I., Pentangelo A., Fogliano V., Graziani G., Monti L. M., Rao R. 2004. DNA fingerprinting and quality traits of corbarino cherry-like tomato landraces. *J. Agric. Food Chem.*, 52:3366-3371.
- Bredemeijer G.M.M., Cooke R.J., Ganai M.W., Peeters R., Isaac P., Noordijk Y., Rendel S., Jackson J., Roder M. S., Wendehake K., Dijcks M., Amelaine M., Wickaert V., Bertrand L., Vosman B. 2002. Construction and testing of a microsatellite database containing more than 500 tomato varieties. *Theor. Appl. Genet.*, 105:1019-1026.
- Caramante M., Rao R., Monti L.M., Corrado G. 2009. Discrimination of 'San Marzano' accessions: A comparison of minisatellite, CAPS and SSR markers in relation to morphological traits. *Scientia Horticulturae*, 120:560-564.
- Faroni I., Rao R., Woeste K. 2005. Characterization of *Juglans regia* L through SSR markers and evaluation of genetic relationships among cultivars and the 'Sorrento landrace'. *J. Hort. Sci. Biotech.*, 80:49-53.
- Faroni I., Woeste K., Monti L.M., Rao R. 2007. Identification of 'Sorrento' Walnut Using Simple Sequence Repeats (SSRs). *Genetic. Resour. Crop Evol.*, 54:1081-1094.
- Lanteri S., Saba E., Cadinu M., Mallica G., Baghino L., Portis E. 2004a. Amplified fragment length polymorphism for genetic diversity assessment in globe artichoke. *Theor. Appl. Genet.*, 108:1534-1544.
- Lanteri S., Acquadro A., Saba E., Portis E. 2004b. Molecular fingerprinting and evaluation of genetic distances among selected clones of globe artichoke (*Cynara scolymus* var. *cardunculus* L.) 'Spinoso sardo'. *J. Hort. Sci. Biotech.*, 79:863-870.
- Martinez I., Aursand M., Erikson U., Singstad T.E., Vellyulin E., van der Zwaag C. 2003. Destructive and non destructive analytical techniques for authentication and composition analyses of foodstuffs. *Trends Food Sci. Technol.*, 14:489-498.

*Ital. J. Agron. / Riv. Agron., 2009, 3 Suppl.:93-99*

- Mauro R., Portis E., Acquadro A., Lombardo S., Mauromicale G., Lanteri S. 2009. Genetic diversity of globe artichoke landraces from Sicilian small-holdings: implications for evolution and domestication of the species. *Conserv. Genet.*, 10:431-440.
- Mazzucato A., Papa R., Bitocchi E., Mosconi P., Nanni L., Negri V., Picarella M.E., Siligato F., Soressi G.P., Tiranti B., Veronesi F. 2008. Genetic diversity, structure and marker-trait associations in a collection of Italian tomato (*Solanum lycopersicum* L.) landraces. *Theor. Appl. Genet.*, 116:657-669.
- Melchiade D., Foroni I., Corrado G., Rao R. 2007. Authentication of the 'Annurca' apple in agro-food chain by amplification of microsatellite loci. *Food Biotechnol.*, 21:33-43.
- Pasqualone A., Montemurro C., Summo C., Sabetta W., Caponio F., Blanco A. 2007a. Effectiveness of microsatellite DNA markers in checking the identity of protected designation of origin extra virgin olive oil. *J. Agric. Food Chem.*, 55:3857-3862.
- Pasqualone A., Montemurro C., Grinn-Gofron A., Sonnante G., Blanco A. 2007b. Detection of soft wheat in semolina and durum wheat bread by analysis of DNA microsatellites. *J. Agric. Food Chem.*, 55:3312-3318.
- Portis E., Barchi L., Acquadro A., Macua J.I., Lanteri S. 2005a. Genetic diversity assessment in cultivated cardoon by AFLP (amplified fragment length polymorphism) and microsatellite markers. *Plant Breed.*, 124:299-304.
- Portis E., Mauromicale G., Barchi L., Mauro R., Lanteri S. 2005b. Population structure and genetic variation in autochthonous globe artichoke germplasm from Sicily Island. *Plant Sci.* 168:1591-1598.
- Rao R., Corrado G., Bianchi M., Di Mauro A. 2006. (GATA)(4) DNA fingerprinting identifies morphologically characterized 'San Marzano' tomato plants. *Plant Breed.*, 125:173-176.
- Salmaso M., Valle R.D., Lucchin M. 2008. Genepool variation and phylogenetic relationships of an indigenous north-east Italian grapevine collection revealed by nuclear and chloroplast SSRs. *Genome*, 51, 10:838-855, doi:10.1139/G08-064.

*Proceedings of the 52nd Italian Society of Agricultural Genetics Annual Congress*  
*Padova, Italy – 14/17 September, 2008*  
ISBN 978-88-900622-8-5

**Poster Abstract – A.54**

---

**VNTR, SSR AND MORPHOLOGICAL CHARACTERIZATION OF  
TOMATO ACCESSIONS SPREAD IN CAMPANIA REGION**

CARAMANTE M., RAO R., MONTI L., CORRADO G.

Department of Soil, Plant, Environmental and Animal Production Sciences, School of  
Biotechnology, University of Naples “Federico II”, Via Università 100, 80055 Portici (Italy)

*Lycopersicon esculentum, GATA, SSR, microsatellites, genetic diversity*

In this study it has been evaluated the ability of two DNA molecular markers to discriminate a set of 16 tomato accessions. Among these accessions, some could not be distinguished on the basis of morphological traits. VNTR and SSR markers were used as a tool to characterise a tomato germplasm collection representative of Italian tomato types that are traditionally cultivated in Campania. The germplasm included material for fresh market and food-processing purposes. VNTR marker revealed that some accessions were not genetically uniform; (GATA)<sub>4</sub> fingerprinting clearly allowed the distinction of contaminating or segregating genotypes, which show different hybridization patterns. The UPGMA hierarchical classification based on a data set of 14 SSR appropriately selected, confirmed the differences observed through VNTR analysis. The genetic diversity identified in the 16 accessions revealed a consistent polymorphism at the analyzed loci, suitable to discriminate all the genotypes. Although both markers effectively discriminated the analysed samples, SSR resulted also suitable for the genetic traceability of tomato varieties along the agro-food chain.

---

*Proceedings of the 52nd Italian Society of Agricultural Genetics Annual Congress*  
*Padova, Italy – 14/17 September, 2008*  
ISBN 978-88-900622-8-5

**Poster Abstract – A.55**

---

**A COMPARISON OF MINISATELLITE AND SSR MARKERS IN  
RELATION TO MORPHOLOGICAL TRAITS FOR DISCRIMINATION  
OF ‘SAN MARZANO’ ACCESSION**

CARAMANTE M., RAO R., CORRADO G.

Department of Soil, Plant, Environmental and Animal Production Sciences, School of  
Biotechnology, University of Naples “Federico II”, Via Università 100, 80055 Portici (Italy)

*tomato, VNTR, SSR, microsatellites, genetic diversity*

‘San Marzano’ is one of the most widely known tomato (*Lycopersicon esculentum* Mill.) cultivars, and is a classic example of a local variety with a premium value. Unfortunately, the original cultivated form is underrepresented in the Protected Denomination of Origin (PDO) area because of the incidence of contaminant morphologically similar genotypes. Our aim was to investigate the ability of three DNA marker systems (minisatellite, CAPS and SSR) to reveal the genetic diversity of tomato accessions that, as indicated by a morphological analysis, are phenotypically very close. The data indicate that CAPSs are the least effective, whereas both minisatellites and SSRs can be used to genetically distinguish the analysed materials and depict relationships consistent with the hierarchal pattern obtained by the morphological data. As locally cultivated tomato accessions are often characterised by some degree of genetic variability, our results will be valuable in facilitating the purification, management and breeding of tomato germplasm. The differences between the marker systems employed are discussed in relation to their usefulness in the agro-food chain.

**DIVERSITA' GENETICA DI ACCESSIONI DI POMODORO CAMPANE**

G. CORRADO\*, D. SCARANO, M. CARAMANTE

Facoltà di Scienze Biotecnologiche, Università degli Studi di Napoli "Federico II"

*Lycopersicum, variabilità, AMOVA*

La competitività, la sostenibilità e la tipicità dell'agricoltura italiana dipende anche dal continuo accesso a nuovo germoplasma, utile al *breeder* per selezionare o migliorare nuove varietà. Nessuna nazione al mondo è autosufficiente da questo punto di vista, ma le varietà tradizionali presenti in ogni regione rappresentano una risorsa troppo spesso inesplorata. Tale materiale possiede il vantaggio di essere già adattato ad un agro-sistema ed ameno a "*low-input practices*", e solitamente presenta caratteristiche qualitative del prodotto edule superiori e facilmente riconoscibili dai consumatori. In questa prospettiva, il pomodoro rappresenta un caso esemplare. La specie è caratterizzata da una ristretta diversità genetica, a causa della sua recente domesticazione e diffusione, ed il panorama varietale è dominato da ibridi esteri che hanno progressivamente scalzato le accessioni locali. La variabilità genetica degli ibridi moderni è limitata e principalmente localizzata nelle regioni genomiche fiancheggianti i geni introgressi da specie selvatiche. Queste ultime infatti, sono la fonte più ricca di geni di resistenza a stress biotici del pomodoro. L'obiettivo di questo lavoro è stato quello di caratterizzare la variabilità genetica di accessioni tradizionali di pomodoro della Campania. Lo studio è stato condotto tramite l'analisi del polimorfismo di 14 loci SSR (*Simple Sequence Repeats*) in elettroforesi capillare. Il confronto tra 15 accessioni Campane e 17 varietà commerciali (*inbred* ed ibridi) tramite analisi gerarchiche, multivariata (PCA) e della varianza molecolare (AMOVA) ha indicato che le accessioni campane sono una popolazione geneticamente distinta dalle comuni varietà coltivate di pomodoro. Tali accessioni devono essere considerate quindi un materiale da proteggere e migliorare che si differenzia significativamente dalle varietà commerciali, ma al tempo stesso anche una probabile risorsa di caratteri di qualità ed adattabilità a specifici ambienti.

**Area tematica:** Biotecnologie, Genetica