Theor Appl Genet (1997) 95:375–383

F. Geuna · H. Hartings · A. Scienza

Discrimination between cultivars of *Vitis vinifera* based on molecular variability concerning 5′ untranslated regions of the *StSy-CHS* genes

Received: 11 April 1996 / Accepted: 14 March 1997

Abstract The degree of polymorphism present in 5′ untranslated regions of stilbene synthase (*StSy*)-like loci was assessed. A ligation-mediated polymerase chain reaction (LM-PCR) cloning strategy was adopted to isolate sequences located immediately upstream of StSy coding regions. Among several clones, 13 randomly chosen fragments were analyzed at the sequence level. Four of the analyzed fragments appeared of particular interest. Two carried sequences reminiscent of micro-satellites, while the remaining fragments contained direct repeats. Oligonucleotides constructed against the specific DNA sequence of these clones disclosed a complex banding pattern when used in polymerase chain reaction (PCR)-analysis of 22 ancient varieties of grapevine. A total of 40 polymorphic bands could be identified and used to calculate coefficients of genetic similarity (GS) between varieties. GS values were used in cluster analysis to differentiate the 22 varieties. The data obtained are in good agreement with available information concerning the relationships between the varieties considered. This suggests the use of the method we have developed in fingerprinting studies of *Vitis vinifera* germ plasma.

Key words Cluster analysis · Genetic variability · Ligation-mediated PCR · Stilbene synthase · *Vitis vinifera*

Communicated by F. Salamini

F. Geuna · A. Scienza (✉)
Istituto di Coltivazioni Arboree, University of Milano, via Celoria 2, 20133 Milano, Italy

H. Hartings
Istituto Sperimentale per la Cerealicoltura, Sezione di Bergamo, via Stezzano 24, 24126 Bergamo, Italy

Introduction

A number of criteria have been used to characterize plant germ plasm-like morphological traits, as well as the nature and extent of secondary compounds and isoenzymes (Tanksley and Orton 1983; Wendel and Weeden 1989). The recent introduction of restriction fragment length of polymorphisms (RFLPs) has facilitated the characterization of genetic diversity at the DNA level. In addition, studies on the complexity of the non-coding regions of the nuclear eukaryotic genomes have made possible the discovery of several classes of genomic repetitive sequences. This knowledge has led to the development of molecular markers that exploit differences in the number and extension of mini- and micro-satellites (Tautz and Renz 1984; Jeffreys et al. 1985; Litt and Luty 1989). PCR technology has also contributed to this field providing RAPDs (random amplified polymorphic DNA) and STS (sequence-tagged site) markers based, respectively, on the amplification of random and of known genomic sequences (Whitkus et al. 1994).

In the genus *Vitis* the above-mentioned molecular tools have been employed for cultivar identification and genome analysis. Microsatellites have been identified for use in fingerprinting and in assisted selection (Thomas and Scott 1993; Cipriani et al. 1994). Polymorphisms mapping near coding regions of the grapevine genome have only been found associated with the intergenic spacer (IGS) of the ribosomal gene cluster (Thomas et al. 1993).

Stilbene synthase (STSy) synthesizes the backbone of the stilbene phytoalexins which have antifungal properties and play a role in the mechanism of plant defense against pathogens (see Schröder et al. 1993 for a review). This enzyme has been reported for peanut (*Arachis hypogaea*; Schröder et al. 1988), pine (*Pinus silvestris*; Fliegmann et al. 1992) and grapevine (*Vitis vinifera*; Melchior and Kindl 1991). The presence of

StSy in unrelated species is thought to be due to independent mutation events of the chalcone synthase (*CHS*) genes which encode a protein catalyzing a reaction similar to the one promoted by StSy. In fact *StSy* genes exhibit a high degree of sequence homology with members of the *CHS* multigene family (Tropf et al. 1994). In grapevine, stilbene synthase is coded by a multigene family of which seven different members have been cloned (Melchior and Kindl 1991; Wiese et al. 1994). In the present paper we describe the existence of polymorphisms for *V. vinifera* which are closely associated with the coding regions of genetic loci belonging to a multigene family. The cloning and sequence analysis of genomic DNA located immediately upstream of StSy-CHS coding regions has allowed the construction of specific oligonucleotides. Their use in PCR analysis on genomic DNA from 22 ancient grapevine cultivars has revealed a high level of polymorphism sufficient for their classification.

Materials and methods

Plant material and DNA extraction

Plant material was obtained from the *Vitis* collection of the University of Milano and comprised 22 ancient varieties, once cultivated in Lombardy, Northern Italy: Thompson seedless, Bianca botticino, Bianca gabana, Brugnera, Corva, Erbamat, Gentile, Groppello Mocasina, Groppellone, Groppello S. Stefano, Invernesca, Marzemino bianco, Maiolina, Moscatello, Malvasia, Negrara, Orsanella, NN (a non-identified variety similar to Orsanella according to ampelographic assays), Pozzolenga, Trebbiano, Villa, and Uva uccellina. The variety "Lambrusco a foglia frastagliata" was used as a DNA source for cloning the *StSy* genes. Cuttings were grown on agrilite, leaves were cut, frozen in liquid nitrogen, and stored at $-80^{\circ}C$ for subsequent DNA extraction. DNA was extracted from three young leaves of each cultivar as described (Steenkamp et al. 1994).

PCR amplification and molecular cloning

PCR amplifications were performed in a reaction volume of 20 µl containing 10 mM Tris.HCl pH 8.8; 1.5 mM $MgCl_2$; 50 mM KCl; 0.1% Triton X-100; 200 mM each of dATP, dTTP, dCTP, and dGTP; 100 ng of template DNA; 1 unit of *Taq* DNA polymerase (Finnzymes Inc). Primers, used alone or in combination, were added at a concentration of 0.5 mM. Oligonucleotide sequences are summarized in Table 1. Reaction mixtures were overlaid with mineral oil and amplification was obtained through 35 cycles with the following temperature profile: 94°C for 30 s, 50°C for 30 s and 72°C for 60 s, followed by a final hold of 6 min at 72°C. When necessary, amplification products were cloned into pGEM5-T vector DNA (Promega, Madison Wis.) following manufacturer's instructions and subjected to direct sequence analysis.

Ligation-mediated PCR cloning

The LM-PCR technique, as described by Fors et al. (1990), was used to clone promoter regions of members of the *StSy-CHS* multigene family. Complementary oligonucleotides A5 and A3S were used to create a common adaptor featuring a sticky end and a blunt end which was used for ligation. The gene-specific oligonucleotide SS3 was used for extension, while oligonucleotide SS1 was used in a nested PCR amplification. Sequence of both oligonucleotides were based on a published *StSy* cDNA sequence (Sparvoli et al. 1994).

Gel electrophoresis and Southern blotting

Aliquots of the amplification products were resolved on 5% polyacrylamide gels made with a modified matrix (Long Ranger, J.T. Baker, Holland), which can effectively separate a wide range of molecular-weight DNA molecules. Gels contained 1.2 × TBE buffer (pH 8.3) and were run at 9 V/cm for 3–4 h with a 0.6 × TBE buffer to enhance resolution.

Alternatively, amplification products were run on 1.6% agarose gels at 5 V/cm for 6 h. When required, agarose gels were blotted against Hybond-N^+ nylon membranes (Amersham), as described by the manufacturer, and hybridized with random labelled probes (Feinberg and Vogelstein 1983).

Cloning amplification products homologous to the *StSy* gene

Genomic DNA of *V. vinifera* cv "Lambrusco a foglia frastagliata" was subjected to PCR amplification using primer combinations SS10/SS3, SS13/SS3, SS16/SS3 and SS19/SS3 with the conditions described above. Amplification fragments were shot-gun cloned into

Table 1 Sequences (5′ to 3′) of oligonucleotides used in this study

Primer	Sequence	Position[a]	Orientation[b]
SS1	GAAGCCATTGATCCCAGC	+8	Antisense
SS3	CCCTTGGCACGTTGAGCG	+44	Antisense
SS10	GCTGCTATGTTTGATCCC	−137	Sense
SS13	GGTCTAAGGAGTACAGAG	−220	Sense
SS16	AGCTTGTCTTGGAGGAAG	−386	Sense
SS19	CACTGCACTAGTTGAGTC	−294	Sense
A5[c]	CAGGTACCTCGAGATCTGCATGC	n.a.	n.a.
A3S[c]	CTAGGCATGCAGATCTCGAGGTACCTG	n.a.	n.a.

[a] Position of the first nucleotide of the primer relative to the ATG start codon of the published *StSy* cDNA sequence (Sparvoli et al. 1994)
[b] Direction of elongation relative to the orientation of the published *StSy* cDNA sequence (Sparvoli et al. 1994)
[c] Complementary oligonucleotides employed to form an adaptor (n.a. = not applicable)

the pGEM-T vector (Promega, Madison Wis.) following the manufacturer's instructions.

Sequence analysis

Nucleotide sequences were determined on double-stranded templates. Direct sequencing of amplified genomic fragments was performed with the use of the Applied Biosystems "*Taq* Dyedeoxy Terminator Cycle Sequence" kit. Sequences were analyzed on an Applied Biosystems model 373A automatic sequencer. In all cases sequences were determined on both strands.

Statistical analysis

Coefficients of genetic similarity (GS) between pairs of cultivars were calculated according to the following equation (Dice 1945; Nei and Li 1979): $GS = 2N_{ij}/(N_i + N_j)$, where GS is the similarity coefficient between cultivars i and j. N_{ij} reflects the number of common bands for i and j, while N_i and N_j reflect the total number of bands detected in cultivar i and j respectively. Hence, identity of two cultivars will result in a GS value of 1, while totally unrelated cultivars will give rise to a GS value of zero. Average linkage cluster analysis (UPGMA), performed on GS values, was used to evaluate possible associations among cultivars.

For comparative purposes, principal component analysis was performed on a [1–0] matrix representing the presence or absence of each polymorphic fragment for each cultivar. Calculations of principal components were based on the co-variance matrix. All necessary computations were performed with the use of the "STATISTICA" software (StatSoft, Inc., 1995).

Results

Isolation of putative *StSy* 5′ untranslated sequences

To isolate 5′ untranslated regions of the *StSy-CHS* genes, an LM-PCR cloning strategy was adopted based on genomic DNA from the variety "Lambrusco a foglia frastagliata". The use of the oligonucleotides SS1 and SS3, complementary to sequences present on a *StSy* cDNA (Sparvoli et al. 1994) and oriented towards the 5′ end of the gene, in combination with a common adaptor, allowed the isolation of 20 independent PCR products, ranging in size from 145 bp to 500 bp.

Sequence analysis of 13 randomly chosen fragments confirmed the presence of both the SS1 and the adaptor sequence in all clones considered. Furthermore, alignment of sequences allowed the identification of homologous (70–90% identity) DNA stretches among the 13 fragments (Fig. 1), while the remaining parts of the DNA sequences were strongly polymorphic. One fragment was identical to the 5′ untranslated region of the *StSy* cDNA clone on which the internal primers used in LM-PCR cloning were designed. Since all fragments carried sequences homologous to the 5′ end of the *StSy* cDNA, and furthermore shared sequence homology, we concluded that these DNA stretches contained 5′ untranslated sequences related to the *StSy-CHS* genes. A high degree of homology was assessed at the SS1 and adaptor sequences.

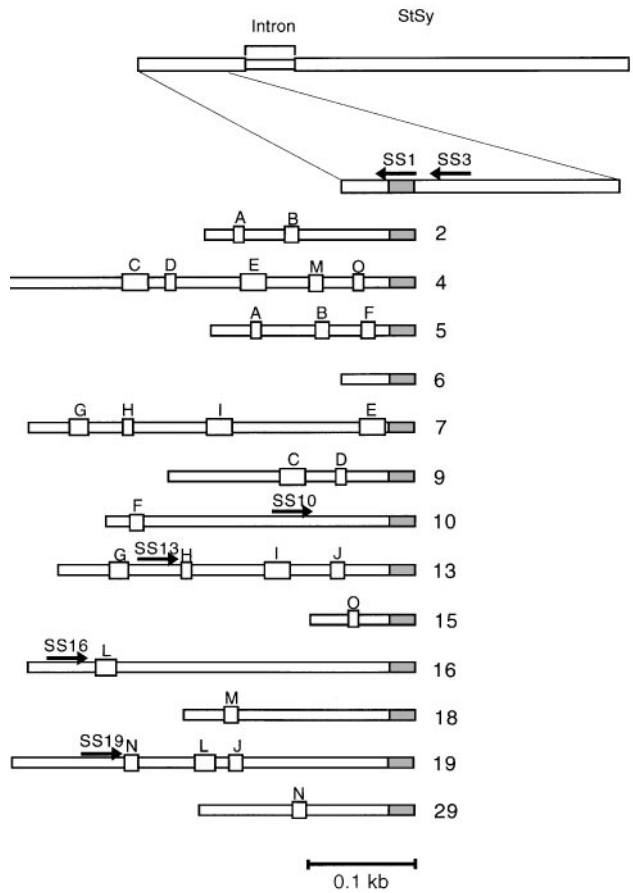

Fig. 1 Sequence alignment of genomic clones obtained by LM-PCR cloning. Regions of sequence similarity between genomic clones are shown. A schematic map of the *StSy* gene and its position relative to the 5′ cloned regions is depicted: the *gray box* marks the 20-bp region spanning the translation start codon of the gene. *Arrows* indicate the position and orientation of the oligonucleotides employed for PCR amplifications

StSy upstream regions disclose a high level of polymorphism

Among the 13 genomic fragments analyzed, four appeared of particular interest (#10, #13, #16 and #19 in Fig. 1). Two fragments (#10 and #16) carried sequences reminiscent of micro-satellites, while fragments #13 and #19 contained direct repeats up to 13 bp in size. The level of polymorphism in the 5′ untranslated regions of the *StSy-CHS* genes was measured using oligonucleotides complementary to the sequence of each of these four cloned fragments. PCR amplification of genomic DNA from Lambrusco with external primers derived from fragments #10, #13 and #16 (SS10, SS13 and SS16 in Table 1), used in combination with the internal primer SS1, revealed a complex banding pattern, while a primer derived from fragment #19 (SS19 in Table 1) revealed a single amplification product. Also PCR reactions based on single primers

produced multiple amplification fragments, suggesting the presence in the genome of several sequences with a sufficient homology to the primers used (Fig. 2).

The reproducibility of amplification results, assayed in multiple runs of identical reactions, was demonstrated by the absence of lot-to-lot amplification differences. Moreover, results varied only marginally with changed template, primer and dNTP concentration (data not shown). Also, the number of amplified fragments could be modulated by modifying the annealing temperature, which in our experiments was kept at 50°C. Amplifications based on the oligonucleotide combination SS10/SS1 demonstrated that an annealing temperature of 52°C reduced the number of bands by approximately 25%, while raising the temperature to 55°C reduced the total number of bands by approximately 40% (data not shown).

StSy-dependent polymorphism existing among grapevine cultivars

The high number of amplification products obtained with one combination of oligonucleotides, or even with a single external primer, prompted us to explore the potential informative content of this phenomenon and its possible exploitation in determining relationships between grapevine cultivars. Twenty two different ancient grapevine varieties, including the cultivar Thompson seedless, which is thought to be only distantly related to the others, were considered. DNAs were subjected to PCR amplification with oligonucleotides SS10, SS13, and SS1. PCR products were separated on 5% Long Ranger polyacrylamide gels as described in Materials and methods.

Primer combinations SS10/SS1 and SS13/SS1 produced a complex amplification pattern for all the varieties tested, with fragments ranging in size from 130 bp to 3000 bp (Fig. 3). The 22 cultivars disclosed a total of

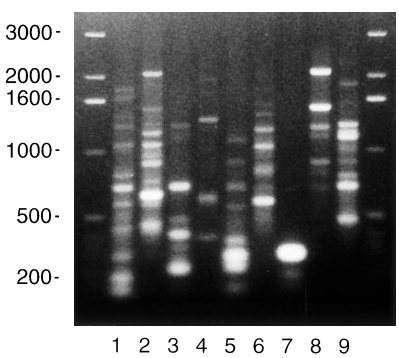

Fig. 2 Amplification products obtained from genomic DNA of Lambrusco with oligonucleotides: *1* SS10 + SS1, *2* SS10, *3* SS13 + SS1, *4* SS13, *5* SS16 + SS1, *6* SS16, *7* SS19 + SS1, *8* SS19, *9* SS1. Molecular weights (bp) are indicated

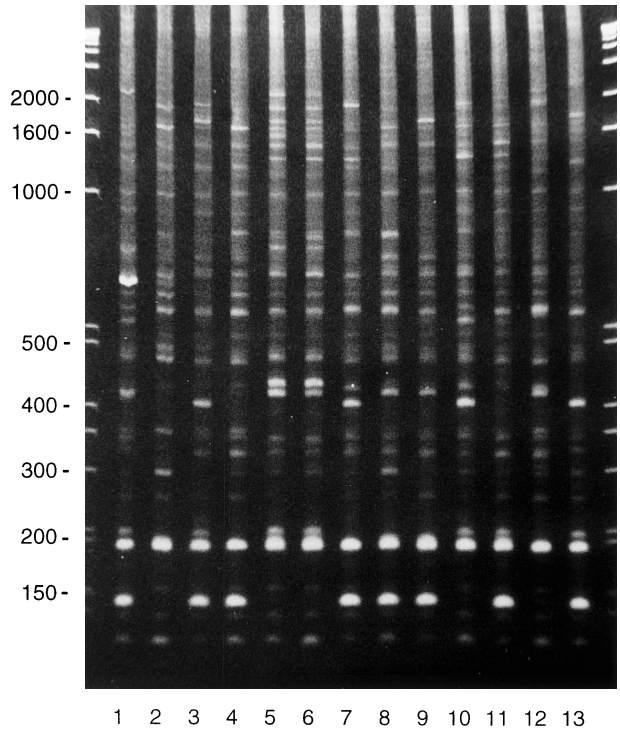

Fig. 3 Amplification products of 13 *Vitis vinifera* cultivars. DNA was amplified with oligonucleotides SS10 + SS1 and the reaction products separated on a 5% polyacrilamide gel as described in Materials and methods. *1* Armenia, *2* Brugnera, *3* Groppello gentile, *4* Marzemino bianco, *5* Moscatello, *6* Malvasia, *7* Negrara, *8* NN, *9* Orsanella, *10* Pozzolenga, *11* Trebbiano Valtenesi, *12* Villa, *13* Uva Uccellina. Molecular markers are run in the flanking lanes. Molecular weights are indicated in bp

11 polymorphic bands upon analysis with the combination SS10/SS1. Among the fragments revealed, two low-molecular-weight bands of major intensity were present. A first band, approximately 200 bp in length, was invariably conserved in all the varieties tested, while a second band of 145 bp was present in 17 out of 22 varieties. Among the amplification products obtained with combination SS13/SS1, four bands of major intensity were detected, one of which, approximately 440 bp in length, was invariably present in all varieties. This combination of oligonucleotides disclosed a total of 13 polymorphic fragments.

Amplification of genomic DNA with primers SS10 and SS1 alone produced a set of bands with molecular weights ranging in size from 500 bp to 2000 bp. The number of polymorphic bands obtained with these primers was six and ten, respectively.

To confirm the relatedness of these fragments with regions coding for the *StSy* genes, PCR products of Lambrusco genomic DNA obtained with primer combinations SS10/SS3, SS13/SS3, SS16/SS3 and SS19/SS3 were hybridized with a *StSy* cDNA probe. Hybridization signals ranging in size from 100 bp up to 2000 bp were detected for all the combinations (Fig. 4).

Products generated with the SS16/SS3 primer combination were cloned and characterized at the sequence level. All fragments showed homology with the *StSy* coding and upstream region (data not shown); one of them, named SS16/3#8, of 278 bp appeared of particular interest since a BLAST search against the GenBank nucleotide database also evidenced a significant homology with the untranslated sequences and promoters of two published stilbene synthase genes (*Vst1* and *Vst2*, Viese et al. 1994; see Fig. 5).

Cluster analysis

A total of 40 polymorphic bands generated by oligonucleotide combinations SS10/SS1 and SS13/SS1, or by SS10, SS13 and SS1 alone, were considered. The presence/absence of amplification fragments in the genotypes considered, gave rise to a matrix used to calculate the coefficient of genetic similarity (GS) among cultivars (Table 2). *Vitis* cultivars revealed an average GS value of 0.634. Individual values calculated on pairs of cultivars ranged from 0.38 (Malvasia–Uva uccellina) to 1.00 (Groppello gentile–Groppellone and Groppello Mocasina–Groppello S. Stefano). Estimates of similarity between cultivars were converted to measures of dissimilarity by the formula $(1 - S_{ij})$. These new values were the elements of a 22×22 proximity matrix which served as an input to cluster analysis. The average Unweighted Pair Group Method Using Arithmetic Average (UPGMA) linkage method was adopted to classify the 22 varieties. Among them, Moscatello ad Malvasia were far more similar to each other than to the other varieties and formed a distinct cluster. Cultivar "Thompson seedless' also showed a low similarity with other cultivars, while the remaining varieties formed two major subgroups (Fig. 6). The results of cluster analysis were supported by the principal component analysis (PCA). The first principal component explained 18% of the variation among varieties and differentiated Moscatello and Malvasia varieties from the remaining cultivars, Approximately 16% of the variation was explained by the second principal component, which separated the same two groups of cultivars, already identified by cluster analysis. The third principal component, which explained 11.5% of total variation, differentiated "Thompson seedless" from other cultivars.

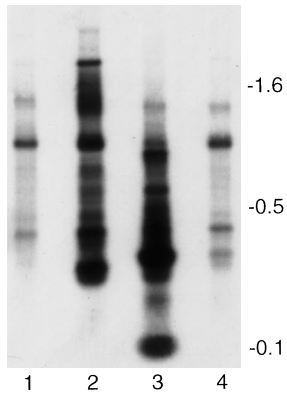

Fig. 4 Southern blot of amplification products obtained from genomic DNA of Lambrusco with different primer combinations. A full-length *StSy* cDNA was used as a molecular probe. *1* SS10/SS3, *2* SS13/SS3, *3* SS16/SS3, *4* SS19/SS3. Molecular weights are indicated in kbp

Fig. 5 Alignment of *StSy* promoters. Promoters of two published *StSy* genes (*Vst1* and *Vst2*, Viese et al. 1994), the untranslated leader sequences of cDNA clones SV21 and SV25 (Melchior and Kindl 1991), and the sequence of clone SS16/3#8 (this paper) are shown. Regions of homology between two or more sequences are shown in *upper case*. TATA-boxes are *underlined*. The determined transcription start of *Vst1* is indicated with an *asterisk*

```
Vst1          attgctctctccttcctgtgggttgcattcatcatatatgggtaggtgaagaacagagag    60
Vst2          cttccaacaaccttacgtgggctaaagacaaaaggattagagttatagagaaagatctaa    60

Vst1          caaaacaaatacattaacacatgcatacatgtctgaaaatccaaaaataaaatcacgcac   120
Vst2          agaatgaaaaccgaagcctaccaagttggaaatgatcaaaacccatgtcactgcagcatg   120
SS16/3#8.      ------------------------------------------------cttgtcttggaggaag    16

Vst1          acaaactttgaagccacctgat---------------------------------------   142
Vst2          gattattcaaatgcacgcctacctacatgcaagtccctgccctttcatgtatcatcatca   180
SS16/3#8      gaggctgaagggccgatgtcatacaaaattgattactttctcagtaatcccttccacgtg    76

Vst1          ------CATTGACTGccga---------------------------TGGATGAGAGTTGG   169
Vst2          agtatcCATTGACTGggcagacaaatcaatttcaatagcgtccaagTGGATGAGAGTTGG   240
SS16/3#8      actcaaatcagcacggaagtactgaacaagaccccaccaccagctacaagtgggttggaa   136

Vst1          TGAGACACAGCtag-------CTTATAAATACCCAACACTCACACCCAGCTTTCTCAAGC   222
Vst2          TGAAACACAGCATTCCTTATGATTATAAATACCaacctcaag-ACACAACTTTTTCCAGC   299
SV25          ------------------------------------------------AGCTTTCTCAAGC    13
SS16/3#8      gtggACACAGCATTTCTGATACTTATAAATACCCGACGCTCACACACTGCTTTGGCAACC   196

                                              *
Vst1          CAGCTCCAAGC--------------------ACTCTTCTCTTTCCTTCCTCAAtc------   257
Vst2          CAGCTCCAAGCACTctgtgctccgaaaCATTCACTTCTCTTTCCTTCTTCAACTT-----   354
SV25          CAACTCCAAGCACT----------------TGAGTTCTCTTTTCTTCCTCAActtaaTC-    56
SV21          ----------------------------cTGAGTTCTCTTTCCTTCCTTCA-----ACt    26
SS16/3#8      CATCTTCAAGCACTagtccg-------CATACTCTTCTCTTTTCTTCTACCGCTT-----   244

Vst1          -----TTCAGCTTCAATTTGAGttcgtagctag----GATCAATG                 293
Vst2          -----TTAAGCTTCGATTTCAGcagtcatcctact---ATCAATG                 391
SV25          -----TTAAGCTTTAtcaATTTCATTACGTATCTagccATCCATG                  96
SV21          taatcTTAGGCTTTA---ATTTGAGTACGTA-GCTggGATCAATG                  67
SS16/3#8      -----TGAAGCTctgtgggggattttactagtacat--A-CAATG                 278
```

Table 2 Coefficients of genetic similarity determined on pairs of cultivars according to Dice (1945). 1–Armenia, 2–Bianca botticino, 3–Bianca gabana, 4–Brugnera, 5–Corva, 6–Erbamat, 7–Gentile, 8–Groppello mocasina, 9–Groppellone, 10–Groppello S. Stefano, 11–Invernesca, 12–Marzemino, 13–Maiolina, 14–Moscatello, 15–Malvasia, 16–Negrara, 17–NN, 18–Orsanella, 19–Pozzolenga, 20–Trebbiano Valtenesi, 21–Villa, 22–Uva uccellina

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1	1.0																					
2	0.69	1.0																				
3	0.67	0.97	1.0																			
4	0.47	0.59	0.63	1.0																		
5	0.63	0.74	0.77	0.55	1.0																	
6	0.59	0.63	0.67	0.74	0.68	1.0																
7	0.53	0.63	0.62	0.55	0.68	0.72	1.0															
8	0.61	0.78	0.81	0.74	0.76	0.80	0.81	1.0														
9	0.53	0.63	0.62	0.55	0.68	0.72	1.0	0.81	1.0													
10	0.61	0.78	0.81	0.74	0.76	0.80	0.81	1.0	0.81	1.0												
11	0.59	0.53	0.57	0.72	0.55	0.74	0.65	0.74	0.65	0.74	1.0											
12	0.51	0.57	0.61	0.59	0.59	0.64	0.63	0.72	0.63	0.72	0.70	1.0										
13	0.53	0.65	0.69	0.39	0.70	0.60	0.70	0.68	0.70	0.68	0.44	0.59	1.0									
14	0.53	0.41	0.46	0.39	0.50	0.47	0.55	0.58	0.55	0.58	0.50	0.49	0.50	1.0								
15	0.50	0.39	0.43	0.42	0.48	0.53	0.62	0.60	0.62	0.60	0.53	0.46	0.53	0.95	1.0							
16	0.63	0.74	0.78	0.70	0.73	0.73	0.73	0.77	0.73	0.77	0.65	0.63	0.65	0.43	0.46	1.0						
17	0.54	0.43	0.47	0.67	0.51	0.70	0.56	0.68	0.56	0.68	0.72	0.70	0.51	0.51	0.54	0.55	1.0					
18	0.63	0.53	0.56	0.65	0.59	0.68	0.64	0.71	0.64	0.71	0.80	0.73	0.55	0.55	0.57	0.63	0.88	1.0				
19	0.55	0.60	0.63	0.76	0.65	0.73	0.65	0.77	0.65	0.77	0.62	0.70	0.57	0.52	0.55	0.65	0.62	0.61	1.0			
20	0.65	0.70	0.74	0.62	0.74	0.78	0.65	0.73	0.65	0.73	0.82	0.75	0.67	0.51	0.49	0.70	0.62	0.74	0.62	1.0		
21	0.65	0.60	0.63	0.62	0.65	0.69	0.65	0.73	0.65	0.73	0.57	0.56	0.57	0.62	0.64	0.60	0.67	0.78	0.67	0.67	1.0	
22	0.58	0.63	0.67	0.65	0.64	0.68	0.55	0.67	0.55	0.67	0.65	0.68	0.60	0.40	0.38	0.68	0.56	0.55	0.70	0.70	0.52	1.0

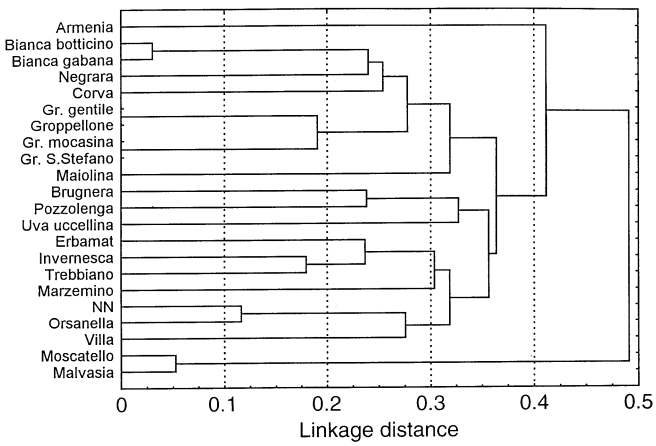

Fig. 6 Phylogenetic tree of 22 *V. vinifera* cultivars. GS coefficients, as summarized in Table 2, were used in cluster analysis based on the UPGMA method to derive the tree reproduced in this figure

Discussion

StSy loci are highly polymorphic in their 5′ untranslated regions

Sequences located upstream of the coding region of the *StSy* and *CHS* genes were isolated and used to assess the degree of polymorphism present in these regions. All recovered fragments revealed similarity to the 5′ end of a cloned *StSy* cDNA, supporting the conclusion that they represent DNA stretches located immediately upstream of *StSy* or *CHS* coding regions. The isolation of a genomic fragment identical in sequence to the 5′ region of a *StSy* cDNA (Sparvoli et al. 1994) further corroborated this hypothesis. The isolated fragments exposed a high degree of polymorphism and, interestingly, the presence of structured repeated sequences, some of which were reminiscent of micro-satellites. PCR analysis of genomic grapevine DNA with oligonucleotides spanning these repeated sequences disclosed a complex number of amplification products. Hybridization with a molecular probe consisting of a full-length *StSy* cDNA resulted in numerous hybridizing fragments, suggesting that these amplification products all comprise the 5′ part of the *StSy-CHS* gene-coding regions. Again, this hypothesis was corroborated by sequence analysis, demonstrating that the hydribizing fragments disclose sequence homology with the *StSy* coding region.

It has been shown that members of multigene families evolved in a concerted manner (Smith 1979; Coen et al. 1982; Arnheim 1983; Matsuo and Yamazaki 1989). Due, most probably, to the need of conservation of gene function, however, concerted evolution applies particularly to coding regions and not to flanking regions of multigene-family components (Hickey et al. 1991; Shibata and Yamazaki 1995). Our data predict the presence in the *StSy-CHS* multigene family of highly polymorphic upstream regions containing repeated DNA stretches. The *StSy* coding regions, on the other hand, appear to exhibit a high level of homology. These findings are in good agreement with concerted evolution described for members of multigene families.

Variability of DNA sequences upstream of coding regions allows one to infer phylogenetic relationships

The high number of amplification products obtained with a single combination of oligonucleotides suggested an investigation of the variability among different cultivars of grapevine of upstream regions associated with the *StSy-CHS* multigene family. Cluster analysis, based on a matrix of dissimilarity coefficients derived from PCR amplifications, revealed the phylogenetic relationships among 22 grapevine cultivars. This classification was, moreover, supported by the results of principle component analysis.

The phylogenetic tree reported in Fig. 5 is supported by relevant, although limited, historical information concerning the 22 varieties considered. Malvasia and Moscatello exhibited a similarity index of 0.95 while being different from all other varieties. This mainly reflects their eastern origin (Levadoux 1956). Considering its origin, the cultivar Thompson seedless was also expected to be clustered independently from the other cultivars (Levadoux 1956). This hypothesis was confirmed by cluster analysis. It is furthermore worth noting that the molecular similarity between Corva and Negrara may reflect an identical origin. In fact, according to information dating back to the 19th century (Acerbi 1825), the two names should be considered as synonyms of the same cultivar. Particularly interesting is the close grouping of all Groppello cultivars: Gentile, Mocasina, Groppellone and S. Stefano. The absence of polymorphism between Groppellone and Groppello gentile is consistent with available information indicating that both names refer to the same variety (Villa 1995).

StSy-promoter PCR as a molecular tool

The use of oligonucleotides homologous to upstream regions of loci belonging to the *StSy-CHS* multigene family has been shown to be a powerful means to investigate the polymorphism between grapevine cultivars by PCR. This method has several advantages over established PCR techniques such as RAPDs and STSs. Because this method is based on a standard PCR amplification, its reproducbility is extremely high. Relevant changes in amplification patterns were not observed, even with changed template, dNTP, and primer concentrations.

Although the number of amplified bands varied according to the external primer selected, a reasonable number of polymorphic fragments could be found with three out of the four primers tested, showing that only a limited effort was necessary to identify oligonucleotides suitable for genotypic fingerprinting. As a matter of fact, the number of polymorphic bands generated by amplification runs based on two external oligonucleotides and one internal oligonucleotide was sufficient to differentiate the 22 *V. vinifera* cultivars.

It was interesting to note that the number of amplified fragments could be modulated by modifying the annealing temperature. An increase in the annealing temperature of 2°C and 5°C reduced the total number of amplified bands by approximately 25% and 40%, respectively. Therefore, identification of polymorphic bands among complex patterns can be facilitated by properly modulating the annealing temperature. Optimization of results should not be difficult, considering both the abundance of polymorphic fragments revealed by a single primer combination and the relatively large temperature shift which can be applied without the loss of information.

The use of *StSy* loci and of the polymorphism they disclose in assisted-selection techniques can be predicted. Stilbene synthase is an enzyme directly involved in the mechanism of resistance of grapevine to fungal pathogens (Kindl 1985 and references therein; Hain et al. 1993). Thus the markers we have developed could be employed in programs of marker-assisted selection for tolerance to biotic stress. Detailed information concerning the regulated induction to *StSy* genes and their response to different types of biotic stress might however be, necessary before engaging in such a delicate type of analysis. Research addressing these topics is in progress and should further elucidate the role that *StSy* genes play in plant resistance to fungal attack.

Acknowledgements The authors are grateful to Prof. F. Salamini for his critical review of the manuscript and many helpful suggestions. The research work was supported by the Italian Ministry of Agriculture in the framework of the project "Resistenze genetiche delle piante agrarie agli stress biotici e abiotici" and by the coordinated project CNR-MARCFRU. F.G. wishes to thank Prof. M. Motto for kind hospitality at the Istituto Sperimentale per la Cerealicoltura.

References

Acerbi G (1825) Delle viti italiane o sia dei materiali per servire alla classificazione. Monografia e sinonimia precedute dal tentativo di una classificazione geografica delle viti. Milano

Arnheim N (1983) Concerted evolution of multigene families. In: Nei M, Koehn K (eds) Evolution of genes and proteins. Sinauer Associates, Sunderland, Massachusetts, pp 38–61

Cipriani G, Frazza G, Peterlunger E, Testolin R (1994) Grapevine fingerprinting using microsatellite repeats. Vitis 33:211–215

Coen ET, Strachan T, Dover G (1982) Dynamics of concerted evolution of ribosomal DNA and histone gene families in the *melanogaster* species subgroup of *Drosophila*. J Mol Biol 158: 17–35

Dice LR (1945) Measures of the amount of ecologic association between species. Ecology 26:297–302

Feinberg AP, Vogelstein B (1983) A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. Anal Biochem 132:6–13

Fliegmann J, Schröder G, Schanz S, Britsch L, Schröder J (1992) Molecular analysis of chalcone and dihydropinosylvin synthase from Scots pine (*Pinus sylvestris*), and differential regulation of these and related enzyme activities in stressed plants. Plant Mol Biol 18:489–503

Fors L, Saavendra RA, Hood L (1990) Cloning of the shark Po promoter using a genomic-walking technique based on the polymerase chain reaction. Nucleic Acids Res 18: 2793–2799

Hain R, Reif HJ, Krause E, Langebartels R, Kindl H, Vornam B, Wiese W, Schmelzer E, Schreier PH, Stocker RH, Stenzel K (1993) Disease resistance results from foreign phytoalexin expression in a novel plant. Nature 361:153–156

Hickey DA, Bally-Cuif L, Bukashawa S, Payant V, Benkel BF (1991) Concerted evolution of duplicated protein-coding genes in *Drosophila*. Proc Natl Acad Sci USA 88:1611–1615

Jeffreys AJ, Wilson V, Thein SL (1985) Hypervariable 'minisatellite' regions in human DNA. Nature 314:67–73

Kindl H (1985) Biosynthesis of Stilbenes. In: Higuchi T (ed) Biosynthesis and biodegradation of wood components. Academic Press, Orlando London, pp 349–378

Levadoux L (1956) Les populations sauvages et cultivées de *Vitis vinifera*. Ann Amelior Plant 1:59–118

Litt M, Luty J (1989) A variable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac-muscle actin gene. Am J Hum Genet 44:397–401

Matsuo Y, Yamazaki T (1989) Nucleotide variation and divergence in the histone multigene family in *Drosophila melanogaster*. Genetics 122:87–97

Melchior F, Kindl H (1991) Coordinate and elicitor-dependent expression of stilbene synthase and phenylalanine ammonia-lyase genes in *Vitis* cv Optima. Arch Biochem Biophys 288: 552–557

Nei M, Li WH (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc Natl Acad Sci USA 76:5269–5273

Schröder G, Brown JWS, Schröder J (1988) Molecular analysis of resveratrol synthase: cDNA, genomic clones and relationship with chalcone synthase. Eur J Biochem 172:161–169

Schröder J, Schanz S, Tropf S, Kärcher B, Schröder G (1993) Phytoalexin biosynthesis: stilbene synthase and co-action of a reductase with chalcone synthase. In: Fritig B, Legrand M (eds) Mechanism of plant defense response. Kluver Academic Publishers, Dordrecht, The Netherlands, pp 257–267

Shibata H, Yamazaki T (1995) Molecular evolution of the *Amy* locus in the *Drosophila melanogaster* species subgroup: concerted evolution only in the coding region and an excess of non-synonymous substitutions in speciation. Genetics 141: 223–236

Smith GP (1979) Evolution of repeated DNA sequences by unequal crossover. Science 191:528–535

Sparvoli F, Martin C, Scienza A, Gavazzi G, Tonelli T (1994) Cloning and molecular analysis of structural genes involved in flavonoid and stilbene biosynthesis in grape (*Vitis vinifera L.*). Plant Mol Biol 24:743–755

Steenkamp J, Wiid I, Lourens A, Van Helden P (1994) Improved method for DNA extraction from *Vitis vinifera*. Am J Enol Vitic 45:102–106

Tanksley SD, Orton TJ (1983) Isozymes in plant genetics and breeding, Part A. Elsevier, Amsterdam

Tautz D, Renz M (1984) Simple sequences are ubiquitous repetitive components of eukaryotic genomes. Nucleic Acids Res 12: 4127–4138

Thomas MR, Scott NS (1993) Microsatellite repeats in grapevine reveal DNA polymorphism when analysed as sequence-tagged sites (STSs). Theor Appl Genet 86:985–990

Thomas MR, Matsumoto S, Cain P, Scott NS (1993) Repetitive DNA of grapevine: classes present and sequences suitable for cultivar identification. Theor Appl Genet 86:173–180

Tropf S, Lanz SA, Rensing J, Schröder J, Schröder G (1994) Evidence that stilbene synthases have developed from chalcone synthases several times in the course of evolution. J Mol Evol 38:610–618

Villa PL (1995) Lo studio e la selezione clonale del Groppello presso il Centro Vitivinicolo Provinciale. Rassegna Vitic 5:14–16

Wendel JF, Weeden NF (1989) Visualization and interpretation of plant isozymes. In: Soltis DE, Soltis PS (eds) Isozymes in plant biology. Dioscorides Press, Portland, Oregon, pp 5–45

Whitkus R, Doebley J, Wendel JF (1994) Nuclear DNA markers in systematics and evolution. In: Phillips RL, Vasil IK (eds) DNA-based markers in plants. Kluver Academic Publishers, The Netherlands, pp 116–141

Wiese W, Vornam B, Krause E, Kindl H (1994) Structural organization and differential expression of three stilbene synthase genes located on a 13-kb grapevine DNA fragment. Plant Mol Biol 26:667–677