Mining and validating grape (Vitis L.) ESTs to develop EST-SSR markers for genotyping and mapping

Hong Huang · Jiang Lu · Zhongbo Ren · Wayne Hunter · Scot E. Dowd · Phat Dang

Abstract Grape expressed sequence tags (ESTs) are a new resource for developing simple sequence repeat (SSR) functional markers for genotyping and genetic mapping. An integrated pipeline including several computational tools for SSR identification and functional annotation was developed to identify 6,447 EST-SSR sequences from a total collection of 215,609 grape ESTs retrieved from NCBI. The 6,447 EST-SSRs were further reduced to 1,701 non-redundant sequences via clustering analysis, and 1,037 of them were successfully designed with primer pairs flanking the SSR motifs. From them, 150 pairs of primers were randomly selected for PCR amplification, polymorphism and heterozygosity analysis in V. vinifera cvs. Riesling and Cabernet Sauvignon, and V. rotundifolia (muscadine grape) cvs. Summit and Noble, and 145 pairs of these primers yielded PCR products. Pairwise comparisons of loci between the parents Riesling and Cabernet Sauvignon showed that 72 were homozygous in both cultivars, while 70 loci were heterozygous in at least one cultivar of the two. Muscadine parents Noble and Summit had 90 homozygous SSR loci in both parents and contained

Electronic supplementary material The online version of this article (doi:10.1007/s11032-010-9477-2) contains supplementary material, which is available to authorized users.
50 heterozygous loci in at least one of the two. These EST-SSR functional markers are a useful addition for grape genotyping and genome mapping.

**Keywords** SSR · EST · Marker · Genotyping · Grape

**Introduction**

Microsatellites, or simple sequence repeats (SSRs), are short (1–6 bp) repeat DNA motifs that are usually single locus markers with characteristics of hypervariability, abundance and reproducibility. The variation of the SSR repeat units can be easily differentiated by PCR products amplified with primers flanking the SSR motif. SSRs have been widely used for bacteria screening (Lin et al. 2005), plant genotyping (Chen et al. 2006), linkage mapping (Zhang et al. 2002), gene tagging (Roy et al. 2002), and map-based gene cloning (Tekeoglu et al. 2002).

The availability of ESTs greatly accelerates the systematic identification of SSRs and corresponding marker development based on computer analytical approaches (Varshney et al. 2002; Gao et al. 2003; Thiel et al. 2003; Chen et al. 2006). EST-derived SSRs have been well documented in some plant species including Arabidopsis (Depeiges et al. 1995), sugarcane (Cordeiro et al. 2001), cereal species (Kantety et al. 2002), cacao (Lima et al. 2008), and rubber tree (Feng et al. 2009). Using homology searches, putative functions can be deduced for the SSRs and thereby provide a new resource that can further aid in genetic and evolutionary studies (Cho et al. 2000; De Keyser et al. 2009).

EST-SSR and genomic SSR markers should be considered as complementary to plant genome mapping, with EST-SSR being less polymorphic but concentrated in the gene-rich regions (Varshney et al. 2006). With hundreds of thousands of ESTs available in the public domain, the process of developing EST-SSR markers has been greatly accelerated by using optimized computational pipelines and high-throughput genotyping techniques.

SSR markers have been widely used in grape genotyping. The high polymorphism of Vitis-derived microsatellite loci has been reported extensively in the literature and used for fingerprinting (Thomas and Scott 1993; Bowers et al. 1996, 1999a, b; Seif et al. 1998; Arroyo García et al. 2002; Di Gaspero et al. 2005, 2007; Merdinoğlu et al. 2005; Lamoureux et al. 2006; Costantini et al. 2007; De Mattia et al. 2007; Cipriani et al. 2008; Bocharova et al. 2009; Riaz et al. 2009). Several publications have also demonstrated transferability of SSR markers across the Vitis genus (Lin and Walker 1998; Tessier et al. 1999; Di Gaspero et al. 2000; Fernández et al. 2008).

SSR markers have been used for construction of grape genetic maps (Dalbò et al. 2000; Doligez et al. 2002; Grando et al. 2003; Adam-Blondon et al. 2004; Doucleff et al. 2004; Riaz et al. 2004, 2006; Lowe and Walker 2006; Di Gaspero et al. 2007; Vezzulli et al. 2008). While the majority of the loci in grape linkage maps are microsatellite markers developed from genomic DNA libraries, the availability of EST-SSRs will serve as new genetic markers to be included into the linkage map (Decroocq et al. 2003; Akkak et al. 2006; Salmaso et al. 2008). EST-SSRs have been reported to be less polymorphic but to have higher transferability than genomic SSRs in grape and other plants because of greater DNA sequence conservation in transcribed regions (Scott et al. 2000; Cho et al. 2000; Chabane et al. 2005).

Traditionally, SSR PCR products are separated by polyacrylamide gels (Thiel et al. 2003) or Metaphor Agarose gels (Chani et al. 2002). The electrophoresis-based technology is low-throughput and comes with imprecise sizing at times. Automatic capillary sequencing using fluorescently-labeled primers (Eujayl et al. 2002) provides more accurate and high-throughput genotyping results but the cost of dye-labeling each forward primer is high. Using M13 universal labeled primers with automatic capillary sequencing can not only reduce the cost but also provide fast and precise genotyping results (Oetting et al. 1995, Chen et al. 2006).

Here we report the identification and characterization of 1,701 unique grape EST-SSRs derived from a total of 215,609 grape ESTs. A set of SSR markers was developed from this analysis and validated by using M13 universal primers and an automatic capillary sequencing system.
Materials and methods

Plant materials

For PCR amplification, genotyping and polymorphism analysis, we selected four genotypes which are parents of two mapping populations: *V. vinifera* cvs. Cabernet Sauvignon and Riesling, and *V. rotundifolia* cvs. Noble and Summit. Genomic DNA was extracted from young leaves/shoot tips of these grape cultivars using a modified CTAB protocol (Qu et al. 1996).

Grape EST and genomic sequences retrieval from NCBI

All grape EST sequences available in the NCBI database on 10 February 2006 were retrieved. Among the total of 215,609 ESTs, 194,200 were from *V. vinifera*, 10,704 were *V. shuttleworthii*, 2,177 were *V. aestivalis*, 1,995 were *V. riparia*, and 6,533 were *V. hybrids* (*V. rupestris* A. de Serres 9 *V. spp. b42-26*).

A total of 31,910 genomics sequences were also retrieved from NCBI on 12 June 06. Among them, 30,832 were from a BAC library of *V. vinifera* cv. Pinot Noir, and 1,078 from *V. vinifera* cvs. Syrah and Maxxa.

Computer programs for mining SSRs from ESTs

A Perl script program named Microsatellite (MISA) developed by Thiel et al. (2003, http://pgrc.ipk-gatersleben.de/misa) was used to identify EST-SSRs. The SSRs are between 2 and 6 nucleotides in size. The minimal length of SSR repeats was defined as $2 \times 9 = 18$ bp for dinucleotides, $3 \times 6 = 18$ bp for trinucleotides, $4 \times 5 = 20$ bp for tetranucleotides, $5 \times 4 = 20$ bp for pentanucleotides, and $6 \times 4 = 24$ bp for hexanucleotides. ESTs containing SSRs were assembled in Sequencher® version 4.2 (Gene Codes, Ann Arbor, Michigan, USA) under criteria of 40% minimum overlap and 90% minimum match percentage. A flow chart for mining and developing the grape EST-SSR markers is provided in Fig. 1.

Functional annotation of EST-SSRs

HTGOFAT, a data mining and annotation tool kit developed in Microsoft NET 2003, was utilized to functionally annotate the assembled EST-SSRs sequences (Dowd and Zaragoza 2005). The putative functional genes were classified using the Munich Information Centre for Protein Sequences (MIPS) *Arabidopsis thaliana* functional catalogue (MATDB, http://mips.gsf.de).

PCR and fragment analysis

EST sequences flanking the microsatellite motifs were used to design PCR primers using the program Primer3®. A total of 150 primer pairs (Table 3) were screened for assessment of polymorphisms among the four parents using a CEQ Genetic Analyzer (Beckman Coulter, California, USA).

To save cost, a 20-bp long universal M13 forward primer sequence GTT GTAAAA CGA CGG CCA GT (Oetting et al. 1995) was added as a common tail to the 5’ end of all 180 SSR forward primers. All SSR primers, including regular and M13-tailed forward primers, were synthesized by Operon Technologies (Huntsville, Alabama, USA). The universal M13 primers were labeled by Sigma-Genosys (USA) and used for CEQ Genetic Analyzer Fragment Analysis.

PCR reactions were performed in a 20-μl reaction mix including 30 ng of genomic DNA, 10 × PCR buffer (Promega), 2 μl of 2 mM dNTP (Promega), 1.0 U Taq DNA polymerase, 2.8 μl of 25 mM MgCl2 (Promega), and 0.3 μM primers. The PCR reactions were carried out in a PTC-200 thermal cycler (MJ Research) with the following thermal profile: 3 min at 94°C followed by 30 cycles of 1 min denaturation at 94°C, 1 min annealing at 48 to 58°C (based on the $T_m$ of the different primer sets), and 2 min extension at 72°C, followed by a final step of 6 min extension at 72°C. The same conditions were also used for labeling the primers.

Fig. 1 Flow chart of *Vitis* EST-SSR identification and validation
For fragment analysis using the CEQ Genetic Analyzer, 0.25 µl of each M13 labeled PCR product was mixed with 40 µl Sample Loading Solution (Beckman Coulter 608087) with 0.2 µl 400-bp DNA size standard (Beckman Coulter 608098) and overlaid with one drop of light mineral oil, then loaded into the 96-well sample microtiter plates (Beckman Coulter 609801). CEQ Sequencing Separation Buffer (Beckman Coulter 608012) were also loaded into the 96-well separation plate (Beckman Coulter 609844). Dye-labeled amplicons were automatically sized by running on “Frag-3” separation and the GenomeLab software (Beckman Coulter) and then visually examined.

Results and discussion

Identification and characterization of grape EST-SSRs

A total of 6,447 out of 215,609 (3%) grape ESTs retrieved from NCBI on 1 February 2006 contained SSRs (Table 1). With some of them having multiple SSR sites, a total of 6,815 SSR motifs were identified among these 6,447 EST sequences. The percentage of EST-SSRs varied slightly among different Vitis species, ranging from 2.98% for V. vinifera (5,782 of 194,200), 3.50% for V. aestivalis (74 of 2,116), 3.55% for V. shuttleworthii (389 of 10,933), to 5.43% for V. riparia (59 out of 1,087). The EST-SSRs accounted for 2.71% for a Vitis hybrid of (V. rupestris A. de Serres × V. spp b42-26) (177 of 6,533; Electronic Supplementary Material 1).

Among the redundant EST-derived SSR repeats, tri-nucleotide, which accounted for 50.2% of total SSRs, was the most abundant repeat unit followed by di (28.5%), hexa (11.4%), penta (9.6%), and tetranucleotide (6.1%; Table 1). These findings are in agreement with previous observations on abundance of SSR repeat units in barley, maize, rice, sorghum, and wheat (Kantety et al. 2002). The dominance of trinucleotide SSRs was viewed as the result of a frame shift in size of one amino acid read, or three nucleotides, a selection against possible frame shift mutations (Metzgar et al. 2000; Toth et al. 2000; Wren et al. 2000; Cordeiro et al. 2001). For the same reason, a higher percentage was also observed in hexanucleotide SSRs than tetra- and penta-repeats. In both non-redundant and redundant EST-SSRs (Table 1), di- and tri-repeats were accounted for about 80% of the total EST-SSRs for each group (redundant: di-28.5%, tri-50.2%; non-redundant: di-39%, tri-41.7%). Interestingly, the proportion of tri repeats dropped from 50.2% in redundant to 41.7% in non-redundant ESTs while di repeats increased from 28.5 to 39.0% after eliminating the redundancy by contig assembling (Table 1). The result was interpreted to suggest that tri-repeat SSRs were mainly found in coding regions (Yu et al. 2004) and many of these redundant EST-SSRs were eliminated because these sequences contain tri-repeats representing putative amino acid runs (Li et al. 2004) as overexpressed ESTs representing the same set of genes. Another

| Redundant SSR-ESTs | Non-redundant SSR-ESTs | Genomic SSR sequences |
|--------------------|------------------------|----------------------|
| **Total %** | 6,447 | 1,701 | 1,346 |
| Di- | 1,835 (28.5%) | 664 (39.0%) | 699 (51.9%) |
| Tri- | 3,235 (50.2%) | 710 (41.7%) | 341 (25.3%) |
| Tetra- | 391 (6.1%) | 125 (7.3%) | 154 (25.2%) |
| Penta- | 618 (9.6%) | 134 (7.9%) | 99 (7.4%) |
| Hexa- | 736 (11.4%) | 179 (10.5%) | 53 (3.9%) |
| **Abundant type** | | | |
| Di- | AG/CT (18.9%) | AG/CT (26.9%) | AT/AT (33.0%) |
| Tri- | AGG/CCT (14.8%) | AAG/CTT (11.3%) | AAT/ATT (18.6%) |
| Tetra- | AAGG/CCTT (1.7%) | AAAT/ATTT (2.4%) | AAAT/ATTT (7.0%) |
| Penta- | AAAAT/ATTTT (2.8%) | AAAAT/ATTTT (2.1%) | AAAAT/ATTTT (3.8%) |
| Hexa- | AGGGTC/AGTCCC (2.6%) | ACCCTG/ACTGGG (1.1%) | AAAAT/TAAAAA (1.7%) |
explanation is the effect of gene duplication and paralogy. Depending on the parameters used for clustering, untranslated regions of paralogous genes, which are more divergent and contain all types of SSR, might have remained separated, while ESTs covering exons of paralogous genes, which are more conserved and highly enriched in tri-nucleotide SSR, might have collapsed more frequently into a “single” redundant EST.

Comparison between genomic and EST derived SSRs

Unlike the tri-nucleotide repeats as the dominant type in SSR-ESTs, the number of genome sequence-derived SSRs were dominated by di-nucleotide repeats that accounted for 51.9% of total genomic SSRs, followed by tri- (25.3%), tetra- (25.2%), penta- (7.4%), and hexa-SSRs (3.9%; Table 1). Similar patterns for EST-SSRs having a higher proportion in tri-repeats than genomic SSRs were reported in the literature (Cardle et al. 2000). Among the top 20 SSRs in ranking, the most abundant di-nucleotide repeat in non-redundant ESTs was AG/CT which accounted for 17.9% of total EST-SSRs, followed by AT/AT (8.4%; Table 2), while the most abundant di-nucleotide repeat in genomic sequences was AT/AT which accounted for 33.0%, followed by AG/CT (15.5%) and AC/GT (3.5%). The most common EST-derived tri-nucleotide repeat was AAG/CTT (14.0%), while AAT/ATT (18.6%) was the most abundant tri-nucleotide SSRs derived from genomic sequences. Among grape genomic sequences, around 67.1% of the SSRs belonged to three types of repeats: AT/AT (33.0%), AAT/ATT (18.6%), and AG/CT (15.5%; Tables 2 and 3).

Functional analysis of EST-SSR sequences

The 1,701 assembled non-redundant EST-SSRs were functionally annotated using the HTGOFAT program (Dowd and Zaragoza 2005). Fifty-eight percent (994

| Table 2 | Top 20 SSR motifs in grape ESTs and genomic sequences |
|---------|---------------------------------|
| EST-SSR repeats | Total | % | Genomic SSR repeats | Total | % |
| AG/CT | 1,217 | 17.9 | AT/AT | 444 | 33.0 |
| AGG/CCT | 957 | 14.0 | AAT/ATT | 251 | 18.6 |
| AT/AT | 575 | 8.4 | AG/CT | 208 | 15.5 |
| AAG/CTT | 547 | 8.0 | AAAAT/ATTT | 94 | 7.0 |
| AAT/ATT | 403 | 5.9 | AAAAT/ATTTT | 51 | 3.8 |
| ACC/GGT | 326 | 4.8 | AAG/CTT | 48 | 3.6 |
| AGC/CGT | 237 | 3.5 | AC/GT | 47 | 3.5 |
| ACG/CTG | 201 | 2.9 | AAAAAT/ATTTTT | 21 | 1.6 |
| AGT/ATC | 201 | 2.9 | AAAAG/CTTTT | 17 | 1.3 |
| AAAAT/ATTTT | 195 | 2.9 | AAAAAG/CTTTTTT | 15 | 1.1 |
| ACT/ATG | 168 | 2.5 | AATT/AAAT | 14 | 1.0 |
| AAC/GTT | 163 | 2.4 | AAAAG/CTTT | 13 | 1.0 |
| AGGGTC/AGTCCC | 144 | 2.1 | AGT/ATC | 12 | 0.9 |
| AAAAG/CTTTTT | 98 | 1.4 | ACAT/ATGT | 10 | 0.7 |
| AAAAC/GTTTT | 97 | 1.4 | ACT/ATG | 9 | 0.7 |
| AAAT/ATTT | 89 | 1.3 | AGAT/ATCT | 8 | 0.6 |
| AAGG/CCTT | 88 | 1.3 | AAC/GTT | 7 | 0.5 |
| AAAGG/CCTTT | 88 | 1.3 | AAAAC/GTTTT | 6 | 0.4 |
| ACCCTG/ACTGGG | 81 | 1.2 | AAAAT/AAATTT | 6 | 0.4 |
| AAAAAG/CTTTTT | 78 | 1.1 | AGG/CCT | 5 | 0.4 |
| Other motifs | 862 | 12.6 | Other motifs | 60 | 4.5 |
| Total | 6,815 | 100.0 | 1,346 | 100.0 |
| Marker ID | Accession no. | Repeat type | Forward primer | Reverse primer | Expected size (bp) | Riesling allele | Cabernet allele | Summit allele | Noble allele | R × C | S × N |
|-----------|---------------|-------------|----------------|----------------|-------------------|----------------|----------------|---------------|-------------|-------|-------|
| FAM01     | CA817092      | (ga)9       | TTACCGACACTGGACAC | ACTTACCAACGAGATGAGG | 310 967/983 976/983 989 989 | ab × ab       |                 |               |             |       |       |
| FAM02     | CB973719      | (tc)10      | GCCCTGACCAACGATATG | CTAAGAAACACCATTCATCAG | 199 218/236 208/216 218 218 | ab × cd       |                 |               |             |       |       |
| FAM03     | CF371950      | (tca)6      | CACCAAGAAACGACAGA | CACCAAGAAACGACAGA | 232 239/266 243/251 262 252/280 | ab × cd       | aa × bc       |               |             |       |       |
| FAM04     | CF405747      | (ctt)10     | TGACCTACAACTCTTTAAC | AAGGAGAGAGAGAGAGAGAGAGAGAG | 180 198 196 189/191 189/191 | aa × bb       | ab × ab       |               |             |       |       |
| FAM05     | CB035928      | (ta)6       | ATTCACATCTGCTACATAA | CCACCTCCTACACATACAC | 331 470 470/476 458 454 | aa × ab       | aa × ab       |               |             |       |       |
| FAM06     | CB915165      | (aag)11     | AGATAAGACGACAGATGAA | CAAACAATCCTACAACCAAC | 296 313 313 304/335 316/322 | ab × cd       | ab × ab       |               |             |       |       |
| FAM07     | CV092730      | (ctt)6      | ACTGTCTCCTCAATCTCATCA | CTCAGACAGGTAGAATAATA | 236 360/369 369 362 362/365 | ab × aa       | aa × ab       |               |             |       |       |
| FAM08     | CV093018      | (tcg)6      | TGTGAATTTCCAATGAT | AAGTGAGAGAGAGAGAGAGA | 180 198 196 189/191 189/191 | ab × ab       | ab × ab       |               |             |       |       |
| FAM09     | CA814604      | (aat)13     | TGAAGCACTGATGCTTATTG | ACAATGTCACACACAAGTG | 126 117 120/147 147 120/147 | aa × bc       | ab × ab       |               |             |       |       |
| FAM10     | CB974681      | (cag)6      | CTTTTTAAAACATGGATCTA | CGAATAACAATCTGGCAT | 325 333/342 342 330 330 | ab × aa       | aa × ab       |               |             |       |       |
| FAM11     | CB348477      | (ctt)8      | TGGTAAATTTCCAAATGAT | TTTGGAGAGTAGGATGA | 113 123 125/131 133/139 NA | ab × bc       | ab × ab       |               |             |       |       |
| FAM12     | CB346585      | (cis)6      | GTTACGAAACGAGGTGGA | CGAATAACAATCTGGCAT | 236 470 470/476 458 454 | aa × ab       | ab × ab       |               |             |       |       |
| FAM13     | CB974681      | (cag)6      | CTTTTTAAAACATGGATCTA | CGAATAACAATCTGGCAT | 286 317 317 317 317 | ab × ac       | ab × ab       |               |             |       |       |
| FAM14     | CT026156      | (aaaat)5    | AGATCCTCCGAAACAAACTT | CAAAGATCAAGGAGAAACTGC | 371 388/392 NA 276 276 | ab × cc       | ab × ab       |               |             |       |       |
| FAM15     | CF512631      | (tatg)7    | TGGCCTTATATGCAGTTTCT | AGGCTCACTGAGATGAGC | 271 281/288 281/288 276 276 | ab × cd       | ab × ab       |               |             |       |       |
| FAM16     | CF515612      | (gttt)5     | CTCTCCACATTACGTGTT | AGGCTCACTGAGATGAGC | 286 317 317 317 317 | ab × ac       | ab × ab       |               |             |       |       |
| FAM17     | CF516198      | (tc)21      | TGGCCTTATATGCAGTTTCT | AGGCTCACTGAGATGAGC | 325 333/342 342 330 330 | ab × aa       | aa × ab       |               |             |       |       |
| FAM18     | BM437681      | (att)6      | TATAGTGGTCAATGCAACCA | GGTGAGTCCACATGGTAAAT | 117 135 135 133 133 | ab × ab       | ab × ab       |               |             |       |       |
| FAM19     | BM437023      | (cac)7      | AAACTGGACTCCACTGT | GGTGAGTGGCAGCTTCTAGTGCT | 211 227 221/227 213 213 | aa × ab       | aa × ab       |               |             |       |       |
| FAM20     | BM437023      | (cac)7      | AAACTGGACTCCACTGT | GGTGAGTGGCAGCTTCTAGTGCT | 211 227 221/227 213 213 | aa × ab       | aa × ab       |               |             |       |       |
| FAM21     | BM437023      | (cac)7      | AAACTGGACTCCACTGT | GGTGAGTGGCAGCTTCTAGTGCT | 211 227 221/227 213 213 | aa × ab       | aa × ab       |               |             |       |       |
| FAM22     | BM437023      | (cac)7      | AAACTGGACTCCACTGT | GGTGAGTGGCAGCTTCTAGTGCT | 211 227 221/227 213 213 | aa × ab       | aa × ab       |               |             |       |       |
Table 3 continued

| Marker ID | Accession no. | Repeat type | Forward primer | Reverse primer | Expected size (bp) | Riesling | Cabernet | Summit | Noble | R × C Allele | S × N Allele |
|-----------|---------------|-------------|----------------|----------------|--------------------|----------|----------|--------|-------|--------------|--------------|
| FAM38     | CB005457      | (ct)10      | GCTTCCATACGAGAAACTCA | TAGGGAATTCACAGATTTGC | 225 235/238 238/242 231/235 231/235 | ab × cd | ab × ab |
| FAM39     | CB915484      | (ag)53      | TAGCTTCTGTTATACCAGGCT | CATATTTCAGTGGGCGACAG | 388 388/414 NA 388 388 | ab × ac |
| FAM40     | CB920839      | (ct)7       | AAGGTAATGCTCTCTCTCCTT | CACCACCTCTCTCTACCAAT | 320 328/334 328/338 326 326 | ab × ac |
| FAM41     | CAG922482     | (ag)6       | CAGAATTTGAGAGATCGAGG | ACTTCTGCCCCAATCTGACT | 175 185/191 185/191 185 185/200 | ab × ab | aa × ab |
| FAM42     | CB916884      | (ga)10      | AAATTGATTTCATCACTG GCC | CTATCATTTGTCGTCCTCCTC | 292 412 397/412 402 402 | aa × ab |
| FAM43     | CB911681      | (ag)6       | AATAAGGGAAAGAGAACG | TCAATGTGAACCAACCAAGTA | 340 489 491 491 | ab × cc |
| FAM44     | CB922984      | (ag)53      | GAGGAGGTGGAGGAAGAGAA | TTTTAGAAGTGGTATGTTGC | 113 129/135 135 135/141 132/141 | ab × aa | ab × ac |
| FAM45     | CD801743      | (ga)10      | TATAGGTACCTGCTCCT | CAAGGATGGAGCACAGTAAG | 379 395 381/398 383 | ab × cc |
| FAM46     | CD801804      | (aaagg)4    | TAACCTCAATCTCATACCT | ATATAGGGCTGTCGCAAT | 144 154/160 154/160 172/182 172/182 | ab × ab | ab × ab |
| FAM47     | CD79867      | (ata)4      | GGGATGAGCTCTACATTCC | CAATATACAGGGCTCCAA | 399 417 417 417 | ab × ab |
| FAM48     | CD798949      | (cac)6      | CTCAACAACGGAATCCACACT | AAAGCATCTTCAAGTGGTT | 241 519 519 523 | ab × ab |
| FAM49     | CD799025      | (aaaac)5    | AGCTGAGAACAAATTTCTT | CAGCAAGAATGAGTGTGA | 217 204/230 204/230 196 | 196 | ab × ab |
| FAM50     | BQ794329      | (ag)14      | CACAAAGCATGTCCATAAC | GGCTTATGCATTACTGGACT | 178 185 185 197 185 | ab × bb | ab × ab |
| FAM51     | BQ798187      | (ct)9       | CCTTGGTCTCCTCACACAAAG | CCAAGCAGTCTGAGGTACT | 101 129/130 129/130 126 | 126 | ab × ab |
| FAM52     | BQ800590      | (ta)21      | AAATCTCAACTACCTATACT | AGGTGCACTGTCGAGTTC | 142 133 NA 134 NA | ab × ab |
| FAM53     | CF603660      | (aaag)4     | AACCTCTGCACCAACACAC | CCACACCTTGCAAGTATCT | 385 394/400 394/400 390 | 390 | ab × ab |
| FAM54     | CF605057      | (cca)7      | ATCTCTGCTCTCTCTTCCTT | ATCAAGAATTATCTCAGGC | 319 326/335 326/335 326 | 326 ab × ab |
| FAM55     | CF605791      | (acc)6      | GCACCACTACCAATGTG | AGGGAAGGAGTGGATGTTG | 271 284/290 290 293 287/293 | ab × aa | aa × ab |
| FAM56     | DT019692      | (ct)6       | GCAATACGAGCCTGCTACAC | CGGTAGTGAAGGAGCTGACG | 356 368 368 368 | ab × ab |
| FAM57     | CB968972      | (ct)16      | CCAACGATACAGCAGTACAC | CGGTAGTGAAGGAGCTGACG | 356 368 368 368 | ab × ab |
| FAM58     | CF403802      | (ct)18      | TAGACGTTTGCCCTATTTGT | GCCTCTAGTGAAGGAGCTGACG | 152 155 173 158/163 171 | aa × bb | ab × cc |
| FAM59     | CB97857      | (gca)7      | GATGGTATACGACGGTGATGTA | AGGTGCACTGTCGAGTTC | 175 186/192 192 180 180 | ab × aa |
| FAM60     | CB916170      | (ca)6       | CCTATGATGCTCTGTCTCT | CGGTAGAATGGGAGCTGACG | 319 143 152/158 132/152 155 | 155 | ab × cd |
| FAM61     | CF403802      | (ct)9      | GCTACTTCTGGGAAGTTCAG | AGGTGCACTGTCGAGTTC | 375 395 395 406 406 | ab × ab |
| FAM62     | CB342303      | (gca)6      | CTAATCTCCAGCAAGAAACAAC | GTACGAATGTTGCTCATTTG | 274 257/289 NA 253/286 286 | ab × aa |
| FAM63     | CB346557      | (ca)6       | AATAAGGCTCTCTCTCTCAT | CGGTAGTGAAGGAGCTGACG | 334 350 350 348/360 348/360 | ab × ab |
| FAM64     | CN007369      | (gca)6      | CTAATCTCCAGCAAGAAACAAC | GTACGAATGTTGCTCATTTG | 274 257/289 NA 253/286 286 | ab × aa |
| FAM65     | DT019756      | (gca)6      | CCTTCTACGCTGCTCCTACAT | TTAATGATGCTTGAGTATT | 129 145/148 145 142 142 | ab × aa |
| FAM66     | BQ794995      | (ata)10     | TTTATCTCAACCTCTACCTACAT | GTTGTAGGAGTAGCCTCC | 200 210 210 210 207/212 | ab × bc |
| FAM67     | CB974335      | (attt)4     | GGACTTCATCTGCTGGTAGA | TTTATCTCAACCTCTACCTACAT | 376 395 395 387 | ab × bc |
| FAM68     | CF607255      | (gaa)6      | AAACCTTACCGAAATGCTCCT | CTTCTCTTCGGCTCCTGTGA | 307 321 321 321 | ab × ab |
| FAM69     | CF608166      | (gaa)6      | CATCCATAGCTCCTGTCGAC | CTTCTCTTCGGCTCCTGTGA | 307 321 321 321 | ab × ab |
| FAM70     | CN545596      | (gaa)10     | TGGGAAATGAGAAGGAGTGT | AACACATCTCCAGATTTGGG | 281 389/392 NA 380/397 397/403 | ab × ac |
| FAM71     | CN548154      | (ata)9     | AGTCTTCTCAGTGGCTCCTCAG | CTTGTCATACGGCAAGAACAA | 180 202 195 195 197 | aa × bb | aa × bb |
| FAM72     | CN549034      | (ct)14     | TCAGTCCAGATCTCTTCCTT | TCATGTCATACGGCAAGAACAA | 171 172/188 172/188 180 170/180 | ab × ab | ab × ab |
| Marker ID | Accession no. | Repeat type | Forward primer | Reverse primer | Expected size (bp) | Riesling | Cabernet | Summit | Noble | R × C | S × N |
|-----------|---------------|-------------|----------------|----------------|-------------------|-----------|-----------|--------|-------|-------|-------|
| FAM73     | CO818811      | (ttc)5      | GCATATGGAAAGGGATAA | ATTTGTGCAATGATCGAAG | 359 | 390/399 | 377/399 | 369 | 369 | ab × ac |
| FAM74     | CO819364      | (ag)8       | GGCCTCAGAGTACAGTAA | GGCCTCCTGCTCATAGAAC | 289 | 307 | 307 | 234/318 | 309 | ab × cc |
| FAM75     | DT004860      | (ct)11      | CCTGTAAGAGGCTCAATCT | ATGGCTGAGCTGATAGAGG | 169 | 184/187 | 187 | 172/175 | 172/175 | ab × aa | ab × ab |
| FAM76     | DT009858      | (cag)6      | ATTAACGGAGATGGTGTTG | AAGAATGCAGATCTCAATCAG | 394 | 409 | 409 | 427 | 432 | 433 | aa × bb |
| FAM78     | DT010742      | (tgaag)5    | TTCATGACAATTGTGTTG | CGGACTCATAGAGAAG | 303 | 322 | 322 | 339 | 339 | ab × aa |
| FAM79     | DT011109      | (ag)12      | GCAGAAGCAGGAAGTGAAG | AGATTCGAAACCCATCTGAAGA | 147 | 162/172 | 162/172 | 153/157 | 153/157 | ab × ab | ab × ab |
| FAM80     | DT016866      | (ct)9       | AACTTCCAAGCAGACCA | AGTATGCTCTGACAGTCAA | 316 | 424 | 424 | 428 | 428 | ab × cd |
| FAM81     | DT01972       | (gcc)8      | TTCTTCAACATAGATGCA | GACTAGTATACCTTGGT | 203 | 218 | 218 | 226 | 226 | ab × cd |
| FAM82     | DT02098       | (tc)8       | AGAAGCAGCTTACATCGAGA | GAAAGGATGAAATCTCTGAAGA | 344 | 357/361 | 361 | 348 | 348 | ab × aa |
| FAM83     | DT02268       | (gagag)4    | CTCCTCTGAGGAAAGGTTF | CATTGCTGACAAACATCG | 249 | 351/275 | 255 | 244/255 | 244/255 | ab × cc | ab × ab |
| FAM84     | DT024858      | (ga)18      | CCAATATCTCACAACAGA | TACCAATATGTGAACTCGT | 177 | 195 | 195/210 | 184 | 184 | aa × ab |
| FAM85     | DT025345      | (ta)10      | GAATTCGAAAGGACAGACAA | TATATCGGAGAACAAACA | 283 | 300 | 300/304 | 308 | 308 | aa × ab |
| FAM86     | DT026122      | (at)11      | AGAAAATCCAGTCGGCTCA | GGAGGAGACCATAGCAGTAG | 267 | 276 | 284 | 277 | 277 | NA | aa × bb |
| FAM87     | DT026426      | (ga)6       | TCCGAAGAAAGAGAAAGAAG | CGTGGCCTATGACATTGAAG | 375 | 388/400 | 388 | 421/436 | 421/433 | ab × aa | ab × ac |
| FAM88     | DV220116      | (at)14      | AATGTCGAAAGCCCAGCAGG | CAGTGCTGACAGCATGAAACA | 237 | 250/242 | 254/256 | 223 | 226 | ab × cd |
| FAM89     | DV220950      | (ag)9       | CCTTGTGCTTGTCTGATG | CTAAGGCTGCTGATGACCT | 193 | 213 | 213 | 211 | 211 | ab × ab |
| FAM90     | DV221456      | (ct)10      | GACCAAAATGTTGCGAGG | AACAGAAGAAACAGAGGTTA | 368 | 358/387 | 387 | 392/400 | 400 | ab × cd |
| FAM91     | DV222762      | (ct)8       | ATTATCGCAACAAAGATGTC | ATACGCTCTGTAACGTGAA | 159 | 176 | 173/176 | 161 | 161 | aa × bc |
| FAM92     | DV39679       | (ta)9       | TTGTATTTTGTGCTACCTTT | ACCCTCTGATAACCTTGGG | 388 | 405 | 405 | 417 | 417 | ab × ac |
| FAM93     | DV940288      | (ag)12      | ATATCTTCTATCCCCTGA | ATGTCGCTAGGAGTCTCA | 242 | 258 | 244/258 | 238/243 | 238/243 | ab × ab | ab × ab |
| FAM94     | CV092439      | (gga)7      | GAGCTTCTTGACCCTCA | CCCATGATAACAGGCTTGA | 247 | 261 | 261 | NA | NA | ab × aa |
| FAM95     | CV092870      | (ct)8       | ATTCCTCGAGATGATGATG | ATGACTGCTGACAGTCAA | 190 | 207 | 207 | 182/191 | 182 | ab × aa |
| FAM96     | CV093018      | (tcg)6      | TCCATGAGGTTGCTGCTAG | CCAAGCTGATGCTGATG | 216 | 231 | 231 | 223 | 223 | ab × ab |
| FAM97     | CV094376      | (cag)11     | ATCAGGTCGAACTGATCTA | GAGAATGGAAATGCTTGAAT | 302 | 326 | 326 | 320/324 | 320/324 | ab × ab | ab × ab |
| FAM98     | CV094448      | (ctt)8      | CTCCTCTGACTTCCACTCA | ATATGCTGAGATGCTGTC | 301 | 326 | 318/326 | 320 | 320/324 | ab × ab | ab × ab |
| FAM99     | CV106105      | (cag)6      | AAAGGGTCTTCTGAGAACCC | TCTCTGACAAACGGAGG | 288 | 395 | 395 | 395 | 395 | ab × ab |
| FAM100    | CV203674      | (tc)13      | CCTTCCAGGCTCTGAAAC | GAGTATGACAGAGTGAAGA | 185 | 186 | 186 | 182/191 | 182 | ab × aa |
| FAM101    | CV204388      | (ta)10      | TGTGACTATGTGTTTCTTGATGT | CAAGATGATGCTGCTGAGT | 251 | 283/286 | 283 | 292/294 | 292/298 | ab × aa | ab × ac |
| FAM102    | CV204610      | (tag)5      | ATGCTACAGATGGCACTG | CAGAGATGGAAGTGTTAGCT | 147 | 154 | 154 | 151 | 151 | ab × ab |
| FAM103    | CV205250      | (ata)6      | AGCGCTGCTTCTGGAAG | GATCAGTACGGCAAGTGGAC | 245 | 347 | 347 | 350 | 344/350 | ab × ab |
| FAM104    | CV205801      | (acc)6      | AAAGGGTCTTCTTCTGAGACC | TCTCTGACAAACGGAGG | 288 | 395 | 395 | 395 | 395 | ab × ab |
| FAM105    | CV206458      | (ga)7       | GAGCTTCTTGACATTCTC | CAAATGCTGAGGCTGCAAGT | 113 | 132 | 132 | 132 | 132 | ab × ab |
| FAM106    | CV206674      | (tc)13      | CCTTCCAGGCTCTGAAAC | GAGTATGACAGAGTGAAGA | 185 | 186 | 186 | 182/191 | 182 | ab × aa |
| FAM107    | CV209318      | (tgg)6      | TTCTTCAATCATCCTCT | GAGCCCTGCTACCTTGGT | 202 | 216 | 216 | 222 | 219/225 | ab × bc |
| FAM108    | CV209444     | (cct)8      | CTCCTCTGACTTCCACTCA | ATGCGTGGAAATGCTTGAAG | 151 | 156 | 156 | 156 | 156 | ab × ab | ab × ab | ab × ab | ab × ab |

*Table 3 continued*
| Marker ID | Accession no. | Repeat type | Forward primer | Reverse primer | Expected size (bp) | Riesling | Cabernet | Summit | Noble | R × C | S × N |
|-----------|---------------|-------------|----------------|----------------|--------------------|-----------|----------|--------|-------|-------|-------|
| FAM109    | CV094765     | (at)9       | GTTAAACTCTAGGCCGTTTGG | GCTTGCACTAGTATTACAGAA | 321                 | 333       | 333       | 333    | 333   |       |       |
| FAM110    | CV095258     | (ct)6       | GGCTATGATTTCAGCCTCTAATC | TACAGGCCTGGTTATTCACC | 192                 | 287/303   | 303       | 290    | 290   | ab × aa |       |
| FAM112    | CV097295     | (ga)8       | AGTTTGGTAATCGAGTCCC | TCTTGAATACGATGAGTTG | 205                 | 206       | 206       | 199    | 199   |       |       |
| FAM115    | CV098232     | (gga)6      | ACCCTATCTAAGCTCTTC | GACCTTACCTAGTACCTT | 246                 | 275       | 275/291   | 260    | 260   | aa × ab |       |
| FAM116    | CV098042     | (agg)6      | AACTAATTGGCAAGGAA | CACACCTTGATGACCATTG | 319                 | 347       | 347       | 345    | 345   |       |       |
| FAM117    | CV099053     | (agg)16     | CTTCTATTCGAGCCACCTTC | TCTTGTGAGGAGAGTGTT | 330                 | 333       | 333/341   | 333    | 333   | aa × ab |       |
| FAM118    | CV097295     | (gaa)8      | AGTTTGGTAATCGAGTCCC | TCTTGAATACGATGAGTTG | 205                 | 206       | 206       | 199    | 199   |       |       |
| FAM119    | CV098402     | (agg)6      | ACCCTATCTAAGCTCTTC | GACCTTACCTAGTACCTT | 246                 | 275       | 275/291   | 260    | 260   | aa × ab |       |
| FAM120    | CV099069     | (gtg)7      | TAGTGAATACGACAGTGAGG | AGTTGCACTATCCAGTACCT | 318                 | 345       | 345       | 333    | 333   | aa × ab |       |
| FAM121    | CV099069     | (gtg)7      | TAGTGAATACGACAGTGAGG | AGTTGCACTATCCAGTACCT | 318                 | 345       | 345       | 333    | 333   | aa × ab |       |
| FAM122    | CV099069     | (gtg)7      | TAGTGAATACGACAGTGAGG | AGTTGCACTATCCAGTACCT | 318                 | 345       | 345       | 333    | 333   | aa × ab |       |
| FAM123    | CV099069     | (gtg)7      | TAGTGAATACGACAGTGAGG | AGTTGCACTATCCAGTACCT | 318                 | 345       | 345       | 333    | 333   | aa × ab |       |
| FAM124    | CV099069     | (gtg)7      | TAGTGAATACGACAGTGAGG | AGTTGCACTATCCAGTACCT | 318                 | 345       | 345       | 333    | 333   | aa × ab |       |
| FAM125    | CV099069     | (gtg)7      | TAGTGAATACGACAGTGAGG | AGTTGCACTATCCAGTACCT | 318                 | 345       | 345       | 333    | 333   | aa × ab |       |
| FAM126    | CV099069     | (gtg)7      | TAGTGAATACGACAGTGAGG | AGTTGCACTATCCAGTACCT | 318                 | 345       | 345       | 333    | 333   | aa × ab |       |
| FAM127    | CV099069     | (gtg)7      | TAGTGAATACGACAGTGAGG | AGTTGCACTATCCAGTACCT | 318                 | 345       | 345       | 333    | 333   | aa × ab |       |
| FAM128    | CV099069     | (gtg)7      | TAGTGAATACGACAGTGAGG | AGTTGCACTATCCAGTACCT | 318                 | 345       | 345       | 333    | 333   | aa × ab |       |
| FAM129    | CV099069     | (gtg)7      | TAGTGAATACGACAGTGAGG | AGTTGCACTATCCAGTACCT | 318                 | 345       | 345       | 333    | 333   | aa × ab |       |
| FAM130    | CV099069     | (gtg)7      | TAGTGAATACGACAGTGAGG | AGTTGCACTATCCAGTACCT | 318                 | 345       | 345       | 333    | 333   | aa × ab |       |
| FAM131    | CV099069     | (gtg)7      | TAGTGAATACGACAGTGAGG | AGTTGCACTATCCAGTACCT | 318                 | 345       | 345       | 333    | 333   | aa × ab |       |
| FAM132    | CV099069     | (gtg)7      | TAGTGAATACGACAGTGAGG | AGTTGCACTATCCAGTACCT | 318                 | 345       | 345       | 333    | 333   | aa × ab |       |
| FAM133    | CV099069     | (gtg)7      | TAGTGAATACGACAGTGAGG | AGTTGCACTATCCAGTACCT | 318                 | 345       | 345       | 333    | 333   | aa × ab |       |
| FAM134    | CV099069     | (gtg)7      | TAGTGAATACGACAGTGAGG | AGTTGCACTATCCAGTACCT | 318                 | 345       | 345       | 333    | 333   | aa × ab |       |
| FAM135    | CV099069     | (gtg)7      | TAGTGAATACGACAGTGAGG | AGTTGCACTATCCAGTACCT | 318                 | 345       | 345       | 333    | 333   | aa × ab |       |
| FAM136    | CV099069     | (gtg)7      | TAGTGAATACGACAGTGAGG | AGTTGCACTATCCAGTACCT | 318                 | 345       | 345       | 333    | 333   | aa × ab |       |
| FAM137    | CV099069     | (gtg)7      | TAGTGAATACGACAGTGAGG | AGTTGCACTATCCAGTACCT | 318                 | 345       | 345       | 333    | 333   | aa × ab |       |
| FAM138    | CV099069     | (gtg)7      | TAGTGAATACGACAGTGAGG | AGTTGCACTATCCAGTACCT | 318                 | 345       | 345       | 333    | 333   | aa × ab |       |
| FAM139    | CV099069     | (gtg)7      | TAGTGAATACGACAGTGAGG | AGTTGCACTATCCAGTACCT | 318                 | 345       | 345       | 333    | 333   | aa × ab |       |
| FAM140    | CV099069     | (gtg)7      | TAGTGAATACGACAGTGAGG | AGTTGCACTATCCAGTACCT | 318                 | 345       | 345       | 333    | 333   | aa × ab |       |
| FAM141    | CV099069     | (gtg)7      | TAGTGAATACGACAGTGAGG | AGTTGCACTATCCAGTACCT | 318                 | 345       | 345       | 333    | 333   | aa × ab |       |
| FAM142    | CV099069     | (gtg)7      | TAGTGAATACGACAGTGAGG | AGTTGCACTATCCAGTACCT | 318                 | 345       | 345       | 333    | 333   | aa × ab |       |
| FAM143    | CV099069     | (gtg)7      | TAGTGAATACGACAGTGAGG | AGTTGCACTATCCAGTACCT | 318                 | 345       | 345       | 333    | 333   | aa × ab |       |
| FAM144    | CV099069     | (gtg)7      | TAGTGAATACGACAGTGAGG | AGTTGCACTATCCAGTACCT | 318                 | 345       | 345       | 333    | 333   | aa × ab |       |
| FAM145    | CV099069     | (gtg)7      | TAGTGAATACGACAGTGAGG | AGTTGCACTATCCAGTACCT | 318                 | 345       | 345       | 333    | 333   | aa × ab |       |
out of 1,701) of the EST-SSRs were annotated and grouped by the Biological Process Classification using the MIPS MATDB Arabidopsis Scheme. The most abundant EST-SSRs belonged to the categories of protein-binding (22%) and subcellular localization (18%; Fig. 2), which demonstrated a similar pattern to wheat, rice, maize and barley (Tang et al. 2006). The 150 validated markers were further functionally annotated (Electronic Supplementary Material 2) and estimation of their genomic/chromosome locations by comparison to grapevine genome assembly (Jaillon et al. 2007; Velasco et al. 2007) is given in Electronic Supplementary Material 3.

SSR marker development and validation

The Beckman CEQ8800 Genetic Analyzer was used for the SSR validation and analysis. This system can detect DNA fragment length polymorphism in a “single base pair”. A set of 150 primer pairs was initially screened for SSR marker development and validation. Parents of two mapping populations, *V. vinifera* Riesling × Cabernet Sauvignon (Riaz et al. 2004) and *V. rotundifolia* Summit × Noble (Ren et al. 2000) were used for the screening. Results showed that 145 out of 150 primers had well-amplified fragments among the four cultivars (Table 3). Some of the fragment sizes exceeded expected sizes possibly due to their having introns within the flanking regions or the length of the repeat being shorter than the source species, and less prone to polymerase slippage. Polymorphisms were found in 66 primer pairs between Riesling and Cabernet Sauvignon, and 40 between Summit and Noble. Only 16 of the polymorphic primers shared the same polymorphic lengths between these two parent pairs, reflecting the fact that the alleles between *V. vinifera* and *V. rotundifolia* grape are distinct (Riaz et al. 2008).

The homo and heterozygosity of these 145 loci were screened in the four testing cultivars; 92 of 144 were identified as homozygous and 52 were heterozygous loci in Riesling, while 86 of 136 were homozygous and 49 were heterozygous loci in Cabernet Sauvignon. Among the 92 Riesling and 86 Cabernet Sauvignon homozygous loci, 68 are common in both parents (Table 4). As for the muscadine grapes Noble and Summit, Noble showed 97 homozygous and 39 heterozygous loci and the respective number for Summit was 108 and 34 (Table 4). Some of those
microsatellite loci were selected for having long stretches in *V. vinifera* grapes, and thus may show more polymorphisms than in muscadines. From this screening of 145 loci, muscadine grapes demonstrated a higher homozygosity compared to *V. vinifera* grapes. Pairwise comparison between Riesling and Cabernet Sauvignon showed that 72 loci were monomorphic with either one allele (60) or two (12). Seventy loci were heterozygous in at least one cultivar with either two (53), three (10), or four alleles (7; Table 5). The muscadine Noble and Summit showed 50 heterozygous loci in at least one parent with either two (31), three (17), or four alleles (2). Ninety homozygous loci were found in both cultivars with either one (84) or two alleles (6; Table 5). According to the results from these 145 EST-SSR loci, the percentage of polymorphisms is about 49% between Riesling and Cabernet Sauvignon, and 29% between Summit and Noble. However, those polymorphic SSRs that are homozygous (e.g. aa × bb) in both parents cannot be mapped in F₁ populations although they are useful for mapping in F₂ or backcross populations (Chen et al. 2006). The heterozygous monomorphic SSRs (e.g. ab × ab) can be used for mapping in F₁ populations (Table 5). As a result, the estimated number of SSRs that can be mapped in the F₁ populations between Riesling and Cabernet Sauvignon is about 46%, which means that out of the total 1,037 SSRs with successful primers designed, around 477 EST-SSR putative markers can be mapped in the F₁ of Summit × Noble.

EST-SSR marker transferability was evaluated and the current research showed a high transferability across species. All but two of the 145 EST-SSR markers

![Functional prediction of 994 grape SSR-EST based on the MIP MATDB classification scheme](image-url)
in *Vitis vinifera* appeared in the muscadine as well. This result indicated that development of EST-SSR markers is a cost-effective method for obtaining additional markers for grape genome typing and gene mapping.

EST-SSRs provided sources of additional markers for marker development. Compared to genomic-derived markers, EST-SSRs are highly transferable for detecting the gene-rich areas within the genome. We can utilize these markers to evaluate marker transferability across taxa, and conduct analysis in comparative mapping and gene functional diversity analysis, in addition to genotyping. The functional EST-SSR markers should be even more useful for developing a linkage map or tagging a viticulturally important trait. In addition, the polymorphic EST-SSR markers are much needed for genotyping, cultivar identification and development of a linkage map in muscadine grapes since they are genetically much less diversified than *Vitis* species.

**Acknowledgments** This work was supported by USDA Capacity Building Grant (#0205031) and FAMU-ARS Science Center for Excellence. The authors thank two anonymous reviewers for invaluable comments on the manuscript. Special thanks go to Ms. Elisa Scott for helping us to prepare the DNA and PCR samples.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

**References**

Adam-Blondon AF, Roux C, Claux D, Butterlin G, Merdinoglu D, This P (2004) Mapping 245 SSR markers on the *Vitis vinifera* genome: a tool for grape genetics. Theor Appl Genet 109(5):1017–1027

Akkak A, Lembo A, Torello Marioni D, Gribaudo I, Botta R (2006) Analysis and transferability of EST-SSR markers in grape. Acta Hort (ISHS) 827:59–62

Arroyo García R, Lefort F, de Andrés M, Ibañez J, Borrego J, Jouve N, Cabello F, Martínez Zapater J (2002) Chloroplast microsatellite polymorphisms in *Vitis* species. Genome 45(6):1142–1149

Bocharova B, Kovaliova I, Mazurenko L (2009) Identification of grapevine clone genotypes by use of microsatellite markers. Cytol Genet 43(6):371–378

Bowers J, Dangl G, Vignani R, Meredith C (1996) Isolation and characterization of additional microsatellite DNA markers for grape. Am J Enol Vitic 47(3):243–246

Cardle L, Ramsey L, Milbourne D, Macaulay M, Marshall D, Waugh R (2000) Computational and experimental characterization of physically clustered simple sequence repeats in plants. Genetics 156:847–854

Chabane K, Ablett G, Cordeiro G, Valkoun J, Henry R (2005) EST versus genomic derived microsatellite markers for genotyping wild and cultivated barley. Genet Res Crop Evol 52:903–909

Chani E, Ashkeni V, Hillel J, Veilleux R (2002) Microsatellite marker analysis of an anther-derived potato family: skewed segregation and gene-centromere mapping. Genome 45(2):236–242

Chen C, Zhou P, Choi Y, Huang S, Gmitter F (2006) Mining and characterizing microsatellites from citrus ESTs. Theor Appl Genet 112(7):1248–1254

Cho Y, Ishii T, Temnykh S, Chen X, Lipovich L, Park W, Ayres N, Cartinhoure S, McCouch S (2000) Diversity of microsatellites derived from genomic libraries and GenBank sequences in rice (*Oryza sativa* L.) Theor Appl Genet 100(5):713–722

| Alleles | R × C<sub>a</sub> | Number | S × N<sub>b</sub> | Number | Mappable in F<sub>1</sub> |
|---------|----------------|--------|----------------|--------|--------------------------|
| 1       | aa × aa        | 57     | aa × aa        | 80     | No                       |
| 2       | aa × bb        | 9      | aa × bb        | 6      | No                       |
| 2       | aa × ab        | 15     | aa × ab        | 13     | Yes                      |
| 2       | ab × aa        | 18     | ab × aa        | 4      | Yes                      |
| 2       | ab × ab        | 12     | ab × ab        | 13     | Yes                      |
| 3       | aa × bc        | 8      | aa × bc        | 4      | Yes                      |
| 3       | ab × cc        | 2      | ab × cc        | 4      | Yes                      |
| 3       | ab × ac        | 6      | ab × ac        | 8      | Yes                      |
| 4       | ab × cd        | 9      | ab × cd        | 1      | Yes                      |
| Total mappable | 70 | 47 | | | |

<sup>a</sup> R × C: Riesling × Cabernet Sauvignon

<sup>b</sup> S × N: Summit × Noble
Cipriani G, Marrazzo M, Di Gaspero G, Pfeiffer A, Morgante M, Testolin R (2008) A set of microsatellite markers with long core repeat optimized for grape (Vitis spp.) genotyping. BMC Plant Biol 8:127

Cordeiro M, Casu R, McIntyre C, Manners J, Henry R (2001) Microsatellite markers from sugarcane (Saccharum spp) ESTs across transferable to erianthus and sorghum. Plant Sci 160(6):1115–1123

Costantini L, Grando M, Feingold S, Ulanovsky S, Mejía N, Depeiges A, Goubely C, Lenoir A, Cocherel S, Picard G, Raynal Decroocq V, Fave M, Hagen L, Bordenave L, Decroocq S De Mattia F, Imazio S, Grassi F, Lovicu G, Tardaguila J, Maitti M, Testolin R (2008) A set of microsatellite markers with functional annotation toolkit (Hit-Go-Fat) utilized for marker assisted grape breeding. Am J Enol Vitic 58:102–111

Dalbo M, Ye G, Weeden N, Steinckelner H, Sefc K, Reisch B (2000) Gene controlling sex in grapevines placed on a molecular marker-based genetic map. Genome 43:333–340

De Keyser E, De Riek J, Van Bockstaele E (2009) Discovery of species-wide EST-derived markers in Rhododendron by intron-flanking primer design. Mol Breed 23(1):171–178

De Mattia F, Imazio S, Grassi F, Lovicu G, Tardaguila J, Maitti C, Scienza A, Labra M (2007) Genetic characterization of Sardinia grapevine cultivars by SSR markers analysis. Int J Vine Wine Sci 41:175–184

Decroocq V, Favé M, Hagen L, Bordenave L, Decroocq S (2003) Development and transferability of apricot and grape EST microsatellite markers across taxa. Theor Appl Genet 106:912–922

Depeiges A, Goubely C, Lenoir A, Cocherel S, Picard G, Raynal M, Grulet F, Delseny M (1995) Identification of the most represented repeated motifs in Arabidopsis thaliana microsatellite loci. Theor Appl Genet 91(1):160–168

Di Gaspero G, Peterlunger E, Testolin R, Edwards K, Cipriani G (2000) Conservation of microsatellite loci within the genus Vitis. Theor Appl Genet 101(1):301–308

Di Gaspero G, Cipriani G, Marrazzo M, Andreotta D, Prado Castro M, Peterlunger E, Testolin R (2005) Isolation of (AC)n-microsatellites in Vitis vinifera L. and analysis of genetic background in grapevines under marker assisted selection. Mol Breed 15(1):11–20

Di Gaspero G, Cipriani G, Adam-Blondon A-F, Testolin R (2007) Linkage maps of grapevine displaying the chromosomal locations of 420 microsatellite markers and 82 markers for R-gene candidates. Theor Appl Genet 114(7):1249–1263

Doligez A, Bouquet A, Danglot Y, Lahogue F, Riaz S, Meredith C, Edwards K, This P (2002) Genetic mapping of grapevine (Vitis vinifera L.) applied to the detection of QTLs for seedlessness and berry weight. Theor Appl Genet 105:780–795

Doucelf M, Jin Y, Gao F, Riaz S, Krivanek A (2004) A genetic linkage map of grape, utilizing Vitis rupestris and Vitis arizonica. Theor Appl Genet 109:1178–1187

Dowd S, Zaragoza J (2005) High throughput gene ontology functional annotation toolkit (Hit-Go-Fat) utilized for animal and plant. Plant and Animal Genome XIII Conference, San Diego, CA

Eujayl I, Sorrells M, Baum M, Wolters P, Powell W (2002) Isolation of EST-derived microsatellite markers for genotyping the A and B genomes of wheat. Theor Appl Genet 104(2):399–407

Feng S, Li W, Huang H, Wang J, Wu Y (2009) Development, characterization and cross-species/genera transferability of EST-SSR markers for rubber tree (Hevea brasiliensis). Mol Breed 23(1):85–97

Fernández M, Núñez Y, Ponz P, Hernández S, Gallego F, Ibáñez J (2008) Characterization of sequence polymorphisms from microsatellite flanking regions in Vitis spp. Mol Breed 22(3):455–465

Gao L, Tang J, Li H, Jia J (2003) Analysis of microsatellites in major crops assessed by computational and experimental approaches. Mol Breed 113(3):163–185

Grando MS, Bellin D, Edwards KJ, Pozzi C, Stefanini M, Velasco R (2003) Molecular linkage maps of Vitis vinifera L. and Vitis riparia Mchx. Theor Appl Genet 106:1213–1224

Jaillon O, Aury JM, Noel B, Policriti C, Clepet C, Casagrande A, Choisne N, Aubourg S, Vitulo N, Jabin C, Vezzi A, Legeai F, Hugueney P, Dasilva C, Horner D, Mica E, Jublot D, Poulain J, Bruyère C, Billaud A, Segurens B, Gouyvenoux M, Ugarté E, Cotonaro F, Anthouard V, Vico V, Del Fabbro C, Alaux M, Di Gaspero G, Dumas V, Felice N, Paillard S, Juman I, Moroldo M, Scalabrin S, Canaguier A, Le Clainche I, Macrì C, Durand E, Pesole G, Laucou V, Chatelot P, Merdinioglu D, Dellemonde M, Pezzotti M, Lecharny A, Scarpelli C, Artiguenave F, Pê ME, Valle G, Morgante M, Caboche M, Adam-Blondon A-F, Weissenbach J, Quétier F, Wincker P (2007) The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. Nature 449:463–467

Kantety R, La Rota M, Matthews D, Sorrells M (2002) Data mining for simple sequence repeats in expressed sequence tags from barley, maize, rice, sorghum and wheat. Plant Mol Biol 48(5):501–510

Lamoureux D, Bernole A, Le Clainche I, Tual S, Thareau V, Paillard S, Legeai F, Dossat C, Wincker P, Oswald M, Merdinioglu D, Vignault C, Delrot S, Caboche M, Chahb H, Adam-Blondon AF (2006) Anchoring of a large set of markers onto a BAC library for the development of a draft physical map of the grapevine genome. Theor Appl Genet 113:344–356

Li Y, Korol A, Fahima T, Nevo E (2004) Microsatellites within genes: structure, function, and evolution. Mol Bio Evol 21(6):991–1007

Lima L, Gramacho K, Gesteira A, Lopes U, Gaiotto F, Zaidan R, Lin H, Civerolo E, Hu R, Barros S, Francis M, Walker M (2005) Multilocus simple sequence repeat markers for differentiating strains and evaluating genetic diversity of Xylella fastidiosa. Appl Environ Microbiol 71(8):4888–4892

Lowe KM, Walker MA (2006) Genetic linkage map of the interspecific grape rootstock cross Ramsey (Vitis x riparia) × Riparia Gloire (Vitis riparia). Theor Appl Genet 112:1582–1592
Meridinoglu D, Butterlin G, Bevilacqua L, Chiquet V, Adam-Blondon AF, Decrooq S (2005) Development and characterization of a large set of microsatellite markers in grapevine \((Vitis vinifera \text{ L.})\) suitable for multiplex. Mol Breed 15(4):349–366

Metzgar D, Bytof J, Wills C (2000) Selection against frame shift mutations limits microsatellite expansion in coding DNA. Genome Res 10(1):72–80

Oetting W, Lee H, Flanders D, Wiesner G, Sellers T, King R (1995) Linkage analysis with multiplexed short tandem repeat polymorphisms using infrared fluorescence and M13 tagged primers. Genomics 30(3):450–458

Qu X, Lu J, Lamikanra O (1996) Genetic diversity in Muscadinia and American bunch grapes based on randomly amplified polymorphic DNA (RAPD) analysis. J Am Soc Hort Sci 121(6):1020–1023

Ren Z, Lamikanra O, Lu J (2000) Identification of a RAPD marker closely linked to the fruit color in muscadine grape \((Vitis rotundifolia)\). Acta Hort 528:263–266

Riaz S, Dangl G, Edwards K, Meredith C (2004) A microsatellite marker based framework linkage map of \(Vitis vinifera \text{ L.}\). Theor Appl Genet 108(5):864–872

Riaz S, Krivanek AF, Xu K, Walker MA (2006) REWned mapping of the Pierce’s disease resistance locus, PdR1, and sex on an extended genetic linkage map of \(Vitis rupestris \times V. arizonica\). Theor Appl Genet 113:1317–1329

Riaz S, Tenschler A, Smith B, Ng D, Walker M (2008) Use of SSR markers to assess identity, pedigree, and diversity of cultivated \(Muscadinia rotundifolia\). J Am Soc Hort Sci 133:559–568

Riaz S, Tenschler A, Graziani R, Krivanek A, Ramming D, Walker A (2009) Using marker-assisted selection to breed Pierce’s disease-resistant grapes. Am J Enol Vitic 60:199–207

Roy J, Balyan H, Prasad M, Gupta P (2002) Use of SAMPL for a study of DNA polymorphism, genetic diversity and possible gene tagging in bread wheat. Theor Appl Genet 104(2):465–472

Salmaso M, Malacarne G, Troggio M, Faes G, Stefanini M, Grando M, Velasco R (2008) A grapevine \((Vitis vinifera \text{ L.})\) genetic map integrating the position of 139 expressed genes. Theor Appl Genet 116:1129–1143

Scott K, Eggerl P, Seaton G, Rossetto M, Ablet E, Lee L, Henry R (2000) Analysis of SSRs derived from grape ESTs. Theor Appl Genet 100(5):723–726

Sefc K, Steinkellner H, Glössl J, Knappler S, Regner F (1998) Reconstruction of a grapevine pedigree by microsatellite analysis. Theor Appl Genet 97(2):227–231

Tang J, Gao L, Cao Y, Jia J (2006) Homologous analysis of SSR-ESTs and transferability of wheat SSR-EST markers across barley, rice and maize. Euphytica 151:1–87

Tekeoglu M, Rajesh N, Muelbauer J (2002) Integration of sequence tagged microsatellite sites to the chickpea genetic map. Theor Appl Genet 105(6):847–854

Tessier C, David J, This P, Boursiquot J, Charrier A (1999) Optimization of the choice of modular markers for varietal identification in \(Vitis vinifera \text{ L.}\). Theor Appl Genet 98(1):171–177

Thiel T, Michalek W, Varshney R, Graner A (2003) Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley \((Hordeum vulgare \text{ L.})\). Theor Appl Genet 106(3):411–422

Thomas M, Scott N (1993) Microsatellite repeats in grapevine reveal DNA polymorphisms when analyzed as sequence-tagged sites (STSs). Theor Appl Genet 86(8):985–990

Toth G, Gaspari Z, Jurka J (2000) Microsatellites in different eukaryotic genomes: survey and analysis. Genome Res 10(7):967–981

Varshney R, Thié L, Stein N, Langridge P, Graner A (2002) In silico analysis on frequency and distribution of microsatellites in ESTs of some cereal species. Cell Mol Biol Lett 7(2A):537–546

Varshney R, Graner A, Sorrells M (2006) Genic microsatellite markers in plants: features and applications. Trends Biotech 23(1):48–55

Velasco R, Zharkikh A, Troggio M, Cartwright D, Cestaro A, Pruss D, Pindo M, FitzGerald L, Vezzulli S, Reid J, Malacarne G, Iliev D, Coppola G, Wardell B, Micheletti D, Macalma T, Facci M, Mitchell J, Perazzolli M, Eldredge G, Gatto P, Oyzerski R, Moretto M, Gutin N, Stefanini M, Chen Y, Segala C, Davenport C, Dematè L, Mraz A, Battilana J, Stormo K, Costa F, Tao Q, Si-Amour A, Harkins T, Lackey A, Perbost C, Taillon B, Stella A, Solovyev V, Fawcett J, Sterck L, Vandepoele K, Grando M, Toppo S, Moser C, Lanchbury J, Bogden R, Skolnick M, Sgaramella V, Bhatnagar S, Fontana P, Gutin A, Van de Peer Y, Salamini F, Viola R (2007) High quality draft consensus sequence of the genome of a heterozygous grapevine variety. PLoS ONE 2:e1326

Vezzulli S, Troggio M, Coppola G, Jermakow A, Cartwright D, Zharkikh A, Stefanini M, Grando M, Viola R, Adam-Blondon AF, Thomas M, This P, Velasco R (2008) A reference integrated map for cultivated grapevine \((Vitis vinifera \text{ L.})\) from three crosses, based on 283 SSR and 501 SNP-based markers. Theor Appl Genet 117:499–511

Wren J, Forgacs E, Fondon J, Pertsemlidis A, Cheng S, Galagan J, Skolnick M, Gonzalez-Maeso E, Lander E, Chang J, Lipman D, Green P, Merrell J, Mathur A, Hillier L, Metzker M, Mardis E, Patrinos A, Tremblay L, Foerster S, Glassbell S, Domin RM, Nebel N, Blumenstock J, Wallis J, Goodhead I, Vingron M, Chang J, Rabinowicz PD, Shumway M, McIrvin E, Flint J, McCombie S, Glazko G, Parkin R, Jhiang S, Serrano E, Glass M, Moore AT, Lewis S, Zhai F, Jensen PE, Hall BD, Batzer M, Scherer S, Green P, Strauss S, Krawczak M, Ng S, Venter E, Weinstock GM, Green ED, Stein KE, et al. (2009) The sequence of the human genome. Nature 451(7183):702–712

Yu J, Dake T, Singh S, Benschker D, Li W, Gill B, Sorrells M (2004) Development and mapping of EST-derived simple sequence repeat markers for hexaploid wheat. Genome 47(5):805–818

Zhang J, Guo W, Zhang T (2002) Molecular linkage map of allotetraploid cotton \((Gossypium hirsutum \text{ L.} \times Gossypium barbadense \text{ L.})\) with a haploid population. Theor Appl Genet 105(8):1166–1174