Theor Appl Genet (2008) 117:499–511
DOI 10.1007/s00122-008-0794-3

ORIGINAL PAPER

A reference integrated map for cultivated grapevine (*Vitis vinifera* L.) from three crosses, based on 283 SSR and 501 SNP-based markers

Silvia Vezzulli · Michela Troggio · Giuseppina Coppola · Angelica Jermakow · Dustin Cartwright · Andrey Zharkikh · Marco Stefanini · M. Stella Grando · Roberto Viola · Anne-Françoise Adam-Blondon · Mark Thomas · Patrice This · Riccardo Velasco

Received: 29 February 2008 / Accepted: 2 May 2008 / Published online: 27 May 2008

Abstract We have developed an integrated map from five elite cultivars of *Vitis vinifera* L.; Syrah, Pinot Noir, Grenache, Cabernet Sauvignon and Riesling which are parents of three segregating populations. A new source of markers, SNPs, identified in ESTs and unique BAC-end sequences was added to the available IGGP reference set of SSRs. The complete integrated map comprises 1,134 markers (350 AFLP®, 332 BESs, 169 ESTs, 283 SSRs) spanning 1,443 cM over 19 linkage groups and shows a mean distance between neighbouring loci of 1.27 cM. Marker order was mainly conserved between the integrated map and the highly dense Syrah × Pinot Noir consensus map except for few inversions. Moreover, the marker order has been validated through the assembled genome sequence of Pinot Noir. We have also assessed the transferability of SNP-based markers among five *V. vinifera* varieties, enabling marker validation across different genotypes. This integrated map can serve as a fundamental tool for molecular breeding in *V. vinifera* and related species and provide a basis for studies of genome organization and evolution in grapevines.

Communicated by H. Nybom.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-008-0794-3) contains supplementary material, which is available to authorized users.

S. Vezzulli (✉) · M. Troggio · G. Coppola · M. Stefanini · M. S. Grando · R. Viola · R. Velasco
IASMA Research Centre, San Michele a/Adige, Italy
e-mail: silvia.vezzulli@iasma.it

A. Jermakow · M. Thomas
CSIRO Plant Industry, Glen Osmond, Australia

D. Cartwright · A. Zharkikh
Myriad Genetics Inc., Salt Lake City, UT, USA

A.-F. Adam-Blondon
UMR 1165, INRA-CNRS-Université d Evry
Génomique Végétale, Paris Cedex, France

P. This
UMR 1097, DIA-PC Equipe génétique Vigne,
INRA Supagro, Montepellier, France

Introduction

Tree breeding is a time-consuming process due to long reproductive cycles, large plant size and an evaluation period for productivity and quality of between 7 and 20 years. Molecular tools may overcome these difficulties and open the way for new efficient breeding strategies (Staub et al. 1996; Morgante and Salamini 2003). In fact, marker applications to assist breeding are reported both for herbaceous (Huang et al. 1996) and woody plants (Gianfranceschi et al. 1994; Akkurt et al. 2007) as concerns disease resistances. In addition to accelerating the characterization of the genetic basis of complex traits, research activities in the development of molecular markers and methodologies have facilitated the construction of genetic linkage maps. In plants, genetic markers provide the framework in breeding programs via marker-assisted selection (MAS, Mazur and Tingey 1995), map-based cloning (Tanksley et al. 1995), and anchoring physical maps (Mun et al. 2006; Troggio et al. 2007). Among the wood plants poplar (*Populus trichocarpa* Torr. & Gray) has been the first to have its genome sequenced (Tuskan et al. 2006), while among the fruit trees a draft genome sequence of grapevine (*Vitis vinifera* L.) has been published for both a near-homozygous line (Jaillon et al. 2007) and a highly

heterozygous clone (Velasco et al. 2007) of the Pinot Noir variety.

Grapevine is a major fruit tree widespread in temperate climates. Among all species within the *Vitaceae* family, *V. vinifera* ($2n = 38$) is the only species extensively used in the wine industry. Thousands of *V. vinifera* cultivars exist but the global market for wine production is dominated by only a few cultivars, such as Chardonnay, Cabernet Sauvignon, Syrah (Shiraz) and Merlot, which has resulted in the decline of cultivars or landraces that represent an important source of genetic variability for traits of interest (This et al. 2006). During the last decade, the effort of the international grape community has been focused on the development of microsatellite or simple sequence repeat (SSR) markers and the construction of several genetic maps (Lodhi et al. 1995; Dalbò et al. 2000; Doligez et al. 2002, 2006; Grando et al. 2003; Adam-Blondon et al. 2004; Doucleff et al. 2004; Fischer et al. 2004; Riaz et al. 2004, 2006; Fanizza et al. 2005; Lowe and Walker 2006; Di Gaspero et al. 2007; Welter et al. 2007; Xu et al. 2008) covering most of the grapevine genome. In contrast, only few single nucleotide polymorphism (SNP) applications have been reported in the grapevine literature (Owens 2003; Salmaso et al. 2004; Lijavetzky et al. 2007; Pindo et al. 2008; Troggio et al. 2008). Along with insertion/deletion (In/del) events that provide reliable locus PCR-based genetic markers, SNPs represent the most frequent genetic differences within various species including plant species (e.g., Rafalski 2002; Batley et al. 2003). Recently, a large set of SNPs was developed from expressed sequence tag (EST) sequences and BAC-end sequences (BESs) and mapped onto a Syrah × Pinot Noir population, providing a comprehensive grapevine genetic map (994 loci; Troggio et al. 2007). Of these SNP-based markers, 95 were also mapped on an interspecific grapevine cross between the Merzling hybrid and the Teroldego variety (246 loci; Salmaso et al. 2008). Moreover, the Syrah × Pinot Noir map by Troggio et al. (2007) has been implemented with 773 SNP markers, developed "ad hoc" based on the heterozygous Pinot Noir genome contig sequence to support the finishing of the genome assembly (1,767 loci; Velasco et al. 2007).

Here we report the inclusion of the SNP-based markers, derived from BESs and ESTs by Troggio et al. (2007), into a novel integrated reference grapevine genetic map based on segregation data from three different crosses derived from five important winegrape varieties 'Syrah' (S), 'Pinot Noir' (P), 'Grenache' (G), 'Cabernet Sauvignon' (CS) and 'Riesling' (R). The linkage analysis was performed simultaneously on these three segregating populations with a novel mapping software tool taking into account genotyping errors. The marker order was validated through the assembled genome sequence of Pinot Noir (accession numbers AM423240-AM489403 at the EMBL/Genbank/DDBJ databases; Velasco et al. 2007). The main objective of this study was to obtain a saturated "species consensus" map, as an improved genetic tool for landrace and evolution studies, quantitative trait loci (QTL) fine mapping and association studies, enabling the bridging from QTL studies to the genome sequence. We also assess the transferability of SNP-based markers among five elite *V. vinifera* varieties, providing basic information for MAS and future breeding programs in grapevine.

Materials and methods

Mapping populations

The S × P (94 individuals at IASMA) and S × G (94 individuals at INRA) full-sib populations are as described in Troggio et al. (2007) and Adam-Blondon et al. (2004), respectively. The third population was derived from a CS × R cross (87 individuals at CSIRO).

DNA extraction

Genomic DNA was extracted from young leaves for S × G and S × P progeny according to the method described in Adam-Blondon et al. (2004) and in Troggio et al. (2007). Given a low quantity of template, for the CS × R cross DNA extraction was followed by whole genome amplification (WGA). DNA was extracted from young leaf material using the DNeasy 96 Plant Kit (Qiagen). WGA was done using Phi 29 DNA Polymerase (New England Biolabs, USA) with Phi29 Random Hexamer Primers (Fidelity Systems, USA). The WGA involved two steps, the first involved mixing together 1 µl of template DNA (1–40 ng), 1 µl of Phi 29 random hexamer primer, 2.5 µl of 2× annealing buffer (80 mM Tris–HCl, pH 8.0; 20 mM $MgCl_2$) and sterile deionised water to 5 µl total volume. This reaction was heated to 94°C for 3 min and then cooled on ice. In the second step 2 µl of Phi 29 10× Buffer, 2 µl of 4 mM dNTPs, 0.5 µl of Phi 29 DNA Polymerase (5 units) and sterile deionised water were mixed in a 15 µl total volume. Mixtures from steps 1 and 2 were combined and incubated at 30°C for 12 h and heat inactivated for 15 min at 65°C. WGA DNA was stored at −20°C.

Genetic markers: development and analysis

SSR analysis

The core of the integrated map was based on the marker-rich S × P map (Troggio et al. 2007) and the SSRs already mapped on S × G were used to provide the framework (Adam-Blondon et al. 2004). For the purpose of this study,

only the CS × R population was screened with microsatellite markers. Available segregation information about heterozygous SSRs in CS and R parental genotypes was exploited (Riaz et al. 2004) and in total 143 SSRs were tested on CS and R. SSR nomenclature and primer sequences have already been reported by Thomas and Scott (1993; VVS), Bowers et al. (1996, 1999; VVMD); Sefc et al. (1999; VRZAG), Scott et al. (2000; SCU), Di Gaspero et al. (2000; VMC), Adam-Blondon et al. (2004; VMC), Arroyo-Garcia and Martinez-Zapater (2004; VMC), Decroocq et al. (2003; VVC), Merdinoglu et al. (2005; VVI), Doligez et al. (2006; A, B, C, GB, GT and TT), Welter et al. (2007); VMC). Most of these markers are described in the NCBI databases dbSTS and UniSTS (http://www.ncbi.nlm.nih.gov/).

SNP development and analysis

Four hundred and forty-four SNP-based markers polymorphic in S × P were derived from 332 BESs and 112 ESTs. All were analyzed on both S × G and CS × R (two parents and six progeny) following the multiplex minisequencing protocol described in Troggio et al. (2008). In addition, 192 regions identified as monomorphic in S × P (Troggio et al. 2007) were re-sequenced in G, CS and R parental genotypes in order to potentially saturate additional regions along the chromosomes with cross-specific SNP markers. Finally, five coding regions involved in metabolic pathway of interest (*VvBURP1*, *VvSP2*, *VvHB13*, *DFR*, *LDOX2*) were sequenced in S, P, CS, and R to identify gene specific SNP markers for mapping. The identified SNPs were then genotyped on the S × P progeny using the multiplex minisequencing protocol. SNP nomenclature is reported in Troggio et al. (2007). SNP marker information has been deposited into the NCBI SNP database (dbSNP accession numbers from 79088086 to 79088470, Build 130).

Integration strategy

In order to merge linkage groups (LGs) between the S × P, S × G and CS × R maps pairwise "bridges" were built and to expand a non-redundant number of SNP markers, the loci common to all three crosses were split into two groups, distributed equally along each of the 19 S × P LGs. One group was extended to S × G only and the other one to CS × R. Unique (cross-specific) markers were screened against the respective populations.

Construction of the three consensus maps

The observed genotypic frequencies in the progeny were tested against the expected segregation ratio using a χ^2 test. A consensus map was built for each progeny using TMAP, a novel mapping software package taking into account genotyping errors (Cartwright et al. 2007; http://math.berkeley.edu/~dustin/tmap). First, phase was inferred using the Phasing algorithm. Second, the groupings and order were determined by the lowest LOD value, which reduced the number of erroneous linkages to a minimum with a maximum distance of 35 cM Kosambi. Finally, LGs were visualized with MapChart v2.1 (Voorrips 2002).

Construction of the integrated map

Comparison of parental recombination rates

Heterogeneity of recombination rates between marker pairs was tested among all three mapping populations using the "Join-combine groups for map integration" function of JoinMap 3.0 (Van Ooijen and Voorrips 2001). This function lists all the pairs of common markers, their recombination frequencies, LOD values, and calculates which pair is showing significant differences in recombination frequency based on a χ^2 test.

Dataset integration and linkage analysis

The observed genotypic frequencies in the whole dataset were tested against the expected segregation ratio using a χ^2 test. The construction of the integrated map was carried out by means of TMAP (Cartwright et al. 2007; http://math.berkeley.edu/~dustin/tmap). First, the three cross outputs from Phasing were merged into a single dataset. Second, groupings were determined using a minimum LOD threshold of 8.0 and a maximum distance of 35 cM Kosambi. Finally, the LGs were visualized with MapChart v2.1 (Voorrips 2002).

Map validation through the grapevine genome sequence

The 11.2× shotgun sequence of the Pinot Noir genome assembled into a total of 2,093 metacontigs covering 504.6 Mbp (Velasco et al. 2007) was used for the map validation. All the sequence-based mapped markers were aligned along the Pinot Noir assembled genome sequence to validate their order.

Results

Genetic markers: development, analysis and integration strategy

Of 143 SSRs tested on CS and R, 104 were found to be polymorphic in at least one parental genotype and were scored on the CS × R population.

Out of 444 SNP markers (327 BES-based and 117 EST-based) segregating in the S x P population, 289 (65.0%) were also informative for the S x G population and 246 (55.4%) for the CS x R population. Within the 327 BES-based SNP markers tested in S x G, 240 (73.4%) were informative, 76 (23.2%) were monomorphic and 11 (3.4%) failed to amplify, whereas within the 117 EST-based SNP markers, 49 (41.9%) were informative, 30 (25.7%) were monomorphic and 32 (27.4%) failed to amplify. Screening the CS × R population found that of the 327 BES-based SNP markers tested, 191 (58.4%) were informative, 116 (35.5%) were monomorphic and 20 (6.1%) failed to amplify, whereas within the 117 EST-based SNP markers, 55 (47%) were informative, 44 (37.6%) were monomorphic and 18 (15.4%) failed to amplify. In total, 190 of the 246 informative SNP markers could be scored for either the S × G progeny or the CS × R progeny that resulted in a non-redundant set of 95 SNP markers well scattered along the 19 LGs for each population.

Furthermore, the 192 BES regions that were previously found monomorphic (129) or failed (63) in both S and P by Troggio et al. (2007) were sequenced in other cultivars: 12 new SNPs were identified in G, 27 in both CS and R, and 19 in G, CS and R genotypes making a total of 58 new SNPs for these varieties. Finally, a SNP was found in each of the *VvBURP1*, *VvSP2*, *VvHB13*, *LDOX2* and *DFR* genes, allowing their mapping in the S × P population with the majority (63.8%) of the observed nucleotide variations corresponded to transitions (A/G or C/T). In addition, only one (0.2%) of the SNPs (1076L08F; dbSNP accession number 79088251) was found to be a triallelic polymorphism (A/C/T) across the three mapping populations. The segregation type of all SNP markers analysed in the present study are reported in Table S1.

In conclusion, our integration strategy has led to 33.3% of the markers being common to at least two of the crosses (Table 1).

Consensus maps

S × P genetic map

Five SNP markers derived from *VvBURP1, VvSP2*, *VvHB13*, *DFR* and *LDOX2* genes were added to the dataset of 1,006 markers already available (Troggio et al. 2007) for a total of 1,011 loci. During the phase analysis by means of Phasing, automatic trials with the parameter *T* showed that $T = 3.1$ was the lowest value for which the algorithm could find a solution. The genes were found to be linked to the markers of LGs 1, 3, 8 and 18 at the minimum LOD of 8.0 and the maximum distance of 35 cM Kosambi and therefore only these four LGs were rebuilt.

S × G genetic map

An additional 166 SNP markers (30 EST-based and 136 BES-based) were added to the already available dataset of 251 SSR markers (Adam-Blondon et al. 2004) for a total of 417 loci. The phase analysis was conducted at $T = 2.4$ and the linkage analysis was performed at the minimum LOD of 6.0 and the maximum distance of 35 cM Kosambi. Fifty markers were unlinked but the remaining 367 markers (22 EST-based SNP markers, 100 BES-based SNP markers, and 245 SSRs) were ordered in 19 LGs, with a total map length of 1,062 cM.

CS × R genetic map

The linkage analysis included 247 markers consisting of 35 EST and 110 BES-based SNP markers, and 102 SSRs. After the Phasing step ($T = 2.0$), the linkage analysis was performed at a minimum LOD of 3.0 and a maximum distance of 35 cM Kosambi. Five markers were found to be unlinked whereas the remaining 242 markers (35 EST-based SNP markers, 106 BES-based SNP markers, and 101

Table 1 Number and type of molecular markers screened in each mapping population

Markers	S × P[a]	S × G[b]	CS × R[c]	Fully shared	S × P and S × G common	S × P and CS × R common	S × G and CS × R common	S × P unique	S × G unique	CS × R unique	Total in the integrated dataset
SSRs	133	251	102	70	31	0	32	32	125	2	292
BESs	316	136	110	0	130	84	0	103	14	35	366
ESTs	179	30	35	0	29	34	0	124	1	2	190
AFLPs	383	0	0	0	0	0	0	383	0	0	383
Total	1,011	417	247	70	190	118	32	642	140	39	1,231

[a] Syrah × Pinot Noir

[b] Syrah × Grenache

[c] Cabernet Sauvignon × Riesling

SSRs) generated a genetic map with 19 LGs covering 1,157 cM.

Integrated map

Three datasets that were merged into one consisted of the segregation of 1,231 markers characterized in up to 275 individuals (Table 1). In only few instances, there were significant differences in recombination rates of marker pairs between two or three populations (data not shown). Thirty markers showing a highly distorted segregation ($P < 0.001$) were discarded from the linkage analysis. The remaining set of 1,201 markers consisted of 284 SSRs, 351 BES-based SNPs, 183 EST-based SNPs and 383 AFLPs. For each cross, the number of markers based on segregation type and the marker class are reported in Table 2.

The complete integrated map (minimum LOD = 8.0 and maximum distance = 35 cM Kosambi), is comprised of 1,134 loci (350 AFLPs, 332 BES-based SNP markers, 169 EST-based SNP markers and 283 SSRs) spanning 1,443 cM across 19 LGs which corresponds to the haploid number of chromosomes in grapevine (Fig. 1). Sixty-seven markers were discarded because they were linked at lower LOD, affecting the neighboring marker order, or because they significantly increased the LG-end distances.

The mean distance between adjacent loci was 1.27 cM and the largest gap was 18.4 cM found in LG 16. Further linkage analysis with the same parameters was performed without the AFLP markers resulting in the map length of 1,487 cM Kosambi with the mean inter-locus distance of 1.90 cM covering 19 LGs.

Alignment of all maps

Initially, for the three consensus maps we considered microsatellite markers as multi-allelic codominant. The 55 common SSRs showed that the marker order was well conserved: for 53 loci (96.3%) the order of the SSR markers was the same and for two other markers the order was slightly different within an interval <6 cM. Since S × G and CS × R did not share any screened SNP-based markers, their consensus maps were compared through the SNP-rich framework of S × P, with 166 and 145 SNP-

Table 2 Marker segregation type in the three mapping populations

Heterozygous state present in							Total
Marker type	Syrah	Pinot Noir	Syrah and Pinot Noir				
	ab × aa	aa × ab	ab × ab	ab × a0	ab × ac	ab × cd	
EST	67	70	39	0	3	0	179
BES	99	100	117	0	0	0	316
SSR	22	13	21	0	54	23	133
AFLP	124	164	95	0	0	0	383
Total	310	347	269	0	57	23	1,011
Marker type	Syrah	Grenache	Syrah and Grenache				Total
	ab × aa	aa × ab	ab × ab	ab × a0	ab × ac	ab × cd	
EST	17	4	10	0	0	0	31
BES	75	20	41	0	0	0	136
SSR	43	51	15	3	72	66	250
AFLP	0	0	0	0	0	0	0
Total	135	75	66	3	72	66	417
Marker type	Cabernet Sauvignon	Riesling	Cabernet Sauvignon and Riesling				Total
	ab × aa	aa × ab	ab × ab	a0 × ab	ab × ac	ab × cd	
EST	8	18	9	0	0	0	35
BES	41	38	31	0	0	0	110
SSR	19	18		1	43	21	102
AFLP	0	0	0	0	0	0	0
Total	68	74	40	1	43	21	247

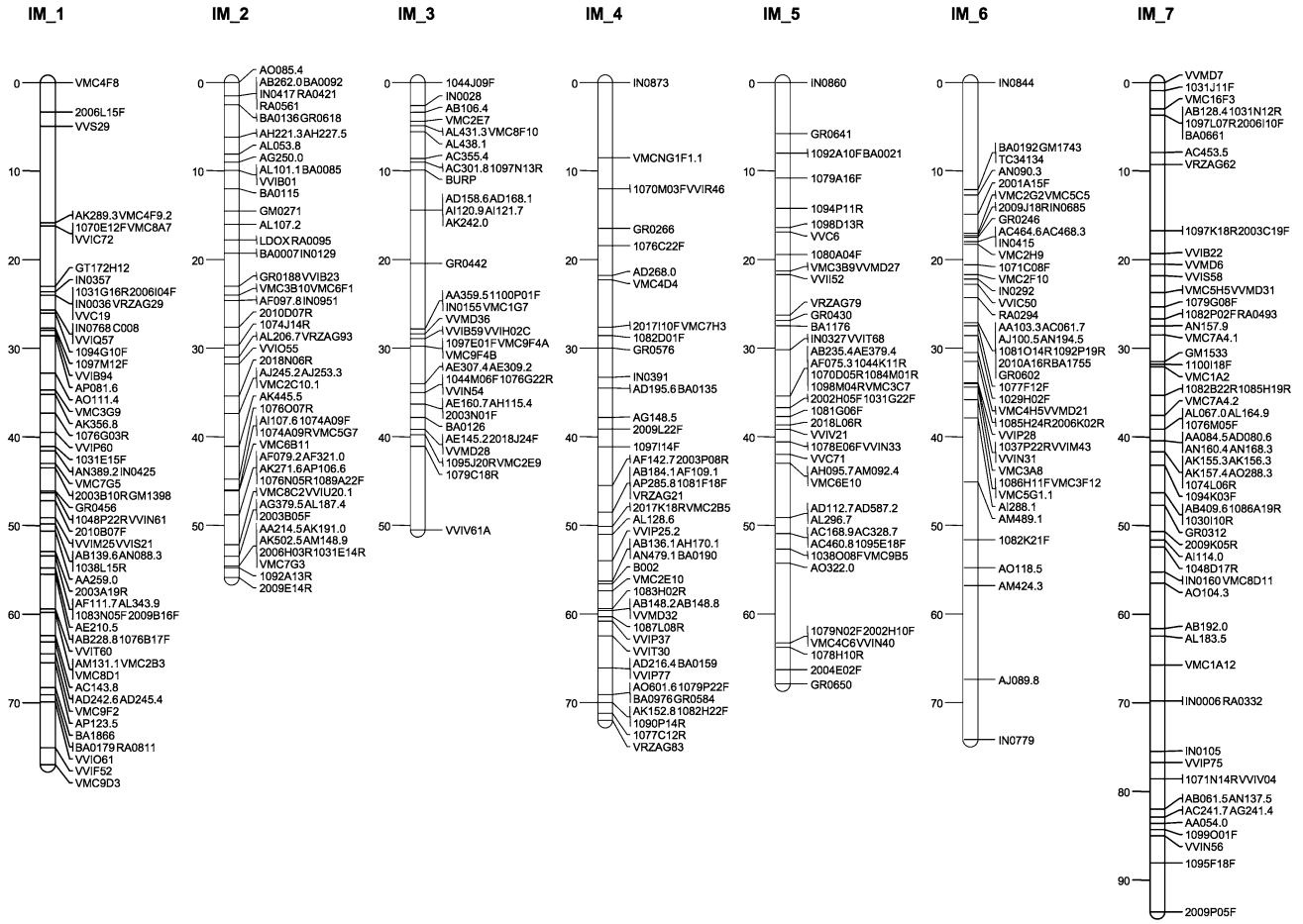

Fig. 1 The integrated map (*IM*) from 5 elite grape cultivars represented by 1,134 mapped loci covering 19 linkage groups

based markers, respectively. Marker order was generally consistent and LG lengths of the three consensus maps were comparable to each other when taking into account the regions covered by homologous markers (Fig. 2 and S1).

Subsequently, the integrated map was compared to the dense consensus map of S × P and the marker distance and order were also consistent between homologous LGs. Compared to the S × P dense map, the integration resulted in similar LG lengths when saturating already covered regions, and only in few cases (LGs 3, 9, 14, and 19) it provided a larger map when covering previously unmapped LG-ends. For few markers, order discrepancies occurred at the end regions of LGs 2, 3, 7, 12, and 16. All pairwise comparisons between the S × P LGs and the integrated LGs are shown in Fig. 3 and S2.

Lastly, the complete integrated map was compared with the integrated map built without AFLP markers. The marker order and distances were generally very well conserved. All pairwise comparisons between LGs built with and without AFLPs are shown in Fig. 3 and S2.

Map validation

The sequences of 671 sequence-associated markers present in the integrated map were aligned to 623 contigs ordered into 178 metacontigs which cover about 360.4 out of 504.6 Mbp of assembled genome sequence (Table S2). This is in a good agreement with independent studies of Lodhi and Reisch (1995) and Thomas et al. (1993) estimating the size of the entire grapevine genome as 475–511 Mbp, which give the range of mapped genomic sequence as high as 70.5–75.9%. The marker order was mainly consistent with the order of genomic contigs (http://genomics.research.iasma.it). Of the 671 sequenced-based markers, 314 belonged to adjacent regions dispersed along the LGs. For 23 pairs of comapping markers, both markers matched to the same genomic contig, which allowed accurate estimation of the physical distance between markers from 9,000 to 79,000 bp. For other 268 pairs, comapping markers matched the sequences of different contigs from the same metacontig. Because of variability in clone sizes used for ordering contigs in metacontigs, the physical distance

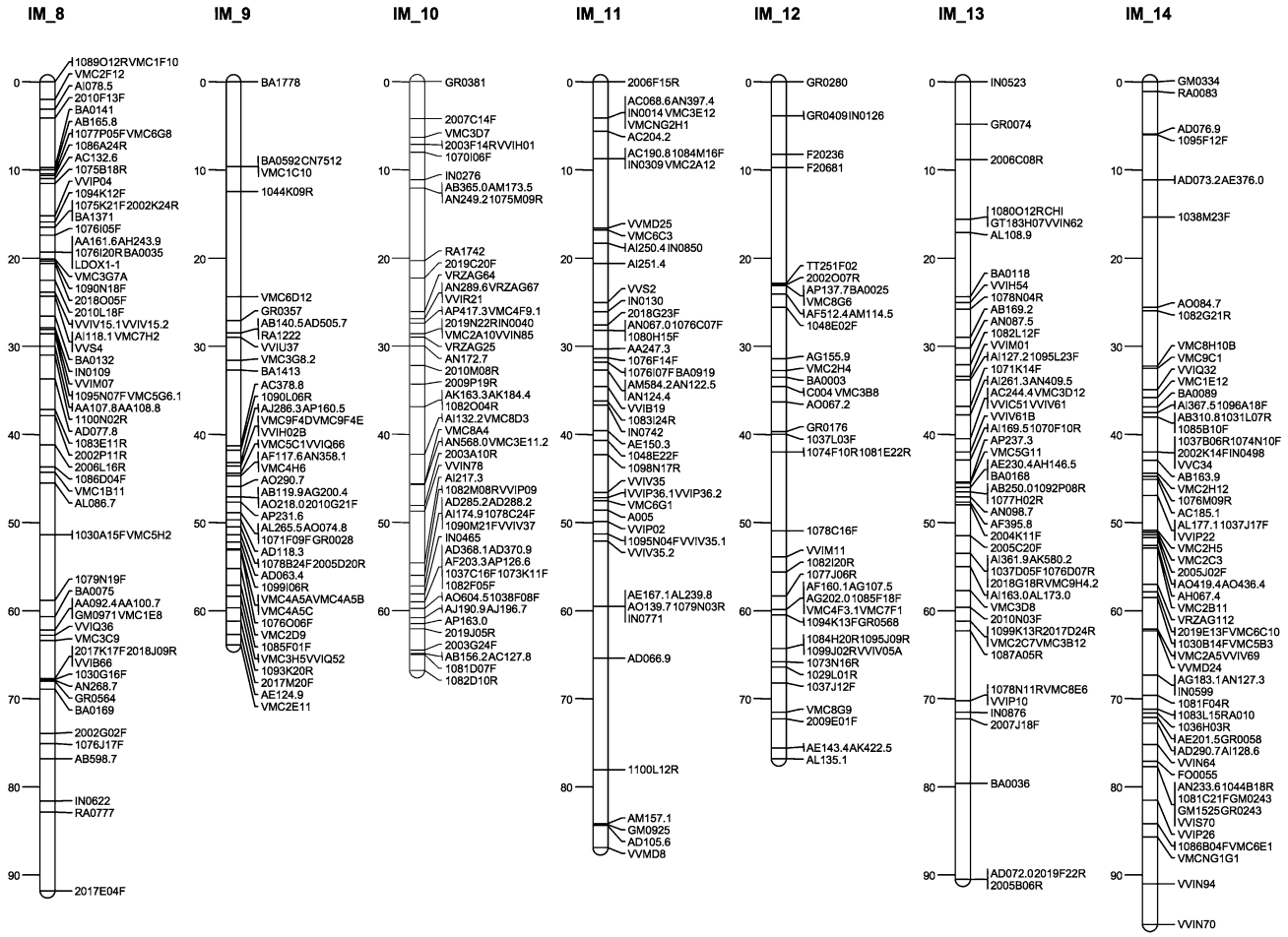

Fig. 1 continued

between markers can be estimated only approximately with an average error of 10% or lower.

Discussion

In this study we report a dense integrated SSR and SNP-based map for five elite cultivars of *V. vinifera*. Mapping based on multiple segregating populations provides several advantages over mapping based on only a single population. In particular, the map integration results in an increased loci density and effective population size, providing a stronger framework for precise mapping of QTLs, association studies based on *linkage disequilibrium* estimation and for alignment with a genome sequence.

Our composite integrated map of 1,134 loci densely covered all 19 LGs of grapevine with the mean distance of 1.27 cM between the adjacent loci. This high marker density is comparable to that of barley where an average distance of 1.3 cM between markers was reached (Qi et al. 1996; Hori et al. 2003). The dense linkage maps are the basis of the fine QTL analysis, map-based gene isolation and integration of physical and genetic maps of genomes (Causse et al. 2004). Moreover, they facilitate comparative study of genomes between distant taxa (Salse et al. 2002). As 250 markers (22% of all mapped loci) corresponded to coding regions, this map has an additional value as a functional integrated map and represents an important genetic tool for future candidate gene studies in grapevine.

Our integration strategy, consisted of extending to the S × G and CS × R populations a non-redundant set of SNP-based markers identified in the S × P population, has led to the construction of pairwise bridges, reducing both costs and time. We also assessed the transferability of SNP-based markers among five *V. vinifera* varieties, enabling marker validation across different genotypes and creating an important background for future applications in marker assisted breeding.

Development and analysis of SNP markers

Until now, microsatellites have represented the basis for cross-talk between maps, due to their multi-allelic nature and abundance. Here, a new source of markers based on

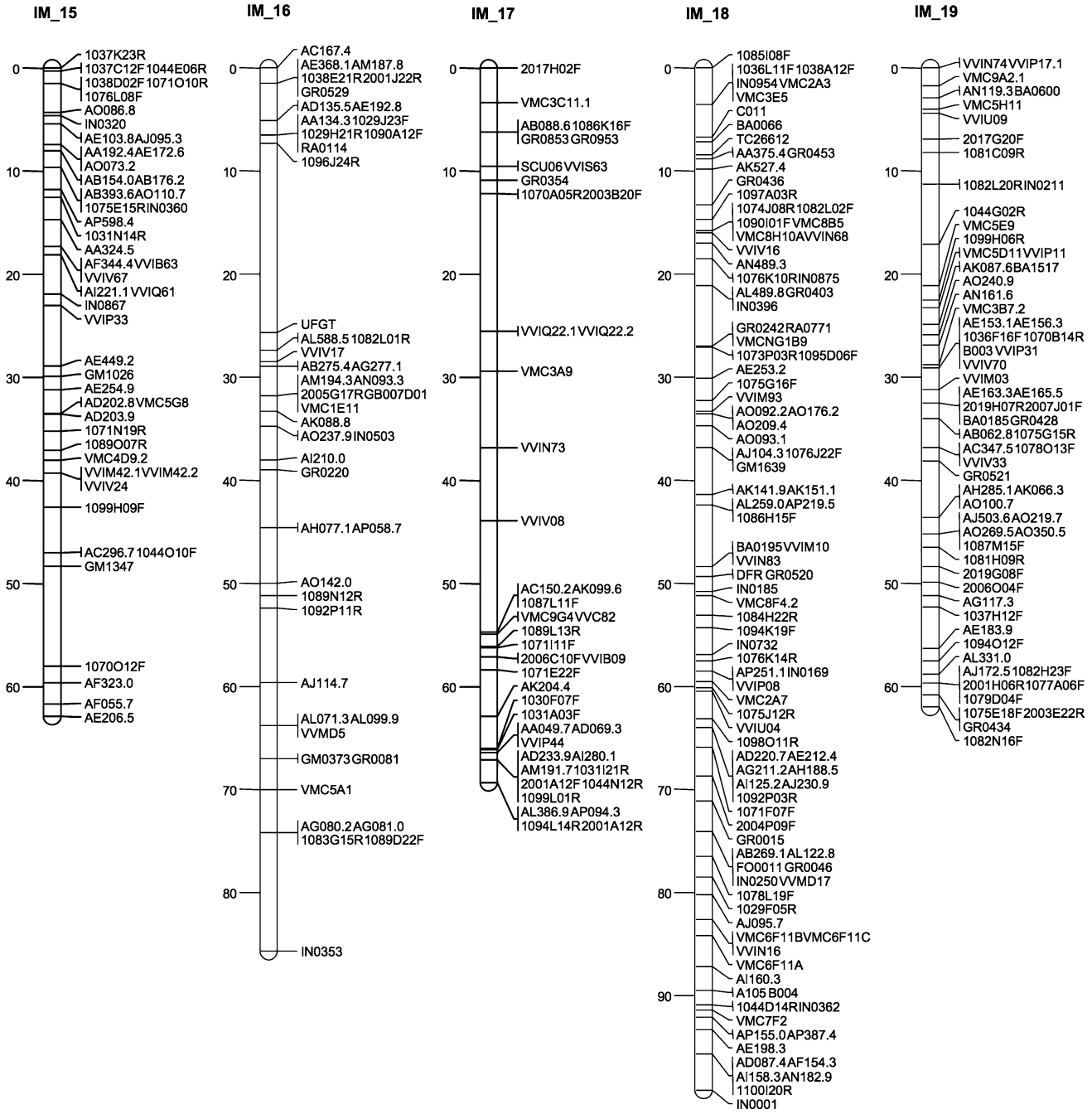

Fig. 1 continued

SNPs, identified in expressed sequences and unique BESs, has been added to the IGGP reference set of SSRs. At the moment, various techniques are available for SNP discovery and genotyping (Gupta et al. 2001). In this study, SNP identification was performed by sequencing and SNP genotyping by multiplex minisequencing (Troggio et al. 2008). This affordable methodology provides up to 7 SNP characterisations per genotype and proved to be highly effective for the transfer of 444 markers previously mapped in SxP together with an additional 58 new markers. The polymorphism detected was mostly due to base transitions which is consistent with previous results in grapevine (Salmaso et al. 2004) and in other organisms (Garg et al. 1999).

Transferability of SNP markers from S × P to S × G (65.0%) was more efficient than for CS × R (55.4%). Since S × P and S × G shared one parental genotype, whereas CS × R did not this difference was expected. Of interest was that the SNP markers from BESs were found to be transferable among mapping populations with a higher efficiency (73.4% in S × G and 58.4% in CS × R) than for

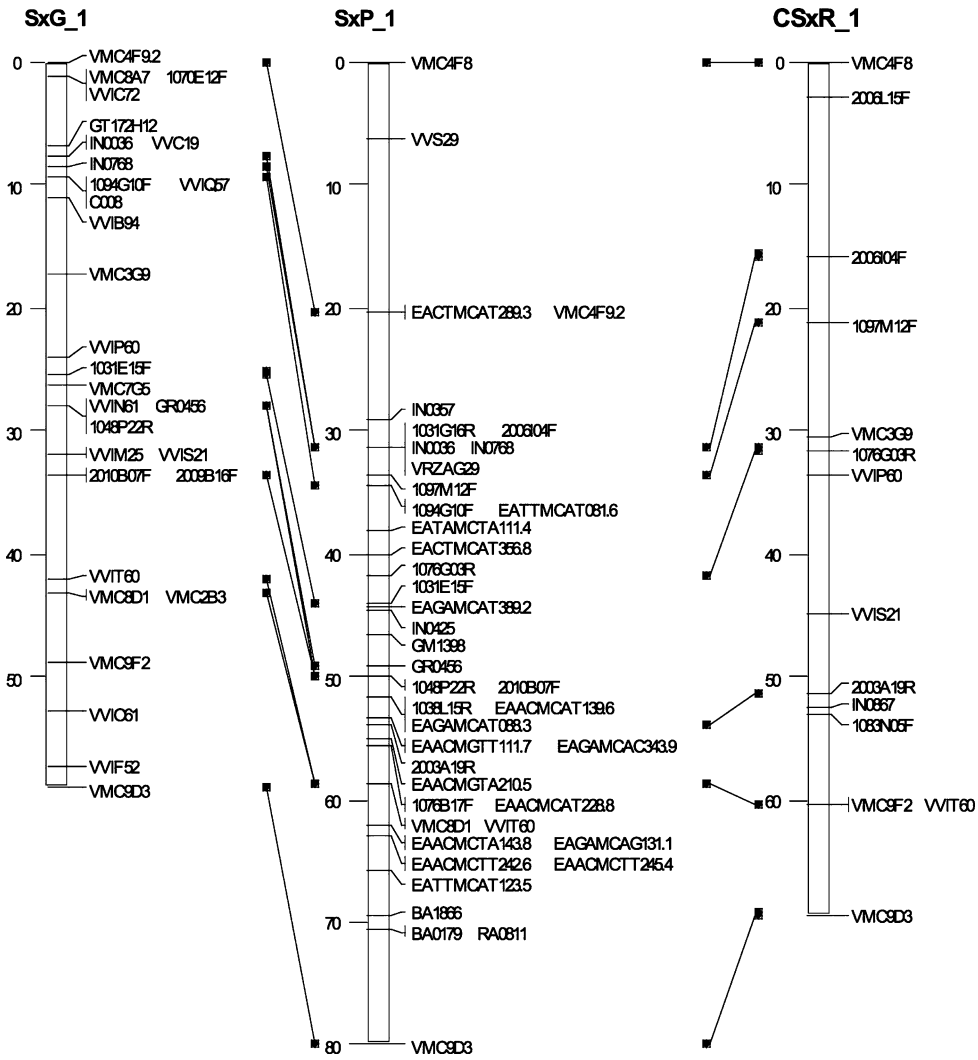

Fig. 2 Comparison of the S × G consensus map, the S × P consensus map, and the CS × R consensus map for linkage group 1. Homologous loci identified. Remaining linkage groups 2–19 are shown in Fig. S1

EST-based SNP markers (41.9% in S × G and 47.0% in CS × R). Furthermore, the number of loci that failed to amplify was similar in S × G (9.7%) and CS × R (8.5%). In particular, the EST-based SNPs showed in both crosses (27.4% in S × G and 15.4% in CS × R) an increased rate of unsuccessful amplification compared to the BES-based SNPs (3.4% in S × G and 6.1% in CS × R). Possible explanations of the higher PCR failure of ESTs could be due to the putative presence of large introns within the PCR amplicons, which would effectively inhibit the amplification, or the positioning of the primer at the intron–exon junctions as was observed in Troggio et al. (2007).

Additionally, among 129 monomorphic BESs of S × P, re-sequencing in G, CS and R identified 58 new SNP markers, while 71 other markers remained monomorphic. Sixty-three problematic BES-based SNP markers identified in S × P also failed in G, CS and R. The monomorphic regions corresponded to the coding sequences which are usually more conserved, whereas the failed sequences were derived from non-coding regions which are mostly variable (Gaafar et al. 2005). Grapevine is a perennial crop which has a high level of heterozygosity (Thomas and Scott 1993). Assuming that the sequences in homozygous regions are accurate, there is still possibility that some results may be artefacts due to preferential PCR amplification of one allele in case of a mismatch between a PCR primer and one of the allelic templates (Walsh et al. 1992).

Marker order and distances

The order of the 889 markers common on the densest consensus map (S × P) and the integrated map was generally consistent except for a few marker inversions at LG-ends. These errors could be explained by a generally increased recombination rate within terminal regions or by recombination rate differences among varieties. Few inversions in the marker order occurred for proximal markers and were mostly caused by changes in the distribution of less informative dominant markers segregating 3:1, such as few SNP-based markers. The linkage analysis was performed

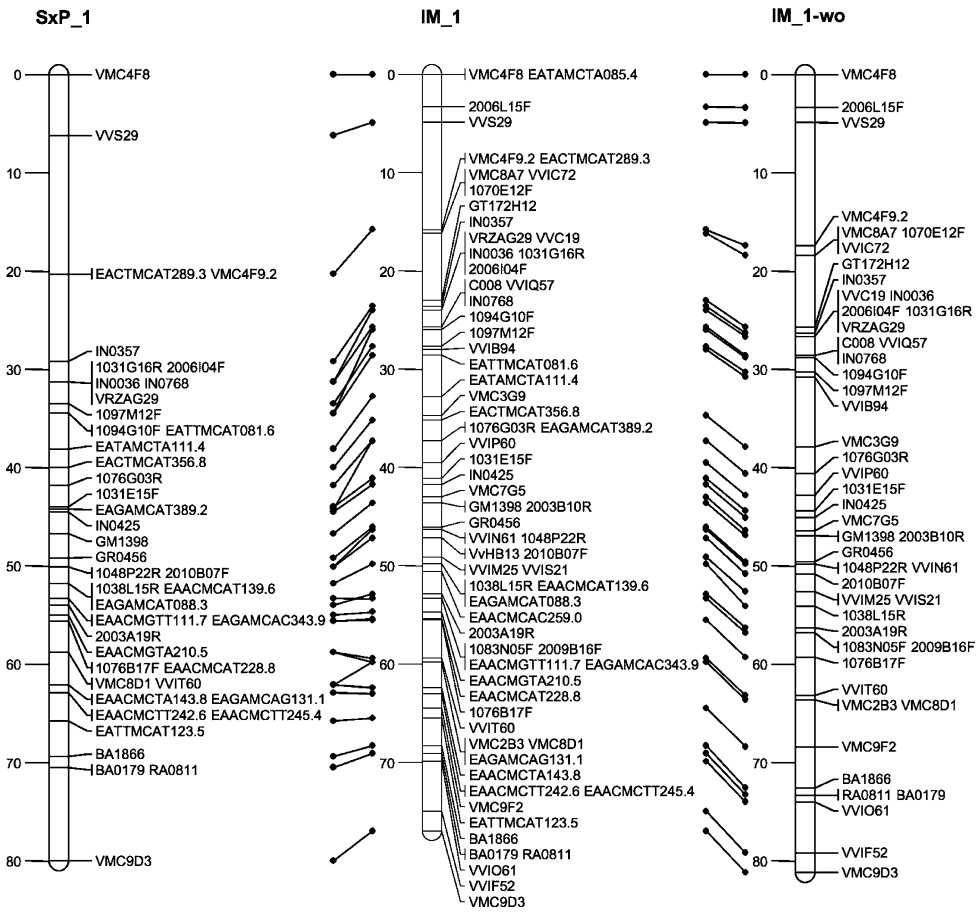

Fig. 3 Comparison of the dense consensus map (S × P), the complete integrated map (*IM*), and the AFLP-less integrated map (IM-wo) for linkage group 1. Homologous loci identified. Remaining linkage groups 2–19 are shown in Fig. S2

using a new mapping software package, TMAP (Cartwright et al. 2007), which takes into account genotyping errors and compensates for distances between close markers. To verify the distance variation with a reduced number of markers, a second integrated map was produced without AFLP markers, as they belonged to only one mapping population (S × P). The reduced integrated map confirmed the validity of the AFLP markers and the marker order within the complete integrated map. Moreover, the distance estimation was generally consistent among LGs. Marker order and distance reliability were supported by the use of TMAP, which considered the genotyping errors. The verification involved removal of every other marker to see how this affected the estimated LG size. With error correction, the LG sizes were very consistent, but in the presence of uncompensated errors removing markers caused distances to shrink (data not shown).

Comparative mapping is a useful technique for investigating chromosomal evolution. It allows importing the genetic information (such as the map positions of qualitative or quantitative traits) obtained in one species to the study of related species. In multiple pedigree mapping studies, the molecular markers are used as "bridges" for merging LGs. For the map comparisons to be meaningful, detection of an orthologous locus in each mapping population can be achieved by searching for DNA sequence homology and for conserved map regions (Brown et al. 2001). Multi-allelic codominant markers, such as microsatellites, are the most efficient for the map comparisons.

The first integrated genetic map of grapevine, comprising 515 loci (502 SSRs and 13 other type PCR-based markers), has been recently published (Doligez et al. 2006). Since their genetic map and our integrated map were constructed by different software, the differences between the two maps were evaluated by using anchor SSR loci. Out of 283 microsatellites positioned on our integrated map, 259 were observed to be shared with the Doligez map. After aligning each LG of our integrated map with the homologous LG from the Doligez map, the microsatellite order appeared to be well-conserved. Therefore, these SSR markers represent a solid framework for our integrated map. LGs 1, 2, 3, 6, 7, 8, 10, 11, 16 and 19 showed a similar (15% length variation) genetic distance. The length of LG 18 was rather different between the two maps; however, the genome coverage was similar between them as estimated by 13 common anchor loci. Our grapevine integrated map was extended by 15 cM at the distal ends of LGs 13 and 15 compared to the Doligez map. In addition, our map had

extra 10 cM regions at the tops of LGs 4, 12, 13, 15 and of 15–18 cM at the tops of LGs 14 and 5, respectively. In contrast, the region of 10 cM at the bottom of LG 9 and the regions of 15 cM at the bottoms of LGs 4 and 18 of the Doligez et al. (2006) map were lacking in our map. In respect to the previous SSR-based integrated map, these findings indicated that our grapevine integrated map has the increased coverage of the whole genome. To date, the highest coverage (477 Mb, 95%) of the grapevine genome was provided by the S × P single cross map (1,767 loci) recently published by Velasco et al. (2007). The order of the 1,006 markers shared between our integrated map and the advanced S × P genetic map of Velasco et al. (2007) was well conserved.

In addition to several inter-species *Vitis* maps, the first genetic linkage map of grape derived from rootstock parents has recently been constructed from a cross of Ramsey (*V. champinii*) × Riparia Gloire (*V. riparia*) (Lowe and Walker, 2006). Given the conservation and the high transferability of SSR markers within the genus *Vitis* it was possible to analyse macro-colinearity between the homologous LGs of their inter-species map and the present integrated map. The order of the common 94 SSRs was well conserved between *Vitis* spp. maps. This shows that the integrated map may be useful as a fundamental tool for molecular breeding not only in *V. vinifera* but also in the related species. To date, only one study (Lijavetzky et al. 2007) has reported on the transferability of SNP markers within *V. vinifera* (*sativa* and *sylvestris* subsp.), while no information has been available for the non-*vinifera* species.

In conclusion, the integrated map represents a substantial resource for molecular breeding programs, as well as trait and QTL marker association. The information reported here may also be useful for comparing the genomes of related species and for supporting the functional genomics studies.

Acknowledgments The authors are grateful to Marco Facci, Jessica Zambanini, Diego Micheletti and Alexandre Fournier-Level for their technical support, and to Alessandro Cestaro, Paolo Fontana and Lorenzo Dematté for their bioinformatics assistance. They also wish to thank Agnès Doligez for helpful discussions. This work was supported in part by the Fondazione Cassa di Risparmio di Trento e Rovereto, the Autonomous Province of Trento, CSIRO Plant Industry and the Commonwealth Cooperative Research Centre Program, and specifically the Cooperative Research Centre for Viticulture (CRCV) and the Grape and Wine Research and Development Cooperation (GWRDC).

References

Adam-Blondon A-F, 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 109:1017–1027

Akkurt M, Welter L, Maul E, Töpfer R, Zyprian E (2007) Development of SCAR markers linked to powdery mildew (*Uncinula necator*) resistance in grapevine (*Vitis vinifera* L. and *Vitis* sp.). Mol Breed 19:103–111

Arroyo-Garcia R, Martinez-Zapater JM (2004) Development and characterization of new microsatellite markers for grape. Vitis 4:175–178

Batley J, Barker G, O'Sullivan H, Edwards KJ, Edwards D (2003) Mining for single nucleotide polymorphisms and insertions/deletions in maize expressed sequence tag data. Plant Physiol 132:84–91

Bowers JE, Dangl GS, Meredith CP (1999) Development and characterization of additional microsatellite DNA markers for grape. Am J Enol Vitic 50:243–246

Bowers JE, Dangl GS, Vignani R, Meredith CP (1996) Isolation and characterization of new polymorphic simple sequence repeat loci in grape (*Vitis vinifera* L.). Genome 39:628–633

Brown GR, Kadel EE, Bassoni DL, Kiehne KL, Temesgen B, van Buijtenen JP, Sewell MM, Marshall KA, Neale DB (2001) Anchored reference loci in loblolly pine (*Pinus taeda* L.) for integrating pine genomics. Genetics 159:799–809

Cartwright DA, Troggio M, Velasco R, Gutin A (2007) Genetic mapping in the presence of genotyping errors. Genetics 176:2521–2527

Causse M, Duffe P, Gomez MC, Buret M, Damidaux R, Zamir D, Gur A, Chevalier C, Lemaire-Chamley M, Rothan C (2004) A genetic map of candidate genes and QTLs involved in tomato fruit size and composition. J Exp Bot 55:1671–1685

Dalbò MA, Ye GN, Weeden NF, Steinkellner H, Sefc KM, Reisch BI (2000) Gene controlling sex in grapevines placed on a molecular marker-based genetic map. Genome 43:333–340

Decroocq V, Favé MG, 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

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

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 114:1249–1263

Doligez A, Bouquet A, Danglot Y, Lahogue F, Riaz S, Meredith CP, Edwards KJ, 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

Doligez A, Adam-Blondon A-F, Cipriani G, Di Gaspero G, Laucou V, Merdinoglu D, Meredith CP, Riaz S, Roux C, This P (2006) An integrated SSR map of grapevine based on five different populations. Theor Appl Genet 113:369–382

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

Fanizza G, Lamaj F, Costantini L, Chaabane R, Grando MS (2005) QTL analysis for fruit yield components in table grapes (*Vitis vinifera*). Theor Appl Genet 111:658–664

Fischer B, Salakhutdinov I, Akkurt M, Eibach R, Edwards KJ, Töpfer R, Zyprian EM (2004) Quantitative trait locus analysis of fungal disease resistance factor on a molecular map of grapevine. Theor Appl Genet 108:501–515

Gaafar RM, Hohmann U, Jung C (2005) Bacterial artificial chromosome-derived molecular markers for early bolting in sugar beet. Theor Appl Genet 110:1027–1037

Garg K, Green P, Nickerson DA (1999) Identification of candidate coding region single nucleotide polymorphisms in 165 human genes using assembled expressed sequence tags. Genome Res 9:1087–1092

Gianfranceschi L, McDermott JM, Seglias N, Koller B, Kellerhals M, Gessler C (1994) Towards a marker assisted breeding for resistance against apple scab. Euphytica 77:93–96

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

Gupta PK, Roy JK, Prasad M (2001) Single nucleotide polymorphisms: a new paradigm for molecular marker technology and DNA polymorphism detection with emphasis on their use in plants. Curr Sci 80:524–535

Hori K, Kobayashi T, Shimizu A, Sato K, Takeda K, Kawasaki S (2003) Efficient construction of high-density linkage map and its application to QTL analysis in barley. Theor Appl Genet 107:806–813

Huang N, Angeles ER, Domingo J, Magpantay G, Singh S, Zhang G, Kumaravadivel N, Bennett J, Khush GS (1996) Pyramiding of bacterial blight resistance genes in rice: marker-assisted selection using RFLP and PCR. Theor Appl Genet 95:313–320

Jaillon O, Aury JM, Noel B, Policriti A, Clepet C, Casagrande A, Choisne N, Aubourg S, Vitulo N, Jubin C, Vezzi A, Legeai F, Hugueney P, Dasilva C, Horner D, Mica E, Jublot D, Poulain J, Bruyère C, Billault A, Segurens B, Gouyvenoux M, Ugarte E, Cattonaro 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, Malacrida G, Durand E, Pesole G, Laucou V, Chatelet P, Merdinoglu D, Delledonne M, Pezzotti M, Lecharny A, Scarpelli C, Artiguenave F, Pè 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

Lijavetzky D, Cabezas J, Ibanez A, Rodriguez V, Martinez-Zapater J (2007) High-throughput SNP discovery and genotyping in grapevine (*Vitis vinifera* L.) by combining a re-sequencing approach and SNPlex technology. BMC Genomics 8:424

Lodhi MA, Daly MJ, Ye GN, Weeden NF, Reisch BI (1995) A molecular marker based linkage map of *Vitis*. Genome 38:786–794

Lodhi MA, Reisch BI (1995) Nuclear DNA content of *Vitis* species, cultivars and other genera of the *Vitaceae*. Theor Appl Genet 90:11–16

Lowe KM, Walker MA (2006) Genetic linkage map of the interspecific grape rootstock cross Ramsey (*Vitis champinii*) × Riparia Gloire (*Vitis riparia*). Theor Appl Genet 112:1582–1592

Mazur BJ, Tingey SV (1995) Genetic mapping and introgression of genes of agronomic importance. Curr Opin Biotechnol 6:175–182

Merdinoglu D, Butterlin G, Bevilacqua L, Chiquet V, Adam-Blondon A-F, Decrooocq S (2005) Development and characterization of a large set of microsatellite markers in grapevine (*Vitis vinifera* L.) suitable for multiplex PCR. Mol Breed 15:349–366

Morgante M, Salamini F (2003) From plant genomics to breeding practice. Curr Opin Biotechnol 14:214–219

Mun JH, Kim DJ, Choi HK, Gish J, Debellé F, Mudge J, Denny R, Endré G, Oliver Saurat, Dudez AM, Kiss GB, Roe B, Young ND, Cook DR (2006) Distribution of microsatellites in the genome of *Medicago truncatula*: a resource of genetic markers that integrate genetic and physical maps. Genetics 172:2541–2555

Owens CL (2003) SNP detection and genotyping in *Vitis*. Acta Hort 603:139–140

Pindo M, Vezzulli S, Coppola G, Cartwright DA, Zharkikh A, Velasco R, Troggio M (2008) SNP-high-throughput screening in grapevine using the SNPlexTM genotyping system. BMC Plant Biol 8:12

Qi X, Stam P, Lindhout P (1996) Comparison and integration of four barley genetic maps. Genome 39:379–394

Rafalski A (2002) Applications of single nucleotide polymorphisms in crop genetics. Curr Opin Plant Biol 5:94–100

Riaz S, Dangl GS, Edwards KJ, Meredith CJ (2004) A microsatellite marker based framework linkage map of *Vitis vinifera* L. Theor Appl Genet 108:864–872

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

Salmaso M, Faes G, Segala C, Stefanini M, Salakhutdinov I, Zyprian E, Toepfer R, Grando MS, Velasco R (2004) Genome diversity and gene haplotypes in the grapevine (*Vitis vinifera* L.) as revealed by single nucleotide polymorphisms. Mol Breed 14:385–395

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

Salse J, Piégu B, Cooke R, Delseny M (2002) New *in silico* insight into the synteny between rice (*Oryza sativa* L.) and maize (*Zea mays* L.) highlights reshuffling and identifies new duplications in the rice genome. Plant J 38:396–409

Scott KD, Eggler P, Seaton G, Rosetto M, Ablett EM, Lee LS, Henry RJ (2000) Analysis of SSRs derived from grape ESTs. Theor Appl Genet 100:723–726

Sefc KM, Regner F, Turetschek E, Glossl J, Steinkellner H (1999) Identification of microsatellite sequences in *Vitis riparia* and their applicability for genotyping of different *Vitis* species. Genome 42:367–373

Staub JE, Serquen FC, Gupta M (1996) Genetic markers, map construction, and their application in plant breeding. HortScience 31:729–741

Tanksley SD, Ganal MW, Martin GB (1995) Chromosome landing: a paradigm for map-based gene cloning in plants with large genomes. Trends Genet 11:63–68

This P, Lacombe T, Thomas MR (2006) Historical origins and genetic diversity of wine grapes. Trends Genet 22:511–519

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

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

Troggio M, Malacarne G, Coppola G, Segala C, Cartwright DA, Pindo M, Stefanini M, Mank R, Moroldo M, Morgante M, Grando MS, Velasco R (2007) A dense single-nucleotide-polymorphism-based genetic linkage map of grapevine (*Vitis vinifera* L.) anchoring Pinot Noir bacterial artificial chromosome contigs. Genetics 176:2637–2650

Troggio M, Malacarne G, Vezzulli S, Salmaso M, Faes G, Velasco R (2008) Comparison of different methods for SNP detection in grapevine. Vitis 47:21–30

Tuskan GA, Di Fazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, Putnam N, Ralph S, Rombauts S, Salamov A, Schein J, Sterck L, Aerts A, Bhalerao RR, Bhalerao RP, Blaudez D, Boerjan W, Brun A, Brunner A, Busov V, Campbell M, Carlson J, Chalot M, Chapman J, Chen G-L, Cooper D, Coutinho PM, Couturier J, Covert S, Cronk Q, Cunningham R, Davis J, Degroeve S, Déjardin A, de Pamphilis C, Detter J, Dirks B, Dubchak I, Duplessis S, Ehlting J, Ellis B, Gendler K, Goodstein D, Gribskov M, Grimwood J, Groover A, Gunter L, Hamberger B, Heinze B, Helariutta Y, Henrissat B, Holligan D, Holt R, Huang W, Islam-Faridi N, Jones S, Jones-Rhoades M, Jorgensen R, Joshi C, Kangasjärvi J, Karlsson J, Kelleher C, Kirkpatrick R, Kirst M, Kohler A, Kalluri U, Larimer F, Leebens-Mack J, Leplé J-C, Locascio P, Lou Y, Lucas S, Martin F, Montanini B, Napoli C, Nelson DR, Nelson C, Nieminen K, Nilsson O, Pereda V, Peter G, Philippe R, Pilate G, Poliakov A, Razumovskaya J, Richardson P, Rinaldi C, Ritland K, Rouzé P, Ryaboy D, Schmutz J, Schrader J, Segerman B, Shin H, Siddiqui H, Sterky F, Terry A, Tsai C-J, Uberbacher E, Unneberg P, Vahala J, Wall K, Wessler S, Yang G, Yin T, Douglas C,

Marra M, Sandberg G, Van de Peer Y, Rokhsar D (2006) The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). Science 313:1596–1604

Van Ooijen JW, Voorrips RE (2001) JoinMap® version 3.0: software for the calculation of genetic linkage maps. Plant Research International, Wageningen

Velasco R, Zharkikh A, Troggio M, Cartwright DA, Cestaro A, Pruss D, Pindo M, FitzGerald LM, Vezzulli S, Reid J, Malacarne G, Iliev D, Coppola G, Wardell B, Micheletti D, Macalma TM, Facci M, Mitchell JT, Perazzolli M, Eldredge G, Gatto P, Oyzerski R, Moretto M, Gutin N, Stefanini M, Chen Y, Segala C, Davenport C, Demattè L, Mraz A, Battilana J, Stormo K, Costa F, Tao Q, Si-Ammour A, Harkins T, Lackey A, Perbost C, Taillon B, Stella A, Solovyev V, Fawcett JA, Sterck L, Vandepoele K, Grando MS, Toppo S, Moser C, Lanchbury J, Bogden R, Skolnick M, Sgaramella V, Bhatnagar SK, 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

Voorrips RE (2002) MapChart: software for the graphical presentation of linkage maps and QTLs. J Hered 93:77–78

Walsh PS, Erlich HA, Higuchi R (1992) Preferential PCR amplification of alleles: mechanisms and solutions. Genome Res 1:241–250

Welter L, Göktürk-Baydar N, Akkurt M, Maul E, Eibach R, Töpfer R, Zyprian E (2007) Genetic mapping and localization of quantitative trait loci affecting fungal disease resistance and leaf morphology in grapevine (*Vitis vinifera* L.). Mol Breed 20:359–374

Xu K, Riaz S, Roncoroni NC, Jin Y, Hu R, Zhou R, Walker MA (2008) Genetic and QTL analysis of resistance to *Xiphinema index* in a grapevine cross. Theor Appl Genet 116:305–311