

An integrated SSR map of grapevine based on five mapping populations

A. Doligez · A. F. Adam-Blondon · G. Cipriani ·
G. Di Gaspero · V. Laucou · D. Merdinoglu ·
C. P. Meredith · S. Riaz · C. Roux · P. This

Received: 8 December 2005 / Accepted: 19 April 2006 / Published online: 24 June 2006
© Springer-Verlag 2006

Abstract A grapevine (mainly *Vitis vinifera* L., $2n = 38$) composite genetic map was constructed with CarthaGene using segregation data from five full-sib populations of 46, 95, 114, 139 and 153 individuals, to determine the relative position of a large set of molecular markers. This consensus map comprised 515 loci (502 SSRs and 13 other type PCR-based markers), amplified using 439 primer pairs (426 SSRs and 13 others) with 50.1% common markers shared by at least two crosses. Out of all loci, 257, 85, 74, 69 and 30 were

mapped in 1, 2, 3, 4 and 5 individual mapping populations, respectively. Marker order was generally well conserved between maps of individual populations, with only a few significant differences in the recombination rate of marker pairs between two or more populations. The total length of the integrated map was 1,647 cM Kosambi covering 19 linkage groups, with a mean distance between neighbour loci of 3.3 cM. A framework-integrated map was also built, with marker order supported by a LOD of 2.0. It included 257 loci spanning 1,485 cM Kosambi with a mean inter-locus distance of 6.2 cM over 19 linkage groups. These integrated maps are the most comprehensive SSR-based maps available so far in grapevine and will serve either for choosing markers evenly scattered over the whole genome or for selecting markers that cover particular regions of interest. The framework map is also a useful starting point for the integration of the *V. vinifera* physical and genetic maps.

Communicated by A. Charcosset

Electronic Supplementary Material Supplementary material is available for this article at <http://dx.doi.org/10.1007/s00122-006-0295-1> and is accessible for authorized users.

A. Doligez (✉) · A. F. Adam-Blondon · V. Laucou ·
C. Roux · P. This
INRA, UMR DGPC 1097, équipe Génétique Vigne,
bâtiment 6, 2 place Viala, 34060 Montpellier Cedex 1, France
e-mail: doligez@ensam.inra.fr

D. Merdinoglu
UMR Santé de la Vigne et Qualité du Vin, INRA-ULP,
INRA, 28 rue Herrlisheim, 68021 Colmar Cedex, France

G. Cipriani · G. Di Gaspero
Dipartimento di Scienze Agrarie e Ambientali,
University of Udine, via delle Scienze, 208, 33100 Udine, Italy

C. P. Meredith · S. Riaz
Department of Viticulture and Enology,
University of California, Davis, CA, 95616, USA

Present Address:
A. F. Adam-Blondon
INRA-URGV, 2 rue Gaston Crémieux, CP 5708,
91057 Evry Cedex, France

Introduction

Grapevine is a crop of major importance worldwide. In such a long-cycled woody plant species, there is a particular expected benefit in developing molecular markers linked to genes of agronomic interest to assist breeding through early seedling selection. Genetic maps are particularly required for this specific goal, but they also contribute to increase knowledge about the structural features of the grapevine genome.

Several genetic linkage maps have already been published for grape (*Vitis* spp.). The older ones were mainly based on RAPD or AFLP markers and aimed at the detection of QTLs for specific traits (Lodhi et al.

1995; Dalbo et al. 2000; Doligez et al. 2002; Grando et al. 2003; Doucleff et al. 2004; Fischer et al. 2004). More recently, two maps containing mainly SSRs were developed to serve as reference maps (Adam-Blondon et al. 2004; Riaz et al. 2004). SSR markers are highly transferable among grapevine genotypes and a large set of these markers are now publicly available (Thomas and Scott 1993; Bowers et al. 1996, 1999; Sefc et al. 1999; Scott et al. 2000; Di Gaspero et al. 2000, 2005; Pellerone et al. 2001; Lefort et al. 2002; Crespan 2003; Decroocq et al. 2003; Adam-Blondon et al. 2004; Arroyo-Garcia and Martinez-Zapater 2004; Merdinoglu et al. 2005; NCBI UniSTS; The Greek *Vitis* Database, <http://www.oldweb.biology.uoc.gr/gvd/contents/general-info/01.htm>).

Since grapevine is highly heterozygous, most existing maps were based on full-sib populations from crosses between two heterozygous parents. Although individual maps were developed for different purposes, all of them shared the pseudo-testcross mapping strategy (Grattapaglia and Sederoff 1994) and partially overlapping sets of SSR markers. This makes it possible to compare marker order between crosses. For this purpose, joining the already available map information of SSR markers scattered in different experimental populations is labour saving compared to increasing the marker density on a single reference cross. Moreover, the number of markers that can be mapped in a single cross is limited to the number of markers that are present in a heterozygous state in either parent of that cross. Therefore, the integration of mapping data from several crosses on a single integrated genetic map would be useful to determine the relative positions of transferable markers. Such an SSR-based and densely covered integrated map will allow to: (1) select markers spanning any particular region of interest for positional gene cloning or marker-assisted selection; (2) choose a batch of markers homogeneously spread over the genome for building framework genetic maps in any other cross useful for QTL detection or for screening germplasm in linkage disequilibrium analyses; (3) compare the genomic localization of genes/QTLs responsible for a given phenotypic variation in different populations; (4) integrate genetic and physical maps.

We report here the first integrated genetic map of grapevine, including 515 loci and based on segregation data from five different crosses simultaneously analysed with a multipoint maximum likelihood method. The primary goal of this study was to place as many transferable markers as possible relative to each other on a single map, to obtain a general idea, rather than a fine resolution, of the order and the distance among available SSR markers.

Materials and methods

Mapping populations

The first population (A1) was a composite 95 full-sib progeny obtained at INRA Montpellier (France) from two reciprocal crosses between the cultivars Syrah and Grenache: 27 offspring were obtained from Syrah × Grenache and 68 offspring from Grenache × Syrah (Adam-Blondon et al. 2004). The second population (A2) was a 114 self-pollinated progeny obtained at INRA Colmar (France) by selfing the cultivar Riesling (Adam-Blondon et al. 2004). The third population (DG) was a 46 full-sib progeny obtained at the University of Udine (Italy) from a cross between the cultivars Chardonnay and Bianca (Di Gaspero et al. 2005). The fourth population (D) was a 139 full-sib progeny obtained at INRA Montpellier from a cross between two table grape genotypes, MTP2223-27 (Dattier de Beyrouth × 75 Pirovano) and MTP2121-30 (Alphonse Lavallée × Sultanine) (Bouquet and Danglot 1996). The fifth population (R) was a 153 full-sib progeny obtained at the University of Davis (CA, USA) from a cross between the cultivars Riesling and Cabernet Sauvignon (Riaz et al. 2004). Among all the parents of these crosses, only Bianca is not pure *Vitis vinifera*. Based on its pedigree (Villard Blanc × Bouvier; Krizsics Csikasz and Kozma 2002), it contains approximately 21% of non-*vinifera* genetic background originating from other *Vitis* species.

Genetic markers

The nature and number of the molecular markers analysed in each population are given in Table 1. They were mainly SSRs from three large series: VMC (Vitis Microsatellite Consortium, Agrogene SA, Moissy Cramayel, France), VVI (Merdinoglu et al. 2005) and UDV (Di Gaspero et al. 2005). Detailed references for SSR markers are given in Electronic Supplementary Material S1. Most markers had already been used for the genotyping of the populations A1 (Adam-Blondon et al. 2004), A2 (Adam-Blondon et al. 2004) and R (Riaz et al. 2004) in previously published individual maps, whereas only 52 markers out of 205 had already been used for the genotyping of the population D (Doligez et al. 2002) and 115 out of 309 for the genotyping of the population DG (Di Gaspero et al. 2005).

For the newly genotyped markers, the genotyping methods used were as follows: the SSR genotyping methods used are described in Adam-Blondon et al. (2004) and Doligez et al. (2002) for the populations A1, A2 and D, in Riaz et al. (2004) for the population R and in Di Gaspero et al. (2005) for the population DG.

Table 1 Number of markers mapped in five full-sib grapevine populations, when each mapping experiment was processed as an independent pseudo-testcross

	A1	A2	DG	D	R
VMC SSRs (including VRZAG and VMCNG)	103 (27)	56 (1)	141 (4)	116	146 (15)
VVI SSRs	106 (3)	50	68 (4)	59	0
UDV SSRs	0	0	63 (3)	0	0
SCU SSRs	0	0	4	1	4
VVMD SSRs	11 (2)	3	14 (1)	16	13 (4)
VVS SSRs	0 (1)	0	7	3	4
VVC SSRs	6	0	0	0	0
A, B, C, GB, GT and TT SSRs	13	0	0	0	0
VH SSR	0	0	0	1	0
VEST ESTP	0	0	0	0	1
Isozymes	0	0	0	3	0
SCARs	0	0	0	3	0
Genes	4	0	0	2	0
Morphological markers	0	0	0	1	1
Total	243 (33)	109 (1)	297 (12)	205	169 (19)

The number of additional markers for which genotypic data were available but could not be mapped is given in parentheses

Five genes (ADH1, ADH2, ADH3, VVAK and SK) were mapped as CAPS markers. Genomic DNA was amplified with primer pairs 1UTR5F/E4R, 2UTR5F/E4R, 3UTR5F/E4R and ATGdir/nmpc rev in the progeny A1 for ADH1, ADH2, ADH3 and VVAK, respectively, and with primer pairs C2/2UTR3R and speATG/500 rev in the progeny D for ADH2 and SK, respectively. Primers E4R, 1UTR5F, 2UTR5F, 3UTR5F and 2UTR3R were designed by Tesnière and Verriès (2000) and ATGdir, nmpc rev, speATG and 500 rev were designed by R. Pratelli (personal communication) from the sequences CS155766 (GenBank accession number) for VVAK (= VVSOR) and AF359521 for SK (= VVSIRK). Ten picomoles of each primer and Qiagen Taq polymerase were used in each PCR mix. PCR conditions were: 94°C for 4 min, 36 cycles of 94°C for 1 min, 55°C for 1 min and 72°C for 1 min and a final extension of 72°C for 6 min. PCR products were then digested with 3 U of the following restriction enzymes: *Hae*III for ADH1, *Dde*I for ADH2, *Msp*I for ADH3 and *Mse*I for VVAK in the progeny A1 and *Msp*I for ADH2 and *Scr*F for SK in the progeny D. Digested fragments were separated on 2% agarose gels and stained with ethidium bromide.

For SCAR markers A27E and A47D, the polymorphism due to the presence or absence of an amplified band was scored for segregation analysis. The markers were amplified in a 25 µl PCR mix containing 10 mM Tris-HCl (pH 9), 50 mM KCl, 1.5 mM MgCl₂, 0.1% Triton X-100 (Appligen), 0.2 mg/ml non-acetylated BSA, 200 µM dNTPs, 0.4 µM each primer, 0.8 U of Taq

polymerase and 20 ng of template DNA. PCR conditions were: 94°C for 4 min, 36 cycles of 94°C for 1 min, 55°C for 1 min, 72°C for 1 min and a final extension of 72°C for 6 min. Fragments were separated on 6% polyacrylamide gel and stained with silver nitrate, as described in Doligez et al. (2002).

Whenever a marker yielded more than one locus in a given progeny, a composite suffix that included the population code (A1, A2, D, DG, R) and a different letter for each locus was added to the marker name. Yet, whenever a marker was mapped on different linkage groups (LGs) in two or more populations, the marker was assumed to detect multiple loci and the population code of the progeny in which that map location was found was suffixed to the marker name.

Construction of the genetic maps for each individual population

All cross-specific maps were built using CarthaGene 0.999R (de Givry et al. 2005). We constructed cross-specific maps also for the populations for which a map had already been published, not only to include the additional markers used to genotype these populations since then, but also to allow comparison between populations by using the same mapping method for all crosses. At the first run of mapping, parental maps were built independently within each cross, using a double pseudo-testcross strategy (Grattapaglia and Sederoff 1994). A consensus map was then constructed for each population using inferred parental phase data. For each locus, the goodness-of-fit of the observed segregation ratio to the appropriate expected ratio was tested using a χ^2 test for both parental and consensus maps.

For parental and consensus maps of each population, LGs were initially determined with a minimum LOD of 3.0 and a maximum distance of 30 cM. Whenever two LGs known to be separate ones according to the two published reference maps (Adam-Blondon et al. 2004; Riaz et al. 2004) were found to stick together using these thresholds, we separated them before ordering markers within LGs. Conversely, when two or more linkage groups that were expected to merge according to the reference maps split at those thresholds, we forced them into the same group before marker ordering. A raw marker order within each linkage group was first determined using a heuristic procedure that incrementally includes each marker by determining its insertion point as the one that yields the highest log-likelihood (“build 5” command). This raw marker order was then improved using an optimization algorithm called “taboo search” with the “greedy 3 1 1 15” command. Finally, local marker order was refined by testing

all possible marker orders within a sliding window of size 5 (“flips 5 2 1” command).

We discarded all loci showing inconsistent positions between parental and/or consensus maps within each progeny, i.e. loci mapped in one parent only whereas they segregate in both and loci mapped in the middle of a LG in one parent but at the end of the LG in the other parent (which was assumed to be due to genotyping errors). We also removed all loci mapping at the end of any LG (> 15 cM far away from any other marker) and all loci that caused large increases in the distance between flanking loci. Finally, we removed the loci mapped at the end of a LG in a given population while in the middle of the LG in other populations, based on the same assumption of genotyping errors as given previously. Map distances were calculated using the Kosambi function, to allow comparison with already published map distances in grape, because the choice of a given mapping function does not affect the determination of marker order but only the final representation of the map. Linkage groups were numbered LG 1 to LG 19, according to Adam-Blondon et al. (2004). Since the haploid chromosome number of *Vitis* spp. is 19, we used this new numbering in place of the former one based on the International Grape Genome Project (IGGP) agreement which included 20 LGs as in the map of Riaz et al. (2004). In the new designation, the former LGs 13 and 18 of the map of Riaz et al. (2004) are merged and the resulting unique group is now designated LG 18; the former LG 20 in Riaz et al. (2004) is now designated LG 13.

Heterogeneity of recombination rate between marker pairs that were linked at $\text{LOD} \geq 3$ was tested among all five populations using the χ^2 test implemented in Joinmap V2.0 module JMHET (Stam and Van Ooijen 1995).

Construction of the integrated genetic maps

Once consensus maps for each population were built and refined, we constructed an integrated genetic map, merging all five progeny datasets with the “dsmergen” command of CarthaGene under the assumption of homogeneous recombination rate between all populations. With this genetic merging method, a single recombination rate is estimated for each given marker pair based on all available meioses, irrespective of which crosses the genotypic data have been derived from. As a consequence, a consensus distance is obtained in addition to a consensus marker order. The commands and parameters used for marker ordering within each LG were “build 5”, “greedy 1 0 1 20” and “flips 5 2 1”. In addition to the complete integrated

map, a framework-integrated map was also built with marker order supported by a LOD of 2.0, using the “buildfw 2 2 {} 1” command.

A test for random distribution of SSR markers (excluding multi-locus markers) along the complete integrated map was carried out as reported in Cervera et al. (2001), based on the coefficient of dispersion (ratio between the variance and the mean) of the number of markers within each 10 cM interval along the whole map.

A different version of the complete integrated map was also built using the “dsmergor” command to merge datasets, followed by the same “build”, “greedy” and “flips” parameters as above for marker ordering. The “dsmergor” command does not rely on the assumption of homogeneous recombination rate between different populations. With this order merging method, a cross-specific recombination rate is estimated for any given marker pair from each separate data set. Therefore, only a consensus marker order is produced, without the calculation of a consensus distance.

We also constructed a composite map using Joinmap V2.0 (Stam and Van Ooijen 1995), which implements a different method to integrate individual maps, based on weighted averages of recombination estimates in the different populations, with a sequential algorithm for ordering markers within each linkage group. Loci with segregation types “ab × a0” or “a0 × ab” had to be discarded since Joinmap does not handle these marker types. We used the following parameters: no fixed order, Kosambi mapping function, use of pairwise recombination estimates ≤ 0.49 with a $\text{LOD} \geq 0.01$ in the JMMAP module, maximum increase of 5 in the goodness-of-fit measure (“jump”) for any new marker to be added, a “triplet” LOD threshold value of 20.0 for automatically considering as fixed orders those triplets more likely than other orders at LOD 20.0 and a moving window of width 5 for local order refinement with the “ripple” function.

Results

Genotypic data were available for a total of 537 loci in 1–5 mapping populations. The cross-specific consensus maps A1, A2, DG, D and R contained 259, 110, 318, 216 and 172 loci, respectively. The main features of each cross-specific consensus map are summarized in Table 2. The exhaustive list of both mapped and unlinked/discarded loci is supplied as Electronic Supplementary Material S1 and a figure showing all five cross-specific consensus maps is provided as Electronic Supplementary Material S2.

Table 2 Main features of the consensus maps obtained for each of the five grapevine populations

	A1	A2	DG	D	R
Unlinked loci ^a	3	1	4	0	4
Discarded loci					
Because of intra-population inconsistency or mapping problems	26	0	13	0	10
Because of inter-population inconsistency	6	0	0	0	5
Segregation types for mapped loci					
ab × aa	47	0	52	30	38
aa × ab	50	0	71	57	32
ab × ab (dominant)	3	2	0	2	0
ab × ab (codominant)	15	108	18	23	6
a0 × ab	3	0	5	1	0
ab × a0	2	0	9	3	0
ab × ac	72	0	111	64	63
ab × cd	67	0	52	36	33
Total	259	110	318	216	172
Mean number of individuals genotyped per locus (minimum–maximum)	90 (64–95)	91 (82–98)	45 (42–46)	137 (91–139)	144 (101–153)
Map features					
Total map length (cM)	1,445.7	1,207.2	1,238.7	1,153.9	1,570.0
Mean distance between loci	6.0	13.3	4.1	5.9	10.3
Number of gaps > 20 cM	8	21	6	4	16
Number of gaps > 30 cM	0	9	0	1	5
Number of distorted loci at $\alpha = 5\%$ (% distorted loci)	27 (10%)	9 (8%)	32 (10%)	15 (7%)	19 (11%)
LGs containing two or more distorted loci	1, 4, 8, 10, 17	7, 18	1, 5, 11, 14	1, 7, 14	1, 7, 8, 19, 20

^a At LOD ≥ 2.0 and recombination rate ≤ 0.5

The percentage of loci showing distorted segregation was similar among the five populations (7–11%). A few regions were largely distorted, on the maps of A1 (LGs 4, 17) and DG (LGs 1, 5, 14). No region harbouring markers with skewed segregation was concomitantly found in two or more populations.

Out of the 537 genotyped loci, 515 were used to build the complete integrated map (Fig. 1), among which 257, 85, 74, 69 and 30 were mapped in 1, 2, 3, 4 and 5 populations, respectively, and 22 were discarded for different reasons (Table 2 and Electronic Supplementary Material S1). The distribution of the number of individuals genotyped per locus is shown in Fig. 2. Only nine markers could not be mapped on the integrated map (UDV034, UDV097, UDV126, VMC1D10, VRZAG26, VVIP25.1, VVIR06.2, VVIR29, VVIV58). For the 515 mapped loci, the pairwise recombination rate could be compared in 756 cases, corresponding to the number of marker pairs that were simultaneously mapped in at least two populations with a LOD ≥ 3.0 . Thirty-seven marker pairs revealed a significantly heterogeneous recombination rate at $\alpha = 1\%$ (Table 3).

The complete integrated map that was obtained at a minimum LOD of 3.0 and a maximum distance of 30 cM consisted of 19 LGs. Marker VMC5A10 remained unlinked under these constraints, but it could be included into LG 19 by increasing the maximum distance threshold to 50 cM. The total length of the map

was 1,646.8 cM and the mean distance between neighbour loci was 3.3 cM. Only one gap larger than 20 cM remained uncovered by any marker on LG 16. The list of all loci with their map position is given in Electronic Supplementary Material S1. Most discrepancies in marker order between the complete integrated map and individual consensus maps (indicated as grey boxes in Electronic Supplementary Material S2) were found in regions where the local order was unsure at LOD 2.0 in the corresponding individual consensus maps, with only a few exceptions (A1 LG 4, A1 LG 6, A1 LG 7, R LG 7, R LG 8, DG LG 9, A1 LG 10, R LG 11, DG LG 12 and A1 LG 18).

The framework-integrated map (Fig. 1) comprised 257 loci spanning 1,485.1 cM over 19 LGs, with a mean inter-locus distance of 6.2 cM and only one gap larger than 20 cM on LG 16. Locus order in the framework map was the same as in the complete map, except for two locus inversions on LG 9 (in the region where a discrepancy between the complete integrated map and the DG map was found) and LG 13.

The ratio between the variance and the mean of the number of single-locus SSR markers within 10 cM intervals of the complete integrated map was 0.96, showing no evidence for a non-random distribution of loci along this map.

The comparison between the locus order within each LG in the complete integrated maps constructed

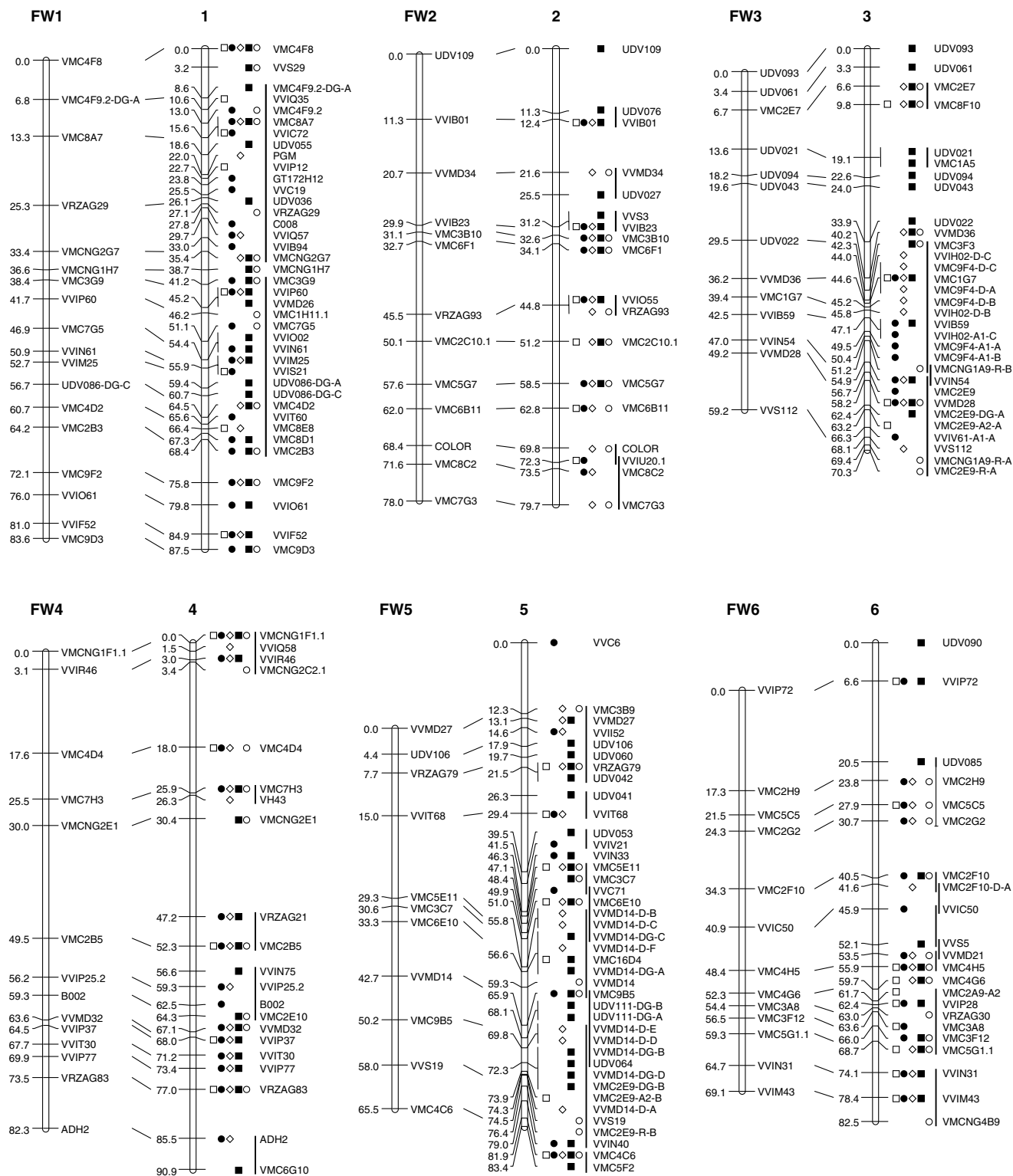


Fig. 1 Complete and framework (FW) integrated maps built from five different grapevine populations. Distances are in cM Kosambi. Symbols at the left of locus names indicate in which individual progeny they were mapped: *open square*, *filled circle*,

open diamond, *filled square* and *open circle* stand for A2, A1, D, DG and R populations, respectively. **Vertical bold lines** indicate groups of loci with local order unsure at LOD 2.0

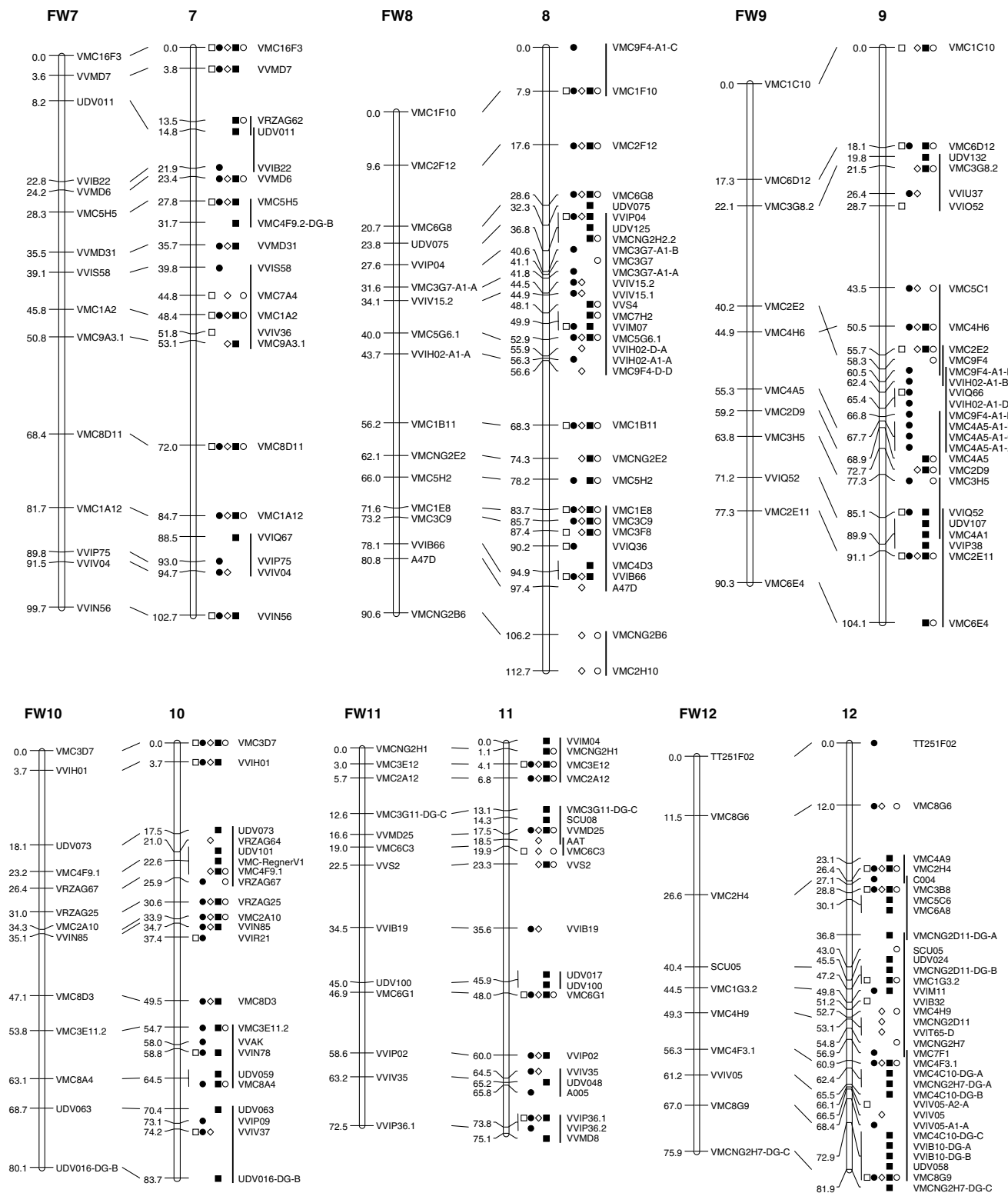


Fig. 1 continued

using CarthaGene and obtained either with “dsmergen” or with “dsmergor” is shown in Electronic Supplementary Material S3. Many marker orders were not

conserved among the two versions of the map, most of which were unsure at LOD 2.0. Conversely, for the framework-integrated maps nearly all orders were

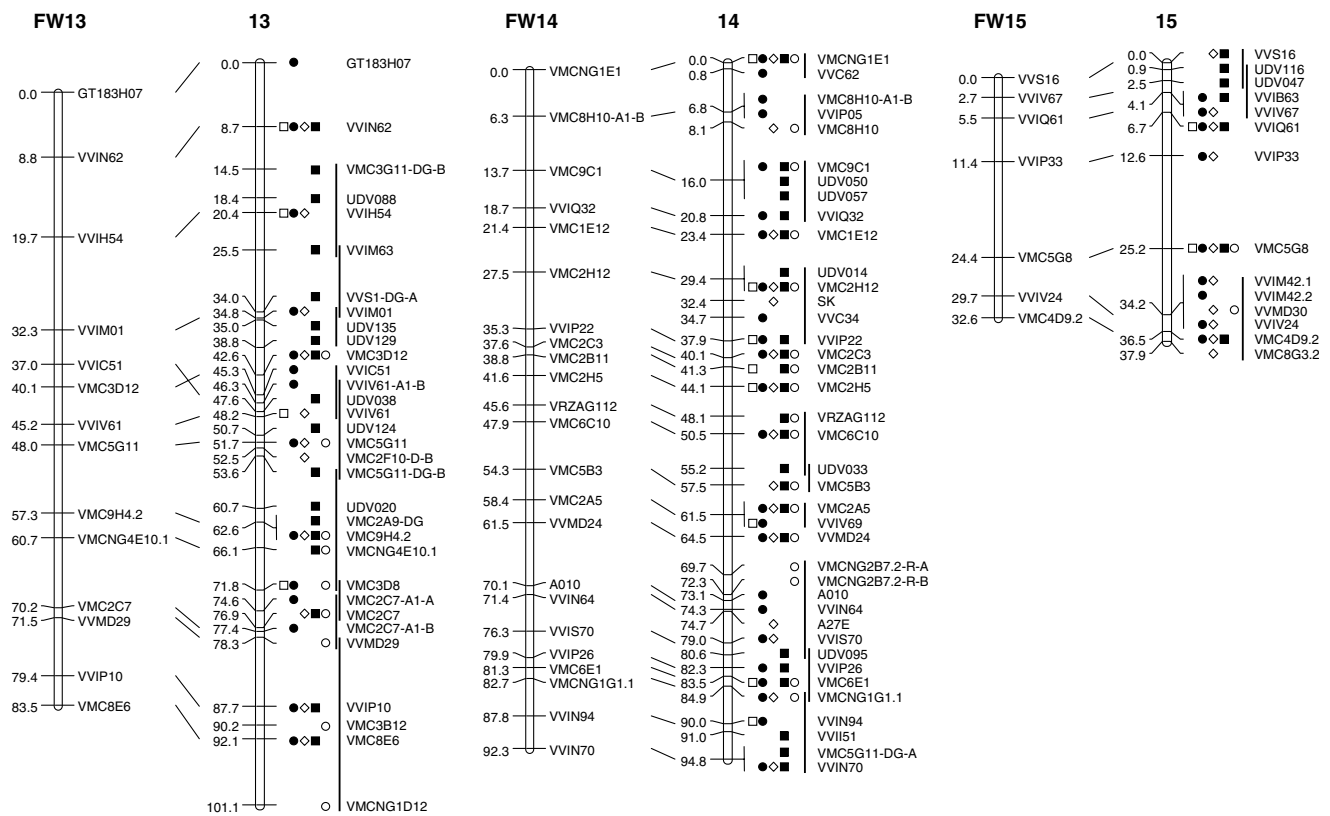


Fig. 1 continued

consistent between the “dsmergen” and “dsmergor” versions, except for a few markers on LG 7 (UDV011-VVIB22), LG 9 (VMC2E11-VMC6E4 and VMC3H5-VVIQ52), LG 11 (VVIB19-UDV100), LG 13 (VVMD29-VVIP10-VMC8E6) and LG 14 (A010-VVIN64-VVIS70). For LG 18, B004 and VVIN16 were included in the framework LG with “dsmergor”, instead of VMC6F11 with “dsmergen”, but orders were not affected.

The integrated map obtained with Joinmap after “round 2” (i.e. after mapping only the loci fitting the “jump” criterion) was consistent with the CarthaGene framework-integrated map (Electronic Supplementary Material S4), except for a number of local inversions of locus order on LG 1 (VMCNG1H7-VMC3G9), LG 2 (VVIB23-VMC3B10-VMC6F1 and VMC5G7-VMC6B11), LG 3 (VMC1G7-VVIB59), LG 4 (VMCNG2E1-VMC4D4), LG 6 (VMC2G2-VMC5C5), LG 9 (VMC3G8.2-VMC6D12), LG 10 (UDV063-VMC3E11.2 and VVIN85-VMC2A10), LG 11 (VMC6G1-UDV100 and VVMD25-VMC3G11-DG-C), LG 14 (A010-VVIS70-VVIP26-VMC6E1), LG 18 (SCC8-VMC7F2) and LG 19 (VVIP34-UDV127 and VVIP31-VMC3B7.2). Local order incon-

sistency between the integrated map obtained with Joinmap after “round 3” (i.e. after mapping also the loci which did not meet the “jump” criterion) and the CarthaGene complete integrated map (obtained with “dsmergen” or “dsmergor”) was much larger (Electronic Supplementary Material S4).

Discussion

Informative content and benefit of an integrated genetic map

We present here an SSR-based integrated genetic map of grapevine, constructed using the segregation data from five full-sib populations of *V. vinifera* (except one population for which the male parent, Bianca, contained about 21% of non-*vinifera* genetic background). The aim of this map was to provide the relative position of a large number of transferable markers initially genotyped in distinct mapping populations and to fill in the major gaps present in each individual map with markers that have been mapped in other mapping experiments.

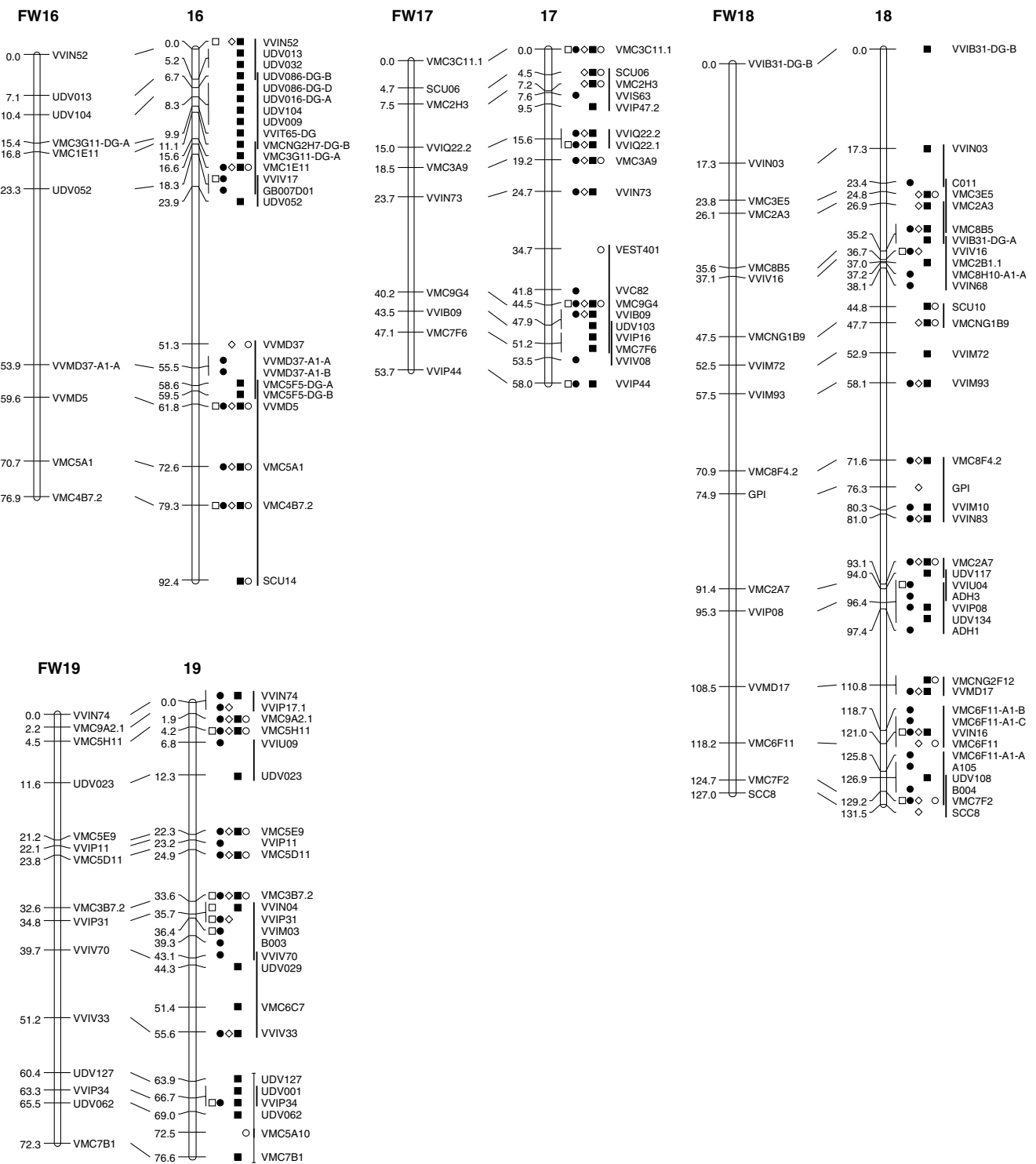


Fig. 1 continued

When compared to the five cross-specific maps presented here and the other grapevine maps containing SSRs that were formerly published (Dalbo et al. 2000; Grando et al. 2003; Doucleff et al. 2004; Fischer et al.

2004), the integrated map densely covered all 19 linkage groups of grapevine, with a mean distance between neighbour loci of 3.3 cM and no evidence of non-random marker distribution over the chromosomes. Out

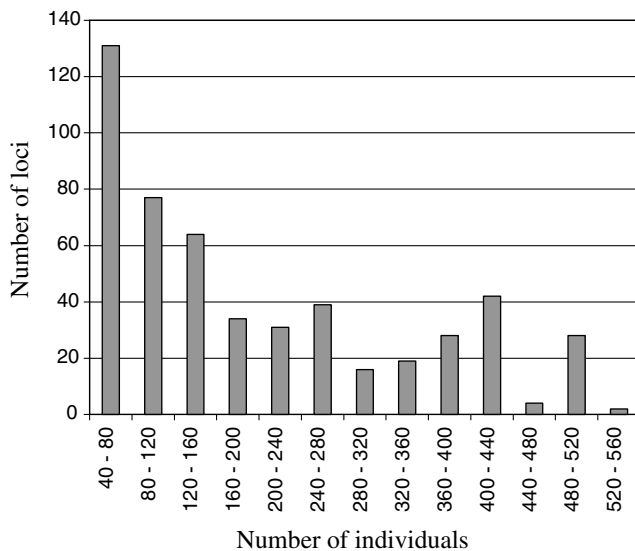


Fig. 2 Distribution of the total number of individuals genotyped per locus, for the 515 loci mapped on the complete integrated grapevine genetic map

of the 502 transferable SSR loci positioned on the integrated map, 16, 45, 9 and 59 were shared between the complete integrated map and the maps of Dalbo et al. (2000), Grando et al. (2003), Doucleff et al. (2004) and Fischer et al. (2004), respectively. Only 11 SSRs that were mapped in Dalbo et al. (2000; VH44, VH444, VRZAG7, VRZAG15, VRZAG26, VRZAG47, VVS13, VVS19 and VVS103) or Grando et al. (2003; SCU11 and VRZAG47) were not present in the integrated map.

In addition to the SSR markers that were already positioned in at least one published map, the integrated map also contained the recently released UDV SSRs (Di Gaspero et al. 2005) whose map position was not known then. In particular, several (AC)*n*-type SSRs of the UDV series proved valuable to saturate some gaps (on LGs 2, 3, 5, 6, 8, 10, 11, 13, 16 and 19) or to extend some LG ends (on LGs 3 and 6) that had remained uncovered with the former SSR series, which were predominantly of (AG)*n*-type. Since microsatellites could display a skewed genomic distribution between gene-rich and heterochromatic regions correlated to the repeat type, mapping of different types of microsatellites may help to better cover the genome with scattered markers. For instance, in *Arabidopsis thaliana*, (AG)*n* repeats are much more frequent in genes than in the rest of the genome and there is a slight reverse tendency for (AC)*n* repeats (Morgante et al. 2002). Our results show that (AG)*n* and (AC)*n* microsatellites provided complementary rather than overlapping map information in grapevine, which is in

favour of a type-specific distribution of SSRs also in this species.

In the integrated map, only one large gap (> 20 cM) remained uncovered on LG 16. This may reflect either a lack of polymorphic markers in a highly homozygous region or alternatively a local increase in recombination rate, combined or not with a low occurrence of SSR markers in that particular region. An answer to this question could possibly be given through the construction of a physical map along this linkage group.

Joining genotypic data sets independently scored in different segregating populations provided us with an effective tool for mapping markers at the scale of whole-genome coverage, even though a disadvantage had to be paid in terms of precise local order of these markers. In order to achieve a reliable fine-scale marker order, it is required to analyse a large number of meioses, which is usually done by genotyping a single large mapping population. Since this process is time-consuming and expensive, fine mapping is most often undertaken only for locally map-based cloning purposes, once a region of interest has already been identified. The complete integrated map presented here is a useful tool for geneticists to identify the markers closely linked to genes/QTLs under study in candidate regions of the grapevine genome. Since they are transferable markers, the SSRs located in a particular region of interest might be easily used in any other cross and more precisely ordered in an extended progeny to fine-map the target locus if required. Alternatively, when searching for novel candidate regions, grapevine geneticists might choose a set of evenly spread markers, like the one provided by the framework-integrated map, for testing the association between markers and a trait of interest at a whole-genome level.

The construction of an integrated map also made it possible to infer some information on the structural features of the grapevine genome. A few regions containing several multi-locus markers could be identified on LGs 3, 5, 9, 12, 13 and 16, predicting the existence of local intra-chromosomal duplications. Segmental intra-chromosomal duplications were found more frequently than genome-wide interspersed duplications in animal and plant genomes (Bailey et al. 2004; Cannon et al. 2004; Zhang et al. 2005) and they were shown to have shaped a gene-rich region also in the grapevine genome (Castellarin et al. 2006). This map also led us to identify marker loci duplicated over different linkage groups, even though more rarely. This feature was not unexpected taking into account the suspected allopolyploid nature of grapevine. Cytological observations of F₁ hybrids between *V. vinifera* (2*n* = 38) and *Muscadinia*

Table 3 Locus pairs showing heterogeneous recombination rate among the five grapevine mapping populations

LG	Locus pair	Populations involved in global comparison	Probability of χ^2 homogeneity test
1	VMC3G9-VMC8D1	A1-DG	0.0083
	VVIM25-VVIP60	D-DG	0.0098
	VMCNG1H7-VMC2B3	DG-R	0.0020
4	VMC2B5-VVIP25.2	A1-D	0.0010
	VRZAG21-VMC7H3	A1-D	0.0022
6	VMC2F10-VVIN31	A1-DG	0.0037
	VMC2H9-VVMD21	A1-D-R	0.0007
	VMC4H5-VMC5C5	A1-A2-D-R	0.0021
	VMC4H5-VVIN31	A1-A2-D-DG	0.0085
	VMC4H5-VVMD21	A1-D-R	0.0000
	VMC4G6-VMC5C5	A2-D-R	0.0073
	VMC4G6-VMC5G1.1	A2-R	0.0017
	VMC1A2-VMC8D11	A1-A2-DG-R	0.0037
	VMC5H5-VVMD31	A1-D	0.0016
7	VMC5H5-VVMD6	A1-D-DG	0.0013
	VMC7A4-VMC8D11	A2-D-R	0.0057
	VMC6G8-VMC1F10	A1-D-DG-R	0.0024
	VMC2H10-VMC3F8	D-R	0.0032
8	VMC5G6.1-VMCNG2H2.2	DG-R	0.0052
	VMC2E11-VMC4H6	A1-D-DG-R	0.0025
	VMC2E11-VMC2E2	A2-D-DG-R	0.0053
	VMC2E11-VMC5C1	D-R	0.0002
	VMC2E2-VMC3G8.2	DG-R	0.0027
9	VMC2A10-VMC3E11.2	A1-DG-R	0.0014
	VMC2A10-VMC8A4	A1-DG-R	0.0018
	VMC3E11.2-VRZAG25	A1-DG-R	0.0007
	VMC8A4-VMC8D3	A1-DG	0.0043
	VMC8A4-VRZAG25	A1-DG-R	0.0005
	VMC2A10-VMC4F9.1	D-DG-R	0.0006
12	VMC3B8-VMC8G6	A1-D-R	0.0016
14	VMC1E12-VMC2H12	A1-D-DG-R	0.0009
15	VMC5G8-VVMD30	D-R	0.0008
16	VVMD5-VMC4B7.2	A1-A2-D-DG	0.0036
	VVMD37-VVMD5	D-R	0.0000
	SCU14-VVMD5	DG-R	0.0006
18	VMC8F4.2-VVIN83	A1-D-DG	0.0072
	VMCNG1B9-VMC3E5	D-DG	0.0086

rotundifolia ($2n = 40$) suggested an allopolyploid origin of the *Vitis* genome (Patel and Olmo 1955). These observations were not supported by the results of in situ hybridization of rDNA probes on the *V. vinifera* chromosomes that showed only one pair of satellite chromosomes (Haas and Alleweldt 2000). Thus polyploidization, if confirmed, is likely to be ancient and to have been followed by many chromosome rearrangements (The Arabidopsis Genome Initiative 2000; Paterson et al. 2004). However, the information provided by this map was not sufficient to recognize precise patterns of inter-chromosomal duplication of large blocks of DNA including several markers.

Reliability of marker order in the integrated map

Marker order was generally well conserved for all linkage groups between the integrated map and the indi-

vidual consensus maps separately constructed on the five populations. No evidence of major chromosomal rearrangements in any parental map emerged at the level of resolution provided by these maps, unlike in other plant species (Lombard and Delourme 2001; Loridon et al. 2005). Local inconsistency in marker order could be partially attributed to a sampling bias mainly in the smallest mapping population (DG, 46 individuals), but less likely in larger ones.

In the present study, a relatively large proportion of loci (50.1%) was common to at least two populations, whereas in integrated maps published for other species (reviewed in Electronic Supplementary Material S5) this proportion was frequently lower than 25%. This allowed a few cases of uncertainty in locus order (at LOD 2.0) that were present in the individual maps to be overcome in the integrated map, thanks to an increased number of meioses statistically supporting

the map position of those loci. However, such an improvement of order confidence was rather rare. Conversely, locus order was uncertain in the integrated map for even more marker pairs than those found in the individual maps, likely due to differences in the recombination rate in different populations and/or to the large number of “private” loci (loci mapped only in a single progeny) especially in supposedly duplicated regions.

Some significant differences in recombination rate between common loci simultaneously scored in different crosses were found and could partially explain the ambiguities in local marker order between pairs of individual maps or between a given individual map and the integrated map. Heterogeneous recombination rate is frequently reported in literature and has been attributed to several environmental and genetic factors (Tulseriam et al. 1992; Fatmi et al. 1993; Causse et al. 1996; Doligez et al. 2002; and references therein). In grapevine, no environmental or genetic factor responsible for recombination heterogeneity has been identified yet. The heterogeneous recombination rate did not seem to be associated with any particular parent of the crosses used in this study, and no apparent relationship between the putative genetic distance between grandparents and a suppression of recombination could be hypothesized. Specifically, no genomic region in the DG population exhibited a recognizable suppression of recombination when compared to the other populations, which could have been suspected due to the partially inter-specific (non-*vinifera*) origin of one of the DG parents (Bianca). According to Lenormand and Dutheil (2005), selection among pollen grains at the haploid level could be a factor responsible for the differences in recombination rates between male and female parents, a case frequently reported for angiosperms. However, in the five mapping populations involved here, no systematic trend for differences between male and female recombination could be found (data not shown).

The impact of recombination rate heterogeneity on the reliability of integrated maps was questioned by Beavis and Grant (1991), who concluded that even so integrated maps could still be useful. Heterogeneity in recombination rates detected with Joinmap JMHEt test may contribute to explain the observed differences in marker order that we found between the “dsmergen” and “dsmergor” versions of the integrated map obtained with CarthaGene for LGs 1, 8, 9, 10 and 16. In addition, distorted segregation for markers linked to deleterious/lethal alleles could lead to artefactual differences in the local recombination rate estimated with CarthaGene, while this does not affect the Join-

map JMHEt test since it is based on a chi-square test for independence. Therefore, distorted segregation detected at some loci on LG 1 and LG 5 could indirectly explain the differences in marker order between the “dsmergen” and “dsmergor” versions of the integrated map for those linkage groups. However, some major differences between the “dsmergen” and “dsmergor” versions observed on LGs 3, 5, 12, 13, 14, 16, 18 and 19 could be related neither to significant differences in recombination rate nor to distorted segregation. The presence of multi-locus markers on some of these LGs could be predictive of the occurrence of segmental intra-chromosomal duplications. Alternative loci of a given marker, positioned slightly apart from each other within a repeated region, could have been scored in individual experimental crosses and caused the order inconsistency between the integrated maps built with “dsmergen” and “dsmergor”. Conversely, significant differences in recombination rate and/or segregation distortion were observed on LGs 4, 6, 7, 11, 15 and 17, but they did not affect the consistency of local marker order between the “dsmergen” and “dsmergor” versions of the map.

CarthaGene and Joinmap produced integrated maps with many differences in marker order. These differences were generally independent from the use of “dsmergen” or “dsmergor” options in CarthaGene (Electronic Supplementary Material S4) and did not always coincide with regions of local order unsure at LOD 2.0. The log-likelihood of the marker order obtained for a few LGs with Joinmap after “round 3”, when calculated with CarthaGene, was much lower ($\text{LOD} < -30.0$) than the log-likelihood of the best order found with CarthaGene (under the “dsmergen” mode). The simulation of a mapping experiment with an artificial data set led Schiex and Gaspin (1997) to state that CarthaGene performed better than Joinmap for identifying the original (true) order of markers in the integrated maps, under a number of variable experimental conditions such as total number of markers, population size, proportion of common markers and probability of missing data. This increased efficiency of the algorithm implemented in CarthaGene was ascribed both to the use of multipoint maximum likelihood criteria simultaneously applied to the genotypic information derived from all the crosses and to the use of efficient local search techniques for order optimization. However, the simulations performed by Schiex and Gaspin (1997) assumed no interference between crossing-over events (use of Haldane mapping function) and no distorted segregation. Both cases are taken into account and managed accordingly by Joinmap but not by CarthaGene. Those simulations also

assumed no difference in recombination rate between individual populations, no genotyping errors and the use of only fully informative 1:1 segregating markers (classical backcross case). Further simulations are needed to fully compare the efficiency of CarthaGene and Joinmap in determining the true marker order in cases of real biological data. Meanwhile, only those marker orders that were consistently conserved whatever the method used to build the integrated map presented here can be safely relied on. In all cases, an interesting feature of CarthaGene is that it yields the whole set of the best maps found rather than the best map alone, thereby allowing to assess how strongly the best order is supported by the data.

Compared to other published grapevine maps constructed using different mapping data from those involved in the present study, marker order in the integrated map was consistent in most linkage groups with the maps of Dalbo et al. (2000) and Grando et al. (2003). Some inconsistency emerged only for LGs 4 and 7 (Dalbo et al. 2000) and LG 11 (Grando et al. 2003). The number of common SSR markers shared between this integrated map and the map of Doucleff et al. (2004) was too low to compare marker order. More inconsistency of marker order emerged from the comparison with the consensus map published by Fischer et al. (2004), which was the only map built with Joinmap while all formerly published ones were constructed with Mapmaker using the same multipoint maximum likelihood criteria as in CarthaGene. Therefore, part of the inconsistency in marker order could result from the different algorithms used. However, since all the mentioned maps (Dalbo et al. 2000; Grando et al. 2003; Fischer et al. 2004) were based on inter-specific crosses involving at least one non-*vinifera* parent, the occurrence of chromosomal rearrangements between some *Vitis* species cannot be excluded as a possible explanation for marker order discrepancy with the integrated map.

In conclusion, with a mean inter-locus distance of 3.3 cM and 50.1% of common loci simultaneously mapped in at least two independent crosses, the grapevine integrated genetic map presented here reached a high informative content with an acceptable marker order confidence. Therefore, it can serve as a good starting point for adding yet unmapped SSRs, in particular those found in sequenced BAC ends of the available BAC libraries and telomeric markers, for the integration of functional markers and for the alignment with the existing physical maps.

Acknowledgements This research was funded by Génoplante grant C11999076 and INRA for genotyping the A1, A2 and D

populations and for building the integrated map, by the Friuli Venezia Giulia Regional Administration for genotyping the population DG and by the American Vineyard Foundation and the United States Department of Agriculture Viticulture Consortium for genotyping the R population. We thank D. Claux, G. Butterlin, A. Fiori, M. Claverie, M. Dardé, C. Sauzet, C. Bisiaux, F. Lahogue for their participation in the genotyping of different populations, as well as R. Pratelli and M. Thomas for personal communication of primer sequences.

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* 109:1017–1027
- Arroyo-Garcia R, Martinez-Zapater JM (2004) Development and characterization of new microsatellite markers for grape. *Vitis* 4:175–178
- Bailey JA, Church DM, Ventura M, Rocchi M, Eichler EE (2004) Analysis of segmental duplications and genome assembly in the mouse. *Genome Res* 14:789–801
- Beavis WD, Grant D (1991) A linkage map based on information from four F₂ populations of maize (*Zea mays* L.). *Theor Appl Genet* 82:636–644
- Bouquet A, Danglot Y (1996) Inheritance of seedlessness in grapevine (*Vitis vinifera* L.). *Vitis* 35:35–42
- 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
- Bowers JE, Dangl GS, Meredith CP (1999) Development and characterization of additional microsatellite DNA markers for grape. *Am J Enol Vitic* 50:243–246
- Cannon SB, Mitra A, Baumgarten A, Young ND, May G (2004) The roles of segmental and tandem gene duplication in the evolution of large gene families in *Arabidopsis thaliana*. *BMC Plant Biol* 4:10
- Castellarin SD, Di Gaspero G, Marconi R, Nonis A, Peterlunger E, Paillard S, Adam-Blondon A-F, Testolin R (2006) Colour variation in red grapevines (*Vitis vinifera* L.): genomic organisation, expression of flavonoid 3'-hydroxylase, flavonoid 3',5'-hydroxylase genes and related metabolite profiling of red cyanidin-/blue delphinidin-based anthocyanins in berry skin. *BMC Genomics* 7:12
- Causse M, Santoni S, Damerval C, Maurice A, Charcosset A, Deatrick J, de Vienne D (1996) A composite map of expressed sequences in maize. *Genome* 39:418–432
- Cervera MT, Storme V, Ivens B, Gusmao J, Liu BH, Hostyn V, van Slycken J, van Montagu M, Boerjan W (2001) Dense genetic linkage maps of three *Populus* species (*Populus deltoides*, *P. nigra* and *P. trichocarpa*) based on AFLP and microsatellite markers. *Genetics* 158:787–809
- Crespan M (2003) The parentage of Muscat of Hamburg. *Vitis* 42:193–197
- Dalbo MA, Ye GN, Weeden NF, Steinkellner H, Sefc KM, Reich BI (2000) A 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 G (2000) Conservation of microsatellite loci within the genus *Vitis*. *Theor Appl Genet* 101:301–308

- Di Gaspero G, Cipriani G, Marrazzo MT, Andreetta D, Prado Castro MJ, 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* 15:11–20
- 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
- Doucleff M, Jin Y, Gao F, Riaz S, Krivanek AF, Walker MA (2004) A genetic linkage map of grape, utilizing *Vitis rupestris* and *Vitis arizonica*. *Theor Appl Genet* 109:1178–1187
- Fatmi A, Poneleit CG, Pfeiffer TW (1993) Variability of recombination frequencies in the Iowa Stiff Stalk Synthetic (*Zea mays* L.). *Theor Appl Genet* 86:859–866
- Fischer BM, Salakhutdinov I, Akkurt M, Eibach R, Edwards KJ, Töpfer R, Zyprian EM (2004) Quantitative trait locus analysis of fungal disease resistance factors on a molecular map of grapevine. *Theor Appl Genet* 108:501–515
- de Givry S, Bouchez M, Chabrier P, Milan D, Schiex T (2005) CarthaGene: multipopulation integrated genetic and radiation hybrid mapping. *Bioinformatics* 21:1703–1704
- 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
- Grattapaglia D, Sederoff R (1994) Genetic linkage maps of *Eucalyptus grandis* and *Eucalyptus urophylla* using a pseudo-test-cross: mapping strategy and RAPD markers. *Genetics* 137:1121–1137
- Haas HU, Alleweldt G (2000) The karyotype of grapevine (*Vitis vinifera* L.). *Acta Hort* 528:247–255
- Krizsics Csikas A, Kozma P Jr (2002) Characterization of fungus resistant grape varieties and candidate varieties in Pécs area. *Acta Hort* 603:763–765
- Lefort F, Kyvelos CJ, Zervou M, Edwards KJ, Roubelakis-Angelakis KA (2002) Characterization of new microsatellite loci from *Vitis vinifera* and their conservation in some *Vitis* species and hybrids. *Mol Ecol Notes* 2:20–21
- Lenormand T, Dutheil J (2005) Recombination difference between sexes: a role for haploid selection. *PLOS Biol* 3(e63):0396–0403
- Lodhi MA, Daly MJ, Ye GN, Weeden NF, Reisch BI (1995) A molecular marker based linkage map of *Vitis*. *Genome* 38:786–794
- Lombard V, Delourme R (2001) A consensus linkage map for rapeseed (*Brassica napus* L.): construction and integration of three individual maps from DH populations. *Theor Appl Genet* 103:491–507
- Loridon K, McPhee K, Morin J, Dubreuil P, Pilet-Nayel ML, Aubert G, Rameau C, Baranger A, Coyne C, Lejeune-Hènaud I, Burstin J (2005) Microsatellite marker polymorphism and mapping in pea (*Pisum sativum* L.). *Theor Appl Genet* 111:1022–1031
- Merdinoglu D, Butterlin G, Bevilacqua L, Chiquet V, Adam-Blondon AF, Decroocq 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, Hanafey M, Powell W (2002) Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes. *Nat Genet* 30:194–200
- Patel GI, Olmo HP (1955) Cytogenetics of *Vitis*: I. The hybrid *Vitis vinifera* × *V. rotundifolia*. *Am J Bot* 42:141–159
- Paterson AH, Bowers JE, Chapman BA (2004) Ancient polyploidization predating divergence of the cereals, and its consequences for comparative genomics. *Proc Natl Acad Sci USA* 101:9903–9908
- Pellerone FI, Edwards KJ, Thomas MR (2001) Grapevine microsatellite repeats: isolation, characterisation and use for genotyping of grape germplasm from Southern Italy. *Vitis* 40:179–186
- Riaz S, Dangi GS, Edwards KJ, Meredith CP (2004) A microsatellite marker based framework linkage map of *Vitis vinifera* L. *Theor Appl Genet* 108:864–872
- Schiex T, Gaspin C (1997) CarthaGene: constructing and joining maximum likelihood genetic maps. In: 5th international conference on intelligent systems for molecular biology, Porto Caras, Halkidiki, Greece, June 1997
- Scott KD, Eggler P, Seaton G, Rossetto 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, Glössl J, Steinkellner H (1999) Identification of microsatellite sequences in *Vitis riparia* and their applicability for genotyping of different *Vitis* species. *Genome* 42:367–373
- Stam P, van Ooijen JW (1995) JoinMap (tm) version 2.0: software for the calculation of genetic linkage maps, Wageningen
- Tesnière C, Verriès C (2000) Molecular cloning and expression of cDNAs encoding alcohol dehydrogenases from *Vitis vinifera* L. during berry development. *Plant Sci* 157:77–88
- The Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408:796–815
- 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
- Tulseriam L, Compton WA, Morris R, Thomas-Compton M, Eskridge K (1992) Analysis of genetic recombination in maize populations using molecular markers. *Theor Appl Genet* 84:65–72
- Zhang LQ, Lu HSS, Chung WY, Yang J, Li WH (2005) Patterns of segmental duplication in the human genome. *Mol Biol Evol* 22:135–141