Construction and characterization of BAC libraries from major grapevine cultivars

Anne-Francoise A.-F. Adam-Blondon, Anne Bernole, G. Faes, Didier
Lamoureux, Stephanie S. Pateyron, M.S. Grando, Michel M. Caboche, R.
Velasco, Paul P. Chalhoub

HAL Id: hal-02681365
https://hal.inrae.fr/hal-02681365

Submitted on 31 May 2020

ORIGINAL PAPER

A.-F. Adam-Blondon · A. Bernole · G. Faes
D. Lamoureux · S. Pateyron · M. S. Grando
M. Caboche · R. Velasco · B. Chalhoub

Construction and characterization of BAC libraries from major grapevine cultivars

Abstract Genome projects were initiated on grapevine (*Vitis vinifera* L., $2n = 38$, genome size 475 Mb) through the successful construction of four bacterial artificial chromosome (BAC) libraries from three major cultivars, Cabernet Sauvignon (Cabernet S), Syrah and two different clones of Pinot Noir (Pinot N). Depending on the library, the genome coverage represented 4.5–14.8 genome equivalents with clones having a mean insert size of 93–158 kb. BAC pools suitable for PCR screening were constructed for two of these BAC libraries [Cabernet S and Pinot N clone (cl) 115] and subsequently used to confirm the genome coverage of both libraries by PCR anchoring of 74 genetic markers sampled from the 19 linkage groups. For ten of these markers, two bands on separate BAC pools were differentiated that could correspond either to different alleles or to a duplication of the locus being studied. Finally, a preliminary assessment of the correspondence between genetic and physical distances was made through the anchoring of all the markers mapped along linkage group 1 of the *V. vinifera* genetic map. A pair of markers, 2.1 cM apart, anchored the same BAC clones, which allowed us to estimate that 1 cM corresponded in this particular region to a maximum length of 130 kb.

A.-F. Adam-Blondon (✉) · A. Bernole · D. Lamoureux
S. Pateyron · M. Caboche · B. Chalhoub
Unité Mixte de Recherches sur les Génomes des Végétaux,
INRA, 2 rue Gaston Crémieux,
5708 91 057 Evry Cedex, France
E-mail: adam@evry.inra.fr
Tel.: +33-160-874534
Fax: +33-160-874549

G. Faes · M. S. Grando · R. Velasco
Istituto Agrario di S.Michele all'Adige,
Via Mach, 1, 38010 S. Michele a/Adige, Italy

Introduction

Vitis vinifera L., the only European representative of the genus *Vitis*, is a crop of great economic importance, with its fruit being mostly used in the preparation of wines and spirits. The gene pool of the cultivated grapevine shows no genetic differentiation, which is consistent with the fact that over a long period of evolutionary time and use its germplasm has been dispersed thoughout all of the grape-growing regions (Bowers et al. 1999; Sefc et al. 2000). The vast majority of cultivars have been propagated vegetatively—in some cases for more than two thousand years—after their first domestication, which was followed by selection from spontaneous crosses (Levadoux 1956; Bowers et al. 1999). Somatic mutations occurred commonly during the generations of vegetative multiplication, leading to the formation of families of "clones" with chimeric structures, among which is the famous Pinot family (Boss and Thomas 2002; Franks et al. 2002; Crespan 2004; Hocquigny et al. 2004). While the molecular markers (simple sequence repeat; SSR) patterns of the clones of one family are highly conserved (Tessier et al. 1999; Franks et al. 2002; Hocquigny et al. 2004), two SSR markers are enough to identify the 44 most important French cultivars, thereby showing the high genetic diversity between cultivars (Tessier et al. 1999; Siret 2001). Grapevine cultivars are highly heterozygous, with an average polymorphic information content (PIC) value of single nucleotide polymorphism-based haplotypes of 0.61 (Salmaso et al. 2004) and an average heterozygosity of SSR markers of 0.77 (Aradhya et al. 2003) to 0.82 (Siret 2001). Consequently, the grapevine genome consists of two very different haplotypes (Siret 2001; Aradhya et al. 2003; Salmaso et al. 2004).

Vitis vinifera has original features, among which are the architecture of its shoots, flowers and fruits and a non-climacteric fleshy fruit, all of which make it very interesting material for comparative functional genomic studies (Boss et al. 2003; Ollat et al. 2002). However,

grape genetics is hampered by a long seed-to-seed cycle and a high susceptibility to inbreeding, both of which have hampered research on the genetics of traits of interest. This problem is being overcome with the development of molecular markers, especially micro-satellite markers, which has enabled the development of genetic maps and the detection of quantitative trait loci (QTLs) for agronomical traits, thereby leading to a better understanding of their genetic determinism (Dal-bó et al. 2001; Doligez et al. 2002; Grando et al. 2003; Fischer et al. 2004; Adam-Blondon et al. 2004; Riaz et al. 2004). Moreover, the small genome size of *V. vinifera* (around 475 Mb, Lodhi and Reisch 1995) and its relative paucity of repetitive sequences (Lodhi and Reisch 1995; Tomkins et al. 2001) are an advantage for gene identification, physical mapping and inter-cul-tivar comparative genome projects.

For such purposes, it is necessary to clone the ge-nomes of important grapevine varieties into bacterial artificial chromosomes libraries (BAC) as this system provides many technical advantages (Sizuya et al. 1992). BAC libraries have now been constructed in a wide variety of plant species, one of which is the hexaploid wheat (Allouis et al. 2003), which has a huge and com-plex genome (16,974 Mb; 130-fold the size of the Ara-bidopsis genome). Although a first BAC library for *V. vinifera* was recently produced and characterized from the Syrah grapevine cultivar (Tomkins et al. 2001), it is important to complete the panel of BAC library resources by constructing additional BAC libraries from other important cultivars and accessions. We report here our effort in constructing and characterizing four other BAC libraries from three important cultivars: a clone of Cabernet Sauvignon (Cabernet S) [chosen as the refer-ence genotype by the International Grape Genome Program (IGGP); http://www.vitaceae.org), two differ-ent clones of Pinot Noir (Pinot N) and a clone of Syrah. Such resources will allow a direct access to genes and alleles for the analysis of sequence variation between the six different haplotypes harbored by these three cultivars.

Materials and methods

BAC library construction

High-molecular-weight (HMW) DNA preparation and BAC library construction (digestion with restriction enzymes, ligation to the vector and bacterial transfor-mation) were carried out according to the procedure described by Chalhoub et al. (2004).

For plant material preparation, around 50 cuttings of each of Cabernet S cl 412, Syrah ENTAV 173, Pinot N clone (cl) 777 and Pinot N cl 115 were rooted in a commercial soil mixture and grown in a growth chamber at 25°(day)/19°C(night) under a 16/8-h (day/night) photoperiod and 50–60% humidity. Before har-vest, cuttings were placed for 1 week in the dark. Very young leaves (1 cm^2) were then harvested, the main vein was cut off and the remaining tissues were immediately frozen in liquid nitrogen. Depending on the material available, different quantities of fresh leaves were used to extract HMW DNA from the four genotypes: 36 g for Cabernet S, 6 g for Pinot cl 777, 10 g for Pinot cl 115 and 6.4 g for Syrah.

Partial digestion of HMW DNA of the different genotypes using *Eco*RI and *Hin*dIII and the recovery of size-fractionated DNA and its electroelution were per-formed as previously described (Chalhoub et al. 2004). Four size-fractions of DNA were recovered from each partial digestion experiment (100–150 kb, 150–250 kb, 250–300 kb, 300–400 kb) and separately ligated into the pIndigoBAC vector (Caltech, Pasadena, Calif.) and prepared for ligation with *Eco*RI or *Hin*dIII according to the single-tube vector preparation method (Chalhoub et al. 2004). A typical ligation experiment consists of 50–200 μg of plant DNA, 8 ng of appropriate vector and 3 U of ligase (Invitrogen, Carslbad, Calif.) in a final volume of 50 μl. Ligated DNA was transformed into *Escherichia coli* DH10B electroMAX-competent cells (Invitrogen). The picking of clones into 384-well plates was carried out using a Qbot robot (Genetix, New Milton, Hampshire, UK) as described by Peterson et al. (2000). Insert sizes were estimated for randomly selected BAC clones from each independent ligation reaction as described by Chalhoub et al. (2004).

Pooling of BAC clones

BAC clones from 48,384-well plates of the Cabernet S BAC library were grown on solid medium. Clones from groups of eight plates, each representing 0.92 genome equivalent, were pooled by plates, lines and columns in water. The cells were precipitated by centrifugation, re-suspended into TE 10-10 and boiled for 5 min at 96°C. The supernatant (called DNA from pools) was diluted to 1/400 in TE 10-0.1 for direct PCR screening. The Pi-not cl 115 BAC library was pooled according to Klein et al. (2000).

PCR primers and screening

We used the universal PCR primers amplifying the CCMP2 and CCMP6 loci (Weising and Gardner 1999) to check the level of contamination with chloroplast (cp) DNA. Both loci were chosen because they are the most distant to each other on the chloroplast genome.

Seventy-four microsatellite loci were chosen from the 19 linkage groups (LGs) of the *V. vinifera* genetic maps published by Riaz et al. (2004) and Adam-Blondon et al. (2004) in order to further assess the genome coverage of the Pinot cl 115 and Cabernet S BAC libraries (Table 1). An additional set of 17 SSR markers (VMC4F8, VMC4F9.2, VMC3G9, VMC7G5, VMC8D1, VMC2B3, VMC9F2, VVIC72, VVIQ57, VVIB94,

Table 1 Screening of the Cabernet S and the Pinot N cl 115 BAC libraries with SSR markers chosen along the 19 linkage groups of the *Vitis vinifera* map (Adam-Blondon et al. 2004; Riaz et al. 2004)

Linkage group	Locus	Origin of the SSR marker	Pinot N Number of hits/marker	Cabernet S Number of hit/marker	Cabernet S Number of hits for band 1[a]	Cabernet S Number of hits for band 2[a]
1	VMC8A7	VMC consortium	5	3		
1	VMC9D3	VMC consortium	6	3		
1	VMC8E8	VMC consortium	2	13		
2	VMC6F1	VMC consortium	1	6		
2	VVMD34	Bowers et al. 1996, 1999	10			
2	VMC5G7	VMC consortium	1	10		
2	VMC8C2	VMC consortium		4		
3	VVMD36	Bowers et al. 1996, 1999	6			
3	VVMD28	Bowers et al. 1996, 1999	4			
3	VMC1G7	VMC consortium		2		
3	VMC8F10	VMC consortium	15	4		
3	VMC2E9	VMC consortium		9	5	4
4	VVMD32	Bowers et al. 1996, 1999	3			
4	VMC7H3	VMC consortium	1	3		
4	VMC2B5	VMC consortium	4	6	2	4
4	VMCNG1F1.1	VMC consortium		7		
5	VRZAG79	VMC consortium	3	6		
5	VVMD27	Bowers et al. 1996, 1999	6	5	3	2
5	VMC4C6	VMC consortium	2	5		
6	VMC4G6	VMC consortium	0	2		
6	VMC5G1.1	VMC consortium	1			
6	VMC2G2	VMC consortium	4	9		
6	VMC3F12	VMC consortium		13	7	6
7	VVMD6	Bowers et al. 1996, 1999	5			
7	VVMD7	Bowers et al. 1996, 1999	8			
7	VVMD31	Bowers et al. 1996, 1999	1	5		
7	VMC5H5	VMC consortium		1		
7	VMC1A2	VMC consortium		4		
8	VMC1F10	VMC consortium	4	6		
8	VVS4	Thomas and Scott 1993	3			
8	VMC1B11	VMC consortium	3	14		
8	VVIB66	Merdinoglu et al. 2005		13		
9	VMC1C10	VMC consortium	3			
9	VMC2E11	VMC consortium	5	12		
9	VMC5C1	VMC consortium	7	8		
9	VMC2D9	VMC consortium		4	3	1
10	VRZAG25	VMC consortium	0	6		
10	VRZAG67	VMC consortium	7			
10	VMC3D7	VMC consortium		6		
10	VMC3E11.2	VMC consortium		6		
11	VVMD25	Bowers et al. 1996, 1999	4	6		
11	VMC6G1	VMC consortium	8	7		
11	VMC3E12	VMC consortium	3	13	7	5
12	VMC2H4	VMC consortium	5	7	2	5
12	VMC4H9	VMC consortium	1			
12	VMC4F3.1	VMC consortium	11	1		
12	VMC8G6	Scott et al. 2000		6		
13	VMC3D12	VMC consortium	1	6	4	2
13	VMC9H4.2	VMC consortium	5	2		
14	VRZAG112	VMC consortium	5			
14	VVMD24	Bowers et al. 1996, 1999	4			
14	VMC2A5	VMC consortium	3	4		
14	VMC2H5	VMC consortium		6	4	2
14	VMC6C10	VMC consortium		6		
15	VMC4D9.2	VMC consortium	1	6		
15	VMC5G8	VMC consortium	3	7		
15	VVIB63			3		
16	VMC4B7.2	VMC consortium	1	11		
16	VVMD5	Bowers et al. 1996, 1999	2			
16	VMC1E11	VMC consortium	8	18		
16	VMC5A1	VMC consortium		14		
17	SCU06	Scott et al. 2000	4	5		
17	VMC9G4	VMC consortium	1	9		

Table 1 (Contd.)

Linkage group	Locus	Origin of the SSR marker	Pinot N Number of hits/marker	Cabernet S Number of hit/marker	Cabernet S Number of hits for band 1[a]	Cabernet S Number of hits for band 2[a]
17	VMC3A9	VMC consortium	4	1		
18	VMCNG1B9	VMC consortium	3			
18	VMC3E5	VMC consortium	1	7		
18	SCU10	Scott et al. 2000	3	5		
18	VMC2A7	VMC consortium	2			
18	VMC7F2	VMC consortium	0	6		
18	VVMD17	Bowers et al. 1996, 1999	2	3		
18	VVIP08	Merdinoglu et al. 2005		9		
19	VMC5E9	VMC consortium	1	4		
19	VMC9A2.1	VMC consortium	1	6	3	3
19	VMC5H11	VMC consortium	4	11		
	Mean number of hit/locus		3.67	6.05		
	Mean number of hit/band				4	

[a] It was possible to distinguish two bands of different sizes for some of the SSR markers

VVIP60, VVIN61, VVIM25, VVIS21, VVIT60, VVIO61 and VVIF52) from LG 1 (Adam-Blondon et al. 2004) was anchored onto the Cabernet S sub-library in order to make a preliminary assessment of the correspondence between the physical and genetic distances in grape.

The PCR mix used for SSR amplifications was as follows: 2 µl of DNA from the pooled BAC library was added to a 10-µl reaction mix containing 1.6 m M MgCl$_2$, 0.5 U AmpliTaq DNA polymerase (Applied Biosystems, Foster City, Calif.), 400 μM dNTP, 1× Taq polymerase buffer, 1.85 ng/µl of each primer and 20% (v/v) of loading buffer [60% (p/v) sucrose, 5 m M Cresol Red in water]. Amplifications were performed using a GeneAmp PCR System 9700 (PE Applied Biosystems) programmed as follows: one cycle of 5 min at 94°C; then 15 cycles of 20 s at 94°C, 20-s touchdown from 65°C to 51°C with a decrease of 1°C at each cycle, 20 s at 72°C; a final 30 cycles of 20 s at 94°C, 20 s at 50°C and 20 s at 72°C. The program ended with a 7-min step at 72°C.

For screening of the Cabernet S library, PCR products were loaded in 1× TAE buffer onto 4% MetaPhor (Cambrex Bio Science Rockland, Rockland, Me.) agarose gels. Electrophoresis was performed at 300 mA for 45 min at 10°C in 1× TAE buffer. The gels were stained with ethidium bromide and photos were realized using a Gel Doc apparatus (Bio-Rad, Hercules, Calif.). For the Pinot N library screening, classical 3% agarose gels were used.

Results

Construction and characterization of the libraries

HMW DNA was partially digested by HindIII. The Cabernet S and Pinot cl 115 DNA was also partially digested by EcoRI. Following size-selection, partially digested DNA from each size-fraction (100–150 kb, 150–250 kb, 250–300 kb, 300–400 kb) was separately eluted, ligated and transformed into *E. coli*. Transformed clones were arranged in 384-well plates. This resulted in four BAC libraries, each consisting of several sub-libraries with different insert sizes (Table 2). To estimate the mean insert size of each library, we extracted plasmid DNA from 789 clones randomly chosen from the different sub-libraries (Table 2). The pattern of the NotI-digested BAC DNA was typical of a dicotyledonous species (only a few bands produced by the NotI enzyme; Fig. 1). In the Cabernet S, Syrah and Pinot cl 777 libraries, the mean size of the inserts increased across the first three size classes (Table 2): 116.2–124.2 kb in E1 and H1, 135.5–156 kb in E2 and H2 and 154.6–182.3 kb in E3 and H3. The insert sizes of the Pinot cl 115 library remained constant at around 90 kb (Table 2). The sizes of the inserts decreased in the last class size in the libraries in which such clones could be obtained (Cabernet S and Pinot cl 777), with mean sizes ranging from 131 kb to 159.3 kb (Table 3). All of the libraries except the one for Pinot cl 115 showed very few clones (1–6%) containing inserts smaller than 100 kb (Fig. 2), which is an important quality criterion. The total number of BAC clones obtained in each library was 44,544 for Cabernet S, 23,040 for Syrah, 48,384 for Pinot cl 777 and 23,424 for Pinot cl 115, representing, respectively, 13.2, 7.7, 14.8 and 4.5 genome equivalents.

The percentage of empty clones was estimated to 0.5% for the Cabernet S, Syrah and Pinot cl 777 libraries, and 2.2% for Pinot cl 115.

The number of clones containing cpDNA was estimated in the Cabernet S and the Pinot cl 115 library. The CCMP2 and CCMP6 loci were successfully amplified on 30 of the 1,152 clones (2.6%) in the Cabernet S library and in 42 of the 1,152 clones (3.6%) in the Pinot cl 115 library.

To further estimate the genome coverage of the two libraries and their value in specifically amplifying different alleles at a given locus, we anchored on BACs

Table 2 Estimation of the mean insert size in each library

Genotype	Ligation name[a]	Number of clones in the library	Number of minipreps	Mean insert sizes (kb)	Number of genome equivalents[b]
Cabernet Sauvignon	E1	5,760	63	122.0	1.5
	E2	1,536	42	147.3	0.5
	E3	3,456	38	154.6	1.1
	H1	8,448	40	116.2	2.1
	H2	11,520	37	135.5	3.3
	H3	8,448	42	163.0	2.9
	H4	5,376	28	159.3	1.8
Whole library		44,544	290	142.2	13.2[b]
Syrah	H2	17,280	69	156.4	5.7
	H3	5,760	44	163	2.0
Whole library		23,040	113	158.0	7.7[b]
Pinot Noir clone 777	H1	20,736	30	124.2	5.4
	H2	8,448	15	143.7	2.5
	H3	14,976	48	182.3	5.7
	H4	4,224	59	131.0	1.2
Whole library		48,384	152	146.1	14.8[b]
Pinot Noir clone 115	H1	15,744	131	93.9	3.0
	H2	4,608	45	93	0.9
	H3	3,072	48	86	0.5
Whole library		23,424	224	92.7	4.5[b]

[a]The targeted size range during the second sizing was 100–150 kb for ligation names E1 and H1, 150–250 kb for ligation names E2 and H2, 250–300 kb for ligation names E3 and H3 and 300–400 kb for ligation name H4

[b]Calculations were carried out under the assumption that the grape genome size is 475 Mb. The number of clones picked in each ligation and the percentage of empty colonies were taken into account but contamination with chloroplasts was not taken into account
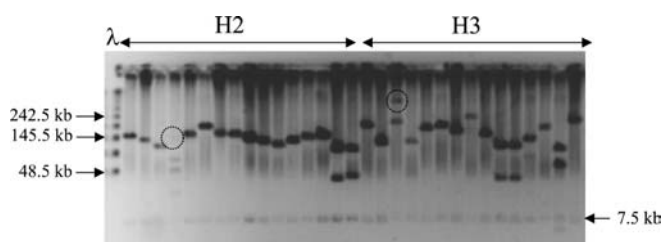

Fig. 1 Insert size of randomly selected BAC clones from ligations of different fractions of partially digested DNA of *Vitis vinifera* cv. Syrah. The BAC plasmid was completely digested by *Not*I. The two *circled areas* represent circular or incompletely digested DNA that was not taken into consideration when calculating the insert size in this experiment. *Arrows* indicate the 7.5-kb common BAC vector band. *λ* is for the pulse field gel electrophoresis lambda size marker

representing six genome equivalents (Cabernet S) and 4.5 genome equivalents (Pinot cl 115) a set of 74 SSR markers that were evenly distributed on the available *V.*

vinifera genetic maps (Table 1). The observed mean numbers of BAC clones anchored per marker were 6.62 for Cabernet S (58 markers anchored) and 3.70 for Pinot cl 115 (57 marker anchored). These values are close to the genome coverage estimated from insert size calculation, always assuming that the markers are mono-locus. The distribution of the number of hits per marker for both libraries is presented in Table 1. We had no case of markers with no hit in the Cabernet S BAC library and three cases with the Pinot cl 115. Based on the Clark and Carbon (1976) method of calculation, the probability of finding no clone carrying a given locus in a BAC library is 0.004 for the Cabernet S subset of clones that was screened and 0.01 for the screened Pinot sub-library.

In the Cabernet S library, ten markers were able to score two different bands. This could correspond to two different alleles as Cabernet S is heterozygous for these mono-locus markers (Adam-Blondon et al. 2004; Riaz

Fig. 2 Insert size distribution of clones from the Cabernet S library (292), Syrah library (113), the Pinot cl 777 library (152) and the Pinot cl 115 library (224)
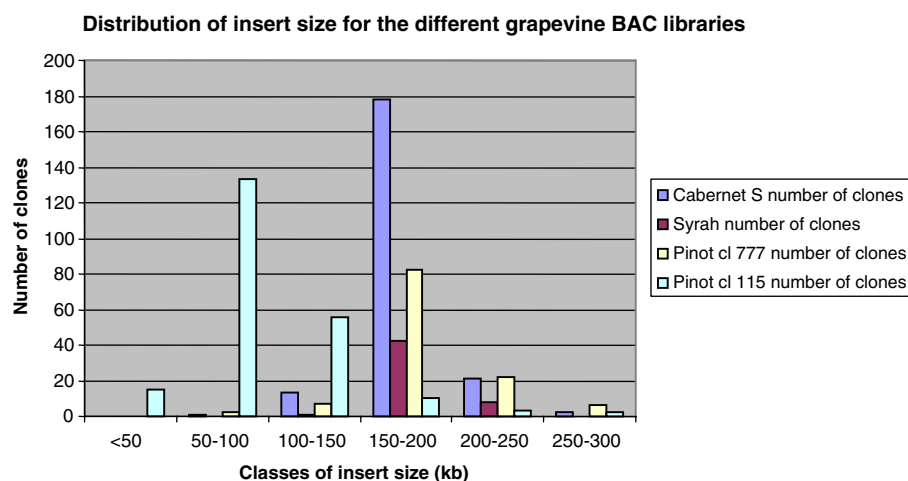
Table 3 Three-dimensional coordinates obtained on the Cabernet S sub-library with the markers from LG 1 of the S×G map (Adam-Blondon et al. 2004)

Marker name	Band[a]	Plate	Row	Column
VMC4F8	a	01, 07	G, I	9, 22
	b	04	D	23
	a	16	G	16
	a	18	C	16
	b	22	N	1
	b	28	B	21
	b	28	B	22
	b	38	D	14
	b	44, 48	A, F	6, 9
	a	46	K	18
VMC4F9.2		21	C	20
		39	C	10
VMC8A7		17	K	6
		41, 47	L, P	17, 24
VVIC72		20	H	16
VVIQ57		05	C, G	3, 14
		17, 20	B, H	4, 24
		30	C	4
		35	C, F	5, 16
VVIB94		03, 04	A, H, P	3, 11, 22
		13	L	9
		21, 24	E, O	5, 7
		27	J, N	16, 17
		33, 35	A, H	4, 23
VMC3G9		02, 03, 05, 06, 07, 08	D, I, J, K, M, O	5, 7, 13, 14, 18, 23
		09	G	7
		17, 18, 23	B, E, P	7, 16, 19, 21
		29	O	19
		34, 37	D, M	2, 21
VVIP60		01	I	14
		12, 13	A, I	9, 15
		18, 19	E, L	14, 20
		41, 45	D, M, O	8, 12, 13
VMC7G5		03	H, P	21, 23
		10, 11, 15	I, K, N	2, 5, 19
		24	N	21
		41, 44	I, M	8
VVIN61		3, 08	H, I, M	2, 4, 14
		15	I	6
		23	I	19
		23	I	9
		26, 31	L, M	16, 21
		41, 47	G, H	2, 10
VVIM25		47	M	3
VVIS21		07	C	23
		12	B	1
		35	C	15
VMC8E8[b]		27	F	8
		3	M	6
		48	E	12
		9	O	1
VVIT60[b]		3	D	14
		15	M	7
		34	J	23
		35	P	23
		44	P	7
VMC8D1[b]		06	I	1
VMC2B3[b]		16	M	4
		17	B	8
		17	J	8
		37	P	16
VMC9F2	a	01, 04	G, K	14, 16
	b	01, 04	A, B	20, 21

Table 3 (Contd.)

Marker name	Band[a]	Plate	Row	Column
	b	30	D, I	10, 16
	a	25	O	1
	b	36, 39	B, L	9, 14
	a	38	H	20
	a	48	J	18
VVIO61		No hit		
VVIF52		02, 04, 06	G, K, O	14, 16, 19
		09, 11, 13	C, F, O	6, 10, 16
		19	J,P	3,12
		46 ,47	A,D,G,O	3, 5, 13, 15
VMC9D3		02	I	9
		26	E	10
		47	I	13

[a]When bands of different sizes were detected, they were labelled "a" for the largest and "b" for the smallest
[b]Markers for which the coordinates obtained after the screening of pools have been checked on individual clones

et al. 2004; Table 1). However, the possibility that some of these markers detected duplicated loci could not be excluded, especially for VMC2E9, VMC3F12 and VMC3E12 for which a significantly higher number of hits than expected was observed. The mean number of positive BAC clones per single band was four, which was slightly more than expected based on our calculation of insert sizes and our assumption that the bands corresponded to different alleles at the same locus.

Correspondence between genetic and physical distances

To evaluate the correspondence between genetic and physical distances along a chromosome, we screened the Cabernet S BAC pools with all of the markers from linkage group 1 (SG1) based on the map published by Adam-Blondon et al. (2004; Fig. 4). The coordinates of clones identified for each marker are shown in Table 3. VVIO61 was not amplified from any of the 18,432 BACs represented in the pools. Two couples of markers showed either exactly the same coordinates for the anchored clones (VMC8D1, VMC2B3) or one possible common coordinate (VMC8E8, VVIT60). These results were checked for both couples of markers on individual clones and were confirmed only for the first pair (Table 3). Thus, in this particular region of SG1, 2.1 cM corresponds to less than 275 kb (the highest size of BAC clones in this library). However, along the linkage group, other couples of markers between which no recombination occurred in the S×G population (VVIC72, VMC8A7 or VVIM25, VVIS21) or couples of markers with similar or lower genetic distances between them, such as (VVIT60, VMD8D1), (VMC4F9.2, VVIC72), (VMC4F9.2, VMC8A7) or (VVIF52, VMC9D3), anchored completely different BAC clones (Table 3). In these cases, the correspondence between physical and genetic distances could be less favorable.

The fact that VVIT60 did not anchor even one clone in common with VMC8D1 is of particular interest as VMC2B3 and VVIT60 are on the two sides of VMC8D1 at the same distance. The coordinates of the clones anchored by the 58 markers evenly distributed on the other linkage groups (Table 1) and anchored on the Cabernet S BAC library were all different with the exception of the three markers chosen on SG1 (data not shown).

Discussion

Four grape BAC libraries were successfully constructed from three major cultivars, Cabernet Sauvignon, Syrah and Pinot Noir (two different clones). These three cultivars are highly different in terms of phenotypic and biochemical characteristics and were assigned to distinct groups through SSR marker-based diversity analysis (Siret 2001; Aradhya et al. 2003). Such BAC libraries should thus provide breeders with access to a variety of alleles for genes of interest.

The Cabernet S BAC library consisted of 44,544 clones representing 13.2 genome equivalents, the Syrah library of 23,040 clones (7.7 genome equivalents), the Pinot cl 777 of 48,384 clones (14.8 genome equivalents) and the Pinot cl 115 of 23,424 clones (4.5 genome equivalents) stored in 384-well plates. Application of the protocol proposed by Chalhoub et al. (2004) successfully resulted in sub-libraries with different mean insert sizes for all of the libraries but that of Pinot cl 115. As already described by Chalhoub et al. (2004) and Noir et al. (2004), the size of the insert obtained was, however, lower than that expected based on size-selection. The percentage of empty clones in all of the libraries was rather low (0.5–2.2%) when compared with other published BAC libraries (Budiman et al. 2000; Yu et al. 2000; Tomkins et al. 2001; Hong et al. 2003; Rajesh et al. 2003; Noir et al. 2004). Conversely, the four libraries showed a rather high rate of chloroplast contamination (2.5–3.6%) when compared to other libraries constructed from plants in other laboratories (Budiman et al. 2000; Yu et al. 2000; Tomkins et al. 2001; Rajesh et al. 2003). One explanation for this may be that some of these chloroplast sequences may correspond to nuclear sequences, as it has been shown that plants have plastid DNA insertions in their nuclear genomes (Shahmuradov et al. 2003). These chloroplast sequences may also represent a valuable resource for the sequencing of the grape plastid genome.

To further check our estimation of the genome coverage, we used 74 genetic markers (SSR), chosen along the 19 linkage groups of the *V. vinifera* map (Adam-Blondon et al. 2004; Riaz et al. 2004), to screen the Cabernet S (subset of 18,432 clones, representing around six genome equivalents) and the Pinot cl 115 libraries. Assuming in a first approximation that most of these markers were mono-locus, we were expecting an average of six and four hits, respectively, per marker,

and we what we observed was 6.62 and 3.70 hits per marker. These results confirm our estimations for the genome coverage of both libraries and indicate that although the cloning was not very successful for the Pinot cl 115 library, there is no strong bias in the genome representation. The observation that one of the 17 additional markers tested on the Cabernet S BAC pools anchored no BACs is still within the range of expectations (1/75 tested markers). The two haplotypes present in Cabernet S may also be well-represented: for ten markers, we detected two bands at the locus, and the mean number of hits per band was four. This result was slightly more than would be expected if each band represented one allele at the locus, and the question that followed was whether it could correspond to different loci amplified by a single primer pair. The loci that were chosen for this study were mono-locus on the two current genetic maps (Adam-Blondon et al. 2004; Riaz et al. 2004). However, the compilation of these two maps together with three others recently showed that VMC2E9 amplifies another locus on group 5 (A. Doligez, personal communication). To date, there is no indication that this is also the case for VMC3F12 and VMC3E12, which also showed a higher number of hits than expected, neither for the other loci tested (A. Doligez, personal communication). Moreover, the sizes of the bands amplified were close in size or identical to those corresponding to the alleles that were mapped in Syrah, and only two bands were observed in this latter genotype on the polyacrylamide gels (Adam-Blondon et al. 2004; data not shown). It is thus highly probable that for most of the nine of the ten remaining loci, the two bands represent the two alleles of the same loci. In all cases it was possible to find at least one positive clone for each band, which is another indication of a correct coverage of the library. Overall, the four BAC libraries represent around 37 *V. vinifera* genome equivalents, and there is a 100% probability of finding a particular sequence among them. The probability of finding a particular allele in one of the two haplotypes of a particular cultivar depends on the size of the related BAC library: 99.8% in Cabernet S, 97.3% in Syrah and 99.9% in Pinot N if we assume that both clones are identical—with the exception of very rare mutations (Franks et al. 2002; Hocquigny et al. 2004).

Based on the last estimation of the grape genome size in centiMorgans (2,200 cM, Adam-Blondon et al. 2004) and in megabases (475 Mbp, Lodhi and Reish 1995), on average, 1 cM should correspond to 216 kb. Two markers located at a distance of less than 1 cM from each other may thus amplify the same BAC clone. To make a first assessment of this correspondence along a chromosome, we decided to anchor all of the markers from SG1 based on the map published by Adam-Blondon et al. 2004. Indeed, this linkage group showed a good marker density: the smallest distance between two markers was 0 cM, the largest distance was 17.5 cM and most of the pairs of markers were spaced at approximately 5 cM from each other (Fig. 3). One pair of

Fig. 3 Physical map obtained
from the anchoring of all the
markers from LG 1 of the S×G
genetic map (Adam-Blondon
et al. 2004). The marker in
italics is distorted (*P* < 0.0001);
those markers that were
ordered at LOD = 2 are
indicated with a *bracket on the
left*; marker distances from the
top are indicated on the *left* in
Kosambi centiMorgans. The
number of anchored clones for
each marker has been deduced
from Table 3. BAC clones are
represented by *grey bars*, all
with the same length (for their
average length). The *black box*
highlights a group of five BACs
anchored with two different
markers
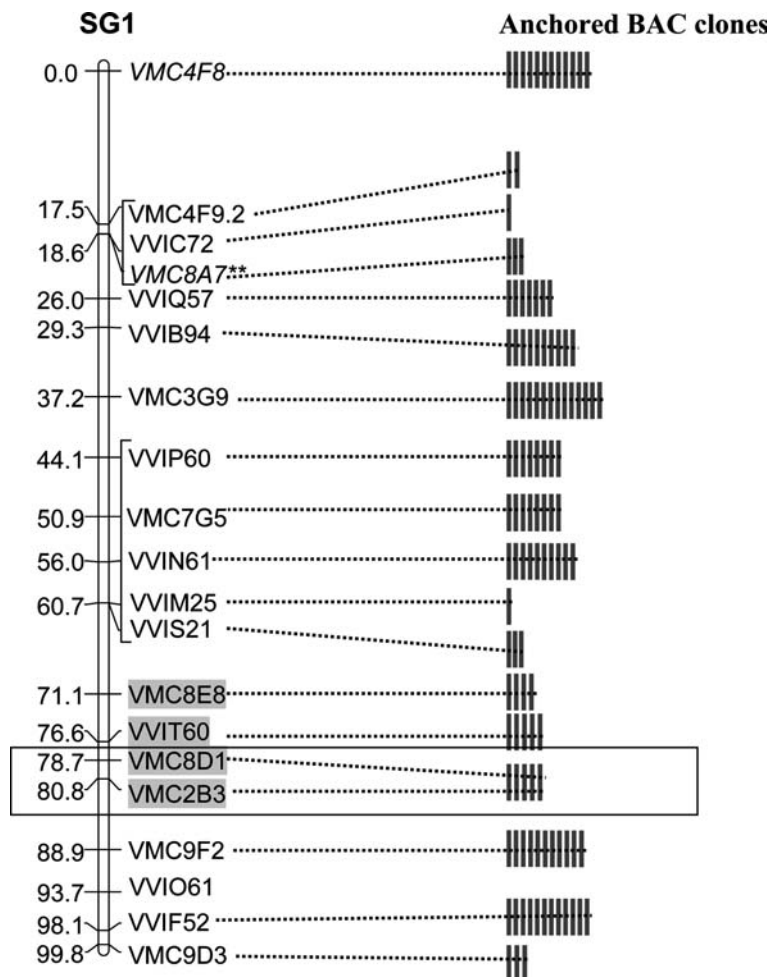
markers, 2.1 cM apart, anchored the same five BAC
clones (VMC8D1, VMC2B3), and we know from their
fingerprint patterns that these clones are different from
each other (S. Paillard, personal communication). Based
on the largest clone size in the Cabernet S library
(275 kb), we can conclude that 1 cM corresponds in this
particular region to a maximum of 130 kb. This very
favorable rate changes very quickly along the chromo-
some as the marker on the other side of VMC8D1,
which is located at the same distance of 2.1 cM, does not
anchor the same clones. It is even worse in the regions
delimited by the markers VMC4F9-2, VVIC72,
VMC8A7 and VVIM25 and VVIS21, respectively, where
the markers are very close to each other but do not
amplify any common clone (Fig. 3). However, this could
be due to a sampling effect, as all of these markers an-
chored fewer clones than expected (Fig. 3). Variations in
the correspondence between physical and genetic dis-
tances along chromosomes are well-known and have
already been very well documented in other species (see
Gorman et al. 1996, where a dramatic change in the
correspondence between physical and genetic distances
was observed in a 610-kb region in tomato).These vari-
ations represent an important parameter for map-based
cloning projects.

The Cabernet Sauvignon library, constructed
with the *Hin*dII and *Eco*RI enzymes, was chosen as
a reference library by the IGGP and will now be used
for the construction of a physical map of the grape
genome.

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

Allouis S, Moore G, Bellec A, Sharp R, Faivre-Rampant P,
Chalhoub B (2003) Construction and characterization of a
hexaploid wheat (*Triticum aestivum* L) BAC library from the
reference germplasm 'Chinese Spring'. Cereal Res Commun
31:331–338

Aradhya MK, Dangl GS, Prins BH, Boursiquot J-M, Walker MA,
Meredith CP, Simon CJ (2003) Genetic structure and differen-
tiation in cultivated grape, *Vitis vinifera* L. Genet Res 81:179–
192

Boss PK, Thomas MR (2002) Association of dwarfism and floral
induction with a grape 'green revolution' mutation. Nature
416:847–849

Boss PK, Buckeridge RJ, Poole A, Thomas MR (2003) New insights into grapevine flowering. Funct Plant Biol 30:593–606

Bowers JE, Dangl GS, Vignani R, Meredith CP (1996) DNA fingerprint characterization of some California grape cultivars. Am J Enol Vit 44:266–274

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

Budiman MA, Mao L, Wood TC, Wing RA (2000) A deep-coverage tomato BAC library and prospects toward development of an STC framework for genome sequencing. Genome Res 10:129–136

Chalhoub B, Belcram H, Caboche M (2004) Efficient cloning of plant genomes into bacterial artificial chromosome (BAC) libraries with larger and more uniform insert size. Plant Biotechnol J 2:181–188

Clark L, Carbon J (1976) A colony bank containing synthetic Col EI hybrid plasmids representative of the entire E coli genome. Cell 9:91–99

Crespan M (2004) Evidence on the evolution of polymorphism of microsatellite markers in varieties of Vitis vinifera L. Theor Appl Genet 108:231–237

Dalbó MA, Ye GN, Weeden NF, Wilcox WF, Reisch BI (2001) Marker-assisted selection for powdery mildew resistance in grapes. J Am Soc Sci 126:83–89

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

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

Franks T, Botta R, Thomas MR, Franks J (2002) Chimerism in grapevines: implications for cultivar identity, ancestry and genetic improvement. Theor Appl Genet 104:192–199

Gorman SW, Banasiak D, Fairley C, McCormick S (1996) A 610 kb YAC clone harbors 7 cM of tomato (Lycopersicon esculentum) DNA that includes the male sterile 14 gene and a hotspot for recombination. Mol Gen Genet 251:52–59

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

Hocquigny S, Pelsy F, Dumas V, Kindt S, Heloir M-C, Merdinoglu D (2004) Diversification within grapevine cultivars goes through chimeric states. Genome 47:579–589

Hong YS, Hogan JR, Wang X, Sarkar A, Sim C, Loftus BJ, Ren C, Huff ER, Carlile JL, Black K, Zhang H-B, Gardner MJ, Collins FH (2003) Construction of a BAC library and generation of BAC end sequence-tagged connectors for the genome sequencing of the African malaria mosquito Anopheles gambiae. Mol Gen Genome 268:720–728

Klein PE, Klein RR, Cartinhour SW, Ulauch PE, Dong J, Obert JA, Morishige DT, Schlueter SD, Childs KL, Ale M, Mullet JE (2000) A high throughput AFLP-based method for constructing integrated genetic and physical maps: progress toward a sorghum genome map. Genome Res 10:789–807

Levadoux L (1956) Les populations sauvages et cultivées de Vitis vinifera L. Ann Amelior Plant 6:59–118

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

Merdinoglu D, Butterlin G, Bevilaqua L, Chiquet V, Adam-Blondon A-F, Decroocq S (2005) Development and characterization of a large set of microsatellite markers in grapevine (Vitis vinifera L) suitable for multiplex PCR. Mol Breed (in press)

Noir S, Pateyron S, Combes M-C, Lashermes P, Chalhou B (2004) Construction and characterization of a BAC library for the genome analysis of the tetraploid coffee species (Coffea Arabica). Theor Appl Genet 109:225–230

Ollat N, Diakou-Verdin P, Carde J-P, Barrieu F, Gaudillère J-P, Moing A (2002) Grape berry development: a review. J Int Sci Vigne Vin 36:109–131

Peterson DG, Tomkins JP, Frisch DA, Wing RA, Paterson AH (2000) Construction of plant bacterial artificial chromosomes (BAC) libraries: an illustrated guide. J Agric Genomics 5 (http://www.ncgr.org/research/jag)

Rajesh PN, Coyne C, Meksem K, Sharma KD, Gupta V, Muehlbauer FJ (2003) Construction of a HindIII bacterial artificial chromosome library and its use in identification of clones associated with disease resistance in chickpea. Theor Appl Genet 108:663–669

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

Salmaso M, Faes G, Segala C, Stefanini M, Salakhutdinov I, Zyprian E, Toepfer R, Grando M S, 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

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

Sefc KM, Lopez MS, Lefort F, Botta R, Roubelakis-Angelakis KA, Ibanez J, Pejie I, Wagner HW, Glössl J, Steinkellner H (2000) Microsatellite variability in grapevine cultivars from different European regions and evaluation of assignment testing to assess the geographic origin of cultivars. Theor Appl Genet 100:498–505

Shahmuradov IA, Akbarova YY, Solovyev VV, Aliyev JA (2003) Abundance of plastid DNA insertions in nuclear genomes of rice and Arabidopsis. Plant Mol Biol 52:923–934

Siret R (2001) Etude du polymorphisme génétique de la vigne cultivée (Vitis vinifera L) à l'aide de marqueurs microsatellites: application à la caractérisation des cépages dans les vins. PhD thesis, University Montpellier I, Montpellier, France

Sizuya H, Birren B, Kim U-J, Mancino V, Slepak T, Tachiiri Y, Simon M (1992) Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in Escherichia coli using a F-factor-based vector. Proc Natl Acad Sci USA 89:8794–9797

Tessier C, David J, This P, Boursiquot J-M, Charrier A (1999) Optimisation of the choice of molecular markers for varietal identification in Vitis vinifera L. Theor Appl Genet 98:171–177

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

Tomkins JP, Peterson DG, Yang TJ, Ablett ER, Henry RJ, Lee LS, Holton TA, Waters D, Wing RA (2001) Grape (Vitis vinifera L) BAC library construction, preliminary STC analysis, and identification of clones associated with flavonoid and stilbene biosynthesis. Am J Enol Vitic 52:287–291

Weising K, Gradner CG (1999) A set of conserved PCR primers for the analysis of simple sequence repeat polymorphisms in chloroplast genomes of dicotyledonous angiosperms. Genome 42:9–19

Yu Y, Tomkins JP, Waugh R, Frisch DA, Kudrna D, Kleinhofs A, Brueggeman RS, Muehlbauer GJ, Wise RP, Wing RA (2000) A bacterial artificial chromosome library for barley (Hordeum vulgare L) and the identification of clones containing putative resistance genes. Theor Appl Genet 101:1093–1099