Bowers JE, Meredith CP. The parentage of a classic wine grape, Cabernet Sauvignon. Nat Genet 16: 84-87

Article *in* Nature Genetics · June 1997

DOI: 10.1038/ng0597-84 · Source: PubMed

2 authors:

John Edward Bowers
University of Georgia
173 PUBLICATIONS **15,504** CITATIONS

SEE PROFILE

Carole P Meredith
University of California, Davis
80 PUBLICATIONS **5,072** CITATIONS

SEE PROFILE

The parentage of a classic wine grape, Cabernet Sauvignon

John E. Bowers & Carole P. Meredith

The world's great wines are produced from a relatively small number of classic European cultivars of *Vitis vinifera* L. Most are thought to be centuries old and their origins have long been the subject of speculation[1,2]. Among the most prominent of these cultivars is Cabernet Sauvignon, described as "the world's most renowned grape variety for the production of fine red wine"[3]. Although now grown in many countries, Cabernet Sauvignon derives its fame from its long association with the Bordeaux region of France, where it has been grown at least since the 17th century[4]. We present microsatellite DNA evidence for the hypothesis that Cabernet Sauvignon is the progeny of two other Bordeaux cultivars, Cabernet franc and Sauvignon blanc. Likelihood ratios support this hypothesis to a very high degree of probability. A close relationship between Cabernet Sauvignon and Cabernet franc has been suspected but the genetic contribution of Sauvignon blanc, despite its similar name, is a surprise.

Because they detect genetic differences at single loci, grape microsatellite markers[5–7] provide a means to detect genetic relationships among varieties. We analysed Cabernet Sauvignon, Cabernet franc and Sauvignon blanc at 30 polymorphic microsatellite loci and at every locus the observed alleles consistent with the hypothesis that Cabernet Sauvignon is the progeny of the other two cultivars (Table 1, Fig.1). Of 51 cultivars analysed (including most major wine cultivars; Table 2), no other pair could be the parents, the nearest possibilities being excluded at least at three loci. For reasons discussed below, six loci were excluded from further analysis.

We investigated the possibility that the alleles of Cabernet Sauvignon were contributed by parents other than the proposed cultivars by means of parentage indices (adapted from paternity indices[8]) and likelihood ratios[9]. For each locus, the likelihood ratio is the ratio of the probability of the observed Cabernet Sauvignon alleles if it had the putative parents to the probability of those alleles if two random cultivars were the parents. The cumulative likelihood ratio is the product of the ratios for each locus[9]. Table 3 presents individual likelihood ratios for four loci and the cumulative ratios for 24 loci.

The cumulative likelihood ratio of the probability of the observed Cabernet Sauvignon alleles if Cabernet franc and Sauvignon blanc were the parents versus the probability of those alleles if two random cultivars were the parents is extremely high — greater than 10^{14}. If Cabernet franc is one of the parents, the likelihood of the observed Cabernet Sauvignon alleles is 3.7×10^7 times higher if Sauvignon blanc were the other parent than if the other parent were a random cultivar. Similarly, if Sauvignon blanc is a parent, the likelihood of the Cabernet Sauvignon alleles if the other parent were Cabernet franc versus a random variety is even higher, greater than 10^{11}. The much higher number is obtained for Cabernet franc because Cabernet Sauvignon and Cabernet franc share several relatively rare alleles (Table 3). The likelihood of the Cabernet Sauvignon alleles is 4×10^3 times higher if Cabernet franc were a parent versus a close relative of Cabernet franc (parent, full sibling, offspring) and more than 500 times higher for Sauvignon blanc versus a Sauvignon blanc close relative (assuming in each case that the other parent is correct).

Allele frequencies were derived from a survey of 51 major wine grape cultivars (Table 2). Because this is a relatively small number of individuals, we also calculated the likelihood ratios with 95% upper confidence limits for the allele frequencies[9]. Even with these higher allele frequencies, the Cabernet Sauvignon alleles are 5.4×10^8 times more likely if its parents were Cabernet franc and Sauvignon blanc than if they were two random cultivars (Table 3). The likelihood ratios for the single parent calculations and the close relatives are similarly still high, the lowest being 101 in the case of Sauvignon blanc versus a Sauvignon blanc relative.

Six of the 30 loci were not used in calculating likelihood ratios but the data from these loci are all consistent with the proposed parentage. Two loci (VVS3, VVS19) that repeatedly failed to amplify in some cultivars appear to have null alleles, for which frequencies could not be calculated[7]. We did not calculate allele frequencies for four other loci (VVMD20, 29, 30, 37) because three or more fragments were amplified per locus in some cultivars.

Additional evidence is found in amplified restriction fragment polymorphism (AFLP), restriction fragment length polymorphism (RFLP) and isozyme data. Of 99 polymorphic AFLP bands observed in a limited study in our laboratory, all 38 bands present in Cabernet Sauvignon were also present in one or both of the putative parents (J.E.B. and C.P.M., unpublished results). Of 44 polymorphic RFLP bands that we analysed, the 23 bands present in Cabernet Sauvignon were all present in one or both of the putative parents[10,11]. In another RFLP study, the 56 polymorphic bands present in Cabernet Sauvignon were also present in Cabernet franc and/or Sauvignon blanc[12]. Data from two isozyme studies, involving seven enzyme systems, are also consistent with the proposed parentage[13,14].

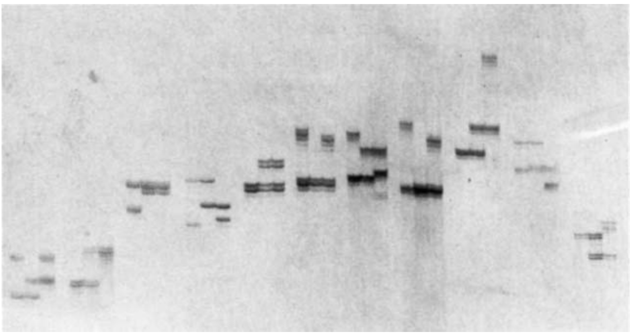

Fig. 1 Microsatellite alleles of Cabernet Sauvignon, Cabernet franc and Sauvignon blanc at 11 loci. Each group of 3 lanes illustrates the alleles at a single locus for all three cultivars. In each case the lanes are Cabernet franc (left), Cabernet Sauvignon (middle) and Sauvignon blanc (right). Loci are, from left to right, *VVMD31, VVMD24, VVMD28, VVMD5, VVMD34, VVMD32, VVMD25, VVMD7, VVMD36, VVMD21, VVMD14*. For each locus, the inheritance in Cabernet Sauvignon of one allele from each parent is evident.

Department of Viticulture and Enology, University of California, Davis, California 95616, USA. Correspondence should be addressed to C.P.M. e-mail: cpmeredith@ucdavis.edu.

Table 1 • Microsatellite alleles of Cabernet Sauvignon and its presumptive parents Cabernet franc and Sauvignon blanc

Locus	Cabernet franc	Cabernet Sauvignon	Sauvignon blanc	Locus	Cabernet franc	Cabernet Sauvignon	Sauvignon blanc
VVMD5	226, 240	232, 240	228, 232	*VVMD27*	181, 189	175, 189	175, 189
VVMD6	205, 211	211, 212	205, 212	*VVMD28*	231, 239	237, 239	237, 239
VVMD7	239, 263	239, 239	239, 257	*VVMD31*	206, 216	206, 210	210, 216
VVMD8	147, 157	143, 157	143, 143	*VVMD32*	241, 259	241, 241	241, 257
VVMD14	228, 235	222, 228	222, 232	*VVMD34*	240, 240	240, 248	240, 248
VVMD16	168,168	168, 168	168, 168	*VVMD36*	254, 254	254, 264	264, 295
VVMD17	212, 221	221, 222	220, 222	*VVS1*	181, 181	181, 181	181, 190
VVMD21	249, 258	249, 258	243, 249	*VVS2*	139, 147	139, 151	133, 151
VVMD23	178, 178	178, 178	178, 178	*VVS4*	167, 175	168, 175	168, 169
VVMD24	210, 210	210, 219	219, 219	*VVS16*	284, 284	284, 284	284, 284
VVMD25	243, 259	243, 253	245, 253	*VVS29*	175, 181	179, 181	171, 179
VVMD26	249, 249	249, 251	251, 251	*VHS43*	210, 213	210, 213	210, 213

Likelihood analyses assume independent loci and Hardy-Weinberg equilibrium. Linkage has not yet been determined for most grape microsatellites, but preliminary results with 18 loci suggest that most may be unlinked (B. Reisch, personal communication). Microsatellite loci mapped in other species are dispersed throughout the genome[15]. Of the 24 loci on which the likelihood ratios are based, 19 are in Hardy-Weinberg equilibrium (5% level, data not shown). The extent to which grape cultivars are derived from the domestication (and subsequent maintenance by vegetative propagation) of individuals resulting from completely random mating in natural (wild) populations versus non-random mating between cultivated vines is not known. Thus random mating cannot be assumed. The grape cultivars we have studied are heterozygous at an average of 67% of the microsatellite loci analysed, indicating that they are likely the result of outcrossing, if not completely random mating.

Like Cabernet Sauvignon, Cabernet franc is also a distinguished red wine cultivar, although it does not enjoy the reputation of Cabernet Sauvignon and is not as widely grown. While both cultivars have been cultivated in Bordeaux since at least the 17th century, there is evidence that Cabernet franc was already being grown there several centuries earlier[2]. The two cultivars are morphologically similar, so much so that they are generally thought to be related[4] and at least one author has wondered if one has arisen from the other by mutation[3].

Table 2 • Cultivars on which allele frequencies are based

Alicante Bouschet
Barbera
Black Malvoisie
Burger
Cabernet franc
Cabernet Sauvignon
Carignane
Chardonnay
Chenin blanc
Dolcetto
Durif
Folle blanche
French Colombard
Gamay noir
Gewürztraminer
Grand Noir
Grenache
Malbec
Malvasia Bianca
Mataro
Melon
Merlot
Mission
Müller Thurgau
Muscat of Alexandria
Muscat Blanc
Napa Gamay
Nebbiolo Lampia
Palomino
Petite Bouschet
Petite Sirah
Petit Verdot
Pinot noir
Pinot St. George
Plavac Mali
Riesling
Ruby Cabernet
Rulˇnder
Sangiovese
Sauvignon blanc
Semillon
St. Emilion
Syrah
Thompson Seedless
Tinta Ameria
Tinta C‹o
Trousseau
Unknown White
Viognier
Xinisteri
Zinfandel

The white wine variety Sauvignon blanc is a major component of the Sauternes wines of Bordeaux and also produces the wines of Sancerre and Pouilly-Fumé in the Loire Valley. This cultivar has existed in diverse regions of both western and eastern France for several centuries and has been known in Bordeaux at least since the early 18th century[4].

Although it is possible that a pigmented form of Sauvignon is the parent, it is unlikely as Sauvignon blanc has long been widely planted in Bordeaux and coloured forms of Sauvignon are either rare or no longer exist[4,16]. Furthermore, some selfed progeny of Cabernet Sauvignon have white berries (M.A. Walker and C.P.M., unpublished results), as would be expected if Cabernet Sauvignon were heterozygous for fruit color by virtue of having a white-berried parent.

It is not likely that the cross between Cabernet franc and Sauvignon blanc was made deliberately. It could have occurred no later than the 17th century, an era that precedes the earliest reports of deliberate plant hybridization[17]. Furthermore, there would have been no rationale for making such a cross. It is more likely that the cross occurred spontaneously between vines in adjacent vineyards or perhaps the same vineyard, as vineyards containing mixtures of red and white cultivars were once common. Grape cultivars are largely self-pollinating, but occasional outcrossing occurs (C.P.M., J.E.B. and S. Yaish, unpublished data). Comparative analysis of chloroplast genomes would establish the direction of the cross. Although some wine cultivars are suggested to have a polyclonal origin (descended from more than one original seedling)[18], we conclude that Cabernet Sauvignon has developed from a single seedling because 11 numbered clones all exhibit identical microsatellite genotypes (J.E.B. and C.P.M., unpublished results).

Does the name Cabernet Sauvignon, in combining elements of the names of each of the parents, reflect prior knowledge of the parentage? While some literature suggests a relationship between Cabernet franc and Cabernet Sauvignon, there is no indication that Sauvignon blanc might be related. Some morphological resemblance between Sauvignon blanc and Cabernet Sauvignon has been suggested as a rationale for the name of Cabernet Sauvignon but the resemblance is considered superficial and not to

Table 3 • Parentage analysis of Cabernet Sauvignon (CS) and its presumptive parents Cabernet franc (CF) and Sauvignon blanc (SB)

		Genotype			Allele frequencies		Likelihood ratios of CF × SB[a] versus alternative parents (observed allele frequencies)					Likelihood ratios of CF × SB versus alternative parents (upper confidence limits of allele frequencies)				
Locus	Allele size (bp)	CF	CS	SB	Observed[b]	Upper 95% confidence limit	X × Y[c,d]	SB × X[e]	CF × X[e]	SB × CF relative[f]	CF × SB relative[f]	X × Y[c,d]	SB × X[e]	CF × X[e]	SB × CF relative[f]	CF × SB relative[f]
VVMD14[g]	222		—	—	0.31	0.39	10.2	12.8	1.6	1.9	1.2	4.5	7.1	1.3	1.8	1.1
	228	—	—		0.04	0.07										
	232			—	0.14	0.19										
	235	—			0.10	0.15										
VVMD24[g]	210	—	—		0.48	0.56	7.6	2.1	7.3	1.4	1.8	4.6	1.8	5.2	1.3	1.7
	219		—	—	0.14	0.19										
VVMD31[g]	206	—	—		0.05	0.08	13.7	10.2	2.7	1.8	1.5	5.9	5.9	2.0	1.7	1.3
	210		—	—	0.19	0.25										
	216	—		—	0.25	0.32										
VVS2[g]	133			—	0.37	0.45	9.0	8.5	2.1	1.8	1.4	4.2	5.1	1.6	1.7	1.2
	139	—	—		0.06	0.10										
	147	—			0.01	0.03										
	151		—	—	0.24	0.30										
Cumulative product for these 4 loci							9525	2301	66.2	8.2	4.3	525	384	22	6.4	3.1
Cumulative product for all 24 loci							1.5×10^{14}	1.6×10^{11}	3.7×10^{7}	3989	576	5.4×10^{8}	1.5×10^{8}	1.6×10^{5}	769	101

[a] No direction is implied by the order of the parents in each cross. [b] Based on 51cultivars (individuals) (unpublished data). [c] Where X and Y are random unrelated cultivars. [d] The identity of both parents is unknown. [e] The identity of one parent is assumed but that of the other is unknown. [f] The identity of one parent is assumed but the possibility is considered that the other parent is actually a close relative of the cultivar proposed as the second parent. [g] Total number of alleles observed in 51 cultivars is 11, 6, 7 and 10 for *VVMD14*, *VVMD24*, *VVMD31* and *VVS2*, respectively.

reflect any genetic relationship[4]. The more commonly accepted explanation for the similarity in names is that they share the word 'sauvignon' (derived from the French word 'sauvage', meaning 'wild') because of the resemblance of both cultivars to wild vines[1].

We present strong genetic evidence for the origin of one of the most celebrated wine cultivars, long the subject of speculation. Continued comparative genetic analysis of wine grapes is likely to shed further light on the history of this culturally and economically significant crop.

Methods

Microsatellite analysis. DNA was extracted from young grapevine leaves as reported[10] from vines in the Foundation Plant Material Services collection at the University of California at Davis. Microsatellite loci were PCR amplified as described[6], separated on denaturing 6% polyacrylamide sequencing gels and visualized by silver staining with a commercial kit (Promega). Thirty polymorphic microsatellite loci were tested, of which eight have been described[5,6], four were provided by M. R. Thomas (personal communication) and eighteen were recently characterized in our laboratory (J.E.B. and C.P.M., unpublished results).

Statistical analysis. Likelihood ratios were calculated from the microsatellite genotypes of the putative parental cultivars and from allele frequencies determined from a sample of 51 wine cultivars (unpublished data). Although *Vitis vinifera* is an outcrossing species in nature, in agriculture each grape cultivar is vegetatively propagated and individual vines within a cultivar are identical at microsatellite loci; thus a cultivar is considered an individual for the purpose of this analysis.

Likelihood ratios. The likelihood of the Cabernet Sauvignon alleles if it had the putative parents (X × Y in Table 3) was estimated largely as done for the identification of the parents of an unknown human murder victim[9]. For each locus, the likelihood ratio was calculated as the quotient of the probability of the 'Cabernet Sauvignon' genotype if it had the presumptive parents and the probability of that genotype if two random cultivars were the parents. If the progeny individual is heterozygous and both putative parents are heterozygous at a locus, the ratio is $1/8pq$, where p and q are the population allelic frequencies, except when all three have identical alleles at the locus, in which case the ratio is $1/4pq$ (this situation was not included in ref. 9). If the progeny individual is heterozygous and one or both putative parents are homozygous, then the ratio is $1/4pq$ or $1/2pq$, respectively. If the progeny individual is homozygous, the ratios are $1/4p^2$, $1/2p^2$ or $1/p^2$, if the putative parents are both heterozygous, one heterozygous or both homozygous, respectively.

If one of the putative parental cultivars is assumed to be correct, the likelihood of the observed Cabernet Sauvignon alleles if the second cultivar were also a parent (SB × X and CF × X in Table 3) was calculated as in human paternity testing[8]. The likelihood ratio was calculated as the quotient of the probability of the Cabernet Sauvignon genotype if it had the presumptive parents and the probability of that genotype if one of the parents were one of the presumptive parents and the other were a random cultivar. The ratios are $1/p$ when the parental cultivar in question is homozygous at a locus, and $1/2p$ when it is heterozygous. However, if the progeny individual is heterozygous and identical to the presumptive parent at that locus, then the ratios are $1/(p+q)$ if the cultivar in question is either homozygous or an identical heterozygote and $1/(2p+2q)$ if the cultivar in question is a different heterozygote.

We also considered the possibility that one of the putative parents is correct and the other is a close relative of the second putative parent (SB × CF relative and CF × SB relative in Table 3). Again, the methodology was that used in human paternity testing[8]. The likelihood ratios were calculated as the quotient of the probability of the Cabernet Sauvignon genotype if it had the putative parents and the probability of that genotype if the parents were one of the presumptive parents and a close relative of the second presumptive parent. A close relative is defined as a parent, progeny or full sibling, sharing half of its alleles with the second cultivar, the other half being determined by the population allelic frequencies. If the cultivar in question is homozygous at a locus, the ratio is $2/(1+p)$. The ratios are $2/(1+p)$ when the parental cultivar in question is homozygous at a locus, and $2/(1+2p)$ when it is heterozygous. However, if the progeny individual is heterozygous and identical to the presumptive parent at that locus, then the ratios are $2/(1+p+q)$ if the cultivar in question is either homozygous or an identical heterozygote and $2/(1+2p+2q)$ if the cultivar in questions is a different heterozygote.

Acknowledgements

We thank A. T. Bowling for valuable guidance regarding parentage analysis and M.R. Thomas and B.I. Reisch for unpublished data. This work was supported by the Fruit Tree, Nut Tree And Grapevine Improvement Advisory Board of the California Department of Food and Agriculture and the American Vineyard Foundation.

Received 5 February; accepted 5 March 1997.

1. Levadoux, L. Les populations sauvages et cultivées de *Vitis vinifera* L. *Annales de l'Amélioration des Plantes* **6**, 59–118 (1956).
2. Bouquet, A. Origine et évolution de l'encépagement français à travers les siècles. *Le progrès agricole et viticole* **99**, 110–121 (1982).
3. Robinson, J. *The Oxford Companion to Wine*. (Oxford University Press, Oxford, 1994).
4. Viala, P. & Vermorel, V. *Ampélographie, Tome II*. (Masson, Paris, 1901).
5. Thomas, M.R. & Scott, N.S. Microsatellite repeats in grapevine reveal DNA polymorphisms when analysed as sequence-tagged sites (STSs). *Theor. Appl.Genet.* **86**, 985–990 (1993).
6. Bowers, J.E., Dangl, G.S., Vignani,R. & Meredith, C.P. Isolation and characterization of new polymorphic simple sequence repeat loci in grape (*Vitis vinifera* L.) *Genome* **39**, 628–633 (1996).
7. Thomas, M.R., Cain, P. & Scott, N.S. DNA typing of grapevines: A universal methodology and database for describing cultivars and evaluating genetic relatedness. *Plant Mol. Biol.* **25**,939–949 (1994).
8. Weir, B.S. *Genetic Data Analysis II*. (Sinauer Associates, Inc., Sunderland, Massachussetts, 1996).
9. Hagelberg, E., Gray, I.C. & Jeffreys, A.J. Identification of the skeletal remains of a murder victim by DNA analysis. *Nature* **352**, 427–429 (1991).
10. Bowers, J.E., Bandman, E.B. & Meredith, C.P. DNA fingerprint characterization of some California winegrape cultivars. *Amer. J. Enol. Vitic.* **44**, 266–274 (1993).
11. Bowers, J.E. *DNA fingerprinting analysis of some winegrape cultivars*. M.S. thesis. (University of California, Davis, 1993).
12. Bourquin, J.-C., Sonko, A., Otten, L. & Walter, B. Restriction fragment length polymorphism and molecular taxonomy in *Vitis vinifera* L. *Theoret. Appl. Genet.* **87**, 431–438 (1993).
13. Wolfe, W.H. Identification of grape varieties by isozyme banding patterns. *Amer. J. Enol. Vitic.* **27**, 68–73 (1976).
14. Parfitt, D.E. & Arulsekar, S. Inheritance and isozyme diversity for GPI and PGM among grape cultivars. *J. Amer. Soc. Hort. Sci.* **114**, 486–491 (1989).
15. Hudson, T.J. *et al.* Isolation and chromosomal assignment of 100 highly informative human simple sequence repeat polymorphisms. *Genomics* **13**, 622–629 (1992).
16. Galet, P. *Cépages et Vignobles de France. Tome II. L'ampélographie Française*. (Imprimerie Charles Dehan, Montpellier, France, 1990).
17. Sturtevant, A.H. *A History of Genetics*. (Harper and Row, New York, 1965.)
18. Rives, M. Bases gènètiques de la sélection clonale chez la vigne. Ann. Amélior. Plantes 11, 337–348 (1961).