Journal of Applied Microbiology

Journal of Applied Microbiology ISSN 1364-5072

ORIGINAL ARTICLE

Genetic relatedness and recombination analysis of *Allorhizobium vitis* strains associated with grapevine crown gall outbreaks in Europe

N. Kuzmanović[1], E. Biondi[2], A. Bertaccini[2] and A. Obradović[1]

1 University of Belgrade-Faculty of Agriculture, Belgrade, Serbia
2 Department of Agricultural Sciences, University of Bologna, Bologna, Italy

Keywords
Agrobacterium, *Allorhizobium*, genetic diversity, grapevine, housekeeping genes, random amplified polymorphic DNA, recombination.

Correspondence
Nemanja Kuzmanović, University of Belgrade-Faculty of Agriculture, Nemanjina 6, 11080 Belgrade, Serbia.
E-mail: kuzmanovic1306@gmail.com

2015/0478: received 8 March 2015, revised 21 May 2015 and accepted 21 May 2015

doi:10.1111/jam.12858

Abstract

Aims: To analyse genetic diversity and epidemiological relationships among 54 strains of *Allorhizobium vitis* isolated in Europe during an 8-year period and to assess the relative contribution of mutation and recombination in shaping their diversity.

Methods and Results: By using random amplified polymorphic DNA (RAPD) PCR, strains studied were distributed into 12 genetic groups. Sequence analysis of *dnaK*, *gyrB* and *recA* housekeeping genes was employed to characterize a representative subcollection of 28 strains. A total of 15 different haplotypes were found. Nucleotide sequence analysis suggested the presence of recombination events in *A. vitis*, particularly affecting *dnaK* locus. Although prevalence of mutation over recombination was found, impact of recombination was about two times greater than mutation in the evolution of the housekeeping genes analysed.

Conclusions: The RAPD analysis indicated high degree of genetic diversity among the strains. However, the most abundant RAPD group was composed of 35 strains, which could lead to the conclusion that they share a common origin and were distributed by the movement of infected grapevine planting material as a most common way of crossing long distances. Furthermore, it seems that recombination is acting as an important driving force in the evolution of *A. vitis*. As no substantial evidence of recombination was detected within *recA* gene fragment, this phylogenetic marker could be reliable to characterize phylogenetic relationships among *A. vitis* strains.

Significance and Impact of the Study: We demonstrated clear epidemiological relationship between majority of strains studied, suggesting a need for more stringent phytosanitary measures in international trade. Moreover, this is the first study to report recombination in *A. vitis*.

Introduction

Allorhizobium vitis is the major causal agent of crown gall of grapevine (*Vitis vinifera*) (Burr *et al.* 1998; Burr and Otten 1999). It is one of the most important bacterial pathogens in grapevine-growing areas worldwide. The systemic movement through the xylem vessels and ability to be latently present within the grapevine facilitates pathogen dissemination by asymptomatic propagation material (Burr *et al.* 1998; Burr and Otten 1999).

Until recently, *A. vitis* was regarded a member of the genus *Agrobacterium*. However, Mousavi *et al.* (2014, 2015) showed that this species is distinct from *Agrobacterium* by using multilocus sequence analysis (MLSA) and proposed its transfer to revived genus *Allorhizobium*. The distinctness of *A. vitis* from the genus *Agrobacterium* was also pointed out by Ramírez-Bahena *et al.* (2014). These authors suggested that members of the genus *Agrobacterium* are characterized by the presence of linear chromosome (chromid) beside circular chromosome. Unlike *Agrobacterium* spp.,

the core genome of *A. vitis* consists of two circular chromosomes (Jumas-Bilak *et al.* 1998; Slater *et al.* 2009). On the other hand, pathogenicity of *A. vitis* strains is mainly determined by the presence of tumour-inducing (Ti) plasmid in their genome.

The genetic diversity of this pathogen has been extensively studied in terms of both Ti plasmid and chromosomal DNA (Burr *et al.* 1998; Burr and Otten 1999). Methods for examining genetic variations and the relationships of phytopathogenic bacteria are particularly important for tracking strains and detecting disease outbreaks in epidemiological studies. In addition, knowledge of pathogen population structure and dynamics is essential for understanding bacterial ecology and evolution. Although evolution of *A. vitis* Ti plasmids was studied extensively (Burr *et al.* 1998; Burr and Otten 1999), little is known about the core genome evolution of this species. Analysis of housekeeping genes, which are responsible for basic cellular functions and relatively conserved, provide data that can be used for these purposes.

Grapevine grown gall has been recorded in many European countries in the last 15 years and caused important economic losses in some vineyards (Genov *et al.* 2006; Bini *et al.* 2008b; Kuzmanović *et al.* 2014). Therefore, the objective of this study was to analyse genetic diversity among 54 strains of *A. vitis* associated with these outbreaks and to evaluate their epidemiological relationships. Initial assessment of genetic diversity among strains was assessed by random amplified polymorphic DNA (RAPD) method. A total of 28 representative strains, at least one per each RAPD group, were characterized by sequence analysis of *dnaK*, *gyrB* and *recA* housekeeping genes. The molecular evolution of the strains was investigated by assessing the relative contribution of mutation and recombination in shaping their diversity.

Materials and methods

Bacterial strains and growth conditions

A total of 63 strains of *A. vitis* originating from eight European countries (58 strains), Africa (two strains), North America (two strains) and Australia (one strain) were included in this study (Table 1). Of 58 European strains, 54 were isolated from diseased grapevine during a period of 8 years, from 2003 until 2011. All the strains harboured Ti plasmid, except the nonpathogenic strain F2/5. Species identities of all strains and presence (or absence) of Ti plasmid in their genome were confirmed by PCR analysis as described by Bini *et al.* (2008b) and Puławska *et al.* (2006).

Strains of *A. vitis* were routinely cultured on yeast mannitol agar (YMA; 10 g l^{-1} mannitol, 1 g l^{-1} yeast extract, 1 g l^{-1} $CaCO_3$, 0·1 g l^{-1} NaCl, 0·5 g l^{-1} K_2HPO_4, 0·2 g l^{-1} $MgSO_4{\cdot}7H_2O$, 18 g l^{-1} agar; pH 7·2) medium at 27°C. For DNA preparation, bacteria were grown on King's medium B (King *et al.* 1954) at 27°C for 24–48 h, to reduce production of polysaccharides. Cultures were stored at 4°C on YMA medium for short-term maintenance and frozen in a −80°C freezer in nutrient broth with 30% glycerol for long-term preservation.

DNA preparation

Total genomic DNA was isolated from bacterial suspensions (approx. 10^8 CFU ml^{-1}) using QIAamp DNA Mini Kit (Qiagen, Hilden, Germany) or Wizard Genomic DNA Purification Kit (Promega Corp., Madison, WI) according to the manufacturer's protocol. The DNA samples were diluted to 10–20 ng μl^{-1} and stored at −20°C for further analysis.

RAPD fingerprinting

RAPD analysis was performed with primers A9, A10 and R13, which were used for typing *A. vitis* strains by other authors (Irelan and Meredith 1996; Momol *et al.* 1998). PCR reactions (25 μl) were performed in 1 × colourless GoTaq Flexi buffer (Promega Corp.), 2 mmol l^{-1} $MgCl_2$, 0·2 mmol l^{-1} of each dNTP, 0·8 μmol l^{-1} of primer, 1·25 U of GoTaq Flexi DNA polymerase (Promega Corp.) and 2 μl of DNA. The thermal profile was as follows: initial denaturation at 95°C for 5 min; 35 cycles of denaturation at 94°C for 30 s, annealing at 36°C for 30 s and extension at 72°C for 1 min. A final extension was conducted at 72°C for 7 min. Electrophoresis was performed in 1 × TAE buffer in a 1·5% agarose gel (w/v) under standardized conditions (80 V, 60 mA, 240 min, room temperature).

Genetic profiles generated by RAPD were transformed into a binary matrix by scoring bands as 1 (present) or 0 (absent). Clustering analysis of the combined binary matrices for all three primers was performed with FREETREE program (Hampl *et al.* 2001), by using the unweighted pair group method with arithmetic mean (UPGMA) and the coefficient of Nei and Li (Nei and Li 1979). Bootstrap analysis was based on 1000 resamplings. The TREEVIEW program (Page 1996) was used for visualization of the dendrogram.

PCR amplification and housekeeping gene sequencing

Gene selection and primer design were conducted using corresponding sequences derived from the complete genome sequences of *A. vitis* strain S4 (CP000633),

Table 1 *A. vitis* strains used in this study

Strain code	Geograpahic origin*	Year of isolation	Source†	Literature	Random amplified polymorphic DNA group
KFB 257‡	Šabac, SRB	2011	KFB	Kuzmanović *et al.* (2014)	1
IPV-BO 6186‡	Tuscany, Italy	2006	IPV-BO	Bini *et al.* (2008a)	2
AB3‡	Balatonboglár, HUN	1982	S. Süle	Szegedi *et al.* (1988)	3
IPV-BO 6207‡	Vršac, SRB	2006	IPV-BO	Bini *et al.* (2008a)	4
IPV-BO 5162‡	FVG, Italy	2003	IPV-BO	Bini *et al.* (2008b)	5
IPV-BO 5761	Moldova	2004	IPV-BO	Bini *et al.* (2008a)	5
IPV-BO 5881	Apulia, Italy	2005	IPV-BO	Bini *et al.* (2008a)	5
IPV-BO 6048A1	Montenegro	2005	IPV-BO	Bini *et al.* (2008a)	5
IPV-BO 6209	Vršac, SRB	2006	IPV-BO	Bini *et al.* (2008a)	5
IPV-BO 6570	Bulgaria§	2006	IPV-BO	Bini *et al.* (2008b)	5
IPV-BO 7104	FVG, Italy	2007	IPV-BO	This study	5
IPV-BO 8463	Morocco	2011	IPV-BO	This study	5
KFB 239‡, KFB 240, KFB 241	Vršac, SRB	2010	KFB	Kuzmanović *et al.* (2014)	5
KFB 242‡, KFB 244	Irig, SRB	2011	KFB	Kuzmanović *et al.* (2014)	5
KFB 245, KFB 246	Negotin, SRB	2011	KFB	Kuzmanović *et al.* (2014)	5
KFB 247, KFB 248	Vršac, SRB	2011	KFB	Kuzmanović *et al.* (2014)	5
KFB 249	Smederevo, SRB	2011	KFB	Kuzmanović *et al.* (2014)	5
KFB 255, KFB 256, KFB 258, KFB 259, KFB 260, KFB 261	Šabac, SRB	2011	KFB	Kuzmanović *et al.* (2014)	5
KFB 262‡, KFB 263	Vladimirci, SRB	2011	KFB	Kuzmanović *et al.* (2014)	5
KFB 267‡, KFB 268	Bujanovac, SRB	2011	KFB	Kuzmanović *et al.* (2014)	5
KFB 269, KFB 270, KFB 271	Vranje, SRB	2011	KFB	Kuzmanović *et al.* (2014)	5
KFB 272, KFB 273, KFB 274	Aleksandrovac, SRB	2011	KFB	Kuzmanović *et al.* (2014)	5
WIN 4.2.3‡, WIN 4.2.4	Poland	2010	J. Puławska	This study	5
IPV-BO 7105‡	FVG, Italy	2007	IPV-BO	This study	6
IPV-BO 8812	ER, Italy	2011	IPV-BO	This study	6
CG102‡	Virginia, USA	1984	IPV-BO	Bazzi *et al.* (1988)	7
K309^T‡	Barmera, AUS	1977	IPV-BO	Ophel and Kerr (1990)	7
KFB 253‡	Smederevo, SRB	2011	KFB	Kuzmanović *et al.* (2014)	7
IPV-BO 5372‡	Italy	2003	IPV-BO	Bini *et al.* (2008b)	8
IPV-BO 6571‡	Bulgaria§	2006	IPV-BO	Bini *et al.* (2008b)	8
KFB 243‡	Irig, SRB	2011	KFB	Kuzmanović *et al.* (2014)	8
S4‡	Orgovány, HUN	1981	S. Süle	Szegedi *et al.* (1988)	9
IPV-BO 5159‡	FVG, Italy	2003	IPV-BO	Bini *et al.* (2008b)	10
IPV-BO 8816‡	ER, Italy	2011	IPV-BO	This study	10
Av2‡	Croatia	2006	IPV-BO	This study	11
IPV-BO 1861-5‡	ER, Italy	1984	IPV-BO	Bini *et al.* (2008b)	12
KFB 254‡	Smederevo, SRB	2011	KFB	Kuzmanović *et al.* (2014)	13
KFB 264‡, KFB 265, KFB 266	Vranje, SRB	2011	KFB	Kuzmanović *et al.* (2014)	14
F2/5‡	Pretoria, SA	UD¶	IPV-BO	Burr *et al.* (1997)	15
KFB 250‡, KFB 251, KFB 252	Smederevo, SRB	2011	KFB	Kuzmanović *et al.* (2014)	16
AB4‡	Balatonboglár, HUN	1982	IPV-BO	Szegedi *et al.* (1988)	17
CG49‡	New York, USA	1979	IPV-BO	Burr *et al.* (1987)	17

*SRB, Serbia; HUN, Hungary; FVG, Friuli Venezia Giulia; ER, Emilia-Romagna; USA, United States of America; AUS, Australia; SA, South Africa.
†KFB, Collection of Phytopathogenic Bacteria, University of Belgrade-Faculty of Agriculture, Belgrade, Serbia; IPV-BO, Plant Pathology Department, University of Bologna, Bologna, Italy; S. Süle, Plant Protection Institute, Hungarian Academy of Sciences, Budapest, Hungary; J. Puławska, Research Institute of Horticulture, Skierniewice, Poland.
‡Strains used in phylogenetic analysis based on housekeeping gene sequences.
§In the paper of Bini *et al.* (Bini *et al.* 2008b) Italy is erroneously listed as a geographic origin of this strain.
¶UD, undocumented.

Agrobacterium tumefaciens genomic species G8 strain C58 (AE007869) and *Rhizobium rhizogenes* strain K84 (CP000628). The following three housekeeping genes located on chromosome 1 of *A. vitis* were included in the study: *dnaK* (encoding 70-kDa heat shock protein), *gyrB* (encoding DNA gyrase subunit B) and *recA* (encoding

Table 2 Primers used for housekeeping gene amplification and sequencing

Target	Primer code	Sequence (5′→3′)	Primer positions*	Ta†	Fragment length (bp)
dnaK	dnaKF1	TGGCAAAAGTAATCGGTATCG	2-22	60	576
	dnaKR1	CGTAAACGGCAATGGTCTT	559-577		
gyrB	gyrBF2	AAGATGTTGTCCAGCCAGGA	1414-1433	60	626
	gyrBR2	TCCTTGACACCACGCACC	2022-2039		
recA	recAF1	AGAGGACAAAACGGTGGATAAAAG	27-50	57	615
	recAR1	AACATCACGCCAATCTTCATAC	620-641		

*The position within the corresponding gene sequences of *A. vitis* S4.
†Annealing temperature (°C).

Table 3 Sequence diversity of the selected loci among the 28 *A. vitis* strains analysed

Locus	Sequence length (bp)	GC%	Number of haplotypes	No. of polymorphic sites (%)	π*	dN/dS†	Tajima's D‡
dnaK	537	58·1	12	52 (9·7)	0·032	0·012	0·676
gyrB	537	59·2	12	55 (10·2)	0·029	0·003	−0·274
recA	564	56·3	12	70 (12·4)	0·041	0¶	0·357
Concat§	1638	57·8	15	177 (10·8)	0·034	NA	NA

NA, not applicable.
*Nucleotide diversity calculated with the Jukes-Cantor correction.
†Ratio between the numbers of nonsynonymous (dN) and synonymous substitutions (dS).
‡No significant ($P > 0{\cdot}10$) deviation from zero was observed for any of the studied data sets.
§Concat, data set of concatenated sequences.
¶Nonsynonymous substitutions were not found.

recombinase A) (Table 2). Primers used for amplification and sequencing are listed in Table 2.

PCR amplifications were performed in a 50 μl reaction mixture containing 1× colourless GoTaq Flexi buffer (Promega Corp.), 1·5 mmol l^{-1} $MgCl_2$, 0·2 mmol l^{-1} of each dNTP, 0·2 μmol l^{-1} (*dnaK* and *gyrB*) or 0·16 μmol l^{-1} (*recA*) of each primer, 1 U (*dnaK* and *recA*) or 0·75 U (*gyrB*) of GoTaq Flexi DNA polymerase (Promega Corp.) and 4 μl of extracted DNA. The thermal profile was as follows: initial denaturation at 95°C for 5 min; 35 cycles of denaturation at 94°C for 45 s, annealing at a gene-specific temperature (Table 2) for 45 s and extension at 72°C for 45 s. A final extension was conducted at 72°C for 10 min.

Purity and yield of PCR products were checked by electrophoresis in 1·7% agarose gel using a phiX174 DNA/*Bsu*RI (*Hae*III) molecular weight marker (Fermentas, Vilnius, Lithuania). The amplification products were commercially sequenced (Macrogen Europe, Amsterdam, the Netherlands) in both directions using forward and reverse primers (Table 2).

Nucleotide sequence analysis

Forward and reverse nucleotide sequences were visually inspected, assembled and edited using FINCHTV 1.4.0 (Geospiza, Inc., Seattle, WA) and MEGA 5.2.2 (Tamura *et al.* 2011). Consensus sequences were translated with the Translate (http://www.expasy.org), back-translated with the Reverse Translate (http://www.bioinformatics.org/sms2/rev_trans.html) and then codon-aligned with the ClustalW algorithm (Higgins *et al.* 1996) implemented in MEGA 5.2.2 (Tamura *et al.* 2011). The size of the codon-aligned sequences used for further analysis is indicated in Table 3. Sequences were manually concatenated following the alphabetic order of the gene names. Analyses were performed on individual gene sequences as well as on their concatenation.

The GC content and nucleotide diversity indexes of the sequences analysed were determined using the DNASP 5.10.2 program (Librado and Rozas 2009). The same software was also used to perform Tajima's D neutrality test (Tajima 1989). The ratio of nonsynonymous to synonymous substitutions (dN/dS) was calculated using the START2 package program (Jolley *et al.* 2001).

The presence of linkage disequilibrium between alleles at all loci was assessed by measuring the standardized index of association (I^S_A) (Haubold *et al.* 1998) using START2 software (Jolley *et al.* 2001). Phylogenetic network analysis was performed with SPLITSTREE 4.13.1 using the Neighbor-Net algorithm (Huson and Bryant 2006). To test the presence of recombination, the pairwise

homoplasy index (Phi) was calculated by the same software (Bruen *et al.* 2006; Huson and Bryant 2006). Furthermore, the detection of potential recombination events was conducted using the RDP 4.27 software package (Martin *et al.* 2010), comprising the following programs: RDP, GENECONV, BOOTSCAN, MAXCHI, CHIMAERA, SISCAN and 3SEQ. The analysis was performed with default settings, including options 'disentangle overlapping signals' and 'auto mask for optimal recombination detection'. Recombination events detected with at least three of seven methods were kept for further analyses to firmly examine their accuracy. The Web-based program GARD (genetic algorithm for recombination detection) (Kosakovsky Pond *et al.* 2006) was used to detect recombination and estimate breakpoint locations.

Two measures of the recombination rate (ρ/θ and r/m) were calculated on the concatenated data (*dnaK*, *gyrB* and *recA*) using CLONALFRAME 1.2 software (Didelot and Falush 2007) as described by Vos and Didelot (2009).

The phylogenetic analysis was performed by maximum likelihood (ML) method using RAXMLGUI 1.3 (Silvestro and Michalak 2012) with 100 rapid bootstraps. The most suitable substitution models were determined by the JMODELTEST 2.1.4 (Darriba *et al.* 2012), according to the Akaike information criterion (AIC) (Akaike 1974). Phylogenetic congruence between ML trees was statistically assessed by performing the Shimodaira–Hasegawa test (Shimodaira and Hasegawa 1999) with RAXMLGUI 1.3 and CONSEL 0.1i software (Shimodaira and Hasegawa 2001).

Nucleotide sequence accession numbers

The DDBJ/EMBL/GenBank accession numbers for the partial sequences used in this study are as follows: for *dnaK*, KF751135 to KF751154, and KJ647313 to KJ647318; for *gyrB*, KF751155 to KF751175, and KJ647319 to KJ647324; and for *recA*, KF751115 to KF751134, and KJ647325 to KJ647330.

Results

RAPD fingerprinting

The RAPD genetic profiles indicated high genetic variability among strains studied, which were differentiated into 17 genetic groups at 90% similarity level (Fig. S1 and Table 1). However, European strains isolated during the 8-year period clustered into 12 genetic groups. A clear correlation between the genetic groups and geographical origin of the strains was not determined. Accordingly, the dominant group 5 was comprised of 36 strains originating from Serbia (27 strains), Italy (three strains), Poland (two strains), Bulgaria (one strain), Moldova (one strain), Montenegro (one strain) and Morocco (one strain) (Table 1). Furthermore, strains isolated from the same geographical regions belonged to different genetic clusters. For example, strains KFB249, KFB 253, KFB 254 and KFB 250 (or KFB 251 and KFB 252) isolated from Smederevo in Serbia belonged to RAPD groups 5, 7, 13 and 16 respectively (Fig. S1 and Table 1). A total of 28 representative strains, at least one per each RAPD group, were subjected to sequence analysis of housekeeping genes.

Nucleotide diversity

The G+C content for all three gene fragments was similar, and ranged from 56·3% (for *recA*) to 59·2% (for *gyrB*). From 28 strains, 15 different haplotypes (sequence types) were found. The number of haplotypes for each locus analysed was 12 (Table 3). The *recA* gene exhibited the highest proportion of polymorphic sites (12·41%), followed by *gyrB* (10·24%) and *dnaK* (9·68%). Nucleotide diversity (π) varied from 0·029 (for *gyrB*) to 0·041 (for *recA*). The dN/dS ratios for the three housekeeping genes were significantly below 1 (Table 3), indicating that these loci were under strong purifying selection. Tajima's D values were not significantly different from zero, supporting a neutral selection of the studied genes (Table 3).

Recombination analysis

The significance of intergenic recombination was estimated by determining the standardized index of association (I^S_A) in which zero indicates linkage equilibrium (freely recombining population), whereas a significant deviation from this value indicates a degree of linkage disequilibrium (clonal population) (Haubold *et al.* 1998). The value of I^S_A was significantly above zero (0·1807; $P < 0·05$), revealing the presence of linkage disequilibrium between the three loci used for the analysis. However, I^S_A value was low, thus suggesting the presence of recombination.

Phylogenetic network analysis (Huson and Bryant 2006) was used to investigate the impact of recombination for all individual genes and their concatenation. Neighbor-Net phylogenetic networks for all datasets showed certain degree of reticulation, which is indicative of recombination events (Figs 1 and S2). However, a lower degree of reticulation was observed for *recA* locus and concatenated dataset.

Furthermore, the Phi test indicated recombination within *dnaK* ($P < 0·005$) and the concatenation of genes ($P < 1E\text{-}15$). The GARD program detected two, three and one potential recombination breakpoint within the *dnaK*, *gyrB* and *recA* genes respectively. Recombination events were searched in all datasets using the seven

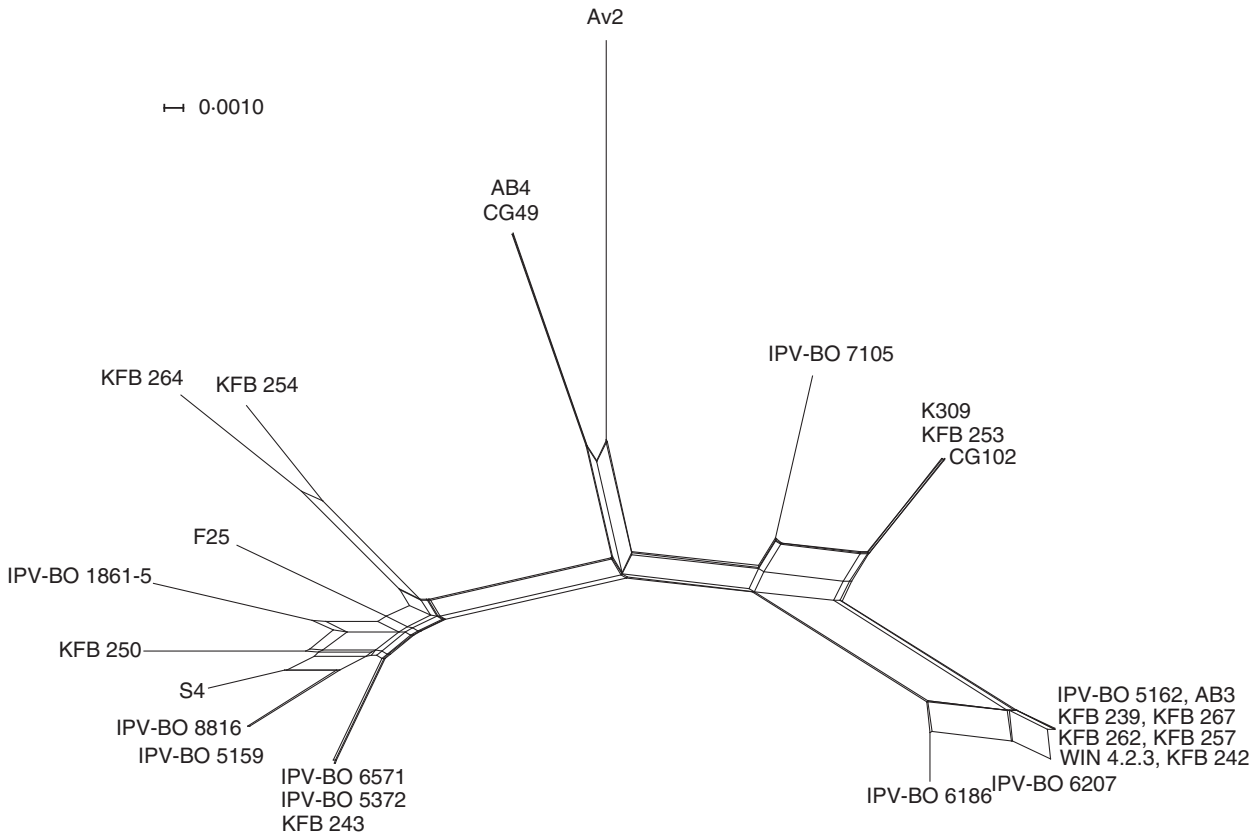

Figure 1 Neighbor-Net phylogenetic network based on a concatenated dataset of three loci *dnaK*, *gyrB* and *recA* showing the relationships between studied *A. vitis* strains. The scale bar represents 0·1% estimated sequence divergence.

Table 4 Detection of possible recombination events among *A. vitis* strains by using RDP4 software

			Potential parents		Breakpoint		
Event	Gene involved	Potential recombinant(s)	Major	Minor	Start	End	A*
1	*dnaK*	KFB 254, KFB 264	Unknown	IPV-BO 8816	94	465	4
2	*dnaK*	AB3, IPV-BO 5162, IPV-BO 6186, IPV-BO 6207, KFB 239, KFB 242, KFB 257, KFB 262, KFB 267	K309	Unknown	308	520	4 (6)†

*A, the number of methods supporting the recombination event.
†Number of methods supporting the recombination event when analysed with sequence of *dnaK* gene alone and concatenated dataset of three studied genes (shown in parentheses).

algorithms implemented in the RDP4 package (Martin *et al.* 2010). Two recombination events which involved *dnaK* gene portion were detected with four methods (Table 4). First recombination event was detected only when analysed sequence of *dnaK* gene alone, while the second one was found in concatenated sequence as well. However, the second recombination event was detected by six methods in concatenated dataset of three studied genes (Table 4).

Using ClonalFrame analysis the effect of recombination and point mutation on population structure was estimated (Didelot and Falush 2007; Vos and Didelot 2009). The ρ/θ value measures the population recombination rate per site (ρ) and the population mutation rate (θ), whereas the r/m value measures the relative impact of recombination (r) and mutation (m) in the genetic diversity of the analysed sequences. The relative rate of recombination to mutation (ρ/θ) was 0·54 for

concatenated dataset, indicating a prevalence of mutation over recombination in the evolution of analysed genes. However, the r/m value was found to be 2·24, suggesting that impact of recombination was about two times greater than the one of mutation in the evolution of the housekeeping genes analysed.

Phylogenetic analysis

The phylogeny of the 28 *A. vitis* strains was analysed by constructing ML trees from the sequences of each locus and concatenated data (Figs 2 and S3). As determined by Shimodaira-Hasegawa's test, the single locus trees were significantly different from each other ($P < 0{\cdot}05$), indicating that studied genes had different evolutionary history (Table 5). Furthermore, unlike the *recA* phylogenetic tree (Fig. 2), the *dnaK* and *gyrB* trees (Fig. S3a,b) were found to be congruent with the tree of concatenated dataset (Fig. S3c). Moreover, as evidence of recombination was found in *dnaK* and partly in *gyrB* gene, phylogenetic analysis based on *recA* gene sequence was used to infer phylogenetic relationship between strains. Four major

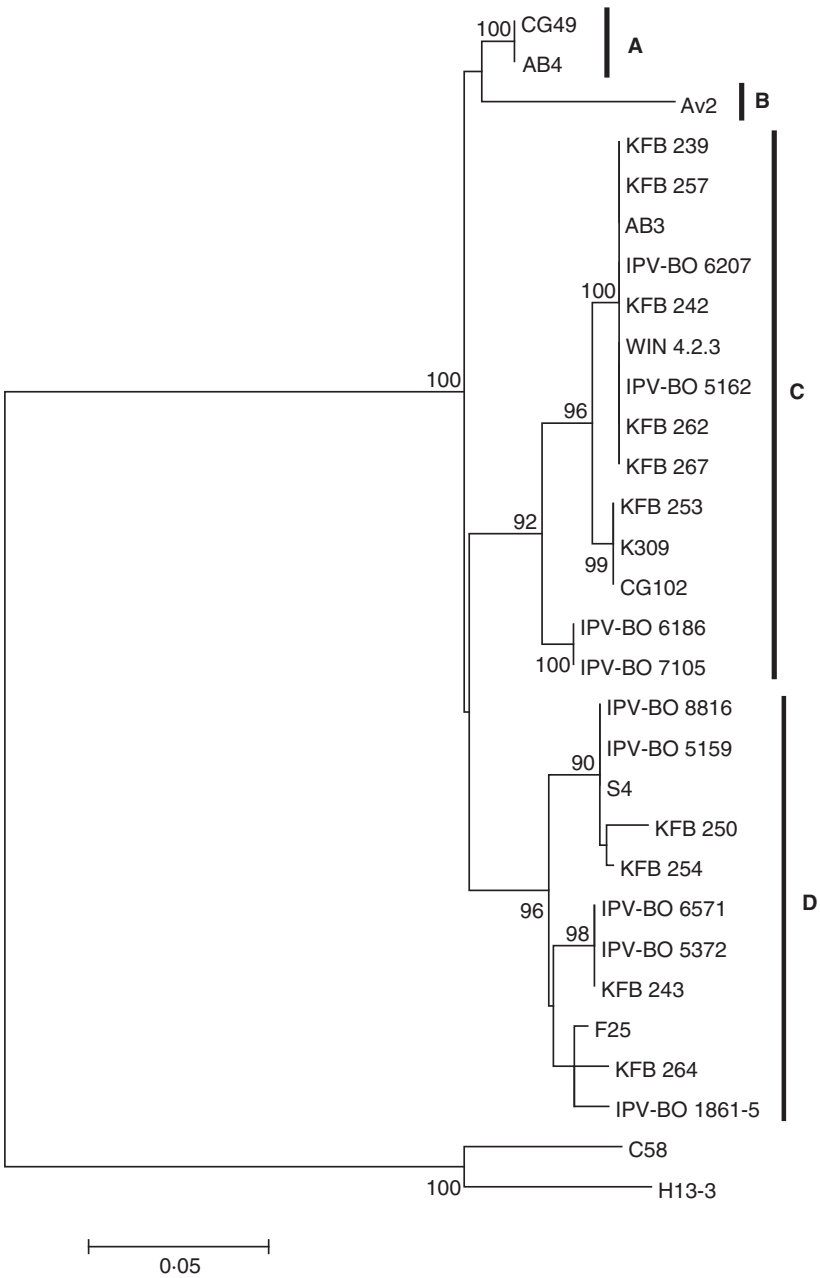

Figure 2 Maximum likelihood tree based on *recA* gene sequence indicating phylogenetic relationships between *A. vitis* strains studied. The tree was constructed using a general time reversible substitution model with a gamma distribution and invariant sites (GTR + G + I). *Agrobacterium tumefaciens* genomic species G1 (strain H13-3) and G8 (strain C58) are used as the outgroup organisms. The scale bar represents the number of substitutions per site.

Table 5 Shimodaira–Hasegawa test for congruency among maximum likelihood tree topologies for the three loci and their concatenated sequence

	P values			
Locus	*dnaK*	*gyrB*	*recA*	Concat
dnaK		0·0060*	0·0000*	0·2700
gyrB	0·0004*		0·0000*	0·0530
recA	0·0000*	0·0000*		0·0000*
Concat	0·0010*	0·2200	0·0340*	

*Tree topologies not supported ($P < 0·05$) by the corresponding dataset.

phylogenetic groups (A, B, C and D) strongly supported by bootstrap values were revealed among *A. vitis* strains studied (Fig. 2). The phylogenetic group A included reference strains AB4 and CG49, while strain Av2 was the only member of group B. The phylogenetic groups C and D were separated into several subgroups (Fig. 2). Phylogenetic groups A, B, C and D were also differentiated on the basis of *dnaK* and *gyrB* partial gene sequences, and concatenated sequence of three housekeeping genes studied (Fig. S3).

Discussion

Genetic diversity and epidemiological relationships among 54 strains of *A. vitis* isolated in eight European countries were analysed. As RAPD fingerprinting has already been proven to be effective for genotyping *A. vitis* (Irelan and Meredith 1996; Momol *et al.* 1998), we used this method to evaluate the extent of genetic diversity among studied strains and to select representative ones for further analysis. The strains studied were distributed into 12 genetic groups indicating high level of genetic diversity among them.

Generally, no clear relationship was found between RAPD groups and geographical origin of the strains. The RAPD group 5 was interesting; it was composed of 35 strains originating from six European countries and one from North Africa (Table 1). As these strains were isolated predominantly from young commercial vineyards during the period of 8 years (from 2003 to 2011), it suggests that they may have a common origin and were distributed following the movement of infected grapevine planting material.

Considerably high genetic diversity was also found in *A. vitis* by sequence analysis of three housekeeping genes (*dnaK*, *gyrB* and *recA*) of 28 representative strains. The percentages of polymorphic sites for the studied loci ranged from 9·68 to 12·41. These values are, however, higher than those reported for related species *Agrobacterium pusense* (i.e. *Ag. tumefaciens* genomic species G2) and *Ag. tumefaciens* genomic species G1, which ranged from 1·8 to 10·2% and 1·2 to 3·9% respectively (Aujoulat *et al.* 2011).

It is particularly important to analyse recombination when inferring phylogeny of bacteria, as recombinant genes can have disruptive influence on phylogenetic analysis (Martin *et al.* 2011). In the present work, several methods for detecting recombination were employed. Moreover, based on Neighbor-Net analysis, it is determined that *dnaK* and *gyrB* loci may be mainly affected by intragenic recombination. The value of r/m calculated for the concatenated dataset (r/m = 2·24) revealed an important role of recombination in the evolution of *A. vitis*, which impacted the sequence diversity two times more than mutations. However, in comparison to some other phytopathogenic bacteria, *A. vitis* has a higher r/m ratio than *Pseudomonas syringae* (r/m = 1·5) (Vos and Didelot 2009), *Ralstonia solanacearum* species complex (r/m = 1·6) (Wicker *et al.* 2012) and *Pseudomonas viridiflava* (r/m = 2·0) (Vos and Didelot 2009), but lower than *Xanthomonas axonopodis* (r/m = 3·18) (Mhedbi-Hajri *et al.* 2013), *Xylella fastidiosa* (r/m = 3·23) (Scally *et al.* 2005) and *Xanthomonas campestris* (r/m = 6·5) (Fargier *et al.* 2011). To the contrary, in the case of *Clavibacter michiganensis* subsp. *michiganensis*, the impact of mutation was greater than the impact of recombination (Jacques *et al.* 2012). Furthermore, recombination in *A. vitis* was found to occur less frequently than mutation ($\rho/\theta < 1$). The ρ/θ value was the same as calculated for *R. solanacearum* species complex ($\rho/\theta = 0·54$) (Wicker *et al.* 2012), contrasting with the results obtained for *Pseudomonas syringae* pv. *tomato* ($\rho/\theta = 5·847$) (Yan *et al.* 2008).

Moreover, recombination breakpoints were detected for all three studied genes by using GARD program. However, evidence of recombination was obtained only for *dnaK* locus by using the Phi test and algorithms available within RDP4 software package. Although recombination analyses included a limited number of strains, it seems that recombination is acting as an important driving force in the evolution of *A. vitis*. Recombination in *A. vitis* is probably facilitated by the biology of the pathogen and agricultural practices that are carried out in the production of grapevine propagation material. In fact, *A. vitis* survives systemically in the grapevine vascular system without showing any symptoms (latent infection) (Burr *et al.* 1998; Burr and Otten 1999). Therefore, utilization of propagation material originating from latently infected mother plants provides a direct contact between strains that may be present in rootstocks and scions. The practice of grafting between diverse scion and rootstock varieties is especially intensive due to agricultural globalization. In this way, opportunity to exchange genetic material

and gene flow between different genetic lineages of *A. vitis* could be high. Recombination events were also found to occur inside some genomic species of the *Ag. tumefaciens* complex (Aujoulat *et al.* 2011). However, I^S_A values calculated for genomic species of *Ag. tumefaciens* complex suggested that recombination rates were low.

Significant linkage disequilibrium among the *A. vitis* sequence types was detected; therefore, it is likely that recombination rate is not strong enough to completely disrupt the linkage between alleles. This may also be due to specificity of some genetic lineages of *A. vitis* that evolved separately, with limited exchange of genetic material. Indeed, the phylogenetic tree based on the concatenated dataset revealed highly homogenous clusters which contained strains with identical sequences for all three studied genes (Fig. S3c). One of these clusters was a subgroup within group D, which included strains isolated in Serbia (KFB 239, KFB 242, KFB 257, KFB 262 and KFB 267), Hungary (AB3), Italy (IPV-BO 5162) and Poland (WIN 4.2.3), over a period ranged from 1982 to 2011. Thus, it is likely these strains represent a clonal complex within *A. vitis*.

The observed incongruences between individual gene phylogenies also suggested the occurrence of recombination in *A. vitis* strains. However, unlike for *dnaK* and partly for *gyrB* locus, substantial evidence that *recA* gene fragment was a subject of recombination was not found. Therefore, *dnaK* and *gyrB* phylogenies, which were also congruent with the one based on the concatenated data set, are probably not suitable for the elucidation of phylogenetic relationship among *A. vitis* strains. On the other hand, *recA* gene could be a reliable phylogenetic marker to characterize relationship among *A. vitis* strains. The use of *recA* as a phylogenetic marker has also been reported as a robust and accurate tool for delineation of *Agrobacterium* spp., mainly strains belonging to *Ag. tumefaciens* genomic species (Costechareyre *et al.* 2010; Shams *et al.* 2013).

In conclusion, our data indicated a remarkable genetic diversity in *A. vitis* strains associated with recent crown gall outbreaks in Europe. However, the majority of geographically remote strains were genotypically related and most likely share a common origin. The results obtained illustrate the importance of pathogen dissemination via latently infected grapevine material and indicate a need for more stringent phytosanitary measures in international trade. However, there is a lack of standardized protocols for diagnosis and detection of *A. vitis* in grapevine. Although we investigated evolutionary history of *A. vitis* by using relatively few phylogenetic markers, substantial evidence for recombination was found. The agricultural globalization provided the possibilities of niche sharing among strains and increasing genetic exchanges between them.

Acknowledgements

This research was supported by the project III46008, financed by Ministry of Education, Science and Technological Development, Republic of Serbia, and EU Commission project AREA, No 316004. Nemanja Kuzmanović was supported by a JoinEU-SEE II scholarship (doctorate exchange) from Erasmus Mundus. The authors gratefully acknowledge Joanna Puławska and Sandor Süle for kindly providing *Agrobacterium vitis* strains. We are grateful to Xavier Didelot for helpful discussions on the ClonalFrame analysis.

Conflict of Interest

No conflict of interest declared.

References

Akaike, H. (1974) A new look at the statistical model identification. *IEEE Trans Automat Contr* **19**, 716–723.

Aujoulat, F., Jumas-Bilak, E., Masnou, A., Salle, F., Faure, D., Segonds, C., Marchandin, H. and Teyssier, C. (2011) Multilocus sequence-based analysis delineates a clonal population of *Agrobacterium* (*Rhizobium*) radiobacter (*Agrobacterium tumefaciens*) of human origin. *J Bacteriol* **193**, 2608–2618.

Bazzi, C., Minardi, P., Burr, T.J., Katz, B.H., Bishop, A.L. and Blanchard, L.M. (1988) Monoclonal and polyclonal antibodies in a comparative serological study of *Agrobacterium* Conn. biovars. *Phytopathol Mediterr* **27**, 51–56.

Bini, F., Geider, K. and Bazzi, C. (2008a) Detection of *Agrobacterium vitis* by PCR using novel *virD2* gene-specific primers that discriminate two subgroups. *Eur J Plant Pathol* **122**, 403–411.

Bini, F., Kuczmog, A., Putnoky, P., Otten, L., Bazzi, C., Burr, T.J. and Szegedi, E. (2008b) Novel pathogen-specific primers for the detection of *Agrobacterium vitis* and *Agrobacterium tumefaciens*. *Vitis* **47**, 181–189.

Bruen, T.C., Philippe, H. and Bryant, D. (2006) A simple and robust statistical test for detecting the presence of recombination. *Genetics* **172**, 2665–2681.

Burr, T.J. and Otten, L. (1999) Crown gall of grape: biology and disease management. *Annu Rev Phytopathol* **37**, 53–80.

Burr, T.J., Bishop, A.L., Katz, B.H., Blanchard, L.M. and Bazzi, C. (1987) A root-specific decay of grapevine caused by *Agrobacterium tumefaciens* and *A. radiobacter* biovar 3. *Phytopathology* **77**, 1424–1427.

Burr, T.J., Reid, C.L., Tagliati, E., Bazzi, C. and Sule, S. (1997) Biological control of grape crown gall by strain F2/5 is not associated with agrocin production or competition for attachment sites on grape cells. *Phytopathology* **87**, 706–711.

Burr, T.J., Bazzi, C., Süle, S. and Otten, L. (1998) Crown gall of grape: biology of *Agrobacterium vitis* and the development of disease control strategies. *Plant Dis* **82**, 1288–1297.

Costechareyre, D., Rhouma, A., Lavire, C., Portier, P., Chapulliot, D., Bertolla, F., Boubaker, A., Dessaux, Y. *et al.* (2010) Rapid and efficient identification of *Agrobacterium* species by *recA* allele analysis: *Agrobacterium recA* diversity. *Microb Ecol* **60**, 862–872.

Darriba, D., Taboada, G.L., Doallo, R. and Posada, D. (2012) jModelTest 2: more models, new heuristics and parallel computing. *Nat Methods* **9**, 772.

Didelot, X. and Falush, D. (2007) Inference of bacterial microevolution using multilocus sequence data. *Genetics* **175**, 1251–1266.

Fargier, E., Fischer-Le Saux, M. and Manceau, C. (2011) A multilocus sequence analysis of *Xanthomonas campestris* reveals a complex structure within crucifer-attacking pathovars of this species. *Syst Appl Microbiol* **34**, 156–165.

Genov, I., Atanassov, I., Tsvetkov, I. and Atanassov, A. (2006) Isolation and characterization of *Agrobacterium* strains from grapevines in Bulgarian vineyards and wild grapes, *V. vinifera* ssp. *silvestris*. *Vitis* **45**, 97–101.

Hampl, V., Pavlicek, A. and Flegr, J. (2001) Construction and bootstrap analysis of DNA fingerprinting-based phylogenetic trees with the freeware program FreeTree: application to trichomonad parasites. *Int J Syst Evol Microbiol* **51**, 731–735.

Haubold, B., Travisano, M., Rainey, P.B. and Hudson, R.R. (1998) Detecting linkage disequilibrium in bacterial populations. *Genetics* **150**, 1341–1348.

Higgins, D.G., Thompson, J.D. and Gibson, T.J. (1996) Using CLUSTAL for multiple sequence alignments. *Methods Enzymol* **266**, 383–402.

Huson, D.H. and Bryant, D. (2006) Application of phylogenetic networks in evolutionary studies. *Mol Biol Evol* **23**, 254–267.

Irelan, N.A. and Meredith, C.P. (1996) Genetic analysis of *Agrobacterium tumefaciens* and *A. vitis* using randomly amplified polymorphic DNA. *Am J Enol Vitic* **47**, 145–151.

Jacques, M.-A., Durand, K., Orgeur, G., Balidas, S., Fricot, C., Bonneau, S., Quillévéré, A., Audusseau, C. *et al.* (2012) Phylogenetic analysis and polyphasic characterization of *Clavibacter michiganensis* strains isolated from tomato seeds reveal that nonpathogenic strains are distinct from *C. michiganensis* subsp. *michiganensis*. *Appl Environ Microbiol* **78**, 8388–8402.

Jolley, K.A., Feil, E.J., Chan, M.S. and Maiden, M.C. (2001) Sequence type analysis and recombinational tests (START). *Bioinformatics* **17**, 1230–1231.

Jumas-Bilak, E., Michaux-Charachon, S., Bourg, G., Ramuz, M. and Allardet-Servent, A. (1998) Unconventional genomic organization in the alpha subgroup of the proteobacteria. *J Bacteriol* **180**, 2749–2755.

King, E.O., Ward, M.K. and Raney, D.E. (1954) Two simple media for the demonstration of pyocyanin and fluorescin. *J Lab Clin Med* **44**, 301–307.

Kosakovsky Pond, S.L., Posada, D., Gravenor, M.B., Woelk, C.H. and Frost, S.D. (2006) GARD: a genetic algorithm for recombination detection. *Bioinformatics* **22**, 3096–3098.

Kuzmanović, N., Ivanović, M., Prokić, A., Gašić, K., Zlatković, N. and Obradović, A. (2014) Characterization and phylogenetic diversity of *Agrobacterium vitis* from Serbia based on sequence analysis of 16S-23S rRNA internal transcribed spacer (ITS) region. *Eur J Plant Pathol* **140**, 757–768.

Librado, P. and Rozas, J. (2009) DNASP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**, 1451–1452.

Martin, D.P., Lemey, P., Lott, M., Moulton, V., Posada, D. and Lefeuvre, P. (2010) RDP3: a flexible and fast computer program for analyzing recombination. *Bioinformatics* **26**, 2462–2463.

Martin, D.P., Lemey, P. and Posada, D. (2011) Analysing recombination in nucleotide sequences. *Mol Ecol Resour* **11**, 943–955.

Mhedbi-Hajri, N., Hajri, A., Boureau, T., Darrasse, A., Durand, K., Brin, C., Saux, M.F.-L., Manceau, C. *et al.* (2013) Evolutionary history of the plant pathogenic bacterium *Xanthomonas axonopodis*. *PLoS ONE* **8**, e58474.

Momol, E.A., Burr, T.J., Reid, C.L., Momol, M.T., Hseu, S.H. and Otten, L. (1998) Genetic diversity of *Agrobacterium vitis* as determined by DNA fingerprints of the 5′-end of the 23S rRNA gene and random amplified polymorphic DNA. *J Appl Microbiol* **85**, 685–692.

Mousavi, S.A., Osterman, J., Wahlberg, N., Nesme, X., Lavire, C., Vial, L., Paulin, L., de Lajudie, P. *et al.* (2014) Phylogeny of the *Rhizobium-Allorhizobium-Agrobacterium* clade supports the delineation of *Neorhizobium* gen. nov. *Syst Appl Microbiol* **37**, 208–215.

Mousavi, S.A., Willems, A., Nesme, X., de Lajudie, P. and Lindström, K. (2015) Revised phylogeny of *Rhizobiaceae*: proposal of the delineation of *Pararhizobium* gen. nov., and 13 new species combinations. *Syst Appl Microbiol* **38**, 84–90.

Nei, M. and Li, W.H. (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc Natl Acad Sci* **76**, 5269–5273.

Ophel, K. and Kerr, A. (1990) *Agrobacterium vitis* sp. nov. for strains of *Agrobacterium* biovar 3 from grapevines. *Int J Syst Bacteriol* **40**, 236–241.

Page, R.D. (1996) TREEVIEW: an application to display phylogenetic trees on personal computers. *Comput Appl Biosci* **12**, 357–358.

Puławska, J., Willems, A. and Sobiczewski, P. (2006) Rapid and specific identification of four *Agrobacterium* species and biovars using multiplex PCR. *Syst Appl Microbiol* **29**, 470–479.

Ramírez-Bahena, M.H., Vial, L., Lassalle, F., Diel, B., Chapulliot, D., Daubin, V., Nesme, X. and Muller, D. (2014) Single acquisition of protelomerase gave rise to speciation of a

large and diverse clade within the *Agrobacterium*/*Rhizobium* supercluster characterized by the presence of a linear chromid. *Mol Phylogenet Evol* **73**, 202–207.

Scally, M., Schuenzel, E.L., Stouthamer, R. and Nunney, L. (2005) Multilocus sequence type system for the plant pathogen *Xylella fastidiosa* and relative contributions of recombination and point mutation to clonal diversity. *Appl Environ Microbiol* **71**, 8491–8499.

Shams, M., Vial, L., Chapulliot, D., Nesme, X. and Lavire, C. (2013) Rapid and accurate species and genomic species identification and exhaustive population diversity assessment of *Agrobacterium* spp. using *recA*-based PCR. *Syst Appl Microbiol* **36**, 351–358.

Shimodaira, H. and Hasegawa, M. (1999) Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Mol Biol Evol* **16**, 1114.

Shimodaira, H. and Hasegawa, M. (2001) CONSEL: for assessing the confidence of phylogenetic tree selection. *Bioinformatics* **17**, 1246–1247.

Silvestro, D. and Michalak, I. (2012) raxmlGUI: a graphical front-end for RAxML. *Org Divers Evol* **12**, 335–337.

Slater, S.C., Goldman, B.S., Goodner, B., Setubal, J.C., Farrand, S.K., Nester, E.W., Burr, T.J., Banta, L. *et al.* (2009) Genome sequences of three agrobacterium biovars help elucidate the evolution of multichromosome genomes in bacteria. *J Bacteriol* **191**, 2501–2511.

Szegedi, E., Czakó, M., Otten, L. and Koncz, C.S. (1988) Opines in crown gall tumours induced by biotype 3 isolates of *Agrobacterium tumefaciens*. *Physiol Mol Plant Pathol* **32**, 237–247.

Tajima, F. (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**, 585–595.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. and Kumar, S. (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* **28**, 2731–2739.

Vos, M. and Didelot, X. (2009) A comparison of homologous recombination rates in bacteria and archaea. *ISME J* **3**, 199–208.

Wicker, E., Lefeuvre, P., de Cambiaire, J.-C., Lemaire, C., Poussier, S. and Prior, P. (2012) Contrasting recombination patterns and demographic histories of the plant pathogen *Ralstonia solanacearum* inferred from MLSA. *ISME J* **6**, 961–974.

Yan, S., Liu, H., Mohr, T.J., Jenrette, J., Chiodini, R., Zaccardelli, M., Setubal, J.C. and Vinatzer, B.A. (2008) Role of recombination in the evolution of the model plant pathogen *Pseudomonas syringae* pv. *tomato* DC3000, a very atypical tomato strain. *Appl Environ Microbiol* **74**, 3171–3181.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Figure S1 Similarity dendrogram based on cluster analysis generated by RAPD-PCR fingerprints of studied *A. vitis* strains.

Figure S2 Neighbor-Net phylogenetic network based on *dnaK* (a), *gyrB* (b) and *recA* (c) gene sequences showing the relationships between studied *A. vitis* strains.

Figure S3 Maximum likelihood trees based on *dnaK* (a) and *gyrB* (b) gene sequences, and concatenated sequences of three housekeeping genes studied (c) indicating phylogenetic relationships between studied *A. vitis* strains.