Adaptation of genetically monomorphic bacteria: evolution of copper resistance through multiple horizontal gene transfers of complex and versatile mobile genetic elements

D. RICHARD,*†‡ V. RAVIGNÉ,* A. RIEUX,* B. FACON,§¶ C. BOYER,* K. BOYER,* P. GRYGIEL,* S. JAVEGNY,* M. TERVILLE,* B. I. CANTEROS,** I. ROBÉNE,* C. VERNIÈRE,†† A. CHABIRAND,‡ O. PRUVOST* and P. LEFEUVRE*

*UMR PVBMT, CIRAD, F-97410 St Pierre, Réunion, France, †Plant Health Laboratory, ANSES, F-97410 St Pierre, Réunion, France, ‡Université de la Réunion, UMR PVBMT, F-97490 St Denis, Réunion, France, §INRA, UMR PVBMT, F-97410 St Pierre, Réunion, France, ¶INRA, UMR CBGP, F-34090 Montpellier, France, **INTA, Estación Experimental Agropecuaria Bella Vista, Bella Vista, Argentina, ††CIRAD, UMR BGPI, F-34398 Montpellier, France

Abstract

Copper-based antimicrobial compounds are widely used to control plant bacterial pathogens. Pathogens have adapted in response to this selective pressure. Xanthomonas citri pv. citri, a major citrus pathogen causing Asiatic citrus canker, was first reported to carry plasmid-encoded copper resistance in Argentina. This phenotype was conferred by the copLAB gene system. The emergence of resistant strains has since been reported in Réunion and Martinique. Using microsatellite-based genotyping and copLAB PCR, we demonstrated that the genetic structure of the copper-resistant strains from these three regions was made up of two distant clusters and varied for the detection of copLAB amplicons. In order to investigate this pattern more closely, we sequenced six copper-resistant X. citri pv. citri strains from Argentina, Martinique and Réunion, together with reference copper-resistant Xanthomonas and Stenotrophomonas strains using long-read sequencing technology. Genes involved in copper resistance were found to be strain dependent with the novel identification in X. citri pv. citri of copABCD and a cus heavy metal efflux resistance–nodulation–division system. The genes providing the adaptive trait were part of a mobile genetic element similar to Tn3-like transposons and included in a conjugative plasmid. This indicates the system’s great versatility. The mining of all available bacterial genomes suggested that, within the bacterial community, the spread of copper resistance associated with mobile elements and their plasmid environments was primarily restricted to the Xanthomonadaceae family.

Keywords: contemporary adaptation, heavy metal resistance, mobile DNA, plasmid, Stenotrophomonas, Xanthomonas

Received 5 June 2016; revision received 28 November 2016; accepted 8 December 2016

Introduction

The dynamics of genome evolution differ considerably across the bacterial diversity spectrum. Some bacteria display a level of polymorphism that makes it possible to distinguish strains from a few housekeeping genes.
On the other hand, genetically monomorphic bacteria appear very similar and require full genome sequencing to uncover some diversity (Achtman 2008). Far from being anecdotic cases, genetically monomorphic bacteria include numerous successful species, including several human pathogens, such as *Mycobacterium tuberculosis*, *Yersinia pestis* or *Escherichia coli* O157:H7 (Achtman 2008), as well as plant pathogens, such as *Pseudomonas syringae* pv. *tomato* (Cai et al. 2011) or *Xanthomonas citri* pv. *citri* (Leduc et al. 2015).

Since the discovery of genetically monomorphic bacteria, they have been considered to be an evolutionary puzzle (Achtman 2008). Genetic variability drives adaptation, which would suggest that these bacteria have a low adaptive potential. However, many have caused worldwide epidemics, demonstrating an ability to thrive in contrasted environmental conditions. For years, a variety of antibiotics have been used to control genetically monomorphic bacteria that are pathogenic to humans. Yet, the bacteria have regularly succeeded in evolving antibiotic resistance through various molecular mechanisms (Davies & Davies 2010). Genome sequencing projects have led to the discovery of the few polymorphisms responsible for some of the bacterial adaptations. For instance, ciprofloxacin resistance in *Y. pestis* and rifampicin resistance in *M. tuberculosis* are due to the mutation of a single gene on their chromosome (Telenti et al. 1993; Hurtle et al. 2003). However, in many cases, the genes responsible for adaptation were actually horizontally acquired from the bacterial mobile gene pool (Galimand et al. 1997; Nonaka et al. 2012; Behlau et al. 2013).

Horizontal gene transfer (HGT, Bryant et al. (2012)) is the incorporation of novel DNA into the host genome. In many bacterial species, HGT is so frequent that the high number of alien genes compared with the number of core genes of a species has led to the development of the concept of an open pan-genome (García-Vallve et al. 2000; Chaudhuri et al. 2010). HGT can occur through conjugation (i.e. DNA transfer via plasmids or integrative conjugative elements, ICE), transformation (i.e. the uptake of naked DNA from the environment), transduction (i.e. the incorporation of DNA through a bacteriophage) and various gene transfer agents (e.g. phage-like structures that are produced by a donor cell and released into the environment (Popa & Dagan 2011)). Although we do not yet fully understand the mechanism involved, nanotubes were recently highlighted because they allow DNA and protein transfer between bacterial cells (Dubey & Ben-Yehuda 2011). Using network theory-based analyses of shared gene content among bacterial species, plasmids were shown to be of prime importance in HGT (Halary et al. 2010). In genetically monomorphic bacteria, alien genes and, therefore, traces of HGT were shown to be largely restricted to bacteriophages and plasmids (Achtman 2012).

Plasmids are composed of backbone genes, essential to their evolutionary dynamics (i.e. genes involved in conjugation, replication control and stable inheritance). The presence of assisted or autonomous mobility genetic elements determines the plasmids’ ability to be mobilizable or conjugative (Smillie et al. 2010). Due to their relative stability in the plasmid genome, backbone genes may show traces of adaptation to bacterial hosts (Norberg et al. 2011). Besides backbone genes, plasmids typically carry accessory genes encoding traits that may be adaptive for the bacterial host (Pansegrau et al. 1994; Sen et al. 2013). Accessory genes are often located on mobile genetic elements themselves, either transposons (Trefault et al. 2004; Haines et al. 2007) or integrons (Tennstedt et al. 2005), defining hot spots of insertions within the plasmid genome. These elements are generally autonomous and can be transferred between genera, families and even kingdoms (Heinemann & Sprague 1989). Plasmid-encoded accessory genes enrich the species’ gene pool and provide new adaptive traits in response to environmental modifications, such as the use of a new antibiotic or the colonization of a new ecological niche (e.g. Hobman & Crossman 2015; Ochman et al. 2000; Weinhall et al. 2007). Therefore, it is now obvious that in genetically monomorphic pathogens as in others, in order to further our understanding of bacterial adaptation and ecology, we need to understand plasmid evolutionary dynamics, that is deciphering the genomic structure of plasmids and the mechanisms involved in plasmid spread among bacteria. While several studies have already documented the history of specific accessory genes (Perry & Wright 2013), very few studies have documented the evolution of the whole plasmid genome for a given ecological function (but see Norberg et al. (2011) for a study of the backbone genome of IncP-1 plasmid family).

Among the traits typically coded by plasmid accessory genes, antibiotic and heavy metal resistance has a tremendous impact on human, animal and plant health (Davies & Davies 2010; Seiler & Berendonk 2012). In agriculture worldwide, fungal and bacterial pathogens that cause damage to vegetable and fruit crops have been controlled chemically for over a century, with the intensive use of copper-based pesticides (Smith 1985). This has led to the development of copper resistance (Cu²⁺) in several plant pathogenic bacterial species, primarily in the *Pseudomonas* and *Xanthomonas* genera, two members of the Gammaproteobacteria class (Cooksey et al. 1990). For example, in most citrus-growing areas, Asiatic citrus canker represents a significant risk that is both direct (yield decrease, alteration of fruit quality and partial defoliation) and indirect (restrictions to...
fresh fruit export because of the quarantine status of the pathogen (Graham et al. 2004; Gottwald et al. 2007). The disease is caused by the genetically monomorphic bacterium X. citri pv. citri, which is disseminated through contaminated water (splashing and wind-driven rains), tools and plant materials (particularly over long distances). Given its Asiatic origin, X. citri pv. citri has geographically expanded and become established in many citrus-producing areas, including Florida (USA), South America and, recently, Africa (Pruvost et al. 2014; Leduc et al. 2015). In regions where X. citri pv. citri has been established for a long time, control typically involves integrated pest management strategies that partly rely on repetitive applications of copper-based pesticides (Graham et al. 2004). X. citri pv. citri copper resistance first emerged in 1994 in Argentina and it has persisted. There were no other reports of X. citri pv. citri copper resistance until it was discovered in the French islands of Réunion (a French territory in the South West Indian Ocean) and Martinique (a French territory in the Eastern Caribbean Sea) in 2014, two decades later (Richard et al. 2016, 2017). The causative genes of Argentinian X. citri pv. citri copper resistance, identified as copLAB, were shown to be located on a large transmissible plasmid (Behlau et al. 2011, 2012), suggesting that HGT has a role. Indeed, exactly the same gene set conferring copper resistance was identified in several other Xanthomonas species and pathovars. In addition, the hypothesis of an HGT is endorsed by evidence of the major incongruence between the phylogeny of the concatenated copLAB genes of Xanthomonas species and the species’ phylogeny (Behlau et al. 2013).

Although it is tempting to conclude that the mechanistic basis of copper resistance has been elucidated in X. citri pv. citri, a recent study in a close relative species has revealed that the genetic basis for copper resistance at the infraspecific level can be variable. In Xanthomonas arboricola pv. juglandis, copper resistance could be conferred by plasmid-borne copLAB (Giovanardi et al. 2016), as well as by a copABCD gene set located on the chromosome (Lee et al. 1994), which is sometimes associated with an ICE structure (Cesbron et al. 2015). The copABCD system has been described in detail on a plasmid from other bacterial species, such as the tomato pathogen P. syringae pv. tomato (Mellano & Cooksey 1988) or E. coli (Brown et al. 1995). The diversity of gene sets involved in copper resistance as well as the diversity of the physical supports of these genes raise the question of the nature and origin of copper resistance in X. citri pv. citri populations.

This study aimed to improve our understanding of the genetic basis for the adaptation to copper in X. citri pv. citri. To achieve this, we gathered an unprecedented collection of strains from the three known outbreaks of copper resistance to date: Argentina, Réunion and Martinique. By combining population genetics based on microsatellite and minisatellite markers, comparative genomics and network analysis, we address the following questions: How do copper-resistant strains from the three outbreaks relate to each other? Are the genes responsible for copper resistance similar for the different outbreaks of X. citri pv. citri? What is the genomic environment of these genes and how does it vary within pathovar citri and among xanthomonads? We discuss the evolutionary implications of our findings and propose future lines of research to increase our understanding of how adaptive genes spread in bacterial communities.

Materials and methods

Bacterial strains

The study included a total of 350 Xanthomonas citri pv. citri CuR strains from the three known outbreaks of copper resistance (Argentina, n = 111; Réunion, n = 219; and Martinique, n = 20). Argentinian strains originated from a collection held at INTA and were collected from 1994 (the first outbreak) to 2015. Starting in early 2014, when copper resistance was discovered in Réunion, strains were collected in the field directly from active epidemics. In addition, a large collection of citrus strains from Réunion (held at the PVBMT laboratory and including strains collected since 1978) was tested for copper resistance. This revealed resistance in 219 CuR strains from all locations in Réunion, as well as the oldest copper-resistant strains found, dating from 2010. Martinique was a disease-free territory until the 2014 outbreak (Richard et al. 2016). Every strain characterized was CuR and 20 were collected from the first disease focus in Morne Rouge.

In addition to CuR haplotypes (n = 115 identified among 219 strains), all known haplotypes corresponding to copper-susceptible (CuS) X. citri pv. citri strains sampled in Réunion (n = 524) from all the citrus-producing areas were also included in the study to assess the genetic relatedness between CuR and CuS strains on the island. This extensive dataset was built from strain collections sampled over two periods, 1978–1997 (historical strains) and 2009–2015 (contemporary epidemics).

Finally, as CuR xanthomonads were originally reported in strains causing bacterial spot of tomato and pepper and no genomic resources were available for such strains, reference CuR strains from the four species causing this disease were included in this study. They originated from Argentina, Mauritius, New Zealand and the USA (Table 1).
Table 1 Characteristics of copper-resistant bacterial strains used for long-read sequencing

<table>
<thead>
<tr>
<th>Strain</th>
<th>Other numbers</th>
<th>Accession</th>
<th>Genus</th>
<th>Species</th>
<th>Pathovar</th>
<th>Copper-resistance location</th>
<th>Country</th>
<th>Date</th>
<th>Host</th>
</tr>
</thead>
<tbody>
<tr>
<td>LMG930</td>
<td></td>
<td>CP018463-CP018467</td>
<td>Xanthomonas</td>
<td>euvesicatoria</td>
<td>Plasmid</td>
<td>USA</td>
<td>1969</td>
<td>Pepper</td>
<td></td>
</tr>
<tr>
<td>LMG911</td>
<td></td>
<td>CP018725-CP018727</td>
<td>Xanthomonas</td>
<td>vesicatoria</td>
<td>Plasmid</td>
<td>New Zealand</td>
<td>1955</td>
<td>Tomato</td>
<td></td>
</tr>
<tr>
<td>LM091</td>
<td></td>
<td>CP017483</td>
<td>Stenotrophomonas sp.</td>
<td>citri</td>
<td>Plasmid</td>
<td>Réunion</td>
<td>2015</td>
<td>Tangor</td>
<td></td>
</tr>
<tr>
<td>LM199</td>
<td>Xcc 15-4632 (INTA)</td>
<td>MSQV00000000</td>
<td>Xanthomonas</td>
<td>citri</td>
<td>Plasmid</td>
<td>Argentina</td>
<td>2015</td>
<td>Orange</td>
<td></td>
</tr>
<tr>
<td>LM180</td>
<td>Xcc 03-1638 (INTA); A44*</td>
<td>MSQW00000000</td>
<td>Xanthomonas</td>
<td>citri</td>
<td>Plasmid</td>
<td>Argentina</td>
<td>2003</td>
<td>Grapefruit</td>
<td></td>
</tr>
<tr>
<td>LM159</td>
<td>Bv-5-4a (INTA)</td>
<td>CP018468-CP018471</td>
<td>Xanthomonas</td>
<td>vesicatoria</td>
<td>Plasmid</td>
<td>Argentina</td>
<td>1987</td>
<td>Pepper</td>
<td></td>
</tr>
<tr>
<td>LL074-4</td>
<td></td>
<td>CP018847-CP018849</td>
<td>Xanthomonas</td>
<td>citri</td>
<td>Plasmid</td>
<td>Martinique</td>
<td>2014</td>
<td>Grapefruit</td>
<td></td>
</tr>
<tr>
<td>LJ207-7</td>
<td></td>
<td>CP018850-CP018853</td>
<td>Xanthomonas</td>
<td>citri</td>
<td>Plasmid</td>
<td>Réunion</td>
<td>2012</td>
<td>Kaffir lime</td>
<td></td>
</tr>
<tr>
<td>LH3</td>
<td></td>
<td>CP018472-CP018476</td>
<td>Xanthomonas</td>
<td>'perforans'*</td>
<td>Plasmid</td>
<td>Maurice</td>
<td>2010</td>
<td>Tomato</td>
<td></td>
</tr>
<tr>
<td>LH276</td>
<td></td>
<td>CP018854-CP018857</td>
<td>Xanthomonas</td>
<td>citri</td>
<td>Plasmid</td>
<td>Réunion</td>
<td>2010</td>
<td>Kaffir lime</td>
<td></td>
</tr>
<tr>
<td>LH201</td>
<td></td>
<td>CP018858-CP018860</td>
<td>Xanthomonas</td>
<td>citri</td>
<td>Plasmid</td>
<td>Réunion</td>
<td>2010</td>
<td>Kaffir lime</td>
<td></td>
</tr>
<tr>
<td>J549-3</td>
<td></td>
<td>CP018728-CP018730</td>
<td>Xanthomonas</td>
<td>gardneri</td>
<td>Plasmid</td>
<td>Réunion</td>
<td>1997</td>
<td>Tomato</td>
<td></td>
</tr>
<tr>
<td>ICMP7383</td>
<td></td>
<td>CP018731-CP018734</td>
<td>Xanthomonas</td>
<td>gardneri</td>
<td>Plasmid</td>
<td>New Zealand</td>
<td>1980</td>
<td>Tomato</td>
<td></td>
</tr>
</tbody>
</table>

ICMP (International Collection of Microorganisms from Plants, Landcare Research, Auckland, New Zealand). BCCM/LMG (Belgian Coordinated Collections of Microorganisms, University of Ghent, Belgium). INTA (Instituto Nacional de Tecnología Agropecuaria).

*As designated by Behlau et al. (2011).

'X. 'perforans' was reclassified as X. euvesicatoria' (Constantin et al. 2016).
Evaluation of copper resistance

To assess copper resistance, we used the growth test on YPGA (yeast extract 7 g/L, casein peptone 7 g/L, glucose 7 g/L, agar 18 g/L, pH 7.2), supplemented with CuSO₄·5H₂O (470 mg/L), and PCR using copL primers. We slightly modified the PCR protocol reported previously (Behlau et al. 2013) using GoTaq Flexi polymerase (Promega), a lower primer concentration (5 pmol/μL) and a higher annealing temperature (66 °C). In the case of negative results using copL primers, copA and copB primers were also used, as reported previously (Behlau et al. 2013).

Minisatellite and microsatellite genotyping

To determine how the studied strains relate to the known diversity of X. citri pv. citri worldwide, a subset of available strains from Argentina (47 CuR and 45 CuS), Martinique (9 CuR) and Réunion (64 CuR and 103 CuS) were genotyped using 31 minisatellites (MLVA-31) developed for global epidemiology analyses and compared with the known worldwide diversity (http://www.biopred.net/MLVA/). Genotyping and discriminant analysis of principal components were performed as reported earlier (Pruvost et al. 2014). MLVA fragment sizes were transformed to random repeat numbers (rounding to the superior integer). Manhattan distances between strains were then calculated and multidimensional scaling (MDS) plots were computed using the BIOS2MDS R package (Pele et al. 2012). Categorical minimum-spanning trees (MST) were built using the algorithm recommended for tandem repeat data combining global optimal EBURST (goEBURST) and Euclidean distances in PHYLLOVIZ v1.0 (Francisco et al. 2012).

The relatedness between strains was assessed at smaller spatiotemporal scales (i.e. at the country level) by examining the diversity of all X. citri pv. citri strains using a set of 14 microsatellites (MLVA-14), which represent the most discriminative X. citri pv. citri genotyping technique available for infrapathotype typing (Bui Thi Ngoc et al. 2009). Population genetics analyses were performed on two different datasets, as explained above. Moreover, allelic richness was computed using a rarefaction method with the HIERFSTAT R package (Goudet 2005). Population pairwise RST values were computed for estimating genetic differentiation among CuR populations at the country level. Significance was tested with 1000 permutations using ARLEQUIN 3.5.2.2 (Excoffier et al. 2005). The first dataset, which consisted of all CuR strains from the three resistance outbreaks, allowed us to decipher the genetic relationships between the three outbreaks. The second dataset, consisting of all CuR and CuS strains from Réunion (the only region where we have extensively sampled local populations), enabled us to examine the relationships between the CuR and CuS strains that were isolated both before and during the CuR X. citri pv. citri epidemics. When comparing CuR and CuS strains from Réunion, the optimal number of clusters in the dataset was assessed by computing silhouette scores from multiple K-means runs using the BIOS2MDS R package.

DNA sequencing and assembly

Based on strain diversity analyses and the Cu profiles (i.e. the combination of the growth test and PCR results), six X. citri pv. citri strains from Argentina, Martinique and Réunion were selected for sequencing (Table 1). In addition, six other reference strains of CuR xanthomonads causing bacterial spot disease on tomato and/or pepper and a CuS strain of Stenotrophomonas sp., a commensal bacterium collected from the citrus phyllosphere in Réunion (i.e. the same ecological niche as X. citri pv. citri), were also selected for genome sequencing. In total, 13 genomes were completely sequenced using the long-read PacBio RSII technology, using one SMRT cell for each strain (Table 1). Assembly of the resulting raw reads was made using SMRT ANALYSIS HGAP v. 2.3 protocol with default parameters. An additional step of contig circularization was conducted using a combination of minimus assembler (Sommer et al. 2007) and the SMRT ANALYSIS resequencing v. 1 protocol.

Sequence annotations

Sequence data obtained from the 13 strains were screened for known copper-resistance systems using BLASTN and BLASTP algorithms. As for most strains, CuR genes were found on a 230-kb plasmid with strong homology. We arbitrarily selected strain LH201 as a reference. Hereafter, this plasmid is referred to as pLH201.1. We carried out pH201.1 CDS prediction and automatic gene annotation using the MaGe genome annotation platform (Vallenet et al. 2006) and then manual curation of the plasmid annotations using the extensive set of tools and databases available via the platform.

Specific plasmid features were characterized in more detail. We conducted a search of plasmid toxin–antitoxin systems using the online databases TADB (Shao et al. 2011) and RASTA (Sevin & Barloy-Hubler 2007). The conjugative apparatus of pLH201.1 (encoded by 16 tra genes dispatched in three clusters) was blasted against amino acid sequences of (i) the plasmid database of NCBI and (ii) the ICEberg database (Wozniak & Waldor 2010). We only retained hits
that displayed homology with at least one gene (on at least 70% of its length) from each of the three gene clusters involved in conjugation. Finally, Isfinder (Siguier et al. 2006) was used to detect known insertion sequences.

We used Circos (Krzywinski et al. 2009) to plot a graphical representation of the plasmid. We computed GC% and GC skew [(G+C)/(G–C)] using a custom perl script available at https://github.com/DamienFr/GC-content-in-sliding-window.

The search for homologous sequences

In our search for sequences that are homologous to pLH201.1 among all the available sequences, we queried the Whole Genome Shotgun (WGS) and Non-Redundant (NR) public nucleotide databases. Results were automatically retrieved using BioPerl modules (SEARCHIO, SEQIO and EUUTILITES).

We first used Cytoscape v. 3.3.0 to build a network of the gene sequences that were homologous with pLH201.1. We retained all the sequences that had a hit with at least one of the pLH201.1 genes (with 95% nucleotide identity over 95% of gene length). All the hits were blasted against each other to obtain the full network structure. The network was then plotted using an edge-weighted spring-embedded layout, which positions the nodes closer if the number of genes they share is higher.

We adopted a clustering strategy based on average nucleotide identity (ANI) to determine the diversity of sequences related to pLH201.1 and eliminate the redundancy of the hit dataset. Nucleotide sequences that shared more than 10 genes with pLH201.1 (with 70% nucleotide identity over 70% of gene length) were clustered using a modified ANI algorithm (available at https://github.com/DamienFr/Clustering_with_ANI) that takes into account the identity of the homologous region in relation to the size of the complete sequence. For each sequence that was representative of a cluster, CDS prediction was performed using Prodigal (Hyatt et al. 2010). Sequences were then compared using a combination of MAUVE (Darling et al. 2004), MAFFT (Katoh & Standley 2013) and the Ape (Popescu et al. 2012) and GenoPlotR (Guy et al. 2010) R packages.

Results

In this study, we combined PCR screening, micro- and minisatellite typing and comparative genomics to place the strains that caused the outbreak into the known diversity and to characterize copper-resistance genetic and genomic support.

Preliminary characterization of CuR outbreak strains and placement within the pathovar diversity

Most CuR Xanthomonas citri pv. citri strains (98%) produced the expected amplicon when assayed by PCR with copL primers. Notably, seven strains from Argentina grew on YPGA supplemented with CuSO4 although they were PCR negative when assayed with copL, copA and copB primers.

Strains sampled from the three outbreaks were assessed by MLVA-31 minisatellites and were all assigned to the same genetic cluster (formerly called DAPC1 in Pruvost et al. (2014)).

At the spatiotemporal evolutionary scale defined by these markers, strains from the three origins formed a single clonal complex (CC), that is a network of haplotypes composed of single-locus variants. A total of 34 haplotypes were delineated with no clear structure in terms of geographical origin or copper phenotype (Fig. S1, Supporting information; information on haplotypes available at http://www.biopred.net/MLVA/). Strains from Martinique were assigned to haplotypes also including strains from Réunion. Some MLVA-31 haplotypes were shared by CuR strains from Martinique and Réunion. Most CuR strains from Réunion were assigned to two frequent haplotypes (#173 and 175), which also contained CuS strains of the same origin. Similarly, haplotype #51, the most frequent haplotype among Argentinean strains, included strains that differed in terms of the copper phenotype and/or cop PCR.

Relationships between outbreaks

Using microsatellite data (MLVA-14), we first characterized the genetic diversity of CuR X. citri pv. citri strains originating from the three resistance outbreaks. A total of 83, eight and 115 haplotypes were identified among the 111, 20 and 219 CuR strains from Argentina, Martinique and Réunion, respectively. Using a rarefaction method (n = 20), strains from Argentina, Martinique and Réunion showed a mean allelic richness of 3.96, 1.71 and 2.76, respectively. Two clusters were identified according to their distribution on the first axis of the MDS plot (68.0% of the total variance). Their identification was endorsed by the silhouette score derived from multiple K-means runs (Fig. 1A). Axes 2–5 contributed much less to total variance, ranging from 6.2% to 2.2%. Cluster 1 included all strains from Martinique and Réunion, while cluster 2 only had Argentinean strains. CuR strains that were PCR negative for copLAB also grouped in cluster 2 without any apparent substructure. Within cluster 1, strains from Martinique and Réunion were moderately but significantly differentiated
In contrast, strains from Argentina (i.e. cluster 2) were strongly differentiated ($R_{ST} = 0.85; P < 0.001$) from strains collected in Martinique and Réunion (i.e. cluster 1).

**Relationships between copper-resistant and copper-susceptible strains**

Microsatellite data (MLVA-14) were also used to decipher the genetic relatedness of the $Cu^R$ and $Cu^S$ strains from Réunion. Silhouette scores derived from multiple K-means runs suggested a lack of strong structure in the dataset. For $K > 2$, strain assignation probabilities to clusters were often low (i.e. < 0.5). We decided to retain $K = 2$ because minimal strain assignation probabilities to clusters were > 0.8. All but two $Cu^R$ strains were assigned to cluster B, including almost exclusively $Cu^S$ strains (Fig. 1B). In a few cases, $Cu^R$ and $Cu^S$ strains shared the same haplotype. Most $Cu^R$ strains were genetically closely related. They were structured as five CCs and 27 singletons, that is haplotypes sharing no single-locus variation with others (Fig. S2, Supporting information). Among these, a major CC (CC1) comprised 64 haplotypes, representing 152 strains (69%). An additional set of 48 strains (37 haplotypes), corresponding to smaller CC and singletons, consisted of double-locus variants of CC1. These strains did not join the main CC, which is probably due to incomplete sampling. Both clusters included historical (1978–1997) and contemporary strains ($\geq 2009$). The $Cu^S$ D07 strain (sampled from satsuma mandarin in 1989) was the closest historical strain to the main group of $Cu^R$ strains (i.e. a double-locus variant).

**Sequencing**

After PacBio reads assembly of the 13 fully sequenced strains, we obtained from one to 18 contigs per strain (see Table S1, Supporting information for details). It is important to note that all the chromosome sequences were successfully circularized and 34 of the remaining 70 contigs were circularized into plasmids. No plasmid was detected in the copper-resistant commensal strain of *Stenotrophomonas* sp. isolated from citrus in Réunion and one to four plasmids were detected in xanthomonads, depending on the strain.

**Features of plasmids associated with copper resistance in xanthomonads**

*Cop* genes were plasmid-borne for all the *Xanthomonas* strains sequenced. Consistent with previous WGS data
from _Stenotrophomonas maltophilia_, these genes were present on the chromosome of the strain of _Stenotrophomonas_ sp. isolated from citrus phyllosphere (Crossman et al. 2008; Davenport et al. 2014; Pak et al. 2015). A MAUVE comparison suggested that, with the exception of _Xanthomonas euvesicatoria_ LMG930, all other plasmids bearing cop genes were genetically related (Fig. 2). Consequently, the plasmid from the _X. citri_ pv. _citri_ strain LH201 was arbitrarily selected as our reference. MAGE annotation of this plasmid revealed two tRNA and 258 CDS, 176 of which (68%) encoded proteins of unknown function (Fig. 3). We identified the plasmid replication initiator gene trfA near a GC-skew switch. This commonly indicates the origin of replication (oriV) (Grigoriev 1998), suggesting that the oriV locus is located nearby. We revealed an 87-amino acid-long conserved domain of a putative HigB-like addiction module killer toxin (e-value < 10\^-10\) (Schuessler et al. 2013). A 222-amino acid-long putative transcriptional regulator of the xenobiotic response element family, which might serve as an antitoxin protein, was found downstream and antisense to the toxin (e-value < 10\^-8\).

The pLH201.1 encodes for all the apparatus required for conjugation with 16 Tra proteins, located at two different regions of the plasmid (region 1: 69 675 – 74 152; region 2: 182 783 – 205 787) and organized in at least three operons (Fig. 3). Using the NCBI plasmid and the ICEberg ICE databases, we conducted a search at the amino acid level, keeping only sequences that matched at least one Tra protein from each of the three pLH201.1 conjugative operons.

These 16 pLH201.1 Tra proteins shared best amino acid identity (AAI) from 20% to 58% and organization with some IncA/C plasmids and SXT/R391-related.
ICE conjugative apparatus (Fricke et al. 2009; Carraro et al. 2014) (Table S2, Supporting information). On the 52 plasmid hits (120 kb to 582 kb) and 16 ICE hits, not a single Tra homologue displayed an AAI superior to 60%. Twelve matched ICEs belonged to the SXT/R391 family and the remaining four were from the SPI-7 family (one sequence) or unclassified (three sequences). Most of the plasmid hits (n = 23) were annotated as IncA/C plasmids, while the remainder were spread in groups H, P, T, J, F, multireplicon and unknown plasmid incompatibility groups. All the plasmids conferred multidrug resistance and were found in five bacterial families (Pseudomonadaceae, Enterobacteriaceae, Aeromonadaceae, Burkholderiaceae and Vibrionaceae) with contrasted ecological niches and geographical origins.
As the protein sequences of conjugative relaxases are useful for plasmid classification, we assessed the AAI of the pLH201.1 relaxase (TraLpLH201.1) to that of each of the six MOB groups defined in the literature (Carcillan-Barea et al. 2009). TraLpLH201.1 solely matched the MOB1 type (e-value < 10^-4) with its typical amino acid signature ([(H)Q]-x2-PASE-x-HDH1-x3-3G-x3-H-x1. and (LV)x-HD-(AVLD)-GK). MOB1 relaxases are scarce and have only been reported in large plasmids (>60 kb) (Smillie et al. 2010) and found in the incompatibility groups IncH, IncJ, IncT, IncP7 and IncA/C (Carcillan-Barea et al. 2009). Of all known MOB1 clades and subclades, the TraLpLH201.1 appeared to be most closely related to the MOB1H21 subclade of MOB1H12 (Alvarado et al. 2012). TraLpLH201.1 also displayed a conserved domain of the PFL_4751 family of ICE relaxases (e-value = 2.20 e^-21), required for transfer of the SXT and R391 ICE types (Daccord et al. 2010). Similarly, the conjugative coupling factor TraD was homologous to the SXT-TraD domain (e-value = 0) found in conjugative transposon-like mobile genetic elements (Beaber et al. 2002) and various groups of plasmids including IncA/C (Fernandez-Alarcon et al. 2011). TraDpLH201.1 also shared good AAI with the TrwB coupling factor (e-value of 2.35 e^-15) from IncW conjugative plasmids (Gomis-Ruth et al. 2002).

Globally, the content and organization of pLH201.1 conjugative apparatus clearly shared similarities with that of IncA/C plasmids and SXT/R391 ICE. However, the low AAI levels indicate that it may be a new system that has not yet been described.

Cop genes are part of a Tn3-like transposon

In IncA/C plasmids, genes associated with adaptive traits are often found as part of complex transposons (Harmer & Hall 2014) and display a higher GC content (Zhang et al. 2014). Globally, pLH201.1 had a GC content of 59.2%, lower than the 64.8% of the chromosome, but displayed a higher local GC content (63.3%) in the 108- to 152-kb region. GC-skew profiles presented several variations. These lines of evidence suggest that the plasmid shows a mosaic structure. On pLH201.1, this region (hereafter referred to as TnpLH201.1, located at 108 034–151 931 bp) was surrounded by two inverted repeats of 34 bp. It contained genes that are syntenic and similar to genes from the plasmid-encoded Xanthomonas TnXo19, a Tn3-like transposon (Niu et al. 2015). This includes a transposon-related cointegrate protein tnpT, a cointegrate protein tnpS, a transposase tnpA, a DNA recombination protein and a DNA helicase, that all display nucleotide identity (NI) between 70% and 90% with their TnXo19 homologues, on 94%, 70%, 26%, 100% and 100% of TnPLH201.1’s gene length, respectively. tnpT, tnpS and tnpA were all shown to be involved in the transposition of some Tn3-family transposons (Tsuda & Iino 1988; Yano et al. 2013). Inside TnpLH201.1, we found an additional copy of the 34-bp repeat, which formed two direct copies separated by 43 897 bp at one extremity of TnpLH201.1 and one inverted copy at the other extremity. This pattern, typical of composite transposons, has also been reported on a Tn3-like transposon from Pseudomonas putida (Lauf et al. 1998).

Gene content of the Tn3-like transposons

The pLH201.1 contained two clusters of heavy metal resistance genes that were included in TnpLH201.1. Apart from LM199, all the strains displayed a TnpLH201.1 homologue with a globally conserved and syntenic gene content, however sometimes showing rearrangements (Figs 2 and S3, Supporting information). The first cluster was delineated by two 34-bp direct repeats and encompassed several genes, including the previously reported copLABMGCDF genes involved in copper resistance in Argentinian X. citri pv. citri populations (Behlau et al. 2011) and cusAB/smmD present in S. maltophilia (Crossman et al. 2008). The copLABMGCDF region was identical (100%) between the five sequenced X. citri pv. citri strains that possessed this system and with the 10 328-bp region, which was described in X. citri pv. citri strain A44 (i.e. LM180 in the present study) and known to be functionally involved in Cu⁶ (Behlau et al. 2011) (Figs S3 and S4, Supporting information). Interestingly, the Cu⁶ X. citri pv. citri strain from Argentina LM199 failed to produce PCR amplicons using primer pairs specific to copL, copA and copB of the copLAB system. Its genome sequence displayed a plasmid backbone extremely similar to that of pLH201.1 but with a distinct copper transposon region (Fig. 2), hereafter called TnpLM199. Indeed, the annotation of TnpLM199 revealed the presence of the alternate copper-resistance system copABCD. TnpLM199 however displayed a transposition apparatus (tnpA, tnpT and TnpS) similar to that of TnpLH201.1 (NI above 90% on more than 80% of each gene’s length) and comprised almost identical 34-bp inverted repeats at its extremities. The nucleotide sequences of copA, copB, copC and copD from LM199 respectively showed a NI of 97, 98, 98 and 98% with those of the known chromosomal system of X. arboricola pv. juglandis Xaj417 isolated from walnut (Pereira et al. 2015). Copper-resistance gene nomenclature is quite ambiguous: whereas copA and copB from copABCD respectively share 98 and 63% NI (both on 74% of the gene length) with their copLAB homologues, CopCcopABCD and CopDcopABCD amino acid sequences only display low
AAI levels with their copLAB counterparts (34% on 97% of gene’s length) (see Fig. S4, Supporting information). PCR primers targeting the four genes were designed (Table S3, Supporting information). The seven Cu\textsuperscript{R} X. citri pv. citri strains from Argentina that were copLAB negative by PCR all produced amplicons of the expected size for copABC. The system found in pLM199, as in X. arboricola pv. juglandis, did not encode for CopRS, unlike E. coli and P. syringae. However, a transcriptional regulator that belongs to the MerR family was found close to the copABCD cluster. This regulator family has been shown to respond to environmental stimuli, including heavy metals. It also controls the expression of copA in E. coli (Stoyanov et al. 2001; Brown et al. 2003).

In TnpLH201.1, genes homologous to cusAB/smmD, a heavy metal efflux resistance-nodulation-division (HME-RND) (<95% AAI) of S. maltophilia, were identified. These sequences corresponded to that of known copper/silver efflux pumps. On all strains from Réunion and Martinique (five X. citri pv. citri, one X. gardneri and one Stenotrophomonas sp.) and in one Argentinian strain (LM180), this cluster was surrounded by two almost perfect 907-bp-long direct repeats. We only found one copy on other Cu\textsuperscript{R}-bearing DNA molecules (plasmids for all strains apart from Stenotrophomonas sp. LM091) for all the other strains and none in LH3. Interestingly, on TnpLH201.1, the two copies of 907 bp comprised an Ile tRNA (UAU anticodon), which is absent on the chromosome of X. citri pv. citri LH201. Codon usage revealed a bias in the use of the ATA codon between chromosomal genes (1.4% of the Ile-encoding codons) and plasmid-borne genes (6.5%) (data not shown).

The second cluster contained genes also involved in heavy metal resistance: czcA, czcB, czcC and czcD (which encode a cobalt/zinc/cadmium efflux system, and three genes involved in arsenic resistance, an arsenate reductase (arsC), a NADPH-dependent FMN reductase (arsH) and an arsenite/antimonite transporter (arsB). Again, czc and ars genes were found to be highly (C-95%) and moderately (30–67%) identical, respectively, to those of S. maltophilia. The latter genes were located downstream of the arsR transcriptional regulator, which is induced by arsenite and antimonite (Wang et al. 2004). This cluster also encoded a copper-dependent transcription regulator hmrR from the MerR family, which is highly identical (97%) to that of S. maltophilia.

The region between the two clusters was surrounded by 66-bp-long direct repeats and encoded a metal chelator protein [also described in TnXo19 from X. oryzae pv. oryzae, (Niu et al. 2015)], a heavy metal efflux protein, a copper-sensing transcriptional repressor (which binds to a gene promoter and to copper with a higher affinity to copper) and a metal binding exoribonuclease (which might help degrade the transcriptional repressor in the presence of copper).

Slight variations were observed between the TnpLH201.1 homologues. For example, pLLO74-4 displayed a 6 945-bp insertion at position 154 316. This insertion contained 40-bp inverted repeats at both ends. We found an identical transposon (100% NI) in a X. citri pv. citri strain jx-6 plasmid pXAC33, which encoded a transposase, a resolvase and the twitching motility protein PilT. In addition, pLMG930 displayed an inserted gene (99% NI with S. maltophilia iron permease), a deletion of the ars cluster and several gene rearrangements in the region between the copLAB and the czcABCD clusters (Fig. S3, Supporting information).

Therefore, TnpLH201.1 appeared to be a hotspot of insertions, deletions and rearrangements, consistent with previous data on Xanthomonas Tn3-like transposons (Ferreira et al. 2015; Niu et al. 2015). Two different Tn3-like transposons encoded for copper resistance. Other genes putatively involved in heavy metal resistance were conserved in the sequenced plasmids, except for a single strain (X. euvesicatoria LMG930).

**TnpLH201.1 is found in various genomic environments**

Homologues of the transposon TnpLH201.1 were found in diverse genomic environments within the other sequenced strains (see Figs 2, S5 and Table S4, Supporting information for values of nucleotide divergence between the blocks of nucleotide identity defined in Fig. 2). First, we found that TnpLH201.1 was integrated in plasmids that were highly homologous and syntenic to the pLH201.1. These conserved plasmids were present in X. citri pv. citri strains LM180, LH276, LJ207-7 and LL74-4 (from the three regions studied) and from other Xanthomonas species pathogenic to solanaceous species: X. gardneri JS749-3 (Réunion) and X. vesicatoria LM159 (Argentina). Then, we found that highly homologous copies of the transposon were also integrated in rearranged pLH201.1 homologues present in strains of other Xanthomonas species pathogenic to solanaceous species: X. ‘perforans’ LH3 (synonym X. euvesicatoria; Mauritius), X. gardneri ICMP7383 (New Zealand) and X. vesicatoria LMG911 (New Zealand) (Table 1). We also observed a highly similar transposon homologue that was integrated in a markedly different plasmid environment (X. euvesicatoria LMG930, USA). The conjugative apparatus of pLMG930 displayed homology to that of pBVI04 from Burkholderia vietnensis G4, an ecologically versatile rice root-associated nitrogen-fixing betaproteobacterium (Chiarini et al. 2006). Indeed, genes of pBVI04 involved in conjugation are located in three...
separate genomic regions, each of which are very well conserved in pLMG930, sharing 93%, 90% and 92% NI, on 7,068, 5,037 and 7,792 bp, respectively. Lastly, TnpLH201.1 was integrated in the chromosome of the citrus-associated strain of Stenotrophomonas sp. LM091 (Réunion) with a NI of 97.7% on 43,276 bp.

Networks of gene sharing

As the clues indicated a mosaic structure for the Cu\textsuperscript{R} plasmid, we searched for pLH201.1 homologues in the public NCBI databases NR and WGS. Networks of gene sharing (Fig. 4) revealed that pLH201.1 homologues present in X. citri pv. citri could only be identified from X. gardneri and X. vesicatoria, consistent with data produced in this study. In contrast, genes homologous to TnpLH201.1 were detected from 14 species included in five genera (Xanthomonas, Stenotrophomonas, Pseudoxanthomonas, Pelomonas and Pseudomonas). Globally, Cu\textsuperscript{R} gene homologues were found further apart in the taxonomy than plasmid backbone homologues.

Thirty percent of the genomes represented on the network only shared one or two genes with pLH201.1, 86% of which (26% of the total) only matched known insertion sequences.

After clustering 180 NCBI sequences sharing more than 10 genes with pLH201.1 (with 70% NI over 70% of gene length), we obtained 62 clusters (Fig. S6, Supporting information). Three patterns of homology emerged. The first pattern (two clusters) comprised homologues to the complete pLH201.1. The second pattern was found with clusters whose backbone region displayed NI with the entire pLH201.1 backbone but differed from TnpLH201.1 (three clusters). Within this group, all clusters displayed a highly similar backbone region and a conserved gene content, which suggests that they are closely related. However, their accessory gene regions were different with the lack of copLABMGCDF (X. 'perforans', contig Accession no. JZUY01000051, that can be circularized with a 21-bp perfect-match overlap), or the incorporation of other gene clusters coding traits such as cobalt efflux or ion transport (X. 'perforans', Accession no. JZVH01000033). In addition, this cluster comprised a contig of a X. euvesicatoria pv. allii strain from Réunion matching the whole backbone region of pLH201.1 (Accession no. JOJQ01000000) that we were unable to circularize. Finally, the third pattern consisted of multiple clusters that only shared NI with TnpLH201.1 or parts of it. However, the distance between them was sufficient to form distinct clusters.

![Network of all NCBI genomes sharing homologous genes with pLH201.1 (nucleotide identity > 95% on 95% of pLH201.1 gene length). Edges appear closer if the number of genes they share is higher and diameter of the nodes is proportional to the number of genes shared with pLH201.1. (A) Nodes are coloured depending on the taxonomy of the organisms; (B) nodes are coloured in green if at least one gene in the sequence is homologous with TnpLH201.1, otherwise they are red.](image-url)
This confirmed that highly similar TnpLH201.1 homologues insert in diverse genomic environments.

Discussion

In response to the use of copper-based antimicrobial compounds to control plant bacterial pathogens, copper-resistant strains have emerged repeatedly in different parts of the world. Determinants of copper resistance have often been reported to be plasmidborne, as in the case of Xanthomonas citri pv. citri, the causal agent of Asiatic citrus canker (Behlau et al. 2012). Until now, the understanding of the ecology of these resistance determinants and the evolution of the associated plasmids has been limited by the lack of genomic data. In the present study, we provide a comparative genomic analysis of plasmids associated with CuR in several Xanthomonas species, including X. citri pv. citri.

Multiple acquisitions of copper resistance in X. citri pv. citri

An unprecedented collection of X. citri pv. citri CuR strains from all the geographical areas where its emergence has been reported (Argentina, Réunion and Martinique) was genotyped using two complementary sets of markers (minisatellites and microsatellites). CuR strains were genetically related, so much so that they were assigned to a single lineage (DAPC1) and formed a single clonal complex based on minisatellite data. This lineage corresponded to the wide host range pathotype A, known to be responsible for the worldwide emergence of Asiatic citrus canker in the 20th century from Asia (i.e. its native origin) (Pruvost et al. 2014). Nevertheless, within this clonal complex, microsatellite analysis revealed a substantial differentiation into two genetic clusters likely with no epidemiological link. One cluster included Argentinian strains, while the other encompassed all strains from Réunion and Martinique. Genomic data confirmed the differentiation between French and Argentinian X. citri pv. citri strains (with intergroup genetic divergence tenfold higher than intragroup one, see Table S4, Supporting information for details). In Réunion, the genetic relatedness between CuR contemporary outbreak strains and CuS ‘historical’ strains isolated two to four decades ago suggested that the establishment of CuS strains on the island was unlikely to be the result of the recent introduction of genetically distinct strains. Instead, our data support the hypothesis that strains, which were genetically similar to the CuS populations characterized in the early years of the disease in Réunion, acquired a plH201.1-like plasmid from a presently unknown source (i.e. Asiatic canker was reported for the first time in Réunion in 1968 – Brun (1971)). On the contrary, the weak genetic divergence between strains from Réunion and Martinique shows a possible epidemiological link between them (Fig. S5 and Table S4, Supporting information).

Contrasting with their epidemiological structure and genetic divergence, the Argentinian LM180 strain and the French ones displayed extremely similar plasmids (Fig. S5 and Table S4, Supporting information), suggesting independent copper-resistance acquisition by these two groups of strains, and showing the mobility of plasmid-encoded adaptive traits at very large geographical scales.

This mobility and the scenario of independent acquisition was further supported by the fact that in Réunion, a CuR X. gardneri strain (a pathogen of tomato and pepper) was found to carry a copy of pLH201.1 (average of 0.05 different nucleotides per kb) and a X. euvesicatoria pv. allii strain contained a contig in its genome (Gagnevin et al. 2014), which corresponded to the plH201.1 backbone. The latter could not be circularized and, therefore, we were unable to confirm that it was a plasmid which lacked TnpLH201.1, despite the fact that the strain was PCR negative for copLAB and had a CuS phenotype. Finally, in Argentina, some CuR strains varying in copLAB PCR amplification were not genetically differentiated based on microsatellite data. This suggests the independent acquisition of two distinct copper-resistance systems within Argentinian lineages.

Several putative copper-resistance systems in X. citri pv. citri

For the Argentinian X. citri pv. citri A44 (LM180), CuR genes primarily shown to be experimentally functional comprise a transcriptional regulator (copL) and two copper-binding proteins (copAB) (Behlau et al. 2011). PCR tests provided evidence that most known CuR X. citri pv. citri strains possess this copLAB system. Using transposon mutagenesis, Behlau et al. (2011) demonstrated that the disruption of the copLAB genes was sufficient to lower the copper-resistance level to that of CuS strains. However, when inserted in a CuS strain of a closely related species, X. perforans, the copLAB system alone did not confer the level of copper resistance of wild-type CuR strains. The authors suggested that this was due to the fact that their recipient strain had a different genetic background (Behlau et al. 2011). In 12 of 13 CuR sequenced strains, we identified an additional gene cluster that could be involved in copper resistance and may explain the partial phenotype restoration observed by Behlau et al. (2011). The HME-RND system, which is also present in S. maltophilia (Crossman et al. 2008) and forms a channel through the periplasm, is composed of an inner membrane pump (here CusA),
a periplasmic protein (CusB) and an outer membrane protein (SmmD) (Routh et al. 2011). The cusAB/smmD is not widespread among xanthomonads. It is not present in the draft genome of the strain complemented with copLAB by Behlau et al. for their functional analysis (data not shown) and, to date, has solely been detected from the draft genome of the Cu\textsuperscript{R} X. vesicatoria ATCC 35937 and mentioned in the NR database in an unknown MOBH\textsubscript{12} plasmid from a marine environment, very broad host range and are found in very diverse plasmids and SXT/R391 ICE. IncA/C plasmids have a tetracystein apparatus clearly have similarities with IncA/C plasmid family. The content and organization of its conjugative apparatus Tn3-like transposon (i.e. containing copABCD) in a genetically related backbone. Both the copLAB and the copABCD systems were reported from distinct strains of the walnut pathogen X. arboricola pv. juglandis on plasmid and chromosome, respectively (Lee et al. 1994; Behlau et al. 2013). Hence, we were able to establish that at least two distinct cop systems were associated with copper resistance in Argentinian X. citri pv. citri. Why polymorphism exists in copper-resistance systems is intriguing. Currently, we lack the necessary elements to test whether it is adaptive, that is occurs in response to environmental variations in copper concentration or fortuitous and driven by bioavailability.

The importance of HGT for the adaptation of genetically monomorphic bacteria

For all the studied X. citri pv. citri strains, copper-resistance systems were found on closely related plasmids of approximately 230 kb in size. Extensive annotation of the Réunion X. citri pv. citri plasmid pLH201.1 revealed that it bears all the genetic elements required for conjugation, confirming \textit{in vitro} tests (data not shown) and previous data on strain A44 (LM189) from Argentina (Behlau et al. 2012). The pLH201.1 showed no strong homology to plasmids described previously. However, its relaxase is such that it belongs to the MOB\textsubscript{112} plasmid family. The content and organization of its conjugative apparatus clearly have similarities with IncA/C plasmids and SXT/R391 ICE. IncA/C plasmids have a very broad host range and are found in very diverse environments and geographical areas. Recently, an unknown MOB\textsubscript{112} plasmid from a marine environment, which also has similarities with IncA/C plasmids and SXT/R391, was reported (Nonaka et al. 2012). This suggests that the MOB\textsubscript{112} plasmid family could be wider than previously thought.

Plasmids can confer a broad range of adaptive traits to their host, such as antibiotic resistance (Ochman et al. 2000), heavy metal resistance (Hobman & Crossman 2015), UV tolerance, hormone production, pathogenicity determinants and toxin production (Vivian et al. 2001; Sundin 2007). These adaptations can lead to the colonization of new ecological niches. They may even be responsible for major evolutionary events, such as the emergence of new pathogenic populations. For example, different allelic forms of the pPATH plasmid have transformed strains of the commensal bacterial species Pantoea agglomerans into gall-forming pathogens of gypsophila and beet (Weinthal et al. 2007). Our results strongly support the acquisition of a new adaptive phenotype through plasmid incorporation by different X. citri pv. citri populations. However, the mechanisms of xanthomonad adaptation through HGT might not be restricted to plasmid acquisition. Indeed, chromosomally encoded resistance was reported on \textit{Stenotrophomonas} and xanthomonads causing bacterial spot of tomato and pepper or bacterial blight of walnut. In addition, several genomic islands, including genes of plasmid origin, were detected on the chromosome of X. citri pv. citri (Gordon et al. 2015).

Importance of plasmid-borne mobile genomic elements

In the present study, we provide evidence that copper-resistance gene clusters on pLH201.1 were encoded on Tn3-like transposon (referred to here as TnpLH201.1). Tn3-like transposons have been reported for other plasmids in \textit{Xanthomonas} (Ferreira et al. 2015; Niu et al. 2015), as well as other genera (Lauf et al. 1998). IncA/C plasmids often carry a complex transposon-based cluster of resistance genes involved in the spread of multidrug resistance between bacteria (Harmer & Hall 2014).

This feature could mitigate the apparently limited host range of pLH201.1, by providing a second layer of mobility. Indeed, three distinct species (\textit{X. euvesicatoria}, \textit{S. maltophilia} and \textit{Stenotrophomonas} sp.) harbour a transposon almost identical to the one hosting the copLAB gene system in \textit{X. citri pv. citri} (>99% NI) in a genomic environment that is markedly different from that of pLH201.1 (i.e. a 179-kb plasmid for \textit{X. euvesicatoria} and the chromosome for the two other species). This supports the hypothesis that the transposon is a source of mobility for the Cu\textsuperscript{R} gene cluster. Moreover, the copABCD system found on LM199 has 98% NI with that encoded on the chromosome of \textit{X. arboricola pv. juglandis}, while encoded on a pLH201.1-related plasmid (NI of 90% on 85% of pLM199 length, see Fig. 2 and Table S4, Supporting information). The pLM199 comprised all the 16 genes from the pLH201.1 conjugative
gene set. In this regard, the transposon TnpLH201.1 can be considered as an autonomous vehicle. Indeed, it encodes for CuR proteins, transcriptional regulators, a transposition apparatus and a single tRNA. The latter, which is required for the transcription of the genes encoded on the transposon, is absent on the chromosome.

To lower the fitness cost of plasmid carriage, chromosomal genome and plasmids co-evolve (Harrison & Brockhurst 2012). This process could limit the spread of entire alien plasmids and, instead, favour the incorporation of the transposon into plasmids that are already present within a restricted taxonomic group and therefore already adapted to their host.

**Barriers to HGT and importance of reservoir bacteria**

Optima of genome functioning leave strong imprints, such as GC content and codon usage. These differences tend to limit exchange of DNA between distantly related bacteria (Popa & Dagan 2011). Indeed, within networks of shared DNA among bacterial genomes (with 95% NI), *Xanthomonas* species tend to form an isolated cluster. Only plasmids with lower NI within the species connect *Xanthomonas* to some distantly related bacterial genera (Halary et al. 2010). Our network approach with pLH201.1 yielded similar results. We identified complete or nearly complete pLH201.1 homologues, as well as genes involved in CuR or conjugation, primarily in the Xanthomonadaceae family.

The present study has provided evidence that distantly related *Xanthomonas* species (e.g. the tomato pathogen *X. gardneri* and the citrus pathogen *X. citri* pv. *citri*) carry the same plasmid. *Xanthomonas* is a bacterial genus largely composed of plant pathogenic bacteria with a high degree of host specialization (Leyns et al. 1984). As *Xanthomonas* lineages that have a different host range colonize distinct ecological niches, they would not be expected to share DNA directly through conjugation because it requires cell-to-cell contact. However, in some agricultural contexts (intercropping, for example), the physical proximity of plant species contaminated with distinct bacterial pathogens could facilitate contact and HGT. In addition, extreme weather events that have been reported to spread bacterial cells over long distances (Irey et al. 2006) may contribute to the mixing of xanthomonad populations. A key factor for gene transfer between populations probably lies in the ability of *Xanthomonas* to survive transiently on plant surfaces, in natural plant openings or even on nonhost plant species (Robinson & Callow 1986). In fact, xanthomonads were reported to form mixed-biofilm structures on plant surfaces (Jacques et al. 2005; Cubero et al. 2011), which have been recognized as highly favourable to HGT within the phyllosphere (Van Elsas et al. 2003).

Indirect transfer of copper resistance between xanthomonads may occur. Different reservoirs of bacteria resistant to antimicrobials can be found in different environmental compartments that interact and share interfaces (Nesme et al. 2014). Indeed, a resistome to environmental or industrial copper does exist and combines different genes associated with copper resistance (He et al. 2010). By tracking the dispersal and availability of this type of resistance in the natural environment or agro-ecosystem and linking it to other settings, we should be able to understand and predict how the ecosystem functions (Vieites et al. 2009).

Following sporadic reporting over several decades (Vauterin et al. 1996), commensal xanthomonads are now being more carefully characterized in terms of taxonomy or taxonomic placement (Triplet et al. 2015) or pathogenicity gene repertoires and mobile genetic elements (Cesbron et al. 2015). The extent to which these commensal *Xanthomonas* strains or commensal bacteria act as reservoirs or hubs for adaptive genes is still unknown. In the context of increased HGT frequencies between phylogenetically related species, the significance of *Stenotrophomonas* (and other genera in the Xanthomonadaceae family) as a major source of adaptive genes for xanthomonads in agricultural ecosystems has largely been underrated. At least two commensal *Stenotrophomonas* species displayed a highly identical copy of TnpLH201.1. Despite its relative individual insignificance as a pathogen, *S. maltophilia* is of major relevance in terms of plant, animal and human health because it constitutes a gene reservoir that is available for gene transfer within the community. Indeed, the panoply of resistance genes that it harbours could provide a source of antibacterial resistance determinants that are transferable to bacterial pathogens, such as the copper-resistance system presented here or other types of resistance relating to human health reported previously (Crossman et al. 2008). Our study highlights the importance of conducting further research on entire microbial communities in order to improve our understanding of the emergence of pathogenic bacteria.

**Acknowledgements**

We would like to thank B. Barrès, B. Doublet and P. Roumagenc for their helpful contribution to discussions, as well as all those who provided help during field prospection. The European Regional Development Fund (ERDF) and European Agricultural Fund for Rural Development (EAFRD), Conseil Départemental de la Réunion, Région Réunion, État français, the French Agropolis Foundation (Labex Agro – Montpellier,
References

Achtman M (2008) Evolution, population structure, and phylo-


Achtman M (2012) Insights from genomic comparisons of

Alvarado A, Carcillan-Barcia MP, de la Cruz F (2012) A degener-

ate primer MOB typing (DPMT) method to classify gamma-proteobacterial plasmids in clinical and environmental settings. PLoS One, 7, e40438.

Beaber JW, Hochhut B, Waldor MK (2002) Genomic and func-
tional analyses of SXT, an integrating antibiotic resistance gene transfer element derived from Vibrio cholerae. Journal of Bacteriology, 184, 4259–4269.


Behlau F, Canteros BI, Jones JB, Graham JH (2012) Copper resistance genes from different xanthomonads and citrus epi-


Bryant J, Chewapreecha C, Bentley SD (2012) Developing insights into the mechanisms of evolution of bacterial pathog-

ens from whole-genome sequences. Future Microbiology, 7, 1283–1296.


Constantin EC, Cleenwerck I, Maes M et al. (2016) Genetic characterization of strains named as Xanthomonas axonopodis pv. dieffenbachiae leads to a taxonomic revision of the X. a xo


nopolis species complex. Plant Pathology, 65, 792–806.


monas maltophilia reveals an organism heavily shielded by drug resistance determinants. Genome Biology, 9, R74.


Daccord A, Caccarelli D, Burrus V (2010) Integrating conjuga-
ye elements of the SXT/K91 family trigger the excision and drive the mobilization of a new class of Vibrio genomic islands. Molecular Microbiology, 78, 576–588.


Davenport KW, Daligault HE, Minogue TD et al. (2014) Complete genome sequence of Stenotrophomonas maltophilia type strain 810–2 (ATCC 13657). Genome Announcements, 2, e00974–e00914.


Dubey GP, Ben-Yehuda S (2011) Intercellular nanotubes medi-

ate bacterial communication. Cell, 144, 590–600.


Fernandez-Alarcon C, Singer RS, Johnson TJ (2011) Compara-
tive genomics of multidrug resistance-encoding IncA/C plas-

mids from commensal and pathogenic Escherichia coli from multiple animal sources. PLoS One, 6, e23415.


© 2017 John Wiley & Sons Ltd


Weinthal DM, Barash I, Panijel M et al. (2007) Distribution and replication of the pathogenicity plasmid pPATH in diverse...


**Data accessibility**

The MLVA-31 and MLVA-14 data generated in this study are available in the *Xanthomonas citri* genotyping database (http://www.biopred.net/MLVA/) and at https://agritrop.cirad.fr/bio82053/, respectively. Sequences produced in this study are deposited in the GenBank database (Table 1).


**Supporting information**

Additional supporting information may be found in the online version of this article.

Fig. S1 Categorical minimum spanning tree of DAPC1 strains from Argentina, Martinique and Réunion, which differ in their susceptibility to copper (268 strains – 34 haplotypes), based on minisatellite data.

Fig. S2 Categorical minimum spanning tree of copper-resistant *Xanthomonas citri* pv. *citri* strains from Réunion based on microsatellite data.

Fig. S3 Alignments between TnpLH201.1 and sequences carrying TnpLH201.1 homologues from strains sequenced in this study and *S. maltophilia* K279a.

Fig. S4 Distance tree of the copper-resistance region of all the sequenced strains as well as some sequences coding for known copper-resistance systems extracted from public database along with and graphical representations of the copper-resistance genes organisation.

Fig. S5 Heatmap representation of the genetic divergence between six sequenced strains.

Fig. S6 Clusters of sequences with homologues to pLH201.1 genes.

Table S1 Size (in bp) of the contigs obtained for each of the 13 sequenced strains

Table S2 plasmids and ICE having similar conjugative apparatus with pLH201.1.

Table S3 Primers and PCR conditions used for amplification of the copABCD system of *Xanthomonas citri* pv. *citri*.

Table S4 Genetic divergence over the homologous sequence tracks defined in Fig. 2.

© 2017 John Wiley & Sons Ltd