



HAL
open science

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

Damien Richard Richard, Virginie Ravigné, Adrien Rieux, Benoit Facon,
Claudine Boyer, Karine Boyer, Pierre Grygiel, Stéphanie Javegny, M. Terville,
B. I. Canteros, et al.

► **To cite this version:**

Damien Richard Richard, Virginie Ravigné, Adrien Rieux, Benoit Facon, Claudine Boyer, et al..
Adaptation of genetically monomorphic bacteria: evolution of copper resistance through multiple
horizontal gene transfers of complex and versatile mobile genetic elements. *Molecular Ecology*, 2017,
26 (7), pp.2131-2149. 10.1111/mec.14007 . hal-01607881

HAL Id: hal-01607881

<https://hal.science/hal-01607881>

Submitted on 7 Jun 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

MR. DAMIEN RICHARD (Orcid ID : 0000-0001-9179-7315)

Received Date : 05-Jun-2016

Revised Date : 28-Nov-2016

Accepted Date : 08-Dec-2016

Article type : Special Issue

This mss is for "Microbial Local Adaptation" special issue

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

D. Richard^{1,2,3,†}, V. Ravigné¹, A. Rieux¹, B. Facon^{4,5}, C. Boyer¹, K. Boyer¹, P. Grygiel¹, S. Javegny¹, M. Terville¹, B. I. Canteros⁶, I. Robène¹, C. Vernière⁷, A. Chabirand², O. Pruvost¹, P. Lefeuvre^{1,†}

Affiliations:

¹ Cirad, UMR PVBMT, F-97410 St Pierre, Réunion, France

² ANSES, Plant Health Laboratory, F-97410 St Pierre, Réunion, France

³ Université de la Réunion, UMR PVBMT, F-97490 St Denis, Réunion, France

⁴ Inra, UMR PVBMT, F-97410 St Pierre, Réunion, France

⁵ Inra, UMR CBGP, F-34090 Montpellier, France

⁶ INTA, Estación Experimental Agropecuaria Bella Vista, Argentina

⁷ Cirad, UMR BGPI, F-34398 Montpellier, France

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/mec.14007

This article is protected by copyright. All rights reserved.

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

†Corresponding authors: damien.richard@cirad.fr; pierre.lefeuvre@cirad.fr

Address: Pôle de Protection des Plantes, UMR PVBMT, Station Ligne Paradis, 7 chemin de l'Irat, 97410 Saint Pierre, Réunion, France

Fax: +262 262 499 293

Short title: Copper resistance in monomorphic bacteria

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 to mobile elements and their plasmid environments was primarily restricted to the Xanthomonadaceae family.

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 (Hurtle *et al.* 2003; Telenti *et al.* 1993). However, in many cases, the genes responsible for adaptation were actually horizontally acquired from the bacterial mobile gene pool (Behlau *et al.* 2013; Galimand *et al.* 1997; Nonaka *et al.* 2012).

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 to the number of core genes of a species has led to the development of the concept of an open pan-genome (Chaudhuri *et al.* 2010; Garcia-Vallve *et al.* 2000). 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 (Haines *et al.* 2007; Trefault *et al.* 2004) 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; Weinthal *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, i.e. 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 had 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 (Cu^{R}) 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 (ban of fresh fruit export because of the quarantine status of the pathogen) (Gottwald *et al.* 2007; Graham *et al.* 2004). 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 (Leduc *et al.* 2015; Pruvost *et al.* 2014). In regions where *X. citri* pv. *citri* has been established for a long time, control typically involves integrated pest management strategies that partly relies 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, Richard *et al.* 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.* 2012; Behlau *et al.* 2011), 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 *X. 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 paper aims 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 *X. citri* pv. *citri* Cu^R 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 Cu^R 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 Cu^R and 20 were collected from the first disease focus in Morne Rouge.

In addition to Cu^R haplotypes (n = 115 identified among 219 strains), all known haplotypes corresponding to copper-susceptible (Cu^S) *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 Cu^R and Cu^S 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 Cu^R xanthomonads were originally reported in strains causing bacterial spot of tomato and pepper and no genomic resources were available for such strains, reference Cu^R 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).

Evaluation of copper resistance

To assess copper resistance, we used the growth test on YPGA (yeast extract 7 g l⁻¹, casein peptone 7 g l⁻¹, glucose 7 g l⁻¹, agar 18 g l⁻¹, pH 7.2), supplemented with CuSO₄·5H₂O (470 mg l⁻¹), and PCR using *copL* primers. We slightly modified the PCR protocol reported previously (Behlau *et al.* 2013) by using GoTaq Flexi polymerase (Promega), a lower primer concentration (5 pmol μl⁻¹) 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 Cu^R and 45 Cu^S), Martinique (9 Cu^R) and Réunion (64 Cu^R and 103 Cu^S) were genotyped using 31 minisatellites (MLVA-31) developed for global epidemiology analyses and compared to 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 tandem 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 PHYLOVIZ v1.0 (Francisco *et al.* 2012).

The relatedness between strains was assessed at smaller spatio-temporal 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 intra-pathotype 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 R_{ST} values were computed for estimating genetic differentiation among Cu^R 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 Cu^R strains from the three resistance outbreaks, allowed us to decipher the genetic relationships between the three outbreaks. The second dataset, consisting of all Cu^R and Cu^S strains from Réunion (the only region where we have extensively sampled local populations), enabled us to examine the relationships between the Cu^R and Cu^S strains that were isolated both before and during the Cu^R *X. citri* pv. *citri* epidemics. When comparing Cu^R and Cu^S 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 were selected for sequencing from Argentina, Martinique and Réunion (Table 1). In addition, six other reference strains of Cu^R xanthomonads causing bacterial spot disease on tomato and/or pepper and a Cu^R 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-reads 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, Cu^R 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 pLH201.1 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 on-line 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 EUTILITIES).

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 Cu^R outbreak strains and placement within the pathovar diversity

Most Cu^R *X. 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 spatio-temporal evolutionary scale defined by these markers, strains from the three origins formed a single clonal complex (CC), i.e. 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 Cu^R strains from Martinique and Réunion. Most Cu^R strains from Réunion were assigned to two frequent haplotypes (#173 and 175), which also contained Cu^S strains of the same origin. Similarly, haplotype #51, the most frequent haplotype among Argentinian 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 Cu^R *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 Cu^R 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 to 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 Argentinian strains. Cu^R 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 ($R_{ST} = 0.22$; $P < 0.001$). 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 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 assignment probabilities to clusters were often low (i.e. < 0.5). We decided to retain $K = 2$ because minimal strain assignment probabilities to clusters were > 0.8 . All but two Cu^R strains were assigned to cluster A, together with Cu^S strains. In contrast, only 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, i.e. 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 (≥ 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 *X. 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 (Carraro *et al.* 2014; Fricke *et al.* 2009) (Table S2, Supporting information). On the 52 plasmid hits (120 kb to 582 kb) and 16 ICE hits, not a single Tra homologue displayed a AAI superior to 60%. Twelve matched ICEs belonged to the SXT/R391 family, 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 multi-drug 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 (TraI_{pLH201.1}) to that of each of the six MOB groups defined in the literature (Garcillan-Barcia *et al.* 2009). TraI_{pLH201.1} solely matched the MOB_H type (e-value < 10⁻⁴) with its typical amino acid signature ((HQ)-x2-PASE-x-HHH-x3-GG-x3-H-x-L and (LV)-x-HD-(AVLI)-GK). MOB_H 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 (Garcillan-Barcia *et al.* 2009). Of all known MOB_H clades and subclades, the TraI_{pLH201.1} appeared to be most closely related to the MOB_{H121} sub-clade of MOB_{H12} (Alvarado *et al.* 2012). TraI_{pLH201.1} 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). TraD_{pLH201.1} also shared good AAI with the TrwB coupling factor (e-value of 2.35 e⁻¹⁹) 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 - 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 homologs, 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 homolog with a globally conserved and syntenic gene content, however sometimes showing rearrangements (Fig. 2 and Fig. 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 (out of six) 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 Cu^R (Behlau *et al.* 2011) (Fig. S3 and Fig. S4, Supporting information). Interestingly, the Cu^R *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* homologs, CopC_{copABCD} and CopD_{copABCD} 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^R *X. citri* pv. *citri* strains from Argentina that were *copLAB* negative by PCR all produced amplicons of the expected size for *copABCD*. 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 reported to respond to environmental stimuli, including heavy metals. It also controls the expression of *copA* in *E. coli* (Brown *et al.* 2003; Stoyanov *et al.* 2001).

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^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 putatively 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 (> 95%) and moderately (30 to 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, pLL074-4 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 Fig. 2, Fig 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 solaneous 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 solaneous 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 pBVIE04 from *Burkholderia vietnamensis* G4, an ecologically versatile rice root-associated nitrogen-fixing betaproteobacterium (Chiarini *et al.* 2006). Indeed, genes of pBVIE04 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^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 paper. In contrast, genes homologous to TnpLH201.1 were detected from 14 species included in five genera (*Xanthomonas*, *Stenotrophomonas*, *Pseudoxanthomonas*, *Pelomonas* and *Pseudomonas*). Globally, Cu^R gene homologues were found further apart in the taxonomy than plasmid backbone homologues.

30% 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 JZUY01000051, that can be circularized with a 21bp perfect-match overlap), or the incorporation of other gene clusters coding traits such as cobalt efflux or ion transport (*X. 'perforans'*, accession 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 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. 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 plasmid-borne, as in the case of *X. 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 Cu^R 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* Cu^R 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). Cu^R 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 Cu^R contemporary outbreak strains and Cu^S 'historical' strains isolated two to four decades ago suggested that the establishment of Cu^R 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 Cu^S 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 show 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 Cu^R *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* contig (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 Cu^S phenotype. In Argentina, some Cu^R 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), Cu^R 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 Cu^R *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 Cu^S strains. However, when inserted in a Cu^S strain of a closely related species, *Xanthomonas* 'perforans', the *copLAB* system alone did not confer the level of copper resistance of wild-type Cu^R 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 out of 13 Cu^R 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^R *X. vesicatoria* ATCC 35937 and mentioned in the NR database in *S. maltophilia* and *S. acidaminiphila*. Here, we added *X. citri* pv. *citri*, *X. gardneri*, *X. euvesicatoria* and *Stenotrophomonas* sp. to the list of species with both the *copLAB* and the *cusAB/smmD* systems.

In contrast to the majority of strains that have been studied, some *X. citri* pv. *citri* strains from Argentina were Cu^R and PCR positive for *copABCD* but not *copLAB*. One of the strains (LM199 from Argentina) was sequenced and revealed that its plasmid hosts a different 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 (Behlau *et al.* 2013; Lee *et al.* 1994). 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, i.e. 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 *in vitro* tests (data not shown) and previous data on strain A44 (LM180) 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_{H12} 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_{H12} 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_{H12} 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 (Sundin 2007; Vivian *et al.* 2001). 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 *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 *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 multi-drug 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 (*X. euvesicatoria*, *S. maltophilia* and *Stenotrophomonas* sp.) harbour a transposon almost identical to the one hosting the *copLAB* gene system in *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 *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^R gene cluster. Moreover, the *copABCD* system found on LM199 has 98% NI with that encoded on the chromosome of *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 Cu^R 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* 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 Cu^R 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 non-host plant species (Robinson & Callow 1986). In fact, xanthomonads were reported to form mixed-biofilm structures on plant surfaces (Cubero *et al.* 2011; Jacques *et al.* 2005), 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 (Triplett *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. Roumagnac 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, E-SPACE project number 1504-004), ANSES and CIRAD provided financial support.

References

- Achtman M (2008) Evolution, population structure, and phylogeography of genetically monomorphic bacterial pathogens. *Annu. Rev. Microbiol.* **62**, 53-70.
- Achtman M (2012) Insights from genomic comparisons of genetically monomorphic bacterial pathogens. *Philos Trans R Soc Lond B Biol Sci* **367**, 860-867.
- Alvarado A, Garcillan-Barcia MP, de la Cruz F (2012) A degenerate 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 functional 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 epiphytic bacteria confer resistance to *Xanthomonas citri* subsp *citri*. *European Journal of Plant Pathology* **133**, 949-963.
- Behlau F, Canteros BI, Minsavage GV, Jones JB, Graham JH (2011) Molecular Characterization of Copper Resistance Genes from *Xanthomonas citri* subsp *citri* and *Xanthomonas alfalfae* subsp *citrumelonis*. *Applied and Environmental Microbiology* **77**, 4089-4096.
- Behlau F, Hong JC, Jones JB, Graham JH (2013) Evidence for acquisition of copper resistance genes from different sources in citrus-associated xanthomonads. *Phytopathology* **103**, 409-418.
- Brown NL, Barrett SR, Camakaris J, Lee BT, Rouch DA (1995) Molecular genetics and transport analysis of the copper-resistance determinant (*pco*) from *Escherichia coli* plasmid pRJ1004. *Mol Microbiol* **17**, 1153-1166.
- Brown NL, Stoyanov JV, Kidd SP, Hobman JL (2003) The MerR family of transcriptional regulators. *FEMS Microbiol Rev* **27**, 145-163.
- Brun J (1971) Le chancre bactérien des Citrus. *Fruits* **26**, 533-540.
- Bryant J, Chewapreecha C, Bentley SD (2012) Developing insights into the mechanisms of evolution of bacterial pathogens from whole-genome sequences. *Future Microbiology* **7**, 1283-1296.
- Bui Thi Ngoc L, Verniere C, Vital K, et al. (2009) Development of 14 minisatellite markers for the citrus canker bacterium, *Xanthomonas citri* pv. *citri*. *Molecular Ecology Resources* **9**, 125-127.

Cai R, Lewis J, Yan S, *et al.* (2011) The Plant Pathogen *Pseudomonas syringae* pv. tomato Is Genetically Monomorphic and under Strong Selection to Evade Tomato Immunity. *PLoS Pathog* **7**, e1002130.

Carattoli A (2009) Resistance plasmid families in Enterobacteriaceae. *Antimicrob Agents Chemother* **53**, 2227-2238.

Carraro N, Sauve M, Matteau D, *et al.* (2014) Development of pVCR94DeltaX from *Vibrio cholerae*, a prototype for studying multidrug resistant IncA/C conjugative plasmids. *Front Microbiol* **5**, 44.

Cesbron S, Briand M, Essakhi S, *et al.* (2015) Comparative genomics of pathogenic and nonpathogenic strains of *Xanthomonas arboricola* unveil molecular and evolutionary events linked to pathoadaptation. *Front. Plant Sci.* **6**, 1126.

Chaudhuri RR, Sebahia M, Hobman JL, *et al.* (2010) Complete genome sequence and comparative metabolic profiling of the prototypical Enteroaggregative *Escherichia coli* strain 042. *PLoS One* **5**, e8801.

Chiarini L, Bevivino A, Dalmastrì C, Tabacchioni S, Visca P (2006) *Burkholderia cepacia* complex species: health hazards and biotechnological potential. *Trends in Microbiology* **14**, 277-286.

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. axonopodis* species complex. *Plant Pathology* **65**, 792-806.

Cooksey DA, Azad HR, Cha JS, Lim CK (1990) Copper resistance gene homologs in pathogenic and saprophytic bacterial species from tomato. *Appl. Environ. Microbiol.* **56**, 431-435.

Crossman LC, Gould VC, Dow JM, *et al.* (2008) The complete genome, comparative and functional analysis of *Stenotrophomonas maltophilia* reveals an organism heavily shielded by drug resistance determinants. *Genome Biol* **9**, R74.

Cubero J, Gell I, Johnson EG, Redondo A, Graham JH (2011) Unstable green fluorescent protein for study of *Xanthomonas citri* subsp. *citri* survival on citrus. *Plant Pathology* **60**, 977-985.

Daccord A, Ceccarelli D, Burrus V (2010) Integrating conjugative elements of the SXT/R391 family trigger the excision and drive the mobilization of a new class of *Vibrio* genomic islands. *Mol Microbiol* **78**, 576-588.

Darling AC, Mau B, Blattner FR, Perna NT (2004) Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Res* **14**, 1394-1403.

Davenport KW, Daligault HE, Minogue TD, *et al.* (2014) Complete genome sequence of *Stenotrophomonas maltophilia* type strain 810-2 (ATCC 13637). *Genome Announc* **2**, e00974-00914.

Davies J, Davies D (2010) Origins and evolution of antibiotic resistance. *Microbiol Mol Biol Rev* **74**, 417-433.

Dubey GP, Ben-Yehuda S (2011) Intercellular nanotubes mediate bacterial communication. *Cell* **144**, 590-600.

Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evol Bioinform Online* **1**, 47-50.

- Fernandez-Alarcon C, Singer RS, Johnson TJ (2011) Comparative genomics of multidrug resistance-encoding IncA/C plasmids from commensal and pathogenic *Escherichia coli* from multiple animal sources. *Plos One* **6**, e23415.
- Ferreira RM, de Oliveira AC, Moreira LM, *et al.* (2015) A TALE of transposition: Tn3-like transposons play a major role in the spread of pathogenicity determinants of *Xanthomonas citri* and other xanthomonads. *MBio* **6**, e02505-02514.
- Francisco AP, Vaz C, Monteiro PT, *et al.* (2012) PHYLOViZ: phylogenetic inference and data visualization for sequence based typing methods. *Bmc Bioinformatics* **13**, 87.
- Fricke WF, Welch TJ, McDermott PF, *et al.* (2009) Comparative genomics of the IncA/C multidrug resistance plasmid family. *Journal of Bacteriology* **191**, 4750-4757.
- Gagnevin L, Bolot S, Gordon JL, *et al.* (2014) Draft Genome Sequence of *Xanthomonas axonopodis* pv. *allii* Strain CFBP 6369. *Genome Announc* **2**, e00727-00714.
- Galimand M, Guiyoule A, Gerbaud G, *et al.* (1997) Multidrug resistance in *Yersinia pestis* mediated by a transferable plasmid. *N Engl J Med* **337**, 677-680.
- Garcia-Vallve S, Romeu A, Palau J (2000) Horizontal gene transfer in bacterial and archaeal complete genomes. *Genome Res* **10**, 1719-1725.
- Garcillan-Barcia MP, Francia MV, de la Cruz F (2009) The diversity of conjugative relaxases and its application in plasmid classification. *FEMS Microbiol Rev* **33**, 657-687.
- Giovanardi D, Bonneau S, Gironde S, *et al.* (2016) Morphological and genotypic features of *Xanthomonas arboricola* pv. *juglandis* populations from walnut groves in Romagna region, Italy. *European Journal of Plant Pathology* **145**, 1-16.
- Gomis-Ruth FX, Moncalian G, de la Cruz F, Coll M (2002) Conjugative plasmid protein TrwB, an integral membrane type IV secretion system coupling protein. Detailed structural features and mapping of the active site cleft. *J Biol Chem* **277**, 7556-7566.
- Gordon JL, Lefeuvre P, Escalon A, *et al.* (2015) Comparative genomics of 43 strains of *Xanthomonas citri* pv. *citri* reveals the evolutionary events giving rise to pathotypes with different host ranges. *Bmc Genomics* **16**, 1098.
- Gottwald TR, Bassanezi RB, Amorim L, Bergamin-Filho A (2007) Spatial pattern analysis of citrus canker-infected plantings in sao paulo, Brazil, and augmentation of infection elicited by the asian leafminer. *Phytopathology* **97**, 674-683.
- Goudet J (2005) Hierfstat, a package for R to compute and test hierarchical F-statistics. *Molecular Ecology Notes* **5**, 184-186.
- Graham JH, Gottwald TR, Cubero J, Achor DS (2004) *Xanthomonas axonopodis* pv. *citri*: factors affecting successful eradication of citrus canker. *Mol. Plant Pathol.* **5**, 1-15.
- Grigoriev A (1998) Analyzing genomes with cumulative skew diagrams. *Nucleic Acids Research* **26**, 2286-2290.
- Guy L, Kultima JR, Andersson SG (2010) genoPlotR: comparative gene and genome visualization in R. *Bioinformatics* **26**, 2334-2335.
- Haines AS, Jones K, Batt SM, Kosheleva IA, Thomas CM (2007) Sequence of plasmid pBS228 and reconstruction of the IncCP-1 alpha phylogeny. *Plasmid* **58**, 76-83.

Halary S, Leigh JW, Cheaib B, Lopez P, Baptiste E (2010) Network analyses structure genetic diversity in independent genetic worlds. *Proceedings of the National Academy of Sciences, USA* **107**, 127-132.

Harmer CJ, Hall RM (2014) pRMH760, a precursor of A/C(2) plasmids carrying blaCMY and blaNDM genes. *Microb Drug Resist* **20**, 416-423.

Harrison E, Brockhurst MA (2012) Plasmid-mediated horizontal gene transfer is a coevolutionary process. *Trends in Microbiology* **20**, 262-267.

He LY, Zhang YF, Ma HY, *et al.* (2010) Characterization of copper-resistant bacteria and assessment of bacterial communities in rhizosphere soils of copper-tolerant plants. *Applied Soil Ecology* **44**, 49-55.

Heinemann JA, Sprague GF (1989) Bacterial conjugative plasmids mobilize DNA transfer between bacteria and yeast. *Nature* **340**, 205-209.

Hobman JL, Crossman LC (2015) Bacterial antimicrobial metal ion resistance. *J Med Microbiol* **64**, 471-497.

Hurtle W, Lindler L, Fan W, *et al.* (2003) Detection and identification of ciprofloxacin-resistant *Yersinia pestis* by denaturing high-performance liquid chromatography. *J Clin Microbiol* **41**, 3273-3283.

Hyatt D, Chen GL, Locascio PF, *et al.* (2010) Prodigal: prokaryotic gene recognition and translation initiation site identification. *Bmc Bioinformatics* **11**, 119.

Irey M, Gottwald TR, Graham JH, Riley TD, Carlton G (2006) Post-hurricane analysis of citrus canker spread and progress towards the development of a predictive model to estimate disease spread due to catastrophic weather events. *Plant Health Prog.*, DOI: 10.1094/PHP-2006-0822-1001-RS.

Jacques MA, Josi K, Darrasse A, Samson R (2005) *Xanthomonas axonopodis* pv. *phaseoli* var. *fuscans* is aggregated in stable biofilm population sizes in the phyllosphere of field-grown beans. *Appl. Environ. Microbiol.* **71**, 2008-2015.

Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution* **30**, 772-780.

Krzywinski M, Schein J, Birol I, *et al.* (2009) Circos: an information aesthetic for comparative genomics. *Genome Res* **19**, 1639-1645.

Lauf U, Muller C, Herrmann H (1998) The transposable elements resident on the plasmids of *Pseudomonas putida* strain H, Tn5501 and Tn5502, are cryptic transposons of the Tn3 family. *Mol Gen Genet* **259**, 674-678.

Leduc A, Traore YN, Boyer K, *et al.* (2015) Bridgehead invasion of a monomorphic plant pathogenic bacterium: *Xanthomonas citri* pv. *citri*, an emerging citrus pathogen in Mali and Burkina Faso. *Environmental Microbiology* **17**, 4429-4442.

Lee YA, Hendson M, Panopoulos NJ, Schroth MN (1994) Molecular-cloning, chromosomal mapping, and sequence analysis of copper resistance genes from *Xanthomonas campestris* pv. *juglandis* - homology with small blue copper proteins and multicopper oxidase. *Journal of Bacteriology* **176**, 173-188.

- Leyns F, De Cleene M, Swings JG, De Ley J (1984) The host range of the genus *Xanthomonas*. *Bot. Rev.* **50**, 308-356.
- Mellano MA, Cooksey DA (1988) Nucleotide sequence and organization of copper resistance genes from *Pseudomonas syringae* pv. *tomato*. *Journal of Bacteriology* **170**, 2879-2883.
- Nesme J, Cecillon S, Delmont TO, *et al.* (2014) Large-scale metagenomic-based study of antibiotic resistance in the environment. *Curr Biol* **24**, 1096-1100.
- Niu XN, Wei ZQ, Zou HF, *et al.* (2015) Complete sequence and detailed analysis of the first indigenous plasmid from *Xanthomonas oryzae* pv. *oryzicola*. *Bmc Microbiology* **15**, 233.
- Nonaka L, Maruyama F, Miyamoto M, *et al.* (2012) Novel conjugative transferable multiple drug resistance plasmid pAQU1 from *Photobacterium damsela* subsp. *damsela* isolated from marine aquaculture environment. *Microbes Environ* **27**, 263-272.
- Norberg P, Bergstrom M, Jethava V, Dubhashi D, Hermansson M (2011) The IncP-1 plasmid backbone adapts to different host bacterial species and evolves through homologous recombination. *Nat Commun* **2**, 268.
- Ochman H, Lawrence JG, Groisman EA (2000) Lateral gene transfer and the nature of bacterial innovation. *Nature* **405**, 299-304.
- Pak TR, Altman DR, Attie O, *et al.* (2015) Whole-genome sequencing identifies emergence of a quinolone resistance mutation in a case of *Stenotrophomonas maltophilia* bacteremia. *Antimicrob Agents Chemother* **59**, 7117-7120.
- Pansegrau W, Lanka E, Barth PT, *et al.* (1994) Complete nucleotide sequence of *Birmingham* IncP-Alpha plasmids - compilation and comparative analysis. *Journal of Molecular Biology* **239**, 623-663.
- Pele J, Becu JM, Abdi H, Chabbert M (2012) Bios2mds: an R package for comparing orthologous protein families by metric multidimensional scaling. *Bmc Bioinformatics* **13**, 133.
- Pereira UP, Gouran H, Nascimento R, *et al.* (2015) Complete genome sequence of *Xanthomonas arboricola* pv. *juglandis* 417, a copper-resistant strain isolated from *Juglans regia* L. *Genome Announc* **3**, e01126-01115.
- Perry JA, Wright GD (2013) The antibiotic resistance "mobilome": searching for the link between environment and clinic. *Front Microbiol* **4**, 138.
- Popa O, Dagan T (2011) Trends and barriers to lateral gene transfer in prokaryotes. *Curr Opin Microbiol* **14**, 615-623.
- Popescu AA, Huber KT, Paradis E (2012) ape 3.0: New tools for distance-based phylogenetics and evolutionary analysis in R. *Bioinformatics* **28**, 1536-1537.
- Pruvost O, Magne M, Boyer K, *et al.* (2014) A MLVA genotyping scheme for global surveillance of the citrus pathogen *Xanthomonas citri* pv. *citri* suggests a worldwide geographical expansion of a single genetic lineage. *Plos One* **9**, e98129.
- Richard D, Boyer C, Javegny S, *et al.* (2016) First report of *Xanthomonas citri* pv. *citri* pathotype A causing asiatic citrus canker in Martinique, France. *Plant Disease*, DOI: 10.1094/PDIS-1002-1016-0170-PDN.
- Richard, D., Tribot, N., Boyer, C., *et al.* (2017) First report of copper-resistant *Xanthomonas citri* pv. *citri*

pathotype A causing Asiatic citrus canker in Réunion, France. *Plant Disease*, DOI: 10.1094/PDIS-1009-1016-1387-PDN.

- Robinson JN, Callow JA (1986) Multiplication and spread of pathovars of *Xanthomonas campestris* in host and non-host plants. *Plant Pathol.* **35**, 169-177.
- Routh MD, Zalucki Y, Su CC, *et al.* (2011) Efflux pumps of the resistance-nodulation-division family: a perspective of their structure, function, and regulation in gram-negative bacteria. *Adv Enzymol Relat Areas Mol Biol* **77**, 109-146.
- Schuessler DL, Cortes T, Fivian-Hughes AS, *et al.* (2013) Induced ectopic expression of HigB toxin in *Mycobacterium tuberculosis* results in growth inhibition, reduced abundance of a subset of mRNAs and cleavage of tmRNA. *Mol Microbiol* **90**, 195-207.
- Seiler C, Berendonk TU (2012) Heavy metal driven co-selection of antibiotic resistance in soil and water bodies impacted by agriculture and aquaculture. *Front Microbiol* **3**, 399.
- Sen DY, Brown CJ, Top EM, Sullivan J (2013) Inferring the evolutionary history of IncP-1 plasmids despite incongruence among backbone gene trees. *Molecular Biology and Evolution* **30**, 154-166.
- Sevin EW, Barloy-Hubler F (2007) RASTA-Bacteria: a web-based tool for identifying toxin-antitoxin loci in prokaryotes. *Genome Biol* **8**, R155.
- Shao Y, Harrison EM, Bi D, *et al.* (2011) TADB: a web-based resource for Type 2 toxin-antitoxin loci in bacteria and archaea. *Nucleic Acids Research* **39**, D606-611.
- Siguier P, Perochon J, Lestrade L, Mahillon J, Chandler M (2006) ISfinder: the reference centre for bacterial insertion sequences. *Nucleic Acids Research* **34**, D32-36.
- Smillie C, Garcillan-Barcia MP, Francia MV, Rocha EPC, de la Cruz F (2010) Mobility of plasmids. *Microbiology and Molecular Biology Reviews* **74**, 434-+.
- Smith IM (1985) *Fungicides for crop protection : 100 years of progress : proceedings of the Bordeaux Mixture Centenary Meeting : Bordeaux, France, 5th-7th September 1985*, Bordeaux, France.
- Sommer DD, Delcher AL, Salzberg SL, Pop M (2007) Minimus: a fast, lightweight genome assembler. *Bmc Bioinformatics* **8**, 64.
- Stoyanov JV, Hobman JL, Brown NL (2001) CueR (YbbI) of *Escherichia coli* is a MerR family regulator controlling expression of the copper exporter CopA. *Mol Microbiol* **39**, 502-511.
- Sundin GW (2007) Genomic insights into the contribution of phytopathogenic bacterial plasmids to the evolutionary history of their hosts. *Annual Review of Phytopathology* **45**, 129-151.
- Telenti A, Imboden P, Marchesi F, *et al.* (1993) Detection of rifampicin-resistance mutations in *Mycobacterium tuberculosis*. *Lancet* **341**, 647-650.
- Tennstedt T, Szczepanowski R, Krahn I, Puhler A, Schluter A (2005) Sequence of the 68,869 bp IncP-1 alpha plasmid pTB11 from a waste-water treatment plant reveals a highly conserved backbone, a Tn402-like integron and other transposable elements. *Plasmid* **53**, 218-238.
- Trefault N, De la Iglesia R, Molina AM, *et al.* (2004) Genetic organization of the catabolic plasmid pJP4 from *Ralstonia eutropha* JMP134 (pJP4) reveals mechanisms of adaptation to chloroaromatic pollutants and evolution of specialized chloroaromatic degradation pathways. *Environmental Microbiology* **6**, 655-668.

- Triplett LR, Verdier V, Campillo T, *et al.* (2015) Characterization of a novel clade of *Xanthomonas* isolated from rice leaves in Mali and proposal of *Xanthomonas maliensis* sp. nov. *Antonie Van Leeuwenhoek* **107**, 869-881.
- Tsuda M, Iino T (1988) Identification and characterization of Tn4653, a transposon covering the toluene transposon Tn4651 on TOL plasmid pWW0. *Mol Gen Genet* **213**, 72-77.
- Vallenet D, Labarre L, Rouy Z, *et al.* (2006) MaGe: a microbial genome annotation system supported by synteny results. *Nucleic Acids Research* **34**, 53-65.
- Van Elsas JD, Turner S, Bailey MJ (2003) Horizontal gene transfer in the phytosphere. *New Phytol.* **157**, 525-537.
- Vauterin L, Yang P, Alvarez A, *et al.* (1996) Identification of non-pathogenic *Xanthomonas* strains associated with plants. *System. Appl. Microbiol.* **19**, 96-105.
- Vieites JM, Guazzaroni ME, Beloqui A, Golyshin PN, Ferrer M (2009) Metagenomics approaches in systems microbiology. *FEMS Microbiol Rev* **33**, 236-255.
- Vivian A, Murillo J, Jackson RW (2001) The roles of plasmids in phytopathogenic bacteria: mobile arsenals? *Microbiology* **147**, 763-780.
- Wang G, Kennedy SP, Fasiludeen S, Rensing C, DasSarma S (2004) Arsenic resistance in *Halobacterium* sp. strain NRC-1 examined by using an improved gene knockout system. *Journal of Bacteriology* **186**, 3187-3194.
- Weinthal DM, Barash I, Panijel M, *et al.* (2007) Distribution and replication of the pathogenicity plasmid pPATH in diverse populations of the gall-forming bacterium *Pantoea agglomerans*. *Appl Environ Microbiol* **73**, 7552-7561.
- Wozniak RA, Waldor MK (2010) Integrative and conjugative elements: mosaic mobile genetic elements enabling dynamic lateral gene flow. *Nat Rev Microbiol* **8**, 552-563.
- Yano H, Genka H, Ohtsubo Y, *et al.* (2013) Cointegrate-resolution of toluene-catabolic transposon Tn4651: determination of crossover site and the segment required for full resolution activity. *Plasmid* **69**, 24-35.
- Zhang WJ, Xu XR, Schwarz S, *et al.* (2014) Characterization of the IncA/C plasmid pSCEC2 from *Escherichia coli* of swine origin that harbours the multiresistance gene *cfr*. *J Antimicrob Chemother* **69**, 385-389.

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> respectively. Sequences produced in this study are deposited in the GenBank database (Table 1).

Author Contributions

PL, VR, CV and OP conceived and designed the study. BIC provided essential bacterial strains and produced phenotypic data on Argentinian strains. DR, CB, KB, PG, SJ, MT, BIC, CV and OP carried out the experiments. OP, CV and VR conducted population biology analyses. DR and PL conducted genomic data analyses. BF, IR and AR helped to design and carry out the study. DR, VR, AR, BF, IR, AC, CV, OP and PL wrote the manuscript. All the authors have read and approved the final manuscript.

Strain	Other numbers	Accession	Genus	Species	Pathovar	Copper resistance location	Country	Date	Host
LMG930		CP018463-CP018467	<i>Xanthomonas</i>	<i>euvesicatoria</i>		plasmid	USA	1969	Pepper
LMG911		CP018725-CP018727	<i>Xanthomonas</i>	<i>vesicatoria</i>		plasmid	New Zealand	1955	Tomato
LM091		CP017483	<i>Stenotrophomonas</i>	sp.		chromosome	Réunion	2015	Tangor
LM199	Xcc 15-4632 (INTA)	MSQV00000000	<i>Xanthomonas</i>	<i>citri</i>	<i>citri</i>	plasmid	Argentina	2015	Orange
LM180	Xcc 03-1638 (INTA); A44 ⁺	MSQW00000000	<i>Xanthomonas</i>	<i>citri</i>	<i>citri</i>	plasmid	Argentina	2003	Pomelo
LM159	Bv-5-4a (INTA)	CP018468-CP018471	<i>Xanthomonas</i>	<i>vesicatoria</i>		plasmid	Argentina	1987	Pepper
LL074-4		CP018847-CP018849	<i>Xanthomonas</i>	<i>citri</i>	<i>citri</i>	plasmid	Martinique	2014	Grapefruit
LJ207-7		CP018850-CP018853	<i>Xanthomonas</i>	<i>citri</i>	<i>citri</i>	plasmid	Réunion	2012	Kaffir lime
LH3		CP018472-CP018476	<i>Xanthomonas</i>	<i>'perforans'#</i>		plasmid	Maurice	2010	Tomato
LH276		CP018854-CP018857	<i>Xanthomonas</i>	<i>citri</i>	<i>citri</i>	plasmid	Réunion	2010	Kaffir lime
LH201		CP018858-CP018860	<i>Xanthomonas</i>	<i>citri</i>	<i>citri</i>	plasmid	Réunion	2010	Kaffir lime
JS749-3		CP018728-	<i>Xanthomonas</i>	<i>gardneri</i>		plasmid	Réunion	1997	Tomato

ICMP7383	CP018730						
	CP018731- CP018734	<i>Xanthomonas</i>	<i>gardneri</i>	plasmid	New Zealand	1980	Tomato

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*

Tables

Table 1 Characteristics of copper-resistant bacterial strains used for long-read sequencing

Figure captions

Fig. 1 Genetic diversity derived from microsatellite data of *Xanthomonas citri* pv. *citri* strains differing in geographical origin and their susceptibility to copper. (A) Manhattan distances represented by multidimensional scaling (MDS) among copper-resistant haplotypes from Argentina, Martinique and Réunion. The MDS plot showing axes 1 and 2 represented 68.0% and 6.2% of the total variation, respectively. Blue and red dots indicate strains bearing the copLAB and copABCD system, respectively, as confirmed by PCR. (B) Manhattan distances represented by multidimensional scaling (MDS) among copper-resistant and copper-susceptible haplotypes from Réunion, sampled between 2009 and 2015. Historical strains sampled between 1978 and 1997 were included as supplementary (suppl.) individuals in the MDS analysis. The MDS plot showing axes 1 and 2 represented 27.5% and 14.8% of the total variation, respectively. The black arrow localizes strain D07, which was identified as the historical strain genetically closest to the main clonal complex of epidemic copper-resistant strains.

Fig. 2 Alignments between pLH201.1 and sequences carrying pLH201.1 homologues from strains sequenced in this study. All sequences are plasmids, unless specified otherwise. On pLH201.1, blocks represent curated gene prediction (MAGE), whereas on other genomes the blocks represent uncurated Prodigal gene prediction. The pLH201.1 genes are categorized into several groups according to the caption. The comparison zone shows homologous nucleotide sequences. Colour varies according to genetic distance between homologous blocks defined by a MAUVE alignment.

Fig. 3 Circular representation of the plasmid pLH201.1. Circles, from the outer to the inner, represent: (1) predicted genes and their function according to the caption at the centre of the circle (putative genes are followed by *); (2) difference between mean GC content of the chromosome of LH201 and pLH201.1's % of GC content in a 2 000 bp sliding window with a 200 bp step; and (3) GC-skew using a 4 000 bp sliding window with a 200 bp step.

Fig. 4 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.

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. These strains were organized as a single clonal complex (i.e. a network of haplotypes linked by single-locus variations). Dot diameter is representative of the number of strains per haplotype. Colour indicates the strain origin and copper phenotype: light green = copper-susceptible Argentina; dark green = copper-resistant Argentina; khaki = copper-resistant non-copLAB Argentina; red = copper-resistant Martinique; light blue = copper-susceptible Réunion; dark blue = copper-resistant Réunion.

Fig. S2 Categorical minimum spanning tree of copper-resistant *Xanthomonas citri* pv. *citri* strains from Réunion based on microsatellite data. These strains were organized as five clonal complexes (i.e. networks of haplotypes linked by single-locus variations) and 28 singletons (i.e. haplotypes with no single-locus variants). Dot diameter represents the number of strains per haplotype. Single and double-locus variations were represented as solid and dotted lines joining haplotypes, respectively.

Fig. S3 Alignments between TnpLH201.1 and sequences carrying TnpLH201.1 homologues from strains sequenced in this study and *S. maltophilia* K279a. All sequences are plasmids, unless specified otherwise. On TnpLH201.1, blocks represent curated gene prediction (Mage), whereas on other genomes the blocks represent uncurated Prodigal gene prediction. The pLH201.1 genes are categorized into several groups according to the caption. The comparison zone shows homologous nucleotide sequences. Colour varies according to genetic distance between homologous blocks defined by a MAUVE alignment.

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 graphical representations of the copper-resistance genes organisation. The alignment used for the tree computation was obtained after the concatenation of alignments of each gene.

Fig. S5 Heatmap representation of the genetic divergence between six sequenced strains. Divergence values (proportions of variable nucleotides between two sequences) were obtained from the comparison of all homologous regions. Whereas the lower triangle represents the divergence between the chromosomes, the upper triangle represents divergence between the plasmids.

Fig. S6 Clusters of sequences with homologues to pLH201.1 genes. Each line represent a cluster of sequences obtained from GenBank that are homologous to at least 10 pLH201.1 genes. Squares represent homologous genes and are ordered as in the pLH201.1 sequence. Squares are coloured according to their level of nucleotide identity with pLH201.1, as indicated on the scale on the right. The names at the left of each line indicate the species of the parental sequence of the cluster, its GenBank identifier (gi) and the number of sequences in the cluster. Psp: *Pseudoxanthomonas* sp.; Sac: *Stenotrophomonas acidaminiphila*; Sma: *Stenotrophomonas maltophilia*; Sni: *Stenotrophomonas nitritireducens*; Ssp: *Streptomyces* sp.; Xar: *Xanthomonas arboricola*; Xax: *Xanthomonas axonopodis*; Xca: *Xanthomonas campestris*; Xci: *Xanthomonas citri*; Xeu: *Xanthomonas euvesicatoria*; Xga: *Xanthomonas gardneri*; Xor: *Xanthomonas oryzae*; Xpe: *Xanthomonas perforans*.







