CRISPR-associated sequence diversity within *Xanthomonas albilineans*, the causal agent of leaf scald disease of sugarcane

I. Pieretti[a], M. Marguerettaz[a], S. Bolot[b], S. Carrère[b], S. Cociancich[a], J. Gouzy[b], P. Rott[a] et M. Royer[a]

[a]*CIRAD, UMR BGPI, TA A-54/K Campus Baillarguet, 34398 Montpellier Cedex 5, France;* [b]*INRA, UMR LIPM, chemin de Borde-Rouge, 31326 Castanet-Tolosan Cedex, France*
isabelle.pieretti@cirad.fr

Xanthomonas albilineans is a xylem-invading pathogen that causes leaf scald, a lethal disease of sugarcane. Unlike other xanthomonads, *X. albilineans* exhibits a large intra-species variability which was previously observed with different genetic markers (PFGE for Pulsed Field Gel Electrophoresis and MLSA for Multi Locus Sequence Analysis). The CRISPR systems (Clustered Regularly Interspaced Short Palindromic Repeats) are repetitive structures in bacteria and Archaea composed of exact 24- to 48-bp repeated sequences (or ”repeats”) separated by unique sequences of similar length (or ”spacers”). Over 40 gene families, which are found nowhere except near these repeats, have been designated collectively as CRISPR-associated (cas) genes. CRISPR/cas systems participate in an antiviral response, probably by an RNA interference-like mechanism. Analysis of the variability of CRISPR spacers is currently used to perform diversity or epidemiological studies in bacteria. The genome sequence of *X. albilineans* revealed the occurrence of two different CRISPR/cas systems in this pathogen. The first system, called CRISPR-1, is associated with seven cas genes and contains repeats of 31 base pairs. It is similar to the CRISPR system found in several sequenced species of *Xanthomonas*. The second system, called CRISPR-2, is associated with six cas genes and contains repeats of 28 base pairs. There is only one *Xanthomonas* pathovar that is known to contain a similar CRISPR-2 system, namely *X. campestris* pv. *raphani*. In this study, we analyzed the polymorphism of the two different CRISPR/cas systems among 21 strains spanning the genetic diversity of *X. albilineans*. We have either sequenced PCR products resulting from amplification of spacers or cas genes, or used sequences from draft genome sequences. Whereas CRISPR-2 is ubiquitous within the 21 strains, CRISPR-1 is absent in three strains. The loss of CRISPR-1 by a common ancestor of these three strains is in accordance with the MLSA phylogeny. As described in other bacteria, we observed a variability of the CRISPR spacers, not only between phylogenetically distant strains, but also between closely related strains (acquisition of new spacers at the 5’ leader-proximal end of CRISPR and deletion or replacement of some spacers in the central region). This polymorphism within *X. albilineans*, which is congruent with previous MLSA and PFGE results, provides a better resolution of the phylogeny of *X. albilineans* strains.