The ISME Journal (2009) 3, 549–562

www.nature.com/ismej

ORIGINAL ARTICLE

Horizontal gene transfer between *Ralstonia solanacearum* strains detected by comparative genomic hybridization on microarrays

Alice Guidot[1], Bénédicte Coupat[2], Saliou Fall[2,3], Philippe Prior[1] and Franck Bertolla[2]

[1]*CIRAD, UMR Peuplements Végétaux et Bioagresseurs en Milieu Tropical, Saint Pierre, La Réunion, France and* [2]*Université de Lyon, Université Lyon 1, Institut Fédératif de Recherche 41, CNRS, UMR5557, INRA, USC 1193, Ecologie Microbienne, Villeurbanne, France*

The plant pathogenic Betaproteobacterium *Ralstonia solanacearum* is a complex species in that most of the strains share the common characteristic of being naturally transformable. In this study, we used a new approach based on comparative genomic hybridization (CGH) on microarrays to investigate the extent of horizontal gene transfers (HGTs) between different strains of *R. solanacearum*. Recipient strains from phylotypes I, II and III were naturally transformed *in vitro* by genomic DNA from the GMI1000 reference strain (phylotype I) and the resulting DNAs were hybridized on a microarray representative of the 5120 predicted genes from the GMI1000 strain. In addition to transfer of the antibiotic resistance marker, in 8 of the 16 tested transformants, CGH on microarrays detected other transferred GMI1000 genes and revealed their number, category, function and localization along the genome. We showed that DNA blocks up to 30 kb and 33 genes could be integrated during a single event. Most of these blocks flanked the marker gene DNA but, interestingly, multiple DNA acquisitions along the genome also occurred in a single recombinant clone in one transformation experiment. The results were confirmed by PCR amplification, cloning and sequencing and Southern blot hybridization. This represents the first comprehensive identification of gene acquisitions and losses along the genome of the recipient bacterial strain during natural transformation experiments. In future studies, this strategy should help to answer many questions related to HGT mechanisms.

The ISME Journal (2009) **3,** 549–562; doi:10.1038/ismej.2009.14; published online 26 February 2009
Subject Category: evolutionary genetics
Keywords: horizontal gene transfer; microarrays; *Ralstonia solanacearum*

Introduction

The gene content of bacterial genomes varies even though bacteria reproduce clonally. Along with the loss of genetic material resulting in genome contraction, the ability of bacterial species to acquire new genetic information from other strains or species by horizontal gene transfer (HGT) is a dominant force of variation in gene content (Koonin and Wolf, 2008). Three mechanisms have been reported to be involved in gene transfer: conjugation, transduction and transformation. During conjugation, DNA is transferred directly from one organism to another, whereas in transduction the DNA is carried by bacteriophages. Transformation is the process by which a competent cell is able to take up DNA molecules directly from its surrounding and incorporate it. It is now known that HGT is one of the main mechanisms involved in the evolution of bacterial genomes (Gogarten and Townsend, 2005).

The major role of HGT in the evolution of bacteria and archaea has recently been recognized as extensive comparison of multiple complete genome sequences became possible (Nakamura *et al.*, 2004; Koonin and Wolf, 2008). One outstanding result from comparative genomic analyses demonstrating the high prevalence of HGT and their major role in bacterial evolution was the discovery of 'pathogeni-

Correspondence: F Bertolla, UMR CNRS 5557 USC INRA 1193, Université Claude Bernard Lyon 1, 16 rue Raphaël Dubois/ Bâtiment Gregor Mendel, F-69622 Villeurbanne, France.
E-mail: bertolla@biomserv.univ-lyon1.fr
[3]Current address: Environmental Microbial Genomics Group, Laboratoire AMPERE UMR CNRS 5005, Ecole Centrale de Lyon et Université de Lyon, 36 avenue Guy de Collongue, 69134 Ecully, France.
Received 27 October 2008; revised 26 January 2009; accepted 26 January 2009; published online 26 February 2009

city islands' in pathogenic bacteria and similar 'symbiosis islands' in symbiotic bacteria (Ochman *et al.*, 2000; Ochman and Moran, 2001). Pathogenic islands are gene clusters that carry pathogenicity determinants, such as genes encoding various toxins or components of type III secretion systems. These islands are typically localized near tRNA genes and contain multiple prophages, suggesting that their insertion in bacterial genomes is mediated by bacteriophages (Hacker and Kaper, 2000). Another demonstration of the contribution of HGT in the evolution of bacteria has been obtained with photosynthetic bacteria. Phylogenetic analysis strongly suggested that photosynthetic gene clusters are complex mosaics of genes assembled through multiple HGT events (Raymond *et al.*, 2002). In addition, the observation that many bacteria developed cellular barriers to transformation such as restriction-modification systems (Kobayashi, 2001) and the methyl-mismatch repair system (Matic, 1995) 'protecting' the host DNA against 'contamination' by foreign sequences is in agreement with the occurrence of HGT in nature (Thomas and Nielsen, 2005).

Methods for detecting HGT in bacterial genomes can be divided into two major categories. The first consists of an analysis of nucleotide sequences of genes. By using bioinformatic tools, putative horizontally transferred genes have been detected by comparing the nucleotide composition, codon usage and amino-acid usage of each open reading frame (ORF) to other ORFs in a genome (Nakamura *et al.*, 2004; Coenye and Vandamme, 2005). However, it has not been unequivocally shown that HGT is the sole cause of unusual compositional bias (Gogarten and Townsend, 2005). Phylogenetic conflicts among loci have also been widely examined to screen for horizontally transferred genes (Comas *et al.*, 2006; Fall *et al.*, 2007). However, not all acquired genes have compositional features that distinguish them from other genes. This is particularly true when donor DNA originates from an organism closely related to the recipient (HGT between strains of the same species). The second method monitors the genomic integration of foreign DNA through experimental approaches. Recombinant genomes are detected through the acquisition of selective markers such as genes coding for the resistance to antibiotics or heavy metals or for the utilization of rare carbon sources (xenobiotics) (Bertolla *et al.*, 1999; Davison, 1999; de Vries and Wackernagel, 2002). However, this method does not allow detection of additional genes transferred into the recipient genome.

In this study, we investigated HGTs between strains of the plant pathogenic Betaproteobacterium *Ralstonia solanacearum*. This bacterium is characterized by a high level of phenotypic and genotypic diversity. The host range of *R. solanacearum* is unusually wide for a plant pathogen, including over 450 host species in 54 botanical families (Allen *et al.*, 2005). Four monophyletic groups of strains, termed phylotypes, have been distinguished through nucleotide sequence analysis of four genes (Fegan and Prior, 2005). These phylotypes are correlated with the geographical origin of the strains. *R. solanacearum* has been described as a highly flexible organism capable of rapid adaptation to environmental changes. Specifically, the ability to rapidly counteract plant resistance and the emergence of new pathological variants of *R. solanacearum*, as recently observed in Martinique (French West Indies), are major issues in agronomic research aimed at controlling the disease (Wicker *et al.*, 2007). The fundamental role of HGT in *R. solanacearum* genome evolution has been supported by the findings of several studies. Bertolla *et al.* (1997, 1999) demonstrated that this bacterium is able to naturally develop the physiological state of competence required to exchange genetic material by transformation under *in vitro* as well as *in planta* conditions. The hypothesis of HGT *in planta* is supported by the observation of mixed infections of *R. solanacearum* strains from two phylotypes in tomato in the field (P Prior, personal communication). Recently, Coupat *et al.* (2008) demonstrated that transformability is ubiquitously shared among the four phylotypes of *R. solanacearum* and that large DNA fragments ranging from 30 to 90 kb can be exchanged. *In silico* analysis of the complete genome sequence of the *R. solanacearum* GMI1000 strain revealed the mosaic structure of both the 3.7 Mb chromosome and the 2.1 Mb megaplasmid that constitute the bacterium's genome (Salanoubat *et al.*, 2002). More than 7% of the genome contains regions with a G + C-biased composition and alternative codon usage regions (ACURs), frequently surrounded by IS and phage elements. The authors assumed that these genomic regions were acquired from other species by HGT (Salanoubat *et al.*, 2002). Complementary analysis based on phylogenetic reconstruction of prokaryote homologous gene families or on nucleotide composition detected that 13–15% of the *R. solanacearum* GMI1000 genes originated from HGT events (Nakamura *et al.*, 2004; Fall *et al.*, 2007). Recently, an analysis of gene distribution among *R. solanacearum* strains by comparative genomic hybridization (CGH) on the microarray identified a list of 2338 variable genes (40% of the genome) within the species (Guidot *et al.*, 2007). These variable genes often cluster within genomic islands, suggesting that they might originate from HGT.

In this study, our aim was to determine the number and functions of genes that can be horizontally transferred between *R. solanacearum* phylotypes in addition to the marker gene during natural transformation. For that purpose, we conducted *in vitro* transformation experiments using genomic DNA from the GMI1000 reference strain (phylotype I) as donor DNA and three other strains from phylotypes I, II and III as recipient strains. Recombinant genomes were analyzed by CGH experiments on the recently described microarray for the GMI1000 strain (Occhialini *et al.*, 2005) to identify all

acquired or lost genes in the recombinant genomes. The hypotheses we tested here were (i) the use of microarrays and CGH is a powerful way to identify gene acquisitions and losses, their number, category, function and localization along the genome of the recipient bacterial strain during natural transformation, (ii) multiple DNA acquisitions along the genome can occur in a single recombinant clone in one transformation experiment and (iii) a high number of genes including virulence genes can be horizontally transferred between *R. solanacearum* phylotypes during a single HGT event, which could play a major role in the adaptive evolution of this plant pathogenic bacterium.

Materials and methods

Bacterial strains and growth conditions

The *R. solanacearum* strain GMI1000, isolated from tomato in Guyana and classified in phylotype I, was selected as the donor strain because its genome is completely sequenced and a microarray representative of the 5120 predicted GMI1000 genes is available. The *R. solanacearum* recipient strains used in this study were CFBP2968, isolated from eggplant in the French West Indies and classified in phylotype I, CFBP2957, isolated from tomato in the French West Indies and classified in phylotype II and NCPPB332, isolated from potato in Zimbabwe and classified in phylotype III (Guidot *et al.*, 2007; Wicker *et al.*, 2007). These recipient strains were selected because they are representative of the three most studied phylotypes of *R. solanacearum* and previous works demonstrated that they have comparable ability to be transformed with GMI1000 genomic DNA (Coupat *et al.*, 2008). The strains were cultured at 28 °C in complete medium B (Boucher *et al.*, 1985). All strains exhibited sensitivity to ampicillin, kanamycin and gentamicin.

Construction of mutant strains

Genomic DNAs from *R. solanacearum* GMI1000 with mutations in *recA*, *pilA*, *comA*, *mutS*, RSc1815, *ftsK*, RSc3252, *vsr*, *popA* and RSp1328 were used as transforming DNA (Table 1). These mutants were constructed by natural transformation with 10 recombinant plasmids containing 2 kb of GMI1000 genomic DNA, including the target gene in which the *aacC3-IV* gene conferring resistance to gentamicin had been inserted in specific insertion site localized approximately in the middle of the target gene (Fall *et al.*, 2007). Genomic DNA from the CFBP2957 *recA*::Gm mutant was also used as homologous DNA control. The *recA*, *mutS* and *vsr* genes were selected because their function is more or less directly involved in controlling genomic integrity. The *comA* gene is required for competence development. The RSp1328 gene is located in a 31-kb tandem region of the megaplasmid. The RSc1815 and RSc3252 genes were identified as being recently acquired by HGT putatively from *Xyllela campestris* and *Chlorobium tepidum*, respectively (Fall *et al.*, 2007). The *ftsK* gene was selected because it is a well-conserved housekeeping gene. The *pilA* and *popA* genes were chosen because they are involved in plant pathogenicity. The recombinant plasmids were extracted and purified by using the QIAprep Mini-prep Kit (Qiagen, Hilden, Germany). Before being used in mutant construction in GMI1000, they were digested with *Sca*I, thus linearizing them without affecting the GMI1000 fragment and the gentamicin resistance gene (marker gene). Genomic DNAs from the GMI1000 mutants, used as transforming DNA in transformation experiments, were extracted using the DNeasy Tissue Kit (Qiagen).

Natural transformation in R. solanacearum

R. solanacearum cells were grown in a minimal medium MM (Bertolla *et al.*, 1997), to an optical density at 580 nm of 0.8. An aliquot of 50 µl of this cell suspension was mixed with 200 ng of transforming DNA. The resulting suspension was incubated on MM plate for 48 h at 30 °C. Each transformation experiment was carried out at least in triplicate and included determination of the spontaneous mutation rate resulting in gentamicin resistance by plating the cells on selective medium containing gentamicin (12 µg ml^{-1}).

Table 1 Description of *Ralstonia solanacearum* genes used as target genes in the transformation experiments

Accession number	*Gene name*	*Function*	*Description*
RSc0551	*recA*	DNA strand exchange and recombination protein	Housekeeping
RSc0558	*pilA*	Signal peptide type 4 fimbrial pilin	Virulence
RSc1120	*comA*	Transmembrane protein	Competence
RSc1151	*mutS*	DNA mismatch repair protein	Housekeeping
RSc1815		Avrbs3 family type III effector protein	ACUR
RSc2341	*ftsK*	Transmembrane cell division protein	Housekeeping
RSc3252		Peptidase M	ACUR
RSc3439	*vsr*	Patch repair protein (DNA mismatch endonuclease)	Housekeeping, ACUR
RSp0877	*popA*	PopA protein	Virulence, effector
RSp1328		Efflux pump antibiotic resistance protein	In tandem region

Abbreviations: ACUR, alternative codon usage region; RSc and RSp, gene identity from the *R. solanacearum* GMI1000 chromosome and megaplasmid, respectively.

Microarray description
The DNA microarray used in these experiments was generated by Occhialini *et al*. (2005). This microarray consists of 5074 65- and 70-mer oligonucleotides representative of the 3442 (RSc0001 to RSc3442) predicted genes from the chromosome and the 1678 (RSp0001 to RSp1678) predicted genes from the megaplasmid of the *R. solanacearum* GMI1000 strain. Each predicted gene in the GMI1000 genome is represented by a single oligonucleotide and each oligonucleotide is spotted twice on the microarray. This microarray also includes, as negative controls, 10 oligonucleotides corresponding to five *Corynebacterium glutamicum* genes and a set of 'blank' controls in which buffer without oligonucleotide was spotted. The additional 26 oligonucleotides corresponding to *R. solanacearum* sequences not present in strain GMI1000 were not considered for the present analysis.

DNA labeling and hybridization
Genomic DNA was fluorescently labeled with Cy3 or Cy5 fluorescent dye (Amersham Biosciences, Freiburg, Gemany) using the BioPrime DNA Labeling System Kit (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's recommendations, as described in Guidot *et al*. (2007). Fluorescently labeled DNA was purified using the CyScribe GFX Purification Kit (Amersham Biosciences) and dissolved in 60 µl of elution buffer. Hybridizations were carried out using a Lucidea automated slide processor (Amersham Pharmacia Biotech, Buckinghamshire, England). Each experiment was run as a competitive hybridization by using Cy3-labeled DNA from one of the recipient strains (wild-type strains or recombinant clones) and Cy5-labeled DNA from the donor GMI1000 wild-type strain. No dye swapping was performed as preliminary experiments had demonstrated that this had no significant impact on final results. Hybridization and washing conditions were as described in Guidot *et al*. (2007). Each hybridization experiment was carried out twice.

Array scanning and analysis
Hybridized microarrays were scanned using GenePix 4000A dual-channel (635 and 532 nm) confocal laser scanner (Axon Instruments, Union City, CA, USA) with a resolution of 10 nm per pixel. The laser power was set at 100 and the photomultiplicator tension was adjusted between 680 and 800 V, according to the average hybridization intensity of each slide to optimize the dynamic range of measurements. Signals from individual arrays were quantified using ImaGene 5.6.1 software and analyzed using Genesight 3.5.2 software (BioDiscovery Inc., El Segundo, CA, USA). Empty spots and spots with impurities, high local background fluorescence or weak intensity compared with the signal observed for hybridization to the negative controls were excluded from the analysis. This concerned fewer than 5% of spots. For each spot, the ratio of the hybridization signal of the recipient strain versus the donor GMI1000 strain was calculated, $\log_2$ transformed, and the values thus obtained were normalized by subtracting the mean $\log_2$ ratio value calculated on a set of 2690 conserved genes in *R. solanacearum* strains, as described in Guidot *et al*. (2007). Finally, the average $\log_2$ ratio value of the four spots representing each gene (two slides with two spots per gene) was calculated and used for further analysis. The lists of GMI1000 genes present in the recipient strains (wild-type strains and recombinant clones) were established by selecting genes for which the calculated average $\log_2$ ratio was above -2 (that is, by excluding genes for which the hybridization signal of the recipient strain was at least fourfold weaker than that of the GMI1000 reference strain). This cutoff value was defined in Guidot *et al*. (2007). The lists of GMI1000 genes transferred into the recombinant clones were established by comparing lists of the GMI1000 genes that were present in the wild-type strains and recombinant clones. In this study, we assumed that the genomic synteny was the same in all tested *R. solanacearum* strains to localize recombination events along the genome.

All primary data from the microarray experiments as well as experimental protocols used are available from the ArrayExpress depository (accession numbers A-MEXP-152, E-MEXP-422 and E-MEXP-2012 at http://www.ebi.ac.uk/arrayexpress/).

Controls for gene acquisition or loss
Recombinant *R. solanacearum* colonies that exhibited resistance to gentamicin were subjected to DNA replacement of the wild-type gene by the chimeric construction. This allelic replacement was verified by PCR for each construction with primers located at each extremity of the target gene (Table 2). Acquisition of genes originating from the GMI1000 strain and loss of genes in the gentamicin-resistant recombinant clones, as detected by microarray hybridization, were checked by PCR amplification, cloning and sequencing and Southern blot hybridization. The PCR primers and probes were used to anneal inside the region selected for the 70–65-mer-long oligonucleotide designed and spotted on the microarray. The list of primers and probes used is given in Table 2. PCR was performed on 200 ng of recombinant DNA with 10% (w/v) of dimethylsulfoxide and Platinum *Taq* DNA polymerase (Invitrogen). Amplification reactions consisted of denaturation at 95 °C for 4 min followed by 30 cycles of denaturation at 95 °C, annealing at the appropriate temperature (Table 2) and elongation at 72 °C for 50 s each. A final elongation was conducted for 4 min at 72 °C. In some cases, when PCR amplification was not sufficient to clearly demonstrate the GMI1000 gene acquisition, PCR products were cloned in pCR2.1 vector with the TOPO-TA Cloning Kit (Invitrogen) and sequenced.

Table 2 List of primers used in this study

Gene	*Names of primers and probes*	*Sequences*	*Annealing temperature (°C)*	*PCR fragment length (bp)*
aacC3-IV	F4164/F4165	GGCCACAGTAACCAACAAAT/CGCTCATCAATCTCCTCAA	50	694
Chromosome				
RSc003–RSc004	F5583/F5584	GCTGCTGTTCCTGCTGCGTGT/GCCGCGGCCTTCTTATTGGTG	60	1962
recA (RSc0551)	F1022/F1023	CTTGCGGCCAGTTTA/CGTACCGGCAAAGCA	55	2388
RSc0555	F5239/F5240	CCTGCTGCTGGGTCCGAACTG/TGCCGCACTCTTCCATAATGT	58	461
RSc0557	F5555/F5556	CGCGGCTTTACATTGATCGAG/GGCACTCCGAGGGCGCGTATT	56	478
pilA (RSc0558)	F2092/F2093	CGCGTCGAACCCAAGAAC/GGCTTCCGCTCATCGTCA	52	2342
comA (RSc1120)	F1024/F1025	CGAGCTGCCCGAAGT/CCTGCGGATGCGGAT	55	2069
mutS (RSc1151)	F530/F1569	GCGTGTGGCTGCGGTTGTGC/GACAAGCACACGCCGATGATG	60	2160
RSc1755	F4864/F4865	GCCGATGCCCGTGTGTATTCC/GCGGCGTTCGTCGGTCAA	59	192
RSc1815	F2253/F2254	AGGCGCAACTGCTGGAG/TGGGCGACACAAACAGGT	56	2287
ftsK (RSc2341)	F2251/F2252	TTCCAGGGATGCGGTAAC/CCCGAGCTTCCACCACT	55	2377
RSc2343	F3508/F3509	CCAGCATCGGCGAAGTGAG/GCGCACCTTGCCCTTGAG	55	400
RSc3252	F2249/F2250	TAGGAACGGCCACAGTCA/CCTCGGCGTCTACAAGC	56	2242
RSc3412	F4975/F4976	GCCGGCGCGACGATGGAAC/TACCGCGCCAGCTTGAACACC	61	248
RSc3413	F4858/F4859	GGCGCATCTGTGGTTCGTCA/TCAGCCCGTCAGTTCGTAGGA	58	511
RSc3417	F4860/F4861	GGCAGCGACCTCCCGTATCTC/GGCGCAAGGTTGACGAT	56	320
RSc3426	F4862/F4863	AGTCCGTGCCCATCTAAGCCA/GCCCGTCTGACCCGTCTC	56	371
vsr (RSc3439)	F4999/F5000	TACGCTCAGCCATGCATC/AGCGAGATTTCTTTACGTGTT	53	295
RSc3440	F4977/F4978	GTCGAGATGGCCGCGTACCTG/CCTGCTCGCGCTGCTTATCC	61	514
RSc3442	F5296/F5297	CAACCCGGGCAAGTCGTACAA/CTCCTCCGGCACCGTCACGTT	58	484
Megaplasmid				
RSp0869	F3514/F3515	GTTGCCGACGACCTCATCA/CCGCCAGCTGGACACTCA	60	457
RSp0875	F4866/F4867	GCCGCGTTCGCTCGTGTTG/TCATCGGTGCGGCCATACTGC	59	305
RSp0876	F3516/F3517	AATCAAGGCTGGCGGACAC/GCTGCAGGGCGATGTTCAT	60	437
popA (RSp0877)	F377/F378	CTCAACATCATCGCAGCCCT/CGCCGCCGAGCTGGCGAGG	55	400
RSp0885	F3518/F3519	GCCAAAGCCAGCATCAAC/GCCTGGCGAAGTGTTCAAT	56	586
RSp1027	F5001/F5002	CGGCTCAACGGCAATATCAC/ATTGAACTTGGGCAGCGTGTC	57	279
RSp1270	F5003/F5004	GCGCCACCTTCTCCGACT/CGTGCGCCATGACAGGTAGAC	60	413
RSp1272	F4870/F4871	TCGCCCGGCTTCCCATCGTGT/GGCTGGCGCACGCGATAGG	62	337
RSp1277	F4879/F4880	GGCCGTGGCATCGACATC/CACCGAGCGCACGAAGCCTAC	57	231
RSp1303	F4995/F4996	TACGGCTTCTGGCGCATTTAC/TTCCCGCGTATGGTTGTCC	56	261
RSp1304	F4850/F4851	CAGCGCAACATGGCCGACAGT/CGCCGCGGCATGAGTACAAGA	56	230
RSp1316	F4852/F4853	CCCGTGCGGAAAGGTGTCGAT/GCGCCGCTCGATGATCT	56	364
RSp1328	F2245/F2246	CCAGGCAGAGGCGACAC/GGCGGGAGTGGATGTTGT	58	2352
RSp1330	F4854/F4855	GCGCACAGCCAGGAGTTAAGG/CGGGCACGCAATGAATAGCAG	56	412
RSp1331	F5551/F5552	AGACCCTGGCGATGGCATTAC/CTTGTCTGCAGCGAGTCGTTC	56	410
RSp1330–1334	F5261/F5262	CTGGACCGTGCTCGCGTTTGC/CAGGCTACTTCGCTTGTCATA	56	3750
RSp1334	F4856/F4857	CAGCGCAACATGGCCGACAGT/CGCCGCGGCATGAGTACAAGA	56	230
RSp1363	F4997/F4998	TGATCGTGCTGGCCGTTCTTG/CTTGCGCGACGACGTGTTGA	60	494
RSp1304	F5094	GCGGCATGAGTACAAGAGTTACCTTTCCCT	60	
RSp1316	F5095	TTCTGCCCGTGACCGGGAACCCGACCATCC	60	
RSp1330	F5096	CTGGTCACGCTATGTGGGGCAGGAGAAGCT	60	

Southern blot hybridizations were performed for the lost DNA regions. Probes were labeled in 5′ termini by forward reaction with T4 polynucleotide kinase (Fermentas, St Rémy Les Chevreuse, France) using [γ-^{32}P]ATP, as recommended by the manufacturer. Hybridizations were performed on *Nco*I-digested genomic DNA.

Results

GMI1000 gene content in the three tested R. solanacearum *wild-type strains*

In the first step, we identified the GMI1000 genes that were also present in wild-type (-WT) strains, CFBP2968, CFBP2957 and NCPPB332. Hybridization of genomic DNA from the CFBP2968-WT strain on the GMI1000 microarray revealed that 98% of the GMI1000 genes were present in this strain. CFBP2968 and GMI1000 are phylogenetically close strains, both belong to phylotype I. Only two clusters of 52 and 48 genes from the GMI1000 chromosome were absent in the CFBP2968 genome. The first cluster (RSc1896–RSc1948) was a DNA region of phage origin. The second cluster (RSc2574–RSc2622) was an ACUR linked to conjugal transfer proteins. Concerning the CFBP2957-WT strain from phylotype II, hybridization of its genomic DNA on the GMI1000 microarray revealed that only 69% of the GMI1000 genes were present in

Table 3 Transformation-recombination frequencies in four strains of *Ralstonia solanacearum* after natural transformation by genomic DNA from different mutants of the GMI1000 strain

Genomic DNA of GMI1000 mutants	*Transformation frequency*[a]			
	GMI1000	*CFBP2968*	*NCPPB332*	*CFBP2957*
recA::Gm	1.0×10^{-5}	4.2×10^{-6}	2.4×10^{-7}	4.0×10^{-7}
pilA::Gm	2.5×10^{-6}	1.3×10^{-6}	3.5×10^{-7}	4.4×10^{-7}
comA::Gm	4.2×10^{-6}	3.0×10^{-6}	1.1×10^{-6}	9.6×10^{-7}
mutS::Gm	8.4×10^{-6}	3.5×10^{-6}	5.9×10^{-7}	6.0×10^{-7}
RSc1815::Gm	1.1×10^{-6}	8.3×10^{-7}	2.4×10^{-8}	1.7×10^{-10}
ftsK::Gm	6.6×10^{-6}	3.4×10^{-7}	3.9×10^{-7}	1.3×10^{-7}
RSc3252::Gm	4.6×10^{-6}	1.3×10^{-6}	$< 1.0 \times 10^{-8}$	$< 1.4 \times 10^{-9}$
vsr::Gm	6.6×10^{-6}	1.6×10^{-6}	5.6×10^{-8}	3.8×10^{-9}
popA::Gm	7.2×10^{-6}	2.0×10^{-6}	2.3×10^{-7}	1.3×10^{-8}
RSp1328::Gm	8.5×10^{-6}	2.4×10^{-6}	1.8×10^{-8}	2.8×10^{-9}
Control[b]	ND	ND	ND	ND

Abbreviations: Gm, gentamicin-resistant gene; ND, gentamicin-resistant clone not detected.
[a]Transformation frequencies are expressed as the number of transformants per recipient cells. The data presented here are the mean of three independent replicates.
[b]Transformation control without DNA. This table was modified with permission from Fall *et al*. (2007).

this strain. A total of 1581 GMI1000 genes were classified as absent or highly divergent in the CFBP2957 genome. Some of these genes were isolated along the chromosome and the megaplasmid but the majority clustered within genomic islands of 3–50 genes. A similar result was obtained for the NCPPB332-WT phylotype III strain in which only 81% of the GMI1000 genes were present. A total of 940 GMI1000 genes were classified as absent or highly divergent in the NCPPB332 genome and most of these genes clustered within genomic islands of 3–50 genes.

Selection of the recombinant clones analyzed

Transformation experiments were conducted using total genomic DNA from different mutants of the GMI1000 strain as donor DNA and CFBP2968, CFBP2957 and NCPPB332 as recipient strains. Transformation frequencies varied from 4.2×10^{-7} to 1.7×10^{-10} (Table 3) depending on the DNA and the recipient strain (Fall *et al*., 2007). To appreciate the effects of heterologous transformation by GMI1000 DNA, a transformation control in the CFBP2957 strain was performed with the homologous genomic DNA from the CFBP2957 *recA*::Gm mutant. This transformation yielded genetic transformants at a frequency 10-fold higher than that of the heterologous genomic DNA *recA*::Gm from the GMI1000 strain, and a minor sequence divergence of the *recA* gene (4%) lead to a significant drop in transfer frequencies.

In this study, we analyzed one recombinant clone for each mutant to get a better overview of the various HGT events that could be detected by CGH on the microarray. A total of 16 recombinant clones were then selected, 2 of which originated from the CFBP2968 strain (the CFBP2968-*mutS* and -*ftsK* recombinant clones), 8 from the CFBP2957 strain (CFBP2957-*recA*, -*comA*, -RSc1815, -*ftsK*, -RSc3252, -*vsr*, -*popA* and -Rsp1328) and 6 from the NCPPB332 strain (NCPPB332-*recA*, -*pilA*, -*mutS*, -*ftsK*, -*vsr* and -*popA*). In all clones, replacement of the target genes by chimeric constructions was verified by PCR amplification (data not shown).

To identify potential new gene integrations or gene losses in addition to transfer of the marker gene during transformation with GMI1000 genomic DNA, we compared the GMI1000 gene content in wild-type and recombinant genomes by CGH on the GMI1000 microarray. This analysis allowed the detection of additional integrations of GMI1000 genes in eight clones (NCPPB332-*pilA*, -*vsr* and -*popA*; CFBP2957-*recA*, -*comA*, -*ftsK*, -*vsr* and -*popA*) and gene losses in one clone (CFBP2968-*mutS*) (Table 4). We did not discover any genomic rearrangement events in addition to replacement of the wild-type gene by the chimeric construction in the seven other gentamicin-resistant recombinant clones.

Description of gene acquisitions in the NCPPB332 recombinant clones

Gene acquisitions in addition to marker gene transfer were observed in three of the six gentamicin-resistant recombinant clones of the NCPPB332 strain (NCPPB332-*popA*, -*vsr* and -*pilA*). Hybridization of genomic DNA from the NCPPB332-*pilA* clone on the GMI1000 microarray revealed acquisition of the GMI1000 RSc0555 gene encoding a succinyl-coA synthetase in addition to the *pilA* gene (RSc0558) (Table 4). PCR amplification confirmed acquisition of this gene from the GMI1000 strain. In this case, the transferred gene was very close to the *pilA* gene, indicating that a cluster of at least four GMI1000 genes (RSc0555–RSc0558) had been transferred into the NCPPB332-*pilA* recombinant genome during the transformation experiment. This hypothesis was confirmed by the sequence of the RSc0557 gene of

Table 4 Detection by comparative genomic hybridization on microarrays of acquired or lost genes in recombinant clones

Recombinant clones	*Localization of genomic modifications*	*Number of ORFs and size (bp) of DNA modified*	*Verification*	*Putative functions*
Acquisition				
NCPPB332-*pilA*	RSc0555–0558	4; 3217	PCR sequencing	Succinyl-coA synthetase, transmembrane protein, signal peptide type 4
NCPPB332-*vsr*	RSc3413–3439–0003	33; 29 131	PCR sequencing	ACURs (24 ORFs), transposase, chemotaxis, hypothetical proteins, DBA methylase, DNA gyrase, DNA polymerase III, replication initiator protein, ribosomal protein, RNA degradation, protein secretion
NCPPB332-*popA*	RSc1755	1; 333	PCR sequencing	Transposase
	RSp0875–0877	3; 4711	PCR	Type III effectors PopA, PopB, PopC
	RSp1027	1; 1575	PCR	Chemotaxis transducer
	RSp1270–1277	8; 11 480	PCR	Ferrisiderophore receptor, helicase II, signal peptide belonging to the *hrpB* pathogenicity regulons, type III effector, hypothetical proteins
CFBP2957-*recA*	RSc0551–RSc0573	23; 22 099	PCR	Molecule metabolism, structural element of cell surface, transport of small molecules, hypothetical proteins
CFBP2957-*comA*	RSc1117–1120–1138	22; 21 587	PCR	Molecule metabolism, structural element of cell surface, cell division, chaperone heat-shock-like protein, protein secretion, hypothetical proteins
CFBP2957-*ftsK*	RSc2341–2343	3; 4308	PCR	Oxidoreductase, hypothetical protein
CFBP2957-*vsr*	RSc3412–3439	27; 22 417	PCR	ACURs (24 ORFs), transposase, molecule metabolism, mobility chemotaxis, hypothetical proteins
CFBP2957-*popA*	RSp0869–0877–0885	17; 27 380	PCR	Pathogenicity-related functions (14 ORFs), molecule metabolism, protein secretion, hypothetical proteins
Loss				
CFBP2968-*mutS*	RSp1304–1330	27; 29 278	PCR Southern blot	ACURs (11 ORFs), transmembrane, hypothetical proteins, small molecule metabolism, transcription regulators, efflux pumps

Abbreviations: ACURs, alternative codon usage regions; ORF, open reading frame; RSc and RSp, gene identity from the *R. solanacearum* GMI1000 chromosome and megaplasmid, respectively.
The marker genes used for selection of recombinant clones are highlighted in grey.

this recombinant clone, which had only 91% identity with the parental NCPPB332-WT strain, whereas 100% identity was found with the donor GMI1000 strain.

For the NCPPB332-*vsr* clone, hybridization of genomic DNA on the GMI1000 microarray revealed the acquisition of a cluster of 27 GMI1000 genes (RSc3413–3439) flanking the *vsr* gene. This gene cluster corresponded to an ACUR (Figure 1). Acquisition of this GMI1000 gene cluster was confirmed by PCR amplification of the RSc3412, RSc3413, RSc3417, RSc3426, RSc3439 and RSc3440 GMI1000 genes. Positive amplification of the expected bands was obtained with both GMI1000-WT and NCPPB332-*vsr* DNA for all tested genes (Figure 1). In agreement with the CGH results, no amplification was obtained for the RSc3413, RSc3417, RSc3426 and RSc3439 genes with the NCPPB332-WT DNA, thus confirming the HGT. Amplification and sequencing of a fragment from the RSc3412 gene demonstrated that this gene in the NCPPB332-*vsr* clone originated from the NCPPB332-WT strain (100% sequence identity with NCPPB332-WT and 98% sequence identity with GMI1000) and could be considered as the gene where the first crossover occurred. Conversely, amplification and sequencing of a fragment from the RSc3440 gene revealed that this gene in the NCPPB332-*vsr* clone still originated from the GMI1000 donor strain (93% sequence identity with NCPPB332-WT and 100% sequence identity with GMI1000). Further amplification and sequencing of fragments from the RSc3442, RSc0003 and RSc0004 genes led us to localize the second crossover in the middle of the RSc0004 gene, which presented a hybrid sequence originating from both the GMI1000 and NCPPB332-WT. In summary, the CGH results combined with PCR and sequencing analyses revealed transfer of a cluster of 33 GMI1000 genes (RSc3413–0003) into the NCPPB332-*vsr* clone (Table 4; Figure 1). Surprisingly enough, this DNA rearrangement was located in the replication region of the chromosome, which is usually considered as a stable DNA region (Hudson *et al.*, 2002). This cluster included essential genes encoding DNA gyrase, DNA polymerase III, the 50S ribosomal protein, a ribonuclease and the RSc3442 chromosomal replication initiator protein (Figure 1).

Concerning the NCPPB332-*popA* clone, hybridization of genomic DNA on the GMI1000 microarray detected gene acquisitions at four apparently

a Gene description

Gene ID	Description	GMI1000 1	NCPPB332-WT 2	NCPPB332-*vsr* 3
RSc3376	Signal peptide	+	-	-
RSc3377	Hypothetical protein	+	-	-
RSc3378	COG1502 protein	+	-	-
RSc3379 to 3392		+	+	+
RSc3393	ISRso5-Transposase	+	-	-
RSc3394 to 3411		+	+	+
RSc3412	Transmembrane methyl-accepting chemotaxis protein I, *cheD4*	+	+	+
RSc3413	Hypothetical protein	+	-	+
RSc3414	Transmembrane	+	-	+
RSc3415	Hypothetical protein	+	-	+
RSc3416	Transmembrane COG0551	+	-	+
RSc3417	Hypothetical protein	+	-	+
RSc3418	Hypothetical protein	+	-	+
RSc3419	Hypothetical protein	+	-	+
RSc3420	Hypothetical protein	+	-	+
RSc3421	ISRso9-Transposase	+	-	-
RSc3422	Hypothetical protein	+	-	+
RSc3423	Hypothetical protein	+	-	+
RSc3424	Hypothetical protein	+	-	+
RSc3425	Hypothetical protein	+	-	+
RSc3426	Transmembrane	+	-	+
RSc3427	Hypothetical protein	+	-	+
RSc3428	Hypothetical protein	+	-	+
RSc3429	Hypothetical protein	+	-	+
RSc3430	VGR-related protein	+	-	+
RSc3431	Hypothetical protein	+	-	+
RSc3432	Hypothetical protein	+	-	+
RSc3433	Hypothetical protein	+	-	+
RSc3434	Hypothetical protein	+	+	+
RSc3435	Hypothetical protein	+	-	+
RSc3436	Hypothetical protein	+	-	+
RSc3437	Hypothetical protein	+	-	+
RSc3438	Site-specific DNA methylase	+	-	+
RSc3439	**Patch repair protein (DNA mismatch endonuclease), *vsr***	+	-	+
RSc3440	DNA gyrase (subunit B), *gyrB*	+	+	+
RSc3441	DNA polymerase III (beta chain), *dnaN*	+	+	+
RSc3442	Chromosomal replication initiator protein, *dnaA*	+	+	+
RSc0001	50S ribosomal protein L34	+	+	+
RSc0002	Ribonuclease P protein component	+	+	+
RSc0003	Hypothetical protein	+	+	+
RSc0004	Transmembrane COG0706, Preprotein translocase subunit YidC	+	+	+
RSc0005 to 0006		+	+	+
RSc0007	Hypothetical protein	+	-	-
RSc0008 to 0018		+	+	+
RSc0019	Hypothetical protein	+	-	-

oriC

b PCR Analysis

Figure 1 Detection of gene cluster acquisition in a *Ralstonia solanacearum* strain by comparative genomic hybridization on microarrays. This gene acquisition was detected in the NCPPB332-*vsr* recombinant clone obtained after DNA transformation and recombination between the NCPPB332 wild-type (−WT) strain and the GMI1000 genomic DNA, which harbored an insertion of the gentamicin resistance gene within the *vsr* gene (in bold). Positive (+) and negative (−) hybridizations on the microarray are reported for each gene and genomic DNA combination tested. The gene cluster suggested to be acquired in the recombinant clone is outlined. Black and gray boxes indicate mobile elements (IS, transposase) and ACURs (alternative codon usage regions), respectively. The acquisition of this gene cluster into the NCPPB332-*vsr* clone was checked by PCR amplification. Arrows indicate the PCR amplification results for each chosen gene. Lane M, 1-kb marker ladder; C, H_2O control; 1, GMI1000 DNA; 2, NCPPB332-WT DNA; 3, NCPPB332-*vsr* DNA; *oriC*, chromosomal replication origin.

separate positions, one in the chromosome and three in the megaplasmid (Table 4). Only the GMI1000 RSc1755 gene was acquired in the chromosome. Nevertheless, as this gene was a transposase, we suggested that its integration in the NCPPB332-*popA* genome could occur anywhere in the chromo-

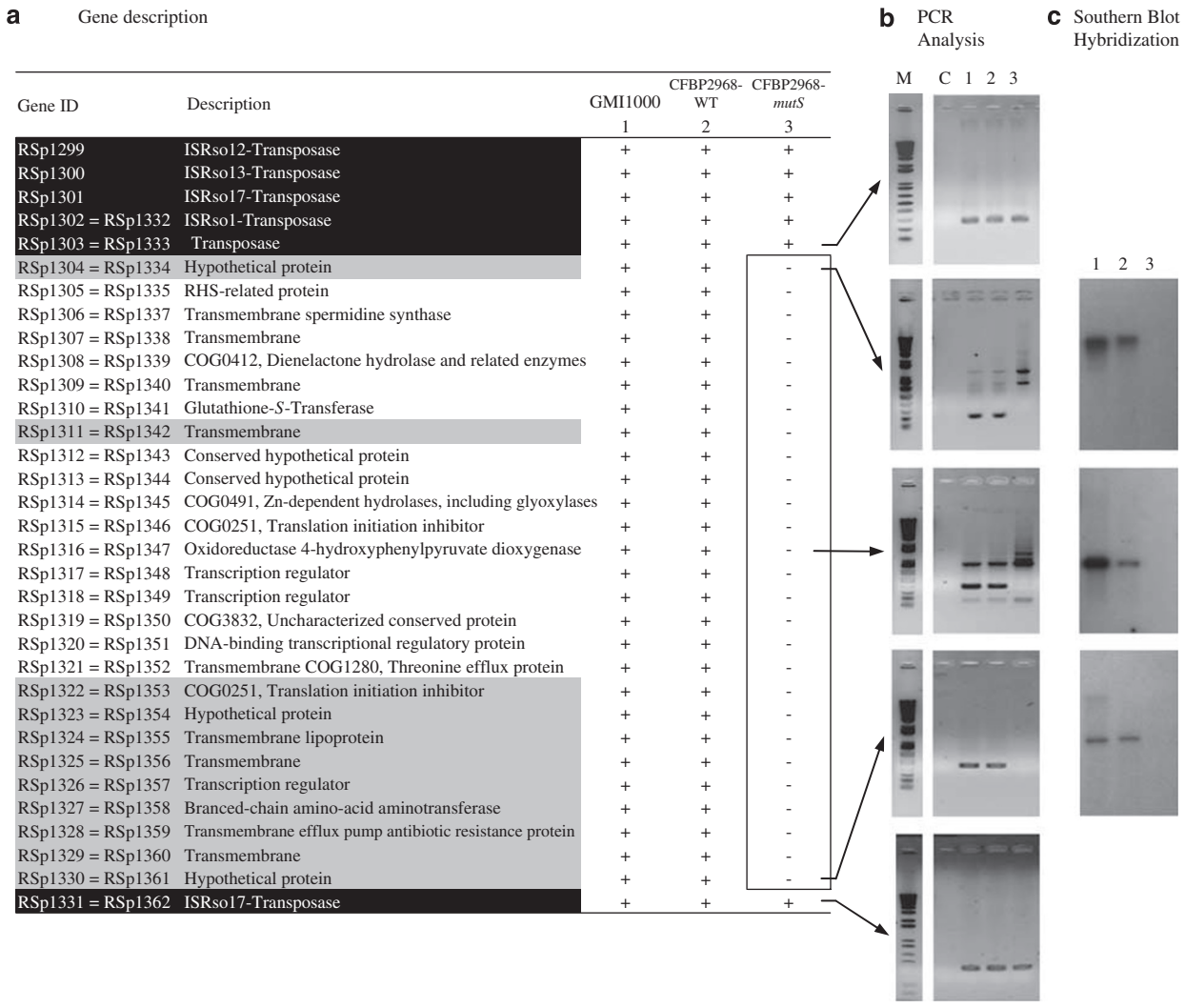
some or in the megaplasmid. This acquisition was confirmed by PCR amplification, cloning and sequencing. Indeed, the RSc1755 PCR fragments amplified from both NCPPB332-WT and -*popA* DNA had 94% and 100% identity with GMI1000 DNA, respectively. In the megaplasmid, the first detected gene acquisition was the *popC* gene from GMI1000 (RSp0875), which was close to the target *popA* gene (RSp0877). As suggested for the NCPPB332-*pilA* recombinant clone, we assumed that a cluster of at least three GMI1000 genes (RSp0875–RSp0877) had been transferred into the NCPPB332-*popA* recombinant genome during the transformation experiment. The second detected gene acquisition in the megaplasmid was the GMI1000 RSp1027 gene, which encodes a transmembrane methyl-accepting chemotaxis transducer. Acquisitions of the *popC* and RSp1027 GMI1000 genes were confirmed by PCR amplification. The third detected gene acquisition in the megaplasmid was a cluster of eight GMI1000 genes (RSp1270–RSp1277), which encodes proteins involved in the virulence of the bacterium (Table 4). Acquisition of this gene cluster was confirmed by PCR amplification of RSp1270, RSp1272 and RSp1277 GMI1000 genes. None of the three genes could be amplified from NCPPB332-WT DNA, whereas positive amplifications of the expected bands were obtained with GMI1000 and NCPPB332-*popA* DNA (data not shown).

Description of gene acquisitions in the CFBP2957 recombinant clones

GMI1000 gene transfers in addition to the marker gene were observed in five of the eight gentamicin-resistant recombinant clones of the CFBP2957 strain (CFBP2957-*recA*, -*comA*, -*ftsK*, -*vsr* and *popA*). Hybridization of their genomic DNA on the GMI1000 microarray revealed that GMI1000 DNA blocks linked to the marker gene ranging from 4.3 to 27.4 kb were transferred (Table 4). These regions encompassed, in addition to the marker gene and depending on the clones, 2–26 GMI1000 genes. Transfer of these GMI1000 DNA blocks into these five CFBP2957 recombinant clones was confirmed by PCR amplification. Insertion of the entire ACUR linked to the *vsr* gene in the GMI1000 genome was detected in the CFBP2957-*vsr* clone. Concerning the other transferred genes, 16 coded for proteins involved in the metabolism of small molecules, 7 for proteins involved in the metabolism of large molecules (synthesis, modification, degradation, energy metabolism; class I and II genes, respectively), 5 for structural elements (cell exterior, cell envelope; class III genes) and 12 for cell processes (cell division, chaperoning, mobility chemotaxis, molecule transport; class IV genes). Half of these transferred genes encoded proteins with unknown functions. Interestingly, 15 of the transferred genes were pathogenicity determinants. All but one were part of a single gene cluster linked to the *popA* gene transferred into the CFBP2957-*popA* clone (Table 4).

Description of gene losses in the CFBP2968-mutS *recombinant clone*

Hybridization of genomic DNA from the CFBP2968-*mutS* clone on the GMI1000 microarray did not reveal any gene acquisition in addition to the *mutS* gene but revealed gene losses at 20 positions in the chromosome and 7 positions in the megaplasmid. However, most of these gene losses could not be confirmed by PCR amplification (data not shown). This suggested that the cutoff value of −2 previously established to define lists of present genes in the tested strains was not strong enough to identify absent genes. This led us to adjust the cutoff value at -3 under which a particular gene could be considered as a lost gene. Using this new cutoff value, the CGH analysis revealed gene losses at one position in the chromosome (RSc0429) and two positions in the megaplasmid (RSp1304–1330 and RSp1334–1361) of the CFBP2968-*mutS* genome. PCR analysis did not confirm loss of the RSc0429 gene but did confirm loss of the two clusters of 27 and 28 genes in the megaplasmid. These clusters corresponded to a tandem repetition of the RSp1304–RSp1330 and RSp1334–RSp1361 genomic regions in the GMI1000 megaplasmid (Salanoubat *et al*., 2002). Even though GMI1000 and CFBP2968 are phylogenetically very close strains, PCR amplification at the junction of the tandem repetition revealed that only one cluster was present in the CFBP2968-WT genome. Indeed, a 3750-bp fragment corresponding to the RSp1330–1334 region could be amplified from GMI1000 DNA but not from CFBP2968-WT DNA (data not shown). Therefore, only one cluster of 29.3 kb harboring 27 genes was lost in the megaplasmid of the CFBP2968-*mutS* clone (Table 4; Figure 2). This deletion was confirmed by PCR amplification of RSp1303, RSp1304, RSp1316, RSp1330 and RSp1331 and by Southern blot hybridization. In the CFBP2968-*mutS* recombinant DNA, positive amplification of the expected fragments was obtained only for the RSp1303 and RSp1331 genes, whereas all tested genes were amplified in GMI1000 and CFBP2968-WT DNA (Figure 2). Southern blot hybridization, using probes derived from the RSp1304, RSp1316 and RSp1330 genes, was performed on DNA from GMI1000-WT, CFBP2968-WT and CFBP2968-*mutS* digested by the *Nco*I restriction enzyme. For wild-type strains, the probes hybridized to fragments of the expected size (20 876, 1784 and 4868 bp, respectively) (Figure 2). For the CFBP2968-*mutS* clone, no hybridization was noted for any probes, thus confirming the microarray and PCR results (Figure 2). This deleted cluster was delimited by the transposases ISRso12, ISRso13, ISRso17 and ISRso1 on one side (RSp1299–1302) and ISRso17 on the other (RSp1331). The other deleted genes coded for transcription regulators, transmembrane and



Gene ID	Description	GMI1000	CFBP2968-WT	CFBP2968-*mutS*
		1	2	3
RSp1299	ISRso12-Transposase	+	+	+
RSp1300	ISRso13-Transposase	+	+	+
RSp1301	ISRso17-Transposase	+	+	+
RSp1302 = RSp1332	ISRso1-Transposase	+	+	+
RSp1303 = RSp1333	Transposase	+	+	+
RSp1304 = RSp1334	Hypothetical protein	+	+	-
RSp1305 = RSp1335	RHS-related protein	+	+	-
RSp1306 = RSp1337	Transmembrane spermidine synthase	+	+	-
RSp1307 = RSp1338	Transmembrane	+	+	-
RSp1308 = RSp1339	COG0412, Dienelactone hydrolase and related enzymes	+	+	-
RSp1309 = RSp1340	Transmembrane	+	+	-
RSp1310 = RSp1341	Glutathione-*S*-Transferase	+	+	-
RSp1311 = RSp1342	Transmembrane	+	+	-
RSp1312 = RSp1343	Conserved hypothetical protein	+	+	-
RSp1313 = RSp1344	Conserved hypothetical protein	+	+	-
RSp1314 = RSp1345	COG0491, Zn-dependent hydrolases, including glyoxylases	+	+	-
RSp1315 = RSp1346	COG0251, Translation initiation inhibitor	+	+	-
RSp1316 = RSp1347	Oxidoreductase 4-hydroxyphenylpyruvate dioxygenase	+	+	-
RSp1317 = RSp1348	Transcription regulator	+	+	-
RSp1318 = RSp1349	Transcription regulator	+	+	-
RSp1319 = RSp1350	COG3832, Uncharacterized conserved protein	+	+	-
RSp1320 = RSp1351	DNA-binding transcriptional regulatory protein	+	+	-
RSp1321 = RSp1352	Transmembrane COG1280, Threonine efflux protein	+	+	-
RSp1322 = RSp1353	COG0251, Translation initiation inhibitor	+	+	-
RSp1323 = RSp1354	Hypothetical protein	+	+	-
RSp1324 = RSp1355	Transmembrane lipoprotein	+	+	-
RSp1325 = RSp1356	Transmembrane	+	+	-
RSp1326 = RSp1357	Transcription regulator	+	+	-
RSp1327 = RSp1358	Branced-chain amino-acid aminotransferase	+	+	-
RSp1328 = RSp1359	Transmembrane efflux pump antibiotic resistance protein	+	+	-
RSp1329 = RSp1360	Transmembrane	+	+	-
RSp1330 = RSp1361	Hypothetical protein	+	+	-
RSp1331 = RSp1362	ISRso17-Transposase	+	+	+

Figure 2 Detection of gene cluster loss in a *Ralstonia solanacearum* strain by comparative genomic hybridization on microarrays. This loss of genes was detected in the CFBP2968-*mutS* recombinant clone obtained after DNA transformation and recombination between the CFBP2968 wild-type (−WT) strain and the GMI1000 genomic DNA, which harbored an insertion of the gentamicin resistance gene within the *mutS* gene. Positive (+) and negative (−) hybridizations on the microarray are reported for each gene and genomic DNA combination tested. The gene cluster suggested to be lost in the recombinant clone is outlined. Black and gray boxes indicate mobile elements (IS, transposase) and ACURs (alternative codon usage regions), respectively. The loss of this gene cluster into the CFBP2968-*mutS* clone was checked by PCR amplification and Southern blot hybridization. Arrows indicate the results for each chosen gene. Lane M, 1-kb marker ladder; C, H_2O control; 1, GMI1000 DNA; 2, CFBP2968-WT DNA; 3, CFBP2968-*mutS* DNA. In CFBP2968-*mutS* clone, positive amplification of the expected fragments was obtained only for the RSp1303 and RSp1331 genes, whereas all tested genes were amplified in GMI1000 and CFBP2968-WT DNA. Despite the high-stringency condition of PCR, 'nonspecific' DNA fragments were amplified with the Rsp1304 and Rsp1316 genes. Southern blot hybridizations were performed using 30-bp probes representative of the RSp1304, RSp1316 and RSp1330 genes on genomic DNA digested by the restriction enzyme *Nco*I.

mobility chemotaxis proteins. Three of them (RSp1317, RSp1323 and RSp1329) belonged to the *hrpB* and *hrpG* pathogenicity regulons (Occhialini *et al.*, 2005; Valls *et al.*, 2006).

Discussion

The use of DNA microarray technology and CGH for detecting HGT

In this study, we investigated HGTs between phylotypes of *R. solanacearum* during natural transformation experiments. We used DNA microarray technology and CGH for detecting these transfers. This technology proved to be a powerful way to detect HGT even if, to our knowledge, it has never been used before for such purpose. Methods described so far to detect HGT events in natural transformation experiments were limited to the analysis of target gene transfers (see for example, Takahata *et al.*, 2007). However, these methodologies do not allow the detection of additional gene transfers or gene losses. Compared with other methodologies, microarray technology provides access to the entire genome of recombinant clones, highlighting then the size and localization of the

DNA blocks acquired or lost in conjunction with transfer of the marker gene. Interestingly, the use of this methodology can detect not only the acquired genes contiguous to the marker genes but also the genes acquired independently from the marker genes in other positions along the genome. In addition, CGH technology on microarrays allowed us to determine the number of genes in each transferred or lost DNA block and the functions encoded by these genes. However, this technology can detect only HGT of new genes or alleles with sufficient nucleotide divergence between the donor DNA and recipient strains. For example, in the NCPPB332-*vsr* clone, PCR amplification and sequencing demonstrated that, in addition to transfer of the gene cluster detected by CGH on microarrays, other genes flanking this cluster were replaced by alleles from the donor strain. In the case of the CFBP2968 strain, no GMI1000 gene integration was detected by CGH on microarrays, probably because this strain is genetically too close to the GMI1000 donor strain. However, it is very likely that DNA exchanges occurred between these strains. In this latter case, in spite of the fact that substitutive recombination might have little effect on the genome plasticity in comparison with acquisition of foreign DNA, the significance of single nucleotide polymorphisms in altering alleles and phenotypes could be increased by the high transformation frequencies among the most closely related strains. In this specific case, an alternative method for detecting allelic exchanges would be to sequence the complete genome of recombinant clones.

Transfer of genes contiguous to the marker gene

Transformation experiments were conducted using total genomic DNA from different mutants of the GMI1000 strain as donor DNA. These mutants varied in the target gene in which the gentamicin resistance gene (marker gene) was integrated. A total of 10 target genes were compared. In this study, we analyzed the number of genes acquired in conjunction with transfer of the marker gene. Interestingly, we detected gene integrations in addition to the marker gene for 8 of the 16 recombinant clones that had integrated the gentamicin resistance gene from GMI1000 mutants. The number of transferred genes linked to the marker gene varied from 2 to 32 depending on the target gene and the recipient strain. These transfers represented DNA blocks ranging from 3.2 to 29.1 kb. Transfers of almost 30-kb DNA blocks from GMI1000 were observed in both phylotype III and II strains. This is in agreement with the results obtained by Coupat *et al.* (2008), which demonstrated that *R. solanacearum* is able to exchange DNA fragments ranging from 30 to 90 kb by homologous recombination. Interestingly in this study, when the *vsr* gene was used as the target gene, the entire ACUR contiguous to the *vsr* gene was transferred to both CFBP2957 and NCPPB332 strains. This supports the hypothesis that ACURs can be transferred between *R. solanacearum* phylotypes. Similar to *Escherichia coli* 'ORFans,' *R. solanacearum* ACURs could be considered as transient genes that have a high turnover rate in the genome (Daubin *et al.*, 2003; Daubin and Ochman, 2004; Gogarten and Townsend, 2005). In future work, it would be interesting to determine the transfer frequencies of ACURs in different *R. solanacearum* strains. More generally, the analysis of the number of transferred contiguous genes in relation to the position of the marker gene along the genome would help to define recombination hotspots in the *R. solanacearum* genome.

HGT between phylotypes of the R. solanacearum *species complex*

Recently, Coupat *et al.* (2008) tested 55 *R. solanacearum* strains selected among the four phylotypes and demonstrated that competence and transformability are ubiquitous physiological traits in the *R. solanacearum* species complex even if it is highly variable among strains. In this study, we were interested in the amount of DNA that can be transferred between phylotypes. Recombinant clones resulting from natural transformation, between donor and recipient strains exhibiting 2%, 19% or 31% gene divergence, confirmed that intraspecific transfers occur between phylotypes of the *R. solanacearum* species complex. As expected, the transformation efficiency dropped with the increase of gene divergence (Table 3) (Fall *et al.*, 2007). These results can be explained by the methyl-mismatch repair system, which acts as the main barrier to recombination between divergent sequences (Mercier *et al.*, 2007). Interestingly, in the *popA* recombinant clones, variable amounts of DNA (4.7–27.3 kb) were transferred to phylotype III and II strains, respectively. Even if it is too early to make any conclusion, this result encourages future works to analyze the impact of genomic divergence between donor and recipient strains on the amount of transferred flanking DNA. The impact of the nucleotide sequence of flanking regions of the marker gene on the recombination sites was reported by Majewski and Cohan (1999), Majewski *et al.* (2000), de Vries and Wackernagel (2002) and Meier and Wackernagel (2005). These authors demonstrated the importance of the presence of short so-called minimum efficiently processed segment sequences or GC-rich microhomologies between the donor and recipient genomes. In *R. solanacearum*, a bioinformatics study also correlated recombination hotspots to the presence of Chi-like signature sequences with which recombination might be preferentially initiated (Fall *et al.*, 2007).

Multiple transformants

Evaluation of the number of DNA regions that can be exchanged by natural transformation is an important

element in assessing the role of HGT in bacterial adaptation and evolution. Very few reports exist on the observation of multiple transformants. Only in the 1970s, multiple transformants, defined as competent cells transformed by two unlinked markers, were reported in *Bacillus subtilis* (Nester and Stocker, 1963). More recently, Coupat *et al.* (2008) also isolated multiple transformants exhibiting resistance to two antibiotics, demonstrating that two independent recombination events can occur by integration of two different DNA fragments in separate regions of the *R. solanacearum* genome. In this study, strength of the approach used is that it allows additional recombination regions to be identified without the use of several selection markers. Then, among the eight recombinant clones that had integrated genes in addition to the marker gene, we could observe that one recombinant, NCPPB332-*popA*, integrated not only genes contiguous to the marker gene but also GMI1000 genes independent from the marker gene at three other positions along the genome (Table 4). PCR verification indicated that this clone was a multiple transformant (data not shown).

Transfer of virulence genes

Interestingly, one of the DNA blocks that had been transferred independently from the marker gene in the NCPPB332-*popA* clone carried virulence genes (a type III effector and the signal peptide RSp1270 belonging to the *hrpB* pathogenicity regulon; Hueck, 1998; Cunnac *et al.*, 2004; Occhialini *et al.*, 2005; Valls *et al.*, 2006). This is in agreement with the classification of virulence genes in the highly transferable gene category (Nakamura *et al.*, 2004; Gogarten and Townsend, 2005). In future work, it would be interesting to determine the natural transformation frequency of virulence genes between *R. solanacearum* phylotypes. Pathogenicity islands can also be transferable by conjugation or transposition. In these cases, the transmissibility of this island has been associated with the presence of conjugative transfer genes or with helper phages (Ruzin *et al.*, 2001) and would be less dependent upon homologous recombination. Horizontal transfers of virulence genes in plant pathogenic bacteria are suggested to be major events in the emergence of new pathotypes (Loria *et al.*, 2006; Wicker *et al.*, 2007). For example, in the plant pathogens *Pseudomonas viridiflava* and *P. syringae*, it has been suggested that HGT and the resulting acquisition and loss of type III effectors constitute an important arms race strategy (Araki *et al.*, 2006; Ma *et al.*, 2006; Sarkar *et al.*, 2006).

Gene losses

To identify gene loss using CGH in the CFBP2968-*mutS* clone, we defined a new cutoff value of −3. Therefore, the loss of a cluster of 29.3 kb harboring 27 genes was detected and confirmed in this clone. This event could have been due to a random event deletion or to the phenotype associated with *mutS* inactivation independently of the transformation process. Indeed, the *mutS* gene is involved in the methyl-mismatch repair system, which acts as the main barrier to recombination between divergent sequences (Rayssiguier *et al.*, 1989). In *R. solanacearum*, experiments recently demonstrated that inactivation of the *mutS* gene leads to a mutator phenotype, increasing the potential for interspecies recombination events and spontaneous mutation frequencies (Mercier *et al.*, 2007). In this study, inactivation of the methyl-mismatch repair system could no longer prevent intra-chromosomal rearrangements between misaligned dispersed repetitive sequences such as IS. Interestingly, the deleted DNA block was surrounded by the ISRso17. This could favor DNA heteroduplex formation and give rise to a stem-loop secondary structure, leading to deletion of the DNA loop (Kirill *et al.*, 1998). The high mutation rate could also explain the false genetic losses using the cutoff value of −2. It seems likely that emergence of additional mismatch in DNA could lead to the non-hybridization of a few genes and thus lead to the detection of a very low number of false genetic losses.

Conclusion

In conclusion, this represents the first comprehensive identification of gene acquisitions and losses, their number, category, function and localization along the genome of the recipient bacterial strain during natural transformation experiments. We demonstrated that the use of microarrays and CGH is an original, accurate and powerful way to detect HGT, especially by providing access to unexpected HGT events. In future works, this strategy should help to answer many questions related to HGT mechanisms such as recombination hotspots or the impact of genomic divergence between donor and recipient strains on the nature and amount of transferred DNA. Here, we showed that DNA blocks up to 30 kb and 33 genes can be transferred between *R. solanacearum* phylotypes and that multiple DNA acquisitions along the genome occurred in a single recombinant strain during a single HGT event. These results support the hypothesis that *R. solanacearum* is capable of rapid adaptation to novel ecological niches.

Acknowledgements

The microarray experiments were conducted in the LIPM (INRA-CNRS) laboratory in Toulouse (France) with Christian Boucher and Stéphane Genin. We acknowledge them for their kind welcome and help. We are grateful to D Bernillon for technical assistance in the Southern blot hybridizations. This study was supported by the Regional Council of La Réunion, the European Community (FEO-

GA) and CIRAD under Research Grant 3P118. Bénédicte Coupat is financially supported by a grant from MENESER.

References

Allen C, Prior P, Hayward AC. (2005). Bacterial Wilt Disease and the *Ralstonia solanacearum* Species complex. APS Press: St Paul, MN, USA.

Araki HD, Goss TEM, Jakob K, Halldorsdottir SS, Kreitman M, Bergelson J. (2006). Presence/absence polymorphism for alternative pathogenicity islands in *Pseudomonas viridiflava*, a pathogen of *Arabidopsis*. *Proc Natl Acad Sci USA* **103**: 5887–5892.

Bertolla F, Frostegård A, Brito B, Nesme X, Simonet P. (1999). During infection of its hosts, the plant pathogen *Ralstonia solanacearum* naturally develops a state of competence and exchanges genetic material. *Mol Plant Microbe Interact* **12**: 467–472.

Bertolla F, Van Gijsegem F, Nesme X, Simonet P. (1997). Conditions for natural transformation of *Ralstonia solanacearum*. *Appl Environ Microbiol* **63**: 4965–4968.

Boucher C, Martinel A, Barberis P, Alloing G, Zischek C. (1985). Virulence genes are carried by a megaplasmid of the plant pathogen *Pseudomonas solanacearum*. *Mol Gen Genet* **205**: 270–275.

Coenye T, Vandamme P. (2005). Displacement of ε-proteobacterial core genes by horizontally transferred homologous genes. *Res Microbiol* **156**: 738–747.

Comas I, Moya A, Azad RK, Lawrence JG, Gonzalez-Candelas F. (2006). The evolutionary origin of Xanthomonadales genomes and the nature of the horizontal gene transfer process. *Mol Biol Evol* **23**: 2049–2057.

Coupat B, Chaumeille-Dole F, Fall S, Prior P, Simonet P, Nesme X *et al.* (2008). Natural transformation in the *Ralstonia solanacearum* species complex: number and size of DNA that can be transferred. *FEMS Microbiol Ecol* **66**: 14–24.

Cunnac S, Occhialini A, Barberis P, Boucher C, Genin S. (2004). Inventory and functional analysis of the large Hrp regulon in *Ralstonia solanacearum*: identification of novel effector proteins translocated to plant host cells through the type III secretion system. *Mol Microbiol* **53**: 115–128.

Daubin V, Lerat E, Perriere G. (2003). The source of laterally transferred genes in bacterial genomes. *Genome Biol* **4**: R57.

Daubin V, Ochman H. (2004). Bacterial genomes as new gene homes: the genealogy of ORFans in *E. coli*. *Genome Res* **14**: 1036–1042.

Davison J. (1999). Genetic exchange between bacteria in the environment. *Plasmid* **42**: 73–91.

de Vries J, Wackernagel W. (2002). Integration of foreign DNA during natural transformation of *Acinetobacter* sp. by homology-facilitated illegitimate recombination. *Proc Natl Acad Sci USA* **99**: 2094–2099.

Fall S, Mercier A, Bertolla F, Calteau A, Gueguen L, Perrière G *et al.* (2007). Horizontal gene transfer regulation in bacteria as a 'spandrel' of DNA repair mechanisms. *PLoS One* **2**: e1055.

Fegan M, Prior P. (2005). How complex is the '*Ralstonia solanacearum* species complex'. In: Allen C, Prior P, Hayward C (ed). *Bacterial Wilt: the Disease and the Ralstonia solanacearum Species Complex*. APS Press: St Paul, MN, USA, pp 449–462.

Gogarten JP, Townsend JP. (2005). Horizontal gene transfer, genome innovation and evolution. *Nat Rev Microbiol* **3**: 679–687.

Guidot A, Prior P, Schoenfeld J, Carrère S, Genin S, Boucher C. (2007). Genomic structure and phylogeny of the plant pathogen *Ralstonia solanacearum* inferred from gene distribution analysis. *J Bacteriol* **189**: 377–387.

Hacker J, Kaper JB. (2000). Pathogenicity islands and the evolution of microbes. *Annu Rev Microbiol* **54**: 641–679.

Hudson RE, Bergthorsson U, Roth JR, Ochman H. (2002). Effect of chromosome location on bacterial mutation rates. *Mol Biol Evol* **19**: 85–92.

Hueck CJ. (1998). Type III protein secretion systems in bacterial pathogens of animals and plants. *Mol Biol Rev* **62**: 379–433.

Kirill SL, Shor BM, Tran HT, Taylor W, Keen JD, Resnick MA *et al.* (1998). Factors affecting inverted repeat stimulation of recombination and deletion in *Saccharomyces cerevisiae*. *Genetics* **148**: 1507–1524.

Kobayashi I. (2001). Behavior of restriction-modification systems as selfish mobile elements and their impact on genome evolution. *Nucleic Acids Res* **29**: 3742–3756.

Koonin EV, Wolf YI. (2008). Genomics of bacteria and archaea: the emerging dynamic view of the prokaryotic world. *Nucleic Acids Res* **36**: 6688–6719.

Loria R, Kers J, Joshi M. (2006). Evolution of plant pathogenicity in *Streptomyces*. *Annu Rev Phytopathol* **44**: 469–487.

Ma W, Dong FFT, Stavrinides J, Guttman DS. (2006). Type III effector diversification via both pathoadaptation and horizontal transfer in response to a coevolutionary arms race. *PLoS Genet* **2**: 2131–2142.

Majewski J, Cohan FM. (1999). DNA sequence similarity requirements for interspecific recombination in *Bacillus*. *Genetics* **153**: 1525–1533.

Majewski J, Zawadzki P, Pickerill P, Cohan FM, Dowson CG. (2000). Barriers to genetic exchange between bacterial species: *Streptococcus pneumoniae* transformation. *J Bacteriol* **182**: 1016–1023.

Matic I. (1995). Interspecies gene exchange in bacteria: the role of SOS and mismatch repair system in evolution of species. *Cell* **80**: 507–515.

Meier P, Wackernagel W. (2005). Impact of *mutS* inactivation on foreign DNA acquisition by natural transformation in *Pseudomonas stutzeri*. *J Bacteriol* **187**: 143–154.

Mercier A, Bertolla F, Passelègue-Robe E, Simonet P. (2007). Natural transformation based foreign DNA acquisition in a *Ralstonia solanacearum mutS* mutant. *Res Microbiol* **158**: 537–544.

Nakamura Y, Itoh T, Matsuda H, Gojobori T. (2004). Biased biological functions of horizontally transferred genes in prokaryotic genomes. *Nat Genet* **36**: 760–766.

Nester EW, Stocker AD. (1963). Biosynthetic latency in early stages of deoxyribonucleic acid transformation in *Bacillus subtilis*. *J Bacteriol* **86**: 785–796.

Ochman H, Lawrence JG, Groisman EA. (2000). Lateral gene transfer and the nature of bacterial innovation. *Nature* **405**: 299–304.

Occhialini A, Cunnac S, Reymond N, Genin S, Boucher C. (2005). Genome-wide analysis of gene expression in *Ralstonia solanacearum* reveals that the *hrpB* gene

acts as a regulatory switch controlling multiple virulence pathways. *Mol Plant Microbe Interact* **18**: 938–949.

Ochman H, Moran NA. (2001). Genes lost and genes found: evolution of bacterial pathogenesis and symbiosis. *Science* **292**: 1096–1099.

Raymond J, Zhaxybayeva O, Gogarten JP, Gerdes SY, Blankenship RE. (2002). Whole-genome analysis of photosynthetic prokaryotes. *Science* **298**: 1616–1620.

Rayssiguier C, Thaler DS, Radman M. (1989). The barrier to recombination between *Escherichia coli* and *Salmonella typhimurium* is disrupted in mismatch-repair mutants. *Nature* **342**: 396–401.

Ruzin A, Lindsay J, Novick RP. (2001). Molecular genetics of SaPI1—a mobile pathogenicity island in *Staphylococcus aureus*. *Mol Microbiol* **41**: 365–377.

Salanoubat M, Genin S, Artiguenave F, Gouzy J, Mangenot S, Arlat M *et al.* (2002). Genome sequence of the plant pathogen *Ralstonia solanacearum*. *Nature* **415**: 497–502.

Sarkar SF, Gordon JS, Martin GB, Guttman DS. (2006). Comparative genomics of host-specific virulence in *Pseudomonas syringae*. *Genetics* **174**: 1041–1056.

Takahata S, Ida T, Senju N, Sanbongi Y, Miyata A, Maebashi K *et al.* (2007). Horizontal gene transfer of *ftsI*, encoding penicillin-binding protein 3, in *Haemophilus influenzae*. *Antimicrob Agents Chemother* **51**: 1589–1595.

Thomas CM, Nielsen KM. (2005). Mechanisms of, and barriers to, horizontal gene transfer between bacteria. *Nat Rev Microbiol* **3**: 711–721.

Valls M, Genin S, Boucher C. (2006). Integrated regulation of the type III secretion system and other virulence determinants in *Ralstonia solanacearum*. *PloS Pathogens* **2**: 798–807.

Wicker E, Grassart L, Coranson-Beaudu R, Mian D, Guilbaud C, Fegan M *et al.* (2007). *Ralstonia solanacearum* strains from Martinique (French West Indies) exhibiting a new pathogenic potential. *Appl Environ Microbiol* **73**: 6790–6801.