

Specific Genes from the Potato Brown Rot Strains of Ralstonia solanacearum and Their Potential Use for Strain Detection

Alice Guidot, Mounira Elbaz, Sebastien Carrere, M.I. Siri, M.J. Pianzzola, Philippe Prior, Christian Boucher

▶ To cite this version:

Alice Guidot, Mounira Elbaz, Sebastien Carrere, M.I. Siri, M.J. Pianzzola, et al.. Specific Genes from the Potato Brown Rot Strains of Ralstonia solanacearum and Their Potential Use for Strain Detection. Phytopathology, 2009, 99 (9), pp.1105-1112. 10.1094/PHYTO-99-9-1105. hal-02668604

HAL Id: hal-02668604 https://hal.science/hal-02668604v1

Submitted on 2 Mar 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.

Specific Genes from the Potato Brown Rot Strains of *Ralstonia solanacearum* and Their Potential Use for Strain Detection

A. Guidot, M. Elbaz, Sébastien Carrère, M. I. Siri, M. J. Pianzzola, P. Prior, and C. Boucher

First and sixth authors: CIRAD, UMR PVBMT, Saint Pierre, La Réunion, F-97410 France; second, third, and seventh authors: CNRS-INRA, UMR LIPM, Castanet Tolosan, F-31326 France; and fourth and fifth authors: Cátedra de Microbiología. Facultad de Química, UDELAR, Gral. Flores 2124. CP 11800, Montevideo, Uruguay. Accepted for publication 29 April 2009.

ABSTRACT

Guidot, A., Elbaz, M., Carrère, S., Siri, M. I., Pianzzola, M. J., Prior, P., and Boucher, C. 2009. Specific genes from the potato brown rot strains of *Ralstonia solanacearum* and their potential use for strain detection. Phytopathology 99:1105-1112.

Ralstonia solanacearum is the agent of bacterial wilt infecting >200 different plant species covering >50 botanical families. The genus *R. solanacearum* can be classified into four phylotypes and each phylotype can be further subdivided into sequevars. The potato brown rot strains of *R. solanacearum* from phylotype IIB, sequevar 1 (IIB1), historically known as race 3, biovar 2 strains, are responsible for important economic losses to the potato industry and threaten ornamental crop production worldwide. Sensitive and specific detection methods are required to control this pathogen. This article provides a list of 70 genes and 15 intergenes specific to the potato brown rot strains of *R. solanacearum*

Ralstonia solanacearum is a gram-negative soilborne bacterial plant pathogen with thousands of distinct strains in a heterogeneous species complex. The bacterium causes bacterial wilt in >200 different plant species covering >50 botanical families, including dicots and monocots, and annual plants as well as trees. Economically important crop hosts include tomato, potato, pepper, tobacco, peanut, ornamentals, banana, plantain, and eucalyptus (21). In European Union (EU) legislation, R. solanacearum was listed as a major quarantine organism to control and eradicate, and the potato brown rot strains of R. solanacearum were placed on the U.S. Bioterrorism Select Agent list. Given the important economic impact of bacterial wilt in addition to the potential threat of this bacterium in several European countries (8,23,26,38) and in the United States (24,33), it is highly desirable to develop methods for specifically detecting strains of R. solanacearum. Such methods, however, need to consider the high phenotypic and genotypic diversity of R. solanacearum strains.

Several studies were conducted to unravel the phenotypic and genotypic diversity within *R. solanacearum*. Strains were first classified into five races according to host range (5) and five biovars according to oxidization of various disaccharides and hexose alcohols (19,22). The race and biovar classifications do not correspond to each other, except that race 3 strains which

Corresponding author: C. Boucher; E-mail address: Christian.Boucher@toulouse.inra.fr

* The *e*-Xtra logo stands for "electronic extra" and indicates that the online version contains a table showing the oligonucleotides representative of the new genes identified in the *R. solanacearum* IPO1609 genome.

from phylotype IIB1. This list was identified by comparative genomic hybridization on microarray and subsequent polymerase chain reaction validation with 14 IIB1 strains against 45 non-IIB1 strains that covered the known genetic diversity in *R. solanacearum*. The microarray used consisted of the previously described microarray representative of the phylotype I strain GMI1000, to which were added 660 70-mer oligo-nucleotides representative of new genomic islands detected in the phylotype IIB1 strain IPO1609. The brown rot strain-specific genes thus identified were organized in nine clusters covering 2 to 29 genes within the IPO1609 genome and 6 genes isolated along the genome. Of these specific genes, 29 were parts of mobile genetic elements. Considering the known instability of the *R. solanacearum* genome, we believe that multiple probes are required to consistently detect all IIB1 strains and we recommend the use of probes which are not part of genetic mobile elements.

represent the potato brown rot strains are generally equivalent to biovar 2 (20). A new classification scheme based on nucleotide sequence analysis of four marker genes was recently proposed (11,12). This classification distinguishes four phylotypes. Phylotype I corresponds to the division 1 of Cook and Sequeira (7) and contains strains belonging to biovars 3, 4, and 5 and to races 1, 4, and 5. Phylotype II corresponds to the division 2 of Cook and Sequeira (7) and contains strains belonging to biovars 1, 2, and 2T and to races 1, 2, and 3. Phylotype III contains strains from Africa and the Indian Ocean belonging to biovars 1, 2, and 2T and no race was attributed to this group of strains. Phylotype IV is reported to be highly heterogeneous; it contains strains from Indonesia, Japan, and Australia belonging to biovars 1, 2, and 2T and to races 1 and 3. Phylotype IV also contains the closely related species R. syzygii and the blood disease bacterium (BDB). Each phylotype can be further subdivided into sequevars based on differences in partial sequence of the endoglucanase gene (egl). This phylotype classification is broadly consistent with the former race or biovar classification (11,12,28) and, in some cases, gives an indication of the geographical origin or pathogenicity of the strains. It is believed that, after the race or biovar classification, the phylotype classification is to become the core organizing principle for assigning a particular strain a phylogenetic position with a predictive value on potential host range (11, 12, 37).

Whole-genome sequencing of a broad host range tomato phylotype I strain (GMI1000) was decisive in unravelling the broad genetic diversity encompassed within *R. solanacearum* (31). A microarray representing the GMI1000 genes was developed (25) and used for comparative genomic hybridizations (CGHs) between 18 strains. Interestingly, this demonstrated that a third of the *R. solanacearum* genome consists of variable genes, probably acquired by horizontal gene transfers (16). The distribution of

doi:10.1094/PHYTO-99-9-1105 © 2009 The American Phytopathological Society

variable genes between strains is related to the phylotype classification (16).

Recent phylogenetic evidence indicated that strains that fit with the definition of the potato brown rot agent were placed into the phylotype IIB sequevar 1 (i.e., the biovar 2 Andean strains of R. solanacearum historically known as race 3, biovar 2). These strains are highly pathogenic to potato and tomato and adapted to highland temperatures. Strains that clustered into three of the four phylotypes of R. solanacearum can wilt potato; however, it is the phylotype IIB sequevar 1 strains (IIB1 strains) that are the most persistent and potentially the most destructive for potato. These IIB1 strains are specifically dangerous because they can cause symptomless latent infection in seed potato tubers (6) but also in geranium (23,24,27,32) and be disseminated worldwide through imported cuttings or nurseries.

Several articles relate to methods for specific detection of potato brown rot strains of R. solanacearum (10,15,36). Serological techniques were developed (15,29); however, cross-reactions with saprophytic bacteria from potato and soil may give false-positive results. Molecular techniques based on polymerase chain reaction (PCR) assays appeared more promising. However, all the PCR methods described so far were based on the amplification of the potato brown rot strain-specific DNA sequence identified by Fegan et al. (10) (patent WO 2004/042016 A2) (36). This sequence encodes a mobile genetic element (i.e., a protein homologous to ORF35 of B3 Mu-like phage present in the Pseudomonas aeruginosa [accession Q7AX27]). However, to ensure specificity and reliability of the detection of a pathogenic microorganism, targeting of mobile genetic elements should be avoided because particular strains of the pathogenic microorganism may lose the element or, conversely, other unrelated microorganisms may harbor the element, thereby yielding false negative and false positive results, respectively.

Recently, DNA microarrays were reported to be powerful tools for identification and detection of microorganisms, given that thousands of probes can be used simultaneously. This technology has already been used for identification and detection of plant-pathogenic microbes (1,13,34), and Aittamaa et al. (2) used it to distinguish several bacterial species pathogenic on potato. *R. solanacearum*, however, was not included in their study. Considering the high genomic plasticity of *R. solanacearum*—specifically, its ability to exchange genetic materials—using only one probe for identification and detection of different pathotypes is not adapted because the targeted probe could potentially move from one pathotype to another. The use of DNA microarrays with thousands of probes should strengthen the reliability of the detection method. For this purpose, a list of specific probes for each *R. solanacearum* pathotype must be established.

Lists of genes specific to two IIB1 strains, UW551 and IVIA 1602, were established by comparative genomic analysis between the completely sequenced genome of phylotype I strain GMI1000 against genome sequences of UW551 and IVIA 1602 (14,35). In total, 402 and 48 genes were detected to be present in strains UW551 and IVIA 1602, respectively, and absent in strain GMI1000. The low number of specific genes obtained for the last strain is explained by the low percentage (6.38%) of genome sequenced (35). However, the specificity of these genes in other IIB1 strains was not validated except for one region detected in strain UW551, where a set of 38 contiguous genes (RRSL02400 to RRSL02437) was present in all 20 IIB1 strains tested and absent in other 36 non-IIB1 strains (14).

In the present work, we established the list of genes that are specific to potato brown rot IIB1 strains of *R. solanacearum*. The list was established by CGH on a microarray representing genes of the reference phylotype I strain GMI1000 and of the phylotype IIB1 strain IPO1609. Subsequently, the specificity of the genes from this list was confirmed by PCR amplification from DNA of 14 IIB1 strains as opposed to no amplification from DNA of 45

non-IIB1 strains. In addition to defining the nucleic acid sequences of all specific gene clusters, we designed primers amplifying fragments of <1 kb from these gene clusters which could be used for specific detection of phylotype IIB1 strains.

MATERIALS AND METHODS

Bacterial strains used in this study. The 14 phylotype IIB sequevar 1 (IIB1) strains and 45 non-IIB1 strains used in this study are listed in Table 1, together with their geographical origin and host of origin.

Whole-genome sequence draft of the IIB1 strain IPO1609. A 20× sequence draft of the genome IPO1609 was established at Genoscope according to the previously described strategy, except that finishing was omitted (31). Assembling of these sequences resulted in 10 supercontigs of 4 to 3,372 kb, with 6 contigs of >10 kb covering >99% of the entire genome sequence. This sequence and its annotation are available from GenBank under accessions CU694431 to CU694431, CU914166, and CU914168 and from our website at http://iant.toulouse.inra.fr/bacteria/ annotation/cgi/ralso.cgi.

Design of oligonucleotides representative of genomic islands specific to strain IPO1609 compared with GMI1000. The predicted proteins from strain IPO1609 were compared with the predicted proteins from the previously sequenced strain GMI1000. In all, 630 proteins which were absent of GMI1000 or did not have a close homolog in this strain (<80% amino acid identity or identity covering <80% of the total length of the two proteins) were identified. A representative 70-mer specific oligonucleotide for each of these 630 proteins was designed using ROSO (30). In addition, 30 oligonucleotides representative of some intergenic regions absent from strain GMI1000 were designed. The list of these oligonucleotides was deposited at the French agency for software protection (Agence de Protection des Progammes) under accession IDDN.FR.001.300024.000.R.P. 2006.000.10300 and is available from the authors upon request. The 660 oligonucleotides thus defined were used to complete the previously designed microarray representative of strain GMI1000 (25) in order to generate the microarray used in the present study.

Genomic DNA extraction, DNA labeling, microarray hybridization, and hybridization signal measurement and analysis. DNA extraction and labeling and microarray hybridization were performed as previously described (16). Standard control DNA used for all genome hybridization experiments consisted of an equimolar combination of the genomic DNA from three strains (GMI1000, IPO1609, and Molk2 which is a phylotype II isolate originating from banana). Analysis was conducted as previously described using ImaGene and GeneSight (BioDiscovery) software (16) except that, in the present analysis, a gene was considered absent in the tested strain when the base 2 logarithm of the ratio of the normalized hybridization signal with the control DNA was lower than the cut-off value of -1 (instead of -2 as described by Guidot et al. [16]).

PCR validation. The list of candidate IIB1-specific genes deduced from CGH experiments was checked by PCR amplification. The PCR primers used (Table 2) were designed to amplify one genomic fragment from each gene. When possible, one of the two primers for each gene was designed within the sequence of the oligonucleotide spotted on the microarray. PCRs were conducted in 25 µl of reaction mixture containing 10 ng of DNA from each tested strain, 25 pmol of each primer (L/R), 1.5 mM MgCl₂, 0.2 mM each of four dNTPs, 0.5 U of Red Gold Star *Taq* DNA polymerase (Eurogentec), and the buffer supplied by the manufacturer. PCR amplifications were performed as follows: an initial denaturation step at 96°C for 5 min; followed by 30 cycles of 94°C for 15 s, 59°C for 30 s, and 72°C for 30 s; with a final extension step of 72°C for 10 min. Negative (PCR reaction mixture without DNA) and positive (IPO1609 DNA) controls were included in each experiment. The multiplex PCR described by Fegan and Prior (11) for the *R. solanacearum* phylotype identification was conducted on each tested DNA as a positive amplification control.

RESULTS

Detection of "candidate IPO1609-specific genes" among a collection of strains representative of the diversity of *R. solanacearum*. CGHs on the GMI1000-IPO1609 microarray designed in the present work were performed to compare the gene

TABLE 1. List of Ralstonia solanacearum strains used in this study^a

content of 11 IIB1 strains and 20 non-IIB1 strains (Table 1). This analysis identified a set of 136 oligonucleotides which were detected as present in at least 10 of the 11 IIB1 strains and absent in at least 19 of the 20 non-IIB1 strains (Supplementary Table 1). These oligonucleotides were representative of 77 genes and 38 intergenic regions from the IPO1609 genome. Therefore, these genes and intergenes were considered as "candidate IIB1-specific genomic regions" (IIB1sgr).

A large proportion of these regions formed clusters in the IPO1609 genome but a few genes mapping within these clusters were not found in the IIB1sgr list. Based on (i) the known mosaic structure of *R. solanacearum* genome, where complete sets of

Strain ID	Host	Origin	Phylotype	Sequevar	Race	Biovar	Genomic analysis
IPO1609	Potato	Netherlands	IIB	1	3	2	CGH and PCR
JT516	Potato	Reunion Is.	IIB	1	3	2	CGH and PCR
CMR34	Tomato	Cameroon	IIB	1	3	2	CGH and PCR
RE	Potato	Uruguay	IIB	1	3	2	CGH and PCR
AP31H	Potato	Uruguay	IIB	1	3	2	CGH
AP42H	Potato	Uruguay	IIB	1	3	2	CGH and PCR
TB1H	Potato	Uruguay	IIB	1	3	2	CGH
TB2H	Potato	Uruguay	IIB	1	3	2	CGH
TC1H	Potato	Uruguay	IIB	1	3	2	CGH and PCR
TB10	Potato	Uruguay	IIB	1	3	2	CGH
ETAC	Potato	Uruguay	IIB	1	3	2	CGH
RM	Potato	Uruguay	IIB	1	3	2	CGH and PCR
PSS525	Potato	Taiwan	IIB	1	3	2	PCR
CMR24	Potato	Cameroon	IIB	1	3	2	PCR
CIP10	Potato	Peru	IIB	2T	3	2T	PCR
NCPPB3987	Potato	Brazil	IIB	2T	3	2T	PCR
Molk2	Banana	Philippines	IIB	3	2	1	CGH
CIP418	Peanut	Indonesia	IIB	3	2	1	CGH
UW9	Heliconia	Costa Rica	IIB	3	2	1	PCR
CFBP1183	Heliconia	Costa Rica	IIB	3	2	1	PCR
UW163	Plantain	Peru	IIB	4	2	1	CGH
Ant75	Heliconia	Martinique	IIB	4	NO	1	CGH
Ant80	Anthurium	Martinique	IIB	4	NO	1	CGH
Ant307	Anthurium	Martinique	IIB	4	NO	1	CGH
IY200	Anthurium	Martinique	IIB	4	NO	1	CGH
IY201	Anthurium	Martinique	IIB	4	NO	1	CGH
Ant1121	Anthurium	Martinique	IIB	4	NO	1	CGH and PCR
CFBP6797	Solanum	Martinique	IIB	4	NO	1	PCR
CFBP7014	Anthurium	Trinidad	IIB	ND	NO	1	PCR
TOM	Tomato	Uruguay	IIA	ND	2	1	CGH and PCR
R34	Banana	Brazil	IIA	24	2	1	CGH and PCR
A 3909	Heliconia	Hawaii	IIA	6	2	1	CGH and PCR
CIP239	Potato	Brazil	IIA	ND	1	1	PCR
CIP301	Potato	Peru	IIA	ND	1	1	PCR
CFBP2957	Tomato	Martinique	IIA	ND	1	1	PCR
CMR39	Tomato	Cameroon	IIA	ND	1	1	PCR
ICMP7963	Potato	Kenva	IIA	7	1	1	PCR
CFRP6942	Solanum	Cameroon	III	, ND	NO	2T	CGH and PCR
CFBP60/1	Tomato	Cameroon	III	ND	NO	21 2T	CGH
CMR43	Potato	Cameroon	III	ND	NO	21 2T	PCR
CIP358	Potato	Cameroon	III	2T	NO	21 2T	PCR
CFRP3050	Faanlant	Burkina Faso	III	ND	NO	1	CGH and PCR
CMR66	Solanum	Cameroon	III	ND	NO	л 2Т	PCR
IT525	Pelargonium	Reunion Is	III	19	NO	1	PCR
IT528	Potato	Reunion Is	III	19	NO	1	PCR
125	Potato	Kenva	III	ND	NO	л 2Т	PCR
NCPPR332	Potato	Zimbabwe	III	ND	NO	1	PCR
GMI1000	Tomato	Guyana	I	ND	1	3	CGH
CMP134	Solanum	Cameroon	I	ND	1	3	CGH
CID365	Dotato	Dhilippines	I	ND	1	3	DCP
R788	Morus alba	China	I		1	5	PCR
N200 DSS259	Tomato	Toiwan	I	ND	1	3	CCH
PSS100	Tomato	Taiwan	I		1	3	CGH
DSS210	Tomato	Taiwall	I		1	3	DCD
F 33219 A CH722	Tomata	Australia	1		I NO	5	PUK DCD
ACH/32 Do:07	Tomata	Australia	1 V		NO	2	PCK
PSIU/	Tomato	Indonesia	1 V		NU	21	PCK
PSI30	Tomato	Indonesia	1 V		NU	21	PCK
MAFF301338	Potato	Japan	1 V		NU	21	PUK
к. syzygu R28	Clove	Indonesia	1V	ND	NO	NO	CGH and PCR

^a CGH = comparative genomic hybridization, PCR = polymerase chain reaction, ND = not determined, and NO = no race assigned.

CCAAGTTCGACTACCT MGATATCCCCTGCAC TTCGAGTACAATGTAGG TGGAGCAAGAACACT TTATTGTGACGATGAGG TGGAGAAAGGATGAG TATTCTGTGGCAGTCATATCAC TGATAGGCATGGAG TGATGGAGGATGAG TGCTACTGGTCAGTCC ^b TATTCTGTGCCCACTACG TAACTGGTGAGGATGAG TACCGCCAGTCATTGACGGATAAT CCTTCACTGGTCAGGATAAT GCTAACGGAGATCAGG AACAGCGGGTCATTGAC TGAAGGGAGATCAGGA AACAGCGGGTCATTGGC TGAAGGGAAACAGA ^b AACCGGTCATTGGC TGAAGGGATCTCGAA AACGCGGTCATTGGC GGTGATACGGATAATCGATA	Magtredacted Matatecettecet Jatatecettecet Jatatecettecet Jatatecettecet Jadaaadent Medaaadent Aathefedaaden Medaadedate Aathefedaaded Medaadedatece Cetteattatea Cetteattatea Cetteattatea Cetteattatea Cetteattatea Medecadeataat Cetteattatea Cetteattatea Medecadeataat Cetteattatea Medecadeataat Cetteattatea Medecadeataatea Cetteattatea Medecadeataatea Getteattatea Get	AGTTCGACTACCT TATTCGAGTACCT TATTCGAGTACAATGTAGG AGCAAAGCATAGAG AGCAAAGCATAGAG AGCAAAGGATA TATTGAAACTCTTCACGAG ATTTCTGTGCCACTACG CCTTGACAACGG TACCGCCAGTCATGAG CCTTCATTATTGAGACGG TACCGCCAGTAATTGAC CCTTCATTGAGATAAT AGGAGGATCAGG ^b ATCGTACAGGATAATTGAG AACGAGGATCAAGGA AACGTACGGAG AACGAGGATCAAGGATAATTGGC AAGGACAAACAGA ^b AACGTACGGGGCCATTGGC AAGGACAAACAGA ^b AACGTACGGGGCCATTGGC AAGGACAAACAGA ^b AACGTACGGGGTCATTGGC AAGGACAAACAGA ^b AACGTACGGGGCCAATTGGC AACGTACCC ^b GGTGATCAATTGGC GGTTCATTCCAATTCGGA GCTTTTTCCGT GGTTGGCTGGTCAATTCGG GGTTATTCGATAATTCGG GGTTATTCGATAATTCGG GGTTATAATTCGGATAATTCGG GGTTTTTCCATTC GTTTTCCATTC AAACCGTCGGTCGATAATTCGG GGTTTTCCATTC AAACGGTCGGGTCAATTCGG	INTEGACTACCT TICGAGTACAATGTAGG INTECCCTGCAC GCAAAGAACACT TATTGAAACTCTTCAGG AAAAGCATAGAG CTTCATTATTGAGACGG CTTCATTGGGAAGGTTATT ACGGCAGGTCATGGAGATAAT ACGGTCAGTCC ^b CCTTCACTTGCAGGAGATAAT ACGGTCAGGAG ACAGGAGGTCATGGAGGGTCATTGGC GAAGAAACAGA AACGCGGGTCATGGAGGG AACGCGGTCATGGAGG AACGCGGTCATGGAGG ACGCTGAAACAGA ACGCTACAGGGTCATTGGC GGTGATCAAGGGTCATTGGC GGTGATAAATCGA GGTGATCAAGGGTCATAGGG GGTGATAAATCGA GGTGCAGGTCAGG GGTGCATACGGATAAATCGG GGTGCCAGGTCAGG GGTGCCAAACGGGTCATAGGG GGTGCAAACGGGTCATAGGGCCGATAA ACCGCTCGATAACGGGTCATAGGG GGTGCAAACGGGTCATAGGG GGTGCAAACGGGTCATAGGG CTTCCATACGGGTCATAGGG CTTCCATAGGGTCATAGGG CTTCCATAGGGTCATAGGG CTTCCATAGGGTCATAGGG CTTCCATACGGGTCATAGGG CTTCCATAGGGTCATAGGG CTTCCATAGGGTCATAGGG CTTCCATAGGGTCATAGGG CTTCCATAGGGTCCGGATAAATCGG GGTTCCAATCGC GGTTCCAACGGTCAGGGTCATAGGGTCGGTCATAGGG CTTCCAATCG CTTCCAATCGGGTCATAGGGTCGGATAAATCGG CTTCCAATCG CTTCCAATCGGGTCATAGGGTCGGTCGGTCATAGGGTCGGGT CTTCCAACCG CTTCCAATCGGGTCATAGGGTCGGGGTCGGTCATAGGGTCGGGT CTTCCAATCGGTCGGGTCGGTCGGTCGGTCGGTCGGTCGG
ICCAAGTTCGACTACCT VAGATATCCCCTGGACA CTGGAGAGAGAGAGGTT CAAGAAAGCATAGAG CTAACAGGGAAAGGTT TGATCAAGGCAAGG	MAGTTCGACTACCT MAGTTCGACTACCT MAGAAAGCATAGAG AGAAAAGCATAGAG AGAAAAGCATAGAG AGAAAAGCATAGAG CTACTGGACAGGCAGGAA INTCAAGGAGATCAAGA ICAAGGAGGATCACGAA ICAAGGAGGATCACGAA ICAAGGAGGATCTCGAA ICAAGGAGGATCTCGAA ICAAGGAGGATCTCGAA ICAAGGAGATCTCGAA ICAAGGAGGATCTCGAA ICAAGGAGGATCTCGAA ICAAGGAGGATCTCGAA ICAAGGAGGATCTCGAA ICAAGGAGGATCTCGAA ICAAGGAGGATCTCGAA ICAAGGAGGATCTCGAA ICAAGGAGGATCTCGAA ICAAGGAGGATCTCGAA ICAAGGAGGATCTCGAA ICAAGGAGATCTCGAA ICAAGGAGGATCTCGAA ICAAGGAGGATCTCGAA ICAAGGAGGATCTCCAAGAACCACAA ICAAGGAGGATCTCCAAGAACCACAA ICAAGGAGGATCTCCAAGAACCACAA ICAAGGAGGATCTCCAAGAACCACAACAACAACCACAA ICAAGGAGGATCTCCAAGACCACAACAACAACAACAACAACAACAACAACAAACAAC	AGCAAGGAACCT TATCCCCTGCAC AGCAAGGAACGACT AGCAAGGCATAGAG CCTTGGAAGGAT TCAAGGCATAGGAA ACTGGTCAGTCC ^b AAGGAGATCAAGA AAGGAATCAGA AAGGAATCAGA AAGGAATCAGA AAGGAATCAGA AACTGTTCC GATATCGAAAC GCTTTCGAAAC GCTTTTCCGTAC GCTTTTCCGTAT CGCCTCGGTAT	INTCGACTACCT INTCCCCTGGAC GCCAAGGAACAGA CTGGGAAGGAAGGAA ACTGGTCAGGCATGGAA ACTGGTCAGGCATGGAA ACTGGTCAGGAA ACTGGTCAGGAA ACTGGAAACAGA AGGACAAACAGA AGGACAAACAGA AGGACAAACAGA GGGCCAGGATCCCAAACAGA GTTTCCAAACAGA GCTTGGAAAC ACTTCCAACAACAGA CCTTTTCCTGT CGCCTCAGTAT GTTTTCCATTC
GCAGAAA GATCAGA TTTTCGAC TGGTGTT ATTAGAC ATGTTTG ATGTTTG AAGAAAT TGAAAT TGAAAT TGAAAT TGAAAT TACAACC CAACCAC	GCAGAAAAA GATCAGATG GATCAGACCA ATGTTTGGACTG ATGTTGTG ATGTTGTG ATGTTGTG ATGACTGG ATGACTGG ATGACTGG AAGGAGG GAAATGCT TACAACCTC CAACCAGAI GAAGCTCGG	GCAGAAAGA GATCAGATGG GATCAGATGG TTTTCGACCAAC ATGTTTGTGC ATGTTTGTGC AAGACCTGA AAGAATGCTC TACAACCTGA TACAACCTGA CAACCGGAC GAAGCTCGGT GAAGCTCGGT GAAGCTCGGT GAAGCTCGGT CCACCAGAAC	GCAGAAAGGA GATCAGATGGA TGGTGTTCTAA ATGTTTGGACCAAG ATGTTTGTGCT AAGGACGAGCT AAAGAATGCTCCA AAAGAAGCTCCA TACAACCTGAA TACAACCTGAA GAAGCTCGGTC GAAGCTCGGTC GAAGCTCGGTC GAAGCTCGGTC GAAGCTCGGTC
CGGAACTAGGACGTCAAGCAGCGCGAAGGAGCCTACATTGTACTCGAAGAACAGGACCAGGACCAGGACCAGGAAGCTATGAGGAGTTGTCGAAGGATTGTGGAGGTTGTCGG GCAGGATAGGATAGGATGCTGGAGGAGCCATCATCACTGATGCAGCAAGCA	GGGAACTAGGCAGTAAGGCAGCCGTGAGGCTACATTTGTACTCGAAGAACAGGGAACTGGG GGAACTAGGCAGTGAGGTTCCTGTCATCTGCGTAATGCACCAACAAAGCTCATGAAGGAACTGGG GCAGGATTACAAGTCACTCACTGCGTAATGCACCAACAAAGCTCATGAAGGAACTGGG GGCGGTTCCTTGCAATGAGGCGAAGCCGTGAAAGCCAACAACACTCAAGGAACTTGGGAACTGGG GGCGGTTCCAATGAAGTCAAGGAGGAAAGCCGAACAACAACACTCAAGGAACTTGGGAAGTGGGACGGGGCAAGGCAAGTCAAGGAAGCTTGAAGGAAG	CGGAACTAGGACATAGGACCCOTCGGAAGCCTACATTTGTACTCGAAGACACGGGACCAGGCAC CGGAACTAGGACCOTGAGGTTCCTGTTCATTCATCGATTTGTACTCGAAGACACTAGGAAGCTTGGGAAGCCTGGGC GAGGATTAGGACCCTTGGAAAGGCCAACCACGACGCACCAACAAGCTAGGAAGGCTTGGAAGC GAGGATTACTTCATCAACCAATTCGCAATTCGCAACTACCAGCCAACGAAGGCTTGGAAGC GAGGATTCATCAACCAATTGTTCGGAAAGGCAGTGGAAGCAATTGCCAACTATGGAATTGGAAGC GAGGATTCTTCATCAACCAATTGTTCGGAAAGGCAGTGGAAGCAATTGGAAAGCATTGGAAAG CGGCTTAACATTCGGAATAGGACGAGTAAGACAGTTCGAAGAAGCATTGAAAGCAATGGAAG CGGCTTAACATTCGGAATAGGACGAGTAAGACAGTTCAACATCAGCAAGGAATGGACATGGAAG CGGCTTAACATTCGGAATAGGACGAGTAAGACAGTTGAAAGCAATTGAAATTCAACAACTAATACAC DAGGCTTTAACATTCGTTGGATAAGACAGTTCAACATCAGCATTGAAATTCAACTTGCAAGGAAGG	CGGAACTAGGCACATIACAAGCACCGCCGTGGGAAGCCCTACATTIGTACTICGAAGAACCGGGACCAGGCGAC GCGCGATTAGAGCCGTGAGGTTCCTGTTCATCTGCGTAATGCAACCATTGCAACCATGGAAGCATTGGAACCTGCGCGGCGACGACCGTGCAACCTCGCTACGGGACGAAGCCAACCACGGGAGCTTCGAACGCCGGGAGCAATTGCAACCATGCAACGTCGAAGGACTTGGAAGCACTGCAACGTCGAATGCAACGACGAAGCAAGTGCAACGACGAAGCAAGC
al protein CGGCTACAGGCAGTCTCAAAGAGACAGGTCTGAAGAGACTATGGCATTCTGGTTATC al protein CGTTGAACATGTGGTAGGCAGGAGGTCTGAAGAGACTATGGCATCTATTCGGTATCT CGTTGAACATGTGGTGGTGGTCAGCAGGAGGTCAGGGCAGTGAAGTGAACTGAACTGAACATGAACTGAACAGAT al protein CCGCGCAATCCCACTAATGTTTGTGCTACTGGTCGGCGATGGAATGTCGACTGAAT cGCGGGAATTCCACTAATGTTTGTGCTACTGGTCGGCGGCGGCGAGGTGAATGGCGCACAACAAT al protein corpresentative oligo on the chip protein) corpresentative oligo on the chip protein corpresentative oligo on the chip corpresentative oligo on the chip protein corpresentative oligo on the chip corpresentative oligo on the chip protein corpresentative oligo on the chip corpresentative oligo on the chip protein corpresentative oligo on the chip corpresentative oligo on the chip protein corpresentative oligo on the chip corpresentative oligo on the chip protein corpresentative oligo on the chip protein corpresentative oligo on the chip corpresentative oligo on the chip protein co	 al protein CGGCTACAAGGCAGTCTCAAAGAGACAGGTCTGAAGAAGACTATGGCATTATGGCTATTG al protein CGTTGAACATGTGGCTGGCCTCCAAGGAGACAGGTCAAGGCATTGAACTTGGCATCATGA al protein CGCTGGTGGTGGCGCCTCCAAGGGGGGGAAAGGCTGGAGTGAAGTCACGGCATCAACTGA al protein CGCGGCGCTGGGCGCTCGAGTGGGTTTGTGCTACGGCGGCAGTGAAGGCTTCACTGGCGCACAAA an protein CGCGGGGATGGAAGGCCGGGGGGGGGGGGGGGGGGGGGG	 al protein cGGCTACAGCATICTCAAGAGACAGTCTGAGTAAGACAGTCATGCCATTATGCCATTATGACTGATTAGACTGATTAGACTGATTAGACTGATTAGACTGATTAGACTGATTAGACTGATTAGACTGATTAGACTGATTAGACTGATTAGACTGATGAGTAAGAGCAGTTAGACACGGCATGATGAGAGAGA	al protein CGCTACAGGCATTCTAAGGGACAGTTGAAGAACTATGCCATTATTCGTATTCAACTTCAACTTCAACTTCAACTGCATCAAGGACATTCAATTCGACTGATTCAACTTCAACTGCACTCAAGGACATTCATT
related protein GGAGTTTGGAGGCTCAGATCGGTGGTGGTGCAGCAGCTGCTGGCTG	Intel protein GGAGGTTTGGAGGCTCAGATCGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTG	atel protein GGAGGTTTGGAGGCTCAGATCGGTGGTTGTCCATGCAGTACCTGCTGTTGCTGCTGGCTG	Interpretein GGAGTTTGGAGGCTCAGATCGGTGGTTTGTCCATGCAGTATCCACCTGCTGTTGCTGCTGGCTG
ne GACCGACTGGAACGAGATGGCCAGGAGCAGGGTGGGCAAGATCGCGAAGATCGGGACTCGAC certext CTTCATGACGGTGTAGCCCTGGACCAGGGTGTCTGATCGCGGACTAAGCACTAGTGGGACTCGAC tweiteal provein GCCGGTGAAGATAGGGTCAGCCTTGACCTTGATCGCGGACTAAGCACACTTGATGCTCT tweiteal provein GCCGGTGAAGATAGGGTCAGCGTCTGAGCAGGGGGCGTGAAATGCTCAAAGACAACAGATCTTCGCGG GCCGGTGAAGATAGGGTCAGCGTCTGAGCAGCGGGGGGGG	c GACCGACTGGAACGAGATGGCCAAGGCTCCGAGCAAGGTCGCGAAAGATCGCGATGGGACTCGAC crTTCATGACGGTTAACCCTGACCCAGGTGTCGACGAAAGATCGCGCGAAAGCAATGATCTTCGCGG GCCGGTGAAGATACGGGTCAAGGCGTCTAGGCGGGCGGGAATGAGCCACAATGATCTTCGCGG tical protein GCCGGTGAAGATCGGGCTCAAGGCGCTCAAGGCGGCACAATGATCTCGCGG GCCGGTGAAGATACGGGTCAGGCGTCTAGGCGGGCGGGGGGGG	Gaccoartisoarcoartesercados antesercados antesercados cal protein Gaccoartisoarcoartesercados antesercados antesercados cal protein Gaccoartisoaren antesercados antesercados antesercados cal protein Gaccoartisoaren antesercados antesercados antesercados antesercados cal protein Gaccoartisoaren antesercados antesercados antesercados antesercados cal protein Gaccoartisoaren antesercados antesetados antesercados antesercados antesercados antese	gaccacrosocoactoradocorcoadecorcoadecorcoaderadocoacrosocoactadocoacrosocoactadocoacrosocoactadocoacrosocoactadocoacrosocoactadocoacrosocoactadocoacrosocoactadocoacrosocoactad
etical protein GGAGAACCTGGACAACCAGCATCTACAAGATCCTAGTCAAGAGGCGAAGGAGAGGGGGG CA netical protein GGCAGATGAATCTTCCTTTGAATTTCCGGTCCGGCAGTCAAGAGGGGAGGC AA netical protein GGCAGATGGACTGGAAGTCTTCCTTGCGGCGGAGGAGGC AA binase GTGCCACTAGCTATTCGTGCGGGGGGGGGGGGGGGGGGG	ical protein GGAGAACCTCGACAACAAGATCCTAGTCAAGATCCTAGTCAAGAGGCGAAGGAGGGGGG CA ical protein GGCAGATGGAATCTTCCTTTGAATTTCACGTCCAGGAAGGA	cal protein GGAGACCTGACAACCAGACATTACAAGATCTTAGTCAATGTCCATCAAGAGGGGAGGGGGGGG	call protein GGAGAACCTCGACAACCAACCAACTACAAGATCCTAGTCAAGTGCATCAAGAGGCGAAGGAGGGGGG CA ical protein GGCAATGGACTCGACAACCAACCAACTACAAGATCCTAGTCGACTCAAGAGGCGAAGGAGGGGGG AA mase GGCAATGGACTCGACAACCAACCAGCGGGGAGAGATTTCAGGCGGAGGGGGGGG
	e CTGTTTACCAACCACACACACACACGGCTGCTGACGGATCAGACTTGGACGAGTCGTG tical protein no representative oligo on the chip GGTGGTCGAGGTGGAAGTGTCTACGACAGCTGTGAAGCCGTTGAGAATCGAATATTGGCAATCAGG GTAT tisal protein GGTGGTCGAGGTGGAAGTGTCTACGACAACCTGTGAAGCCGTTGAGAATCGAATATTGGCAATCAGG GTAT nine-specific no representative oligo on the chip GTCGTTGAGAAGTCTACGACAAGCCGTGGAAGCCGTTGAGGAATCGAATCGGCAATCAGG GTAT nine-specific no representative oligo on the chip GTCATTCGTGAAGGCTTCTGATTGCTTACAACAGGCTACTTCATGGTGGCGGCGGCGGCGGCGCGCGC	clicitity CTGTTTACCAACCATAGCCTCACACGGGGGGGGGGGGGG	CTGTTTACCAACCACATAGCCTCACACGGGGGTGGGGATCAGACTTGGACGACTACTTGAGTCGTG ical protein no representative oligo on the chip GGTGGTGAGAGGTGAGAAGTCTCACGACAACCTGTGAAAGCCGTTGAGAATCGAATCAGG GTAT ical protein GGTGGTGAGAAGTCTCACGACAACCTGTGAAAGCCGTTGAGAATCGAATCAGG GTAT ine-specific no representative oligo on the chip GTATCGTGAGAAGTCTCACGACAACCTGTGAAAGCCGTTGAGAATCGAATCAGGCAATCAGG GTAT tethylase CGTCATTCGTGAGAACCTGATCAAAGGCTTCTGATGGCGGGAAACGGTGGTGGCGCGCAATCAGGCGGCAATCAGGCGGCAATCAGGCGCGGAAACGCTAGGGGAAACGGTGGTGGGCGGGAAACGCTAGGGGAAACGGTGGTGGGCGGGGAAACGGTGGAAGCGTGGAAGGCTGGAGGGCTGGGGGGGG

7	IPO_02148	Bacteriophage tail sheath	TTCCTGCTACATCAAGCGGCTGATTTCCATGTACCTGCATCGCTCCGATGGCGGCGGGGGGGG	CTGATTTCCATGTACCTGCATC ^b	ACATCAGCACGTTCTTGTAGAGC	235
7	IPO_02149	Hypothetical protein	CGTCGGTGGCCAAGTCCTGGAGTACACCGACGGCGATTTTGTACTGGCGCGCGATCGCTACGAGTTCGACGCC	GACGATGAGTTGCTGAAGTACC	TGAAGGTAATGGTCACAGCTTTT	237
7	IPO_02150	Hypothetical protein	CCGCATCCGGATGATCCGGTTCCTTCTGATGATCTGGCCATCATCGAAAGCGCACCGCGCCGCCGCTTTGGCCG	ATCCGGTTCCTTCTGATGATCT ^b	CTTCAACTTCACGTGTTCAATCA	124
	IPO_02151 IPO_02152	Hypothetical protein Mu phage prophage-	CGGAAGACIGGAAGATCGCCAAGCIAGICACCAGIACCGGCAAATCCAACACCIACGAATGGCIGAG GGTGAAGTGCTCGAAAATCCTTTCGGTCGCCACCAACACCCGCCCTGGATGGGCTAGAAGCCGTGG	GCTAGTCCACGCGCGACAAAATCAT AGGTGAAGTGCTCGAAATCCT ^b	GTGACTAGCTTGGGGGATCTTCT* CTTCTTCCTTCTCGATGCCTTC	163 284
t		related protein				000
- 1	CC120_041	Transmomhene	CUCCUAU IUUU IUUUU IUIAUUUUAI UAI UUUUUUAI IUAUUAU IUUAU IUUUU IUUUU PATETEVETEVETEVETEVEVEVEVEVEVEVEVEVEVEVEV	CATTATOTOCITICICIOUCAAIA	ACAACUICCAUCUICAIC ACAECCAAACTEACAAAEATCE ^b	877
	IPO_02154/02155	Intergene	CGTGGTAGACCCGGCTGCCGGTGAGATGTGTGCTGGAGATGTGGCCGGGGGGGG	CALIFACCICICACITIC	ACAUCCANACICANONICO	C/1
7	IPO_02155	Hypothetical protein	GCGCGTCAAGTGGGCCGTCACGTGCCGGACTTGGACAGCTTGGCCGTAGCTGAATGCCAAACGATTGCCG	GCATGAAAATGTCTACGTTCCTC	ATGGTGATAGTGCGGAATGAC	491
7	IPO_02156	Transmembrane	GCAGACATGTTCTGCGAAGGATGGGTGTTTGCCGTGCGGGGGGGG	CAGACATGTTCTGCGAAGGAT ^b	GAGGAACGTAGACATTTTCATGC	283
7	IPO_02157	DNA-binding protein HU-beta (NS1) (HU-1)	GACCGAAACCATTCTTGACACCGTGCGCGGCGGTAACGAGCTGACGATCACCGGGCAAGTTCGGC	TGTAATGACCAAGCAAGAACTCA	CACGGTGTCAAGAATGGTTTC ^b	112
7	IPO_02158	Hypothetical protein	TCACCTTGAGCAGGTTCGACGGTCGCAGGTAGTTCGTGTCCCAGGCCGAGAACATCGCTTTTCGATGA	ATTICAAACGCCAAGCCITTAAC	GTTCATCGAAAGCGATGTTCTC ^b	165
7	IPO_02159	Hypothetical protein	CCGCTGCCTGAAGTCGTCTGCCCGAACTGCCGGGGTGGGGGGGG	GCCTATTTGGACCGAAGAAGAC	ACTTCTTGAGGCATCTCGGTTT	394
7	IPO_02159/02160	Intergene	GAAACAGCTAGCAGGCATCACCCCAGGGACCTGTTGCAGCCGAACTGCTGGCGGGGGGGCGTGGCGCCGAATCTGC			
L	09160 Odl	I amhda rennesor-like	CULTULTIGATICUCATUCUCAAUGGUCAAUGUCUCAAUGUCUGUCUGUTTUGUCUGUAUGUAGU CTTTGGAACACGACACTTGGTTCATGGATGAGGGCGGAAAGGCTCAAGGUCGACAAGGTTCAAGATCAAG	GCAGCATTGATGACA AGTTCC	AGTATEGETA & AGGCGTTGC AC	330
		DNA-binding				
r 1	IPO_02161°	Hypothetical protein	no representative oligo on the chip			012
-	IPO_02162	MuA-transposase/ repressor protein	COCTOTACATCOCCOCOCOMANTGOACOOCTICOCOOLAGOACOACTACCAGTACCAG	AAGGCTAAGGGGGGGGGGGAGTAAGTCAT	AGGIAGTIGCGIACTIGGICGIA	548
		CI-related protein				
r 1	IPO_02163	Hypothetical protein	CAAGCAAGCAAGATATCGACCTGGGCATCACGGAAGGCGCTCTAGTTCACGCCACGCAGACGGACG	AGCAACCGAAGATATCGACCT ^b	GCCCAAGCGAAATCAACTCTT	284
-	+0170 ⁻ 041	Friage related DNA- binding protein	no representative ongo on the chip			
7	IPO_02165	Helix-turn-helix motif	GCCGACACGACGAAGCAACACACCACCCCATACGAAACCGACCG	CTITCCCAGTCAATACATTCCAG	CGTTATCACATAAGCCACATCAA	220
r 1	IPO_02166	Lipocalin	GGCGATCTTGGAATGTCGGGCGAAGCGGCGAACAACACGCGCTGTGAGCGCATGAACAAGCGTTTC	AAACAACACGCTGTTGAGCAT ^b	CTTCTTTGCGCACCAATAATC	195
	IPO_02168 ^c	Lytic transplycosylase.	IN TEPRESENTATIVE ONLO ON UNE CHIP GGACTGGGGCCACTGGGCAACAGGGAAGCGGGCCACGGTGGGGGCCTGGGGCTGACGGCGCAACCATCGATAGCG			
		catalytic				
٢	IPO_02168/02169	Intergene	GTCGCTAGCTCGGATCAACCACGCCAAAGGCCAATACGAAATCTGCCGGGGTCGGACTCGGCGCACTGATCG GGAAGGTCTATCACGATCATCTCAGTGACCCTAACAAGAGGCCACCAGCGAACCTGGAACTTGTTGCTGC			
0	1DO 03300	Hunothation] motain	OCTOR A REAL ACTOR A CONTRACTOR OCTOR OCTOR ACTOR AC	CETETATATCECCAETCAAEA	CTTO A GA A A TO A A GOTTO DA A TO	LCC
0 00	IPO_03303 ^c	Hypothetical protein	bacconconconconcontraction of the chip			177
8	IPO_03304°	Transmembrane protein	no representative oligo on the chip			
8	IPO_03304/03305	Intergene	CGTTG6GAAGACAAAGAAAGAACGTCTTTCTCTACGCATCAGGATTTCTTGCCGCAATGGATTTCCTCAGGC			
~	IPO 03305°	Hypothetical protein	Decondent to the chip of the chip of the chip of the control of the chip of th			
8	IPO_03306	Hypothetical protein	CAAGGCAATTTGATTGTGGAGGATCGTATTCAGGTTCACATGAGCTTGCGACTTCTGCATATTCCAACGG	GGATGACTTGTGTTAGCGACTCT	GAATACGATCCTCCACAATCAAA ^b	171
~	IPO_03306/03307	Intergene	GTCG6GTTGTTGTACCTCTACGGCGAGCAAACCGGTCATCCATCAGCGGCTGGTGGCGCTCTACTGCTGT			
6	IPO_04521	Hypothetical protein	CGGCGTTCTACCAGCCAAGGAAGTGCTTTCCACTGCGAGATCTGTAGGCACATGGACATGGAGACACAAG	TGAAGCACTGTCTATCAACCAGA	TITGTCCTAGTCACAGCACTGAA	222
6	IPO_04522°	Hypothetical protein	no representative oligo on the chip			
6	IPO_04523	Hypothetical protein	CGGCAGCACTCATGCCAGCGGTTCATTGCCTCTCTACGTCCCAAGCTCAACTTCATTGTCACTGATACG	GACTTCGGCTATCTGGAGAAAAT	TCTTAGCAGGTTTAGGCTGAGTG	217
6 0	IPO_04524 IPO_04525	Hypothetical protein Helicase-like motein	CGGCGTTCGTGGGGTACGTGCGTCACTGGTGGGCAGCAATTCGGCCTGGTCACTTCTTCATCCTCCAC GTGGTGAGATTCCGGTCATCAAGAATGGCGAGATCAAGCAGGAAGGTGGACAAAACGCTTTGTTCGCATGGCCG	CATCTTCAAGGATGACTCTCTGG ACTCAGTGACGAAGAGGTTGAAG	GAAGAGTGACCAGGCTGAATTT [®] ACGAGTAGCTTCAATGGTGTCTT	235 250
6	IPO_04526	Hypothetical protein	CCACGAACTGGATGTGGTGGTTGGGCAGTGCCACCAGATTAAGCCGTTCTACAGCGATACCGAACGCTTG	ACACTATGCCTGCTGACTTGAA	AGGGACTTCAACAATGTTAGGC	169
IIB1-sp	secific individual genes					
10	IPO_01058	Hypothetical protein	ATCATGGCTGAGATCCAGCGTATTGCCGCAACCGAGCTCATCGTTATCGACATGGGCGACAGGCTTGGCA	CGAGCTCATCGTTATCGACAT ^b	AAGCTCTTGGACTAGGACGATCT	140
1 2	IPO_01362 IPO_02090	Hypothetical protein Hynothetical protein	CAATTGGGTATACGTGATCTGTGGGGTGCCTGATGGGGATTCGGCTGGGCTGGGATTGCCATGCTCTGGCTC GACGTCAAGCTGATCGGGCAGGTATGCCCAGTGTCGTTGAACATCCTTAATCAAGGGGGAGGGGGGGG	AATTGGGTATACGTGATCTGTGG CAATAGAAATTGCCGAGGTGATA	TCGGGTAAGACGAAGCTGACTA CCTTGATAAGGATGTTCAACGAC ^b	280
13	IPO_03123	Transposase	CGCGTGACCAAACGCCCAGCGCAGGAGATGGAGGAGGAGGTCTTTG	GAACGGAGCCATAGTGATGAAG	AGAGTCGCTAACACAAGTCATCC	946
14	IPO_03132 IPO_04004	Helicase,6 related protein Hynothetical protein	CAAACATTGCCTTCACGGAAGATGAGCAGCATCTGTACAACCTGAGGCTTGATGAGCTGGGCGACGAGGAGG GACCAAATACCGCAAGGACGACCAAGATGAACAACAAGCTGGCCTGGAGGGCGACGACGAGG	AGGGAATCAAATCGCTCATCT CTTGACGTCTGACAACCAAGTAG	AGAAGAAGCCCATGATGACAGAG ATAAGATAAAGGTCGGCCTTC	226 347
	LOOLO-O H	moord monomodify				410
^a Dark	and light gray boxes inc	dicate mobile genetic element	s and ACURs, respectively.			

Data and ugur gray rows more mean generations and receive sequences. Repeated on the microarray. ^b Primer sequences correspond to part of the oligonucleotide spotted on the microarray. ^c Added genes correspond to the genes that were included into the list based on their location within clusters of IIB1-specific genes.

genes could have been acquired through horizontal gene transfers or lost "en bloc" through deletions (16,31); (ii) the consideration that some genes might be missing from this IIB1sgr list due to lack of hybridization data; and (iii) taking into account the possibility that certain genes that score next to the cut-off value were miscategorized, we included into the IIB1sgr list all the sets of one or two contiguous genes that were not originally detected as being IIB1 specific but that are located within a IIB1-specific gene cluster. Based on these criteria, 15 additional genes (identified with a "c" in Table 2) were included in the list of IIB1specific genes.

When compared with the sequence of the predicted proteins from strains GMI1000 using BlastP, five of these genes were found to have a counterpart in this strain and, therefore, were eliminated from the final list of IIB1-specific genes presented in Table 2.

Validation of candidate IIB1sgr. The list of candidate IIB1sgr detected by CGH was established based on the analysis of a limited number of strains. Therefore, we validated this list on a larger collection of strains representative of the diversity found in *R. solanacearum*.

In total, 8 IIB1-strains (6 of which were also used for CGH experiments) and 32 non-IIB1 strains (7 of which were also used for CGH experiments) were used for this validation (Table 1). This validation was conducted by PCR amplification of one genomic fragment from each candidate IIB1-specific gene. All tested DNA could be amplified using the multiplex PCR for the *R. solanacearum* phylotype identification as described by Fegan and Prior (11), thereby confirming that the tested strains actually correspond to *R. solanacearum* isolates (Fig. 1).

Each primer pair that gave a positive amplification from non-IIB1 strains or a negative amplification from IIB1 strains was excluded from the list of IIB1sgr. The results validated the specificity of nine gene clusters and six individual genes (Table 2). Among these, 29 genes were predicted to be parts of mobile genetic elements (bacteriophage or insertion sequences) and 34 were classified as alternate codon usage regions (ACURs) according to Salanoubat et al. (31).

DISCUSSION

In the present work, we identified a set of 70 genes and 15 intergenes that were specific to potato brown rot strains from R. solanacearum phylotype IIB1 historically known as race 3, biovar 2 strains. Within the IPO1609 genome, these brown rot strainspecific genes are organized in nine clusters encompassing 2 to 29 genes and in 6 single genes. Among the 70 specific genes, 29 (41%) are parts of mobile genetic elements such as transposases, recombinases, or phage proteins and 34 (49%) were classified as ACURs. This suggested that most of these specific genes originate from acquisition of foreign genes through lateral gene transfers. We hypothesized that the occurrence of these genomic islands within a R. solanacearum strain is related to the ability to cause potato brown rot disease and that they could be considered as pathogenicity islands (18). However, the functions encoded by these specific genes could not be related to known pathogenic determinants. Among these brown rot strain-specific genes, 37 (53%) encoded proteins with unknown functions, 6 encoded proteins involved in DNA synthesis and repair (helicase, thymidylate synthase, N-6 adenine-specific DNA methylase, DNAbinding protein, and part-B like nuclease), 2 encoded transmembrane proteins, 1 encoded a tetratricopeptide repeat (TPR)domain protein which is involved in protein-protein interactions, and 1 encoded a lipocalin described to play an important role in membrane biogenesis and repair and to be implicated in the dissemination of antibiotic resistance genes and in the activation of immunity (4).

The DNA sequences identified in this study to be specific for potato brown rot strains of *R. solanacearum* are potentially useful for the development of diagnostic assays. We designed primers



Fig. 1. Validation of the specificity of the IPO_00043, IPO_02090, and IPO_02155 genes for IIB1 strains of *Ralstonia solanacearum*. This validation was conducted by polymerase chain reaction (PCR) amplification of one genomic fragment from these three genes using the PCR primers given in table 2 and genomic DNA from 8 IIB1-strains, 32 non-IIB1 strains, the IPO1609 strain as a positive control, and H_2O as a negative control. The ID and phylotype of the strains are given for each lane. **A**, Multiplex PCR as described by Fegan and Prior (11) for the *R. solanacearum* phylotype identification. **B**, PCR amplification of a fragment of the IPO_00043 gene. **C**, PCR amplification of a fragment of the IPO_02090 gene. **D**, PCR amplification of a fragment of the IPO_02155 gene. These three last fragments could be amplified from IIB1 genomic DNA only. M = molecular weight marker.

from these sequences which amplified a PCR fragment from all 14 IIB1 strains tested in the present work whereas they did not amplify any fragment from the 45 non-IIB1 strains tested. Interestingly, the potato brown rot strain-specific primers 630 and 631 developed by Fegan et al. (10) amplified a sequence of 357 bp between the IPO_02173 and IPO_02174 genes from the IPO1609 genome. This sequence is located not far from the extremity of the potato brown rot strain-specific 29-gene cluster number 7 identified in the present work. This sequence was also detected in the sequenced potato brown rot strain UW551 of R. solanacearum (14). The authors found that the sequence amplified by the 630 and 631 primers corresponded to the extremity of the potato brown rot strain-specific 38-gene cluster RRSL02400-RRSL02437 of UW551. However, most genes of this specific cluster encode various bacteriophage proteins which, therefore, have the potential to move from one bacterium to another. For this reason, detection methods based on this gene cluster may not be reliable in the long term. In the present work, 41% of the potato brown rot strain-specific genes were parts of mobile genetic elements. Some articles relate the utilization of mobile elements such as insertion sequences (IS) for specific detection of bacterial pathogens (2). However, it is now well established that such elements can be deleted from the genome or can be transmitted from one strain to another by plasmid conjugation, DNA transformation, or transduction by bacteriophages (17,18). R. solanacearum is known to have a genome capable of rapid evolution because of its natural competence for transformation and the occurrence of horizontal gene transfers between strains (including the transfer of mobile genetic elements) (3,9,31,37). Thus, using a mobile element as marker for detection of particular R. solanacearum strains would decrease the reliability of the detection method by giving rise to false-positive or false-negative results. In the present work, we identified 85 specific sequences for the potato brown rot IIB1 strains of R. solanacearum, 52 of which were not predicted to be part of mobile genetic elements and 10 of which were also not predicted to be part of ACURs. Detection methods based on one of these last sequences or, preferably, on a combination of them should allow identification of IIB1 strains with a high level of specificity and reliability.

ACKNOWLEDGMENTS

This work was supported by the Regional Council of La Réunion, the European Community (FEOGA), CIRAD under Research Grant 3P118 and INRA Division of Plant Health and Environment. We thank P. Winker, C. Dossat, and B. Ségurens from Genoscope, who generated the draft sequence of strain IPO1609; and E. Wicker for critical reading of the manuscript. A. Guidot and M. Elbaz contributed equally to this work.

LITERATURE CITED

- 1. Afshar, B., Pitcher, D., Nicholas, R. A. J., and Miles, R. J. 2008. An evaluation of PCR methods to detect strains of *Mycoplasma fermentans*. Biologicals 36:117-121.
- Aittamaa, M., Somervuo, P., Pirhomen, M., Mattinen, L., Nissinen, R., Auvinen, P., and Valkonen, J. P. T. 2008. Distinguishing bacterial pathogens of potato using a genome-wide microarray approach. Mol. Plant Pathol. 9:705-717.
- Bertolla, F., Frostegård, A., Brito, B., Nesme, X., and 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.
- Bishop, R. E. 2000. The bacterial lipocalins. Biochim. Biophys. Acta (BBA) Prot. Struct. Mol. Enzymol. 1482:73-83.
- Buddenhagen, I., and Kelman, A. 1964. Biological and physiological aspects of bacterial wilt caused by *Pseudomonas solanacearum*. Annu. Rev. Phytopathol. 2:203-230.
- Ciampi, L., Sequeira, L., and French, E. R. 1980. Latent infection of potato tubers by *Pseudomonas solanacearum*. Am. Potato J. 57:377-386.
- Cook, D., and Sequeira, L. 1994. Strain differentiation of *Pseudomonas* solanacearum by molecular genetic methods. Pages 77-94 in: Bacterial Wilt—The Disease and Its Causative Agent, *Pseudomonas solanacearum*.

A. C. Hayward, and G. L. Hartman, eds. CAB International, Wallingford, UK.

- Elphinstone, J. G., Stanford, H. M., and Stead, D. E. 1998. Survival and transmission of *Ralstonia solanacearum* in aquatic plants of *Solanum dulcamara* and associated surface water in England. Bull. OEPP 28:93-94.
- Fall, S., Mercier, A., Bertolla, F., Calteau, A., Gueguen, L., Perrière, G., Vogel, T. M., and Simonet, P. 2007. Horizontal gene transfer regulation in bacteria as a "Spandrel" of DNA repair mechanisms. PLOS One 2:e1055. doi:10.1371/journal.pone.0001055.
- Fegan, M., Holoway, G., Hayward, A. C., and Timmis, J. 1998. Development of a diagnostic test based on the polymerase chain reaction (PCR) to identify strains of *R. solanacearum* exhibiting the biovar 2 genotype. Pages 35-43 in: In: Bacterial Wilt Disease. Molecular and Ecological Aspects. P. Prior, C. Allen, and J. Elphinstone, eds. Springer-Verlag, Berlin.
- Fegan, M., and Prior, P. 2005. How complex is the "Ralstonia solanacearum species complex"? Pages 449-462 in: Bacterial Wilt: the Disease and the Ralstonia solanacearum Species Complex. C. Allen, P. Prior, and C. Hayward, eds. American Phytopathological Society Press, St. Paul, MN.
- Fegan, M., and Prior, P. 2006. Diverse members of the *Ralstonia* solanacearum species complex cause bacterial wilts of banana. Australas. Plant Pathol. 35:93-101.
- Fessehaie, A., De Boer, S. H., and Lévesque, C. A. 2003. An oligonucleotide array for the identification and differentiation of bacteria pathogenic on potato. Phytopathology 93:262-269.
- 14. Gabriel, D. W., Allen, C., Schell, M., Denny, T. P., Greenberg, J. T., Duan, Y. P., Flores-Cruz, Z., Huang, Q., Clifford, J. M., Presting, G., González, E. T., Reddy, J., Elphinstone, J., Swanson, J., Yao, J., Mulholland, V., Liu, L., Farmerie, W., Patnaikuni, M., Balogh, B., Norman, D., Alvarez, A., Castillo, J. A., Jones, J., Saddler, G., Walunas, T., Zhukov, A., and Mikhailova, N. 2006. Identification of open reading frames unique to a select agent: *Ralstonia solanacearum* race 3 biovar 2. Mol. Plant-Microbe Interact. 19:69-79.
- Griep, R. A., van Twisk, C., van Beckhoven, J. R. C. M., van der Wolf, J. M., and Schots, A. 1998. Development of specific recombinant monoclonal antibodies against the lipopolysaccharide of *Ralstonia solanacearum* race 3. Phytopathology 88:795-803.
- Guidot, A., Prior, P., Schoenfeld, J., Carrere, S., Genin, S., and Boucher, C. 2007. Genomic structure and phylogeny of the plant pathogen *Ralstonia solanacearum* inferred from gene distribution analysis. J. Bacteriol. 189:377-387.
- Haagensen, J. A. J., Hansen, S. K., Johansen, T., and Molin, S. 2002. In situ detection of horizontal transfer of mobile elements. FEMS Microbiol. Ecol. 42:261-268.
- Hacker, J., Hochhut, B., Middendorf, B., Schneider, G., Buchrieser, C., Gottschalk, G., and Dobrindt, U. 2004. Pathogenomics of mobile genetic elements of toxigenic bacteria. Int. J. Med. Microbiol. 293:453-461.
- Hayward, A. C. 1964. Characteristics of *Pseudomonas solanacearum*. J. Appl. Bacteriol. 27:265-277.
- Hayward, A. C. 1991. Biology and epidemiology of bacterial wilt caused by *Pseudomonas solanacearum*. Annu. Rev. Phytopathol. 29:65-87.
- Hayward, A. C. 2000. *Ralstonia solanacearum*. Pages 32-42 in: Encyclopaedia of Microbiology. J. Lederberg, ed. Academic Press, San Diego, CA.
- He, L. Y., Sequeira, L., and Kelman, A. 1983. Characteristics of strains of Pseudomonas solanacearum from China. Plant. Dis. 67:1357-1361.
- Janse, J. D., Beld, H. E. V. D., Elphinstone, J., Simpkins, S., Tjou-Tam-Sin, N. A. A., and Vaerenbergh, J. V. 2004. Introduction to Europe of *Ralstonia solanacearum* biovar 2, race 3 in *Pelargonium zonale* cuttings. J. Plant Pathol. 86:147-155.
- 24. Kim, S. H., Olson, T. N., Schaad, N. W., and Moorman, G. W. 2003. *Ralstonia solanacearum* Race 3, biovar 2, the causal agent of brown rot of potato, identified in geranium in Pennsylvania, Delaware, and Connecticut. Plant Dis. 87:450.
- 25. Occhialini, A., Cunnac, S., Reymond, N., Genin, S., and Boucher, C. 2005. Genome-wide analysis of gene expression in *Ralstonia solana-cearum* reveals that the *hrp*B gene acts as a regulatory switch controlling multiple virulence pathways. Mol. Plant-Microbe Interact. 18:938-949.
- Palomo, J. L., Caruso, P., Gorris, M. T., Lopez, M. M., and Garcia-Benavides P. 2000. Comparacion de metodos de detección de *Ralstonia solanacearum* en aguas superficiales. Page 120 in: Abstr. 10th Congr. Soc. Espanola Fitopatol., Valencia, Spain.
- Perez S. A., Mejia, L., Fegan, M., and Allen, C. 2008. Diversity and distribution of *Ralstonia solanacearum* strains in Guatemala and rare occurrence of tomato fruit infection. Plant Pathol. 57:320-331.
- Prior, P., and Fegan, M. 2005. Recent development in the phylogeny and classification of *Ralstonia solanacearum*. Acta Hortic. 695:127-136.
- 29. Priou, S., and Gutarra, L. 1999. Highly sensitive detection of Ralstonia

solanacearum in latently infected potato tubers by post-enrichment enzyme-linked immunosorbent assay on nitrocellulose membrane. Bull. OEPP 29:117-125.

- Reymond, N., Charles, H., Duret, L., Calevro, F., Beslon, G., and Fayard, J. M. 2004. ROSO: optimizing oligonucleotide probes for microarrays. 1. Bioinformatics 20:271-273.
- 31. Salanoubat, M., Genin, S., Artiguenave, F., Gouzy, J., Mangenot, S., Arlat, M., Billault, A., Brottier, P., Camus, J. C., Cattolico, L., Chandler, M., Choisne, N., Claudel-Renard, C., Cunnac, S., Demange, N., Gaspin, C., Lavie, M., Moisan, A., Robert, C., Saurin, W., Schiex, T., Siguier, P., Thébault, P., Whalen, M., Wincker, P., Levy, M., Weissenbach, J., and Boucher, C. A. 2002. Genome sequence of the plant pathogen *Ralstonia solanacearum*. Nature 415:497-502.
- Swanson, J. K., Montes, L., Mejia, L., and Allen, C. 2007. Detection of latent infections of *Ralstonia solanacearum* race 3 biovar 2 in geranium. Plant Dis. 91:828-834.
- Swanson, J. K., Yao J., Tans-Kersten J., and Allen C. 2005. Behavior of *Ralstonia solanacearum* race 3 biovar 2 during latent and active infection of geranium. Phytopathology 95:136-143.

- Tambong, J. T., de Cock, A. W. A. M., Tinker, N. A., and Lévesque, C. A. 2006. Oligonucleotide array for identification and detection of *Pythium* species. Appl. Environ. Microbiol. 72:2691-2706.
- Terol, J., Munoz-Soriano, V., Arahal D. R., Buades, C., Pérez-Pérez, A., Llop, P., Belmonte, U. C. F., Lopez, M. M., and Pérez-Alonso, M. 2006. Random genome sequencing of *Ralstonia solanacearum* strain IVIA 1602 and comparative analysis with strain GMI1000. J. Phytopathol. 154:556-564.
- Weller, S. A., Elphinstone, J. G., Smith, N. C., Boonham, N., and Stead, D. E. 2000. Detection of *Ralstonia solanacearum* strains with a quantitative, multiplex, real-time, fluorogenic PCR (TaqMan) assay. Appl. Environ. Microbiol. 66:2853-2858.
- Wicker, E., Grassart, L., Coranson-Beaudu, R., Mian, D., Guilbaud, C., Fegan, M., and Prior, P. 2007. *Ralstonia solanacearum* strains in Martinique (French West Indies) exhibiting a new pathogenic potential. Appl. Environ. Microbiol. 73:6790-6801.
- Wood, J. R., Breckenridge, K., and Chard, J. M. 2002. Eradicating Ralstonia solanacearum from Scottish rivers. Page 31 in: Abstr. 3rd Int. Bacterial Wilt Symp. (IBWS), Nelspruit, South Africa.