AMERICAN SOCIETY FOR MICROBIOLOGY | Infection and Immunity

Genomic Changes Associated with the Loss of *Nocardia brasiliensis* Virulence in Mice after 200 *In Vitro* Passages

Carolina Gonzalez-Carrillo,[a] Cassandra Millan-Sauceda,[a] Hector Gerardo Lozano-Garza,[a] Rocio Ortiz-Lopez,[b,c] Ramiro Elizondo-Gonzalez,[c] Oliverio Welsh,[a] Jorge Ocampo-Candiani,[a] Lucio Vera-Cabrera[a]

Laboratorio Interdisciplinario de Investigación Dermatológica, Servicio de Dermatología, Hospital Universitario, U.A.N.L., Monterrey, Mexico[a]; Universidad Autónoma de Nuevo León, Departmento de Bioquímica y Medicina Molecular, Monterrey, Mexico[b]; Universidad Autónoma de Nuevo León, Centro de Investigación y Desarrollo en Ciencias de la Salud, Monterrey, Mexico[c]

***Nocardia* species, particularly *Nocardia brasiliensis*, are etiologic agents of mycetoma, a chronic subcutaneous infection. Until now, little has been known about the pathogenic mechanisms involved in nocardial infection. Traditionally, subculture in rich media has been a simple way to induce attenuation. In this work, we report the changes in virulence toward mice and in genomic constitution of *N. brasiliensis* produced after 200 continuous subcultures in brain heart infusion (BHI) medium (P-200 strain). The ability of the *N. brasiliensis* P-200 strain to produce experimental infection was tested using BALB/c mice. P-200 was also used to immunize mice to determine whether it could induce resistance against a challenge with a nonsubcultured isolate (P-0). Comparative proteomic analysis between *N. brasiliensis* P-0 and P-200 was performed by two-dimensional (2-D) electrophoresis, and the genome sequence was obtained through Roche 454 sequence analysis. Virulence in BALB/c mice was completely lost, and BALB/c mice immunized with P-200 bacterial cells were resistant to mycetoma production by the nonsubcultured strain. Whole-genome sequence analysis revealed that P-200 lost a total of 262,913 bp distributed in 19 deleted regions, involving a total of 213 open reading frames (ORFs). The deleted genes included those encoding bacterial virulence factors, e.g., catalase, nitrate reductase enzymes, and a group of mammalian cell entry (MCE) family proteins, which may explain the loss of virulence of the isolate. Thus, completely attenuated *N. brasiliensis* was obtained after 200 passages in BHI medium, and putative *Nocardia* virulence genes were identified for the first time.**

Mycetoma is a chronic subcutaneous infection reported mostly in countries with tropical and subtropical weather (1); it is characterized by tumefaction of the anatomical site affected (particularly the upper and lower limbs), the production of subcutaneous abscesses, and fistulae. It is a deforming and stigmatizing disease caused by fungi or soil actinobacteria. The etiologic agent is inoculated through the skin via minor trauma with wood splinters or thorns contaminated with soil or organic matter (1). In Mexico, the most commonly isolated agents are *Nocardia brasiliensis* and *Actinomadura madurae*. The first agent causes approximately 70% of cases in the country and more than 90% of cases in the state of Nuevo Leon (1).

N. brasiliensis is a species of Gram-positive, partially acid-fast filamented bacilli that belongs to the *Corynebacterineae* suborder, a group of bacteria characterized by a type IV cell wall (which possesses an arabinogalactan cell wall polysaccharide) and by the presence of abundant mycolic acids and trehalose-derived lipids (2). Little is known about the pathogenic properties of *Nocardia* spp.; several molecules have been reported among the putative virulence factors, including superoxide dismutase (SOD) and catalase, enzymes that may decrease the ability of phagocytes to destroy bacteria by O_2-derived mechanisms (3). An intact *N. brasiliensis* cell wall also appears to be important to avoid intracellular destruction either by polymorphonuclear leukocytes or by rabbit alveolar or human-derived THP-1 macrophages. The removal of its outer layer appears to decrease the virulence of *N. brasiliensis* (4).

Recently, the complete genome sequence of *N. brasiliensis* strain HUJEG-1 became available (5). Bioinformatic analysis revealed the presence of an extensive synthetic machinery for lipid compounds, nonribosomal protein synthases (NRPS), hydrolases, lipases, and proteases that might be important for nocardial virulence. In this work, we induced attenuation of *N. brasiliensis* HUJEG-1 by subculturing it 200 times in brain heart infusion (BHI) medium, and we determined putative virulence-associated genes by using whole-genome sequence analysis.

MATERIALS AND METHODS

Subculture method. *N. brasiliensis* HUJEG-1 (ATCC 700358), which has been utilized in previous assays (6, 7), was used for these experiments. Bacterial cultures obtained from mouse lesions were kept frozen at −70°C in 20% skim milk and represented the parental strain (P-0). From these stocks, bacteria were grown on Sabouraud agar at 30°C for 4 to 7 days, and a single colony was then placed in a 7-ml sterile Eveljham-Potter device. We added 2.5 ml of sterile saline, ground the bacterial mass to obtain a homogeneous suspension, and adjusted the turbidity to a McFarland standard of 1. We inoculated a 125-ml Erlenmeyer flask containing 33 ml of previously sterilized liquid BHI medium with 0.1 ml of this suspension. We then incubated the culture with constant agitation at 110 rpm and 37°C. After 72 h, the bacterial mass was harvested by centrifugation at

Received 15 April 2016 **Returned for modification** 1 June 2016
Accepted 18 June 2016
Accepted manuscript posted online 27 June 2016
Citation Gonzalez-Carrillo C, Millan-Sauceda C, Lozano-Garza HG, Ortiz-Lopez R, Elizondo-Gonzalez R, Welsh O, Ocampo-Candiani J, Vera-Cabrera L. 2016. Genomic changes associated with the loss of *Nocardia brasiliensis* virulence in mice after 200 *in vitro* passages. Infect Immun 84:2595–2606. doi:10.1128/IAI.00329-16.
Editor: A. Camilli, Tufts University School of Medicine
Address correspondence to Lucio Vera-Cabrera, luvera_99@yahoo.com.

FIG 1 Biological changes observed in *N. brasiliensis* after 200 passages in BHI medium. (Left) Growth of P-0 in BHI medium as small microcolonies of bacteria. (Middle) After 200 passages, the isolate grew as a homogenous suspension. (Right) Kinyoun staining of P-200, showing a negative reaction.

2,500 rpm for 3 min, washed, and ground as described above. A new Erlenmeyer flask was inoculated with 0.1 ml of this suspension. These steps were repeated until 200 subcultures (P-200) had been reached. Samples were taken every 10 passages (including P-0) and kept frozen at −70°C. The entire process took approximately 6 years to complete.

Experimental mycetoma in a murine model. Cultures were obtained from aliquots of the P-200 and P-0 strains, which were stored in a deep freezer. The inoculums were prepared using a previously published technique (8) and adjusted to 20 mg (wet weight) of *N. brasiliensis* in 50 μl of saline solution. Female 8- to 12-week-old BALB/c mice were injected with 50 μl of nocardial suspension in the right footpad, and the development of lesions was scored from 0 (no inflammatory changes) to 4+ (extension of the lesions beyond the ankle of the animal, with extensive production of inflammation and abscesses), as previously described (8). The thickness of each lesion was measured with calipers every week for 12 weeks.

The study was approved by the Comité Local de Investigación en Salud 1906, Centro de Investigación Biomédica del Noreste, IMSS. The animal handling was performed according to the NORMA Oficial Mexicana NOM-062-ZOO-1999 (Especificaciones técnicas para la producción, cuidado y uso de los animales de laboratorio [Technical specifications for the production, care and handling of laboratory animals]).

Induction of infection resistance in a murine model. To examine whether infection with subcultured *N. brasiliensis* produced a state of immune resistance, a group of animals ($n = 20$) were inoculated in the right footpad with *N. brasiliensis* that had been subcultured 200 times (P-200). After 12 weeks, the left footpad was inoculated with nonsubcultured bacteria (P-0). As a control, we inoculated a group of animals of the same age ($n = 20$) with the nonsubcultured isolate in the right footpad. In all cases, the development of lesions was scored and measured as described above.

Proteomic analysis. Both *N. brasiliensis* strains (P-0 and P-200) were cultured on RPMI 1640 medium with agitation for 2 weeks at 37°C and 110 rpm. The supernatant or culture filtrate protein (CFP) was collected by centrifugation and concentrated by lyophilization. To obtain the intracellular proteins, a 3-day culture in BHI medium was disrupted with a fast-prep system using zirconia beads. The debris was eliminated by centrifugation, and the supernatant was dialyzed against phosphate-buffered saline (PBS). In both cases, the proteins were quantified by the Bradford method. Approximately 100 μg of each antigen was analyzed by two-dimensional sodium dodecyl sulfate-polyacrylamide gel electrophoresis (2-D SDS-PAGE). We initially used a pH range of 1 to 10 and later narrowed it to 4 to 7. The gels were stained with Coomassie blue R-250. Spots missing among the P-200 antigens were cut out of the P-0 gels by use of a scalpel, placed in an Eppendorf tube, covered with 100 μl of MilliQ water, and sent to Applied Biomics (Hayward, CA) for amino acid sequencing. The resulting sequences were analyzed using the BLAST program at the NCBI website.

Whole-genome sequencing. A suspension of *N. brasiliensis* HUJEG-1 subjected to 200 passages in BHI (P-200) was plated on a BHI agar plate to obtain separated colonies. After incubation for 6 days at 37°C, a colony was picked, a suspension in saline was prepared, and the procedure was repeated 4 more times to try to avoid DNA heterogeneity in the sample. The DNA was extracted from the last clone and subjected to mass sequencing using a Roche/454 GS (FLX Titanium) sequencing platform (8-kb library). The Roche/454 GS reads were assembled using Newbler 2.5.3 software (Roche Diagnostics, Branford, CT). The obtained contigs were compared to those already published for P-0 (GenBank accession number NC_018681.1) by using Sequencher software (Gene Codes, CA), and the presence of genetic changes, single nucleotide polymorphisms (SNPs), deletions, and duplications was scored.

Accession number(s). The data from this whole-genome shotgun project have been deposited at DDBJ/EMBL/GenBank under accession number LRRM00000000. The version described in this paper is version LRRM01000000.

RESULTS

Biological changes produced by continuous passaging. Bacteria were subcultured every 48 to 72 h, with constant agitation, in BHI medium; after 200 passages, changes in the macroscopic morphology were observed (Fig. 1). Instead of the bacteria growing as a tight cumulus of entangled filaments, a more disperse suspension was observed. Additionally, when the bacteria were stained with Kinyoun stain, P-200 cells showed lighter staining than P-0 cells, and when the nocardial mass was suspended in chloroform-methanol, differences in cell density were observed ($CHCl_3$ density = 1.48 g/ml; CH_3OH density = 0.791 g/ml). For P-0, the bacterial mass was observed at the bottom after centrifugation; in contrast, the P-200 culture separated the two solvents after centrifugation because it possessed an intermediate density (Fig. 2). Thin-layer chromatography (TLC) analysis of the chloroform-methanol extracts of P-200 and P-0 showed a marked decrease in the acyl-glycerol fraction in P-0 (Fig. 2).

Virulence of *N. brasiliensis* P-200. Virulence was tested in BALB/c mice. When mice were injected with the parental strain, P-0, 90% of the animals developed lesions after 12 weeks of infection (Fig. 3A). In contrast, only 10% of the animals infected with P-200 developed small lesions at this time, and after 2 weeks, the lesions completely disappeared. Fifteen animals from the P-200 strain-infected group, with absolutely no lesions, were challenged in the contralateral footpad with P-0 (Fig. 3B). Fifteen weeks after inoculation, all mice, except one which developed a 1+ lesion, presented no lesions (not shown). We analyzed the lesions histologically 5 weeks after infection with P-0, and we observed the presence of a strong mononuclear infiltrate in the lesions, along

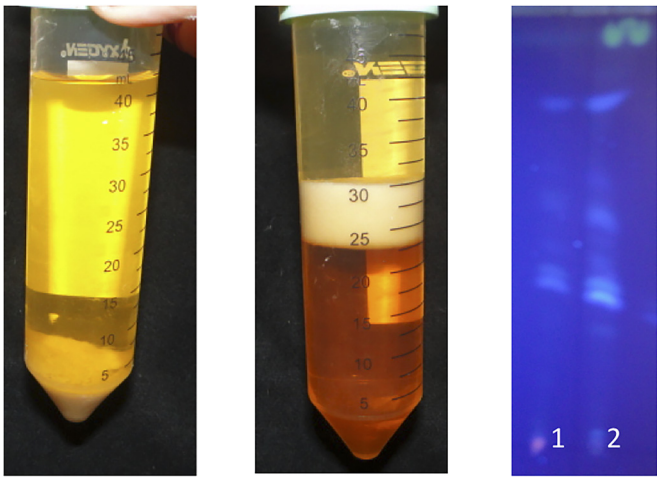

FIG 2 Changes in cell wall composition of *N. brasiliensis* P-200. (Left) Suspension of P-0 in chloroform-methanol. The bacterial cell mass stayed in the bottom of the tube. (Middle) Suspension of P-200 in the same solvent mix. In this case, the bacterial mass was located between both solvents. (Right) TLC analysis of chloroform-methanol extracts of P-0 (lane 1) and P-200 (lane 2).

with the presence of granules in several stages of apparent destruction (Fig. 3C).

Genomic mass sequencing of P-200. The genome sequence of the P-200 strain was determined using a Roche/454 GS (FLX Titanium) sequencing platform (8-kb library). A total of 44,023 reads were obtained, providing approximately 24× genome coverage. The Roche/454 GS reads were assembled into 70 scaffolds by use of Newbler 2.5.3 software (Roche Diagnostics, Branford, CT).

Comparing the sequences of the *N. brasiliensis* HUJEG-1 P-0 (accession number NC_018681.1) and P-200 scaffolds revealed the presence of 17 deleted regions (Table 1), with sizes ranging from 8 to 216,368 bp, for a total of 262,913 bp affecting 213 genes, including 107 exclusive to *N. brasiliensis*.

On analyzing the clusters of orthologous groups of proteins (COGs) of the deleted genes, we observed 156 different NCBI COGs. The most frequent COGs, with four genes each, were COG583 (transcriptional regulators), COG596 (predicted hydrolases or acyltransferases), COG1463 (ABC-type transport system involved in resistance to organic solvents, periplasmic component [secondary metabolite biosynthesis, transport, and catabolism]), and COG2207 (AraC-type DNA-binding domain-containing proteins). There were six other COGs associated with three genes, and the rest were associated with one or two genes.

In Table 1, we list the hypothetical open reading frames (ORFs) in the deleted DNA regions. The putatively important lost virulence genes included genes encoding a catalase, an SOD, several proteases and peptidases, and a mammalian cell entry (MCE) operon. Other single-nucleotide indels are summarized in Table 2.

A comparison analysis revealed 36 SNPs (Table 3), including 5 synonymous SNPs, 15 nonsynonymous SNPs, and 16 SNPs located in intergenic regions (Table 3). Some putative affected proteins involved in virulence included several peptidases and certain enzymes involved in cell wall peptidoglycan synthesis, such as O3I_042080, a peptidoglycan lipid II flippase and an ortholog of MurJ of *Escherichia coli*. This orthologous protein is an important enzyme in peptidoglycan translocation to the bacterial periplasm.

Other genetic changes observed in P-200 included short nucleotide sequence duplications. At nucleotide 7,459,445, we observed 12 duplications of the sequence 5′-TGGCCGGGC-3′, interrupting gene 03I-032985 (from nucleotides 7,459,087 to 7,460,010), which encodes a hypothetical protein. BLAST analysis of the protein sequence encoded by this ORF showed homology to an orthologous protein sequence present in several *N. brasiliensis* strains but not in other *Nocardia* species. It also showed homology to a streptomyces RNA polymerase sigma factor. Another duplication of 77 nucleotides occurred at position 6,504,536, in an intergenic zone (not shown).

Proteomic analysis of P-200. To study the putative proteomic changes derived from the genetic changes, we analyzed the protein content of a cellular extract (CE) obtained by breakage of nocardial cells by use of a fast-prep system, and we also analyzed proteins secreted into the medium (culture filtrate proteins [CFPs]). The extracts were analyzed by SDS-PAGE with a 12% gel and/or by 2-D electrophoresis using a pH gradient of 4 to 7 (Bio-Rad) and a 12% SDS-PAGE gel. The CE of P-0 showed abundant dots in the pH range of 4 to 7. Comparison with the P-200 CE dot pattern

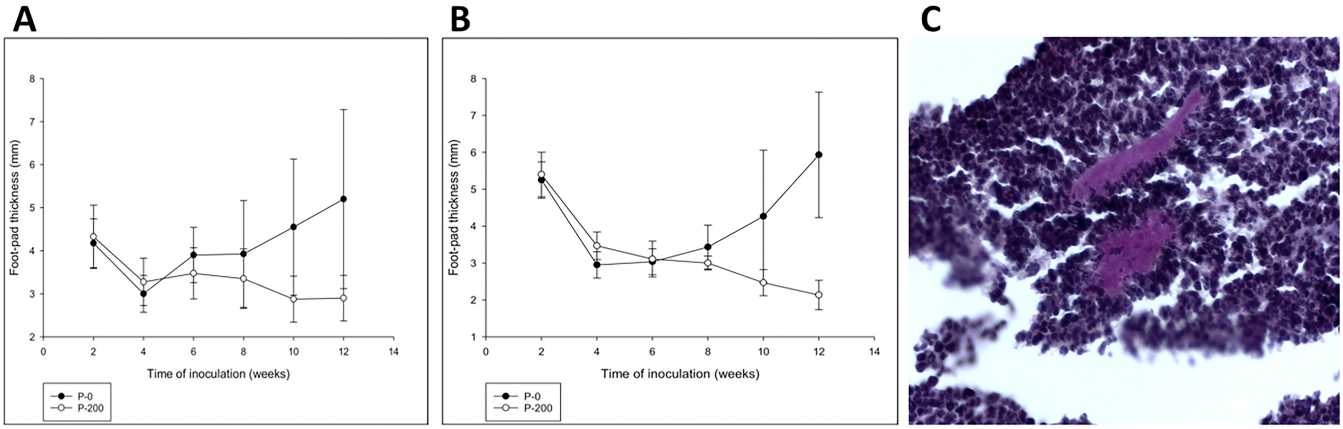

FIG 3 Virulence assay of P-0 and P-200 in BALB/c mice. (A) Female animals were inoculated with either P-0 or P-200, and the footpad thickness was measured every week. (B) Twelve weeks later, the animals inoculated with P-200 were challenged with P-0. As a control, a naive group of animals of the same age were inoculated with P-0. (C) Histological findings in a footpad biopsy specimen from a mouse inoculated with P-200 and reinoculated with P-0, showing polymorphonuclear leukocytes destroying the *Nocardia* granules.

TABLE 1 Deleted regions in *N. brasiliensis* P-200[a]

Region no.	Deleted region				Corresponding region in reference sequence		Organism(s) with ortholog	NCBI COG
	Start position	End position	Gene	Protein	Start position	End position		
1	541011	541445	03I_002340	Clp protease	540929	543247	*Nocardia cyriacigeorgica*, *Nocardia farcinica*, *Rhodococcus jostii*	COG0542
2	1449785	1450050	Intergenic region					
3	1453274	1453310	Intergenic region					
4	1501900	1533608	03I_006545	Putative drug resistance transporter	1500162	1502150	*N. cyriacigeorgica*, *N. farcinica*	COG0477
			03I_006550	PadR family transcriptional regulator	1502234	1502887	*N. cyriacigeorgica*, *N. farcinica*	COG1695
			03I_006555	Monooxygenase, FAD-binding protein	1503068	1504273	*N. cyriacigeorgica*, *N. farcinica*	COG0654
			03I_006560	ATP-binding transport protein NatA	1504341	1505108	*N. brasiliensis*	COG1131
			03I_006565	Hypothetical protein	1505093	1506658	*N. brasiliensis*	
			03I_006570	LysR family transcriptional regulator	1506719	1507639	*N. brasiliensis*	COG0583
			03I_006575	XRE family transcriptional regulator	1507644	1508447	*N. cyriacigeorgica*, *R. jostii*	COG1396
			03I_006580	Hypothetical protein	1508578	1509024	*N. cyriacigeorgica*	
			03I_006585	Hypothetical protein	1509060	1509791	*N. cyriacigeorgica*, *R. jostii*	
			03I_006590	Haloalkane dehalogenase	1510380	1511252	*N. cyriacigeorgica*	COG0596
			03I_006595	Putative RNA polymerase ECF-type sigma factor	1511337	1512200	*N. cyriacigeorgica*	COG1595
			03I_006600	C_4-dicarboxylate transporter/malic acid transport protein	1512157	1513347	*N. cyriacigeorgica*, *R. jostii*	COG1275
			03I_006605	LysR family transcriptional regulator	1513418	1514326	*N. cyriacigeorgica*, *R. jostii*	COG0583
			03I_006610	Phosphate uptake regulator PhoU	1514345	1515028	*N. cyriacigeorgica*, *N. farcinica*	COG0704
			03I_006615	Transglutaminase domain-containing protein	1515153	1515956	*R. jostii*	COG1305
			03I_006620	Na-Ca exchanger/integrin-beta4	1515963	1516991	*N. brasiliensis*	
			03I_006625	Pyridoxamine 5′-phosphate oxidase-related FMN-binding protein	1518688	1519095	*N. cyriacigeorgica*	COG3467
			03I_006630	Transporter	1519200	1520300	*N. farcinica*, *R. jostii*	COG0387
			03I_006635	YrbE family protein	1520752	1520752	*N. cyriacigeorgica*, *N. farcinica*	COG0767
			03I_006640	YrbE family protein	1521601	1522467	*N. cyriacigeorgica*, *N. farcinica*	COG0767
			03I_006645	Mce family protein	1522498	1522498	*N. cyriacigeorgica*, *N. farcinica*	COG3008
			03I_006650	Mce family protein	1523496	1524497	*N. cyriacigeorgica*, *N. farcinica*	COG1463
			03I_006655	Mce family protein	1524485	1525519	*N. cyriacigeorgica*, *N. farcinica*	COG1463
			03I_006660	Mce family protein	1525489	1526604	*N. cyriacigeorgica*, *N. farcinica*	COG1463
			03I_006665	Mce family protein	1526601	1526601	*N. cyriacigeorgica*, *N. farcinica*	COG0282
			03I_006670	Mce family protein	1527728	1527728	*N. cyriacigeorgica*, *N. farcinica*	COG1463
			03I_006675	Hypothetical protein	1528699	1529298	*N. cyriacigeorgica*, *N. farcinica*	
			03I_006680	Hypothetical protein	1529433	1529433	*N. brasiliensis*	
			03I_006685	Aminoglycoside *O*-phosphotransferase	1530067	1530951	*N. cyriacigeorgica*, *N. farcinica*, *R. jostii*	COG3570
			03I_006690	Short-chain dehydrogenase/reductase SDR	1530978	1530978	*N. brasiliensis*	COG1028
			03I_006695	Signal transduction histidine kinase	1532065	1533246	*N. brasiliensis*	COG4585
			03I_006700	Two-component LuxR family transcriptional regulator	1533243	1533890	*N. brasiliensis*	COG2197
5	2271950	2277016	03I_Or42277	16S rRNA	2271959	2273488		
6	2372465	2372842	03I_010455	Glyoxylate carboligase	2372219	2373994		COG3960
7	2578134	2578295	Intergenic region					

(Continued on following page)

TABLE 1 (Continued)

Region no.	Deleted region		Gene	Protein	Corresponding region in reference sequence		Organism(s) with ortholog	NCBI COG
	Start position	End position			Start position	End position		
8	2933006	2933045	Intergenic region					
9	3597423	3597782	03I_016110	DNA-directed RNA polymerase subunit beta (RpoB)	3596307	3599822	*N. cyriacigeorgica*, *N. farcinica*, *R. jostii*	
10	3901876	3903739	03I_Or42279	16S rRNA	3902217	3903746		
11	3904013	3907228	Intergenic region					
12	4429801	4429809	Intergenic region					
13	4795932	5012300	03I_021225	Gamma-glutamyltransferase	4795342	4797141	*N. brasiliensis*	COG0405
			03I_021230	Methyltransferase	4797178	4798242	*N. brasiliensis*	COG0500
			03I_021235	Cytochrome P-450	4798239	4799441	*R. jostii*	COG2124
			03I_021240	Alcohol dehydrogenase GroES domain-containing protein	4799438	4800493	*N. brasiliensis*	COG1064
			03I_021245	Aldehyde dehydrogenase	4800513	4802000	*N. farcinica*, *R. jostii*	COG1012
			03I_021250	Putative amidotransferase	4802029	4803891	*R. jostii*	COG0367
			03I_021255	ABC transporter	4803936	4806227	*N. cyriacigeorgica*, *N. farcinica*, *R. jostii*	COG0178
			03I_021260	Short-chain fatty acid MFS superfamily protein	4806415	4807854	*R. jostii*	COG2031
			03I_021265	Hypothetical protein	4807919	4808521	*N. brasiliensis*	
			03I_021270	Hypothetical protein	4808582	4809127	*N. farcinica*	
			03I_021275	Hypothetical protein	4809244	4810527	*N. brasiliensis*	
			03I_021280	Hypothetical protein	4810655	4811164	*N. farcinica*	COG2259
			03I_021285	Hypothetical protein	4811244	4812002	*N. farcinica*, *R. jostii*	COG3384
			03I_021290	MarR family transcriptional regulator	4812093	4812554	*N. farcinica*, *R. jostii*	COG1846
			03I_021295	NADP-dependent oxidoreductase domain-containing protein	4812632	4813675	*R. jostii*	COG0667
			03I_021300	Abortive infection protein	4813887	4814675	*N. brasiliensis*	COG1266
			03I_021305	Dioxygenase	4814780	4815688	*N. farcinica*	COG2175
			03I_021310	Transporter	4815883	4816527	*N. farcinica*	COG1174
			03I_021315	ABC transporter ATP-binding protein	4816524	4817804	*N. farcinica*	COG1125
			03I_021320	ABC transporter permease	4817801	4818538	*N. farcinica*	COG1174
			03I_021325	Transporter permease	4818535	4819515	*N. farcinica*	COG1732
			03I_021330	Hypothetical protein	4819550	4820365	*N. farcinica*, *R. jostii*	
			03I_021335	Hypothetical protein	4820488	4821048	*N. brasiliensis*	
			03I_021340	Nonribosomal peptide synthetase	4821049	4841793	*N. cyriacigeorgica*, *N. farcinica*, *R. jostii*	COG1020
			03I_021345	Hypothetical protein	4841793	4857035	*N. cyriacigeorgica*, *N. farcinica*, *R. jostii*	COG1020
			03I_021350	SARP family transcriptional regulator	4857537	4860785	*N. brasiliensis*	COG0745
			03I_021355	Hypothetical protein	4860727	4861287	*N. brasiliensis*	
			03I_021360	Hypothetical protein	4861313	4862404	*N. brasiliensis*	COG0451
			03I_021365	Nonribosomal peptide synthetase	4862392	4863786	*N. brasiliensis*	COG1020
			03I_021370	Acetylornithine deacetylase or succinyl-diaminopimelate desuccinylase	4863783	4864988	*R. jostii*	COG0624
			03I_021375	UbiE/COQ5 family methyltransferase	4865045	4865869	*N. brasiliensis*	COG0500
			03I_021380	Alcohol dehydrogenase	4866058	4867254	*N. farcinica*, *R. jostii*	COG1454
			03I_021385	Sensor histidine kinase	4867244	4868566	*N. brasiliensis*	
			03I_021390	D-3-Phosphoglycerate dehydrogenase	4868598	4869608	*R. jostii*	COG0111
			03I_021395	Phosphoenolpyruvate phosphomutase	4869662	4870588	*N. brasiliensis*	COG2513
			03I_021400	Phosphonopyruvate decarboxylase	4870585	4871736	*N. brasiliensis*	COG4032
			03I_021405	Aldehyde dehydrogenase	4871726	4873117	*N. brasiliensis*	COG1012
			03I_021410	Short-chain dehydrogenase/reductase SDR	4873151	4873954	*N. brasiliensis*	COG1028
			03I_021415	Polyketide synthase	4873956	4874171	*N. brasiliensis*	COG3321
			03I_021420	AMP-dependent synthetase and ligase	4874171	4875613	*N. brasiliensis*	COG0318
			03I_021425	Class III aminotransferase	4875616	4876941	*N. brasiliensis*	COG0160
			03I_021430	Hypothetical protein	4876976	4877728	*N. brasiliensis*	COG0842
			03I_021435	ABC transporter-like protein	4877725	4878678	*N. brasiliensis*	COG1131

(Continued on following page)

TABLE 1 (Continued)

Region no.	Deleted region		Gene	Protein	Corresponding region in reference sequence		Organism(s) with ortholog	NCBI COG
	Start position	End position			Start position	End position		
			03I_021440	Penicillin-binding protein	4878669	4879817	*N. brasiliensis*, *R. jostii*	COG1680
			03I_021445	Pantoate–beta-alanine ligase PanC	4879877	4880716	*N. brasiliensis*	COG0414
			03I_021450	LuxR family transcriptional regulator	4880906	4881586	*N. brasiliensis*	COG2197
			03I_021455	Integral membrane sensor signal transduction histidine kinase	4881576	4882817	*N. brasiliensis*	COG4585
			03I_021460	Hypothetical protein	4882983	4883552	*N. brasiliensis*	
			03I_021465	Hypothetical protein	4883543	4883869	*N. brasiliensis*	
			03I_021470	Hypothetical protein	4883931	4884668	*N. brasiliensis*, *N. farcinica*, *R. jostii*	COG3393
			03I_021475	LuxR family transcriptional regulator	4884766	4887378	*N. brasiliensis*, *R. jostii*	COG1066
			03I_021480	Hypothetical protein	4887544	4887990	*N. brasiliensis*, *R. jostii*	
			03I_021485	Hypothetical protein	4888075	4888605	*N. brasiliensis*	
			03I_021490	Hypothetical protein	4888830	4889687	*N. brasiliensis*, *N. farcinica*	COG0693
			03I_021495	HTH-type transcriptional regulator GlxA	4889733	4890743	*N. brasiliensis*, *N. cyriacigeorgica*	COG4977
			03I_021500	Exonuclease SbcC	4890883	4891464	*N. brasiliensis*, *N. farcinica*, *R. jostii*	
			03I_021505	Helix-turn-helix domain-containing protein	4891566	4892396	*N. brasiliensis*, *N. cyriacigeorgica*	COG2207
			03I_021510	Hypothetical protein	4892421	4892924	*N. brasiliensis*, *N. cyriacigeorgica*	
			03I_021515	Hypothetical protein	4893061	4893711	*N. brasiliensis*	
			03I_021520	Putative transcriptional regulator TetR	4893718	4894305	*N. brasiliensis*, *N. farcinica*	COG1309
			03I_021525	Phytanoyl-coenzyme A (CoA) dioxygenase (PhyH) family protein	4894403	4895257	*N. brasiliensis*, *N. farcinica*, *R. jostii*	COG5285
			03I_021530	Hypothetical protein	4895383	4896189	*N. brasiliensis*	
			03I_021535	Peptidase M14 carboxypeptidase	4896301	4897695	*N. brasiliensis*	COG2866
			03I_021540	Glutamate-cysteine ligase GCS2	4897981	4899078	*N. brasiliensis*	COG2170
			03I_021545	Hypothetical protein	4899169	4899396	*N. brasiliensis*	
			03I_021550	Short-chain dehydrogenase	4899479	4900318	*N. brasiliensis*, *N. cyriacigeorgica*, *N. farcinica*, *R. jostii*	COG1028
			03I_021555	FAD-binding monooxygenase	4900447	4902174	*N. brasiliensis*	
			03I_021560	Luciferase	4902171	4903256	*N. brasiliensis*	COG2141
			03I_021565	Pyridoxal-5′-phosphate-dependent protein subunit beta	4903272	4904462	*N. brasiliensis*	COG0498
			03I_021570	Opine dehydrogenase	4904459	4905574	*N. brasiliensis*, *N. farcinica*	
			03I_021575	Acyl-CoA reductase	4905577	4907022	*N. brasiliensis*	
			03I_021580	Hypothetical protein	4907010	4908308	*N. brasiliensis*	
			03I_021585	Ferredoxin	4908610	4908825	*N. brasiliensis*	COG1141
			03I_021590	LuxE bioluminescence protein	4908850	4909983	*N. brasiliensis*	COG1541
			03I_021595	Acyltransferase LuxD	4910022	4910987	*N. brasiliensis*	
			03I_021600	Sodium/hydrogen exchanger	4910945	4912375	*N. brasiliensis*	COG0475
			03I_021605	Fluorinating enzyme	4912446	4913309	*N. brasiliensis*	COG1912
			03I_021610	*S*-Adenosyl-L-homocysteine hydrolase	4913437	4914891	*N. brasiliensis*, *N. cyriacigeorgica*, *N. farcinica*, *R. jostii*	COG0499
			03I_021615	Methylthioadenosine phosphorylase	4915122	4916000	*N. brasiliensis*	COG0005
			03I_021620	Adenine phosphoribosyltransferase	4916036	4916611	*N. brasiliensis*	COG0503
			03I_021625	Translation initiation factor, aIF-2BI family protein	4916581	4917585	*N. brasiliensis*	COG0182
			03I_021630	Serine hydroxymethyltransferase GlyA	4917783	4919630	*N. brasiliensis*, *R. jostii*	COG0112
			03I_021635	DNA-binding regulatory protein	4919638	4920213	*N. brasiliensis*	COG1396

(Continued on following page)

TABLE 1 (Continued)

Region no.	Deleted region		Gene	Protein	Corresponding region in reference sequence		Organism(s) with ortholog	NCBI COG
	Start position	End position			Start position	End position		
			03I_021640	Hypothetical protein	4920272	4921222	*N. brasiliensis*	COG0697
			03I_021645	Two-component system response regulator	4921827	4922435	*N. brasiliensis*	COG2197
			03I_021650	Universal stress protein UspA-like protein	4922436	4922879	*N. brasiliensis*	COG0589
			03I_021655	Hypothetical protein	4922918	4923145	*N. brasiliensis*	
			03I_021660	Hypothetical protein	4923186	4923701	*N. brasiliensis*	COG2606
			03I_021665	OsmC-like protein	4923894	4924364	*N. brasiliensis*, *R. jostii*	COG1764
			03I_021670	Peptidase M15A	4924375	4926096	*N. brasiliensis*	COG3108
			03I_021675	Peptidase M15A	4926115	4928226	*N. brasiliensis*	
			03I_021680	Hypothetical protein	4928280	4929614	*N. brasiliensis*	
			03I_021685	Hypothetical protein	4929628	4930167	*N. brasiliensis*	
			03I_021690	Hypothetical protein	4930190	4931440	*N. brasiliensis*	
			03I_021695	Hypothetical protein	4931440	4933926	*N. brasiliensis*	
			03I_021700	Hypothetical protein	4933923	4934150	*N. brasiliensis*	
			03I_021705	SARP family transcriptional regulator	4934354	4935067	*N. brasiliensis*	COG3629
			03I_021710	Hypothetical protein	4934996	4936714	*N. brasiliensis*	
			03I_021715	SARP family transcriptional regulator	4936750	4937550	*N. brasiliensis*	COG3629
			03I_021720	Alpha/beta-fold hydrolase	4937701	4938486	*N. brasiliensis*	COG0596
			03I_021725	Oxidoreductase	4938581	4939534	*N. brasiliensis*	COG0604
			03I_021730	TetR family transcriptional regulator	4939638	4940246	*N. brasiliensis*	COG1309
			03I_021735	FAD-linked oxidase domain-containing protein	4940308	4941654	*N. brasiliensis*	COG0277
			03I_021740	Hypothetical protein	4941757	4942299	*N. brasiliensis*, *N. cyriacigeorgica*, *R. jostii*	COG5485
			03I_021745	4-Carboxymuconolactone decarboxylase	4942296	4942679	*N. brasiliensis*	COG0599
			03I_021750	3-Oxoadipate enol-lactonase	4942676	4943452	*N. brasiliensis*, *R. jostii*	COG0596
			03I_021755	3-Carboxy-*cis*,*cis*-muconate cycloisomerase	4943449	4944795	*N. brasiliensis*, *R. jostii*	COG0015
			03I_021760	Protocatechuate 3,4-dioxygenase alpha subunit	4944788	4945312	*N. brasiliensis*, *R. jostii*	COG3485
			03I_021765	Protocatechuate 3,4-dioxygenase beta subunit	4945305	4946045	*N. brasiliensis*, *R. jostii*	COG3485
			03I_021780	CoA transferase B subunit	4947234	4948013	*N. brasiliensis*, *N. farcinica*, *R. jostii*	COG2057
			03I_021785	CoA transferase A subunit	4948010	4948825	*N. brasiliensis*, *N. farcinica*, *R. jostii*	COG1788
			03I_021790	4-Hydroxybenzoate 3-monooxygenase	4948828	4950030	*N. brasiliensis*, *R. jostii*	COG0654
			03I_021795	Transmembrane transport protein	4950027	4951259	*N. brasiliensis*, *R. jostii*	COG0477
			03I_021800	LysR family transcriptional regulator	4951362	4952264	*N. brasiliensis*, *R. jostii*	COG0583
			03I_021805	Glycosidase	4952276	4954621	*N. brasiliensis*, *R. jostii*	COG1554
			03I_021810	Hydrolase	4954618	4955349	*N. brasiliensis*, *R. jostii*	COG0637
			03I_021825	Hypothetical protein	4956335	4957054	*N. brasiliensis*, *N. cyriacigeorgica*, *N. farcinica*	COG2129
			03I_021830	DsbA oxidoreductase	4957051	4957644	*N. brasiliensis*	
			03I_021835	Transcriptional regulator	4957793	4958872	*N. brasiliensis*	COG2207
			03I_021840	Cell surface protein	4959086	4959961	*N. brasiliensis*	
			03I_021845	Hypothetical protein	4960040	4960426	*N. brasiliensis*	
			03I_021850	Putative cysteine synthase	4960527	4961639	*N. brasiliensis*, *N. cyriacigeorgica*, *N. farcinica*	COG0031
			03I_021855	Putative MFS transporter	4961636	4962907	*N. brasiliensis*, *N. cyriacigeorgica*, *N. farcinica*	COG0477
			03I_021860	Amino acid binding protein	4962998	4964092	*N. brasiliensis*, *R. jostii*	
			03I_021865	Sigma factor	4964772	4965512	*N. brasiliensis*	COG1191

(Continued on following page)

TABLE 1 (Continued)

Region no.	Deleted region		Gene	Protein	Corresponding region in reference sequence		Organism(s) with ortholog	NCBI COG
	Start position	End position			Start position	End position		
			03I_021870	Hypothetical protein	4965603	4965887	*N. brasiliensis*	
			03I_021875	AraC family transcriptional regulator	4965976	4966569	*N. brasiliensis*	COG2207
			03I_021880	Hypothetical protein	4966735	4966926	*N. brasiliensis*	
			03I_021885	5,10-Methylenetetrahydrofolate reductase	4967383	4968366	*N. brasiliensis*, *N. cyriacigeorgica*, *N. farcinica*, *R. jostii*	COG0685
			03I_021890	Hypothetical protein	4968457	4968840	*N. brasiliensis*, *N. farcinica*, *R. jostii*	
			03I_021895	Glyoxalase/bleomycin resistance protein/dioxygenase	4968853	4969248	*N. brasiliensis*	COG0346
			03I_021900	Thimet oligopeptidase	4969268	4971199	*N. brasiliensis*	COG0339
			03I_021905	PadR-like family transcriptional regulator	4971239	4971832	*N. brasiliensis*	COG1695
			03I_021910	Ferredoxin	4971845	4973587	*N. brasiliensis*, *R. jostii*	COG1018
			03I_021915	Hypothetical protein	4973826	4974689	*N. brasiliensis*, *N. cyriacigeorgica*, *N. farcinica*	COG0596
			03I_021920	Hypothetical protein	4974798	4975796	*N. brasiliensis*, *N. cyriacigeorgica*, *N. farcinica*, *R. jostii*	
			03I_021925	Methyltransferase	4975793	4976491	*N. brasiliensis*, *N. cyriacigeorgica*, *N. farcinica*, *R. jostii*	COG2890
			03I_021930	Hypothetical protein	4976529	4976678	*N. brasiliensis*, *N. cyriacigeorgica*, *N. farcinica*	COG3369
			03I_021935	Oxidoreductase	4976757	4978205	*N. brasiliensis*, *N. cyriacigeorgica*	COG0665
			03I_021940	Hypothetical protein	4978223	4978432	*N. brasiliensis*	
			03I_021945	Catalase	4978659	4980830	*N. brasiliensis*, *N. cyriacigeorgica*, *N. farcinica*, *R. jostii*	COG0753
			03I_021950	Cytochrome P450 monooxygenase	4980837	4982192	*N. brasiliensis*	COG2124
			03I_021955	DNA ligase	4982302	4984293	*N. brasiliensis*, *N. farcinica*	COG0272
			03I_021960	Propionyl-CoA carboxylase beta chain (PCCase) (propanoyl-CoA:carbon dioxide ligase)	4984482	4986032	*N. brasiliensis*, *N. cyriacigeorgica*, *N. farcinica*, *R. jostii*	COG4799
			03I_021965	Hypothetical protein	4986188	4986607	*N. brasiliensis*	
			03I_021970	Hypothetical protein	4986612	4987046	*N. brasiliensis*	
			03I_021975	Serine/threonine protein kinase	4987060	4988448	*N. brasiliensis*	COG0515
			03I_021980	Hypothetical protein	4988845	4990197	*N. brasiliensis*, *R. jostii*	
			03I_021985	Type 11 methyltransferase	4990248	4991027	*N. brasiliensis*	COG0500
			03I_021990	Hypothetical protein	4991125	4991433	*N. brasiliensis*	
			03I_021995	Hypothetical protein	4991467	4993941	*N. brasiliensis*, *R. jostii*	
			03I_022000	DNA-binding protein	4994209	4995144	*N. brasiliensis*	COG1396
			03I_022005	Monooxygenase FAD-binding protein	4995520	4997040	*N. brasiliensis*	COG0654
			03I_022010	Hypothetical protein	4997078	4997308	*N. brasiliensis*	
			03I_022015	LysR family transcriptional regulator	4997443	4998345	*N. brasiliensis*	COG0583
			03I_022020	Cobalt ABC transporter ATPase	4998379	4999098	*N. brasiliensis*, *R. jostii*	COG1122
			03I_022025	ABC transporter permease	4999095	4999856	*N. brasiliensis*, *R. jostii*	COG0619
			03I_022030	Cobalamin (vitamin B_{12}) biosynthesis CbiM protein	4999857	5000921	*N. brasiliensis*, *R. jostii*	COG0310
			03I_022035	ArsR family transcriptional regulator	5001072	5001416	*N. brasiliensis*	COG0640

(Continued on following page)

TABLE 1 (Continued)

Region no.	Deleted region		Gene	Protein	Corresponding region in reference sequence		Organism(s) with ortholog	NCBI COG
	Start position	End position			Start position	End position		
			03I_022040	Hypothetical protein	5001435	5002505	*N. brasiliensis*	COG3236
			03I_022045	Cupin	5002682	5003389	*N. brasiliensis*, *R. jostii*	COG2140
			03I_022050	RNA polymerase factor sigma 70	5003373	5004428	*N. brasiliensis*	COG1595
			03I_022055	NADPH:quinone reductase and related Zn-dependent oxidoreductase	5004425	5005423	*N. brasiliensis*	COG0604
			03I_022060	Hypothetical protein	5005604	5006257	*N. brasiliensis*, *N. cyriacigeorgica*, *N. farcinica*	
			03I_022065	Putative AraC family transcriptional regulator	5006364	5007230	*N. brasiliensis*, *N. cyriacigeorgica*	COG2207
			03I_022070	Hydrolase	5007230	5008423	*N. brasiliensis*	COG1680
			03I_022075	Hypothetical protein	5008639	5009004	*N. brasiliensis*	COG0346
			03I_022080	Hypothetical protein	5009015	5009701	*N. brasiliensis*	COG5479
			03I_022085	MarR family transcriptional regulator	5009886	5010323	*N. brasiliensis*, *N. cyriacigeorgica*, *R. jostii*	COG1846
			03I_022090	Hypothetical protein	5010400	5011047	*N. brasiliensis*, *N. cyriacigeorgica*	
			03I_022095	Hypothetical protein	5011268	5011873	*N. brasiliensis*	
			03I_022100	Putative membrane porin	5011922	5012581	*N. brasiliensis*	COG0094
14	5025398	5027080	03I_022175	Nitrate reductase Z subunit beta (NarY)	5024840	5026579	*N. cyriacigeorgica*, *N. farcinica*	COG1140
15	5029253	5029670	03I_022190	Nitrate reductase Z subunit alpha (NarZ)	5029170	5030117		COG5013
16	5393100	5393780	03I_023565	Putative peptidase	5392840	5393805	*N. brasiliensis*	COG3590
17	5867835	5868021	Intergenic region					

[a] The data were prepared by aligning the contigs obtained for P-200 with the reference sequence (accession number NC_18681.1) by using Sequencer software. The orthologous genes were determined by using BLAST searches.

showed that some protein dots were missing for P-200 (Fig. 4), including one for a low-molecular-weight protein. Since low-molecular-weight proteins have been associated with attenuation of *Mycobacterium bovis* BCG, we performed an amino acid sequence analysis of this dot. The results showed that this protein corresponds to a hypothetical protein (encoded by gene 03I_038645 in the *N. brasiliensis* genome [accession number NC_018681.1]) with a calculated molecular weight of 15,308. BLAST analysis with the whole-protein sequence showed high conservation among other *Nocardia* species, although it was not associated with any protein of known function. The 2-D analysis of the P-0 CFPs showed a smaller number of protein dots; two low-molecular-weight proteins were clearly lost (Fig. 4). Amino acid sequence analysis followed by BLAST searching identified

TABLE 2 Single-nucleotide indels

Indel no.	Position of indel	Original sequence[a]	Nucleotide change[a]	Protein designation	Gene	ORF location
1	19559	:	G		Intergenic region	
2	516378	C	:	Hypothetical protein	O3I_002245	515626–516600
3	645912	G	:		Intergenic region	
4	1451834	:	A		Intergenic region	
5	2086472	:	G		Intergenic region	
6	2277031	G	:		Intergenic region	
7	3631248	:	C		Intergenic region	
8	4429802	C	:		Intergenic region	
9	4429803	C	:		Intergenic region	
10	4429804	G	:		Intergenic region	
11	4429805	C	:		Intergenic region	
12	4429806	A	:		Intergenic region	
13	4429807	G	:		Intergenic region	
14	4429808	C	:		Intergenic region	
15	6723711	:	C	Hypothetical protein	O3I_029575	6723369–6723722
16	6777743	:	C	Hypothetical protein	O3I_029815	6777516–6777797
17	7492119	G	:		Intergenic region	
18	8590058	:	G		Intergenic region	
19	8590058	:	C		Intergenic region	

[a] :, no nucleotide at indicated position.

TABLE 3 Locations of SNPs in the *N. brasiliensis* P-200 genome[a]

SNP no.	Location	Original codon (amino acid)	Nucleotide change (amino acid)	Gene	Protein designation	Location
1	263186	CGC (R)	CTC (L)	O3I_001145	Hypothetical protein	261412–263706
2	307213	CAG (Q)	AAG (R)	O3I_001340	Hypothetical protein	305875–307494
3	307214	CAG (Q)	AGG (R)	O3I_001340	Hypothetical protein	305875–307494
4	415131	GCG (A)	ACG (T)	O3I_001810	Putative trypsin-like serine protease	414578–415738
5	442835	ATG (M)	GTG (V)	O3I_001920	DNA topoisomerase I subunit omega	441014–443857
6	939436	GCG (A)	GCA (A)	O3I_004090	Putative nonribosomal peptide synthetase (modular protein)	925811–942625
7	1022704	GCT (A)	GCC (A)	O3I_004395	Putative peptidase	1021063–1023102
8	1022963	GCG (A)	GTG (V)	O3I_004395	Putative peptidase	1021063–1023102
9	1214362	CGG (R)	TGG (W)	O3I_005190	Homoserine *O*-acetyltransferase	1213901–1215061
10	1240877	GAA (E)	GAT (D)	O3I_005320	D-Alanyl-D-alanine carboxypeptidase	1240385–1241674
11	1429121	GCC (A)	GTC (V)	O3I_006200	Homoserine kinase	1428787–1429752
12	1451830			Intergenic region		
13	1451835			Intergenic region		
14	1451836			Intergenic region		
15	2082888	GTC (V)	TTC (F)	O3I_009135	*O*-Dimethylpuromycin-*O*-methyltransferase	2082487–2083548
16	2277017			Intergenic region		
17	3631212			Intergenic region		
18	3631314			Intergenic region		
19	3631353			Intergenic region		
20	3631428			Intergenic region		
21	3631431			Intergenic region		
22	3656905	GTC (V)	GTG (V)	O3I_016400	FAD-dependent oxidoreductase	3656411–3657805
23	3864067	TCC (S)	TTC (F)	O3I_017285	Acyl-CoA dehydrogenase	3863076–3864872
24	5105062			Intergenic region		
25	6684580			Intergenic region		
26	6684588			Intergenic region		
27	6777353			Intergenic region		
28	7492092			Intergenic region		
29	7492093			Intergenic region		
30	7794000	GGC (G)	GGG (G)	O3I_034520	Membrane-bound C_5 sterol desaturase Erg3	7793862–7794752
31	7924188			Intergenic region		
32	7997064	GCG (A)	GCA (A)	O3I_035455	Hypothetical protein	7996777–7997214
33	8020777	GAC (D)	AAC (N)	O3I_035540	Alpha-ketoglutarate decarboxylase	8019496–8023245
34	8296987	GAA (E)	TAA (stop)	O3I_036955	Transcriptional regulator	8296886–8297533
35	9157861	TGG (W)	TGA (stop)	O3I_040880	Hypothetical protein	9157676–9158764
36	9419279	AGT (S)	GGT (G)	O3I_042080	Peptidoglycan lipid II flippase	9415559–9419329

[a] Nucleotide positions were determined by aligning P-200 contig sequences against the *N. brasiliensis* reference sequence (accession no. NC_18681.1).

one of them as the 10-kDa cochaperonin GroES. This protein, together with GroEL or HSP65, is important for the proper folding of many proteins; its deletion in *Mycobacterium tuberculosis* results in a nonviable clone (9). The GroES gene was complete in P-200; the product of the CDS coding for the GroES protein in *M. tuberculosis* produces several spots (10). Therefore, it is possible that this lost dot represents an isoform of the protein.

DISCUSSION

Continuous passaging has been used to attenuate microorganisms for a long time. In 1885, Pasteur and Roux attenuated rabies virus by subculturing a sample from wild rabies in rabbit brains (11, 12). Years later, Calmette and Guérin subcultivated a virulent isolate of *Mycobacterium bovis* to obtain an avirulent bacterium that is still used as a vaccine (bacillus Calmette-Guérin [BCG]) (13). However, in most of these cases, the molecular events underlying the attenuation were unknown at the time that the vaccines were developed. Molecular biology techniques, such as mass sequencing, are powerful tools by which to determine the nature of these biological changes. In the case of BCG, the loss of specific genes after approximately 230 continuous subcultures, leading to a loss of virulence, is already known (14); however, the original isolate was lost, thus making a comparative analysis impossible. Even as few as 15 to 20 subcultures can result in lost biological properties, such as virulence, in the case of fungi, protozoa, or viruses (15–17). The most extreme case of continuous passaging was reported for *E. coli* after 6,000 daily passages, equivalent to 40,000 generations (18). Few changes were observed before 20,000 generations; however, after that point, mutations accumulated rapidly in subsequent generations, resulting in a total loss of 1.2% of the genome, with 627 SNPs and 26 deletion or insertion changes after 40,000 generations. Most of the deletions were related to insertion sequences and were less than 25,000 bp long. A large inversion (1,493,854 bp) occurred as early as 5,000 passages. We previously reported a decrease in virulence of *N. brasiliensis* after fewer than 130 passages (27); however, the loss of virulence was not complete, and the genomes of the subcultured strains were not obtained. In

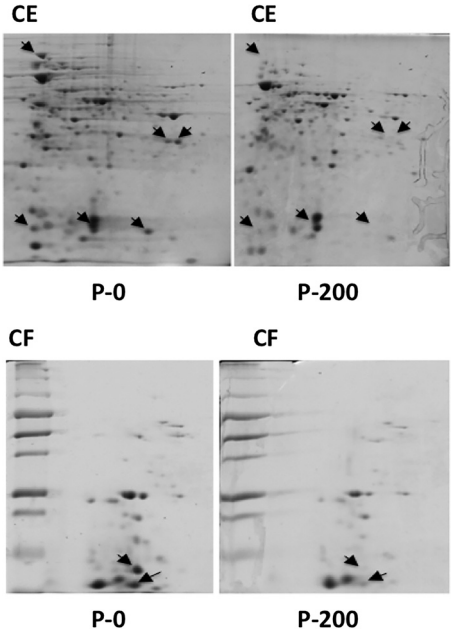

FIG 4 Analysis by 2-D gel electrophoresis of cell extracts (CE) and culture filtrates (CF) of *Nocardia brasiliensis*. A pH gradient of 4 to 7 was used for isoelectric focusing, and a 12% gel was used for the SDS-PAGE analysis. Clear differences in the spot patterns are indicated by arrows.

this work with P-200, we observed that *N. brasiliensis* lost a very large DNA fragment (262,913 bp) and that 36 SNPs occurred in a smaller number of duplications. Although in the case of *N. brasiliensis* the number of generations per subculture was not calculated (because it grows as entangled filaments rather than in cell suspensions as *E. coli* does), it took 6,000 subcultures to achieve 40,000 generations for *E. coli*, and therefore 200 subcultures may correspond to approximately 1,333 generations.

In a similar number of subcultures (270 passages), *Salmonella enterica* lost approximately 224,873 bp after 1,500 generations (19), which is more comparable to our case, in which 262,913 bp were lost after 200 subcultures. However, in the case of *S. enterica*, we do not know whether other biological changes associated with subculturing occurred.

A bovine isolate of *M. bovis* was used by Calmette and Guérin to produce an attenuated bacterial strain after 230 subcultures (13); the genes associated with virulence were identified by comparing the genome of BCG against the genome map of *M. tuberculosis*, with subsequent knockout analysis (14). Although *M. tuberculosis* and *M. bovis* are similar species, they are not identical and have important biological differences (20). The attenuated strain *M. bovis* BCG has been used as a vaccine since its production in 1921 (21); however, protection levels vary greatly (22). A possible explanation for the low level of protection by BCG is overpassaging (23). Because Calmette and Guérin did not have the tools to determine the point at which virulence was lost, it is possible that the bacteria underwent genomic changes (and the loss of important antigens) before Calmette and Guérin decided that virulence was completely attenuated. In our work, we kept samples obtained every 5 to 10 passages and can thus determine in future assays the times when the genetic changes occurred.

The large deletion in P-200 is located at nucleotide position 4,795,932. In comparing the genome of *N. brasiliensis* with those of other *Nocardia* species, we observed higher synteny in the first 2,000 Mb before and after the *oriC* site (*dnaJ* gene). Independently of genome size, bacteria try to protect the DNA around *oriC*, avoiding a loss of DNA material and the presence of transposons or insertion sequences that may introduce changes in this important region; this zone has been termed the nucleocore of the genome. It appears that during growth in a rich medium, such as BHI medium, *N. brasiliensis* can lose a considerable amount of DNA material outside the nucleocore. *N. brasiliensis* is a species of soil bacteria with a genome size more similar to those of other soil bacteria, such as *Actinomadura*, *Saccharopolyspora*, and *Streptomyces* (approximately 8 to 10 Mb) (24, 25), than to those of human pathogens (approximately 4 Mb). This large chromosome allows soil inhabitants to encode pathways necessary for growth by use of simple compounds as carbon and nitrogen sources, including even aromatic compounds, by use of enzymes such as protocatechuate or homogentisate oxidase (26). In prolonged culture in rich media, it seems that these enzymes are dispensable.

In addition to the loss of DNA, other biological changes occurred. The bacteria grew more rapidly, and they grew in the form of a bacterial suspension instead of the typical entangled filaments. Furthermore, the bacterial cell density decreased, possibly because of biochemical changes in the cell wall. In the host (mice), all these changes were reflected by an inability to produce subcutaneous lesions. With this new genomic information, we can now track the specific gene or genes responsible for *Nocardia* pathogenesis.

ACKNOWLEDGMENT

This work fulfills in part the requirements for a doctoral degree for Carolina Gonzalez-Carrillo.

FUNDING INFORMATION

This work, including the efforts of Lucio Vera-Cabrera, was funded by CONACYT (155892).

REFERENCES

1. **López-Martínez R, Méndez-Tovar LJ, Bonifaz A, Arenas R, Mayorga J, Welsh O, Vera-Cabrera L, Padilla-Desgarennes MC, Contreras Pérez C, Chávez G, Estrada R, Hernández-Hernández F, Manzano-Gayosso P.** 2013. Update on the epidemiology of mycetoma in Mexico. A review of 3933 cases. Gac Med Mex **149:**586–592.
2. **Brown-Elliott BA, Conville P, Wallace RJ, Jr.** 2006. Current status of nocardia taxonomy and recommended identification methods. Clin Microbiol Newsl **37:**25–32.
3. **Beaman BL, Black CM, Doughty F, Beaman L.** 1985. Role of superoxide dismutase and catalase as determinants of pathogenicity of *Nocardia asteroides*: importance in resistance to microbicidal activities of human polymorphonuclear neutrophils. Infect Immun **47:**135–141.
4. **Trevino-Villarreal JH, Vera-Cabrera L, Valero-Guillén PL, Salinas-Carmona MC.** 2012. *Nocardia brasiliensis* cell wall lipids modulate macrophage and dendritic responses that favor development of experimental actinomycetoma in BALB/c mice. Infect Immun **80:**3587–3601. http://dx.doi.org/10.1128/IAI.00446-12.
5. **Vera-Cabrera L, Ortiz-Lopez R, Elizondo-Gonzalez R, Ocampo-Candiani J.** 2013. Complete genome sequence analysis of *Nocardia brasiliensis* HUJEG-1 reveals a saprobic lifestyle and the genes needed for human pathogenesis. PLoS One **8:**e65425. http://dx.doi.org/10.1371/journal.pone.0065425.
6. **Meester I, Rosas-Taraco AG, Salinas-Carmona MC.** 2014. *Nocardia brasiliensis* induces formation of foamy macrophages and dendritic cells in vitro and in vivo. PLoS One **9:**e100064. http://dx.doi.org/10.1371/journal.pone.0100064.
7. **Salinas-Carmona MC, Rosas-Taraco AG, Welsh O.** 2012. Systemic increased immune response to *Nocardia brasiliensis* co-exists with local im-

munosuppressive microenvironment. Antonie Van Leeuwenhoek **102:** 473–480. http://dx.doi.org/10.1007/s10482-012-9779-y.

8. **González-Martínez NA, Lozano-Garza HG, Castro-Garza J, De Osio-Cortez A, Vargas-Villarreal J, Cavazos-Rocha N, Ocampo-Candiani J, Makarov V, Cole ST, Vera-Cabrera L.** 2015. In vivo activity of the benzothiazinones PBTZ169 and BTZ043 against *Nocardia brasiliensis*. PLoS Negl Trop Dis **9:**e0004022. http://dx.doi.org/10.1371/journal.pntd.0004022.
9. **Griffin JE, Gawronski JD, Dejesus MA, Ioerger TR, Akerley BJ, Sassetti CM.** 2011. High-resolution phenotypic profiling defines genes essential for mycobacterial growth and cholesterol catabolism. PLoS Pathog **7:**e1002251. http://dx.doi.org/10.1371/journal.ppat.1002251.
10. **Mollenkopf HJ, Jungblut PR, Raupach B, Mattow J, Lamer S, Zimny-Arndt U, Schaible UE, Kaufmann SH.** 1999. A dynamic two-dimensional polyacrylamide gel electrophoresis database: the mycobacterial proteome via Internet. Electrophoresis **20:**2172–2180. http://dx.doi.org/10.1002/(SICI)1522-2683(19990801)20:11<2172::AID-ELPS2172>3.0.CO;2-M.
11. **Geison GL.** 1978. Pasteur's work on rabies: reexamining the ethical issues. Hastings Cent Rep **8:**26–33.
12. **Pasteur L.** 1885. Méthode pour prévenir la rage après morsure. Comptes Rendus Acad Sci **101:**765–774.
13. **Calmette A, Guérin C.** 1920. Nouvelles recherches expérimentales sur la vaccination des bovidés contre la tuberculose. Ann Inst Pasteur **34:**553–560.
14. **Pym AS, Brodin P, Brosch R, Huerre M, Cole ST.** 2002. Loss of RD1 contributed to the attenuation of the live tuberculosis vaccines *Mycobacterium bovis* BCG and *Mycobacterium microti*. Mol Microbiol **46:**709–717. http://dx.doi.org/10.1046/j.1365-2958.2002.03237.x.
15. **Shah FA, Allen N, Wright CJ, Butt TM.** 2007. Repeated in vitro subculturing alters spore surface properties and virulence of *Metarhizium anisopliae*. FEMS Microbiol Lett **276:**60–66. http://dx.doi.org/10.1111/j.1574-6968.2007.00927.x.
16. **Songe MM, Thoen E, Evensen Ø, Skaar I.** 2014. In vitro passages impact on virulence of Saprolegnia parasitica to Atlantic salmon, Salmo salar L. parr. J Fish Dis **37:**825–834. http://dx.doi.org/10.1111/jfd.12175.
17. **Druelle J, Sellin CI, Waku-Kouomou D, Horvat B, Wild FT.** 2008. Wild type measles virus attenuation independent of type I IFN. Virol J **5:**22. http://dx.doi.org/10.1186/1743-422X-5-22.
18. **Barrick JE, Yu DS, Yoon SH, Jeong H, Oh TK, Schneider D, Lenski RE, Kim JF.** 2009. Genome evolution and adaptation in a long-term experiment with *Escherichia coli*. Nature **461:**1243–1247. http://dx.doi.org/10.1038/nature08480.
19. **Nilsson AI, Koskiniemi S, Eriksson S, Kugelberg E, Hinton JC, Andersson DI.** 2005. Bacterial genome size reduction by experimental evolution. Proc Natl Acad Sci U S A **102:**12112–12116. http://dx.doi.org/10.1073/pnas.0503654102.
20. **Medina E, Ryan L, LaCourse R, North RJ.** 2006. Superior virulence of *Mycobacterium bovis* over *Mycobacterium tuberculosis* (Mtb) for Mtb-resistant and Mtb-susceptible mice is manifest as an ability to cause extrapulmonary disease. Tuberculosis (Edinb) **86:**20–27. http://dx.doi.org/10.1016/j.tube.2005.04.003.
21. **World Health Organization.** 2004. BCG vaccine. WHO position paper. Wkly Epidemiol Rec **79:**27–38.
22. **Roy A, Eisenhut M, Harris RJ, Rodrigues LC, Sridhar S, Habermann S, Snell L, Mangtani P, Adetifa I, Lalvani A, Abubakar I.** 2014. Effect of BCG vaccination against *Mycobacterium tuberculosis* infection in children: systematic review and meta-analysis. BMJ **349:**g4643. http://dx.doi.org/10.1136/bmj.g4643.
23. **Behr MA, Small PM.** 1997. Has BCG attenuated to impotence? Nature **389:**133–134.
24. **Vera-Cabrera L, Ortiz-Lopez R, Elizondo-González R, Campos-Rivera MP, Gallardo-Rocha A, Molina-Torres CA, Ocampo-Candiani J.** 2014. Draft genome sequence of *Actinomadura madurae* LIID-AJ290, isolated from a human mycetoma case. Genome Announc **2:**e00201-14. http://dx.doi.org/10.1128/genomeA.00201-14.
25. **Oliynyk M, Samborskyy M, Lester JB, Mironenko T, Scott N, Dickens S, Haydock SF, Leadlay PF.** 2007. Complete genome sequence of the erythromycin-producing bacterium *Saccharopolyspora erythraea* NRRL23338. Nat Biotechnol **25:**447–453. http://dx.doi.org/10.1038/nbt1297.
26. **Zeinali M, Vossoughi M, Ardestani SK.** 2008. Degradation of phenanthrene and anthracene by *Nocardia otitidiscaviarum* strain TSH1, a moderately thermophilic bacterium. J Appl Microbiol **105:**398–406. http://dx.doi.org/10.1111/j.1365-2672.2008.03753.x.
27. **Almaguer-Chávez JA, Welsh O, Lozano-Garza HG, Said-Fernández S, Romero-Díaz VJ, Ocampo-Candiani J, Vera-Cabrera L.** 2011. Decrease of virulence for BALB/c mice produced by continuous subculturing of *Nocardia brasiliensis*. BMC Infect Dis **11:**290. http://dx.doi.org/10.1186/1471-2334-11-290.