



dnaJ: a New Approach to Identify Species within the Genus *Enterobacter*

Enrique Hernandez-Alonso,^{a,b} Simon Barreault,^{a,c} Luis A. Augusto,^a Pierre Jatteau,^b Millie Villet,^b Pierre Tissieres,^{a,c,d}
 Florence Doucet-Populaire,^{a,b} Nadege Bourgeois-Nicolaos^{a,b} on behalf of the SENSE Group

^aUniversité Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC), Gif-sur-Yvette, France

^bDepartment of Bacteriology-Hygiene, AP-HP Université Paris-Saclay, Antoine Beclere Hospital, Clamart, France

^cDepartment of Pediatrics Intensive Care and Neonatal Medicine, AP-HP Université Paris-Saclay, Bicêtre Hospital, Paris, France

^dFHU Sepsis, AP-HP/Université Paris-Saclay/Inserm, Le Kremlin-Bicêtre, France

ABSTRACT The taxonomy of the genus *Enterobacter* can be confusing and has been considerably revised in recent years. We propose a PCR and amplicon sequencing technique based on a partial sequence of the *dnaJ* gene for species assignment consistent with DNA-DNA digital hybridization (dDDH) and pairwise average nucleotide identity (ANI). We performed a validation of the method by comparing the type strains of each species, sequences obtained from the GenBank database, and clinical specimens. Our results show that the polymorphism of the target sequence of *dnaJ* allows the identification of species. Using this gene, we assigned the species to 100 strains deposited in the GenBank database that were consistent with the species assignment by dDDH and ANI. The analysis showed that using the partial *dnaJ* sequence is congruent with WGS as far as correct identification of *Enterobacter* species is concerned. Finally, we applied our *dnaJ* method on a national collection of 68 strains identified as *Enterobacter* isolated from the blood cultures of premature babies using an algorithm based on a type-strain library and the SeqScape software. For the first time, we identified *Enterobacter quasihormaechei* in blood cultures from four neonatal sepsis cases. We also noticed a higher prevalence of *E. bugandensis* (36.3%; 32/88) and *E. xiangfangensis* (46.5%; 41/88). *E. bugandensis* is a novel species recently described specifically in instances of neonatal sepsis. In conclusion, sequencing a part of the *dnaJ* gene could be a quick, more economical, and highly discriminating method of identifying *Enterobacter* species in clinical practice and research.

IMPORTANCE We propose a new approach for *Enterobacter* species identification based on the diversity of the gene encoding the heat shock protein DnaJ. This new tool can be easily implemented in clinical laboratories in addition to identification by MALDI-TOF.

KEYWORDS *Enterobacter bugandensis*, *Enterobacter cloacae*, *Enterobacter quasihormaechei*, *Enterobacter*, SeqScape software, *dnaJ* gene, gene PCR-sequencing, neonates, sepsis, species identification

Enterobacter species are increasingly involved in human infections particularly among critically ill patients with sepsis. The incidence of these often-nosocomial infections is particularly high in neonatology where they are facilitated by immune incompetence linked to the immaturity of premature infants. (1, 2). The taxonomy of the genus *Enterobacter* has been reevaluated in the last 10 years. Some species have been reclassified into other genera such as *Enterobacter aerogenes* which has been moved to the genus *Klebsiella* (3). Recently, Wu et al. suggested that *Enterobacter timonensis* should have been moved to a new genus called *Pseudenterobacter* (4). Moreover, the concept of subspecies in the genus *Enterobacter* has been a source of confusion and is no longer recommended (4). In recent years the use of tools such as digital DNA-DNA hybridization (dDDH) and pairwise average nucleotide identity (ANI) have allowed a more accurate classification of this genus. In particular, the

Editor N. Esther Babady, Memorial Sloan Kettering Cancer Center

Copyright © 2021 Hernandez-Alonso et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Florence Doucet-Populaire, florence.doucet-populaire@universite-paris-saclay.fr.

The authors declare no conflict of interest.

Received 11 August 2021

Accepted 19 November 2021

Published 22 December 2021

DNA-DNA hybridization (DDH) 70% cutoff is now regarded as the “gold standard” for species delineation (5). In practice, thresholds of 96.0% or 70.0% with ANI and dDDH (comparable with DDH for all intents and purposes) have been used, respectively, for species identification using Whole-Genome Sequencing (WGS) (6). The most recent taxonomic study of the genus *Enterobacter* according to the overall genome relatedness index (OGRI) using dDDH and ANI values between the type strains showed that *E. dissolvens*, *E. hoffmannii*, *E. hormaechei*, and *E. xiangfangensis* are independent species. Furthermore, *E. hormaechei* subsp. *oharae* and *E. hormaechei* subsp. *steigerwaltii* have been reclassified as *E. xiangfangensis* (4). Nowadays, the genus *Enterobacter* consists of 22 species (Table 1). In a recent study Wu et al. showed that *E. asburiae*, *E. bugandensis*, *E. cancerogenus*, *E. chengduensis*, *E. cloacae*, *E. hoffmannii*, *E. huaxiensis*, *E. kobei*, *E. ludwigii*, *E. quasiroggenkampii*, *E. roggkampii*, and *E. xiangfangensis* are frequently isolated from blood cultures (4).

Correct species identification within the genus *Enterobacter* is still a challenge. Matrix-assisted laser desorption ionization–time of flight mass spectrum (MALDI-TOF MS) is the major tool used in clinical laboratories for bacterial identification. However, species identification using MALDI-TOF MS may have its limits (7). In clinical laboratories, tests based on the phenotype and the MALDI-TOF MS usually identify *Enterobacter* isolates as *E. cloacae*, thereby neglecting the clinical importance of other species in the hospital environment (8). Nevertheless, a correct identification on the species-level is crucial for epidemiology, pathogenesis, diagnosis, treatment, prognosis, and prevention (9, 10). According to Wu et al.’s precise species study, among the 1,960 *Enterobacter* strains submitted in the GenBank, only 80 (14.8%) of the 540 genomes labeled *E. cloacae* actually belonged to these species (4). Several other species of the genus *Enterobacter* have also been associated with infections, particularly *E. bugandensis* in neonatal sepsis (11).

Although the WGS allows correct identification on the species-level, this tool is still far from routine in clinical practice. It requires highly trained personnel for data assembly and analysis, in addition to being too expensive to use in many countries. Faster and more affordable methods of the species identification correlated with WGS as regards to species identification, are essential in clinical microbiology.

The *dnaJ* housekeeping gene that encodes the heat shock protein 40 (Hsp40) has been shown to have a high discriminating power in *Enterobacterales* showing better resolution than 16S rRNA, *tuf*, *hsp60*, or *atpD* (12, 13). The *dnaJ* gene contains an immutable sequence within each species of the genus *Enterobacter* that can be used for accurate identification (14, 15).

In this study, we propose a PCR and amplicon sequencing method to correctly identify species within the genus *Enterobacter* using a partial sequence of the *dnaJ* gene. Our method provides a highly discriminating tool in clinical practice.

RESULTS

***dnaJ* polymorphism analysis in *Enterobacter* type strains.** The genomes of the 22 type strains were obtained from GenBank (Table 1). We extracted the partial sequence of the *dnaJ* gene from position 385 to 1098 (714 bp), a region previously described for the analysis of species in *Enterobacterales* (14). The DNA fragment is characterized by the presence of two highly variable regions but identical within the same species (positions 541 to 558 and positions 1009 to 1027). We performed the alignment of the 22 sequences of the type strains. The polymorphism within the sequence was visually inspected (Fig. 1). Visual analysis showed a variation from 3.8% to 19.3% (corresponding to 27–138 different nucleotides). Using the alignment of the partial sequences of the *dnaJ* gene of the 22 type strains, we identified a percentage of similarity ranging from 87.8% to 95.9% (Table S1). The phylogenetic tree based on the target sequence of the *dnaJ* gene showed that each species of the genus forms an independent branch separated from the others that allows for species discrimination (Fig. 2). In addition, we determined the synonymous and nonsynonymous substitutions for a better understanding of the evolutive molecular dynamics of the sequence (Table 1). In all type strains, the d_N/d_S ratio was well below 1, clearly indicating a negative selection.

***In silico* validation of species-level identification using *dnaJ*.** We downloaded the genome of the 100 *Enterobacter* strains of interest that were taxonomically corrected in

TABLE 1 Classification of the 22 type strains of the genus *Enterobacter* used in this study

Species (n = 22)	Accession	Type strain	Substitution rates ^a
<i>Enterobacter asburiae</i>	CP011863.1	JCM 6051	S = 183.3 N = 530.7 $d_s = 0.2624$ $d_N = 0.0039$ $d_N/d_s = 0.0147$
<i>Enterobacter bugandensis</i>	LT992502.1	EB-247	S = 187.8 N = 526.2 $d_s = 0.2265$ $d_N = 0.0019$ $d_N/d_s = 0.0086$
<i>Enterobacter cancerogenus</i>	ERR1854846	ATCC 35316	S = 190.2 N = 523.8 $d_s = 0.3061$ $d_N = 0.0098$ $d_N/d_s = 0.0322$
<i>Enterobacter chengduensis</i>	CP043318.1	WCHECI-C4	S = 175.8 N = 538.2 $d_s = 0.2814$ $d_N = 0.0038$ $d_N/d_s = 0.0136$
<i>Enterobacter chuandaensis</i>	GCF_003594915.1	090028	S = 213.6 N = 500.4 $d_s = 0.2007$ $d_N = 0.0021$ $d_N/d_s = 0.0103$
<i>Enterobacter cloacae</i>	CP001918.1	ATCC 13047	S = 186.7 N = 527.3 $d_s = 0.3218$ $d_N = 0.0019$ $d_N/d_s = 0.0060$
<i>Enterobacter dissolvens</i>	WJWQ01000001.1	ATCC 23373	S = 192.5 N = 521.5 $d_s = 0.3190$ $d_N = 0.0020$ $d_N/d_s = 0.0061$
<i>Enterobacter hoffmannii</i>	CP017186.1	DSM 14563	S = 180.2 N = 533.8 $d_s = 0.2423$ $d_N = 0.0019$ $d_N/d_s = 0.0079$
<i>Enterobacter hormaechei</i>	MKEQ01000001.1	ATCC 49162	S = 182.6 N = 531.4 $d_s = 0.2536$ $d_N = 0.0019$ $d_N/d_s = 0.0076$
<i>Enterobacter huaxiensis</i>	QZCT01000001.1	090008	S = 181.6 N = 532.4 $d_s = 0.3211$ $d_N = 0.0097$ $d_N/d_s = 0.0301$
<i>Enterobacter kobei</i>	CP017181.1	ATCC BAA-260	S = 189.0 N = 525.0 $d_s = 0.2929$ $d_N = 0.0019$ $d_N/d_s = 0.0066$
<i>Enterobacter ludwigii</i>	CP017279.1	EN-119	S = 185.6 N = 528.4 $d_s = 0.3056$ $d_N = 0.0039$ $d_N/d_s = 0.0126$
<i>Enterobacter mori</i>	AEXB00000000.1	LMG 25706	S = 189.0 N = 525.0

(Continued on next page)

TABLE 1 (Continued)

Species (<i>n</i> = 22)	Accession	Type strain	Substitution rates ^a
<i>Enterobacter oligotrophica</i>	AP019007.1	CCA6	$d_s = 0.1990$
			$d_N = 0.0039$
			$d_N/d_s = 0.0198$
			$S = 183.8$
			$N = 530.2$
<i>Enterobacter quasihormaechei</i>	SJON01000001.1	WCHEs120003	$d_s = 0.2296$
			$d_N = 0.0039$
			$d_N/d_s = 0.0169$
			$S = 181.4$
			$N = 532.6$
<i>Enterobacter quasimori</i>	RXRX00000000.1	090044	$d_s = 0.2401$
			$d_N = 0.0058$
			$d_N/d_s = 0.0242$
			$S = 181.1$
			$N = 532.9$
<i>Enterobacter quasihormaechei</i>	LFDQ00000000.2	WCHECL1060	$d_s = 0.2380$
			$d_N = 0.0019$
			$d_N/d_s = 0.0081$
			$S = 191.5$
			$N = 522.5$
<i>Enterobacter roggkampii</i>	CP017184.1	DSM 16690	$d_s = 0.2672$
			$d_N = 0.0020$
			$d_N/d_s = 0.0074$
			$S = 191.4$
			$N = 522.6$
<i>Enterobacter sichuanensis</i>	POVL01000001.1	WCHECI1597	$d_s = 0.2424$
			$d_N = 0.0020$
			$d_N/d_s = 0.0081$
			$S = 194.7$
			$N = 519.3$
<i>Enterobacter soli</i>	LXE501000001.1	ATCC BAA-2102	$d_s = 0.3025$
			$d_N = 0.0020$
			$d_N/d_s = 0.0065$
			$S = 202.9$
			$N = 511.1$
<i>Enterobacter wuhouensis</i>	S JOO01000001.1	WCHEs120002	$d_s = 0.4126$
			$d_N = 0.0101$
			$d_N/d_s = 0.0244$
			$S = 187.3$
			$N = 526.7$
<i>Enterobacter xiangfangensis</i>	CP017183.1	LMG 27195	$d_s = 0.2919$
			$d_N = 0.0059$
			$d_N/d_s = 0.0201$
			$S = 191.6$
			$N = 522.4$
			$d_s = 0.1984$
			$d_N = 0.0020$
			$d_N/d_s = 0.0099$

^aThe parameters were calculated using the web server PAL2NAL (18). *S*, number of synonymous sites; *N*, number of nonsynonymous sites; d_s , synonymous substitution rate; d_N , nonsynonymous substitution rate.

2020 to determine the relationship between their genomes and that of the type strains by dDDH (Table S2). We found an interval between 74.4% and 100% as did other studies (3). Next, we extracted the 714 bp partial sequence of the *dnaJ* gene from those 100 strains and established the phylogenetic relationship in conjunction with the 22 *dnaJ* partial sequences of the type strains. The resulting phylogenetic tree showed separate branches for each species as was the case with dDDH (Fig. 3).

Species-identification using PCR and amplification sequencing of *dnaJ* is consistent with the results of the ANI and dDDH methods. First, we used the same 20 *Enterobacter* clinical strains previously sequenced by WGS in our laboratory (BioProject access number: PRJNA770343) in order to perform the PCR and amplicon sequencing of the *dnaJ* gene. After amplification, the 714 bp amplicon of each strain was sequenced, and the alignment

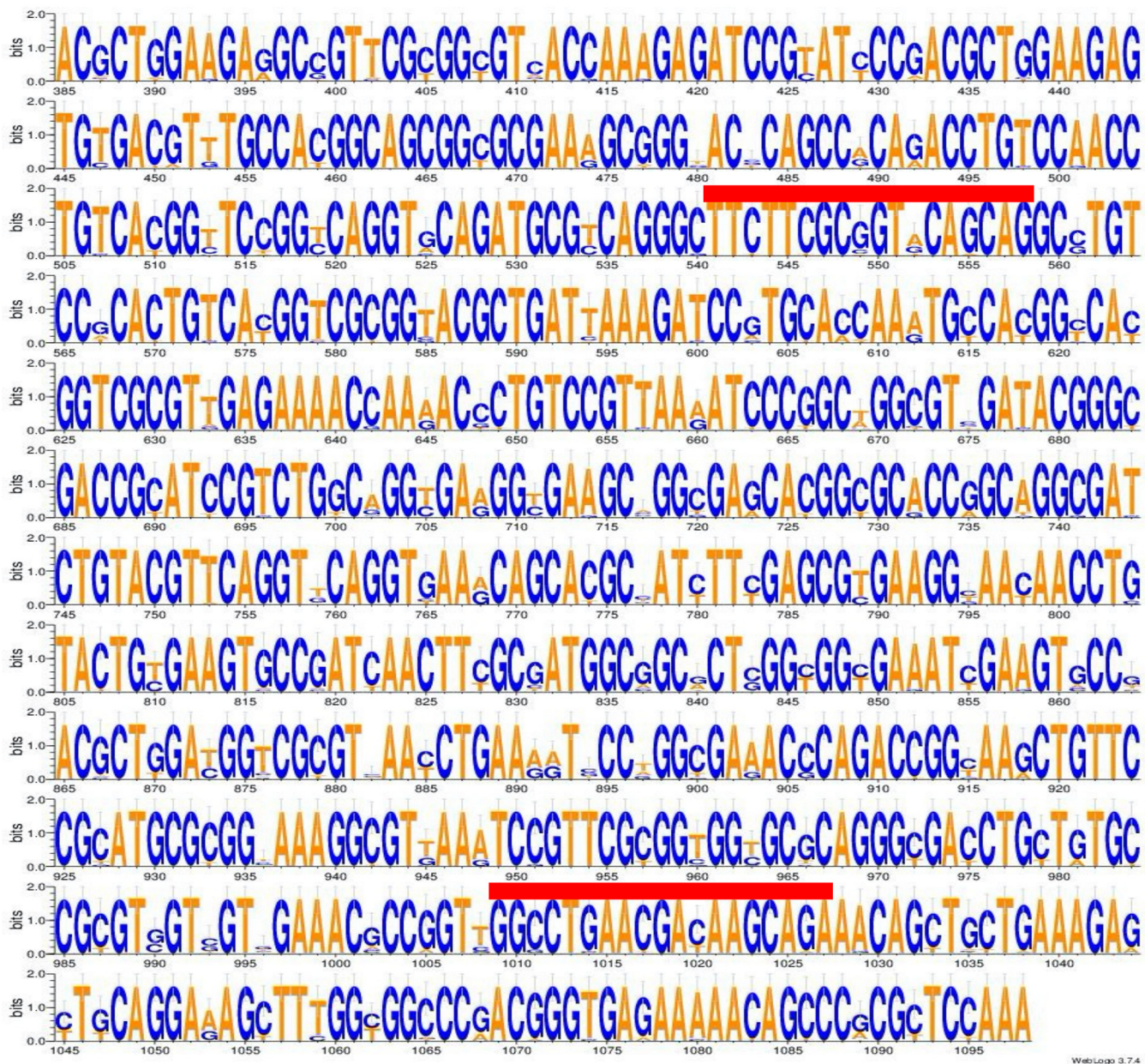


FIG 1 Logo plot of partial sequence of the *dnaJ* gene from position 385 to 1098 (714 bp). Sequence sets are shown relative to a *dnaJ* sequence of 22 different type strains of the genus *Enterobacter* (Table S2). The DNA fragment is characterized by the presence of two regions highly variable but conserved intraspecies (positions 541 to 558 and positions 1009 to 1027, showed by red line) described previously (14). All DNA sequences are listed 5' to 3'.

and construction of the phylogenetic tree was carried out in conjunction with the type strains. The tree shows a clear distribution of the different species based on the sequence of the *dnaJ* gene (Fig. 4). According to the results of the phylogenetic analysis, the species were assigned as *E. bugandensis*, *E. cancerogenus*, *E. cloacae*, *E. hoffmannii*, *E. quasihormaechei*, and *E. xiangfangensis* (Table 2).

Second, we used the WGS data to confirm the species using ANI and dDDH values. We identified a range from 98.4% to 99.9% and 87.0% to 96.8% for ANI and dDDH, respectively (Table 2). The results showed a correct species assignment using *dnaJ* confirmed by the ANI and dDDH methods for 100% of the strains studied.

Application of the *dnaJ* target sequencing to clinical specimens. First, to verify the usefulness of this new technique based on the PCR and amplicon sequencing of the target *dnaJ* sequence as a diagnostic tool, we retrospectively analyzed 68 blood cultures isolates (named SENSE1-68) previously identified using MALDI-TOF MS as *E.*

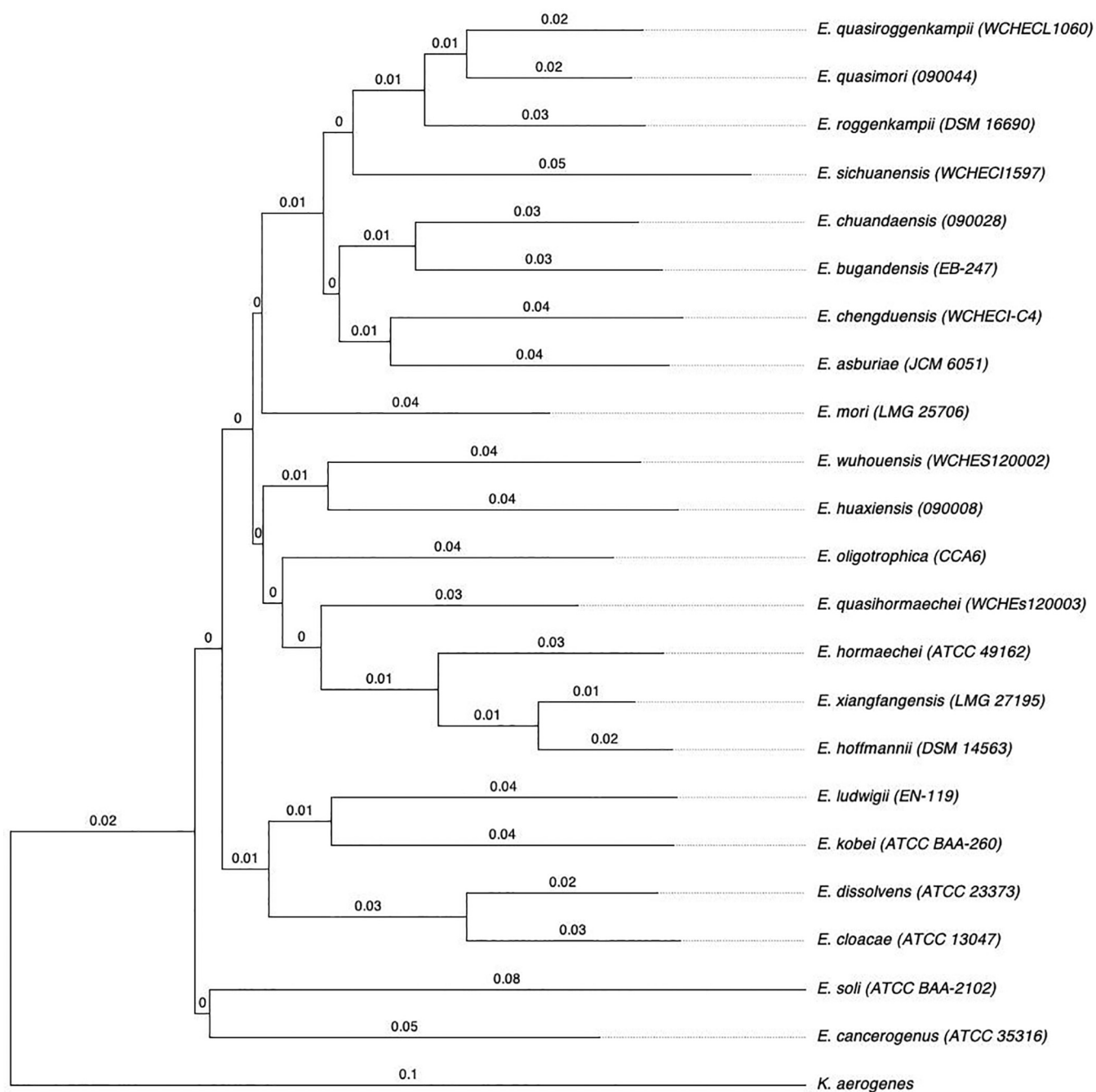


FIG 2 Molecular phylogenetic analysis of the partial sequence of the *dnaJ* gene of the type strains. The tree was inferred using the Neighbor-Joining method under the Tamura–Nei model with a 1000-bootstrap test. The analysis involved 22 partial nucleotide sequences (714 bp) of the *dnaJ* gene. The tree was rooted with the *dnaJ* sequence of *K. aerogenes* download from GenBank (AB008141.1). The *dnaJ* gene partial sequences of the type strains used are available in the Table S2.

cloacae complex. We found 9 different species: *E. bugandensis* ($n = 24$), *E. cancerogenus* ($n = 1$), *E. dissolvens* ($n = 2$), *E. hoffmannii* ($n = 7$), *E. kobei* ($n = 2$), *E. ludwigii* ($n = 1$), *E. mori* ($n = 1$), *E. quasihormaechei* ($n = 3$), and *E. xiangfangensis* ($n = 27$). The most prevalent was *E. xiangfangensis* (40.0%) followed by *E. bugandensis* (35.0%) (Fig. 5).

Second, we developed a local library using all the 22 partial sequences of *dnaJ* of the type strains in order to implement this method in the microbiology department of our hospital. This library was uploaded in the SeqScape software that provides library functions for comparison to a known group of sequences. After Sanger sequencing of partial sequence of *dnaJ* of 68 *Enterobacter* clinical strains from a national collection

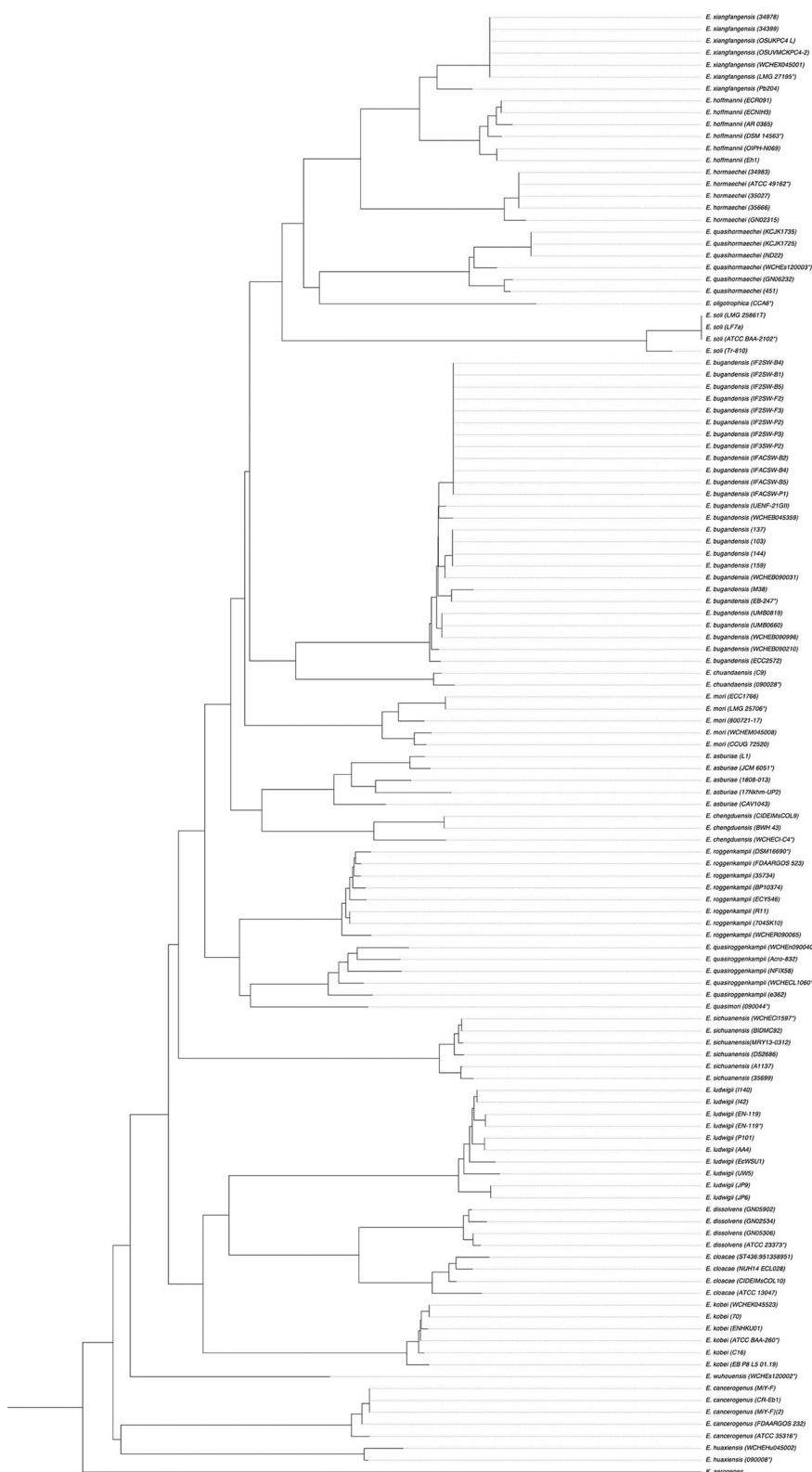


FIG 3 Molecular phylogenetic analysis of the partial sequence of the *dnaJ* gene between GenBank strains with the type strains (*). The tree was inferred using the Neighbor-Joining method under the Tamura–Nei model with a 1000-bootstrap test. The analysis involved 122 partial nucleotide sequences (714 bp) of the *dnaJ* gene. The tree was rooted with the *dnaJ* sequence of *K. aerogenes* download from GenBank (AB008141.1). All 122 *dnaJ* gene partial sequences are available in the Table S2.

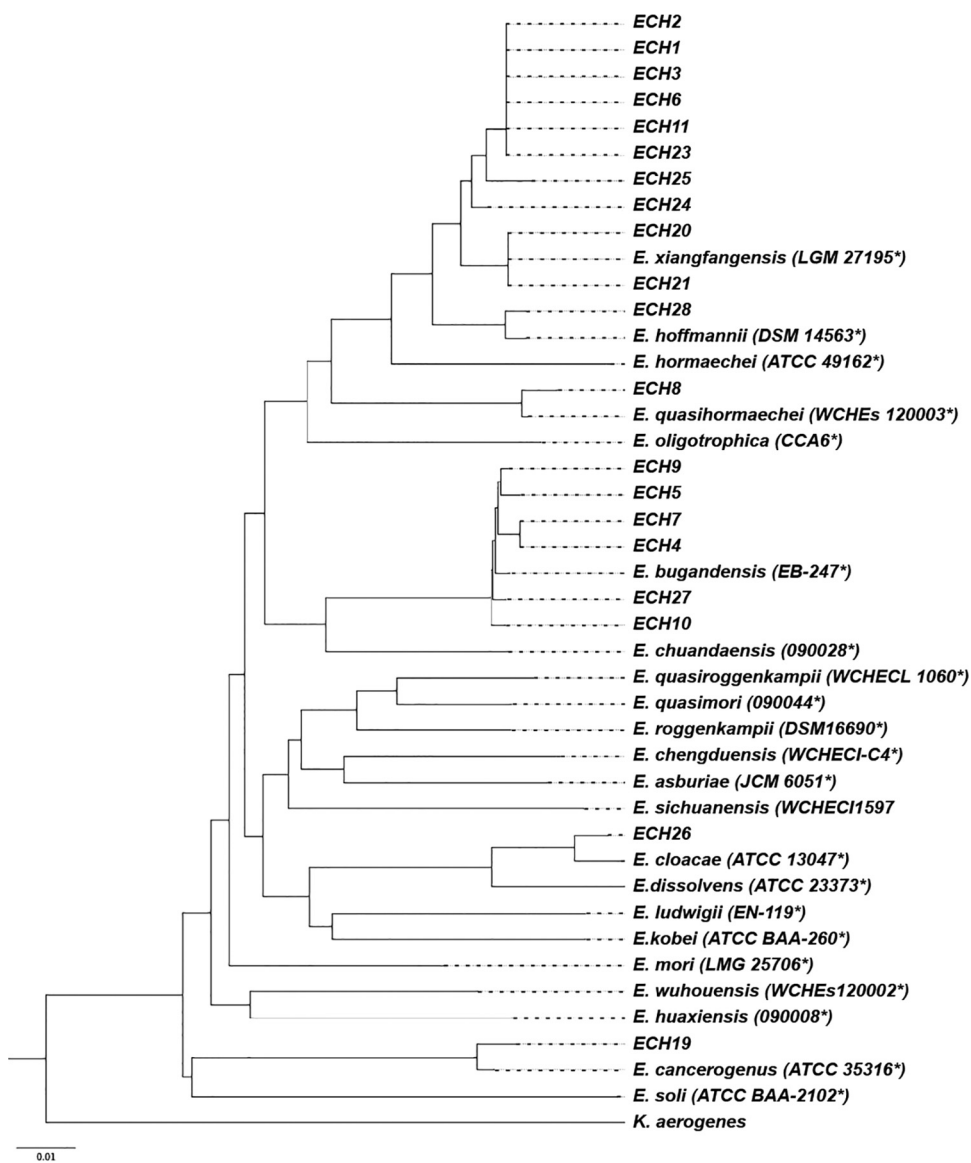


FIG 4 Molecular phylogenetic analysis of the partial sequence of the *dnaJ* gene of type strains and 20 relevant clinical strains from blood culture sequenced by WGS. The tree was inferred using the Neighbor-Joining method under the Tamura–Nei model with a 1000-bootstrap test. The analysis involved 42 partial nucleotide sequences (714 bp) of the *dnaJ* gene. The tree was rooted with the *dnaJ* sequence of *K. aerogenes* download from GenBank (AB008141.1). The *dnaJ* gene partial sequences of the type strains (*) used are available in table Supplementary text file 1 and the 20 clinical strains sequences are available in GenBank with the access number from MT665008.1 to MT665027.1.

(16) all sequencing files were compared with our library. All the 68 strains (100.0%) matched with the species previously identified by phylogenetic analysis.

DISCUSSION

The correct identification of bacterial isolates is crucial to patient care. It is essential in predicting the clinical prognosis as well as in choosing an adapted antibiotic regimen. However, for the genus *Enterobacter*, species identification is complicated. The traditional methods used in clinical laboratories, fail to correctly identify the species. Therefore, other methods are needed. Pavlovic et al. observed that MALDI-TOF MS was inadequate to differentiate *E. asburiae*, *E. hormaechei*, *E. kobei*, and *E. ludwigii* from *E. cloacae* (15). We developed a method based on a partial sequence of the *dnaJ* gene for species identification within the genus *Enterobacter*.

TABLE 2 Comparison of species identification of *Enterobacter* clinical strains by MALDI-TOF MS and *dnaJ* gene sequencing as well as OGRI

Strain	MALDI-TOF MS		<i>dnaJ</i> Organism identified ^a	OGRI	
	Organism identified	Score		ANI	dDDH
ECH1	<i>E. cloacae</i>	2.12	<i>Enterobacter xiangfangensis</i>	99.00	91.70
ECH2	<i>E. cloacae</i>	2.09	<i>Enterobacter xiangfangensis</i>	98.98	91.00
ECH3	<i>E. cloacae</i>	2.08	<i>Enterobacter xiangfangensis</i>	99.01	91.60
ECH4	<i>E. cloacae</i>	2.19	<i>Enterobacter bugandensis</i>	98.69	92.70
ECH5	<i>E. asburiae</i>	2.00	<i>Enterobacter bugandensis</i>	98.53	95.80
ECH6	<i>E. cloacae</i>	2.23	<i>Enterobacter xiangfangensis</i>	99.99	91.60
ECH7	<i>E. cloacae</i>	2.05	<i>Enterobacter bugandensis</i>	98.68	93.80
ECH8	<i>E. cloacae</i>	2.14	<i>Enterobacter quasihormaechei</i>	98.76	95.80
ECH9	<i>E. asburiae</i>	2.02	<i>Enterobacter bugandensis</i>	98.70	95.40
ECH10	<i>E. asburiae</i>	2.11	<i>Enterobacter bugandensis</i>	98.37	96.80
ECH11	<i>E. cloacae</i>	2.12	<i>Enterobacter xiangfangensis</i>	99.00	91.40
ECH19	<i>E. cancerogenus</i>	2.12	<i>Enterobacter cancerogenus</i>	99.14	91.30
ECH20	<i>E. cloacae</i>	2.24	<i>Enterobacter xiangfangensis</i>	99.26	93.90
ECH21	<i>E. cloacae</i>	2.21	<i>Enterobacter xiangfangensis</i>	99.24	94.00
ECH23	<i>E. cloacae</i>	2.16	<i>Enterobacter xiangfangensis</i>	98.99	91.20
ECH24	<i>E. cloacae</i>	2.18	<i>Enterobacter xiangfangensis</i>	99.03	91.70
ECH25	<i>E. cloacae</i>	2.18	<i>Enterobacter xiangfangensis</i>	99.19	93.50
ECH26	<i>E. cloacae</i>	2.09	<i>Enterobacter cloacae</i>	98.68	88.90
ECH27	<i>E. kobei</i>	2.06	<i>Enterobacter bugandensis</i>	98.44	87.00
ECH28	<i>E. cloacae</i>	2.16	<i>Enterobacter hoffmannii</i>	98.44	93.30

^aSpecies assignment according to the most recent classification. OGRI, overall genome relatedness index; ANI, average nucleotide identity; dDDH, digital DNA-DNA hybridization.

In silico analysis of the partial sequence of the *dnaJ* gene of the type strains showed a high interspecies polymorphism verified by the analysis of the phylogenetic evolution of the partial sequence of the genus *Enterobacter* ($d_N/d_S < 1$). We observed that the polymorphism of the *dnaJ* target sequence allows precise *Enterobacter* species identification.

As dDDH values have often been used in taxonomic studies, 100 recent genomes of different *Enterobacter* species were compared with the type strains to verify the correct species assignment. We observed a range from 74.4% to 100.0%. Our results were consistent with the Wu et al. study (4). We used these genome data to extract the partial sequence of the *dnaJ* gene for phylogenetic analysis and species assignment. The results show 100% accurate species assignment of the strains analyzed using *dnaJ* verified by dDDH.

A phylogenetic analysis based on the *dnaJ* gene supports the idea that, using only this partial sequence, it is possible to obtain correct species identification in the genus *Enterobacter* (Fig. 2 and 3). The method described here allows for a correct and rapid species identification as valid as OGRI.

Using our PCR-sequencing *dnaJ* method, we observed that 18/20 (86.0%) of the whole-genome sequenced *Enterobacter* clinical strains were misidentified by MALDI-TOF MS. Only one strain among 15 (6.0%) corresponded to *E. cloacae*. 3 *E. asburiae* and 1 *E. kobei* were identified as *E. bugandensis*, 11 *E. xiangfangensis* were misidentified as *E. cloacae*, and one *E. cloacae* was identified as *E. hoffmannii* (Table 2). Those results were correlated with OGRI analysis and dDDH.

Then we applied our method on a historical sampling of 68 blood culture *Enterobacter* isolates from neonates. We found a high prevalence (40.0%) of *E. xiangfangensis*. We also identified *E. quasihormaechei*, a novel species reported in 2020 (17). *E. quasihormaechei* could share the same virulence and pathogenicity mechanisms as *E. xiangfangensis*, previously identified in other studies as *E. hormaechei*, a species that often exhibits resistance to multiple commonly used antibiotics and persists in nosocomial environments (18). In this study, we report for the first time to our knowledge, four *E. quasihormaechei* isolates from blood cultures in patients with neonatal sepsis.

We also detected a high prevalence of *E. bugandensis*, another novel enterobacterial species associated with severe clinical infections, particularly with neonatal sepsis, suggesting that this element of the genus *Enterobacter* is an emergent species in Europe

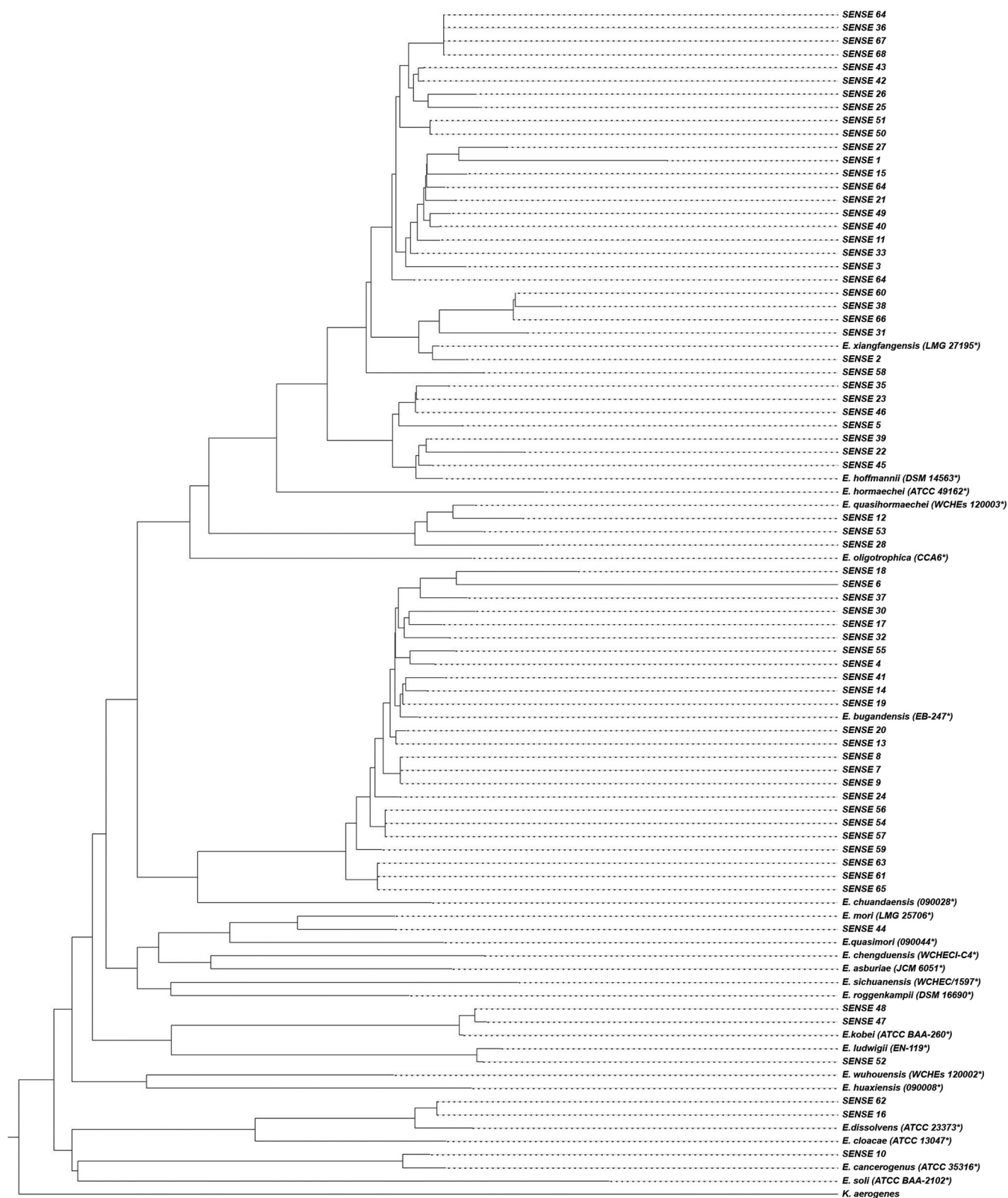


FIG 5 Molecular phylogenetic analysis of the partial sequence of the *dnaJ* gene between SENSE strains with the type strains (*). The tree was inferred using the Neighbor-Joining method under the Tamura-Nei model with a 1000-bootstrap test. The analysis involved 68 partial nucleotide sequences (714 bp) of the *dnaJ* gene. The tree was rooted with the *dnaJ* sequence of *K. aerogenes* download from GenBank (AB008141.1).

(19, 20). *E. bugandensis* is considered to be a highly pathogenic *Enterobacter* species with specific virulence mechanisms and usually a multidrug resistant profile (7, 10). Our preliminary analysis showed that our isolates identified as *E. bugandensis* are susceptible to third generation cephalosporins revealing an antibiotic resistance profile different from that previously reported (data not shown) (19).

Several epidemiological studies reported a high prevalence of *E. cloacae* in clinical isolates which is contrary to our findings where *E. cloacae* is less prevalent. The high prevalence of *E. cloacae* in other studies could be explained by species misidentification in the genus *Enterobacter* (21–23). In clinical practice bacterial identification is necessary for a better understanding of pathogenesis, virulence, and resistance to antibiotics. The fact is that identifying *Enterobacter* species has considerable medical significance since each species behaves differently: for instance, some are more likely to lead to nosocomial outbreaks while other are more prone to develop resistance to antibiotics. For example, in a recent study, we showed that *E. bugandensis* is more virulent in neonates than other species due to the presence of specific factors of virulence (24). Moreover, *E. xiangfangensis* and *E. steigerwaltii* have the highest prevalence of resistance of third generation cephalosporins and these two species were found to be especially prone to produce carbapenemases (25). In the future, the correct identification of species of the genus *Enterobacter* will also improve the understanding of their epidemiology in the clinical environment and will permit implementation of better strategies in the prevention of health care-associated infections. It will also improve patients' prognosis and treatment.

In conclusion, our method, based on a partial *dnaJ* gene PCR and amplicon sequencing, could be used in clinical practice as a specific, rapid, and highly discriminating tool, with similar results to those obtained using OGRI methods, for the correct identification of *Enterobacter* species. This improvement in correct identification can also facilitate the management of nosocomial outbreaks and promote rapid environmental monitoring.

MATERIALS AND METHODS

Phylogenetic and *in silico* analysis of the partial sequence of the *dnaJ* gene in type strains. The 22 type strain genomes were downloaded from GenBank for *in silico* analysis. We aligned the *dnaJ* gene (1146 bp) and then extracted a partial sequence from position 385 to 1098 (714 bp). The alignment and pairwise distancing were performed using the MUSCLE method (26). The Neighbor-Joining phylogenetic tree was constructed with the Tamura–Nei model and a 1000-bootstrap test using MegAlign Pro (DNASTAR) and MEGAX (27). We determined the numbers of synonymous (S) and nonsynonymous (N) sites as well as the rates of synonymous (d_s) and nonsynonymous (d_n) substitutions using the web site <http://abacus.gene.ucl.ac.uk/software/paml.html> (28).

Determination of overall genome relatedness. The dDDH among 100 genomes selected from GenBank (Table S1) and type strains of *Enterobacter* species were determined using the web-service <http://ggdc.dsmz.de> (formula 2). The pairwise ANI and dDDH among 20 clinically relevant bacterial isolates (Table 2) and type strains of *Enterobacter* species were determined using the web-service <http://enve-omics.ce.gatech.edu/ani/> for ANI and the web-service GGDC for dDDH. A $\geq 70\%$ dDDH or a $\geq 96\%$ ANI value was used as the cutoff to species assignment (29). The genome of the 20 clinically relevant strains conveniently used in this study have been deposited at DDBJ/ENA/GenBank under the BioProject accession number PRJNA770343.

Designing of primers, PCR, and amplicon sequencing of the partial *dnaJ* gene. Degenerate primers Hsp40-Fw (5'-GACCTGCGCTACAACATGGAKCT-3') and Hsp40-Rv (5'-CCGCGYTCCAAAAGCTTCTTYGAT-3') were visually designed and analyzed (OligoAnalyzer) according to the alignment of the *dnaJ* gene sequences of the type strains. The primers were used to amplify a fragment of 750 bp. The amplification reaction was performed with Master Mix GoTaq (PROMEGA, USA), with 0.5 μ l of each primer (0.2 μ M final PCR concentration), and 2.5 μ l of DNA template in a final reaction volume of 25 μ l. PCR amplification was carried out in a thermal cycler (VERITI, Applied Biosystems) as follows: 4 min denaturation step at 94°C, followed by 30 cycles at 94°C for 50 s, 60°C for 35 s, and 72°C for 1 min, with a final extension step of 5 min at 72°C. PCR products were verified by agarose gel electrophoresis, purified using WizardSV gel and PCR clean Up System (Promega), and then sequenced using the Sanger sequencing technology with Hsp40-Fw/Hsp40-Rv primers by BigDye Terminator Cycle Sequencing Ready Reaction kit (Applied Biosystems, USA) according to the manufacturer's instructions. After sequencing, primer sequences were removed and partial sequences of the *dnaJ* gene (714 bp) were confirmed by at least two chromatograms (forward and reverse) for the phylogenetic analysis.

Bacterial strains. Eighty-eight *Enterobacter* spp. strains isolated from blood cultures of neonate patients with sepsis were included. The strains were divided in 2 groups. The first 20 strains, already sequenced by WGS were used, to validate our method (BioProject access number: PRJNA770343) (24).

The second group included 60-eight *Enterobacter* spp. clinical strains collected from 8 French neonatal intensive care units (NICUs) known as the SENSE group between 2016 to 2019. They were grown on tryptic soy agar at 37°C for 24 h in aerobic conditions, and all isolates were identified twice as *E. cloacae*

complex in the respective hospitals and in our laboratory by MALDI-TOF MS (Bruker, Leipzig, Germany). Microbial suspension at 1 on the McFarland scale was employed and 200 μ l were used for DNA extraction following the easyMAG protocol (bioMérieux, France) according to the manufacturer's instructions.

Species identification in clinical practice. We developed a local library using the partial sequence of *dnaJ*. This library allows comparison of the *dnaJ* sequence of clinical isolates with the sequences of the type strains. The local library was uploaded in the SeqScape software v4.0 (Applied Biosystems, Courtaboeuf, France). The analyze protocol was generated according to the manufactured recommendations. After Sanger sequencing, the amplicon sequences were analyzed with SeqScape Software v4.0. We employed original sequencing files and corrected sequencing files. The species identification of a submitted clinical isolate sequence was given when the software algorithm assigns to each possible way of arrangement a score, which is defined as the standard measure of sequence similarity within a set of aligned sequences (30). A score of 800 was considered a cutoff value corresponding to 98.0–100.0% of homology between type strains and clinical isolates (Fig. S1).

Data availability. Sequence data generated by this study is available at BioProject PRJNA770343 (Whole Genome Shotgun project has been deposited at DDBJ/ENA/GenBank. The version described in this paper is version JAJAPC000000000-JAJAPV000000000).

The GenBank accession numbers for the sequences reported in this paper for the *dnaJ* gene of these 20 *Enterobacter* isolates range from MT665008.1 to MT665027.1.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

SUPPLEMENTAL FILE 1, PDF file, 0.4 MB.

SUPPLEMENTAL FILE 2, XLSX file, 0.03 MB.

ACKNOWLEDGMENTS

This work was supported by AP-HP and the Consejo Nacional de Ciencia y Tecnología, México with a fellowship to EHA from CONACYT (2019-000004-01EXTF-00028). The founders had no role in study design, interpretation, or the decision to submit the work for publication.

We thank the members of the SENSE Group: K. Allouche and D. De Luca (Antoine-Béclère, Clamart), S. Corvec and C. Flamant (Nantes), E. Bille and A. Lapillonne (Necker-Enfants malades, Paris), S. Bonacorci and L. Chafa (Robert-Debré, Paris), C. Chaplain and P. Bolot (Saint-Denis), O. Dawalder and M. Butin (Lyon), N. Fortineau and N. Le Saché (Le Kremlin-Bicêtre), F. Garnier and A. Bedu (Limoges), P. Lanotte and G. Favrais (Tours), R. Ruimy and S. Eleni Dit Trolli (Nice), M.N. Noulard and L. Desnoullez-Osselin (Arras), T. Guillard and N. Bednarek Weirauch (Reims).

We also thank Peggy Hanemann-Castex for English editing.

The study was approved by ethical committee of the French Society of Intensive Care (CE SRLF 19–40).

REFERENCES

- Kasim K, El Sadak AA, Zayed K, Abdel-Wahed A, Mosaad M. 2014. Nosocomial infections in a neonatal intensive care unit. *Middle-East J Sci Res* 19.
- Davin-Regli A, Lavigne JP, Pagès JM. 2019. *Enterobacter* spp.: update on taxonomy, clinical aspects, and emerging antimicrobial resistance. *Clin Microbiol Rev* 32:1–32. <https://doi.org/10.1128/CMR.00002-19>.
- Tindall BJ, Sutton G, Garrity GM. 2017. *Enterobacter aerogenes* Hormaeche and Edwards 1960 (approved lists 1980) and *Klebsiella mobilis* Bascomb et al. 1971 (Approved Lists 1980) share the same nomenclatural type (ATCC 13048) on the Approved Lists and are homotypic synonyms, with consequences for the name *Klebsiella mobilis* Bascomb et al. 1971 (approved lists 1980). *Int J Syst Evol Microbiol* 67:502–504. <https://doi.org/10.1099/ijsem.0.001572>.
- Wu W, Feng Y, Zong Z. 2020. Precise Species identification for *Enterobacter*: a genome sequence-based study with reporting of two novel species, *Enterobacter quasiroggenkampii* sp. nov. and *Enterobacter quasimori* sp. nov. *mSystems* 5:e00527-20. <https://doi.org/10.1128/mSystems.00527-20>.
- Richter M, Rosselló-Móra R. 2009. Shifting the genomic gold standard for the prokaryotic species definition. *Proc Natl Acad Sci U S A* 106:19126–19131. <https://doi.org/10.1073/pnas.0906412106>.
- Ciufo S, Kannan S, Sharma S, Badretidin A, Clark K, Turner S, Brover S, Schoch CL, Kimchi A, DiCuccio M. 2018. Using average nucleotide identity to improve taxonomic assignments in prokaryotic genomes at the NCBI. *Int J Syst Evol Microbiol* 68:2386–2392. <https://doi.org/10.1099/ijsem.0.002809>.
- De Florio L, Riva E, Giona A, Dedej E, Fogolari M, Cella E, Spoto S, Lai A, Zehender G, Ciccozzi M, Angeletti S. 2018. MALDI-TOF MS identification and clustering applied to *Enterobacter* species in nosocomial setting. *Front Microbiol* 9:1885. <https://doi.org/10.3389/fmicb.2018.01885>.
- Mshana SE, Gerwing L, Minde M, Hain T, Domann E, Yamuya E, Chakraborty T, Imirzalioglu C. 2011. Outbreak of a novel *Enterobacter* sp. carrying blaCTX-M-15 in a neonatal unit of a tertiary care hospital in Tanzania. *Int J Antimicrob Agents* 38:265–269. <https://doi.org/10.1016/j.ijantimicag.2011.05.009>.
- Tamburini FB, Andermann TM, Tkachenko E, Senchyna F, Banaei N, Bhatt AS. 2018. Precision identification of diverse bloodstream pathogens in the gut microbiome. *Nat Med* 24:1809–1814. <https://doi.org/10.1038/s41591-018-0202-8>.
- Campos LC, Lobianco LF, Seki LM, Santos RMR, Asensi MD. 2007. Outbreak of *Enterobacter hormaechei* septicaemia in newborns caused by contaminated parenteral nutrition in Brazil. *J Hosp Infect* 66:95–97. <https://doi.org/10.1016/j.jhin.2007.02.013>.
- Falgenhauer J, Imirzalioglu C, Falgenhauer L, Yao Y, Hauri AM, Erath B, Schwengers O, Goesmann A, Seifert H, Chakraborty T, Dojjad S. 2019. Whole-genome sequences of clinical *Enterobacter bugandensis* isolates from Germany. *Microbiol Resour Announc* 8. <https://doi.org/10.1128/MRA.00465-19>.
- McLean K, Rosenthal CA, Sengupta D, Owens J, Cookson BT, Hoffman NG, Salipante SJ. 2019. Improved species-level clinical identification of *Enterobacteriaceae* through broad-range *dnaJ* PCR and sequencing. *J Clin Microbiol* 57:e00986-19. <https://doi.org/10.1128/JCM.00986-19>.

13. Hoffmann H, Roggenkamp A. 2003. Population genetics of the nomenspecies *Enterobacter cloacae*. *Appl Environ Microbiol* 69:5306–5318. <https://doi.org/10.1128/AEM.69.9.5306-5318.2003>.
14. Pham HN, Ohkusu K, Mishima N, Noda M, Monir Shah M, Sun X, Hayashi M, Ezaki T. 2007. Phylogeny and species identification of the family *Enterobacteriaceae* based on *dnaJ* sequences. *Diagn Microbiol Infect Dis* 58: 153–161. <https://doi.org/10.1016/j.diagmicrobio.2006.12.019>.
15. Pavlovic M, Konrad R, Iwobi AN, Sing A, Busch U, Huber I. 2012. A dual approach employing MALDI-TOF MS and real-time PCR for fast species identification within the *Enterobacter cloacae* complex. *FEMS Microbiol Lett* 328: 46–53. <https://doi.org/10.1111/j.1574-6968.2011.02479.x>.
16. Hernandez-Alonso E, Barreault S, Allouche K, Jatteau P, Augusto L, Tissieres P, Doucet-Populaire F, Bourgeois-Nicolaos N, and the SENSE study group. 2021. *Enterobacter cloacae* complex associated with neonatal sepsis, a retrospective multicenter study in France. *Abstr 31st European Congress of Clinical Microbiology & Infectious Diseases*. European Society of Clinical Microbiology and Infectious Diseases. Vienna, Austria.
17. Wang C, Wu W, Wei L, Feng Y, Kang M, Xie Y, Zong Z. 2020. *Enterobacter wuhouensis* sp. nov. and *Enterobacter quasihormaechei* sp. nov. recovered from human sputum. *Int J Syst Evol Microbiol* 70:874–881. <https://doi.org/10.1099/ijsem.0.003837>.
18. Monahan LG, DeMaere MZ, Cummins ML, Djordjevic SP, Roy Chowdhury P, Darling AE. 2019. High contiguity genome sequence of a multidrug-resistant hospital isolate of *Enterobacter hormaechei*. *Gut Pathog* 11:3. <https://doi.org/10.1186/s13099-019-0288-7>.
19. Doijad S, Imirzalioglu C, Yao Y, Pati NB, Falgenhauer L, Hain T, Foessel BU, Abt B, Overmann J, Mirambo MM, Mshana SE, Chakraborty T. 2016. *Enterobacter bugandensis* sp. nov., isolated from neonatal blood. *Int J Syst Evol Microbiol* 66:968–974. <https://doi.org/10.1099/ijsem.0.000821>.
20. Pati NB, Doijad SP, Schultze T, Mannala GK, Yao Y, Jaiswal S, Ryan D, Suar M, Gwozdzinski K, Bunk B, Mraheil MA, Marahiel MA, Hegemann JD, Spröer C, Goesmann A, Falgenhauer L, Hain T, Imirzalioglu C, Mshana SE, Overmann J, Chakraborty T. 2018. *Enterobacter bugandensis*: a novel enterobacterial species associated with severe clinical infection. *Sci Rep* 8:5392. <https://doi.org/10.1038/s41598-018-23069-z>.
21. Paauw A, Caspers MP, Leverstein-van Hall MA, Schuren FH, Montijn RC, Verhoef J, Fluit AC. 2009. Identification of resistance and virulence factors in an epidemic *Enterobacter hormaechei* outbreak strain. *Microbiology (Reading)* 155:1478–1488. <https://doi.org/10.1099/mic.0.024828-0>.
22. Morand PC, Billoet A, Rottman M, Sivadon-Tardy V, Eyrolle L, Jeanne L, Tazi A, Anract P, Courpied JP, Poyart C, Dumaine V. 2009. Specific distribution within the *Enterobacter cloacae* complex of strains isolated from infected orthopedic implants. *J Clin Microbiol* 47:2489–2495. <https://doi.org/10.1128/JCM.00290-09>.
23. Liu WY, Wong CF, Chung KM, Jiang JW, Leung FC. 2013. Comparative genome analysis of *Enterobacter cloacae*. *PLoS One* 8:e74487. <https://doi.org/10.1371/journal.pone.0074487>.
24. Augusto L, Bourgeois-Nicolaos N, Breton A, Barreault S, Hernandez-Alonso E, Gera S, Faraut-Derouin V, Semaan N, De Luca D, Chaby R, Doucet-Populaire F, Tissieres P. 2021. Presence of 2-hydroxymyristate on endotoxins is associated with death in neonates with *Enterobacter cloacae* complex septic shock. *iScience* 24:102916. <https://doi.org/10.1016/j.isci.2021.102916>.
25. Lazarovitch T, Amity K, Coyle JR, Ackerman B, Tal-Jasper R, Ofer-Friedman H, Hayakawa K, Bogan C, Lephart PR, Kaplansky T, Maskit M, Azouri T, Zaidenstein R, Perez F, Bonomo RA, Kaye KS, Marchaim D. 2015. The complex epidemiology of carbapenem-resistant *Enterobacter* infections: a multicenter descriptive analysis. *Infect Control Hosp Epidemiol* 36:1283–1291. <https://doi.org/10.1017/ice.2015.186>.
26. Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32:1792–1797. <https://doi.org/10.1093/nar/gkh340>.
27. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 28:2731–2739. <https://doi.org/10.1093/molbev/msr121>.
28. Yang Z, Nielsen R. 2000. Estimating synonymous and nonsynonymous substitution rates under realistic evolutionary models. *Mol Biol Evol* 17: 32–43. <https://doi.org/10.1093/oxfordjournals.molbev.a026236>.
29. Chun J, Rainey FA. 2014. Integrating genomics into the taxonomy and systematics of the Bacteria and Archaea. *Int J Syst Evol Microbiol* 64: 316–324. <https://doi.org/10.1099/ijms.0.054171-0>.
30. Foulet F, Botterel F, Buffet P, Morizot G, Rivollet D, Deniau M, Pratlong F, Costa JM, Bretagne S. 2007. Detection and identification of *Leishmania* species from clinical specimens by using a real-time PCR assay and sequencing of the cytochrome B gene. *J Clin Microbiol* 45:2110–2115. <https://doi.org/10.1128/JCM.02555-06>.