



# *In Silico* Identification of Three Types of Integrative and Conjugative Elements in *Elizabethkingia anophelis* Strains Isolated from around the World

✉ Jiannong Xu,<sup>a</sup> Dong Pei,<sup>a\*</sup> Ainsley Nicholson,<sup>b</sup> Yuhao Lan,<sup>a</sup> Qing Xia<sup>a</sup>

<sup>a</sup>Biology Department, New Mexico State University, Las Cruces, New Mexico, USA

<sup>b</sup>Special Bacteriology Reference Laboratory, Bacterial Special Pathogens Branch, Division of High-Consequence Pathogens and Pathology, Centers for Disease Control and Prevention, Atlanta, Georgia, USA

**ABSTRACT** *Elizabethkingia anophelis* is an emerging global multidrug-resistant opportunistic pathogen. We assessed the diversity among 13 complete genomes and 23 draft genomes of *E. anophelis* strains derived from various environmental settings and human infections from different geographic regions around the world from 1950s to the present. Putative integrative and conjugative elements (ICEs) were identified in 31/36 (86.1%) strains in the study. A total of 52 putative ICEs (including eight degenerated elements lacking integrases) were identified and categorized into three types based on the architecture of the conjugation module and the phylogeny of the relaxase, coupling protein, TraG, and TraJ protein sequences. The type II and III ICEs were found to integrate adjacent to tRNA genes, while type I ICEs integrate into intergenic regions or into a gene. The ICEs carry various cargo genes, including transcription regulator genes and genes conferring antibiotic resistance. The adaptive immune CRISPR-Cas system was found in nine strains, including five strains in which CRISPR-Cas machinery and ICEs co-exist at different locations on the same chromosome. One ICE-derived spacer was present in the CRISPR locus in one strain. ICE distribution in the strains showed no geographic or temporal patterns. The ICEs in *E. anophelis* differ in architecture and sequence from CTnDOT, a well-studied ICE prevalent in *Bacteroides* spp. The categorization of ICEs will facilitate further investigations of the impact of ICE on virulence, genome epidemiology, and adaptive genomics of *E. anophelis*.

**IMPORTANCE** *Elizabethkingia anophelis* is an opportunistic human pathogen, and the genetic diversity between strains from around the world becomes apparent as more genomes are sequenced. Genome comparison identified three types of putative ICEs in 31 of 36 strains. The diversity of ICEs suggests that they had different origins. One of the ICEs was discovered previously from a large *E. anophelis* outbreak in Wisconsin in the United States; this ICE has integrated into the *mutY* gene of the outbreak strain, creating a mutator phenotype. Similar to ICEs found in many bacterial species, ICEs in *E. anophelis* carry various cargo genes that enable recipients to resist antibiotics and adapt to various ecological niches. The adaptive immune CRISPR-Cas system is present in nine of 36 strains. An ICE-derived spacer was found in the CRISPR locus in a strain that has no ICE, suggesting a past encounter and effective defense against ICE.


**KEYWORDS** CRISPR-Cas, *Elizabethkingia anophelis*, comparative genomics, genome epidemiology, integrative and conjugative element

The genus *Elizabethkingia* belongs to the family *Flavobacteriaceae* in the phylum *Bacteroidetes*. It was separated from the genus *Chryseobacterium* in 2005 (1). The two member species initially recognized were *Elizabethkingia meningoseptica*, named based on its initial isolation as the causative agent for neonatal meningitis (2), and

**Citation** Xu J, Pei D, Nicholson A, Lan Y, Xia Q. 2019. *In silico* identification of three types of integrative and conjugative elements in *Elizabethkingia anophelis* strains isolated from around the world. *mSphere* 4:e00040-19. <https://doi.org/10.1128/mSphere.00040-19>.

**Editor** Michael Koomey, University of Oslo  
This is a work of the U.S. Government and is not subject to copyright protection in the United States. Foreign copyrights may apply.  
Address correspondence to Jiannong Xu, [jxu@nmsu.edu](mailto:jxu@nmsu.edu).

\* Present address: Dong Pei, Department of Biostatistics, University of Kansas Medical Center, Kansas City, Kansas, USA.

 Integrative and conjugative elements (ICEs) were identified in various strains of *Elizabethkingia anophelis* around the world.

**Received** 22 January 2019

**Accepted** 6 March 2019

**Published** 3 April 2019

*Elizabethkingia miricola*, an isolate obtained from condensation water from Space Station Mir (3). A third species, *Elizabethkingia anophelis*, was proposed in 2011 (4) based on the description of the type strain R26<sup>T</sup> that was originally isolated from the midgut of *Anopheles gambiae* mosquitoes maintained in Stockholm University in Sweden (5). In 2015, a new species, *Elizabethkingia endophytica* was proposed (6), but whole-genome sequence (WGS)-based genome comparison revealed it to be a homotypic synonym of *E. anophelis* (7, 8). Additional species have since been added to the genus (8).

In 2011, the first human infection attributed to *E. anophelis* was documented in Central Africa Republic, neonatal meningitis caused by the bacterium (9). Later, an *E. anophelis* outbreak in an intensive-care unit in Singapore in 2012 was reported (10), followed by the worrying account of *E. anophelis* transmission from a mother to her infant in Hong Kong (11). In 2016, a large *E. anophelis* outbreak occurring in Wisconsin in the United States was unusual in that a substantial proportion of patients were not already hospitalized and were instead admitted directly from their homes. No indication of human-to-human transmission was found (12–14). Human infection cases of *E. anophelis* have been reported with increasing frequency around the world (13, 15–17), aided in part by improved identification methods. Several strains previously described as *E. meningoseptica* were determined to actually belong to the *E. anophelis* species (18). In addition to human-derived strains, multiple strains of the species have been isolated from its namesake, the mosquito, and their genomes have been sequenced (19–21).

Genome comparison of pathogenic bacteria in general has greatly increased our understanding of the evolution, pathogenesis, and epidemiology in many pathogen outbreak investigations (22–26). Genome analysis of *E. anophelis* strains has revealed substantial genetic diversity (11, 16, 27–31), much of it mediated by mobile genetic elements (MGEs). Horizontal transfer of MGEs between bacterial strains occurs frequently and is a major source of genetic variation in bacteria (32). Therefore, an improved understanding of the MGEs in *E. anophelis* should facilitate the genomic epidemiological studies of pathogenic strains of *E. anophelis* in the future.

One type of modular mobile genetic elements, known as integrative and conjugative elements (ICEs), is capable of transferring between bacteria horizontally via conjugation. ICEs integrate into the host chromosome and replicate along with the genome, and they can excise, form a plasmid, and relocate to another site in the genome or transfer to another bacterial cell many generations later (33). The cargo genes that are brought in by ICEs may endow the recipient bacteria with new phenotypes (32–35). For example, in *Pseudomonas syringae* pv. actinidiae, the resistance to copper, arsenic, and cadmium was attributed to the resistant genes that were carried in ICEs (36). In *Helicobacter pylori*, ICE *tfs4* has been identified in certain strains mediating gastric disease. The presence of the *tfs4* ICE has been associated with virulence, although mechanisms behind the virulence phenotype remain unknown (37). Several ICEs have been identified in taxa of *Bacteroides*, including CTnDOT (38, 39), CTnERL (40, 41), and CTnGERM1 (42). In *E. anophelis*, an ICE named ICEEa1 has been identified in the strains associated with the outbreak in Wisconsin and in unrelated strains from the outbreak in Singapore (43).

In this study, we searched for ICEs in 13 complete genomes and 23 draft genomes of *E. anophelis* strains around the world. Based on the architecture of conjugation modules and associated signature genes, three types of ICEs were recognized. As *E. anophelis* is recognized as an environmental bacterium, the high prevalence and diversity of ICEs as well as the mosaic integration pattern in different strains suggest that ICEs play a significant role in shaping its genome to adapt to different environmental niches.

## RESULTS

**Origin and geographic distribution of the strains.** In this study, 13 complete genomes and 23 draft genomes of *E. anophelis* were compared. These strains were

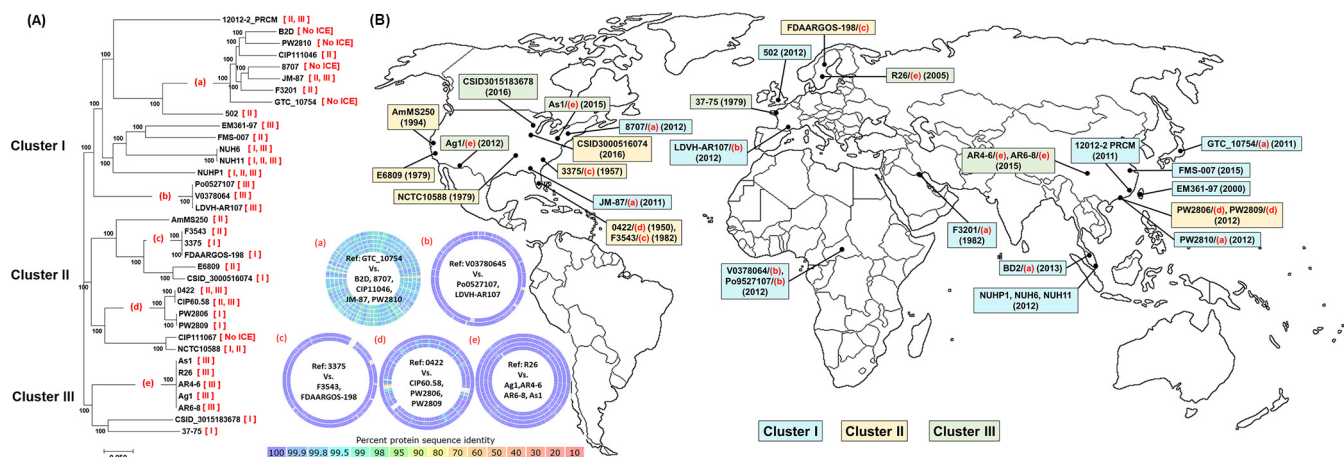
**TABLE 1** *Elizabethkingia anophelis* strains used in this study

WGS accession no.	Level	Cluster	Strain	Source	Region	Collection time	ICE type(s) (n) <sup>a</sup>	CRISPR (n) <sup>b</sup>
ERS1197909 <sup>c</sup>	Draft	II	AmMS250/CIP104057	Human patient	US	1994	II (1)	No
CP023010.1	Complete	II	FDAARGOS-198	Human patient	Sweden	Unknown	I (2)	No
MAHA01	Draft	II	CSID 3000516074	Human patient	Illinois, US	2016	I (1)	No
CP016373.1	Complete	II	3375	Human patient	South Carolina, US	1957	I (1)	No
MAHS01	Complete	II	E6809	Human patient	California, US	1979	II (1)	No
CP014340.1	Complete	II	F3543	Human patient	Florida, US	1982	II (1)	No
CP016370.1	Complete	II	0422	Human patient	Florida, US	1950	II (1), III (1)	No
FTQY00000000.1	Draft	II	CIP60.58	Unknown	Unknown	Unknown	II (1), III (1)	No
CBYD000000000.1	Draft	II	PW2806	Human patient	Hong Kong	2012	I (1)	No
CBYE000000000.1	Draft	II	PW2809	Human patient	Hong Kong	2012	I (1)	No
ERS605480 <sup>c</sup>	Draft	II	NCTC10588	Human patient	US	1959	I (1), II (1)	No
FTQZ00000000.1	Draft	II	CIP111067	Unknown	Unknown	Unknown	No ICE	Yes (37)
CP014805.2	Complete	III	CSID 3015183678	Human patient	Wisconsin, US	2016	I (1)	No
ERS1197911 <sup>c</sup>	Draft	III	37-75/CIP79.29	Human patient	St. Nazaire, France	1979	I (1)	No
CP023401.1	Complete	III	R26	Mosquito	Stockholm, Sweden	2005	III (2)	No
CP023402.1	Complete	III	Ag1	Mosquito	New Mexico, US	2012	III (2)	No
CP023404.1	Complete	III	AR4-6	Mosquito	Sichuan, China	2015	III (2)	No
CP023403.1	Complete	III	AR6-8	Mosquito	Sichuan, China	2015	III (2)	No
LFKT01	Draft	III	As1	Mosquito	Pennsylvania, US	2015	III (2)	No
CP007547.1	Complete	I	NUHP1	Human patient	Singapore	2012	I (1), II (1), III (4)	No
ASYJ01	Draft	I	NUH6	Human patient	Singapore	2012	I (1), III (3)	No
ASYK01	Draft	I	NUH11	Human patient	Singapore	2012	I (1), II (1), III (1)	No
CCAC01	Draft	I	Po0527107	Human patient	Central African Republic	2006	III (1)	Yes (21)
CCAB000000000.1	Draft	I	V0378064	Human patient	Central African Republic	2011	III (1)	Yes (23)
FTPG00000000.1	Draft	I	LDVH-AR107	Common carp	Montpellier, France	2004	III (1)	Yes (42)
CP006576.1	Complete	I	FMS-007	Human patient	Xiaoshan, China	2015	II (1)	Yes (15)
LWDS00000000.1	Draft	I	EM361-97	Human patient	Taiwan	2000s	III (1)	No
AVCQ00000000.1	Draft	I	502	Human patient	Birmingham, UK	2012	II (1)	No
CP016374.1	Complete	I	F3201	Human patient	Kuwait	1982	II (1)	No
CP016372.1	Complete	I	JM-87	Corn <i>Zea mays</i>	Alabama, US	2011	II (1), III (1)	No
ERS1197907 <sup>c</sup>	Draft	I	8707/CIP78.9	Human patient	New York, US	1962	No ICE	Yes
JNCG00000000.1	Draft	I	B2D	Human patient	Malaysia	2013	No ICE	No
FTRB00000000.1	Draft	I	CIP111046	Human patient	Unknown	Unknown	II (1)	Yes (6)
CBYF00000000.1	Draft	I	PW2010	Human patient	Hong Kong	2012	No ICE	Yes (27)
DRS013860	Draft	I	GTC_10754	Unknown	Japan	2014	No ICE	Yes (32)
LPXG00000000.1	Draft	I	12012-2 PRCM	Human patient	Fujian, China	2009	II (1), III (1)	No

<sup>a</sup>n is the number of ICEs in the type.<sup>b</sup>n is the number of spacers in the CRISPR locus.<sup>c</sup>There are no assemblies in the NCBI database for these four strains. We assembled the genomes from Illumina reads directly.

collected from mosquitoes, human patients, and environment globally across Asia, Europe, Africa, and North America between 1950s to the present (Table 1 and Fig. 1).

**Core genome-based phylogeny of the strains.** The pan-genome of a bacterial species is comprised of a core genome and an accessory genome. The core genome is a set of genes that are conserved and shared by all strains, while the accessory gene set is shared by only some strains (44, 45). Because the core genome contains essential genes and their inheritance is necessitated, they carry more reliable evolutionary information than accessory genes do for inferring phylogenetic relationships. Hence, core genome-based single nucleotide polymorphism (SNP) typing has become a recognized method for accurate evolutionary reconstructions (46, 47). The phylogenetic relationship of the core genome from these strains was reconstructed using Parsnp from the Harvest suite (47). Across the 36 genomes, Parsnp recognized 37,781 anchors and 985 maximal unique matches (MUMs) that anchored the genome align-



**FIG 1** Evolutionary relationship and geographic locations of the strains. (A) The phylogenetic tree was derived from the core genome SNP comparison. The type of ICE is shown in red brackets after the strain name. The five circles labeled (a) to (e) in red demonstrate the protein identity between the genomes in the corresponding clades. The color bar represents the percent identity when a genome was compared to the reference genome. (B) Geographic distribution of the strains. Clusters were color coded, and the strain name, collection time, and clade were specified in the map. The letters correspond to the clades in panel A. Strain information was presented in Table 1.

ment. The genome alignment revealed that 68% of the genome regions were shared by all 36 genomes, so these regions were set as core genome. A phylogenetic tree was constructed using the SNPs in the core genome. The tree topology appeared similar regardless of which complete genome—R26<sup>T</sup>, 0422, CSID3015183678, or NUH1—was used as the reference for core genome alignment (data not shown). Based on the tree topology, the strains were sorted into three clusters (Fig. 1 and Table 1). Cluster I contained 14 strains with the genome of strain 12012-2 PRCM at the basal position. Cluster II consisted of two clades, with six strains in each clade. The strains in this cluster were globally distributed and collected over decades starting in 1950s. Cluster III had two clades as well. One clade contained all five mosquito-derived strains, while the other clade contained strain 37-75 obtained from a human infection in France in 1979 and strain CSID3015183678, which was a representative of the strains associated with the Wisconsin outbreak in 2016 (43).

There was no clear indication of geography-based clustering. Figure 1B presents the distribution of strains with their cluster/clade designation and collection date. Among the 14 strains in cluster I, nine strains were collected in Asia, two in Europe, and three in Africa. Among nine strains in cluster II, seven were collected in North America, one in Europe, and one in Asia (Fig. 1 and Table 1). Closely related species that formed the clades a, b, c, d, and e were scattered in different geographic locations (Fig. 1). The clustering of these strains was corroborated by the higher percent identity in the protein sequences between the genomes. Overall, it appears that closely related strains can be widely separated both temporally and spatially.

**Identification of ICEs in the genomes.** A genome region was considered a putative ICE if it harbored genes encoding an integrase, a relaxase (which nicks the DNA strand at the origin of transfer *oriT*), a coupling protein VirD4 ATPase, T4CP, (which couples the relaxosome to the type IV secretion system [T4SS]), and several Tra proteins in the T4SS. These components are involved in integration, excision, and/or conjugation process, which have been used as markers for identifying putative ICEs in genomes (48, 49). So far, only one plasmid has been isolated and sequenced (GenBank accession number CP016375) which is associated with strain F3201. There are no conjugation genes annotated in the complete plasmid sequence, so the plasmid is not a conjugative plasmid. Therefore, the ICE signature genes identified are a part of the genome in these ICE-containing genomes. Cargo genes are carried in the unit, many are flanking the region with the signature genes. Insertion sites of putative ICE loci were recognized and mapped in genomes based on pairwise comparison with closely related genomes

**TABLE 2** Types of ICEs identified in the strains

Element	Strain(s)	Integration site	Integrase(s) <sup>a</sup> (n)	Size (nt) <sup>b</sup>	ICE TPA GenBank accession no.
ICEEal(1)	3375	Siroheme synthase and HP	Tyr (1)	63,549	BK010586
ICEEal(1)	CSID3015183678	Inside <i>mutY</i>	Tyr (1)	62,894	BK010587
ICEEal(1)	FDAARGOS-198	EngB and ElaA	Tyr (1)	62,860	BK010588
ICEEal(1)	NUHP1	SlyD and HP	Tyr (1)	62,960	BK010589
ICEEal(2)	FDAARGOS-198	TonB receptor and HP	Tyr (3)	69,029	BK010590
ICEEal(3)	37-75	Inside <i>Ahp</i>	Tyr (1)	65,419	Not available <sup>c</sup>
ICEEal(4)	CSID3000516074	HP and LuxR TF	Tyr (1), DDE(1)	97,160	BK010591
ICEEal(5)	NCTC10588	Macrolide-efflux protein and LytR/AlgR TR	Tyr (1)	81,809	Not available <sup>c</sup>
ICEEal(6)	PW2806, PW2809	Efflux protein and peptidase	Tyr (1)	79,842	BK010592
ICEEal(7)	NUH6	TomB receptor and HP	Tyr (1)	59,635	BK010593
ICEEal(8)	NUH11	tRNA-Asp-GTC <sup>a</sup>	Tyr (2)	95,803	BK010594
ICEEal(1)	0422	tRNA-Leu-CAA	No Int	63,840	BK010595
ICEEal(1)	CIP60.58	tRNA-Leu-CAA	No Int	63,840	BK010596
ICEEal(2)	502	tRNA-Leu-CAA	IS481 (2)	38,146	BK010597
ICEEal(3)	12012-2 PRCM	tRNA-Leu-CAA	No Int	>56,734	BK010598
ICEEal(4)	AmMS250	tRNA-Leu-CAA	No Int	>64,773	Not available <sup>c</sup>
ICEEal(5)	CIP111046	tRNA-Leu-CAA	No Int	87,438	BK010599
ICEEal(6)	F3201	tRNA-Leu-CAA	No Int	91,608	BK010600
ICEEal(7)	F3543	tRNA-Leu-CAA	No Int	104,603	BK010601
ICEEal(8)	FMS-007	tRNA-Leu-CAA	DDE (1)	73,167	BK010602
ICEEal(9)	NCTC10588	tRNA-Leu-CAA	Ser (2)	67,087	Not available <sup>c</sup>
ICEEal(10)	JM-87	tRNA-Leu-CAA	Ser (1), IS481 (1)	36,967	BK010603
ICEEal(11)	NUH11	tRNA-Leu-CAA	No Int	94,112	BK010604
ICEEal(12)	NUHP1	tRNA-Ser-GGA	Tyr (1)	71,591	BK010605
ICEEal(1)	R26, Ag1, Ar4-6, AR6-8	tRNA-Ser-TGA	Tyr (4), DDE (1)	101,692	BK010606
ICEEal(2)	R26, Ag1, Ar4-6, AR6-8	tRNA-Arg-ACG	Tyr (1), DDE (1)	77,358	BK010607
ICEEal(3)	As1	tRNA-Arg-ACG	Tyr (1), Ser (1), DDE (1)	>44,888	BK010608
ICEEal(4)	As1	tRNA-Ser-TGA	DDE (1)	>31,354	BK010609
ICEEal(5)	NUHP1	tRNA-Glu-TTC	Tyr (4)	60,900	BK010610
ICEEal(6)	NUHP1	tRNA-Glu-TTC	Tyr (1)	74,499	BK010611
ICEEal(7)	NUHP1	tRNA-Asp-GTC	Tyr (1)	116,331	BK010625
ICEEal(8)	NUHP1	tRNA-Asp-GTC	Tyr (3)	109,040	BK010626
ICEEal(9)	NUH6	tRNA-Gln-TTG	Tyr (1)	>30,166	BK010612
ICEEal(10)	NUH6	tRNA-Glu-TTC	Tyr (1)	>33,929	BK010623
ICEEal(11)	NUH6	tRNA-Asp-GTC	Tyr (1)	>98,049	BK010613
ICEEal(12)	NUH11	Beta-lactamase <sup>d</sup>	Tyr (2) <sup>d</sup>	84,534	BK010614
ICEEal(13)	0422	tRNA-Gln-TTG	Tyr (2), DDE (4)	67,662	BK010615
ICEEal(13)	CIP60.58	tRNA-Gln-TTG	Tyr (2), DDE (4)	66,636	BK010616
ICEEal(14)	Po0527107	tRNA-Glu-TTC	Tyr (1)	>31,721	BK010617
ICEEal(14)	V0378064	tRNA-Glu-TTC	Tyr (1)	>37,189	BK010618
ICEEal(15)	JM-87	tRNA-Ser-TGA	Tyr (1)	73,828	BK010624
ICEEal(16)	12012-2 PRCM	tRNA-Glu-TTC	Tyr (1)	>31,732	BK010619
ICEEal(17)	LDVH-AR107	tRNA-Asp-GTC	Tyr (2)	>84,845	BK010620
ICEEal(18)	EM361-97	tRNA-Arg-ACG	Tyr (1), DDE (2)	69,393	BK010621
ICEEal(19)	EM361-97	tRNA-Glu-TTC	DDE (1)	>28,887	BK010622

<sup>a</sup>Tyr, tyrosine type; Ser, serine type; DDE, DDE transposase; IS481, IS481 family of transposases; Int, integrase.

<sup>b</sup>Size (in nucleotides [nt]) is shown. >, elements were partially assembled.

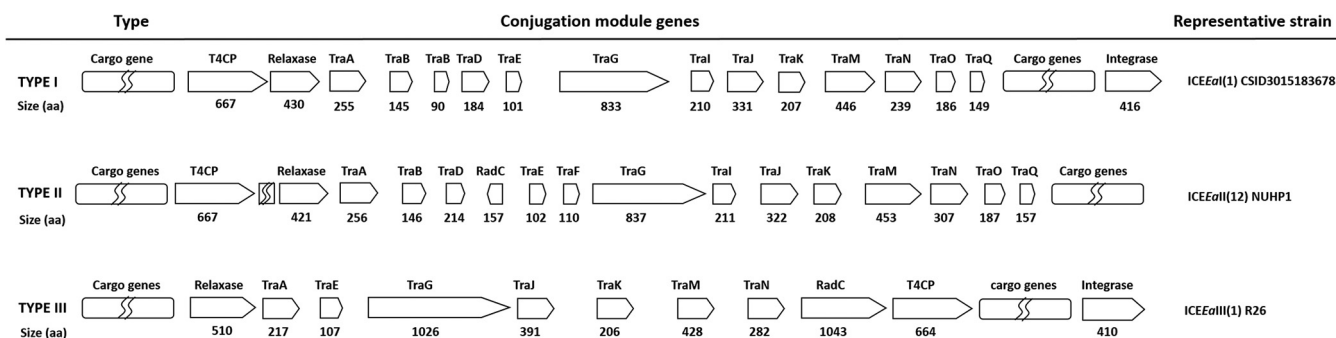
<sup>c</sup>There are no assemblies in the NCBI database for these four strains, so we did not submit third party annotation (TPA) for these genomes to GenBank.

<sup>d</sup>Two elements, ICEEal(8) and ICEEal(12) in strain NUH11 combined and integrated between tRNA-Asp-GTC and beta-lactamase. These two elements may share the same integrase.

lacking the putative ICE. Among the 36 strains, 31 strains contained at least one putative ICE, and a total of 52 putative ICEs were detected. There were eight ICEs in which no integrase was found in the element. Since integrase is necessary for integration and excision, the unit lacking integrase could not conduct integration or excision. The integrase could have been lost after integration into the genome. These elements were defined as degenerated ICEs (Table 2). According to the architecture of the modular genes, these putative ICEs were classified into three types (Fig. 2). For the ease of readability, we omit “putative” before ICE in the remaining text.

Type I ICEs were featured by 13 *tra* genes (*traABBDEGIJKMNOQ*) in addition to genes coding for T4CP, relaxase, and integrase. In most cases, the T4CP and relaxase genes

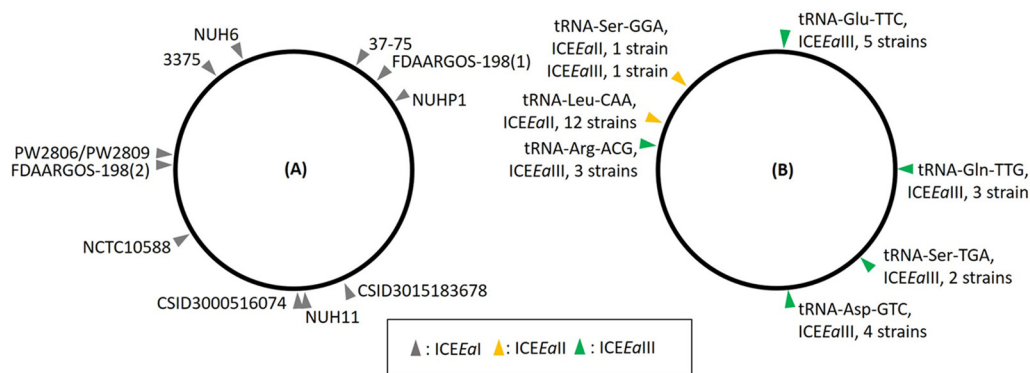




**FIG 2** Schematic view of the architecture of conjugation modular genes in the three types of ICEs.

were located tandemly upstream of the *tra* gene cassette, and the integrase gene was located downstream of the *tra* gene cassette (Fig. 2). The two *traB* genes in tandem were independent from each other, both with complete CDS. Such *traB* genes in tandem were predicted in a few genomes in *E. meningoseptica* and *Chryseobacterium gleum* as well. A total of eight type I elements were found in 11 strains (Table 2). ICEEal(1) was described previously and named ICEEal in strain CSID3015183678 (Wisconsin, US, 2016). It inserted into and disrupted the gene encoding MutY, which is an adenine DNA glycosylase that is required for repairing G-A mismatches, causing the strain to be more prone to mutation (43). ICEEal(1) was found between genes in three additional strains, NUHP1 (Singapore, 2012), 3375 (South Carolina, US, 1957), and FDAARGOS-198 (Sweden, collection time unknown). Each of the four ICEEal(1) elements has a distinct insertion site in the respective genome (Fig. 3). Other type I ICEs inserted into intergenic regions, except for the one in strain 37-75, which inserted into the gene encoding an alkyl hydroperoxide reductase (*Ahp*) between codons K207 and I208. This protein is a primary scavenger of H<sub>2</sub>O<sub>2</sub> in *Escherichia coli* (50), so its disruption by this ICE may result in a loss of function and make the strain vulnerable to oxidative stress. Integration sites of these elements were listed in Table 2 and marked in Fig. 3.

Twelve type II ICEs were identified in 13 strains (Table 2). Unlike type I ICEs, the *T4CP* and *relaxase* genes in type II ICEs were separated by one or more predicted CDS. The cassette of *tra* genes consisted of *traABCDEFGHIJKMNOQ*. There was a gene located between *traD* and *traE*, encoding a RadC domain-containing protein (Fig. 2). All ICEEalI ICEs integrated next to a tRNA gene. In 12 strains, the elements resided at the 3' end of the tRNA-Leu-CAA. Only ICEEalI(12) in strain NUHP1 was located after the tRNA-Ser-GGA (Table 2 and Fig. 3). Integrases were found in four of the 12 elements (see below). Integrases are required for integration and excision, so the absence of an integrase in the elements suggests that these elements may have degenerated and may not be mobilizable. These eight elements were defined as degenerated ICEs.



**FIG 3** Integration sites of ICEs in different strains. (A) Locations of the ICEEal integration sites. (B) Locations of the tRNA genes where ICEEalI and ICEEalIII integrated. The ICE types were color coded. Refer to Table 2 for strain and ICE information.

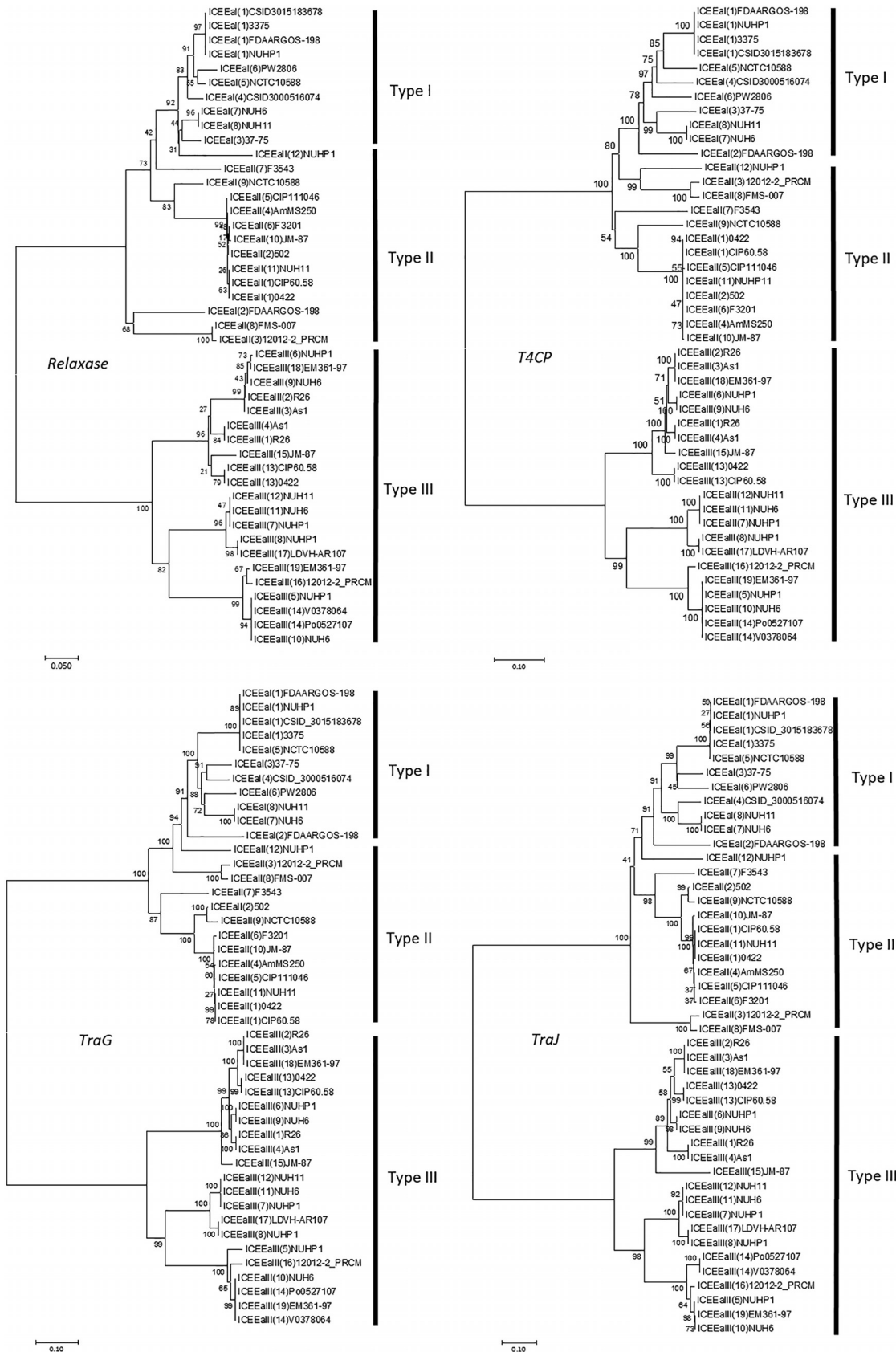
Type III ICEs were recognized in 16 strains. The structure was quite different from that in type I and II ICEs, with the presence of only seven *tra* genes, *traAEGJKMN*. The *relaxase* and *T4CP* genes flanked the *tra* genes. In 11 of 16 elements, a gene was present after the *traN*, encoding a large protein (791 to 1,177 amino acids [aa]) with a RadC domain (Fig. 2). The type III ICEs integrated after a tRNA gene. Five tRNA genes, tRNA-Arg-ACG, tRNA-Gln-TTG, tRNA-Asp-TGA, tRNA-Ser-TGA, and tRNA-Glu-TTC, were targeted by ICEEall3s (Fig. 3). In strain NUHP1, two type III elements, ICEEall3(7) and ICEEall3(8), were collocated tandemly between two tRNA-Asp-GTC and separated by a tRNA-Asp-GTC. In strain NUH11, ICEEal(8) and ICEEall3(12) combined as one segment, which was localized between the tRNA-Asp-GTC and the gene encoding a beta-lactamase. In most type III elements, at least one integrase gene was present.

In sum, eight type I ICEs, 12 type II ICEs (including eight degenerated elements), and 19 type III ICEs were identified, and 15 elements in 11 draft genomes were partially assembled (Table 2). The size of ICEs was up to 116.3 kb in complete genomes. Annotated ICEs were listed in Table S1 in the supplemental material, which included predicted CDS and gene functions. In 36 genomes examined, 17 carried one ICE and 14 harbored more than one ICE. No ICEs were found in these five strains: PW2810, B2D, 8707, CIP11067, and GTC\_10754 (Tables 1 and 2).

**Phylogenetic relationship of the ICEs.** To track the evolutionary history of these ICEs, the nucleotide sequences of the genes encoding relaxase, T4CP, TraG, and TraJ from each ICE were compared. As shown in Fig. 4, the type I and II elements were more closely related, forming a clade distinct from the type III clade. This pattern was consistent with the type classification based on the structure of the conjugation module (Fig. 2). The sequences of type I and II *relaxase* and *TraJ* genes had less resolution than the T4CP and TraG genes to separate type I from type II elements. The conservative nature of the *relaxase* and *TraJ* genes suggests that both may have been under functional constraint with limitation for changes during evolution. The *relaxase* gene of ICEEal(2) in strain FDAARGOS-198 was an outlier in the type I elements. In Fig. 1, the ICE type of each strain was marked in red brackets on the SNP core genome tree. There was no clear association pattern between ICE types and core genome clusters, and no geographic or temporal patterns were observed in the distribution of the ICE types.

**Integrases.** Integrases are required for ICE integration and excision. A total of 43 tyrosine recombinases (TRs) were found in 8/8 type I ICEs, 1/12 type II ICEs, and 18/19 type III ICEs. One ICE may carry up to four tyrosine type recombinases (Table 2). A phylogenetic tree of tyrosine recombinases was constructed using protein sequences from the *E. anophelis* ICEs as well as several *Bacteroides* transposons. As shown in Fig. 5, TRs in type I ICEs form a clade with IntDOT at the basal position, and two TRs of type III ICEs are included in the clade. The clade has strong bootstrap support. However, the protein sequence identity between the TRs in type I ICEs and IntDOT is only about 41%. The ICEEal(2)FDAARGOS-198\_2 is an outlier. The TRs in type III ICEs show greater sequence divergence and split into several clades, with most clades receiving strong bootstrap support. One of these clades contained IntN1 which is associated with NBU1, a mobilizable transposon, in *Bacteroides* (51). DDE transposases were identified in 8/19 type III ICEs and one type I ICE. In type II ICEs, tyrosine recombinase was found only in one type II ICE, ICEEall3(12) NUHP1 (Fig. 5 and Table 2). Serine recombinases were found in ICEEall3(10) JM-87 and ICEEall3(9) NCTC10588. In addition, transposases in the IS481 family (52) are present in ICEEal(2) 502 and ICEEall3(10) JM-87 (Table 2). So, in 12 type II ICEs, only four carry an integrase. No integrases were identified in the remaining eight elements. These elements may have degenerated and become functionally deficient.

**Cargo genes.** In addition to the module genes that define an ICE, there are a diverse cargo of genes that are carried in ICEs. The cargo genes in each ICE are presented in Table S1. Restriction-modification (R-M) system components such as DNA modification methylases, type I R-M systems, and type II restriction enzymes were prevalent in the ICEs of all three types. In addition, the anti-restriction protein ArdA was detected in



**FIG 4** Phylogenetic relationship of the *T4CP*, *relaxase*, *TraG*, and *TraJ* genes associated with ICEs. The nucleotide sequences from different ICEs were aligned, and the evolutionary history was inferred using the neighbor-joining method. The evolutionary distances (Continued on next page)



three type II ICEs and one type III ICE. A beta-lactamase gene was found in two ICEs in two strains. The component genes of tripartite nodulation-division RND complex multidrug efflux pump system (53) were carried in the type I ICEs in six strains, one of the three type III ICEs in strain NUHP1, and the type II ICE in strain F3543 (Table S1). ABC transporters for various substrates, such as manganese, potassium, and oligopeptides, were found in all ICEs. TonB-dependent receptors for siderophore import or carbohydrate uptake (SusC) were present in several ICEs. There were various transcriptional regulators in the ICEs, such as AraC family, ArsR family, MarR family, HxlR family, and TetR family. The AraC family transcriptional regulators (AFTRs) were very prevalent; a total of 36 of these were found in the ICEs in 19 strains. In addition, a two-component regulatory system is present in the type I ICEs in three strains. Some ICEs also carried DNA topoisomerases, helicases, primases, and polymerases, presumably for independent replication while in plasmid form.

**CRISPR-Cas loci.** A type II-C CRISPR-Cas system (54) containing a CRISPR array and the genes encoding Cas9, Cas1, and Cas2 proteins was detected in nine strains (Table 1) with a prevalence of 25% (9/36). The unit is located downstream of the gene encoding a cobalt-zinc-cadmium resistance protein CzcD. The number of spacers varies between 6 in strain CIP111046 to 47 in strain CIP111067. LDVH-AR107 was isolated from the internal organ of a common carp *Cyprinus carpio* collected in 2004 in Montpellier, France. It has two CRISPR loci, each with 21 spacers. Interestingly, strains LDVH-AR107, Po0527107, and V0378064 share 12 spacers, despite the difference of geographic origins and sampling time (Table 1), suggesting that these CRISPR loci may have been derived from a common ancestor. The CRISPR locus of strain GTC\_10754 carried 32 spacers, and the spacer at the far end of the CRISPR array might be derived from the gene encoding a hypothetical protein located in the region between TraG and Tra J in a type III ICE. The predicted spacer (CTTATTTAAATATTATGCTAACAGAGATAA) was 30 nt long, the 2 to 30 nucleotides of the spacer, TTATTTAAATATTATGCTAACAGAGATAA, had a perfect match with the target sequence, or corresponding protospacer, in five type III ICEs, ICE*Ea*III(1), -(3), -(6), -(9), and -(15), and the same spacer had one mismatch with the protospacer in four other elements, ICE*Ea*III(2), -(4), -(13), and -(18). These protospacer carrying ICEs are present in five strains derived from mosquitoes (R26<sup>T</sup>, Ag1, As1, AR4-6, and AR6-8), three strains derived from human infection (0422, CIP60.85, and EM361-97), and one strain derived from corn (JM-87). It is conceivable that the spacer might have been acquired from previous encounters with the type III ICEs and inserted into the CRISPR locus. Theoretically, strain GTC\_10754 can target the ICEs with the protospacer. In fact, no ICEs are present in strain GTC\_10754 (Table 1). The CRISPR loci in the other eight strains do not contain any spacers derived from the ICEs in this study.

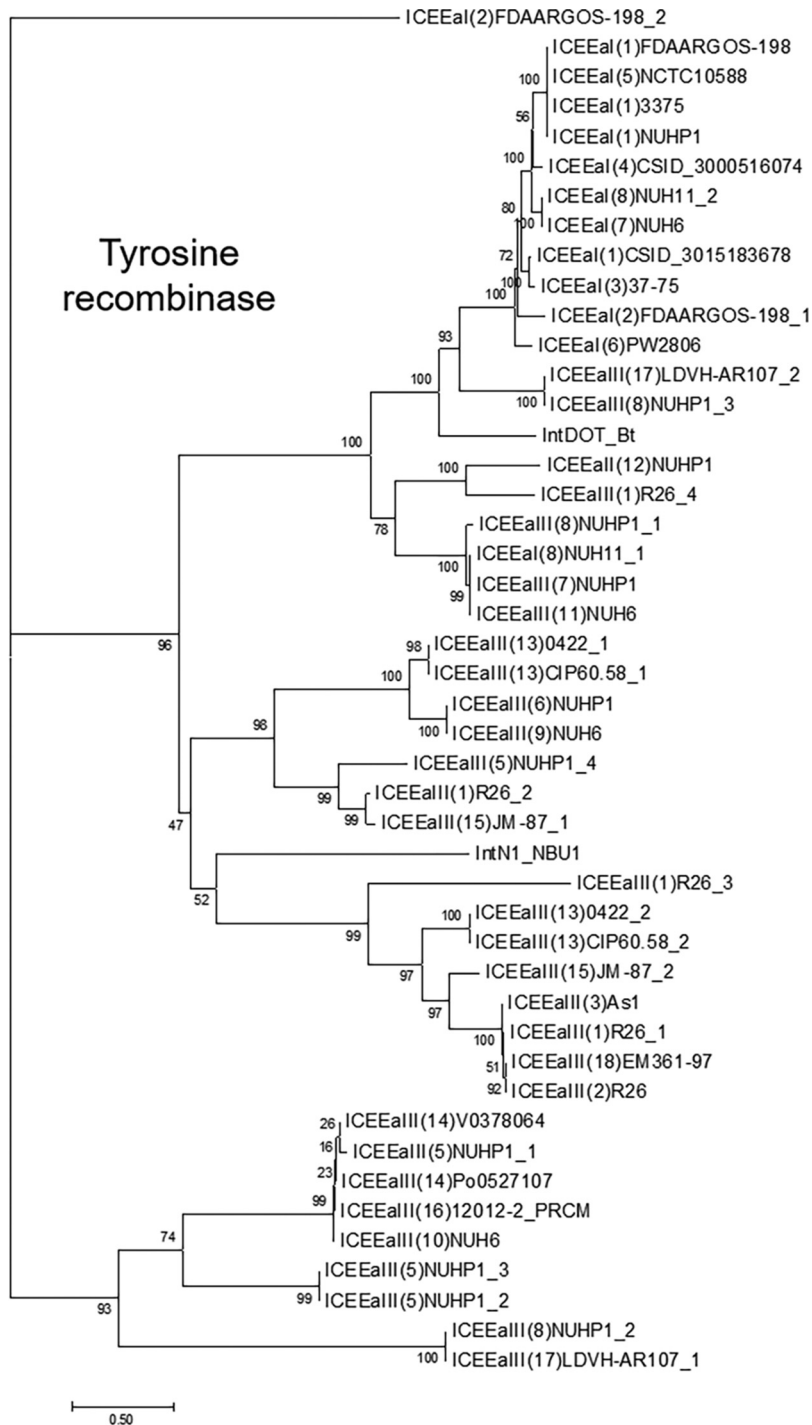
## DISCUSSION

Many genomes have been sequenced for *E. anophelis* strains derived from different sources, including mosquitoes, human infections, hospital environments, fish, and corn stems. This genome availability enabled a comparative genomics approach to investigate the genetic architecture and repertoire of the *E. anophelis* population.

**Core genome-based phylogeny.** The core genome is shared by all strains and consists of essential genes that are vertically transmitted, while genes in the accessory genome are present only in a subset of strains. The sum of the core and accessory genomes in all strains constitutes a pan-genome for the species (45, 55, 56). Due to its inheritability, the core genome is intrinsically suitable for inferring phylogenetic relationships (47). In the present study, the core genome SNP analysis separated the *E. anophelis* strains into three clusters (Table 1 and Fig. 1); however, no correlation patterns were found between genetic relatedness and spatial distribution. The *E.*

### FIG 4 Legend (Continued)

were computed using different models to reconstruct the phylogenetic trees with the bootstrap test using 1,000 replicates, which generated similar tree topology. The consensus trees generated using the Kimura two-parameter model were presented. The bootstrap values are shown on the nodes.



**FIG 5** Phylogenetic relationship of the tyrosine recombinase proteins associated with ICEs. The protein sequences of tyrosine recombinases were aligned, and the evolutionary history was inferred using the neighbor-joining method. The evolutionary distances were computed using different models to reconstruct the phylogenetic trees with the bootstrap test using 1,000 replicates, which generated similar tree topology. The consensus tree generated using the JTT model was presented. The bootstrap values are shown on the nodes.

*anophelis* outbreak in Singapore in 2012 was caused by isolates derived from the hospital environment (10, 27). The Wisconsin outbreak in 2016 showed a different pattern; many patients were likely infected separately before they were hospitalized (12). Recently, an epidemiological genomic survey using core genome comparison was

reported on the transmission pattern of global spread of aggressive nontuberculous *Mycobacterium abscessus* (57). Originally, human infections were caused by *M. abscessus* isolates acquired from the environment. However, the dominant circulating strains within the global *M. abscessus* patient community may be mediated by nosocomial fomite spread (57). Some cases may result from hospital-based cross-infection (58). The core genome data from the present study suggest that *E. anophelis* infections have been caused by diverse strains acquired from the environment. The mosquito-derived strains from three mosquito species were clustered together and distinct from the strains derived from human infections (Fig. 1). Association of *E. anophelis* has been documented in several different mosquito species, including *Anopheles gambiae* (59), *Culex quinquefasciatus* (60), and *Aedes aegypti* (61), and a role for the bacteria in larval development has been hypothesized. The core genome phylogeny suggests that mosquito-associated strains and strains derived from human infections were not epidemiologically linked.

**Diverse types of ICEs in the strains.** Genome diversification promotes bacterial adaptation and evolution. ICE, a type of mobile genetic element, contributes significantly to the pan-genome reservoir with potentially adaptive genes (33, 48). Several ICEs have been found in the taxa of *Bacteroides* (42, 49, 62–64), which belong to the family *Bacteroidaceae* in the phylum *Bacteroidetes*. Of these elements, CTnDOT in *Bacteroides thetaiotaomicron* has been studied extensively (39, 65–67). The ICEs identified in *E. anophelis* are structurally distinct from CTnDOT. Conserved features of relaxase, coupling protein, and conjugative elements have been used for identifying ICEs in a wide variety of genomes (48, 49). Using this approach, we categorized the ICEs in *E. anophelis* strains into three types based on the architecture of the elements and the phylogeny of sequences from four genes in the elements (Fig. 2 and 4). Type I elements are closer to type II elements. Diverse integrases were associated with ICEs. Tyrosine recombinases are present in all type I and III ICEs. Serine recombinase, tyrosine recombinase, or transposase was found in 4/12 type II ICEs, and the remaining eight elements lack an integrase. In that integrase is essential for integration and excision of an ICE, the absence of integrase in the ICEs suggests that these elements had degenerated and were no longer mobile. Type II and III ICEs tended to integrate adjacent to a tRNA gene, while the type I ICEs used a non-tRNA gene region for integration (Fig. 3). Integration into a gene could result in a loss of function. For example, all strains attributed to the 2015 Wisconsin *E. anophelis* outbreak, including strain CSID3015183678, contained an ICE disrupting the *mutY* gene, which is required for excision of G-A mismatch (68, 69) and consequently demonstrated a high rate of nonsynonymous to synonymous substitutions compared to strains with an intact mismatch repair system (43). The presence of various element-associated integrases leads to diverse integration loci (including tRNA gene and non-tRNA loci) which in turn increases the ability of ICEs to contribute their diverse and dynamic genetic repertoire to *E. anophelis* pan-genome (48, 64).

Among the cargo genes carried by the ICEs for which a function could be identified, most are involved with host defense, nutrient acquisition, or transcription regulation. Innate immune mechanisms like restriction-modification systems provide host defense by protecting against invading DNA (70, 71). RND type multidrug efflux pumps enable the cell to defend against environmental toxins by recognizing and expelling various structurally diverse compounds (72), including antibiotics, as shown previously in strain R26<sup>T</sup> (29) and in the strains associated with the outbreak in Singapore (27, 28). Regarding nutrient acquisition, a variety of ABC transporters and TonB-dependent receptors are present in the ICEs, potentially improving the host's ability to satisfy its nutrient needs in a nutrient-limited environment. The ICEs also equip the host with a variety of transcriptional regulators, with those in the AraC family being particularly prevalent (see Table S1 in the supplemental material). Transcription regulators of this sort can sense various chemical signals involved in carbon metabolism, quorum-sensing signaling, virulence, and stress response, modifying expression of a global gene

repertoire (73–76). Thus, the genetic flexibility conferred by ICEs enables the host to quickly adapt to eco-niche changes, which may contribute to establishment of an infection in humans as well.

**CRISPR loci.** The CRISPR-Cas system is an adaptive immune mechanism against invading nucleic acids. In this study, the type II-C CRISPR-Cas system was identified in nine strains of *E. anophelis* (Table 1). Interestingly, three strains that were collected separately share 12 identical spacers in their CRISPR loci, suggesting a common origin of the CRISPR. Besides, one of the 32 spacers in the CRISPR array of strain GTC\_10754 might have been derived from a gene present in nine of the type III ICEs. The presence of this spacer suggests that the CRISPR-Cas system may have encountered and eliminated a type III ICE in the past. The defense is target specific. In fact, no ICE is present in strain GTC\_10754. It has been demonstrated that a type II-C CRISPR-Cas system from *Riemerella anatipestifer* was able to acquire spacers from a transformed exogenous plasmid and target the plasmid (77). Similarly, the CRISPR-Cas system in *Enterococcus faecalis* functions as immune barriers to the acquisition of a conjugative plasmid (78). On the other hand, five CRISPR-carrying strains harbored ICEs as well (Table 1). However, the CRISPR loci in the strains do not contain any spacers derived from the ICEs in this study, so the CRISPR-Cas systems in these strains are not able to recognize and target these ICEs. The coexistence of mobile elements and CRISPR defense system demonstrate an evolutionary balance between the stability and plasticity of the genome (79, 80). A well-characterized example of genome stability mediated by the CRISPR-Cas systems can be found in *Flavobacterium columnare*, where they were shown to play a role in host-phage interactions, driving long-term genome coevolution through arms race-like mechanisms between the host and phages (81). Conversely, the CRISPR-Cas machinery appeared not to play a role as a resistance mechanism against phage for the fish pathogen *Flavobacterium psychrophilum*, which suggests we still have much to learn about the dynamics of CRISPR-Cas systems. Availability of the *E. anophelis* strains that possess CRISPR-Cas with and without ICEs enables further studies on coevolution of mobile elements and host defense systems in the bacteria.

**Conclusion.** In this study, we identified ICEs in the genomes of 31 of 36 *E. anophelis* strains isolated from a wide array of hosts, collected at diverse geographic locations over several decades. The ICEs can be categorized into three types based on the distinctive architecture of the conjugation module and integration sites. The identification of the ICEs warrants further studies on the genetic diversity of the pan-genome and its impact on the virulence and pathogenesis of this global opportunistic pathogen.

## MATERIALS AND METHODS

**Bacterial strains.** In this study, the complete genomes from 13 strains and draft genomes from 23 strains were analyzed (Table 1). Mosquito-derived strains were isolated from *Anopheles gambiae* (R26<sup>T</sup> and Ag1), *Anopheles stephensi* (As1), and *Anopheles sinensis* (AR4-6 and AR6-8). Strain CSID3015183678 was one of the strains associated with the 2015–2016 Wisconsin outbreak, which has been characterized previously (43), and strains NUHP1, NUH6, and NUH11 were representative strains isolated from the Singapore outbreak; the genomes of these strains have been described (27). Strain LDVH-AR107 was derived from a common carp *Cyprinus carpio*. Strain JM-87 was isolated from maize *Zea mays* and was originally described as “*E. endophytica*” until genome comparison identified it as belonging to the species *E. anophelis* (7). The other strains were all derived from human patients or hospital environments, collected from 1950 to 2016. The draft genomes of strains LDVH-AR107, 8707, NCTC10588, AmMS250, and 37-75 were reassembled using Illumina reads downloaded from the SRA database. The reads were *de novo* assembled by CLC genomics workbench v 10.1.1. The 36 genomes were annotated using the SEED and Rapid Annotations using Subsystems Technology (RAST) at the RAST server (82).

**Phylogenetic relationship of the strains based on core genome comparison.** The Parsnp module in the Harvest suite (47) was used for estimating the phylogenomic relationships of the strains. The core genome that is shared by all 36 strains was identified by multiple genome alignment, and SNPs in the core genome were typed to infer phylogenetic relationships; the process was implemented by using Parsnp. Four complete genomes from strains R26<sup>T</sup>, 0422, CSID3015183678, and NUHP1, were used as reference for the process, respectively. Each complete genome was used iteratively as the reference for tree construction, and tree topology of all four core genome trees was unaffected by choice of reference genome. The tree with R26<sup>T</sup> as the reference was depicted in Fig. 1.

**Identification of ICEs.** To identify an ICE, the RAST annotation of each genome was searched for clusters of genes coding for an integrase, relaxase, coupling protein (T4CP), and transfer (Tra) proteins, including a VirB4 ATPase (TraG) in the conjugation module. These proteins are the key components of an ICE (48, 83). In type II ICEs, eight elements lack any type of integrases; these elements were defined as degenerated elements. The boundary of an element was delimited as between the two open reading frames (ORFs) that flank the ICE. The genome of the type strain R26<sup>T</sup> was used as a reference to mark the integration sites. Each ICE was categorized into one of the three types based on its gene structure (Fig. 2). The ICEs were named based mainly on the nomenclature proposed by Burrus et al. (34): the acronym ICE was followed by the initials of the name of the bacterium (*Ea*), a Roman numeral as type, a strain name and a ordinal number in brackets to identify the same sequence if it is encountered in a different strain. For example, the type I ICE found in both NUHP1 and CSID3015183678 strains would be named ICE*Eal*(1)\_CSID3015183678 and ICE*Eal*(1)\_NUHP1, respectively. Each type has its set of numbers; for example, the four type III ICEs in strain NUHP1 are designated ICE*EaIII*(1)\_NUHP1 through ICE*EaIII*(4)\_NUHP1.

**Inference of phylogenetic relationships of selected genes.** The phylogenetic relationships of the *relaxase*, *T4CP*, *TraG*, and *TraJ* genes were inferred using nucleotide sequences. For tyrosine type recombinases, protein sequences were used for inference. The sequences were aligned using ClustalW, and the sequence alignments were used to infer evolutionary history using the neighbor-joining method. The evolutionary distance was computed using different models, which resulted in similar tree topology. The trees using Kimura two-parameter model for nucleotide sequences were presented in Fig. 4, and the tree using the Jones-Taylor-Thornton (JTT) model for protein sequences was presented in Fig. 5. A bootstrap test with 1,000 replicates was implemented to determine the percentage of clustering of associated taxa. The phylogenetic analyses were conducted in MEGA7 (84). IntDOT and IntN1 were included in the analysis of tyrosine recombinases. The GenBank accession numbers of IntDOT and IntN1 follow: IntDOT-Bt, [CAC47921](https://doi.org/10.1093/ajcp/31.3.241); IntN1\_NBU1, [AF238307.1](https://doi.org/10.1093/ajcp/31.3.241).

**Identification of CRISPR sequences.** Genome sequences were searched for CRISPR using CRISPR-*R*finder (85). The spacers located were then used as query to search against ICEs to identify protospacer sequences.

**Accession number(s).** The genome assemblies of 31 strains were used for ICE identification and annotation. The GenBank accession numbers are listed in Table 1. The draft genomes of strains LDVH-AR107, 8707, NCTC10588, AmMS250, and 37-75 were reassembled using Illumina reads downloaded from the SRA database. The SRA accession numbers are listed in Table 1. ICEs were reannotated in the respective genomes. The nucleotide sequence data reported are available in the Third Party Annotation Section of the DDBJ/ENA/GenBank databases under the ICE TPA accession numbers [BK010586](https://doi.org/10.1093/ajcp/31.3.241) to [BK010626](https://doi.org/10.1093/ajcp/31.3.241), which are listed in Table 2.

**Ethics statement.** The work was conducted following the policy of New Mexico State University Institutional Biosafety Committee.

## SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at <https://doi.org/10.1128/mSphere.00040-19>.

**TABLE S1**, XLSX file, 1.4 MB.

## ACKNOWLEDGMENTS

This work was supported by the National Institutes of Health (SC1A112786 to J.X.), National Science Foundation (1633330 to J.X.), and CDC program funds designated for the study of emerging infectious agents. Funding for the open-access charge was provided by the National Institutes of Health (SC1A112786).

The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health, National Science Foundation, and Centers for Disease Control and Prevention. Mention of company names or products does not constitute endorsement.

J.X. conceived of and designed the study. J.X., D.P., Y.L., and Q.X. collected genome data and performed data analysis. J.X. and A.N. wrote the manuscript.

We declare that we have no conflicts of interest.

## REFERENCES

- Kim KK, Kim MK, Lim JH, Park HY, Lee ST. 2005. Transfer of *Chryseobacterium meningosepticum* and *Chryseobacterium miricola* to *Elizabethkingia* gen. nov. as *Elizabethkingia meningoseptica* comb. nov. and *Elizabethkingia miricola* comb. nov. *Int J Syst Evol Microbiol* 55:1287–1293. <https://doi.org/10.1099/ijs.0.63541-0>.
- King EO. 1959. Studies on a group of previously unclassified bacteria associated with meningitis in infants. *Am J Clin Pathol* 31:241–247. <https://doi.org/10.1093/ajcp/31.3.241>.
- Li Y, Kawamura Y, Fujiwara N, Naka T, Liu H, Huang X, Kobayashi K, Ezaki T. 2003. *Chryseobacterium miricola* sp. nov., a novel species isolated from condensation water of space station Mir. *Syst Appl Microbiol* 26:523–528. <https://doi.org/10.1078/072320203770865828>.
- Kampfer P, Matthews H, Glaeser SP, Martin K, Lidders N, Faye I. 2011. *Elizabethkingia anophelis* sp. nov., isolated from the midgut of the mosquito *Anopheles gambiae*. *Int J Syst Evol Microbiol* 61:2670–2675. <https://doi.org/10.1099/ijs.0.026393-0>.



5. Lindh JM, Borg-Karlson AK, Faye I. 2008. Transstadial and horizontal transfer of bacteria within a colony of *Anopheles gambiae* (Diptera: Culicidae) and oviposition response to bacteria-containing water. *Acta Trop* 107:242–250. <https://doi.org/10.1016/j.actatropica.2008.06.008>.
6. Kampfer P, Busse HJ, McInroy JA, Glaeser SP. 2015. *Elizabethkingia endophytica* sp. nov., isolated from *Zea mays* and emended description of *Elizabethkingia anophelis* Kampfer et al. 2011. *Int J Syst Evol Microbiol* 65:2187–2193. <https://doi.org/10.1099/ijs.0.000236>.
7. Doijad S, Ghosh H, Glaeser S, Kampfer P, Chakraborty T. 2016. Taxonomic reassessment of the genus *Elizabethkingia* using whole-genome sequencing: *Elizabethkingia endophytica* Kampfer et al. 2015 is a later subjective synonym of *Elizabethkingia anophelis* Kampfer et al. 2011. *Int J Syst Evol Microbiol* 66:4555–4559. <https://doi.org/10.1099/ijsem.0.001390>.
8. Nicholson AC, Gulvik CA, Whitney AM, Humrighouse BW, Graziano J, Emery B, Bell M, Loparev V, Juieng P, Gartin J, Bizet C, Clermont D, Criscuolo A, Brisse S, McQuiston JR. 2018. Revisiting the taxonomy of the genus *Elizabethkingia* using whole-genome sequencing, optical mapping, and MALDI-TOF, along with proposal of three novel *Elizabethkingia* species: *Elizabethkingia bruuniana* sp. nov., *Elizabethkingia ursingii* sp. nov., and *Elizabethkingia occulta* sp. nov. *Antonie Van Leeuwenhoek* 111:55–72. <https://doi.org/10.1007/s10482-017-0926-3>.
9. Frank T, Gody JC, Nguyen LBL, Berthet N, Le Fleche-Mateos A, Bata P, Rafai C, Kazanji M, Breurec S. 2013. First case of *Elizabethkingia anophelis* meningitis in the Central African Republic. *Lancet* 381:1876. [https://doi.org/10.1016/S0140-6736\(13\)60318-9](https://doi.org/10.1016/S0140-6736(13)60318-9).
10. Teo J, Tan SY, Tay M, Ding Y, Kjelleberg S, Givskov M, Lin RT, Yang L. 2013. First case of *E anophelis* outbreak in an intensive-care unit. *Lancet* 382:855–856. [https://doi.org/10.1016/S0140-6736\(13\)61858-9](https://doi.org/10.1016/S0140-6736(13)61858-9).
11. Lau SK, Wu AK, Teng JL, Tse H, Curreem SO, Tsui SK, Huang Y, Chen JH, Lee RA, Yuen KY, Woo PC. 2015. Evidence for *Elizabethkingia anophelis* transmission from mother to infant, Hong Kong. *Emerg Infect Dis* 21: 232–241. <https://doi.org/10.3201/eid2102.140623>.
12. Figueroa Castro CE, Johnson C, Williams M, VanDerSlik A, Graham MB, Letzer D, Ledebner N, Buchan BW, Block T, Borlaug G, Munoz-Price LS. 2017. *Elizabethkingia anophelis*: clinical experience of an academic health system in southeastern Wisconsin. *Open Forum Infect Dis* 4:ofx251. <https://doi.org/10.1093/ofid/oxf251>.
13. Navon L, Clegg WJ, Morgan J, Austin C, McQuiston JR, Blaney DD, Walters MS, Moulton-Meissner H, Nicholson A. 2016. Notes from the field: investigation of *Elizabethkingia anophelis* cluster – Illinois, 2014–2016. *MMWR Morb Mortal Wkly Rep* 65:1380–1381. <https://doi.org/10.15585/mmwr.mm6548a6>.
14. Nicholson AC, Whitney AM, Emery BD, Bell ME, Gartin JT, Humrighouse BW, Loparev VN, Batra D, Sheth M, Rowe LA, Juieng P, Knipe K, Gulvik C, McQuiston JR. 2016. Complete genome sequences of four strains from the 2015–2016 *Elizabethkingia anophelis* outbreak. *Genome Announc* 4:e00563-16. <https://doi.org/10.1128/genomeA.00563-16>.
15. Hu S, Jiang T, Zhang X, Zhou Y, Yi Z, Wang Y, Zhao S, Wang M, Ming D, Chen S. 2017. *Elizabethkingia anophelis* isolated from patients with multiple organ dysfunction syndrome and lower respiratory tract infection: report of two cases and literature review. *Front Microbiol* 8:382. <https://doi.org/10.3389/fmicb.2017.00382>.
16. Lin JN, Lai CH, Yang CH, Huang YH, Lin HH. 2017. Genomic features, phylogenetic relationships, and comparative genomics of *Elizabethkingia anophelis* strain EM361-97 isolated in Taiwan. *Sci Rep* 7:14317. <https://doi.org/10.1038/s41598-017-14841-8>.
17. Lau SK, Chow WN, Foo CH, Curreem SO, Lo GC, Teng JL, Chen JH, Ng RH, Wu AK, Cheung IY, Chau SK, Lung DC, Lee RA, Tse CW, Fung KS, Que TL, Woo PC. 2016. *Elizabethkingia anophelis* bacteremia is associated with clinically significant infections and high mortality. *Sci Rep* 6:26045. <https://doi.org/10.1038/srep26045>.
18. Janda JM, Lopez DL. 2017. Mini review: new pathogen profiles: *Elizabethkingia anophelis*. *Diagn Microbiol Infect Dis* 88:201–205. <https://doi.org/10.1016/j.diagmicrobio.2017.03.007>.
19. Kukutla P, Lindberg BG, Pei D, Rayl M, Yu W, Steritz M, Faye I, Xu J. 2013. Draft genome sequences of *Elizabethkingia anophelis* strains R26T and Ag1 from the midgut of the malaria mosquito *Anopheles gambiae*. *Genome Announc* 1:e01030-13. <https://doi.org/10.1128/genomeA.01030-13>.
20. Pei D, Nicholson AC, Jiang J, Chen H, Whitney AM, Villara A, Bell M, Humrighouse B, Rowe LA, Sheth M, Batra D, Juieng P, Loparev VN, McQuiston JR, Lan Y, Ma Y, Xu J. 2017. Complete circularized genome sequences of four strains of *Elizabethkingia anophelis*, including two novel strains isolated from wild-caught *Anopheles sinensis*. *Genome Announc* 5:e01359-17. <https://doi.org/10.1128/genomeA.01359-17>.
21. Raygoza Garay JA, Hughes GL, Koundal V, Rasgon JL, Mwangi MM. 2016. Genome sequence of *Elizabethkingia anophelis* strain EaAs1, isolated from the Asian malaria mosquito *Anopheles stephensi*. *Genome Announc* 4:e00084-16. <https://doi.org/10.1128/genomeA.00084-16>.
22. Georgiades K. 2012. Genomics of epidemic pathogens. *Clin Microbiol Infect* 18:213–217. <https://doi.org/10.1111/j.1469-0691.2012.03781.x>.
23. Raskin DM, Seshadri R, Pukatzki SU, Mekalanos JJ. 2006. Bacterial genomics and pathogen evolution. *Cell* 124:703–714. <https://doi.org/10.1016/j.cell.2006.02.002>.
24. Worby CJ, Lipsitch M, Hanage WP. 2017. Shared genomic variants: identification of transmission routes using pathogen deep-sequence data. *Am J Epidemiol* 186:1209–1216. <https://doi.org/10.1093/aje/kwx182>.
25. Hu R, Yuan J, Meng Y, Wang Z, Gu Z. 2017. Pathogenic *Elizabethkingia miricola* infection in cultured black-spotted frogs, China, 2016. *Emerg Infect Dis* 23:2055–2059. <https://doi.org/10.3201/eid2312.170942>.
26. Peter S, Oberhettinger P, Schuele L, Dinkelacker A, Vogel W, Dorfel D, Bezdán D, Ossowski S, Marschal M, Liese J, Willmann M. 2017. Genomic characterisation of clinical and environmental *Pseudomonas putida* group strains and determination of their role in the transfer of antimicrobial resistance genes to *Pseudomonas aeruginosa*. *BMC Genomics* 18:859. <https://doi.org/10.1186/s12864-017-4216-2>.
27. Teo J, Tan SY, Liu Y, Tay M, Ding Y, Li Y, Kjelleberg S, Givskov M, Lin RT, Yang L. 2014. Comparative genomic analysis of malaria mosquito vector-associated novel pathogen *Elizabethkingia anophelis*. *Genome Biol Evol* 6:1158–1165. <https://doi.org/10.1093/gbe/evu094>.
28. Li Y, Liu Y, Chew SC, Tay M, Salido MM, Teo J, Lauro FM, Givskov M, Yang L. 2015. Complete genome sequence and transcriptomic analysis of the novel pathogen *Elizabethkingia anophelis* in response to oxidative stress. *Genome Biol Evol* 7:1676–1685. <https://doi.org/10.1093/gbe/evv101>.
29. Kukutla P, Lindberg BG, Pei D, Rayl M, Yu W, Steritz M, Faye I, Xu J. 2014. Insights from the genome annotation of *Elizabethkingia anophelis* from the malaria vector *Anopheles gambiae*. *PLoS One* 9:e97715. <https://doi.org/10.1371/journal.pone.0097715>.
30. Chew KL, Cheng B, Lin RTP, Teo JWP. 2018. *Elizabethkingia anophelis* is the dominant *Elizabethkingia* species found in blood cultures in Singapore. *J Clin Microbiol* 56:e01445-17. <https://doi.org/10.1128/JCM.01445-17>.
31. Breurec S, Criscuolo A, Diancourt L, Rendueles O, Vandenbogaert M, Passet V, Caro V, Rocha EP, Touchon M, Brisse S. 2016. Genomic epidemiology and global diversity of the emerging bacterial pathogen *Elizabethkingia anophelis*. *Sci Rep* 6:30379. <https://doi.org/10.1038/srep30379>.
32. Bellanger X, Payot S, Leblond-Bourget N, Kjedon G. 2014. Conjugative and mobilizable genomic islands in bacteria: evolution and diversity. *FEMS Microbiol Rev* 38:720–760. <https://doi.org/10.1111/1574-6976.12058>.
33. Wozniak RA, Waldor MK. 2010. Integrative and conjugative elements: mosaic mobile genetic elements enabling dynamic lateral gene flow. *Nat Rev Microbiol* 8:552–563. <https://doi.org/10.1038/nrmicro2382>.
34. Burrus V, Pavlovic G, Decaris B, Guedon G. 2002. Conjugative transposons: the tip of the iceberg. *Mol Microbiol* 46:601–610. <https://doi.org/10.1046/j.1365-2958.2002.03191.x>.
35. Johnson CM, Grossman AD. 2015. Integrative and conjugative elements (ICEs): what they do and how they work. *Annu Rev Genet* 49:577–601. <https://doi.org/10.1146/annurev-genet-112414-055018>.
36. Colombi E, Straub C, Kunzel S, Templeton MD, McCann HC, Rainey PB. 2017. Evolution of copper resistance in the kiwifruit pathogen *Pseudomonas syringae* pv. *actinidiae* through acquisition of integrative conjugative elements and plasmids. *Environ Microbiol* 19:819–832. <https://doi.org/10.1111/1462-2920.13662>.
37. Sugimoto M, Watada M, Jung SW, Graham DY, Yamaoka Y. 2012. Role of *Helicobacter pylori* plasticity region genes in development of gastroduodenal diseases. *J Clin Microbiol* 50:441–448. <https://doi.org/10.1128/JCM.00906-11>.
38. Cheng Q, Paszkiet BJ, Shoemaker NB, Gardner JF, Salyers AA. 2000. Integration and excision of a *Bacteroides* conjugative transposon, CTnDOT. *J Bacteriol* 182:4035–4043. <https://doi.org/10.1128/JB.182.14.4035-4043.2000>.
39. Wood MM, Gardner JF. 2015. The integration and excision of CTnDOT. *Microbiol Spectr* 3:MDNA3. <https://doi.org/10.1128/microbiolspec.MDNA3-0020-2014>.
40. Li LY, Shoemaker NB, Salyers AA. 1995. Location and characteristics of the transfer region of a *Bacteroides* conjugative transposon and regu-

- lation of transfer genes. *J Bacteriol* 177:4992–4999. <https://doi.org/10.1128/jb.177.17.4992-4999.1995>.
41. Whittle G, Hamburger N, Shoemaker NB, Salyers AA. 2006. A *Bacteroides* conjugative transposon, CTnERL, can transfer a portion of itself by conjugation without excising from the chromosome. *J Bacteriol* 188:1169–1174. <https://doi.org/10.1128/JB.188.3.1169-1174.2006>.
  42. Wang Y, Wang GR, Shelby A, Shoemaker NB, Salyers AA. 2003. A newly discovered *Bacteroides* conjugative transposon, CTnGERM1, contains genes also found in gram-positive bacteria. *Appl Environ Microbiol* 69:4595–4603. <https://doi.org/10.1128/AEM.69.8.4595-4603.2003>.
  43. Perrin A, Larssonneur E, Nicholson AC, Edwards DJ, Gundlach KM, Whitney AM, Gulvik CA, Bell ME, Rendueles O, Cury J, Hugon P, Clermont D, Enouf V, Loparev V, Juieng P, Monson T, Warshauer D, Elbadawi LI, Walters MS, Crist MB, Noble-Wang J, Borlaug G, Rocha EPC, Criscuolo A, Touchon M, Davis JP, Holt KE, McQuiston JR, Brisse S. 2017. Evolutionary dynamics and genomic features of the *Elizabethkingia anophelis* 2015 to 2016 Wisconsin outbreak strain. *Nat Commun* 8:15483. <https://doi.org/10.1038/ncomms15483>.
  44. Medini D, Donati C, Tettelin H, Massignani V, Rappuoli R. 2005. The microbial pan-genome. *Curr Opin Genet Dev* 15:589–594. <https://doi.org/10.1016/j.gde.2005.09.006>.
  45. Tettelin H, Massignani V, Cieslewicz MJ, Donati C, Medini D, Ward NL, Angiuoli SV, Crabtree J, Jones AL, Durkin AS, Deboy RT, Davidsen TM, Mora M, Scarselli M, Margarit y Ros I, Peterson JD, Hauser CR, Sundaram JP, Nelson WC, Madupu R, Brinkac LM, Dodson RJ, Rosovitz MJ, Sullivan SA, Daugherty SC, Haft DH, Selengut J, Gwinn ML, Zhou L, Zafar N, Khouri H, Radune D, Dimitrov G, Watkins K, O'Connor KJ, Smith S, Utterback TR, White O, Rubens CE, Grandi G, Madoff LC, Kasper DL, Telford JL, Wessels MR, Rappuoli R, Fraser CM. 2005. Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: implications for the microbial “pan-genome”. *Proc Natl Acad Sci U S A* 102:13950–13955. <https://doi.org/10.1073/pnas.0506758102>.
  46. Leekitcharoenphon P, Kaas RS, Thomsen MC, Friis C, Rasmussen S, Aarestrup FM. 2012. snpTree—a web-server to identify and construct SNP trees from whole genome sequence data. *BMC Genomics* 13(Suppl 7):S6. <https://doi.org/10.1186/1471-2164-13-S7-S6>.
  47. Treangen TJ, Ondov BD, Koren S, Phillippy AM. 2014. The Harvest suite for rapid core-genome alignment and visualization of thousands of intraspecific microbial genomes. *Genome Biol* 15:524. <https://doi.org/10.1186/s13059-014-0524-x>.
  48. Cury J, Touchon M, Rocha E. 2017. Integrative and conjugative elements and their hosts: composition, distribution and organization. *Nucleic Acids Res* 45:8943–8956. <https://doi.org/10.1093/nar/gkx607>.
  49. Guglielmini J, Quintais L, Garcillan-Barcia MP, de la Cruz F, Rocha EP. 2011. The repertoire of ICE in prokaryotes underscores the unity, diversity, and ubiquity of conjugation. *PLoS Genet* 7:e1002222. <https://doi.org/10.1371/journal.pgen.1002222>.
  50. Seaver LC, Imlay JA. 2001. Alkyl hydroperoxide reductase is the primary scavenger of endogenous hydrogen peroxide in *Escherichia coli*. *J Bacteriol* 183:7173–7181. <https://doi.org/10.1128/JB.183.24.7173-7181.2001>.
  51. Rajeev L, Salyers AA, Gardner JF. 2006. Characterization of the integrase of NBU1, a *Bacteroides mobilizabilis* transposon. *Mol Microbiol* 61:978–990. <https://doi.org/10.1111/j.1365-2958.2006.05282.x>.
  52. Register KB, Sanden GN. 2006. Prevalence and sequence variants of IS481 in *Bordetella bronchiseptica*: implications for IS481-based detection of *Bordetella pertussis*. *J Clin Microbiol* 44:4577–4583. <https://doi.org/10.1128/JCM.01295-06>.
  53. Venter H, Mowla R, Ohene-Agyei T, Ma S. 2015. RND-type drug efflux pumps from Gram-negative bacteria: molecular mechanism and inhibition. *Front Microbiol* 6:377. <https://doi.org/10.3389/fmicb.2015.00377>.
  54. Mir A, Edraki A, Lee J, Sontheimer EJ. 2018. Type II-CRISPR-Cas9 biology, mechanism, and application. *ACS Chem Biol* 13:357–365. <https://doi.org/10.1021/acscchembio.7b00855>.
  55. Bentley SD, Parkhill J. 2015. Genomic perspectives on the evolution and spread of bacterial pathogens. *Proc Biol Sci* 282:20150488. <https://doi.org/10.1098/rspb.2015.0488>.
  56. Mira A, Martin-Cuadrado AB, D'Auria G, Rodríguez-Valera F. 2010. The bacterial pan-genome: a new paradigm in microbiology. *Int Microbiol* 13:45–57. <https://doi.org/10.2436/20.1501.01.110>.
  57. Bryant JM, Grogono DM, Rodriguez-Rincon D, Everall I, Brown KP, Moreno P, Verma D, Hill E, Drijkoningen J, Gilligan P, Esther CR, Noone PG, Giddings O, Bell SC, Thomson R, Wainwright CE, Coulter C, Pandey S, Wood ME, Stockwell RE, Ramsay KA, Sherrard LJ, Kidd TJ, Jabbour N, Johnson GR, Knibbs LD, Morawska L, Sly PD, Jones A, Bilton D, Laurenson I, Ruddy M, Bourke S, Bowler IC, Chapman SJ, Clayton A, Cullen M, Daniels T, Dempsey O, Denton M, Desai M, Drew RJ, Edenborough F, Evans J, Folb J, Humphrey H, Isalska B, Jensen-Fangel S, Jönsson B, Jones AM, Katzenstein TL, et al. 2016. Emergence and spread of a human-transmissible multidrug-resistant nontuberculous mycobacterium. *Science* 354:751–757. <https://doi.org/10.1126/science.aaf8156>.
  58. Bryant JM, Grogono DM, Greaves D, Foweraker J, Roddick I, Inns T, Reacher M, Haworth CS, Curran MD, Harris SR, Peacock SJ, Parkhill J, Floto RA. 2013. Whole-genome sequencing to identify transmission of *Mycobacterium abscessus* between patients with cystic fibrosis: a retrospective cohort study. *Lancet* 381:1551–1560. [https://doi.org/10.1016/S0140-6736\(13\)60632-7](https://doi.org/10.1016/S0140-6736(13)60632-7).
  59. Wang Y, Gilbreath TM, III, Kukutla P, Yan G, Xu J. 2011. Dynamic gut microbiome across life history of the malaria mosquito *Anopheles gambiae* in Kenya. *PLoS One* 6:e24767. <https://doi.org/10.1371/journal.pone.0024767>.
  60. Telang A, Skinner J, Nemitz RZ, McClure AM. 2018. Metagenome and culture-based methods reveal candidate bacterial mutualists in the Southern house mosquito (Diptera: Culicidae). *J Med Entomol* 55:1170–1181. <https://doi.org/10.1093/jme/tjy056>.
  61. Terenius O, Lindh JM, Eriksson-Gonzales K, Bussiere L, Laugen AT, Bergquist H, Titanji K, Faye I. 2012. Midgut bacterial dynamics in *Aedes aegypti*. *FEMS Microbiol Ecol* 80:556–565. <https://doi.org/10.1111/j.1574-6941.2012.03137.x>.
  62. Burrus V, Waldor MK. 2004. Shaping bacterial genomes with integrative and conjugative elements. *Res Microbiol* 155:376–386. <https://doi.org/10.1016/j.resmic.2004.01.012>.
  63. Malanowska K, Salyers AA, Gardner JF. 2006. Characterization of a conjugative transposon integrase, IntDOT. *Mol Microbiol* 60:1228–1240. <https://doi.org/10.1111/j.1365-2958.2006.05164.x>.
  64. Nguyen M, Vedantam G. 2011. Mobile genetic elements in the genus *Bacteroides*, and their mechanism(s) of dissemination. *Mob Genet Elements* 1:187–196. <https://doi.org/10.4161/mge.1.3.18448>.
  65. Laprise J, Yoneji S, Gardner JF. 2013. IntDOT interactions with core sites during integrative recombination. *J Bacteriol* 195:1883–1891. <https://doi.org/10.1128/JB.01540-12>.
  66. Whittle G, Shoemaker NB, Salyers AA. 2002. Characterization of genes involved in modulation of conjugal transfer of the *Bacteroides* conjugative transposon CTnDOT. *J Bacteriol* 184:3839–3847. <https://doi.org/10.1128/JB.184.14.3839-3847.2002>.
  67. Waters JL, Salyers AA. 2013. Regulation of CTnDOT conjugative transfer is a complex and highly coordinated series of events. *mBio* 4:e00569-13. <https://doi.org/10.1128/mBio.00569-13>.
  68. Au KG, Clark S, Miller JH, Modrich P. 1989. *Escherichia coli* mutY gene encodes an adenine glycosylase active on G-A mispairs. *Proc Natl Acad Sci U S A* 86:8877–8881. <https://doi.org/10.1073/pnas.86.22.8877>.
  69. Michaels ML, Pham L, Nghiem Y, Cruz C, Miller JH. 1990. MutY, an adenine glycosylase active on G-A mispairs, has homology to endonuclease III. *Nucleic Acids Res* 18:3841–3845. <https://doi.org/10.1093/nar/18.13.3841>.
  70. Murray NE. 2000. Type I restriction systems: sophisticated molecular machines (a legacy of Bertani and Weigle). *Microbiol Mol Biol Rev* 64:412–434. <https://doi.org/10.1128/MMBR.64.2.412-434.2000>.
  71. Oliveira PH, Touchon M, Rocha EP. 2014. The interplay of restriction-modification systems with mobile genetic elements and their prokaryotic hosts. *Nucleic Acids Res* 42:10618–10631. <https://doi.org/10.1093/nar/gku734>.
  72. Nikaido H, Pages JM. 2012. Broad-specificity efflux pumps and their role in multidrug resistance of Gram-negative bacteria. *FEMS Microbiol Rev* 36:340–363. <https://doi.org/10.1111/j.1574-6976.2011.00290.x>.
  73. Gallegos MT, Schleif R, Bairoch A, Hofmann K, Ramos JL. 1997. AraC/XylS family of transcriptional regulators. *Microbiol Mol Biol Rev* 61:393–410.
  74. Yang J, Tauschek M, Robins-Browne RM. 2011. Control of bacterial virulence by AraC-like regulators that respond to chemical signals. *Trends Microbiol* 19:128–135. <https://doi.org/10.1016/j.tim.2010.12.001>.
  75. Tan A, Petty NK, Hocking D, Bennett-Wood V, Wakefield M, Praszkiar J, Tauschek M, Yang J, Robins-Browne R. 2015. Evolutionary adaptation of an AraC-like regulatory protein in *Citrobacter rodentium* and *Escherichia* species. *Infect Immun* 83:1384–1395. <https://doi.org/10.1128/IAI.02697-14>.
  76. Tobes R, Ramos JL. 2002. AraC-XylS database: a family of positive transcriptional regulators in bacteria. *Nucleic Acids Res* 30:318–321. <https://doi.org/10.1093/nar/30.1.318>.
  77. He Y, Wang M, Liu M, Huang L, Liu C, Zhang X, Yi H, Cheng A, Zhu D, Yang Q, Wu Y, Zhao X, Chen S, Jia R, Zhang S, Liu Y, Yu Y, Zhang L. 2018.

- Cas1 and Cas2 from the type II-C CRISPR-Cas system of *Riemerella anatipestifer* are required for spacer acquisition. *Front Cell Infect Microbiol* 8:195. <https://doi.org/10.3389/fcimb.2018.00195>.
78. Price VJ, Huo W, Sharifi A, Palmer KL. 2016. CRISPR-Cas and restriction-modification act additively against conjugative antibiotic resistance plasmid transfer in *Enterococcus faecalis*. *mSphere* 1:e00064-16. <https://doi.org/10.1128/mSphere.00064-16>.
79. Koonin EV, Makarova KS. 2017. Mobile genetic elements and evolution of CRISPR-Cas systems: all the way there and back. *Genome Biol Evol* 9:2812–2825. <https://doi.org/10.1093/gbe/evx192>.
80. Varahan S, Hancock LE. 2016. To defend or not to defend: that's the question. *mSphere* 1:e00127-16. <https://doi.org/10.1128/mSphere.00127-16>.
81. Laanto E, Hoikkala V, Ravannti J, Sundberg LR. 2017. Long-term genomic coevolution of host-parasite interaction in the natural environment. *Nat Commun* 8:111. <https://doi.org/10.1038/s41467-017-00158-7>.
82. Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, Edwards RA, Gerdes S, Parrello B, Shukla M, Vonstein V, Wattam AR, Xia F, Stevens R. 2014. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Res* 42:D206–D214. <https://doi.org/10.1093/nar/gkt1226>.
83. Guglielmini J, Neron B, Abby SS, Garcillan-Barcia MP, de la Cruz F, Rocha EP. 2014. Key components of the eight classes of type IV secretion systems involved in bacterial conjugation or protein secretion. *Nucleic Acids Res* 42:5715–5727. <https://doi.org/10.1093/nar/gku194>.
84. Kumar S, Stecher G, Tamura K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Mol Biol Evol* 33:1870–1874. <https://doi.org/10.1093/molbev/msw054>.
85. Grissa I, Vergnaud G, Pourcel C. 2007. CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res* 35:W52–W57. <https://doi.org/10.1093/nar/gkm360>.