JOURNAL OF CLINICAL MICROBIOLOGY, Mar. 1999, p. 581–590 Vol. 37, No. 3
0095-1137/99/$04.00+0

Genetic Diversity and Population Structure of *Vibrio cholerae*

PILAR BELTRÁN,[1] GABRIELA DELGADO,[1] ARMANDO NAVARRO,[1] FRANCISCA TRUJILLO,[1] ROBERT K. SELANDER,[2]* AND ALEJANDRO CRAVIOTO[1]

Departamento de Salud Pública de la Facultad de Medicina, Universidad Nacional Autónoma de México, México, D.F., México,[1] and Institute of Molecular Evolutionary Genetics, Pennsylvania State University, University Park, Pennsylvania 16802[2]

Received 9 September 1998/Returned for modification 17 November 1998/Accepted 8 December 1998

Multilocus enzyme electrophoresis (MLEE) of 397 *Vibrio cholerae* isolates, including 143 serogroup reference strains and 244 strains from Mexico and Guatemala, identified 279 electrophoretic types (ETs) distributed in two major divisions (I and II). Linkage disequilibrium was demonstrated in both divisions and in subdivision Ic of division I but not in subdivision Ia, which includes 76% of the ETs. Despite this evidence of relatively frequent recombination, clonal lineages may persist for periods of time measured in at least decades. In addition to the pandemic clones of serogroups O1 and O139, which form a tight cluster of four ETs in subdivision Ia, MLEE analysis identified numerous apparent clonal lineages of non-O1 strains with intercontinental distributions. A clone of serogroup O37 that demonstrated epidemic potential in the 1960s is closely related to the pandemic O1/O139 clones, but the nontoxigenic O1 Inaba El Tor reference strain is not. A strain of serogroup O22, which has been identified as the most likely donor of exogenous *rfb* region DNA to the O1 progenitor of the O139 clone, is distantly related to the O1/O139 clones. The close evolutionary relationships of the O1, O139, and O37 epidemic clones indicates that new cholera clones are likely to arise by the modification of a lineage that is already epidemic or is closely related to such a clone.

For *Vibrio cholerae*, the causative agent of cholera and a natural inhabitant of aquatic environments, the conventional method of identifying and classifying strains is a serotyping scheme in which nearly 200 serogroups (or serovars) have been distinguished on the basis of epitopic variation in the cell surface lipopolysaccharide (LPS) (52). From an epidemiological standpoint, the species has been divided into serogroup O1 and serogroup non-O1 strains, which were long believed to differ in ability to cause epidemic cholera (4, 10). Historically, O1 strains have been responsible for all major epidemics, including seven pandemics (19), but in 1992 an epidemic clone of serogroup O139 (Bengal) appeared in southern India (30). It rapidly spread throughout much of Southeast Asia (1, 9) and reached western Africa in 1994 (37).

The emergence of the O139 clone with pandemic potential stimulated increased interest in the molecular basis of pathogenesis in *V. cholerae* and the degree to which genes determining serotype and virulence properties are subject to horizontal transfer and recombination among strains (17). Molecular genetic studies have shown that the origin of the O139 clone involved a complex rearrangement of the *rfb* region in a strain of O1 El Tor, which included deletion of genes responsible for the biosynthesis and assembly of the side chains of the O1 cell surface LPS and insertion of exogenous DNA mediating synthesis of the O139 LPS core (2, 3, 40–42) and a capsule (12, 45). The observations that strains with identical nucleotide sequences of the aspartate semialdehyde dehydrogenase gene (*asd*) may express different O antigens and that O1 isolates are heterogeneous in sequence provided further evidence of the horizontal transfer of genes mediating O-antigen synthesis (19). Most surprisingly, it was discovered that the CTX element, which includes the structural genes (*ctxA* and *ctxB*) for the subunits of cholera toxin, is the integrated genome of a filamentous bacteriophage, CTXø, and is transmissible (31, 46). Moreover, the bacterial receptor for CTXø, the toxin-coregulated pilus, is encoded by an operon (*tcp*) that is part of a transmissible pathogenicity island (20, 21). These findings raise the possibility that all strains of *V. cholerae* have the potential to become agents of epidemic cholera.

Previous research on the evolutionary genetics of *V. cholerae* has been primarily concerned with the identification and epidemiology of O1 and O139 strains that are responsible for cholera epidemics and pandemics. This work includes the application of multilocus enzyme electrophoresis (MLEE) to assess genotypic diversity in a collection of 181 O1 and 79 non-O1 strains (33) and the extensive use of this technique, in conjunction with ribotyping and restriction fragment length polymorphism analysis of the *ctxA* gene, to study various aspects of the molecular epidemiology of cholera in Latin America and elsewhere (8, 15, 43, 44). Additionally, ribotyping and comparative sequence analysis of the *asd* gene have been employed to reconstruct the evolutionary history of O1 clones involved in the sixth and seventh pandemics (18, 19).

We report here the results of an analysis of 397 isolates of *V. cholerae* by MLEE undertaken to determine the extent of genetic diversity in the species as a whole, the relationships of the epidemic O1 and O139 clones to strains of other serogroups, and the genetic population structure of the non-O1 segment of the species.

MATERIALS AND METHODS

Bacterial strains. This study was based on 397 strains received as *V. cholerae*. The sample included 143 strains in the serogroup reference collection maintained at the National Institute of Infectious Diseases in Japan (52). These strains were isolated from worldwide sources in the period from 1932 to 1993; 117 of them were recovered from humans, 13 from animals, 6 from river water, 3 from seawater, and 3 from unknown sources. The reference strains for serogroups O1 through O83 were provided by T. Cheasty, and those for serogroup O155 and serogroups O84 through O140, together with strain CA-385 (rough), which is used in serotyping, were obtained from T. Shimada.

A collection of 191 strains from Sonora, Tabasco, and 14 other states in Mexico was provided by the Instituto Nacional de Diagnóstico y Referencia

* Corresponding author. Mailing address: Institute of Molecular Evolutionary Genetics, Mueller Laboratory 516, Pennsylvania State University, University Park, PA 16802. Phone: (814) 234-8997. Fax: (814) 863-4706. E-mail: rks3@psu.edu.

Epidemiológicos, the Laboratorio Estatal de Salud Pública (Sonora), and the Laboratorio Regional de Salud Pública (Tabasco). A sample of 53 strains recovered from humans in Guatemala was obtained from the Instituto de Nutrición de Centroamérica y Panamá. Of the total of 244 strains from Mexico and Guatemala, 172 were recovered from humans, 41 from well and sewage water, 15 from fish, 7 from other environmental sources, 2 from food, and 7 from unspecified sources.

Five serogroup O139 strains from Thailand were furnished by P. Echeverria. From the Centers for Disease Control and Prevention (CDC), we received single isolates of O1 El Tor from Australia, Romania, Peru, and Louisiana and an O139 isolate from an imported human case of cholera in the United States in 1993.

The strains from Mexico and Guatemala, almost all of which were isolated in the period from 1991 to 1995, were serotyped in our laboratory by the standard method of Sakazaki and Donovan (32). Eight of these strains were of a serotype that was not then represented in the reference collection; for purposes of this paper, we have designated this serotype OA.

All of the strains used in this study have been deposited in the collection of the Facultad de Medicina, Universidad Nacional Autónoma de México (FMU).

MLEE. MLEE was performed by the methods described by Selander et al. (34). Seventeen enzyme loci were assayed for allelic variation: 6PG (6-phosphogluconate dehydrogenase), G6P (glucose 6-phosphate dehydrogenase), IDH (isocitrate dehydrogenase), NSP (nucleoside phosphorylase), ALD (alanine dehydrogenase), SHK (shikimate dehydrogenase), CAT (catalase), LAP (leucine aminopeptidase), GOT (glutamic-oxalacetic transaminase), ME (malic enzyme), MDH (malate dehydrogenase), PLP (phenylalanyl-leucine peptidase), PGI (phosphoglucose isomerase), HEX (hexokinase), PGM (phosphoglucomutase), IPO (indophenol oxidase), and THD (threonine dehydrogenase).

Electromorphs (mobility variants) were equated with alleles, an absence of enzyme activity was scored as a null allele, and distinctive allele profiles for the 17 loci were designated electrophoretic types (ETs).

Statistical analysis. From the allele profiles of the ETs, mean genetic diversity per locus (H) and pairwise genetic distance were calculated as described by Selander et al. (34), and dendrograms were constructed by the unweighted pair-group method with arithmetic mean (UPGMA). As a measure of multilocus population structure (linkage disequilibrium), we calculated an index of association of alleles (I_A) by using the equation $1 - V_O/V_E$, where V_O is the variance of the observed distribution of number of mismatched alleles between ET pairs and V_E is the mismatch variance expected when allele associations are random (linkage equilibrium) (24, 49). Calculations were made with computer programs written by T. S. Whittam (Pennsylvania State University, University Park).

Detection of the *ctxA* gene. Tests for the presence of the *ctxA* gene, which encodes a subunit protein of cholera toxin, were performed by colony blot assay with a digoxigenin-labeled oligonucleotide probe, according to the procedure of Maniatis et al. (23), and by PCR amplification. The nucleotide sequences of the probe and the primers were those specified by Shirai et al. (39).

RESULTS

Serogroups. Of the 244 strains from Mexico and Guatemala, 230 were of 59 serogroups and 14 were serologically nontypeable (NT). The reference collection currently includes nearly 200 serogroups (52), and thus about 30% of the described serotypic diversity of the species is exhibited by the Mexico-Guatemala strains.

In this paper, individual reference strains are referred to by their serogroup designations, marked with an asterisk (e.g., O155*).

Multilocus enzyme genotypes. All 17 loci assayed were polymorphic, with an average of 9.5 alleles per locus and a range of 4 (IPO) to 15 (CAT and LAP). Among the total sample of 397 isolates, 279 ETs, each representing a distinctive allele profile, were distinguished. The 142 serogroup reference strains were of 133 ETs, and the 244 Mexico-Guatemala strains were assigned to 154 ETs.

For each of the 17 enzymes assayed, an absence of activity was rare. Among the 279 ETs, only 26 null alleles, distributed over 12 of the loci, were recorded. Thus, only 0.5% of a total of 4,743 (17 × 279) alleles scored were nulls, and 10 of them were represented in the profiles of the two strains (ETs 276 and 277) that form evolutionary subdivision x (see below).

The mean genetic diversities per locus were 0.436 for all ETs and 0.430 for the ETs of the reference strains (Table 1).

Estimates of genetic distance among the 279 ETs are summarized in the dendrograms presented in Fig. 1 and 2. All but four of the ETs are members of two major divisions (designated I and II) that diverge from one another at a distance of about 0.7, which roughly corresponds to an 11-locus difference (Fig. 1). In addition, there are two lineages (labeled x and y) that branch from the major divisions and from one other at distances greater than 0.9. Each of these lineages includes two distantly related ETs.

TABLE 1. Measures of mean genetic diversity per locus and multilocus linkage equilibrium, based on ETs

Sample	No. of:		Mean no. of alleles	H^a	$I_A{}^b$
	Isolates	ETs			
All strains					
Total	397	279	9.5	0.436	1.248 ± 0.082*
Divisions I and II	392	275	7.5	0.421	0.728 ± 0.083*
Division I	380	263	6.7	0.397	0.237 ± 0.085*
Subdivision Ia					
Group A	103	37	3.4	0.297	0.248 ± 0.228
Group B	148	123	4.8	0.309	−0.048 ± 0.125
Group C	55	44	3.3	0.274	0.053 ± 0.210
Subdivision Ic	47	40	3.2	0.384	0.931 ± 0.217*
Division II	12	12	2.8	0.375	1.204 ± 0.399*
Reference strains[c]					
Total	142	133	7.5	0.441	1.759 ± 0.120*

[a] H, mean genetic diversity per locus.
[b] I_A, index of association of alleles. *, $P \leq 0.05$.
[c] Strain CA-385 (rough) not included.

Division II consists of 12 ETs, each of which is represented by a single serogroup reference strain. Seven of these strains are from India and the Far East, as follows: India, two strains, both from humans; China, one human isolate; the Philippines, one human strain; and Japan, two strains, from a crab and from river water. Three strains (O114*, O115*, and O116*) were cultured from river water (1979), a human (1980), and seawater (1978) in the United States; one isolate (O71*) was recovered from a bird in Denmark (1978); and one strain (O29*) was isolated from a human at an unspecified locality (1968).

The vast majority of strains are of ETs in division I, in which for purposes of reference we have designated six subdivisions (Ia to If), each consisting of a single ET or a group of related ETs.

Most of the 51 strains of the 40 ETs that form subdivision Ic are from Mexico and Guatemala, but 8 strains (representing 8 ETs) are from the serogroup reference collection, as follows: O50*, O75*, O78*, O82*, and O126* from India; O155* from Thailand; and O107* and O92* from Japan.

Subdivision Ia includes 82% of the reference collection strains and 78% of the Mexico-Guatemala strains. All but 10 of the 214 ETs in this subdivision form three branches (designated A, B, and C in Fig. 1) that diverge from one another at a genetic distance of about 0.4. With the singular exception of the O1 Inaba El Tor reference strain (ET 259, in subdivision Ie), all O1 strains in our sample, together with all O139 isolates, are of 4 ETs that form a tight cluster in group A of subdivision Ia, which consists of a total of 37 ETs (Fig. 2).

Multilocus population structure. Estimates of multilocus association of alleles for all 279 ETs and for the ETs in several segments of the dendrograms (Fig. 1 and 2) are shown in Table

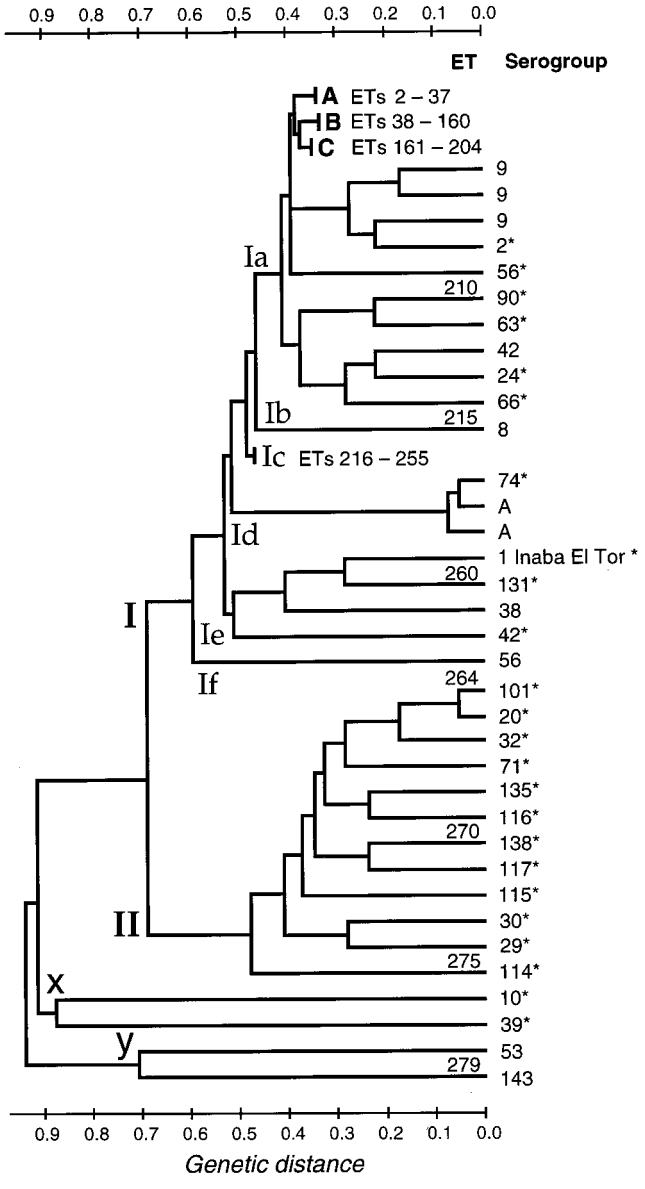

FIG. 1. Dendrogram showing genetic relationships of the ETs of *V. cholerae*, based on MLEE analysis (17 loci). The dendrogram was constructed from a matrix of pairwise genetic distances by the UPGMA method. The lineages of subdivision Ic and of groups A, B, and C of subdivision Ia are truncated. The relationships of the 37 ETs in group A are shown in the dendrogram in Fig. 2.

1. For the total sample of 279 ETs, the 133 ETs of the reference strains, and the ETs of each of the divisions I and II and subdivision Ic, there is evidence of significant nonrandom associations of alleles (linkage disequilibrium), but allele associations are not demonstrably nonrandom for the ETs of groups A, B, and C in subdivision Ia.

Relationships of O1 and O139 epidemic strains. With allowance for a difference in the panels of enzymes employed in MLEE analysis, our findings for the epidemic O1 and O139 strains are fully consistent with those reported by Evins et al. (15) and in earlier studies by Wachsmuth's group at the CDC (43, 44). These workers assayed variation in 16 enzymes, only 9 of which were included in our panel of 17 enzymes. In the interest of consistency, we have numbered the ETs of our O1 and O139 isolates to correspond to the ET designations of the CDC group (Table 2; Fig. 2).

ET 4 is represented by 13 isolates of O1 Inaba El Tor. Eleven of these strains are from Mexico (Quintana Roo, Campeche, Tabasco, Veracruz, Puebla, and Hidalgo), one is from Guatemala, and one is from Peru. ET 4 marks the original, or first-wave, Latin American epidemic clone (15).

ET 3 includes 2 strains of O1 Inaba El Tor, 38 strains of O1 Ogawa El Tor, and 7 strains of O139. The two Inaba El Tor strains (FMU strains 90501 and 90500) were recovered from humans in Tabasco in 1991 and 1993. The sample of O1 Ogawa El Tor isolates includes the reference strain, 35 isolates from Mexico (Tabasco, Morelos, and the state of México), a strain from Australia, and an isolate from Romania. ET 3 is the seventh-pandemic type, a clone that in Latin America was first identified in Mexico in 1991 and is now widely distributed in

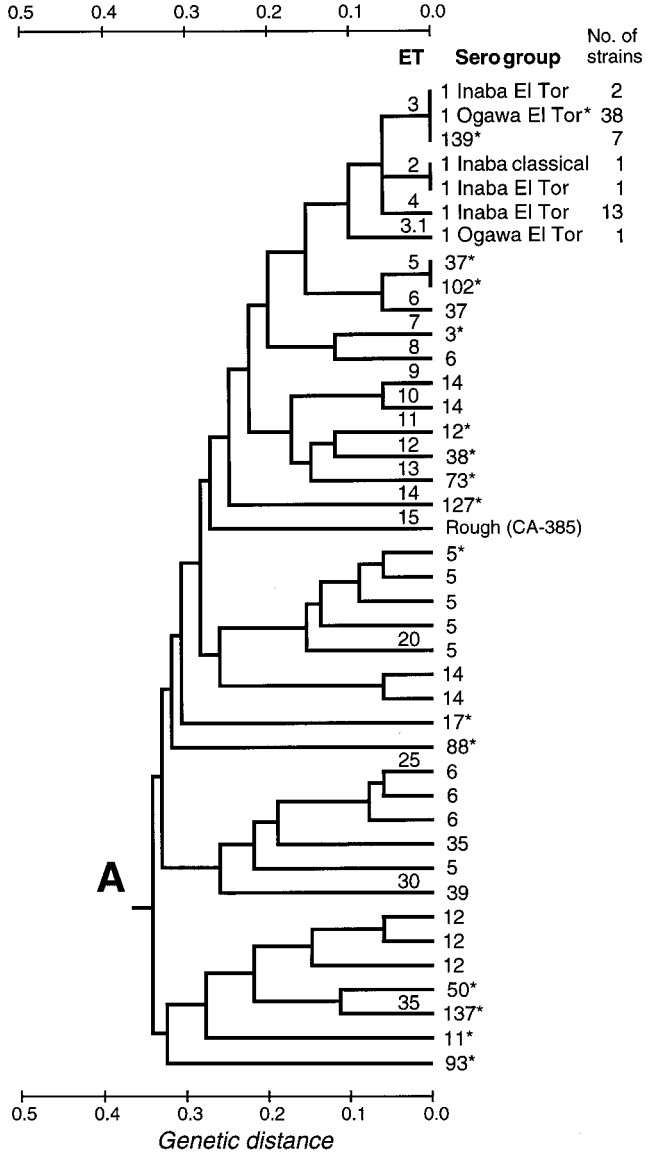

FIG. 2. Dendrogram showing genetic relationships of the 37 ETs of *V. cholerae* in group A of subdivision Ia of division I, based on MLEE analysis (17 loci). The dendrogram was constructed from a matrix of pairwise genetic distances by the UPGMA method.

TABLE 2. Allele profiles of 16 ETs in division I

ET	Serogroup or reference strain	*n*	Allele at indicated enzyme locus[a]																
			6PG	G6P	IDH	NSP	ALD	SHK	CAT	LAP	GOT	ME	MDH	PLP	PGI	HEX	PGM	IPO	THD
Subdivision Ia, group A																			
3	O1[b]/O139	47	3	4	3	3	3	1	3	3	3	3	2	4	**3**	3.5	3	3	1
2	O1[c]	2	3	4	3	3	3	1	3	**4**	3	3	2	4	**3**	3.5	3	3	1
4	O1[d]	13	3	4	3	3	3	1	3	**4.5**	3	3	2	4	**3**	3.5	3	3	1
3.1	O1[e]	1	3	**3.8**	3	3	3	1	3	3	3	3	2	4	**3**	3.5	3	3	1
5	O37*/O102*	2	3	4	3	3	3	1	3	**4**	3	3	2	4	4	3.5	3	3	1
6	O37	1	3	4	3	3	3	1	3	**4**	**4**	3	2	4	4	3.5	3	3	1
7	O3*	1	3	4	3	3	3	1	3	**3.5**	3	3	2	4	4	3.5	**5**	3	1
8	O6	1	**2**	4	3	3	3	1	3	3	3	3	2	4	4	3.5	**5**	3	1
9	O14	1	3	4	3	3	3	1	3	3	3	3	**3**	4	4	**4**	**3.5**	3	1
10	O14	1	3	4	3	3	3	1	3	3	3	3	**3**	4	4	**4**	3	3	1
11	O12*	1	3	4	3	3	3	1	3	3	3	3	2	**5.5**	4	3.5	3	3	1
12	O38*	1	3	4	3	3	3	1	3	3	3	3	2	4	4	**3**	3	3	1
13	O73*	1	3	4	3	3	3	1	3	3	3	3	**3**	**6**	4	**3**	3	3	1
14	O127*	1	**4**	4	3	3	3	1	3	3	3	**2**	2	4	4	**4**	3	3	1
15	Rough[f]	1	**2**	4	3	3	3	1	3	**4**	**4**	3	2	4	4	**3**	3	3	1
Subdivision Ie																			
259	O1*[g]	1	**3.5**	**2**	3	3	**0**	1	**1**	**4.5**	3	3	2	**5**	**3**	**1**	3	3	1

[a] Minority alleles are indicated in boldface type.
[b] Thirty-nine isolates of O1 Ogawa El Tor and 2 isolates of O1 Inaba El Tor.
[c] One isolate of O1 Inaba classical and 1 isolate of O1 Inaba El Tor.
[d] Thirteen isolates of O1 Inaba El Tor from Mexico, Guatemala, and Peru.
[e] One isolate of O1 Ogawa El Tor from Tabasco.
[f] Strain CA-385.
[g] O1 Inaba El Tor reference strain.

Mexico and Central America (15). The Australian isolate (CDC 2463-88) was distinguished, as ET 1, from isolates of ET 3 by Evins et al. (15) on basis of its carrying a different allele for diaphosrase 1. ET 1 marks a distinctive Australian toxigenic clone (13, 43).

Included in ET 2 are an O1 Inaba El Tor isolate (CDC 2164-78) collected in Louisiana in 1978 and an O1 Inaba classical strain (FMU 87295/0) recovered from a tourist returning to the United States from Cancún, Quintana Roo, in 1983. ET 2 marks a toxigenic clone that is endemic to the Gulf Coast of Mexico and the United States (15).

With our panel of 17 enzymes, the sole basis for distinguishing ETs 2, 3, and 4 is allelic variation at the LAP locus (Table 2). One O1 Ogawa El Tor isolate (from Tabasco) is ET 3.1, which differs from ET 3 in having a distinctive G6P allele. This variant genotype was not detected in previous studies.

Relationship of serogroup O37 strains. The reference strains O37* (India, 1969) and O102* (China, 1988), both of which are of ET 5 (Fig. 2), differ from strains of the O1/O139 cluster (ETs 2 to 4) at a single locus, that for PGI, and share the LAP 4 allele with strains of ET 2 (Table 2). A second O37 strain (from Guatemala) is of ET 6, which differs from ET 5 in having a 4 allele rather than a 3 allele at the GOT locus. Two other O37 isolates in the collection, both of which were cultured from well water in Campeche, represent ETs 75 and 149, which are in group B of subdivision Ia.

Serotypic diversity among non-O1 strains of the same ET. In addition to ETs 2 to 4 of the epidemic O1 and O139 clones, MLEE identified nine ETs that are each represented by non-O1 isolates from different continents or other major land masses (Table 3). In all cases, the strains from different continents are of different serogroups. For example, ET 196 (a member of group C of subdivision Ia) is represented by a total of seven isolates, including four (O15*, O47*, O51*, and O53*) recovered from humans in India between 1968 and 1974, a strain (O68*) obtained from seawater in Japan in 1978, and two isolates (both O7) cultured from well water in Campeche and a fish in Sonora in 1992 and 1993. The inference is that ET 196 marks a widely distributed clone that had persisted for at least 24 years, with several modifications in serotype through mutation or recombination of genes of the *rfb* region.

ETs 128 and 131 are each represented by a strain collected in India in 1968 and an isolate recovered in Campeche or Guatemala in the early 1990s. It is noteworthy that ET 131 differs at only a single locus (that for LAP) from ET 132 of the O14 reference strain, which was recovered in India in 1964 (see Table 5). Thus, ETs 131 and 132 are members of a clonal lineage that was extant for at least 28 years.

Nine ETs are represented by pairs or multiple strains from different states in Mexico (Table 4). Five of these ETs are represented by strains of the same serogroup, and four of them are represented by strains of different serogroups. For example, ET 16 was represented by seven O5 isolates from Hidalgo, Tabasco, Yucatán, and an unspecified locality in Mexico. Significantly, ET 16 differs by only one locus (that for PLP) from ET 17 of the O5 reference strain, which was collected in India in 1964 (see Table 5).

Serotypic diversity among strains of closely related ETs. We identified 19 cases in which pairs of ETs that differ at a single enzyme locus are represented by strains collected on different continents or major land masses (Table 5). In four of these cases, the strains are of the same serotype; these are O5* from India (1964) and seven O5 isolates from Hidalgo and other states in Mexico (1991 and 1992), O14* from India (1964) and O14 from Guatemala (1993), O37* from India (1969) and O37 from Guatemala (1993), and O44* from India (1973) and O44 from Chiapas (1991). In all other cases, the strains are of different serotypes or one of them was NT.

TABLE 3. ETs represented by non-O1 isolates from different continents

ET	No. of isolates	Serogroup	Locality (date)	Source	FMU strain no.
196	7	O15*	India (1968)	Human	88554
		O47*	India (1973)	Human	88586
		O51*	India (1973)	Human	88590
		O53*	India (1974)	Human	88592
		O68*	Japan (1978)	Seawater	88607
		O7	Campeche (1992)	Well water	87242
		O7	Sonora (1993)	Fish	88354
256	6	O74*	India (1979)	Human	88613
		OA	Hidalgo (1991)	Human	87250
		OA	Hidalgo (1991)	Human	87256
		OA	Veracruz (1991)	Human	87264
		OA	Puebla (1991)	Human	87246
		OA	Guatemala (1993[a])	Human	88778
128	2	O22*	India (1968)	Human	88561
		O6	Campeche (1991)	Human	87268
131[b]	2	O34*	India (1968)	Human	88573
		O14	Guatemala (1993[a])	Human	88729
247	2	O92*	Japan (1987)	River water	88631
		O68	Tabasco (1992)	Human	87311
73	2	O23*	India (1971)	Human	88562
		O26*	Philippines (1972)	Human	88565
5	2	O37*	India (1969)	Human	88576
		O102*	China (1988)	Human	88641
129	2	O130*	India (1981)	Human	88669
		O104*	China (1988)	Human	88643
65	2	O125*	India (1981)	Human	88664
		O132*	Thailand (1981)	Human	88671

[a] Date of receipt at the Departamento de Salud Pública de la Facultad de Medicina.

[b] ET 131 differs from ET 132 of strain O14* (India, 1964) at a single locus (that for LAP) (Table 5).

Genetic diversity within serogroups. MLEE analysis demonstrated that strains of the same serogroup may belong to two or more widely divergent ET lineages. Thus, for example, strains of serogroup O29 were assigned to four ETs that occur in division I (ET 50 and ET 52 in group B and ET 173 in group C) and division II (ET 274), and O53 strains are found in group C of division I (ET 196) and also in lineage y (ET 278). Some estimated levels of genetic diversity between or among the ETs of strains of the same serogroup are shown in Table 6. In several cases, the estimated diversity is at least equivalent to that obtained for the 279 ETs identified among all 397 isolates.

Distribution of the *ctxA* gene. When tested with the *ctxA* probe, 13 of the 143 reference strains were positive, as shown in Table 7. With PCR amplification of the gene, the same strains were positive, with the exception of O1 Inaba El Tor*. Among 104 of 254 nonreference strains that were randomly selected for testing, 2 isolates of O1 Ogawa El Tor, 3 isolates of O1 Inaba El Tor, an isolate of O1 Inaba classical, and an isolate of O6 were positive for *ctxA* by both colony blot assay and PCR amplification.

All but 3 of the total of 21 *ctxA*-positive strains are members of subdivision Ia of division I (Fig. 1 and 2); the exceptions are the reference strains O1 Inaba El Tor*, in subdivision Ie, and O135* and O138*, in division II.

DISCUSSION

Species limits. Strains of ETs 276 to 279 in the deep lineages x and y (Fig. 1) are sufficiently differentiated from all other strains as to raise the question of whether they should be included in the species *V. cholerae*. It is likely that assessment of genomic relatedness by DNA-DNA hybridization would show relative degrees of annealing with other strains somewhat below the 70% standard adopted for species inclusion by the CDC (6).

Genetic diversity. The estimate of 0.436 for the mean genetic diversity per locus among the 279 ETs of *V. cholerae* represented in the present study is larger than the comparable value of 0.343 reported for the *Escherichia coli* reference collection (35) but smaller than the corresponding value of 0.627 obtained for *Salmonella enterica* (36). In a previous MLEE study of allelic variation at 13 enzyme loci among 260 isolates of *V. cholerae* (most of which were serogroup O1), Salles and Momen (33) detected an average of 4.3 alleles per locus, identified 73 ETs, and estimated the mean genetic diversity per locus as 0.326.

Genetic structure of populations. Comparisons of the observed and expected variances of the mismatch distributions for ETs at several hierarchical levels of dendrogram structure yielded only limited evidence of linkage disequilibrium (Table 1). The cases in which the observed variance exceeded the upper 95% confidence limit of the variance expected under random association of alleles were those involving all 279 ETs, the 275 ETs of divisions I and II combined, the 263 ETs of division I, the 12 ETs of division II, and the 40 ETs of subdivision Ic of division I. Within each of the groups A, B, and C of subdivision Ia, which include 204 ETs (Fig. 1), significant levels of nonrandom association were not demonstrable. The inference is that, at least among the strains of ETs in subdivision Ia, the rate of horizontal transfer and recombination of housekeeping enzyme genes is sufficiently high to prevent the development and long-term maintenance of distinctive allele complexes. On the whole, it seems likely that the frequency of recombination, both intragenic and assortative (50), of housekeeping genes in *V. cholerae* is somewhat higher than in either *E. coli* (48) or *S. enterica* (36), a conclusion also reached by Karaolis et al. (19) from a comparative sequence analysis of the *asd* gene. But even within subdivision Ia, clonal lineages may persist for periods of time measured in at least decades. The most obvious examples are the epidemic and pandemic strains of serogroups O1 and O139 (ETs 2 to 4), but our analysis identified numerous clones and clonal lineages of non-O1 strains with widespread, if not global, distributions (Tables 3 to 5).

A factor that has not been evaluated in studies of the genetic structure of bacterial populations on the basis of MLEE data is the convergent evolution of electromorphs, which cannot be equated with isoalleles. Similarity in electrophoretic mobility resulting from convergence in the net electrostatic charge of an enzyme will lessen the likelihood that linkage disequilibrium is detected from the analysis of MLEE data. Studies of sequence variation in several housekeeping enzymes among multiple strains of *E. coli* and *S. enterica* (5, 26–28, 47) have shown that individual electromorphs may exhibit substantial heterogeneity in amino acid sequence, much of which clearly stems from convergence rather than mutational divergence from a common ancestral sequence.

There was already evidence for recombination of genes of

TABLE 4. ETs represented by isolates from two or more states in Mexico

ET	No. of isolates	Serogroup	Locality (date)	Source	FMU strain no.
Of same serogroup					
16[a]	7	O5	Hidalgo (1991)	Human	87243
		O5	Hidalgo (1991)	Human	87259
		O5	Tabasco (1991)	Human	87139
		O5	Tabasco (1991)	Human	87672
		O5	Mexico (1991)	Human	87288
		O5	Tabasco (1992)	Human	87306
		O5	Yucatán (1992)	Human	87291
228	2	O149	Guanajuato (1991)	Human	87297
		O149	Tabasco (1991)	Human	87673
219	2	O149	Yucatán (1991)	Human	87262
		O149	Zacatecas (1991)		87299
107	2	O64	Veracruz (1991)	Human	87257
		O64	Hidalgo (1991)	Human	87289
124	2	O24	Veracruz (1991)	Human	87282
		O24	Tabasco (1992)	Human	87434
Of different serogroups					
181	4	O35	Sonora (1993)	Sewage water	88374
		O35	Sonora (1993)	Sewage water	88375
		O42	Sonora (1993)	Fish	88351
		NT	Guanajuato (1991)	Human	87287
52	3	O79	Zacatecas (1991)	Human	87295
		O43	Tabasco (1992)	Human	87307
		O29	Sonora (1993)	Sewage water	88371
234	2	O5	Campeche (1992)	Well water	87240
		O62	Tabasco (1992)	Human	87304
171	2	O41	Sonora (1993)	Septic tank	88366
		NT	Guerrero (1991)		87271

[a] ET 16 differs from ET 17 of strain O5* (India, 1964) at a single locus (that for PLP) (Table 5).

the *rfb* region of *V. cholerae*, based on studies of relatively small numbers of strains and serogroups. Our observation that strains of the same serogroup frequently are found in divergent, even distantly related, lineages supports earlier evidence (2, 19) that the *rfb* genes are subject to horizontal transfer and further suggests that this process occurs with relatively high frequency. Convergence in serotype is, of course, an alternative explanation, but reasoning by analogy from the lack of evidence for convergence in epitope structure in the serologically diverse flagellins of *S. enterica* (22), we favor the first hypothesis. The issue can be settled by comparative sequencing of the epitope-encoding segments of the *rfb* region.

The fact that strains of the same ET may express different O antigens can be explained by recombination or by spontaneous mutation of the genes encoding O somatic properties.

Epidemic non-O1 clones. There are two examples of epidemic *V. cholerae* expressing a non-O1 antigen. The first is the serogroup O139 clone, which emerged in India and Bangladesh through modification of the El Tor O1 pandemic strain by acquisition of genes mediating the synthesis of the O139 LPS and a polysaccharide capsule. The second is the O37 strain that reportedly was responsible for a large outbreak of cholera in the Sudan in 1968 (54). By IS*1004* fingerprinting, Bik et al. (3) determined that an O37 isolate from the Sudan is closely related to classical O1 strains. The O37 reference strain (ET 5), which was recovered from a patient in India in 1969, presumably represents the same clone as the O37 Sudan strain. As determined by MLEE, it is closely related to O1 El Tor and other epidemic O1 strains (Fig. 2), thus confirming the result obtained by IS*1004* fingerprinting. In fact, ET 5 is distinguished from ETs 2 to 4 of the O1/O139 cluster solely by possession of a 4 allele (versus a 3 allele) at the PGI locus. It carries the *ctxA* gene and expresses cholera toxin. Yamamoto et al. (53) reported that the amino acid sequence of the cholera toxin produced by O37 strain S7 differs from that of most O1 strains in having single substitutions in both the CtxA and CtxB segments (29), which are presumed to cause the formation of an unusually large subunit B oligomer. Recently, Karaolis et al. (20) reported that, almost uniquely among non-O1 strains, a Sudan 1968 outbreak strain carries a chromosomal pathogenicity island that is characteristic of epidemic and pandemic strains.

Honma et al. (16) studied an O37 isolate that produces a hemolysin (O37-Hly) that is antigenically similar to O1 El Tor hemolysin (El Tor-Hly) but different in molecular size, hemolytic activity, and glucose-binding capacity. The gene encoding O37-Hly differs from that encoding O1 El Tor-Hly by the presence of a 4-bp insertion that generates a premature stop codon in the downstream region. Thus, the O37-Hly is a truncated derivative of O1 El Tor-Hly, sharing 90% of the N-terminal region.

In the Mexico-Guatemala collection, there are three O37

TABLE 5. Serogroups and geographic sources of non-O1 strains representing pairs of ETs that differ at a single locus

Strains from different continents					Strains from same continent				
ET	Serogroup	Locality (date)	FMU strain no.	Locus	ET	Serogroup	Locality (date)	FMU strain no.	Locus
17	O5*	India (1964)	88544	PLP	194	O4*	India (1932)	88543	LAP
16	O5	Hidalgo (1991)[a]	87243	PLP	193	O96*	India (1976)	88635	LAP
132	O14*	India (1964)	88553	LAP	166	O6*	India (1962)	88545	PGM
131	O14	Guatemala (1993[b])	88729	LAP	165	O79*	India (1976)	88618	PGM
5	O37*	India (1969)[c]	88576	GOT	132	O14*	India (1964)	88553	LAP
6	O37	Guatemala (1993[b])	88777	GOT	131	O34*	India (1968)	88573	LAP
109	O44*	India (1973)	88583	PLP	196	O15*	India (1968)	88554	MDH
108	O44	Chiapas (1991)	87245	PLP	195	O35*	India (1969)	88574	MDH
102	O18*	India (1964)	88557	PGI	115	O54*	India (1974)	88593	LAP
101	O14	Campeche (1992)	87440	PGI	114	O94*	India (1976)	88633	LAP
152	O16*	India (1971)	88555	PGI	168	O95*	India (1976)	88634	CAT
151	O48	Tabasco (1991)	87675	PGI	167	O129*	India (1981)	88668	CAT
					169	O134*	India (1991)	88673	CAT
122	O49*	India (1974)	88588	6PG					
121	NT	Michoacan (1992)	87293	6PG	155	O112*	Japan (1989)[h]	88651	IDH
123	NT	Sonora (1993)	88352	6PG	154	O113*	Japan (1989)[h]	88652	IDH
192	O60*	India (1975)	88599	IDH	31	O12	Tabasco (1992)	87432	G6P
191	O68	Sonora (1993)	88392	IDH	32	O12	Guatemala (1993[b])	88750	G6P
134	O55*	India (1975)	88594	LAP	107	O64	Veracruz (1991)	87257	LAP
133	NT	Hidalgo (1991)	87263	LAP	106	O64	Chiapas (1991)	87272	LAP
186	O81*	India (1978)	88620	6PG	229	O149	Guanajuato (1991)	87290	CAT
185	O80	Campeche (1992)	87437	6PG	228	O149	Guanajuato (1991)	87297	CAT
					230	O155	Campeche (1992)	90594	CAT
73	O23*	India (1971)[d]	88562	G6P					
72	O93	Guatemala (1993[b])	88773	G6P	22	O14	Campeche (1992)	87661	PGM
					21	O14	Campeche (1992)	87778	PGM
256	O74*[c]	India (1979)	88613	IPO					
257	OA	Guatemala (1993[b])	88733	IPO	59	O44	Sonora (1993)	88380	PGM
					58	O5	Sonora (1993)	88382	PGM
247	O92*	Japan (1987)[f]	88631	CAT					
246	O31	Tabasco (1992)	87303	CAT	157	O141	Hidalgo (1991)	87274	LAP
					156	O82	Sonora (1993)	88389	LAP
136	O120*	Japan (1991)	88659	PGM					
135	O52	Sonora (1993)	88356	PGM	233	NT	Hidalgo (1991)	87298	CAT
					234	O5	Campeche (1992)	87240	CAT
53	O108*	Japan (1989)	88637	PGM					
52	O79	Zacatecas (1991)	87295	PGM	225	O41	Tabasco (1992)	87309	LAP
54	O105	Sonora (1993)	88387	PGM	226	NT	Tabasco (1992)	87305	LAP
98	O122*	Romania (1980)	88661	CAT	221	O35	Veracruz (1991)	87279	THD
97	O97	Veracruz (1991)	87283	CAT	220	O151	Sonora (1993)	88367	THD
265	O20*	India (1962)	88559	IDH	51	O43	Guatemala (1993[b])	88727	CAT
264	O101*	China (1988)	88640	IDH	50	O40	Guatemala (1993[b])	88744	CAT
128	O22*	India (1968)[g]	88561	PGM					
127	O83*	India (1978)	88622	PGM					
129	O104*	China (1988)	88643	PGM					
200	O86*	Philippines (1981)	88625	GOT					
199	O133*	India (1991)	88672	GOT					

[a] There are six additional O5 isolates of ET 16 from Hidalgo, Yucatán, Tabasco, and "Mexico" (Table 4).
[b] Date of receipt at the Departamento de Salud Pública de la Facultad de Medicina, UNAM.
[c] Strain O37* (India, 1969) is of the same ET as O102* from China (1988) (Table 3).
[d] Strain O23* (India, 1971) is of the same ET as strain O16* from the Philippines (1972) (Table 3).
[e] Five additional OA isolates from Mexico and Guatemala (1991 to 1993) are of the same ET as strain O74* from India (1979) (Table 3).
[f] Strain O92* (Japan, 1987) is of the same ET as a strain of serotype O68 from Tabasco (1992) (Table 3).
[g] Strain O22* (India, 1968) is of the same ET as an O6 strain from Campeche (1991) (Table 3).
[h] Source: rat.

TABLE 6. Mean genetic diversity per locus among multiple ETs of the same serogroup

Serogroup	No. of:		Mean genetic diversity per locus
	Isolates	ETs	
O12	4	4	0.206
O37	4	4	0.314
O6	12	10	0.316
O8	14	4	0.333
O5	14	8	0.345
O38	4	2	0.412
O155	9	7	0.423
O42	3	3	0.490
O29	4	4	0.559
O39	5	4	0.618
O30	2	2	0.824
O53	3	2	1.000

isolates, one of which (ET 6) is almost identical in MLEE genotype (it carries a GOT 4 rather than a GOT 3 allele) to the O37 reference strain (ET 5) from India but apparently lacks the *ctxA* gene. It was isolated from a patient in Guatemala. The two other O37 isolates, both of which were cultured from well water in Campeche, are distantly related (six- and seven-locus differences) to both the reference O37 and Guatemala O37 strains, as well as to one another (four-locus difference), and neither one carries the *ctxA* gene.

It is noteworthy that strain O102*, which was recovered from a patient with diarrhea in China in 1988, is identical in MLEE genotype (ET 5) to strain O37* but apparently does not carry the *ctxA* gene.

In sum, there is a clone of serogroup O37 that has epidemic potential and was present in Africa and India in 1968 and 1969. Because it is closely related to O1 El Tor and the other O1 pandemic clones, it apparently represents a case similar to that of the O139 clone, in which an already-established pathogenic lineage of serogroup O1 acquired a new serotype by horizontal DNA transfer and rearrangement of the *rfb* region genes. The O37 strain from Guatemala may be an offshoot of this clone in which the CTX genetic element has been deleted. The two O37 strains from Campeche presumably have independent acquisitions of the O37 polysaccharide gene region.

O1 Inaba El Tor reference strain. The O1 Inaba El Tor reference strain (ET 259), which does not produce cholera toxin although it carries at least part of the *ctxA* gene, is not closely related to the O1/O139 cluster of pandemic strains or to the toxigenic O37 clone (ET 5). According to T. Shimada (38a), the reference strain is the NIH 35-a-3 isolate listed by Burrows et al. (7) as one of the strains used for vaccine preparation by the U.S. Army in the early 1940s. It was received from the Central Research Institute in Kasauli, India, in 1942, without indication of collection date or source of isolation. Perhaps it is related to the O1 strain that caused a cholera-like disease in Hong Kong in the 1950s (43).

Relationships of serogroup O1 strains. Colwell et al. (11) hypothesized that non-O1 cells may convert to the O1 serotype and vice versa under suitable conditions, a possible strategy for survival in the environment. As noted earlier, O37* and O102* (both of ET 5) may represent cases in which O1 clones have acquired new serotypes. In our collection, the only apparent case of conversion of a non-O1 strain to the O1 serotype (apart from O139) is the O1 Inaba El Tor reference strain (ET 259), which occurs in subdivision Ie (Fig. 1) and is distantly related to the O1 epidemic strains (ETs 2 to 4) in group A of subdivision Ia (Fig. 2).

Source of *rfb* region DNA in the emergence of the epidemic O139 clone. The putative source of the exogenous *rfb* region DNA that was involved in the transformation of an O1 El Tor strain to the epidemic O139 clone has been identified as a strain of serogroup O22, O141, or O155 on the basis of serotypic cross-reactions with O139 (2, 38, 42). Molecular analysis

TABLE 7. Strains testing positive for *ctxA*[a]

Serogroup	FMU strain no.	ET	Division or subdivision (group)	Source
Reference				
O1 Ogawa El Tor*		3	Ia (A)	India (1941), human
O1 Inaba El Tor*		259	Ie	India (1942), human
O19*		41	Ia (B)	
O33*		204	Ia (C)	India (1968), human
O37*		5	Ia (A)	India (1969), human
O43*		43	Ia (B)	India (1973), human
O44*		109	Ia (B)	India (1972), human
O54*		115	Ia (B)	India (1974), human
O105*		116	Ia (B)	India (1988), human
O106*		64	Ia (B)	China (1988), human
O135*		268	II	India (1992), human
O138*		270	II	Japan (1992), crab
O139*		3	Ia (A)	India (1993), human
Nonreference				
O1 Ogawa El Tor	87668	3	Ia (A)	Morelos, human
O1 Ogawa El Tor	90334	3	Ia (A)	Mexico, human
O1 Inaba El Tor	87269	4	Ia (A)	Campeche, human
O1 Inaba El Tor	88696	2	Ia (A)	Louisiana, human
O1 Inaba El Tor	90500	3	Ia (A)	Tabasco, human
O1 Inaba classical	87395/0	2	Ia (A)	Quintana Roo, human
O6	88751	27	Ia (A)	Guatemala, human
O139	88230	3	Ia (A)	United States, imported case

[a] All strains were positive when tested with the *ctxA* probe and by PCR, with the exception of O1 Inaba El Tor*, which was probe positive but PCR negative.

showed that, in common with O139, they have two open reading frames in the *rfaD* region that are lacking in O1 strains.

Through study of the gene content and organization of the *rfb* region adjacent to IS*1358* in strains of O139 and 13 other serogroups, including O22 and O155, Dumontier and Berche (14) recently identified the clone represented by strain O22* (Shimada strain 169-68, from India) as the most likely donor, although the possibility of a multistep rearrangement in the recipient O1 strain cannot be excluded. As determined by MLEE analysis, O22* (ET 128) falls in group B of subdivision Ia and differs in ET from the epidemic O1 and O139 strains at four or five loci.

In our sample of strains, there were nine serogroup O155 isolates belonging to seven ETs. ET 224 of the O155 reference strain (Thailand, 1993) and five other ETs (represented by isolates from Tabasco and Campeche) are in subdivision Ic, where, however, they do not form a tight cluster. The remaining ET, which is represented by two isolates from Sonora, is in group B of subdivision Ia. Thus, strains of serogroup O155 belong to a moderately diverse group of ETs, none of which is closely related to the ETs of the epidemic O1 and O139 clones. This suggests that genes mediating expression of the O155 LPS antigen are transferred with relatively high frequency.

Genesis of epidemic clones. Because genes for the major virulence factors can be transferred horizontally and antigenic conversion can be achieved by the acquisition and loss of *rfb* genes, there is the formal possibility that any *V. cholerae* cell could be transformed into a virulent strain, even an epidemic one (19, 25). However, the close evolutionary relationships of the O1, O139, and O37 epidemic clones indicate that new epidemic or other strongly virulent clones are likely to arise by the modification of a lineage that is already epidemic or is closely related to such a clone. Thus, O139 evolved from an El Tor O1 clone by acquisition of a transposon carrying genes for the O139 LPS and a polysaccharide capsule and deletion of most of the genes mediating synthesis of the serotype O1 LPS (see the review by Rubin et al. [31]). Also, the toxigenic O37 clone that caused outbreaks in the Sudan and India in the late 1960s is closely related to the cluster of O1/O139 epidemic clones. Analogously, *E. coli* O157:H7, which emerged as an agent of hemorrhagic colitis by acquisition of the O157 antigen, a Shiga-like toxin, and the enterohemorrhagic *E. coli* plasmid, is an evolutionary derivative of an O55:H7 clone that is associated with infantile diarrhea (51).

ACKNOWLEDGMENTS

We thank Thomas Cheasty, Peter Echeverria, Alma Rosa González, Sergio León, Claudio Lezana, José Luis Navarro-Heinze, and Toshio Shimada for supplying strains and Delia Licona and José Luis Méndez for technical assistance in the laboratory.

This research was supported by grants from the Consejo Nacional de Ciencia y Tecnología (project 2397PB); the Dirección General de Apoyo al Personal Académico, UNAM (project IN211496); and the National Institutes of Health (AI-22144).

REFERENCES

1. **Albert, M. J., A. K. Siddique, M. S. Islam, A. S. Faruque, M. Ansaruzzaman, S. M. Faruque, and R. B. Sack.** 1993. Large outbreak of clinical cholera due to *Vibrio cholerae* non-O1 in Bangladesh. Lancet **341:**704.
2. **Bik, E. M., A. E. Bunschoten, R. D. Gouw, and F. R. Mooi.** 1995. Genesis of the novel epidemic *Vibrio cholerae* O139 strain: evidence for horizontal transfer of genes involved in polysaccharide synthesis. EMBO J. **14:**209–216.
3. **Bik, E. M., R. D. Gouw, and F. R. Mooi.** 1996. DNA fingerprinting of *Vibrio cholerae* strains with a novel insertion sequence element: a tool to identify epidemic strains. J. Clin. Microbiol. **34:**1453–1461.
4. **Blake, P. A., R. E. Weaver, and D. G. Hollis.** 1980. Diseases of humans (other than cholera) caused by vibrios. Annu. Rev. Microbiol. **34:**341–367.
5. **Boyd, E. F., K. Nelson, F.-S. Wang, T. S. Whittam, and R. K. Selander.** 1994. Molecular genetic basis of allelic polymorphism in malate dehydrogenase (*mdh*) in natural populations of *Escherichia coli* and *Salmonella enterica*. Proc. Natl. Acad. Sci. USA **91:**1280–1284.
6. **Brenner, D. J., A. G. Steigerwalt, P. Epple, W. F. Bibb, R. M. McKinney, R. W. Starnes, J. M. Colville, R. K. Selander, P. H. Edelstein, and C. W. Moss.** 1988. *Legionella pneumophila* serogroup Lansing 3 isolated from a patient with fatal pneumonia, and descriptions of *L. pneumophila* subsp. *pneumophila* subsp. nov., *L. pneumophila* subsp. *fraseri* subsp. nov., and *L. pneumophila* subsp. *pascullei* subsp. nov. J. Clin. Microbiol. **26:**1695–1703.
7. **Burrows, W., A. N. Mather, M. E. Elliott, and S. M. Wagner.** 1946. Studies on immunity to Asiatic cholera. J. Infect. Dis. **79:**159–167.
8. **Chen, F., G. M. Evins, W. L. Cook, R. Almeida, N. Hargrett-Bean, and K. Wachsmuth.** 1991. Genetic diversity among toxigenic and nontoxigenic *Vibrio cholerae* O1 isolated from the Western Hemisphere. Epidemiol. Infect. **107:**225–233.
9. **Chongsa-nguan, M., W. Chaicumpa, P. Moolasart, P. Kandhasingha, T. Shimada, H. Kurazono, and Y. Takeda.** 1993. *Vibrio cholerae* O139 Bengal in Bangkok. Lancet **342:**430–431.
10. **Colwell, R. R.** 1996. Global climate and infectious disease: the cholera paradigm. Science **274:**2025–2031.
11. **Colwell, R. R., A. Huq, M. A. R. Chowdhury, P. R. Brayton, and B. Xu.** 1995. Serogroup conversion of *Vibrio cholerae*. Can. J. Microbiol. **41:**946–950.
12. **Comstock, L. E., D. Maneval, Jr., P. Panigrahi, A. Joseph, M. M. Levine, J. B. Kaper, J. G. Morris, Jr., and J. A. Johnson.** 1995. The capsule and O antigen in *Vibrio cholerae* O139 Bengal are associated with a genetic region not present in *Vibrio cholerae* O1. Infect. Immun. **63:**317–323.
13. **Desmarchelier, P. M., H. Momen, and C. A. Salles.** 1988. A zymovar analysis of *Vibrio cholerae* isolated in Australia. Trans. R. Soc. Trop. Med. Hyg. **82:**914–917.
14. **Dumontier, S., and P. Berche.** 1998. *Vibrio cholerae* O22 might be a putative source of exogenous DNA resulting in the emergence of the new strain of *Vibrio cholerae* O139. FEMS Microbiol. Lett. **164:**91–98.
15. **Evins, G. M., D. N. Cameron, J. G. Wells, K. D. Greene, T. Popovic, S. Giono-Cerezo, I. K. Wachsmuth, and R. V. Tauxe.** 1995. The emerging diversity of the electrophoretic types of *Vibrio cholerae* in the Western Hemisphere. J. Infect. Dis. **172:**173–179.
16. **Honma, Y., K. Yamamoto, and M. Iwanaga.** 1995. Aberrant gene for El Tor hemolysin from *Vibrio cholerae* non-O1, N037. FEMS Microbiol. Lett. **133:**151–154.
17. **Kaper, J. B., J. G. Morris, Jr., and M. M. Levine.** 1995. Cholera. Clin. Microbiol. Rev. **8:**48–86.
18. **Karaolis, D. K. R., R. Lan, and P. R. Reeves.** 1994. Molecular evolution of the seventh-pandemic clone of *Vibrio cholerae* and its relationship to other pandemic and epidemic *V. cholerae* isolates. J. Bacteriol. **176:**6199–6206.
19. **Karaolis, D. K. R., R. Lan, and P. R. Reeves.** 1995. The sixth and seventh cholera pandemics are due to independent clones separately derived from environmental, nontoxigenic, non-O1 *Vibrio cholerae*. J. Bacteriol. **177:**3191–3198.
20. **Karaolis, D. K. R., J. A. Johnson, C. C. Bailey, E. C. Boedeker, J. B. Kaper, and P. R. Reeves.** 1998. A *Vibrio cholerae* pathogenicity island associated with epidemic and pandemic strains. Proc. Natl. Acad. Sci. USA **95:**3134–3139.
21. **Kovach, M. E., M. D. Shaffer, and K. M. Peterson.** 1996. A putative integrase gene defines the distal end of a large cluster of ToxR-regulated colonization genes in *Vibrio cholerae*. Microbiology **142:**2165–2174.
22. **Li, J., K. Nelson, A. C. McWhorter-Murlin, T. S. Whittam, and R. K. Selander.** 1994. Recombinational basis of serovar diversity in *Salmonella enterica*. Proc. Natl. Acad. Sci. USA **91:**2552–2556.
23. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
24. **Maynard Smith, J., N. H. Smith, M. O'Rourke, and B. G. Spratt.** 1993. How clonal are bacteria? Proc. Natl. Acad. Sci. USA **90:**4384–4388.
25. **Mekalanos, J. J., E. J. Rubin, and M. K. Waldor.** 1997. Cholera: molecular basis for emergence and pathogenesis. FEMS Immunol. Med. Microbiol. **18:**241–248.
26. **Nelson, K., T. S. Whittam, and R. K. Selander.** 1991. Nucleotide polymorphism and evolution in the glyceraldehyde-3-phosphate dehydrogenase gene (*gapA*) in natural populations of *Salmonella* and *Escherichia coli*. Proc. Natl. Acad. Sci. USA **88:**6667–6671.
27. **Nelson, K., and R. K. Selander.** 1992. Evolutionary genetics of the proline permease gene (*putP*) and the control region of the proline utilization operon in populations of *Salmonella* and *Escherichia coli*. J. Bacteriol. **174:**6886–6895.
28. **Nelson, K., and R. K. Selander.** 1994. Intergeneric transfer and recombination of the 6-phosphogluconate dehydrogenase gene (*gnd*) in enteric bacteria. Proc. Natl. Acad. Sci. USA **91:**10227–10231.
29. **Olsvik, Ø., J. Wahlberg, B. Petterson, M. Uhlen, T. Popovic, I. K. Wachsmuth, and P. I. Fields.** 1993. Use of automated sequencing of polymerase chain reaction-generated amplicons to identify three types of cholera toxin subunit B in *Vibrio cholerae* O1 strains. J. Clin. Microbiol. **31:**22–25.
30. **Ramamurthy, T., S. Garb, R. Sharma, S. K. Bhattacharya, G. B. Nair, T. Shimada, T. Takeda, T. Karasawa, H. Kurazano, A. Pal, and Y. Takeda.** 1993. Emergence of novel strain of *Vibrio cholerae* with epidemic potential in

southern and eastern India. Lancet **341:**703–704.
31. **Rubin, E. J., M. K. Waldor, and J. J. Mekalanos.** 1998. Mobile genetic elements and the evolution of new epidemic strains of *Vibrio cholerae*, p. 147–161. *In* R. M. Krause (ed.), Emerging infections. Academic Press, San Diego, Calif.
32. **Sakazaki, R., and T. J. Donovan.** 1984. Serology and epidemiology of *Vibrio cholerae* and *Vibrio mimicus*. Methods Microbiol. **16:**271–289.
33. **Salles, C. A., and H. Momen.** 1991. Identification of *Vibrio cholerae* by enzyme electrophoresis. Trans. R. Soc. Trop. Med. Hyg. **85:**544–547.
34. **Selander, R. K., D. A. Caugant, H. Ochman, J. M. Musser, M. N. Gilmour, and T. S. Whittam.** 1986. Methods of multilocus enzyme electrophoresis for bacterial population genetics and systematics. Appl. Environ. Microbiol. **51:**873–884.
35. **Selander, R. K., D. A. Caugant, and T. S. Whittam.** 1987. Genetic structure and variation in natural populations of *Escherichia coli*, p. 1625–1648. *In* F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology, vol. 2. American Society for Microbiology, Washington, D.C.
36. **Selander, R. K., J. Li, and K. Nelson.** 1996. Evolutionary genetics of *Salmonella enterica*, p. 2691–2707. *In* F. C. Neidhardt, R. Curtiss III, J. L. Ingraham, E. C. C. Lin, K. B. Low, B. Magasanik, W. S. Reznikoff, M. Riley, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella*: cellular and molecular biology, 2nd ed., vol. 2. ASM Press, Washington, D.C.
37. **Sharma, C., A. Ghosh, A. Dalsgaard, A. Forslund, R. K. Ghosh, S. K. Bhattacharya, and G. B. Nair.** 1998. Molecular evidence that a distinct *Vibrio cholerae* O1 biotype El Tor strain in Calcutta may have spread to the African continent. J. Clin. Microbiol. **36:**843–844.
38. **Shimada, T., E. Arakawa, K. Itoh, T. Nakazato, T. Okitsu, S. Yamai, M. Kusum, G. B. Nair, and Y. Takeda.** 1994. Two strains of *Vibrio cholerae* non-O1 possessing somatic (O) antigen factors in common with *V. cholerae* serogroup O139 synonym "Bengal." Curr. Microbiol. **29:**331–333.
38a. **Shimada, T.** Personal communication.
39. **Shirai, H., M. Nishibuchi, T. Ramamurthy, S. K. Bhattacharya, S. C. Pal, and Y. Takeda.** 1991. Polymerase chain reaction for detection of the cholera enterotoxin operon of *Vibrio cholerae*. J. Clin. Microbiol. **29:**2517–2521.
40. **Stroeher, U. H., K. E. Jedani, B. K. Dredge, R. Morona, M. H. Brown, L. E. Karageorgos, M. J. Albert, and P. A. Manning.** 1995. Genetic rearrangements in the *rfb* regions of *Vibrio cholerae* O1 and O139. Proc. Natl. Acad. Sci. USA **92:**10374–10378.
41. **Stroeher, U. H., G. Parasivam, B. K. Dredge, and P. A. Manning.** 1997. Novel *Vibrio cholerae* O139 genes involved in lipopolysaccharide biosynthesis. J. Bacteriol. **179:**2740–2747.
42. **Vimont, S., S. Dumontier, V. Escuyer, and P. Berche.** 1997. The *rfaD* locus: a region of rearrangement in *Vibrio cholerae* O139. Gene **185:**43–47.
43. **Wachsmuth, I. K., G. M. Evins, P. I. Fields, Ø. Olsvik, T. Popovic, C. A. Bopp, J. G. Wells, C. Carrillo, and P. A. Blake.** 1993. The molecular epidemiology of cholera in Latin America. J. Infect. Dis. **167:**621–626.
44. **Wachsmuth, K., Ø. Olsvik, G. M. Evins, and T. Popovic.** 1994. Molecular epidemiology of cholera, p. 357–370. *In* I. K. Wachsmuth, P. A. Blake, and Ø. Olsvik (ed.), *Vibrio cholerae* and cholera: molecular to global perspectives. ASM Press, Washington, D.C.
45. **Waldor, M. K., R. Colwell, and J. J. Mekalanos.** 1994. The *Vibrio cholerae* O139 serogroup antigen includes an O-antigen capsule and lipopolysaccharide virulence determinants. Proc. Natl. Acad. Sci. USA **91:**11388–11392.
46. **Waldor, M. K., and J. J. Mekalanos.** 1996. Lysogenic conversion by a filamentous phage encoding cholera toxin. Science **272:**1910–1914.
47. **Wang, F.-S., T. S. Whittam, and R. K. Selander.** 1997. Evolutionary genetics of the isocitrate dehydrogenase gene (*icd*) in *Escherichia coli* and *Salmonella enterica*. J. Bacteriol. **179:**6551–6559.
48. **Whittam, T. S.** 1996. Genetic variation and evolutionary processes in natural populations of *Escherichia coli*, p. 2708–2720. *In* F. C. Neidhardt, R. Curtiss III, J. L. Ingraham, E. C. C. Lin, K. B. Low, B. Magasanik, W. S. Reznikoff, M. Riley, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella*: cellular and molecular biology, 2nd ed., vol. 2. ASM Press, Washington, D.C.
49. **Whittam, T. S., H. Ochman, and R. K. Selander.** 1983. Multilocus genetic structure in natural populations of *Escherichia coli*. Proc. Natl. Acad. Sci. USA **80:**1751–1755.
50. **Whittam, T. S., and S. E. Ake.** 1993. Genetic polymorphisms and recombination in natural populations of *Escherichia coli*, p. 223–245. *In* N. Takahata and A. G. Clark (ed.), Mechanisms of molecular evolution. Sinauer Associates, Inc., Sunderland, Mass.
51. **Whittam, T. S., E. A. McGraw, and S. D. Reid.** 1998. Pathogenic *Escherichia coli* O157:H7: a model for emerging infectious diseases, p. 163–183. *In* R. M. Krause (ed.), Emerging infections. Academic Press, San Diego, Calif.
52. **Yamai, S., T. Okitsu, T. Shimada, and Y. Katsube.** 1997. Distribution of serogroups of *Vibrio cholerae* non-O1 non-O139 with specific reference to their ability to produce cholera toxin, and addition of novel serogroups. Kansenshogaku Zasshi **71:**1037–1045. (In Japanese.)
53. **Yamamoto, K., G. R. F. Do Valle, M. Xu, T. Miwatani, and T. Honda.** 1995. Amino acids of the cholera toxin from *Vibrio cholerae* O37 strain S7 which differ from those of strain O1. Gene **163:**155–156.
54. **Zinnaka, Y., and C. C. Carpenter, Jr.** 1972. An enterotoxin produced by noncholera vibrios. Johns Hopkins Med. J. **131:**403–411.

AUTHOR'S CORRECTION

Genetic Diversity and Population Structure of *Vibrio cholerae*

PILAR BELTRÁN, GABRIELA DELGADO, ARMANDO NAVARRO, FRANCISCA TRUJILLO, ROBERT K. SELANDER, AND ALEJANDRO CRAVIOTO

Departamento de Salud Pública de la Facultad de Medicina, Universidad Nacional Autónoma de México, México, D.F., México, and Institute of Molecular Evolutionary Genetics, Pennsylvania State University, University Park, Pennsylvania 16802

Volume 37, no. 3, p. 581–590, 1999. In our article, we suggested that non-O1 strains of identical or closely similar electrophoretic type (ET) collected on different continents represent clonal lineages (Tables 3 and 5 in the original article). Reconsideration of our data, however, indicates that for many of these sets of strains an equally plausible if not more likely explanation of multilocus genotypic similarity is the independent recombinational assembly of common alleles. This is the case for ET 196 and ET 128 (Table 3), each of which has a genotype consisting of alleles that occur in high or moderate frequency in populations, as well as for many of the ETs listed in Table 5, which pertains to sets of strains differing at single loci.

However, for a number of ETs of non-O1 strains, a clonal relationship is probable because their genotypes include unique or rare alleles. Thus, for example, the genotype of ET 256, which was represented by a strain collected in India in 1979 and five strains recovered from patients in Mexico and Guatemala in the early 1990s, includes a unique allele of leucine aminopeptidase, an extremely rare indophenol oxidase allele, and uncommon alleles at four additional loci. Similarly, a clonal relationship is indicated for a pair of serotype O44 strains from India (1973) and Mexico (1991) that share a unique allele of nucleoside phosphorylase and a rare allele of phosphoglucomutase.

Multilocus enzyme electrophoresis is of limited use in identifying clonal lineages because, as we have noted, electromorphs cannot be equated with isoalleles and convergence in electrophoretic mobility of an enzyme is not infrequent. To determine clonal relationships among strains with a high degree of confidence, sequence data for multiple housekeeping genes will be required. This reinterpretation of the likely status of certain groups of non-O1 strains does not affect any other aspect of the work reported in our study.