Genomic profiles of clinical and environmental isolates of *Vibrio cholerae* O1 in cholera-endemic areas of Bangladesh

Young-Gun Zo*, Irma N. G. Rivera*†, Estelle Russek-Cohen‡, M. Sirajul Islam§, A. K. Siddique§, M. Yunus§, R. Bradley Sack¶, Anwar Huq*, and Rita R. Colwell*‖**

*Center of Marine Biotechnology, University of Maryland Biotechnology Institute, Baltimore, MD 21202; ‡Biometrics Program, Department of Animal and Avian Sciences, and ‖Department of Cell Biology and Molecular Genetics, University of Maryland, College Park, MD 20742; §International Centre for Diarrhoeal Disease Research, Dhaka 1000, Bangladesh; and ¶Department of International Health, School of Hygiene and Public Health, Johns Hopkins University, Baltimore, MD 21205

Contributed by Rita R. Colwell, July 17, 2002

Diversity, relatedness, and ecological interactions of toxigenic *Vibrio cholerae* O1 populations in two distinctive habitats, the human intestine and the aquatic environment, were analyzed. Twenty environmental isolates and 42 clinical isolates were selected for study by matching serotype, geographic location of isolation in Bangladesh, and season of isolation. Genetic profiling was done by enterobacterial repetitive intergenic consensus sequence–PCR, optimized for profiling by using the fully sequenced *V. cholerae* El Tor N16961 genome. Five significant clonal clusters of haplotypes were found from 57 electrophoretic types. Isolates from different areas or habitats intermingled in two of the five significant clusters. Frequencies of haplotypes differed significantly only between the environmental populations (exact test; $P < 0.05$). Analysis of molecular variance yielded a population genetic structure reflecting the differentiating effects of geographic area, habitat, and sampling time. Although a parameter confounding the latter differences explained 9% of the total molecular variance in the entire population ($P < 0.01$), the net effect of habitat and time could not be separated because of the small number of environmental isolates included in the study. Five subpopulations from a single area were determined, and from these we were able to estimate a relative differentiating effect of habitat, which was small compared with the effect of temporal change. In conclusion, the resulting population structure supports the hypothesis that spatial and temporal fluctuations in the composition of toxigenic *V. cholerae* populations in the aquatic environment can cause shifts in the dynamics of the disease.

Historically, seven pandemics of cholera have been recorded, the worst being the recent and continuing seventh pandemic that began in 1961, after a hiatus of 33 years (1). Cholera affects more than 75 countries and every continent (Communicable Disease Surveillance and Response, World Health Organization, www.who.org). Unlike earlier pandemics, the seventh has been persistent, spanning more than 40 years. The recurrent cholera epidemics in Africa and Latin America raise the question of endemicity of cholera versus geographical expansion, i.e., spread from the Bay of Bengal, where it is endemic, to other geographical areas otherwise believed to be free of endemic cholera, vs. *Vibrio cholerae* being autochthonous to estuarine, riverine, and coastal sites, as has been recently demonstrated (2–4).

A characteristic of cholera that has been studied in depth is the remarkable evolutionary capacity of its etiologic agent, *V. cholerae* (5). It appears that clonal shifts in the causative agent of cholera represent a fundamental characteristic of cholera dynamics (see *Background*, which is published as supporting information on the PNAS web site, www.pnas.org). The serotype shift from O1 to O139 was notable in its effect on the differential age structure of cholera patients (6). Recently, Karaolis *et al.* (7) found a hypervariable region in the DNA sequence of the virulence factor TcpA, suggesting that a structural shift in the pilin protein might be driven by selective pressure for the lysogenic phage VPIΦ to avoid intestinal immunity against *V. cholerae* (7). Such antigenic shifts can alter the interaction between the cholera pathogen and host immune system; hence the dynamics of cholera can be affected by a change in the susceptibility of the individual host (8). In addition, replacement of the predominant biotype from classical to El Tor and widespread occurrence in Africa and South America as well were regarded as a consequence of the more successful survival capacity of the El Tor biotype in the environment (3). Thus, clonal shifts in the composition of a toxigenic *V. cholerae* population are hypothesized to be significant in changes that occur in the progress of cholera epidemics, especially if a change occurs in the survival capacity of each *V. cholerae* clone combating intestine immunity or stresses in environmental habitats.

It is now widely accepted that the aquatic environment is the reservoir of toxigenic *V. cholerae* (3, 9–11) and also is a habitat that can support survival of toxigenic clones by providing protective or nutrient microhabitats, e.g., particulate organic detritus and biomass of aquatic fauna and flora (12, 13). Supported by the finding that climate forcing, such as El Niño, is a modulator of interannual cholera dynamics in cholera endemic areas (14–16), and that environmental factors, notably water temperature, are strongly correlated with the seasonal dynamics of cholera (14, 17, 18), it is hypothesized that the dynamics of *V. cholerae* populations in the aquatic environment contribute significantly to variation in cholera epidemics. However, comprehensive studies of the underlying mechanisms of how aquatic environmental factors modulate the dynamics of cholera epidemics have yet to be done. In this study, we chose to investigate compositional change in *V. cholerae* populations in the aquatic environment as a part of such a mechanism and used molecular identification methods to the microbial ecology of *V. cholerae*.

It is now known that *V. cholerae* O1 or O139 isolates from cholera patients form tight clusters within the species (19–21). Furthermore, although nontoxigenic clones from the aquatic environment contribute to most of the clonal diversity, the species demonstrates a panmictic population genetic structure, with genetic information actively exchanged between clones (19, 20, 22). Clusters of cholera pathogenicity genes, which are, in fact, the lysogenic bacteriophages CTXΦ (23) and VPIΦ (24), can be transferred to diverse *V. cholerae* clones in the aquatic environment, making the aquatic environment a source of novel

MICROBIOLOGY

Abbreviations: ERIC, enterobacterial repetitive intergenic consensus sequence; AMOVA, analysis of molecular variance.

†Present address: Department of Microbiology, Biomedical Science Institute, São Paulo University, São Paulo, Brazil.

**To whom reprint requests should be addressed. E-mail: colwell@umbi.umd.edu.

pathogenic *V. cholerae* outbreak strains (9, 11). A survey carried out in West Bengal, India (25), revealed a wide distribution of cholera toxin genes among diverse non-O1 *V. cholerae* strains in surface waters. In addition, Faruque *et al.* (26) demonstrated that the transduction process in the natural environment can result in nontoxigenic environmental strains becoming cholera toxin gene-carrying strains.

On the basis of this information, we hypothesize that quantitative analyses of the diversity, relatedness, and interactions of toxigenic *V. cholerae* within and between two distinctive habitats, the aquatic environment and the human intestine, can be used to elucidate the contribution of the aquatic environment to fluctuations in cholera epidemics in a cholera endemic geographical location. To substantiate this hypothesis, we synchronized surveys made in both clinical and environmental habitats for toxigenic *V. cholerae* within a defined spatial and temporal frame. Results of previous studies (19–21, 27–31), in which toxigenic environmental *V. cholerae* isolates were compared with clinical isolates, have yielded information limited by lack of a temporal and spatial frame for comparison, as well as excessive broadness of the frame (e.g., the global population structure of phylogenetic groups) and/or lack of quantitative measurements. In the study reported here, we carried out biweekly clinical surveys synchronized with environmental surveys in three cholera-endemic areas of Bangladesh over a 2.5-yr time frame. *V. cholerae* O1 El Tor was successfully isolated from the aquatic environment in each area. The genome of all isolates was profiled by using enterobacterial repetitive intergenic consensus sequence (ERIC)–PCR analysis, and the data were subjected to population genetic analyses.

Materials and Methods

Survey and Isolation of Toxigenic *V. cholerae* Strains. During the period from July 1997 to December 1999, stools of 20% of all diarrhea patients administered in the clinics of Bakergonj (22.5°N 90.2°E), Chhattak (24.9°N 91.6°E), and Chaughacha (23.2°N 89.0°E), Bangladesh, were screened at 15-d intervals for *V. cholerae*. In parallel with the clinical work, we also screened for *V. cholerae* in surface waters of the ponds and rivers used by the patients in each area for various purposes, including drinking water. Enrichment plating methods, using alkaline peptone water and thiosulfate–citrate–bile salts–sucrose agar, as well as taurocholate–tellurite–gelatin agar, were used for isolation of *V. cholerae*. Following standard reference protocols (32), we used a set of biochemical and serological (coagglutination) tests to confirm *V. cholerae* colonies to species and serogroup. A further description of identification procedures used is published as supporting *Materials and Methods* on the PNAS web site.

Genomic Profiling. ERIC-PCR fingerprinting, as described by Rivera *et al.* (33), was used, with modification to enhance resolution (see supporting *Materials and Methods*).

Population Genetic Analyses. To examine the differential distribution of haplotypes between clinical and environmental isolates from the three areas, we separated isolates into six populations by three areas and two habitat types. The phylogenetic relationship among genotypes was calculated as the Euclidean distance ($d = \sqrt{1 - S_{SM}}$, where S_{SM} is the simple match coefficient) from the binary coded data of ERIC-PCR band patterns (34). From the pairwise distance matrix, we constructed unrooted trees, using the neighbor-joining (NJ) procedure in the PHYLIP version 3.5 package (35). The techniques for evaluation of the resulting tree topology are published as supporting *Materials and Methods*. Examination of the population genetic structure, by analysis of molecular variance (AMOVA) (36) and exact test, was carried out using procedures in the software ARLEQUIN (37). The Euclidean distance matrix was used to estimate variances attributable to differences of geographic area and habitat (i.e., environmental vs. clinical).

Results

Characteristics of Isolates. Throughout the study period, four surveys produced toxigenic *V. cholerae* strains from diverse environmental samples (see Table 2, which is published as supporting information on the PNAS web site), and the clinical survey yielded *V. cholerae* O1 isolates throughout at least one full season in a given year. All environmental *V. cholerae* O1 isolates were El Tor biotype and belonged to the Inaba or Ogawa serovar. To compare the environmental and clinical isolates, we selected clinical isolates by matching biotype and serotype (see Table 3, which is published as supporting information on the PNAS web site). For the Inaba serotype populations from Chhattak, Bangladesh, all clinical isolates at four times of isolation were selected. For other areas, we randomly selected one isolate at times when matching serotypes were present. All were positive by PCR for the pathogenicity island genes *ctxA* and *tcpI*.

ERIC-PCR Patterns. ERIC-PCR produced more than 40 bands for each isolate (see Fig. 2, which is published as supporting information on the PNAS web site). From the bands, those ranging from 100 to 588 bp were selected for genomic profiling. The size range selected avoided the occasional smearing of strong intensity bands and gave consistent resolution, from which duplicated PCR and electrophoresis produced 99% band matches.

When formation of PCR products of 100–588 bp was simulated from primer-binding sites along the two chromosomal DNA sequences of *V. cholerae* O1 El Tor N16961 (38), 105 fragments with T_m values higher than 0°C were generated (see Fig. 3, which is published as supporting information on the PNAS web site). The maximum T_m of predicted PCR products was 30.4°C, and most of the products had a T_m below 12°C. Considering the temperature (65°C) used in the elongation step of the ERIC-PCR, occurrence of primer-binding sites in the genome of *V. cholerae* to produce bands of 100–588 bp was random rather than specifically organized. When we tested the hypothesis of even distribution of primer-binding sites and PCR product loci on the two chromosomes of *V. cholerae* N16961, the hypothesis was not rejected [Rao's spacing test for uniformity (39); $P > 0.05$ for fragments with $T_m > 12$°C]. Therefore, our fingerprints of 100- to 588-bp range were consistent with random occurrence of primer-binding sites evenly distributed throughout the entire region of the two chromosomes. An additional aspect to be noted from the distribution of primer-binding sites was that the affinity of the primer binding was significantly reduced, e.g., $T_m < 0$°C, by a single point mutation in most of the loci, especially when it occurred at the 3′ end of the primer. Therefore, the sensitivity of PCR product formation can be considered to be at the level of a single point mutation.

One of our goals in the simulation of PCR by calculating ERIC primer-binding affinity was to detect differences in the chromosomal regions of genes coding the Inaba or Ogawa serovar determinant. However, we did not find primer-binding sites that produced the fragments in the vicinity of the *wbeT* gene (40). The closest was a locus 22 kb upstream and 7 kb downstream of the structural and regulatory nucleotide sequences of the *wbeT* gene (see Fig. 2). Thus, ERIC-PCR could not discriminate strains by Inaba vs. Ogawa serovar determinants.

From the genomic fingerprints produced by ERIC-PCR for the 63 *V. cholerae* O1 isolates, 71 band loci producing 57 haplotypes were identified. When pairwise mismatches among the isolates were counted, a unimodal pattern was observed among the clinical isolates and in the clinical–environmental comparison; however, a unimodal distribution of mismatches

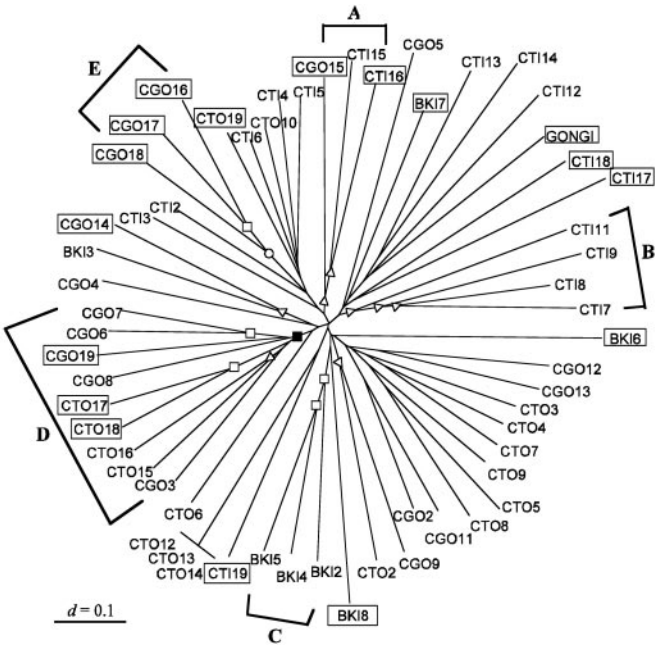

Fig. 1. Unrooted neighbor-joining tree based on Euclidean distances (*d*) among ERIC-PCR band patterns of *V. cholerae* O1 isolates. Isolates from environmental sources are enclosed in rectangles. When a branching point is significant by having bootstrap support value from 0.5 to 0.9, it is marked as a open circle. When a branching point is significant by appearing consistently in the trees by unweighted pair group method with arithmetic averages (UPGMA) or maximum parsimony (MP), it is marked as triangle. When a branching point is significant by both methods, it is marked as a open square. When a branching point is significant both by bootstrap support >0.9 and by consensus occurrence, it is marked as a solid square. Labeling of the strains follows that of Tables 2 and 3, except for GONGI, a *V. cholerae* O1 El Tor Inaba strain isolated from coastal water off Chittagong, Bangladesh (22.1°N 91.8°E).

was not characteristic of the environmental isolates (see Table 4, which is published as supporting information on the PNAS web site). Although the modal number of mismatches was 25, another mismatch peak of 17 was observed. Unlike the mode at 25 mismatches, the secondary peak at 17 mismatches was mostly (75 vs. 50%) from pairs of environmental strains isolated from different geographical areas.

Cluster Analyses. The phylogenetic relationship of the haplotypes is shown in Fig. 1. Clustering of isolates into different lineages occurred, and the bootstrapping support of the clusters resulted in one distinct cluster (Cluster D in Fig. 1) with >0.99 support value. Although most of the lineages in Fig. 1 were not highly supported by bootstrapping, the appearance of other clusters was consistent with that of other tree-building methods, such as unweighted pair group method with arithmetic averages (UPGMA) and maximum parsimony (MP). Although Clusters B, C, and E showed clustering of isolates of the same serotype from the same habitat of a geographic area, haplotypes in Clusters A and D were intermingled, regardless of source of isolation or whether Inaba vs. Ogawa serotype.

Pairwise Comparison of Clinical and Environmental Populations. Haplotypic compositions of the six populations, differentiated by area of isolation and habitat of isolation, were compared by the exact test. Pairwise comparisons between the populations from the two habitats for each area did not show significant difference (exact test; $P = 0.06$ in Bakergonj, $P = 0.51$ in Chhattak, and $P = 0.98$ in Chaughacha). Compositional differences among the clinical populations from the three areas were also not significant (exact test; $P > 0.35$). However, the haplotype frequency of the environmental population in Bakergonj was significantly different from those of the environmental populations from the other two areas (exact test; $P < 0.05$).

AMOVA. As shown in Table 1, we decomposed the total variance of the PCR bands by difference in geographic area and then nested by the two types of habitats within each geographic area (Structure I). Although isolates from the Chhattak area could be divided into Inaba and Ogawa serogroups, we disregarded serotype differences because the exact test on haplotype distribution between serotypes did not show a significant difference ($P = 0.22$). The observation that isolates of different serotypes

Table 1. Estimates from a hierarchical AMOVA of all populations of *V. cholerae* O1, based on different areas of isolation (Bakergonj, Chhattak, or Chaughacha) and habitats (clinical versus environmental) of isolation

		Observed partition			
Structure*	Variance component	Variance	% total	*P*	*F*
I. Habitats under areas	Among areas	0.27	2	0.08	$F_{CT} = 0.024$
	Among habitats in an area	1.02	9	<0.01	$F_{SC} = 0.093$
	Within populations	9.96	89	<0.01	$F_{ST} = 0.114$
II. Time under habitats in Chhattak area	Among habitats	0.08	1	0.53	$F_{CT} = 0.007$
	Among temporally separated subpopulations	1.20	10	<0.01	$F_{SC} = 0.130$
	Within subpopulations isolated on the same day	10.53	89	<0.01	$F_{ST} = 0.108$
III. Habitat–time complex in Chaughacha	Among populations separated by habitat type and time (3 mo)	0.95	9	0.12	$F_{ST} = 0.092$
	Within subpopulations isolated on the same day	9.36	91		

*Structure I assumed *V. cholerae* populations are separated among different areas, and the clinical population is separated from the environmental population in each area. In Structure II, the three clinical *V. cholerae* populations from Chhattak (CTIC9711, CTIC9712, and CTOC9810) were compared to the two environmental populations from the same area (CTOE9909 and CTIE9911). Structure III was AMOVA application to CGOC9903 and CGOE9906, which were separated by the difference of habitats and 3 months in the time of isolation from Chaughacha.

intermingled in clustering (Fig. 1) also justified merging the Chhattak populations.

Quantitatively, the population of *V. cholerae* O1 in this study had an F_{ST} value of 0.11 when divided into the six area–habitat subpopulations (Table 1). According to the qualitative classification by Wright (41), an F_{ST} value within the range of 0.05–0.15 indicates moderate differentiation, within 0.15–0.25 great differentiation, and greater than 0.25 indicates very great differentiation. Therefore, the level of divergence among *V. cholerae* O1 populations distributed in the three areas in Bangladesh over the 2.5-yr time period was a typical moderate differentiation. Approximately 90% of the variance was explained by isolate-to-isolate variance within a population. Habitat difference accounted for the rest of the variance, and the estimate for the effect of geographic area was not significant. However, we recognized that the comparison of subpopulations by habitat type involves confounding its effect with that of temporal variation in subpopulations, because the environmental strains and clinical strains were not isolated at the same time. Therefore, genetic divergence among the six area–habitat subpopulations in the AMOVA result (Structure I, Table 1) was regarded to be the effect of habitat–time complex.

Because of this habitat–time confounding, it was not possible to examine the effect of habitat and time separately. Nevertheless, we estimated the relative magnitude of habitat effect to temporal change in populations by examining the population structure of each geographical area. In Structure II of Table 1, we analyzed the genetic variance in a part of the Chhattak populations, finding three clinical populations separated by 1 mo to 1 yr and two environmental populations separated by 2 mo. The level of fixation within the subpopulations (F_{ST} in Structure II) was comparable to that of whole populations (F_{ST} in Structure I), indicating the subpopulations used in Structure II represents well the whole population in Structure I. The hierarchs of the populations in Structure II comprised by the two factors: (*i*) a habitat–time complex in which difference of habitat types and >11-mo difference in time of isolation were confounded, and (*ii*) a pure time factor, which is nested under the habitat–time complex and consisted of 1- to 11-mo difference in the time of isolation among clinical strains and 2-mo difference in the time of isolation among environmental strains. The AMOVA result shows that the second factor alone can explain about 10% of the total variance and habitat–time complex at the top level of the hierarchy is not significant ($P = 0.53$) as a factor causing differentiation among the subpopulations. In the case of the Chaughacha strains, the subpopulation of environmental strains (CGOE9906; see Table 2) were compared with a clinical subpopulation (CGOC9906; see Table 3) isolated with a finite time difference of 3 mo, and the analysis in Structure III (Table 1) shows the two subpopulations were not significantly diverged.

Discussion

Clonality and Diversity of *V. cholerae* O1 El Tor. The species level population structure of *V. cholerae* and the relatedness of toxigenic *V. cholerae* strains to nontoxigenic *V. cholerae* have been studied extensively by multilocus enzyme electrophoresis (19, 20) and by DNA sequence analyses of a single housekeeping gene, the *recA* gene (21), or multiple genes (42). In a study done by Beltran *et al.* (19), 244 strains of *V. cholerae* (17% of which were environmental) isolated from Mexico and Guatemala during 1991 and 1995 were analyzed, along with 143 serogroup reference strains. Most of the electrophoretic types comprised non-O1/non-O139 strains, whereas the O1 and O139 strains formed a tight cluster. Farfán *et al.* (20) analyzed a collection of ≈100 *V. cholerae* strains isolated from the environment in diverse geographic locations during an unspecified time period. Their results showed that the entire *V. cholerae* population was more diverse than the set analyzed by Beltran *et al.* (19), as were the O1 and O139 subpopulations. Stine *et al.* (21) reported nucleotide sequences for the single housekeeping gene, *recA*. It was calculated that toxigenic O1/O139 El Tor strains form a phylogenic lineage that distinguishes itself from other populations of the species *V. cholerae*.

Besides clonality of the O1 El Tor strains, results of previous studies also revealed significant diversity within the lineage of toxigenic O1 El Tor strains. Seventeen housekeeping genes of the *V. cholerae* O1 El Tor population examined by Beltran *et al.* (19) showed genetic distances of up to about 10% of the maximum genetic distance calculated from the total *V. cholerae* population. The multilocus enzyme electrophoresis data of Farfán *et al.* (20) also demonstrated significant diversity among O1 El Tor strains, with a mean genetic diversity per locus (H) as high as 0.4, whereas the total *V. cholerae* population showed an H value of 0.5. The lack of congruence between the phylogenetic relationships of pathogenic clones deduced from the four housekeeping genes studied by Byun *et al.* (42), and another housekeeping gene *asd* (43) offers molecular proof for potentially significant microevolution among *V. cholerae* O1 strains, especially via a high level of gene recombination. Recently, a powerful microarray assay of a set of seventh pandemic strains revealed that the clonality within *V. cholerae* El Tor strains is indeed mediated by a set of clone-specific genetic elements (44). Lan and Reeves (2) also report further diversity and clonality within the seventh pandemic *V. cholerae* O1 El Tor cluster, which are strong enough to be related to temporal and geographical divergence of the clones in the cluster.

In this study, we analyzed diversity among populations of *V. cholerae* O1 El Tor isolates within a specifically defined temporal and geographical frame. Although the frame was much narrower than those of previous studies, we observed significant diversity among the isolates (Fig. 1). When significance of clonality among the diverse *V. cholerae* O1 El Tor electrophoretic types was tested, three clusters (Clusters C–E in Fig. 1) were significant by the bootstrapping procedure (35). We interpret this low support of clusters as the effect of violation of the independence assumption in bootstrapping. Randomly amplified polymorphic DNA fingerprinting techniques, like ERIC-PCR, may have a strong tendency to violate the assumption in several ways, e.g., formation of a single band from several primer-binding sites and participation of one primer-binding site in producing several bands of different sizes. An alternative to bootstrapping, consensus of tree topology by different tree construction methods can be used to examine the significance of a cluster (45, 46). When we examined the matrix of pairwise genetic distances by three different methods [unweighted pair group method with arithmetic averages (UPGMA), neighbor-joining (NJ), and maximum parsimony (MP)], two additional clusters appeared as consensus (Clusters A and B in Fig. 1), suggesting additional clonality among the *V. cholerae* O1 El Tor isolates.

Within the significant clusters, an intermingling of Inaba strains with Ogawa strains (Cluster A) was observed. The Inaba serotype can be generated by a recessive mutation of the *wbeT* gene of the Ogawa serotype (40, 47). Therefore, occurrence of two different serovars in an identical phylogenic lineage is highly plausible. In Clusters A and D, intermingling of strains from two different geographical areas, Chhattak and Chaughacha, was observed, implying weakness in spatial separation between the two areas. In addition, environmental isolates and clinical isolates from Chhattak and Chaughacha belonged to the same lineage (Clusters A and D). Therefore, toxigenic *V. cholerae* O1 El Tor strains in cholera-endemic areas in Bangladesh are concluded to evolve to form several clonal lineages. At least five lineages observed in this study were significant. Furthermore, environmental and clinical strains were found to belong to the same lineage. In addition, the unimodal distribution of mismatches among the clinical isolates and clinical-to-environmental isolates (see Table 4) that was observed in this study implies that panmictic gene exchange between lineages may weaken such clonality (48).

Effect of Geographic Isolation. Application of molecular epidemiological methods revealed that toxigenic *V. cholerae* O1 strains from the seventh pandemic or from localized outbreaks could be differentiated according to geographical location of the outbreak (2, 31). Ribotyping, plasmid carriage, and phage typing data suggested continent-scale differentiation of *V. cholerae* strains. Geographical isolation is considered to be one of the factors in establishing different lineages of *V. cholerae*.

In this study, we obtained genomic profiles of *V. cholerae* O1 El Tor populations from two distinctive habitats in three separate geographic areas. Thus, we were able to examine the genetic structure of toxigenic *V. cholerae* populations differentiated by spatial isolation on a regional scale, as well as by habitat. Of the population structures shown in Table 1, the one that accounts for the differentiation of *V. cholerae* O1 populations by geographic isolation, and then by habitat, was valid (Structure I in Table 1). Although the estimate of the significance of the contribution of the difference in geographical areas was marginal ($P = 0.08$), heterogeneity among the toxigenic *V. cholerae* populations across geographical areas was also supported by other aspects of the data. As assessed by exact test of the distribution of haplotypes, the composition of the environmental populations was significantly different, whereas those among the clinical populations were not. The lack of clarity of unimodal mismatch distribution (Table 4) among environmental isolates also revealed heterogeneity in the distribution of genetic exchange among environmental strains. Therefore, it is concluded that environmental populations of toxigenic *V. cholerae* O1 in Bangladesh may be separated by geographical isolation, whereas the composition of clinical populations is relatively homogeneous.

Effect of Habitat and Time. Environmental populations of *V. cholerae* O1 in each area were isolated on only four occasions throughout the survey period (see Table 2), and the time of isolation of the environmental strains was when a cholera outbreak had not occurred, with respect to the clinical survey, imposing a temporal uncoupling of more than one month in the time scale. The nature of the survey not only placed a limit in the resolution of our analysis on the effect of geographic area or habitat difference in the total variation of population composition but also caused a confounding of the effect of habitat difference and of temporal dynamics in the composition of the clinical and environmental *V. cholerae* populations. Therefore, the effect of the habitat–time complex, which explained about 9% of the total variance as shown Structure I in Table 1, indicates that the effect of habitat actually is less than 9%.

The clinical and environmental populations from Chhattak, however, provided an opportunity to estimate the relative magnitude of the effect of habitat in the molecular variance analyses (Table 1). The difference of time, in the scale of up to 11 mo, had a significant variance (10% of the variance in the total five subpopulations from Chhattak) in the clinical populations. It was concluded that the effect of habitat–time complex observed for the entire population (Structure I, Table 1) was mostly the effect of the temporal difference between the two habitat samples and that the habitat difference does not impose a measurable magnitude of difference in the genetic structure of the populations in the two habitats.

It should be noted that the existence of a time effect implies that changes in the population composition of *V. cholerae* O1 in a cholera-endemic area occur over time. In the recent amplified fragment length polymorphism analyses by Lan and Reeves (2), clonal shifts in the seventh pandemic strains over a time scale of a decade was well demonstrated. In this study, the time scale for a measurable change was 1 mo or less among the clinical cholera cases. This time scale is supported by our finding of a significant fixation index among the subpopulation of clinical isolates (Structure II, Table 1), as well as from the dramatic change in the relative prevalence of O1 and O139 serogroups in cholera cases of the study areas within the time scale of 15 d (data not shown). A rapid microevolution, or more likely a selection of clones from a variety of clones, in the environment, corresponding to fluctuations in the environment or to immunity in the host populations, may cause such dynamic change. Because we did not find any significant difference in the genomic profiles of the *V. cholerae* O1 El Tor strains in the two habitats, we conclude seasonal changes in the aquatic environment cause temporal change by placing a selective pressure on environmental clones of *V. cholerae*.

Several examples of microevolution are known for *V. cholerae.* Horizontal gene transfer of the O antigen genes has been shown to occur in the generation of O139 and O37 serogroups (40, 49), and variability in the *tcpA* gene is believed to be caused by homologous recombination (7). The unimodal distribution of mismatches found in this study and in an earlier study by Farfán *et al.* (20) suggests that panmictic recombination drives rapid microevolution among *V. cholerae* strains and that gene flow is not restricted or bottlenecked between environmental and clinical habitats. In summary, environmental *V. cholerae* O1 El Tor populations demonstrate significant geographical isolation, but barriers between the clinical habitat and aquatic environment are not significant. In addition to spatial variance, temporal variance is a significant factor, explaining total genomic variances among toxigenic *V. cholerae* populations.

We conclude that the composition of environmental populations of toxigenic *V. cholerae* is identical to that of *V. cholerae* populations causing endemic cholera. Although the composition of *V. cholerae* O1 El Tor populations causing endemic cholera in Bangladesh undergoes dynamic change, *V. cholerae* populations in the two distinctive habitats achieve a dynamic equilibrium. Rapid transfer between habitats or panmictic geneflow by active intermingling of clinical and environmental *V. cholerae* strains is suggested to be the mechanism of the dynamic equilibrium between the two distinctive habitats. Finally, we also conclude that the aquatic environment is the cholera reservoir and is associated with shifts in the dynamics of the disease by causing spatial and temporal fluctuations in the composition of toxigenic *V. cholerae* inhabitants. Any change in the composition of *V. cholerae* populations in the aquatic environment that may be driven by seasonal fluctuation in the environment or by introduction of new strains through microevolution or being imported from other systems can cause coupled changes in composition and behavior of the clinical populations, leading to a shift in the dynamics of cholera epidemics. Because most cholera endemic areas occur in developing countries where there is poor sanitary control of drinking water, our findings for Bangladesh are applicable to other cholera endemic areas of the world.

This research was supported by National Institutes of Health Grant 1R01A139129-01.

1. Barua, D. (1992) in *Cholera*, eds. Barua, D. & Greenough, W. B. (Plenum Medical Book, New York), pp. 1–35.
2. Lan, R. T. & Reeves, P. R. (2002) *J. Clin. Microbiol.* **40,** 172–181.
3. Kaper, J. B., Morris, J. G., Jr., & Levine, M. M. (1995) *Clin. Microbiol. Rev.* **8,** 48–86.
4. Colombo, M. M., Mastrandrea, S., Leite, F., Santona, A., Uzzau, S., Rappelli, P., Pisano, M., Rubino, S. & Cappuccinelli, P. (1997) *FEMS Immunol. Med. Microbiol.* **19,** 33–45.
5. Mooi, F. R. & Bik, E. M. (1997) *Trends Microbiol.* **5,** 161–165.
6. Cholera Working Group, International Centre For Diarrhoeal Disease Research, Bangladesh (1993) *Lancet* **342,** 387–390.
7. Karaolis, D. K., Lan, R., Kaper, J. B. & Reeves, P. R. (2001) *Infect. Immun.* **69,** 1947–1952.
8. Codeço, C. T. (2001) *BMC Infect. Dis.* **1,** 1.
9. Colwell, R. R., Kaper, J. & Joseph, S. W. (1977) *Science* **198,** 394–396.
10. Colwell, R. R. & Huq, A. (1994) *Ann. N.Y. Acad. Sci.* **740,** 44–54.

11. Faruque, S. M., Albert, M. J. & Mekalanos, J. J. (1998) *Microbiol. Mol. Biol. Rev.* **62,** 1301–1314.
12. Colwell, R. R. & Spira, W. M. (1992) in *Cholera*, eds. Barua, D. & Greenough, W. B. (Plenum Medical Book, New York), pp. 107–127.
13. Islam, M. S., Drasar, B. S. & Sack, R. B. (1994) *J. Diarrhoeal Dis. Res.* **12,** 87–96.
14. Colwell, R. R. (1996) *Science* **274,** 2025–2031.
15. Pascual, M., Rodo, X., Ellner, S. P., Colwell, R. & Bouma, M. J. (2000) *Science* **289,** 1766–1769.
16. Speelmon, E. C., Checkley, W., Gilman, R. H., Patz, J., Calderon, M. & Manga, S. (2000) *J. Am. Med. Assoc*. **283,** 3072–3074.
17. Nair, G. B., Sarkar, B. L., De, S. P., Chakrabarti, M. K., Bhadra, R. K. & Pal, S. C. (1988) *Microbiol. Ecol.* **15,** 203–215.
18. Lobitz, B., Beck, L., Huq, A., Wood, B., Fuchs, G., Faruque, A. S. & Colwell, R. (2000) *Proc. Natl. Acad. Sci. USA* **97,** 1438–1443.
19. Beltran, P., Delgado, G., Navarro, A., Trujillo, F., Selander, R. K. & Cravioto, A. (1999) *J. Clin. Microbiol.* **37,** 581–590.
20. Farfán, M., Minaña, D., Fusté, M. C. & Lorén, J. G. (2000) *Microbiology* **146,** 2613–2626.
21. Stine, O. C., Sozhamannan, S., Gou, Q., Zheng, S., Morris, J. G., Jr., & Johnson, J. A. (2000) *Infect. Immun.* **68,** 7180–7185.
22. Faruque, S. M. & Nair, G. B. (2002) *Microbiol. Immunol.* **46,** 59–66.
23. Waldor, M. K. & Mekalanos, J. J. (1996) *Science* **272,** 1910–1914.
24. Karaolis, D. K., Somara, S., Maneval, D. R., Jr., Johnson, J. A. & Kaper, J. B. (1999) *Nature (London)* **399,** 375–379.
25. Chakraborty, S., Mukhopadhyay, A. K., Bhadra, R. K., Ghosh, A. N., Mitra, R., Shimada, T., Yamasaki, S., Faruque, S. M., Takeda, Y., Colwell, R. R., *et al*. (2000) *Appl. Environ. Microbiol.* **66,** 4022–4028.
26. Faruque, S. M., Asadulghani, Rahman, M. M., Waldor, M. K. & Sack, D. A. (2000) *Infect. Immun.* **68,** 4795–4801.
27. Bik, E. M., Gouw, R. D. & Mooi, F. R. (1996) *J. Clin. Microbiol.* **34,** 1453–1461.
28. Huq, A., Parveen, S., Qadri, F., Sack, D. A. & Colwell, R. R. (1993) *J. Trop. Med. Hygiene* **96,** 86–92.
29. Jiang, S. C., Matte, M., Matte, G., Huq, A. & Colwell, R. R. (2000) *Appl. Environ. Microbiol.* **66,** 148–153.
30. Singh, D. V., Matte, M. H., Matte, G. R., Jiang, S., Sabeena, F., Shukla, B. N., Sanyal, S. C., Huq, A. & Colwell, R. R. (2001) *Appl. Environ. Microbiol.* **67,** 910–921.
31. Wachsmuth, I. K., Olsvik, O., Evins, G. M. & Popovic, T. (1994) in *Vibrio cholerae and Cholera: Molecular to Global Perspectives*, eds. Wachsmuth, I. K., Blake, P. A. & Olsvik, O. (Am. Soc. Microbiol., Washington, DC), pp. 357–370.
32. Tison, D. L. (1999) in *Manual of Clinical Microbiology*, eds. Murray, P. R., Baron, E. J., Pfaller, M. A., Tenover, F. C. & Yolken, R. H. (Am. Soc. Microbiol., Washington, DC), pp. 497–504.
33. Rivera, I. G., Chowdhury, M. A., Huq, A., Jacobs, D., Martins, M. T. & Colwell, R. R. (1995) *Appl. Environ. Microbiol.* **61,** 2898–2904.
34. Sneath, P. H. A. & Sokal, R. R. (1973) *Numerical Taxonomy* (Freeman, San Francisco).
35. Felsenstein, J. (1993) PHYLIP: Phylogenetic Inference Package (Dept. of Genetics, University of Washington, Seattle), Version 3.5c.
36. Excoffier, L., Smouse, P. E. & Quattro, J. M. (1992) *Genetics* **131,** 479–491.
37. Schneider, S., Soessli, D. & Excoffier, L. (2000) ARLEQUIN: A Software for Population Genetics Data Analysis (Genetics and Biometry Laboratory, Dept. of Anthropology, University of Geneva, Geneva), Version 2.0.
38. Heidelberg, J. F., Eisen, J. A., Nelson, W. C., Clayton, R. A., Gwinn, M. L., Dodson, R. J., Haft, D. H., Hickey, E. K., Peterson, J. D., Umayam, L., *et al.* (2000) *Nature (London)* **406,** 477–483.
39. Jammalamadaka, S. R. & SenGupta, A. (2001) *Topics in Circular Statistics* (World Scientific, Singapore).
40. Stroeher, U. H., Jedani, K. E. & Manning, P. A. (1998) *Gene* **223,** 269–282.
41. Wright, S. (1978) *Variability Within and Among Natural Populations* (Univ. Chicago Press, Chicago).
42. Byun, R., Elbourne, L. D., Lan, R. & Reeves, P. R. (1999) *Infect. Immun.* **67,** 1116–1124.
43. Karaolis, D. K., Lan, R. & Reeves, P. R. (1995) *J. Bacteriol.* **177,** 3191–3198.
44. Dziejman, M., Balon, E., Boyd, D., Fraser, C. M., Heidelberg, J. F. & Mekalanos, J. J. (2002) *Proc. Natl. Acad. Sci. USA* **99,** 1556–1561.
45. Kim, J. (1993) *Syst. Biol.* **42,** 331–340.
46. Hilali, F., Ruimy, R., Saulnier, P., Barnabé, C., Lebouguénec, C., Tibayrenc, M. & Andremont, A. (2000) *Infect. Immun.* **68,** 3983–3989.
47. Sack, R. B. & Miller, C. E. (1969) *J. Bacteriol.* **99,** 688–695.
48. Whittam, T. S. (1995) in *Population Genetics of Bacteria*, eds. Baumberg, S., Young, J. P., Wellington, E. M. H. & Saunders, J. R. (Cambridge Univ. Press, Cambridge, U.K.), pp. 217–245.
49. Sozhamannan, S., Deng, Y. K., Li, M., Sulakvelidze, A., Kaper, J. B., Johnson, J. A., Nair, G. B. & Morris, J. G., Jr. (1999) *Infect. Immun.* **67,** 5033–5040.