

## **CHAPTER 4**

# **POPULATION DYNAMICS OF *V. CHOLERAE* IN THE VAAL BARRAGE CATCHMENT**

### **4.1 INTRODUCTION**

Cholera remains to be one of the most severe diseases of our time, with seven pandemics being recorded since 1817. The Seventh pandemic is still raging since its start in Asia in 1961, and it has led to major epidemics in developing African countries after its introduction into Africa in 1971 (Lan and Reeves, 2002). South Africa has not escaped and had two major epidemics since 1980 (Department of Health, South Africa 2004). Due to the large amount of people infected during epidemics the disease has a huge socio-economic impact on regions; this along with the potential for rapid spread calls for effective management strategies to alleviate risk factors.

*Vibrio cholerae* is indigenous to the aquatic environment, preferring salt water, but can also be readily isolated from fresh water bodies (Halpern *et al.*, 2004). These aquatic systems may serve as possible reservoirs for not only enterotoxigenic *V. cholerae*, which are capable of epidemic spread, but also for non-enterotoxigenic environmental strains. Non-enterotoxigenic strains should also be regarded as clinically important due a variety of secondary pathogenetic mechanisms that enable them to cause disease in humans, though not to the extent of enterotoxin producing strains (Chakraborty *et al.*, 2000). Environmental *V. cholerae* strains should also be regarded as a potential health risk in the

light of the discovery of the lysogenic filamentous phage CTX $\phi$ , which has been shown to be able to donate enterotoxin genes to non-toxigenic strains. The receptor for CTX $\phi$  invasion (the toxin co-regulated pili) is encoded for by an operon, which forms part of the cholera pathogenicity island. This pathogenicity island has been found to be present in some non-O1 and non-O139 environmental strains suggesting that it can be transferred among *V. cholerae* strains within a population (Karaolis *et al.*, 1998). Not only could environmental *Vibrio cholerae* strains serve as progenitors for the evolution of epidemic strains, but some strains have also been shown to be genetically closely related to clinical O1 and O139 strains (Thompson *et al.*, 2003). These factors make environmental isolates an ideal model for the study of *Vibrio cholerae* O1 and O139 survival and transport in water systems.

An understanding of *Vibrio cholerae* population structure and cholera molecular epidemiology may give insight into the spread of the disease as well as the possible emergence of novel epidemic strains. After the emergence of the O139 epidemic strain, which served as an example of how novel epidemic strains can arise, many researchers are studying the genetic diversity of both epidemic as well as environmental *V. cholerae* strains. Multilocus enzyme electrophoreses was found to be of more use with environmental strains than with clinical strains obtained from epidemics, due to the fact that epidemic strains mainly belong to the same electrophoretic type (Chen *et al.*, 1991; Beltran *et al.*, 1998). Single-locus sequence typing of the cholera toxin sub-unit B was shown to have low discriminatory ability (Olsvik *et al.*, 1993) whilst multi-locus sequence typing was shown to have higher resolution (Thompson *et al.*, 2004).

Kotetishvili *et al.* (2003) found the technique to compare favourably with pulsed field gel electrophoreses (PFGE). Many researchers used PFGE to characterize *V. cholerae* as the technique exhibits good discrimination between strains (Choudhury *et al.*, 1994; Cameron *et al.*, 1994; Popovic *et al.*, 1993). The one disadvantage is, however, the inability of this technique to type strains that exhibit endonuclease activity (Pichel *et al.*, 2003). At present, the technique to be of value for genetic diversity studies is most likely amplified fragment length polymorphism fingerprinting (AFLP). This technique has been used successfully by various researchers to type both clinical and environmental *V. cholerae* strains (Jiang *et al.*, 2000 b; Lan and Reeves, 2001; Thompson *et al.*, 2003). AFLP has the advantage that no prior sequence knowledge is required, it is robust and reliable due to the stringent reaction conditions used, and it is sensitive due to the incorporated PCR (Vos *et al.*, 1995). The ability of this technique to discriminate between closely related strains makes it well suited for the typing of epidemic as well as environmental strains.

The aim of this study is determine the genetic diversity of *Vibrio cholerae* in the Vaal Barrage catchment using the AFLP technique. The results obtained could provide insight into the population structure and dynamics of environmental *V. cholerae* with reference to South African conditions.

## 4.2 MATERIALS AND METHODS

### 4.2.1 Acquisition and Maintenance of Cultures

Environmental *Vibrio cholerae* isolates were obtained from Rand Water (Vereeniging, South Africa) (Table 9). The strains were isolated over a two-year period from 14 different sampling sites in the Vaal Barrage Catchment, South Africa (Figure 1, Table 5, see chapter 3). An Enterotoxigenic strain was obtained from the National Collection of Type Cultures, Public Health Laboratory Services, London (strain NCTC 5941). Three other *Vibrio cholerae* O1 clinical strains (isolated from cholera cases in the Kwazulu-Natal and Nelspruit regions) were obtained from the National Health Laboratory Services, Johannesburg, South Africa.

*Vibrio cholerae* were grown on Nutrient Agar (Biolab, Merck C1) or alternatively on Nutrient Broth No. 2 (Oxoid CM67). Agar plates, which were incubated at 37 °C for 24 hours to obtain bacterial growth, were stored at room temperature and were subcultured on a fortnightly basis. Long-term storage of strains at -70 °C (up to three years) was achieved using Microbank™ Cryobeads (Davies Diagnostics).

**Table 9. Isolates typed with AFLP**

SAMPLING SITE	SAMPLING DATE
A18	
53	5-Feb-2001
235	7-May-2001
B10	
1	20-Dec-2000
178	2-Apr-2001
455	26-Nov-2001
515	14-Jan-2002
598	25-Feb-2002
599	25-Feb-2002

SAMPLING SITE	SAMPLING DATE
	655 25-Mar-2002
	1029 14-Oct-2002
<b>K18</b>	
	46 29-Jan-2001
	373 27-Aug-2001
	502 7-Jan-2002
	602 25-Feb-2002
	632 11-Mar-2002
	1059 20-Oct-2002
	1193 13-Jan-2003
	1246 10-Feb-2003
	1302 10-Mar-2003
<b>K19</b>	
	520 14-Jan-2002
	521 14-Jan-2002
	576 11-Feb-2002
	605 25-Feb-2002
<b>K25</b>	
	195 9-Apr-2001
	226 30-Apr-2001
	240 7-May-2001
	505 7-Jan-2002
	522 14-Jan-2002
	523 14-Jan-2002
	606 25-Feb-2002
	607 25-Feb-2002
	1250 10-Feb-2003
	1307 10-Mar-2003
	1317 10-Mar-2003
	1411 29-Apr-2003
<b>M-CANAL</b>	
	535 14-Jan-2002
	618 25-Feb-2002
	619 25-Feb-2002
	1265 10-Feb-2003
<b>N8</b>	
	3 20-Dec-2000
	209 23-Apr-2001
	456 26-Nov-2001
	501 7-Jan-2002
	628 11-Mar-2002
	661 25-Mar-2002
	1355 24-Mar-2003
	1412 29-Apr-2003
<b>R5</b>	
	446 12-Nov-2001

SAMPLING SITE	SAMPLING DATE
506	7-Jan-2002
637	11-Mar-2002
746	13-May-2002
1011	30-Sep-2002
1012	30-Sep-2002
1152	9-Dec-2002
<b>R6</b>	
2	20-Dec-2000
171	26-Mar-2001
241	7-May-2001
269	21-May-2001
391	9-Sep-2001
610	25-Feb-2002
1155	9-Dec-2002
1256	10-Feb-2003
1362	31-Mar-2003
1420	29-Apr-2003
<b>RV2</b>	
256	14-May-2001
448	12-Nov-2001
529	14-Jan-2002
612	25-Feb-2002
613	25-Feb-2002
640	11-Mar-2002
1072	30-Oct-2002
<b>S1</b>	
449	12-Nov-2001
615	25-Feb-2002
643	11-Mar-2002
671	25-Mar-2002
672	25-Mar-2002
<b>S2</b>	
11	8-Jan-2001
160	19-Mar-2001
174	26-Mar-2001
244	7-May-2001
510	7-Jan-2002
616	25-Feb-2002
617	25-Feb-2002
644	11-Mar-2002
645	11-Mar-2002
995	23-Sep-2002
1158	9-Dec-2002
<b>VRB30</b>	
1039	14-Oct-2002
1040	14-Oct-2002

SAMPLING SITE	SAMPLING DATE
1267	12-Feb-2003
1321	12-Mar-2003
1424	30-Apr-2003
1425	30-Apr-2003
VRB47	
13	15-Jan-2001
622	27-Feb-2002
650	13-Mar-2002
1025	2-Oct-2002
1053	14-Oct-2002
1054	14-Oct-2002
1322	12-Mar-2003
165	19-Mar-2001

#### 4.2.2 DNA Extraction

Whole genomic DNA was extracted and prepared using Dneasy Tissue Kit (Qiagen 69504). The manufacturer's protocol was followed, with final elution in 100 $\mu$ L buffer to obtain a high concentration of DNA. The concentration of DNA in the final elute was determined using a spectrophotometer (Nanodrop ND1000, Nanodrop Technologies, Wilmington, USA.)

#### 4.2.3 Restriction Enzyme Digestion of Genomic DNA and Adapter ligation

Restriction enzymes *MseI* [*Tru9*] (Roche 1464817) and *EcoRI* (Roche 1175084) were used to generate fragments from whole genomic DNA. Digestions were performed in 15 $\mu$ L reactions containing 3.0 $\mu$ L 5X Reaction buffer (50mM Tris HAc, 50mM MgAc, 230mM KAc and 25mM DTT), 12 units *EcoRI* and 8 units *MseI* restriction enzymes, and 100 ng template DNA. The digestions were carried out at 37°C for two hours, after which the restriction enzymes were deactivated at 70°C for 10 minutes.

Adapters as described by Lan and Reeves (2002) (Table 10) were ligated to the terminal ends of the generated fragments by adding a reaction mix containing the adapters to the restriction enzyme digestion reaction after completion. The adapter ligation reaction mix contained 1X reaction buffer (50mM Tris HAc, 50mM MgAc, 230mM KAc and 25mM DTT), 5 pmol EcoRI and 50 pmol MseI adapters, 0.3 mM ATP and 1 unit T4 DNA ligase (Roche 481220). After addition the adapters were left to ligate at room temperature for two hours. The final product was stored at -20°C for no more than 24 hours before use as PCR template. Extended storage was found to lead to DNA degradation.

**Table 10. Adapters used in AFLP ligation reaction**

Name	Sequence
925 EcoRI	5'-AATTGGTACGCAGTCAGTGAGGTTTCATTACCATCC-3'
931 MseI	5'-CCTGATTGCTACAACGATGAGTCCTGAG-3'

#### 4.2.4 PCR primers

The primers used in this study are shown in Table 11. Primers Eco-G was labeled with a fluorescent dye (Cye 5.5) in order to visualize the fragments using an automated system. All primers were synthesized by Inqaba Biotech, South Africa.

**Table 11. Primer sequences for pre-amplification and selective-amplification**

Amplification step	Primer	Sequence
Pre-amplification	Eco-O	5' – GACTGCGTACCAATTC – 3'
	Mse-O	5' – GATGAGTCCTGAGTAA – 3'
Selective-amplification	Eco-G	5' – GACTGCGTACCAATTCG – 3'
	Mse-T	5' – GATGAGTCCTGAGTAAT – 3'



#### 4.2.5 PCR amplification of fragments

Two PCR amplification steps were carried out, the first was the pre-amplification where all terminally heterogeneous fragments were amplified (fragments generated by *EcoRI* on one end and *MseI* on the other). This was followed by a selective-amplification step where both primers have an additional 3' base, which gives a theoretical 16-fold reduction in the number of fragments amplified. The template for the pre-amplification, adapter ligation reaction from 4.2.3, was diluted ten fold with nuclease free water (Whitehead Scientific P1193). Each of the pre-amplification reactions contained 2 $\mu$ L of the template (1:10 Dilution of adapter ligation), 0.2 $\mu$ L of each of the dNTP's (Abgene Ab-11.24), 2nM MgCl<sub>2</sub> (JMR 801), 1X Reaction Buffer (JMR801), 1 $\mu$ M of each of the primers (Table 11) and 1 unit Taq DNA polymerase (SuperTherm Taq, JMR 801). The reaction was made up to 25 $\mu$ L with nuclease free water (Whitehead scientific P1193), and the PCR carried out in a Perkin-Elmer 2700 Gene-amp (Applied Biosystems). The PCR parameters for pre-amplification were: Initial denaturation at 94°C for 3 minutes, 20 cycles of amplification consisting of denaturation at 94°C for 30 seconds, primer annealing at 56°C for 1 minute and DNA extension 72°C for 1minute. Synthesis was completed at 72°C for 7 minutes. The PCR product was then diluted 50 fold (with nuclease free water) for use as template in the selective-amplification. Each of the selective-amplification reactions contained 5 $\mu$ L of the template (1:50 Dilution of pre-amplification), 0.2 $\mu$ L of each of the dNTP's (Abgene Ab-11.24), 2nM MgCl<sub>2</sub> (JMR 801), 1X Reaction Buffer (JMR801), 1 $\mu$ M of each of the primers (Table 8) and 1 unit Taq DNA polymerase (SuperTherm Taq, JMR 801). The reaction was made up to 20 $\mu$ L

with nuclease free water (Whitehead scientific P1193), and the 'touch-down' PCR carried out in a Perkin-Elmer 2700 Gene-amp (PE Applied Biosystems, Weiterstadt, Germany). The PCR parameters for selective-amplification were: initial denaturation at 94°C for 3 minutes, the first cycle consisted of denaturation at 94°C for 30 seconds, primer annealing at 65°C for 30 seconds and DNA extension at 72°C for 1 minute. After the first cycle the annealing temperature was lowered with 1°C per cycle until it reached 56°C and the amplification was therefore completed with the following parameters: 23 cycles of amplification consisting of denaturation at 94°C for 30 seconds, primer annealing at 56°C for 30 seconds, DNA extension at 72°C for 1 minute, and completion of synthesis at 72°C for 7 minutes. To verify that amplification took place the pre-amplification and selective amplification products were electrophoresed on a 1% agarose gel (Hispanagar D-1 LE) and visualized with ethidium bromide staining (10mg/ml). Successful amplification produced smears.

#### **4.2.6 Polyacrylamide Gel Electrophoreses and visualization**

The AFLP amplification reactions were separated using an automated PAGE system (Licor Global IR2 DNA analyzer, Licor Inc. Nebraska, USA). Two µL of the selective-amplification product was mixed with 2 µL loading dye prior to denaturation at 90°C for 10 minutes, thereafter the sample was kept on ice until loaded. A pre-run (10 Watt for 30 minutes) was done to homogenize the gradients within the PAGE gel before the samples were loaded onto the gel. Gel electrophoreses was carried out at 35 Watt for four hours, the temperature being controlled at 55°C. A sizing standard of between 50-700bp (Li-Cor 4200-60[700]) was used as reference. The Automated Li-Cor system generated

digitized fingerprints (16 Bit TIFF images) during the gel run, which were used in analysis with compatible analysis software.

#### **4.2.7 Analysis of fingerprints**

Digitized fingerprints were analyzed using GelCompar II software (Applied Maths, Kortrijk, Belgium) following the manufacturers instructions. Images were straightened and normalized by alignment to molecular size standards loaded on each gel. Curve-based dendograms were generated using Pearson Correlation Coefficient with 1% optimization and 0.14 position tolerance values.

### **4.3 RESULTS**

Restriction enzymes EcoRI and MseI were used to generate DNA fragments from 104 environmental and clinical *Vibrio cholerae* isolates. The AFLP patterns generated with primers listed in Table 11 contained 45 to 60 bands per isolate. Diverse banding patterns were observed within environmental isolates, while clinical strains grouped loosely together. Similarity analysis with Pearsons correlation coefficient yielded a number of clusters, with environmental isolates grouping in various clusters, while the four clinical isolates together with two environmental strains were grouped in a separate cluster. The clinical isolates were not of the same origin, two originated from the Nelspruit region, one from the Kwazulu-Natal province, and the last being a *Vibrio cholerae* reference isolate (NCTC 5941). Environmental isolates grouped in clusters showing no or little resemblance to their isolation site and or isolation date (Fig 5). Thus there was no correlation between site of isolation and date of isolation with isolates obtained from the various different river systems over a period of three years.

Using a Pearson correlation value of 90% or higher, clusters were further investigated (Table 12). In 6 of the 17 clusters created by this cut-off value (clusters consisted of 2 or more strains, outbreak strains were excluded) the isolates grouped together were of the same sampling site or geographically closely related sampling sites. Sampling site N8 and R5 were regarded as being closely related, with the R5 sampling site situated less than 2km upstream of the convergence of the Rietspruit and Natalspruit rivers, which represents site N8. These 6 clusters contained strains isolated on different sampling dates, with the dates of isolation differing with up to seven months. Five of the 17 clusters only contained isolates sampled at the same site on the same date. These isolates were disregarded, as they may be duplicates of one strain. This effectively reduced the clusters of interest to 12. Isolates from unrelated sampling sites and different sampling dates made up the remainder of the 17 clusters.

**Table 12. Isolates present in Clusters showing a Pearson correlation of 90% or higher**

Isolate no.	Sampling site	Sampling date
<b>Correlation of sampling site</b>		
622	VRB47	27-Feb-2002
1054	VRB47	14-Oct-2002
1053	VRB47	14-Oct-2002
616	S2	25-Feb-2002
645	S2	11-Mar-2002
174	S2	26-Mar-2002
617	S2	25-Feb-2002
644	S2	11-Mar-2002
672	S1	25-Mar-2002
501	N8	7-Jan-2002

Isolate no.	Sampling site	Sampling date
506	R5	7-Jan-2002
521	K19	14-Jan-2002
576	K19	11-Feb-2002
195	K25	23-Apr-2001
226	K25	30-Apr-2001
<b>Correlation of sampling site and date - Possible duplicates</b>		
598	B10	25-Feb-2002
599	B10	25-Feb-2002
1424	VRB30	30-Apr-2003
1425	VRB30	30-Apr-2003
612	RV2	25-Feb-2002
613	RV2	25-Feb-2002
1039	VRB30	14-Oct-2002
1040	VRB30	14-Oct-2002
1011	R5	30-Sep-2002
1012	R5	30-Sep-2002
<b>Diverse sampling sites/dates</b>		
256	RV2	14-May-2001
510	S2	7-Jan-2002
610	R6	25-Feb-2002
618	M-Canal	25-Feb-2002
1411	K25	29-Apr-2003
1412	N8	20-Apr-2003
160	S2	19-Mar-2001
619	M-canal	25-Feb-2002
171	R6	26-Mar-2001
615	S1	25-Feb-2002
3	N8	20-Dec-2000
13	VRB47	15-Jan-2001

#### 4.4 DISCUSSION

Molecular techniques are an invaluable tool for determining population structure and dynamics of pathogenic bacteria. Researchers have used a variety of these techniques to study the genetic diversity of *Vibrio cholerae*, including pulsed field gel electrophoreses, restriction fragment length polymorphisms, sequencing of genes, ribotyping and amplified fragment length polymorphism fingerprinting (Cameron *et al.*, 1994; Wachsmuth *et al.*, 1991; Dalsgaard *et al.*, 1995; Olsvik *et al.*, 1993; Kotetishvilli *et al.*, 2003; Jiang *et al.*, 2000 a, Lan and Reeves, 2002). Most of these studies focused on clinical strains, where the epidemic as well as pandemic spread of cholera was the main issue of interest. Only in recent years have researchers turned to studying the genetic diversity of environmental *Vibrio cholerae* isolates (Jiang *et al.*, 2000 a; Jiang *et al.*, 2000 b; Thompson *et al.*, 2003). Amplified fragment length polymorphism DNA fingerprinting has the advantage that it allows the differentiation of highly related bacterial genomes in a highly reproducible fashion. The technique also does not require prior sequence knowledge, and combines the power of RFLP analysis with the flexibility associated with PCR-based technology (Janssen *et al.*, 1996). The combination of these characteristics makes AFLP a highly suited tool for determining genetic diversity within environmental *Vibrio cholerae* populations. This study focused on closely-related environmental *V. cholerae* strains and a primer pair with high discriminatory ability was sought. The discriminatory power of a typing technique can be measured by the Simpson index of diversity (D), the higher the D values the less likely the probability of type I errors (where highly related strains are not shown to be as such). Lan and Reeves (2002)

showed that the combination of primers Eco-G and Mse-T could obtain a D-value of 0.837 when discriminating closely related seventh pandemic strains. The decision was therefore made to use this primer pair in this study because they showed the highest D-value within the Eco/Mse group.

*Vibrio cholerae* is an autochthonous inhabitant in the Vaal Barrage Aquatic system, with environmental strains being isolated on a regular basis. Factors that have been shown to influence the survival of *V. cholerae* in inland water systems include water temperature, water pH, salinity and associations with water fauna (Jiang *et al.*, 2000 a; Huq *et al.*, 1983; Huq *et al.*, 1984; Halpern *et al.*, 2003). The conditions in the Vaal Barrage system is favourable for *Vibrio cholerae* propagation and survival with water temperatures ranging between 17°C and 23°C, pH values averaging around 8.0, and Na present in concentrations ranging between 5 and 300 mg/ml (as measured from Jan. 2001 to Oct. 2003). These favourable conditions have lead to the establishment of diverse *V. cholerae* populations in the river system. The clusters generated using AFLP patterns and a cut-off value of 90% contained isolates from different isolation sites and sampling dates. This indicates that *V. cholerae* strains with similar genotypes are spread throughout the aquatic system, and are able to survive for extended periods. It was seen that the population is genetically diverse. Many of the clusters were only distantly related to other *Vibrio cholerae* clusters, indicating how well established and diversified the *V. cholerae* population is. This diversity is not only observed when looking at the entire aquatic system, but also when focusing on specific sampling sites. Isolates from the same sampling site rarely grouped closely together. Where clusters do contain highly related

isolates from the same sampling site the sampling dates do not differ by more than a few months. The similarity is possibly lost due to small genome rearrangements over time. The clinical isolates used in this study grouped together, even though originating from different sources. The clinical strains originating from Nelspruit and Kwazulu-Natal showing an extremely high level of similarity suggesting that they were derived from the same strain. The inclusion of an environmental isolate in the same cluster, but at a lower similarity, as the previously mentioned clinical isolates suggests that there could be a high degree of relatedness between some of the enterotoxin producing strains and environmental non-O1, non-O139 strains.

With the high level of genetic diversity seen in this study we believe that the *Vibrio cholerae* population in the Vaal Barrage system is not a product of one or two strains that have adapted to local conditions, nor does it consist of diverse clones that only occupy specific niches in the aquatic system. The population in the Vaal Barrage is rather made up of highly diverse clones that constantly compete, resulting in genetic shifts only perceivable within short time frames and localized regions. This constant adaptation with the associated genetic rearrangements gives an indication of the genetic flexibility of environmental *Vibrio cholerae* strains, and may portray the potential of environmental populations to serve as reservoirs for future epidemic strains. Using environmental *Vibrio cholerae* as a model, we suggest that enterotoxigenic strains may exhibit the same degree of persistence and survival in inland aquatic systems, being able to survive for extended periods and posing as a potential future health risk.





**Fig 5.** Dendogram derived from AFLP patterns of 104 *Vibrio cholerae* strains. Dendogram was constructed using the Pearson coefficient and unweighted pair group method of arithmetic averages.

