CHAPTER 3

RESULTS

3.1 Comparison of DNA extraction methods

DNA templates prepared by both boiling and phenol-chloroform extraction methods could be amplified (Fig. 3.1). However, there were some differences in the efficiency of PCR amplifications. For example, when DNA prepared by the boiling method was used, 26 out of 46 strains showed the presence of both the *set1A* and *set1B* amplicons. When DNA from the phenol-chloroform extraction method was used, only 16 and 14 strains showed amplifications for both the above genes respectively. Similarly, amplifications for the *ial*, *ipaH* and *sen* genes were also more successful using boiled lysates (52.2%, 100% and 28.3% respectively) than the phenol-chloroform extracted templates (43.5%, 82.6% and 6.5% respectively). Reproducible results were obtained when the PCR assays were repeated twice for each method.

Paired samples t-test was performed on both sets of results. The hypotheses tested were:

H_o : There is no difference between the presence of genes in DNAs prepared by the boiling and phenol-chloroform extraction methods

H_a : There is a difference between the presence of genes in DNAs prepared by the boiling and phenol-chloroform extraction methods

The test showed that differences of results using boiling and phenol-chloroform extraction methods were significant ($p = 0.00$) by rejecting the null hypothesis at a 0.05 significance level (Table 3.1).

As the boiling method was a simple and rapid means to extract DNA for amplification, the method was then adopted for subsequent analyses.

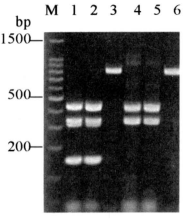

Fig 3.1

Amplicons generated from DNAs prepared by the boiling and phenol-chloroform extraction methods

Lane M: 100-bp marker
1: mPCR result from boiled lysate of TH13/00
2: mPCR result from phenol-chloroform extracted template of TH13/00
3: monoplex PCR result from boiled lysate of TH13/00
4: mPCR result from boiled lysate of TH2/00
5: mPCR result from phenol-chloroform extracted template of TH2/00
6: monoplex PCR result from boiled lysate of TH2/00

Table 3.1
Prevalence of virulence-associated genes in DNAs prepared by the boiling and phenol-chloroform extraction methods

Method	Total strains	*set1B* (%)	*set 1A* (%)	*ial* (%)	*ipaH* (%)	*sen* (%)
Boiling	46	26 (56.5)	26 (56.5)	24 (52.2)	46 (100)	13 (28.3)
Phenol-chloroform	46	16 (34.8)	14 (30.4)	20 (43.5)	38 (82.6)	3 (6.5)

Paired samples t-test: $\alpha = 0.05$, $p = 0.00$

3.2 Optimization of monoplex PCRs

A monoplex PCR for each primer set was initially carried out based on a published report by Vargas *et al.* (1999). Although the concentrations of $MgCl_2$ (3mM), dNTP (400 μM each) and primers (1 μM each) were used as recommended, unspecific bands were present together with bright primer-dimers. In order to reduce the background noise and the presence of primer-dimers, concentrations of 0.5 μM of each primer and 200 μM of each dNTP were utilized. Despite obliterating unspecificity at *set1A*, *set1B* and *ipaH* amplifications, the presence of unspecific bands remained in *ial* and *sen* amplifications. Following optimization of the concentrations of $MgCl_2$ (2 to 4 μM) and dNTP (100, 130 and 150 μM each), equally intense amplicons were produced with a clean background (Fig 3.2). Similar amplification conditions were carried out throughout the optimization processes (Section 2.2.3.1).

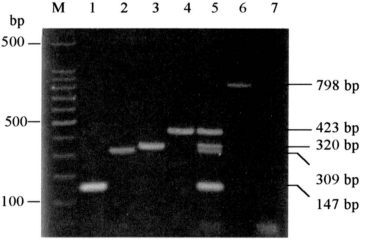

Fig 3.2

Optimized monoplex PCR and mPCR

Lane	M:	100-bp marker	4:	*ipaH*	(423 bp)
	1:	*set1B* (147 bp)	5:	mPCR	
	2:	*set1A* (309 bp)	6:	*sen*	(798 bp)
	3:	*ial* (320 bp)	7:	negative control	

3.3 Optimization of multiplex PCR (mPCR)

Initial attempts to amplify all the virulence-associated genes in a single reaction using the same reaction condition as in monoplex PCR, were not successful. The amplification for *sen* (798 bp) was inhibited whilst intense amplicons were seen for both *ipaH* (423 bp) and *set1B* (147 bp) (Fig 3.3). Consequently, optimizations for various concentrations of mPCR mixture and amplification conditions were performed.

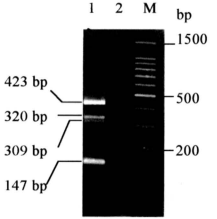

Fig 3.3

Amplicons generated by a combination of 5 primer sets at an equimolar concentration of 0.5 μM each in mPCR

Lane 1: positive control strain (TH13/00)
2: negative control
M: 100-bp marker

3.3.1 Different primer concentrations

A common practice in mPCRs involving any non-amplification of a required gene (deemed as the "weak" locus) is to increase the amount of primers of the gene at the same time with a decrease of the amount of primers for all loci that can be amplified, especially those with strong amplifications. Hence, the concentrations of primers for both *ipaH* (Shig) and *set1B* (ShET1B) were reduced to 0.4 μM each, and the primers for *sen* (ShET2) were increased in a range of 0.6 to 0.8 μM. The remaining primer sets, ShET1A and ial, were at 0.5 μM each. Only PCR products of the *set1B*, *set1A*, *ial* and *ipaH* genes were generated. No *sen* band was seen on the gel, regardless of the increased concentration of ShET2 (Fig 3.4).

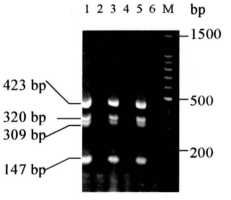

Fig 3.4

Optimization of primer concentrations in mPCR

Lane 1: ShET2 at 0.6 μM; Shig and ShET1B at 0.4 μM each; ShET1A and ial at 0.5 μM each
2: negative control
3: ShET2 at 0.7 μM; Shig and ShET1B at 0.4 μM each; ShET1A and ial at 0.5 μM each
4: negative control
5: ShET2 at 0.8 μM; Shig and ShET1B at 0.4 μM each; ShET1A and ial at 0.5 μM each
6: negative control
M: 100-bp marker

The next attempt at primer concentration optimization involved increasing ShET2 concentrations from 0.6 to 0.9 μM simultaneously with the reduction of Shig and ShET1B concentrations from 0.3 to 0.1 μM (Fig 3.5). A band of ~798 bp could be detected in the presence of unspecific bands. Although the band was stronger in signal intensity as ShET2 concentration was increased, so did the brightness of unspecific bands of ~ 500, 700 and above 800 bp in sizes. The brightness of *ipaH*, *ial*, *set1A* and *set1B* bands remained the same.

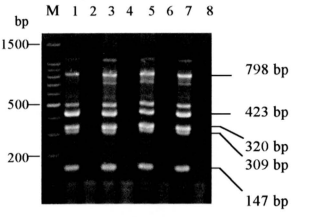

Fig 3.5

Further optimization of primer concentrations in mPCR

Lane M: 100-bp marker
1: ShET2 at 0.6 μM; Shig and ShET1B at 0.3 μM each; ShET1A and ial at 0.5 μM each
2: negative control
3: ShET2 at 0.7 μM; Shig and ShET1B at 0.2 μM each; ShET1A and ial at 0.5 μM each
4: negative control
5: ShET2 at 0.8 μM; Shig and ShET1B at 0.1 μM each; ShET1A and ial at 0.5 μM each
6: negative control
7: ShET2 at 0.9 μM; Shig and ShET1B at 0.1 μM each; ShET1A and ial at 0.5 μM each
8: negative control

3.3.2 Different buffer concentrations

Buffer concentration may affect mPCR amplification in spite of the fact that it is seldom considered in monoplex optimization (Section 1.7.1.2). Thus, the optimization of buffer concentrations was carried out with a range of 1.4X to 2.4X (Fig 3.6). As the buffer concentration was increased, unspecific bands together with the *sen* product disappeared. The other bands however, were still visible until a 2.0X buffer concentration (lane 4). Since buffer concentrations of 1.6X and 1.8X gave similar results, they were used to optimize the annealing temperature parameter, separately (Section 3.3.3).

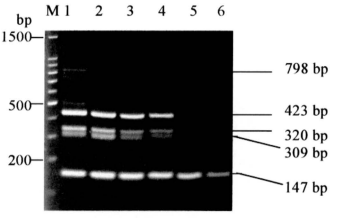

Fig 3.6

Optimization of buffer concentration in mPCR

Lane M: 100-bp marker
1: 1.4X
2: 1.6X
3. 1.8X
4. 2.0X
5. 2.2X
6. 2.4X

3.3.3 Different annealing temperatures

A range of annealing temperatures from 49 to 59°C was tested under two separate buffer concentrations: 1.6X and 1.8X (Fig 3.7). A band of ~ 798 bp was seen in all lanes in both gels. At 1.6X buffer concentration, there was strong presence of unspecific bands (gel L). At 1.8X buffer concentration, the 798 bp band became more intense with the increase of annealing temperatures (gel R). An unspecific band of ~700 bp also became more intense as the temperatures were increased. The signal intensity of the other four bands remained unaffected. Although lane 1 had all the desired amplifications without unspecificity (annealing temperature of 49°C), the faintness of 798-bp amplicon may be missed during screenings of strains. An optimal annealing temperature of 55°C (lane 4) with a buffer concentration of 1.8X was chosen for further optimization work with *Taq* DNA polymerase and dNTP concentrations.

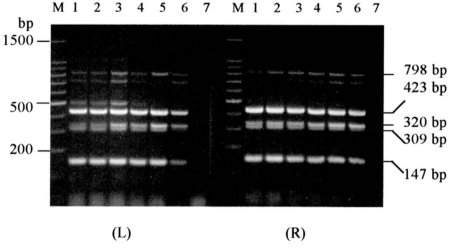

Fig 3.7

Optimization of annealing temperature in mPCR

1.6X buffer concentration (L)
1.8X buffer concentration (R)

Lane M: 100-bp marker
1: 49 °C
2: 51 °C
3: 53 °C
4: 55 °C
5: 57 °C
6: 59 °C
7: negative control

3.3.4 Different *Taq* DNA polymerase concentrations

A range of *Taq* DNA polymerase concentrations from 0.6 to 4U were evaluated. Fig 3.8 shows that at high *Taq* DNA polymerase concentrations of 2.0 to 4.0U per reaction, background products became as intense as the targeted PCR products. At the highest *Taq* concentration of 4.0U, the amplification efficiency of the *setIA* (309 bp) gene was decreased. Nevertheless, at lower enzyme concentrations, the presence of the 798 bp band was unstable or very faint. The other 4 bands remained unaffected.

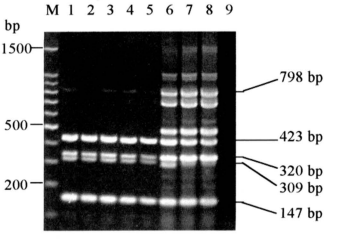

Fig 3.8

Optimization of *Taq* DNA polymerase concentration in mPCR

Lane M: 100-bp marker
1: 0.6 U
2. 0.7 U
3. 0.8 U
4. 0.9 U
5. 1.0 U
6. 2.0 U
7. 3.0 U
8. 4.0 U
9. negative control

3.3.5 Different dNTPs concentrations

In this study, regardless of the dNTP concentration increase from 130 to 220 μM each (at a constant $MgCl_2$ concentration of 4mM), there was a bare presence of the recalcitrant *sen* in all lanes (Fig 3.9).

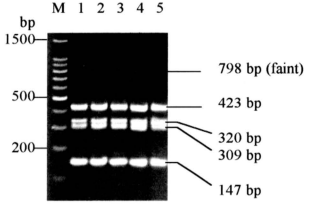

Fig 3.9

Optimization of dNTPs concentration in mPCR

Lane M: 100-bp marker
1: 130 μM
2: 160 μM
3. 180 μM
4. 200 μM
5. 220 μM

Following the extensive experiments to co-amplify the *sen* product (798 bp) efficiently with the other products of interest, it was observed that the band was weakly amplified, its presence was inhibited altogether or its amplification was associated with unwanted background products. Hence, for the remaining amplification experiments in this study, *set1B*, *set1A*, *ial* and *ipaH* would be detected in a single mPCR with *sen* in a separate monoplex PCR (Fig 3.2 lanes 5 and 6).

The reaction mixture for the mPCR is as follows:

Stock solution	Volume (µl)	Final concentration (µM)
10X PCR buffer B	4.50	1.8X
25 mM $MgCl_2$	4.00	4.00 mM
5 mM dNTP	0.65	130.00
50 µM ShET1B (F + R)	0.15 + 0.15	0.30 each
ShET1A (F + R)	0.25 + 0.25	0.50 each
ial (F + R)	0.25 + 0.25	0.50 each
Shig1 + Shig 2	0.15 + 0.15	0.30 each
Sterile ddH_2O	12.05	-
Taq DNA polymerase (5U/µl)	0.20	1U
DNA template	2.00	-
Total	25.00	

The reaction mixture for the monoplex PCR is as shown below:

Stock solution	Volume (µl)	Final concentration (µM)
10X PCR buffer B	2.50	1X
25 mM $MgCl_2$	4.00	4.00 mM
5 mM dNTP	0.65	130.00
50 µM primer (F)	0.25	0.50
(R)	0.25	0.50
sterile ddH_2O	15.15	-
Taq DNA polymerase (5U/µl)	0.20	1U
DNA template	2.00	-
Total	25.00	

PCR amplifications for both reactions were carried out at the following conditions: a cycle of initial denaturation at 95^0C for 5 min; 30 cycles of denaturation at 95^0C for 50 sec, annealing at 55^0C for 1.5 min, extension at 72^0C for 2 min; a cycle of final extension at 72^0C for 7 min.

3.4 Sequencing

To verify the identities of the PCR amplicons of the virulence-associated genes, PCR products were gel-purified and sequenced. Standard nucleotide-nucleotide BLAST search software was used to evaluate the sequences of PCR amplicons with published sequences in the GenBank (Appendix 2).

The DNA fragment of *set1B* amplicon shared a 99% identity with published gene sequence in the GenBank (accession number Z47381) (Fig 3.10i).

set1A amplicon had a 100% sequence identity with its counterpart in the GenBank (accession number Z47381).

The *ial* gene showed a 99% sequence identity with the published sequence (GenBank accession number D13663). There was a nucleotide substitution occurring at nucleotide 5553 (T → C) (Fig 3.10ii).

The *ipaH* gene registered a 100% sequence identity with the published gene sequence (GenBank accession number M76445).

The gene sequence from *sen* amplicon shared a 100% identity with the referred sequence (GenBank accession number Z54211).

The high percentage of sequence similarities showed that all the amplicons in this study were identical or almost identical to the referred sequences in the GenBank. Hence, they were not mere artefacts that happened to share similar molecular weights as the published sequences.

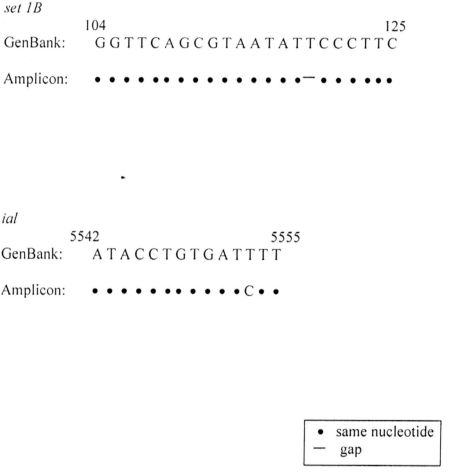

Fig 3.10

Polymorphic site(s) within the virulence-associated genes in *Shigella* spp.

3.5 Reproducibility

A total of 28 *Shigella* strains were chosen randomly from 1997 to 2000 to determine the reproducibility of the amplification results. Both PCR assays were carried out twice. When the second set of results differed from the first, a third assay was performed. All the virulence-associated genes amplified in the mPCR were 100% reproducible, however only 78.6% of the *sen* amplification could be reproduced (Table 3.2).

Table 3.2
Reproducibility of amplification results

Year of isolation	Number of strains	*set1B*	*set1A*	*ial*	*ipaH*	*sen*
1997	4	++	++	++	++	+x+
		++	++	++	++	++
		--	--	--	++	--
		++	++	--	++	--
1998	10	--	--	++	++	--
		++	++	--	++	--
		++	++	++	++	++
		++	++	--	++	--
		--	--	--	++	--
		--	--	--	++	--
		--	--	--	++	--
		++	++	++	++	++
		--	--	--	++	--
		--	--	--	++	--
1999	4	++	++	--	++	--
		--	--	--	++	--
		++	++	--	++	--
		++	++	++	++	--
2000	10	++	++	++	++	++
		--	--	++	++	++
		++	++	++	++	+xx
		++	++	++	++	--
		++	++	--	++	--
		--	--	++	++	+xx
		++	++	++	++	+xx
		--	--	++	++	+xx
		++	++	++	++	+xx
		--	--	++	++	+xx
Reproducibility %		100	100	100	100	78.6

+ : presence of gene
- : absence of gene
x : not detected in repeats

3.7 Sensitivity

The mPCR assay was tested on 10-fold dilutions of an overnight culture of *S. flexneri* 2a. All the four virulence-associated genes were detected until 10^{-3} dilution (Fig 3.12). This was equivalent to 2.45 X 10^5 cfu/ml of lysate or a minimum of 490 cfu of shigellae per mPCR reaction.

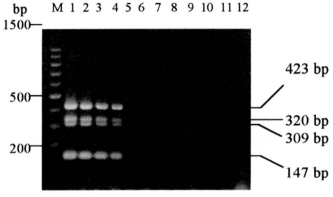

Fig 3.12

Sensitivity result of mPCR assay by using DNA templates from bacterial culture at different dilutions

Lane M: 100-bp marker
1: undiluted bacterial culture (4 bands)
2: 10^{-1} dilution (4 bands)
3: 10^{-2} dilution (4 bands)
4: 10^{-3} dilution (4 bands)
5: 10^{-4} dilution
6: 10^{-5} dilution
7: 10^{-6} dilution
8: 10^{-7} dilution
9: 10^{-8} dilution
10: 10^{-9} dilution
11: 10^{-10} dilution
12: negative control

3.8 Faecal-spiking and sensitivity

A prior experiment using the similar protocol (Section 2.2.3.8) with a smaller amount of faecal sample (~ 0.05 g) and no additional BHI for preincubation, only managed to detect all the virulence-associated genes until 10^{-4} dilution (Fig 3.13).

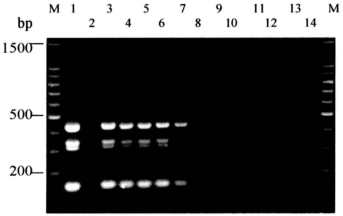

Fig 3.13

Faecal-spiking and sensitivity result of mPCR (without additional BHI)

Lane				
M:	100-bp marker			
1:	TH13/00 (positive control)			
2:	unspiked faeces (negative control)			
3:	undiluted spiked faeces (4 bands)			
4:	10^{-1} dilution (4 bands)	9:	10^{-6} dilution	
5:	10^{-2} dilution (4 bands)	10:	10^{-7} dilution	
6:	10^{-3} dilution (4 bands)	11:	10^{-8} dilution	
7:	10^{-4} dilution (4 bands)	12:	10^{-9} dilution	
8:	10^{-5} dilution (2 bands)	13:	10^{-10} dilution	
		14:	"water blank"	

When more faecal sample (~0.2g) was used and preincubated in additional BHI (500 µl) for 4h, the detection limit was to 10^{-5} dilution (Fig 3.14), which was equivalent to 5.0 X 10^4 cfu/ml or 100 cfu shigellae per mPCR assay. Its sensitivity was 10-fold more than the initial experiment, despite more faeces being used and therefore, a larger amount of PCR inhibitors would be present during mPCR amplification. The test was carried out twice and the average detection sensitivity limit was as reported above.

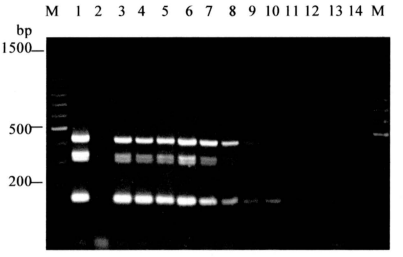

Fig 3.14

Faecal-spiking and sensitivity result of mPCR (with additional BHI)

Lane				
M:	100-bp marker			
1:	TH13/00 (positive control)			
2:	unspiked faeces (negative control)			
3:	undiluted spiked faeces (4 bands)			
4:	10^{-1} dilution (4 bands)	9:	10^{-6} dilution (2 bands)	
5:	10^{-2} dilution (4 bands)	10:	10^{-7} dilution (1 band)	
6:	10^{-3} dilution (4 bands)	11:	10^{-8} dilution (1 band)	
7:	10^{-4} dilution (4 bands)	12:	10^{-9} dilution (1 band)	
8:	10^{-5} dilution (4 bands)	13:	10^{-10} dilution (1 band)	
		14:	"water blank"	

3.9 Prevalence of virulence-associated genes in Malaysian strains

One hundred and ten strains of *Shigella* spp. obtained from IMR, Malaysia at the time of study were used to test the applicability of both amplification systems. Statistical analyses were also performed to determine if there was a significant difference amongst the prevalence of genes detected in the study as well as to confirm the almost-exclusive presence of *set1A* and *set1B* in *S. flexneri* 2a serotype as proposed by Noriega *et al.* (1995).

All the strains (110/110) were positive for *ipaH* (Table 3.3). The next most prevalent genes detected in the collection were *set1B*, *set1A* and *ial* (~ 41%). There was also a complete correlation in the occurrence of *set1B* and *set1A* whereby any strain that had a *set1B* gene would also give a positive result for *set1A* (Appendix 1). The *sen* gene registered the lowest prevalence (34/110 strains or ~ 31%).

Table 3.3
Prevalence of all the virulence-associated genes in Malaysian *Shigella* strains

Virulence-associated genes	presence (%)	absence (%)
set1B	45 (40.9)	65 (59.1)
set1A	45 (40.9)	65 (59.1)
ial	45 (40.9)	65 (59.1)
ipaH	110 (100)	0 (0)
sen	34 (30.9)	76 (69.1)

Paired samples t-test was used to compare the differences in prevalence of virulence-associated genes at a 5% significance level. The hypotheses were set as:

H_0: There is no significant difference between the prevalence of gene *a* and gene *b*

H_a: There is a significant difference between the prevalence of gene *a* and gene *b*

As there was a 100% correlation in the prevalence of *set1B* and *set1A*, all the statistical analyses in this study were conducted using the results of *set1A* to represent both genes.

Results of the test are presented in Table 3.4. As expected, there was a significant difference between the prevalence of *ipaH* and *set1A*, *ipaH* and *ial*, and *ipaH* and *sen* (p=0.00) in the Malaysian strains. The prevalence of *ial* had a slightly

significant difference from *sen* (p=0.048). However, there was no significant difference in the prevalence of *set1A* and *ial* (p=1.00) and, *set1A* and *sen* (p=0.10).

Table 3.4
Paired samples t-test results

Pair of genes	p-value
set1A-ial	1.00
set1A-ipaH	0.00
set1A-sen	0.10
ial-ipaH	0.00
ial-sen	0.048
ipaH-sen	0.00

$\alpha = 0.05$

3.9.1 Distribution of virulence-associated genes according to species and serotype

All the four species were represented in the IMR collection although the number of strains per species differed from each other. There were 84 strains of *S. flexneri* (8 serotypes), 15 strains of *S. sonnei* (1 serotype), 10 strains of *S. dysenteriae* (2 serotypes) and 1 strain of *S. boydii* (1 serotype) (Table 3.5). *S. flexneri* had the largest number of serotypes with *S. flexneri* 2a being the most common (47 strains), followed by serotype 3a (18 strains) and serotype 3c (10 strains).

Nearly all the *S. flexneri* 2a (41/47) strains had *set1B* and *set1A*. These genes, encoding for ShET1 enterotoxin, were also detected in *S. flexneri* 3a (3 strains) and *S. flexneri* 4a (1 strain) serotypes.

The *ial* gene was present in ~ 40% (19/47), ~ 67% (12/18) and 80% (8/10) strains of *S. flexneri* 2a, *S. flexneri* 3a and *S. dysenteriae* 2.

The *ipaH* gene was present in all species and serotypes. However, *sen* gene was only present in *S. flexneri* 2a (19 strains), *S. flexneri* 3a (8 strains), *S. dysenteriae* 2 (5 strains), *S. flexneri* 3c (1 strain) and *S. flexneri* 6 (1 strain) serotypes.

Table 3.5
Prevalence of virulence-associated genes according to species

Species	Total strains	*set1B* (%)	*set1A* (%)	*ial* (%)	*ipaH* (%)	*sen* (%)
S. flexneri 1a	3	0 (0)	0 (0)	1 (33.3)	3 (100)	0 (0)
1b	3	0 (0)	0 (0)	0 (0)	3 (100)	0 (0)
2a	47	41 (87.2)	41 (87.2)	19 (40.4)	47 (100)	19 (40.4)
3a	18	3 (16.7)	3 (16.7)	12 (66.7)	18 (100)	8 (44.4)
3c	10	0 (0)	0 (0)	1 (10)	10 (100)	1 (10)
4a	1	1 (100)	1 (100)	1 (100)	1 (100)	0 (0)
6	1	0 (0)	0 (0)	1 (100)	1 (100)	1 (100)
y	1	0 (0)	0 (0)	0 (0)	1 (100)	0 (0)
S. sonnei	15	0 (0)	0 (0)	2 (13.3)	15 (100)	0 (0)
S. dysenteriae 2	10	0 (0)	0 (0)	8 (80)	10 (100)	5 (50)
S. boydii 6	1	0 (0)	0 (0)	0 (0)	1 (100)	0 (0)
Total	110	45	45	45	110	34

Chi-square (χ^2) test was conducted to determine if there was a significant association between the prevalence of genes and *S. flexneri* 2a serotype (α =0.05). The hypotheses were:

H_o: The prevalence of gene *a* is not significantly associated with *S. flexneri* 2a serotype

H_a: The prevalence of gene *a* is significantly associated with *S. flexneri* 2a serotype

There was an association in the prevalence of *set1B* and *set1A* with *S. flexneri* 2a serotype (χ^2 =72.85, p =0.00) (Table 3.6). Cramer's V, which measures the strength of association from 0 (no association) to 1 (perfect association), had indicated a strong association between the variables (0.81). On the other hand, the prevalence of *ial* and *sen* was not significantly associated to *S. flexneri* 2a serotype (both p-values were >0.05, Cramer's V <0.20). As *ipaH* was amplified in all the strains (ie. a constant), no statistical analysis was performed.

Table 3.6
Chi-square test results for the determination of a significant association between the prevalence of virulence-associated genes and *S. flexneri* 2a serotype

Virulence-associated genes	χ^2		Cramer's V
	value	p	
set1B	72.85	0.00	0.81
set1A	72.85	0.00	0.81
ial	0.01	0.93	0.01
sen	3.48	0.06	0.18

α =0.05

3.9.2 Distribution of virulence-associated genes according to year of isolation

Sixty-five strains of *Shigella* were isolated in 1998, 16 strains each in 1997 and 2000, and the remaining 13 strains from 1999 (Table 3.7). This uneven distribution of strains per year was due to the outbreaks of shigellosis in Perak and Kelantan in 1998 (Appendix 1).

There was no fixed pattern of the prevalence of virulence-associated genes in *Shigella* spp. over the years. Twenty-four strains yielded positive results for both *set1B* and *set1A* in 1998, followed by 8 strains in 1997, 7 strains in 2000 and 6 strains in 1999. The majority of *ial*-positive strains were isolated in 1998 (n=24), followed by years

2000 (n=10), 1997 (n=6) and 1999 (n=5). As for *ipaH*, 65 strains were from 1998, 16 strains each from 1997 and 2000, and 13 strains from 1999. There were 19 strains with *sen* amplification in 1998, 9 strains in 2000, 6 strains in 1997 and none in 1999.

Table 3.7
Prevalence of virulence-associated genes according to year of isolation

Gene / Year	1997 (%)	1998 (%)	1999 (%)	2000 (%)	Total (%)
set1B	8 (17.8)	24 (53.3)	6 (13.3)	7 (15.6)	45 (100)
set1A	8 (17.8)	24 (53.3)	6 (13.3)	7 (15.6)	45 (100)
ial	6 (13.3)	24 (53.3)	5 (11.1)	10 (22.2)	45 (100)
ipaH	16 (14.6)	65 (59.1)	13 (11.8)	16 (14.6)	110 (100)
sen	6 (17.7)	19 (55.9)	0 (0)	9 (26.5)	34 (100)
Total strains	16	65	13	16	100

An in-depth examination of the prevalence of plasmid-encoded virulence-associated genes, *ial* and *sen* revealed that more strains yielded positive results for *ial* (n=45) than *sen* (n=34) (Table 3.8). More strains were detected for *ial* presence from 1998 to 2000 as compared to *sen* presence during the similar period of time. From 1997 to 1999, there were also fewer *Shigella* strains having both *ial* and *sen*.

Table 3.8
Comparison of prevalence of *ial* and *sen* with year of isolation

	1997 (%)	1998 (%)	1999 (%)	2000 (%)	Total (%)
Positive for *ial*, Negative for *sen*	1 (4.8)	14 (66.7)	5 (23.8)	1 (4.8)	21 (100)
Negative for *ial*, Positive for *sen*	1 (10.0)	9 (90.0)	0 (0)	0 (0)	10 (100)
Positive for both	5 (20.8)	10 (41.7)	0 (0)	9 (37.5)	24 (100)
Negative for both	9 (16.4)	32 (58.2)	8 (14.5)	6 (10.9)	55 (100)

Due to the unequal data distribution amongst the years, statistical analysis could not be carried out to determine if there was a significant association between prevalence of genes with year of isolation.

3.9.3 Analysis of the profiles of virulence markers (pathotypes) in Malaysian *Shigella* spp.

The results of mPCR and monoplex PCR assays were reorganized into Table 3.9, which shows the association of different combinations of virulence markers with positive PCR results. Among the 110 strains of *Shigella* spp. analyzed, eight different profiles were observed. The most frequent profile, H, was shown by 34 out of 110 strains with a single virulence-associated gene, *ipaH*. This was followed by profile E (21/110 strains) with the combination of *set1B/set1A/ipaH*. The combination of all five virulence markers (profile A) was only detected in 12 out of 110 strains. A total of 55 strains did not have the presence of either *ial* or *sen*.

Table 3.9
Pathotypes of *Shigella* spp. based on the presence of virulence-associated genes

Profile	Markers	Number of strains (%)
A	*set1B/set1A/ial/ipaH/sen*	12 (10.9)
B	*set1B/set1A/ial/ipaH*	7 (6.4)
C	*set1B/set1A/ipaH/sen*	5 (4.6)
D	*ial/ipaH/sen*	12 (10.9)
E	*set1B/set1A/ipaH*	21 (19.1)
F	*ial/ipaH*	14 (12.7)
G	*ipaH/sen*	5 (4.6)
H	*ipaH*	34 (30.9)
	Total strains	110 (100)

3.10 Clinical specimens

Ten stool specimens from diarrhoea patients from the University of Malaya Medical Centre (UMMC) were processed according to the protocol in Section 2.2.3.9. None of the *SS* agar plates had colourless colonies. All the MacConkey agar plates had pink lawns. The mPCR assay did not amplify any virulence-associated gene in the stool specimens even though both positive controls generated the targeted PCR products.