称号及び氏名	博士(獣医学)　Nityananda Chowdhury
学位授与の日付	平成２２年３月３１日
論　文　名	Genetic diversity of superintegron and its implication in the development of molecular typing tools for epidemiological studies of *Vibrio cholerae* （スーパーインテグロンの多様性とそれを利用したコレラ菌の分子疫学的解析法の開発）
論文審査委員	主査　山崎　伸二 副査　笹井　和美 副査　三宅　眞実

論文要旨

Introduction

Vibrio cholerae is a Gram-negative bacterium. It is an autochthonous inhabitant of aquatic environments and can cause cholera in human. Cholera is a dreadful diarrheal disease killing millions of people since ancient time and is characterized by rice-water stool, vomiting etc. Even having more than 200 serogroups of *V. cholerae*, only two of them O1 and O139 are so far known to be responsible for epidemics of cholera. The O1 serogroup has two biotypes, El Tor and classical. Previous six pandemics (1817-1923) were presumably caused by classical biotype which is now extinct, whereas El Tor biotype is responsible for the ongoing 7th pandemic starting from 1961. *V. cholerae* O139 synonym Bengal emerged in 1992 and was thought as an initiator of the 8th pandemic. The non-O1/non-O139 strains are noticed to cause sporadic cases of gastroenteritis.

To trace the epidemic strains rapidly and appropriately, suitable molecular typing tools for *V. cholerae* are essential. There are several molecular typing methods like ribotyping, pulsed-field gel electrophoresis (PFGE), multilocus sequence typing (MLST) etc. These methods have good discriminative capacity but

are time consuming, complicated and expensive. On the contrary, PCR-based typing methods like amplified fragment length polymorphism (AFLP), restriction fragment length polymorphism (RFLP), etc. are rapid, simple and also cost-effective.

Several genetic islands which are present in the genome of *V. cholerae* play vital role in its evolutionary fitness. One of them is a distinctive class of integron which is also known as superintegron (SI). The integron is a genetic system which can capture exogenous DNA and aid in their expression in bacteria. It possesses three key elements such as (i) a gene encoding for an integrase (*intI*), (ii) a primary recombination site (*attI*) and (iii) a strong promoter. In a *V. cholerae* O1 El Tor strain (N16961) whose whole genome sequence is now available; the SI is ca 126 kb in size. It is composed of a class 4 integrase gene (*intI*4) and 179 gene cassettes which constitutes 3% of the whole genome. Each of the gene cassettes is followed by short (123 to126 bp) repeat sequence, called *V. cholerae* repeat (VCR). SI is considered as a flexible platform where new gene cassettes can be inserted or deleted. This flexible nature in gene cassette capturing may result in SI diversity. However, little is known about the genetic diversity of SI and its implication in the pathophysiology of *V. cholerae*.

In this study, SI in various strains of *V. cholerae* from both clinical and environmental origins, having differences in serogroup, biotype, geographic background and isolation time, were analyzed to understand the genetic diversity of SI, and to develop simple and rapid PCR-fingerprinting methods for *V. cholerae* on the basis of SI diversity. The PCR-based methods were also evaluated by comparing with PFGE. Moreover, the effect of repeated *in vitro* subculturing of *V. cholerae* strains on the stability of SI and PFGE profiles was monitored to evaluate the reliability or usefulness of the PCR-fingerprinting methods in comparison to PFGE.

Chapter 1. Genetic diversity of superintegron among various *V. cholerae* strains

To check the genetic diversity of SI, a total of 16 *V. cholerae* strains including El Tor (n=7), classical (n=5), O139 (n=2) and non-O1/non-O139 (n=2) isolated from different geographic origin, source and time were selected. Before proceeding for diversity analysis, Southern hybridization of *BglI*-digested genomic DNA of these selected *V. cholerae* strains was carried out using ^{32}P-labelled *intI*4 specific gene probe. Hybridization results confirmed the existence of *intI*4 gene among all the strains tested. This indicates that the test strains may possess SI island in their genome. From the genome sequence of strain N16961 (O1, El Tor biotype), total SI island (ca 126 kb) was extracted, divided into 12 segments and overlapping LA (long and accurate)-PCR primers were designed to scan the whole SI island. All the El Tor and O139 strains produced amplicons in respective to the primers of most of the segments but amplicons were either identical or different in size in comparison to the control strain N16961. However, primers of most of the SI segments did not

work for *V. cholerae* classical and non-O1/non-O139 strains. This might be due to the presence of extensive diverse SI island in these strains. PCR amplicons of identical sizes were checked for RFLP using suitable restriction enzymes (*Bgl*I, *Hin*dIII, *Mlu*I, or *Sau3A*I). PCR amplicons which were different in size were sequenced and analyzed. After careful analysis, it was revealed that the possible diversity of SI segments were due to the deletion, insertion or rearrangement of gene cassettes. Possible mechanisms behind gene cassette deletion were homologous recombination between (i) *attI* & VCR, (ii) VCR & VCR or (iii) homologous ORF & ORF. Genomic diversity of these strains revealed by PFGE, were well correlated with SI diversity among El Tor and O139 strains. It might be also true for classical and non-O1/non-O139 strains, but amplification of more SI segments are required for better comparison.

PCR scanning, RFLP or sequencing of SI segments revealed that strains having differences in serogroup, biotype, geographic origin and isolation time had positive correlation with the diversity of their SI. These findings raised the possibility to implement the diversity of SI in the development of molecular typing methods for *V. cholerae*. On the other hand, the gene rearrangement and the absence or presence of same gene cassettes in SI island among different *V. cholerae* strains raised the question about the stability of SI island as well.

Chapter 2. Development of simple and rapid PCR-fingerprinting methods (PCR-RFLP & VCR-AFLP) for *Vibrio cholerae*

A total of 177 *V. cholerae* including El Tor (n=94), O139 (n=54) and classical (n=29) strains isolated between 1948 and 2006 were used in this study. For PCR-RFLP, comparative ORFs from SI sequences of N16961 (El Tor) and O395 (classical strain whose genome sequence is also available) were analyzed and based on this analysis, forward and reverse primers were designed from *intI*4 gene and a common ORF, respectively. Then PCR products were digested with *Bgl*I and RFLP patterns were analyzed. Out of 94 El Tor strains, 86 strains could generate expected amplicons. In case of O139 and classical strains, 39 out of 54 and 25 out of 29, respectively, produced PCR amplicons. PCR-RFLP with *Bgl*I was able to generate 8 different fingerprints for El Tor strains, designated arbitrarily as ER1 (**E**l Tor **R**FLP type **1**) to ER8. For O139 strains, total 5 different fingerprints BR1 (**B**engal **R**FLP type **1**) to BR5, were observed and 2 different RFLP types CR1 (**C**lassical **R**FLP type **1**) and CR2 were produced by classical strains. In all cases, the strains that did not produce any amplicon in LA-PCR were considered to produce at least a single but different RFLP pattern. As a result, a total of 9 different RFLP patterns for El Tor, 6 for O139 and 3 for classical strains were obtained.

For the development of VCR-AFLP, a total of 309 VCR sequences including 178 from the SI sequence of N16961 (O1, El Tor) and 131 from that of O395 (O1, classical) were aligned and two outward oriented primers were designed. All the

strains used for PCR-RFLP analysis, were also tested by VCR-AFLP assay. Five different fingerprints were obtained from 94 El Tor strains. These patterns were designated arbitrarily as EA1 (**E**l Tor **A**FLP type **1**) to EA5. Fifty four O139 strains produced 2 different patterns, BA1 (**B**engal **A**FLP type **1**) and BA2, whereas 29 classical strains were differentiated into 9 patterns, CA1 (**C**lassical **A**FLP type **1**) to CA9.

PCR-RFLP and VCR-AFLP data obtained from each strain were overlapped and merged into a single fingerprint identity to gain better resolution of molecular epidemiological analysis for *V. cholerae* strains. For example, an El Tor strain that produced RFLP pattern ER1 and AFLP pattern EA1 were considered as a single fingerprint and it was denoted as e-I. In this system, all the El Tor strains (n=94) were found to generate a total of 11 fingerprints which were denoted as e-I to e-XI. Similarly, the O139 strains (n=54) produced 7 fingerprints, b-I to b-VII, whereas the classical strains (n=29) produced 10 different fingerprints, c-I to c-X.

To check the feasibility of PCR-RFLP and VCR-AFLP as molecular typing tools, the strains (n=177) analyzed by PCR-based methods were also analyzed by PFGE and comparative studies were done. Out of 94 El Tor strains, PFGE could generate 11 different types, EP1 (**E**l Tor **P**FGE type **1**) to EP11 and 14 subtypes. In case of the O139 strains (n=54), only 2 different PFGE types, BP1 (**B**engal **P**FGE type **1**) and BP2 could be differentiated, and BP1 type had 19 subtypes. PFGE could distinguish 29 classical strains into 5 different fingerprints, CP1 (**C**lassical **P**FGE type **1**) to CP5 and 13 subtypes. Considering Tenover's criteria, the PFGE subtypes in El Tor, O139 and classical strains might be merged with their respective parental type. Thus, only 11, 2 and 5 PFGE types could be considered for El Tor, O139 and classical strains, respectively.

In a comparative analysis between PCR-based methods and PFGE, the combined PCR-based assays produced 11 types which were equal to the number of PFGE types (11 types) for El Tor strains. In case of O139 and classical strains, the PCR-based methods produced comparatively more fingerprints (7 and 10 types, respectively) than those of PFGE (2 and 5 types, respectively). Moreover, PCR-based methods were effective to distinguish some PFGE types or subtypes into 4 to 6 groups. Conversely, some of the PCR-fingerprints in El Tor, O139 and classical strains were discriminated more by PFGE. Comparative analysis indicated that PCR-based fingerprints data were well correlated with those of PFGE.

Chapter 3. Effect of repeated subculturing of *V. cholerae* strains on the stability of superintegron (SI) island: evaluation of the SI-based PCR-fingerprinting methods

The SI island is sedentary and not associated with mobile genetic elements. However, rearrangement, absence or presence of same gene cassettes in SI island among different *V. cholerae* strains have raised the possibility of possessing mobile

gene cassettes that may influence the fingerprinting profiles. In case of Shiga toxin-producing *Escherichia coli* (STEC) the variations in PFGE patterns of the same clone have been reported due to *in vivo* or *in vitro* passage. There was no such report about *V. cholerae* previously. Therefore, the effect of repeated *in vitro* subculturing of *V. cholerae* strains on the stability of the SI and change in PFGE profile (if any) was examined to evaluate the usefulness or reliability of the newly developed PCR-fingerprinting methods.

Every 3- to 4-day interval repeated subculturing of 3 *V. cholerae* strains N16961 (representative of O1 El Tor), O395 (representative of O1 classical), and AP32549 (representative of O139) up to 25-week (total 50 times subculturing) was carried out. Every 5th week interval, 14 colonies from a single strain were randomly analyzed by PFGE as well as by newly developed PCR-RFLP and VCR-AFLP assays. In El Tor strain N16961, in addition to the original ep-Ia (**E**l Tor **P**FGE type-**Ia**) profile, another 3 different profiles (ep-Ib to ep-Id) were observed. The PFGE profiles differed from each other by 1 to 3 DNA fragments. The original PFGE profile in classical strain O395 was denoted as cp-Ia (**C**lassical **P**FGE type-**Ia**). Like El Tor strain, the classical strain also produced 3 more PFGE profiles (cp-Ib to cp-Id). The PFGE profiles of classical strain (cp-Ia to cp-Id) differed from each other by 1 to 6 DNA fragments. In case of O139 strain AP32549, the change in PFGE profile was also observed but it produced 2 more PFGE profiles (bp-Ib and bp-Ic) in addition to the original bp-Ia (**B**engal **P**FGE type-**Ia**). The PFGE profiles in O139 strain were shown to differ from each other by 1 or 2 DNA fragments.

In contrast to PFGE analysis, there was no change in PCR-RFLP fingerprints and thus the clones of *V. cholerae* always produced the original fingerprints er-I (**E**l Tor **R**FLP type-**I**), cr-I (**C**lassical **R**FLP type-**I**), and br-I (**B**engal **R**FLP type-**I**) by El Tor, classical and O139 strains, respectively. Similarly, VCR-AFLP assay always produced the original fingerprints ea-I (**E**l Tor **A**FLP type-**I**), ca-I (**C**lassical **A**FLP type-**I**), and ba-I (**B**engal **A**FLP type-**I**) by El Tor, classical and O139 strains, respectively. There was no visible change in fingerprints due to the repeated subculturing of *V. cholerae* strains regardless of their serogroups or biotypes and PFGE profiles. The combination of both the PCR-assays is considered to count the diversity of total SI island. Therefore, no change in PCR-fingerprints by PCR-RFLP and VCR-AFLP indicates the stability of SI in a single strain. It also indicates that the change in PFGE profiles may be due to the instability of the part of *V. cholerae* genome other than SI island.

Conclusions

1. In *Vibrio cholerae*, there is significant genetic plasticity in the structure of superintegron (SI) owing to the loss, acquisition and rearrangements of gene cassettes. Therefore, SI island can be targeted for the development of molecular typing tools for *V. cholerae*.

2. PCR-fingerprinting methods (PCR-RFLP & VCR-AFLP) developed on the basis of SI diversity: (i) have good discriminating power and well correlations with PFGE, (ii) are simple, rapid, and cost-effective in comparison to other complicated typing methods such as PFGE, MLST etc. Therefore, newly developed PCR-fingerprinting methods may be applicable for molecular epidemiological studies of *V. cholerae*.
3. The change in PFGE profiles but not in PCR-RFLP or VCR-AFLP fingerprints due to the repeated subculturing of *V. cholerae* strains indicates that the SI island in a single strain is stable under laboratory conditions. Therefore, PCR-RFLP and VCR-AFLP assays developed on the basis of genetic diversity of the SI can be more practical and reliable than PFGE, when analyzing the genetic lineage of *V. cholerae* strains in an etiological manner.

審査結果の要旨

コレラ菌は、汽水域に生息するグラム陰性桿菌であり、ヒトにコレラを引き起こす。コレラは古代から多くの人々の命を奪ってきた非常に恐れられている病気である。現在までに 200 種類以上の O 群血清型が報告されているが、コレラの原因となるのはコレラ毒素を産生する 01 と 0139 の O 群血清型のみであると考えられていた。01 血清型にはさらにエルトール型と古典型の 2 つの生物型に分けられる。第１次から第６次の世界流行では古典型 01 コレラ菌（古典型 01）が、1961 年から始まり今日尚続いている第７次ではエルトール型 01 コレラ菌（エルトール型 01）が関わっている。1992 年に突如として新型コレラ菌 0139（0139）が出現し、コレラの大流行が引き起こされ、第 8 次のコレラの世界流行かと思われた。一方、01 と 0139 以外のコレラ菌は non-01、non-0139 コレラ菌（non-01/019）と呼ばれ、散発性下痢症の原因菌として問題となっている。

流行型の菌を迅速に検出するためには分子型別が重要である。リボタイピング、PFGE や MLST 等の分子型別法が開発され、優れた解像度を有している。しかしながら、手間、時間やコストがかかること等の問題点がある。一方、PCR 法に基づく分子型別法は簡便、迅速で比較的低コスト等の利点がある。コレラ菌は、スーパインテグロンと呼ばれる全ゲノムの 3%に相当する約 126 kb の遺伝子領域を普遍的に保持している。インテグロンは、インテグラーゼ、*attI* サイトと呼ばれる特定の遺伝子カセットを取り込む領域と、遺伝子カセットの転写に関わるプロモータの３つからなっており、インテグラーゼ近傍の *attI* サイトに次々に遺伝子を取り込むことから、遺伝子カセットの年輪の様な役割を果たしている。エルトール型 01 の代表菌株である N16961 株の染色体中 126 kb の領域に 179 個の遺伝子カセットと遺伝子カセットの間に *Vibrio cholera* repeat（VCR）と呼ばれる多様性のある繰り

返し配列を有している。本研究では、スーパーインテグロンの多様性を解析し、コレラ菌スーパーインテグロンを標的とした PCR 法に基づく簡便で迅速な分子型別法を開発した。さらに開発した PCR 法の解像度と再現性について PFGE と比較した。

第 1 章では、様々な年代と様々な国から分離されたエルトール型01、古典型01、non-01/0139 を用いて、スーパーインテグロン領域を 12 等分したPCRスキャニング法でその多様性の解析を行った。その結果、増幅産物の分子量に差が認められ、また増幅産物が得られない株や領域があり、コレラ菌のスーパーインテグロンにはかなりな多様性があることが明らかになった。分子量の異なるPCR産物の塩基配列を解析した結果、VCRや同じORFの相同組み換えによる遺伝子の挿入や脱落がこの領域の多様性に関与していることを明らかとした。

第２章では、第 1 章で得られた塩基配列の解析結果やGenBankに登録されているコレラ菌スーパーインテグロンの塩基配列情報を基に、コレラ菌スーパーインテグロンを標的としたPCR-RFLP法及びVCRを標的としたPCR-AFLP法を実施するために複数のPCRプライマーを設計し、エルトール型01、94 株、0139、54 株、古典型01、29 株、合計 177 株のコレラ菌を用いて評価した。最も解像度の高かったPCRプライマーを用いた結果と 177 株のPFGEの結果を比較したところ、PFGEではエルトール型01 を 11 タイプに、0139 を 2 タイプに、古典型01 を 5 タイプにしか分類できなかった。一方、PCR-RFLP法とPCR-AFLP法の組み合わせでは、エルトール型01 で 11 タイプ、0139 で 7 タイプ、古典型01 で 10 タイプと、PFGEよりも高い解像度で分子型別可能となることを明らかとした。

第３章では、同一のクローンであっても実験室内で継代する際にPFGEのPFLPパターンが変化する現象がコレラ菌でも観察されるかどうか、また、本PCR法がその現象により影響を受けるかどうかについて解析した。エルトール型01、0139 及び古典型01 の代表株をそれぞれ用いて解析した結果、PFGEでは継代に伴うPFGEパターンの変化が見られたが、PCR-RFLP法やPCR-AFLP法では菌の継代によってパターンが変化することなく、本法はコレラ菌の分子疫学的解析法として有用であることを明らかとした。

以上の結果は、コレラ菌スーパーインテグロンに多様性のあること、またコレラ菌スーパインテグロンがPCR法による分子型別法の標的となりうることを示した。さらに、スーパーインテグロンを標的としたPCR-RFLP法及びPCR-AFLP法を開発し、PFGEに勝るとも劣らない解像度と再現性を有することを明らかとした。これらの成果は、コレラ菌のスーパーインテグロンに多様性があるという新たな知見のみでなく、コレラ菌の新たな分子疫学的解析法の開発としても期待できるものであり、獣医学の分野のみならず医学領域においても多大な貢献をすると考えられる。従って、最終試験の結果と併せて、博士（獣医学）の学位を授与することを適当と認める。