Journal of General Virology (2010), **91**, 1164–1172 DOI 10.1099/vir.0.018028-0

Evaluation of white spot syndrome virus variable DNA loci as molecular markers of virus spread at intermediate spatiotemporal scales

Bui Thi Minh Dieu,[1,2] Hendrik Marks,[1]† Mark P. Zwart[3] and Just M. Vlak[1]

Correspondence
Just M. Vlak
just.vlak@wur.nl

[1]Laboratory of Virology, Wageningen University, Radix building (#107), Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands

[2]Biotechnological Research and Development Institute, Can Tho University, 3/2 Street No. 1, 0084710 Can Tho City, Vietnam

[3]Quantitative Veterinary Epidemiology, Marijkeweg 40, 6709 PG Wageningen, The Netherlands

Variable genomic loci have been employed in a number of molecular epidemiology studies of white spot syndrome virus (WSSV), but it is unknown which loci are suitable molecular markers for determining WSSV spread on different spatiotemporal scales. Although previous work suggests that multiple introductions of WSSV occurred in central Vietnam, it is largely uncertain how WSSV was introduced and subsequently spread. Here, we evaluate five variable WSSV DNA loci as markers of virus spread on an intermediate (i.e. regional) scale, and develop a detailed and statistically supported model for the spread of WSSV. The genotypes of 17 WSSV isolates from along the coast of Vietnam – nine of which were newly characterized in this study – were analysed to obtain sufficient samples on an intermediate scale and to allow statistical analysis. Only the ORF23/24 variable region is an appropriate marker on this scale, as geographically proximate isolates show similar deletion sizes. The ORF14/15 variable region and variable-number tandem repeat (VNTR) loci are not useful as markers on this scale. ORF14/15 may be suitable for studying larger spatiotemporal scales, whereas VNTR loci are probably suitable for smaller scales. For ORF23/24, there is a clear pattern in the spatial distribution of WSSV: the smallest genomic deletions are found in central Vietnam, and larger deletions are found in the south and the north. WSSV genomic deletions tend to increase over time with virus spread in cultured shrimp, and our data are therefore congruent with the hypothesis that WSSV was introduced in central Vietnam and then radiated out.

Received 29 October 2009
Accepted 19 January 2010

INTRODUCTION

White spot syndrome virus (WSSV), the major cause of large economic losses in shrimp farming, is a large, double-stranded DNA (dsDNA) virus belonging to the family *Nimaviridae*, genus *Whispovirus* (Vlak *et al.*, 2005). WSSV was first discovered in 1992 in Taiwan and China, and the virus subsequently spread quickly to most countries in South-East Asia, the Indian subcontinent, and North and South America. In addition to shrimp, WSSV can infect a broad range of crustaceans, including crabs and crayfish. This broad host range is thought to be a major cause of the rapid and extensive spread of WSSV (Flegel, 1997).

Molecular methods for genotyping WSSV isolates are powerful tools for understanding virus spread and epidemiology (Dieu *et al.*, 2004; Marks *et al.*, 2004; Pradeep *et al.*, 2008a, b). Initial studies using molecular methods to compare WSSV isolates suggested that genetic differences between various isolates were small, as these studies used insensitive techniques such as restriction fragment-length polymorphism (RFLP) (Lo *et al.*, 1996a, b, 1999; Nadala & Loh, 1998; Wang *et al.*, 2002; Moon *et al.*, 2003). However, differences between WSSV isolates could be detected readily by using more sensitive methods, such as PCR. Lan *et al.* (2002), for example, found host-dependent differences among WSSV isolates with a PCR-based method. Marks *et al.* (2004) aligned three completely sequenced WSSV isolates, originating from Taiwan (WSSV-TW) (Wang *et al.*, 1995), China (WSSV-CN) (Yang *et al.*, 2001) and Thailand (WSSV-TH) (van Hulten *et al.*, 2001). Although the overall nucleotide identity was >99 %, five variable loci were identified, consisting of two

†Present address: Department of Molecular Biology, Faculty of Science, Nijmegen Centre for Molecular Life Sciences, Radboud University Nijmegen, 6500 HB Nijmegen, The Netherlands.

A supplementary table showing primers used in PCR analysis for the variable loci of WSSV is available with the online version of this paper.

regions with genomic deletions (ORF23/24 and ORF14/15 variable regions) and three loci with a variable number of tandem repeats (VNTR) (ORF75, ORF94 and ORF125) (Marks *et al.*, 2004).

A number of subsequent studies on WSSV epidemiology have used the regions with genomic deletions (Musthaq *et al.*, 2006; Waikhom *et al.*, 2006; Pradeep *et al.*, 2008b) or one or more VNTR loci (Hoa *et al.*, 2005; Kiatpathomchai *et al.*, 2005; Kang & Lu, 2007; Pradeep *et al.*, 2008a; Tan *et al.*, 2009), or both (Dieu *et al.*, 2004; Marks *et al.*, 2005; Pradeep *et al.*, 2009), as genetic markers to characterize WSSV variants. VNTRs appear to be more variable than the deletions (Dieu *et al.*, 2004). High degrees of polymorphism for VNTR-like loci have been reported in various other large dsDNA viruses, such as cytomegaloviruses (Davis *et al.*, 1999). This suggests that, whilst VNTRs may be useful for studying WSSV spread on small spatiotemporal scales, genomic deletions are more suitable for studying spread on intermediate and large scales. We use the following terms to describe different scales on which WSSV has spread: (i) very small: spread between ponds and farms (10 km); (ii) small: spread between clusters of shrimp farms and villages (100 km); (iii) intermediate: regional spread within and between countries (1000 km); (iv) large: continental and global spread (10 000 km). A systematic comparison of the two approaches – using VNTRs or genomic deletions – to studying spread on these different spatial scales has not been reported.

The presence of WSSV in Vietnam was first confirmed by PCR analyses on samples collected in 1997 (Corsin *et al.*, 2001), but it is not clear how many times the virus was introduced from abroad. Also, it is unclear where WSSV originated and how it subsequently spread to other regions in Vietnam from the original introduction site(s). Epidemiological studies have been faced with numerous design and execution problems, making it difficult to infer WSSV spread based on farmer reports of shrimp health and screening ponds for WSSV (Corsin *et al.*, 2002). Finding suitable methods to identify and discriminate WSSV strains – and infer their origin – is therefore important for WSSV forensics and epidemiology. We previously reported a preliminary study of genomic variation in central Vietnam, based on the mapping of deletions and VNTRs of eight WSSV isolates. Vietnam is an ideal location to study the spatiotemporal spread and evolution of WSSV because of the relatively late introduction of large-scale shrimp culture, the moderate socio-economic development around the turn of the millennium, small-size farming operations and an accurate WSSV-reporting system. Our previous results suggested that WSSV originated from a common ancestor – reported in Taiwan – and subsequently spread to Vietnam through multiple introductions (Dieu *et al.*, 2004). However, WSSV isolates from important shrimp-production regions in northern and southern Vietnam were not available during the previous study. Isolates originating from these regions are important to develop a statistically supported model of the introduction and spread of WSSV in Vietnam. This analysis is now even more relevant as WSSV has become more virulent over time, which could be attributed to the above genomic mutations (Marks *et al.*, 2005).

Here, we study genomic variation in WSSV isolates obtained from all important shrimp-production regions in Vietnam. We performed genomic analysis on two northern, one central and six southern WSSV isolates. For our analysis, we used the five variable loci described by Marks *et al.* (2004). These loci were characterized for each Vietnamese (VN) isolate by PCR amplification, cloning and sequencing. We could further test and validate our previous hypothesis on the spread of WSSV in Vietnam (Dieu *et al.*, 2004). Finally, we could, for the first time, systematically evaluate the utility of each of the variable regions as genetic markers for studying WSSV spread and epidemiology on an intermediate spatiotemporal scale.

RESULTS

Description of WSSV isolates

WSSV-infected *Penaeus monodon* shrimp were collected from nine shrimp-culture ponds in Vietnam in 2004 and subsequently genotyped (Table 1). The ponds from which samples were collected were distributed over nine different provinces, covering 2500 km of the Vietnamese coast (Fig. 1). Juvenile shrimp were selected for WSSV testing if (i) there was a WSSV outbreak in the pond or (ii) the shrimp showed reduced feeding. All chosen shrimp tested positive for the presence of WSSV using a single-step PCR for *VP26*. One WSSV-infected shrimp from each pond was chosen randomly for further analysis, and assumed to be representative of that pond and region.

Variable region ORF23/24

In order to map the ORF23/24 locus, we first performed a PCR with the VR23/24-screen primers on all samples (Supplementary Table S1, available in JGV Online). These primers were used previously to detect deletions in six WSSV isolates from central Vietnam (Dieu *et al.*, 2004). Only isolate H – from central Vietnam – tested positive, rendering a 548 bp amplicon (Fig. 2). Cloning and sequencing of this PCR fragment indicated that isolate H was identical to the other VN-central WSSV isolates, with a deletion of about 8539 bp compared with WSSV-TW. New primers were developed to genotype variable region ORF23/24 for the other isolates from across Vietnam. VN isolate ND gave a product of approximately 3.8 kb with primer set VR23/24-ND; HP gave a product of approximately 850 bp with primer set VR23/24-HP. Both PCR products were sequenced to reveal the exact identity of the deletions up to the nucleotide level (Fig. 2). Overall, the data indicate increasing deletion size of the isolates from central Vietnam relative to the ND and HP isolates from the north (Fig. 2).

Table 1. Origins of the Vietnamese WSSV isolates used in this study

Region/pond	Place (district)	Province	Origin of post larvae	Date of collection	Abbreviation
North Vietnam					
Hai Phong	Do Son	Hai Phong	Central region	14 Aug 2004	HP
Nam Dinh	Nam Dinh	Nam Dinh	Central region	01 Sep 2004	ND
Central Vietnam					
Hue	Phu Vang	Hue	Central region	28 Jun 2004	H
South Vietnam					
Ba Ria	Xuyen Moc	Ba Ria	Unknown	22 Feb 2004	BR
Tra Vinh b	Duyen Hai	Tra Vinh	Local	10 Feb 2004	Tv-b
Soc Trang	My Xuyen	Soc Trang	Unknown	05 Mar 2004	ST
Bac Lieu	Vinh Loi	Bac Lieu	Unknown	15 Feb 2004	BL
Ca Mau	Tan Thanh	Ca Mau	Local	20 Feb 2004	CM
Ha Tien	Thuan Yen	Kien Giang	Central region	10 Feb 2004	HT

Isolates from southern Vietnam produced unique PCR amplicons with a different set of primers (VR23/24-south), indicating a larger deletion. ST, BL, CM and HT all gave the same PCR product of approximately 400 bp, which, after sequencing, indicated that they contained the same deletion of 11 866 bp compared with the WSSV-TW sequence. Isolate BR produced an approximately 400 bp amplicon with this primer set. Sequencing indicated a 13 210 bp deletion, identical to WSSV-TH (Dieu *et al.*, 2004). PCR on the Tv-b isolate using the same primers resulted in an approximately 1.6 kb amplicon, similar in size to Tv, a previously analysed isolate but from a different district in Tra Vinh province. Restriction enzyme analysis of the PCR products confirmed that the amplified sequences are the same (data not shown). Together, the data indicated that WSSV isolates show an increasing deletion size from central to southern Vietnam (Fig. 2).

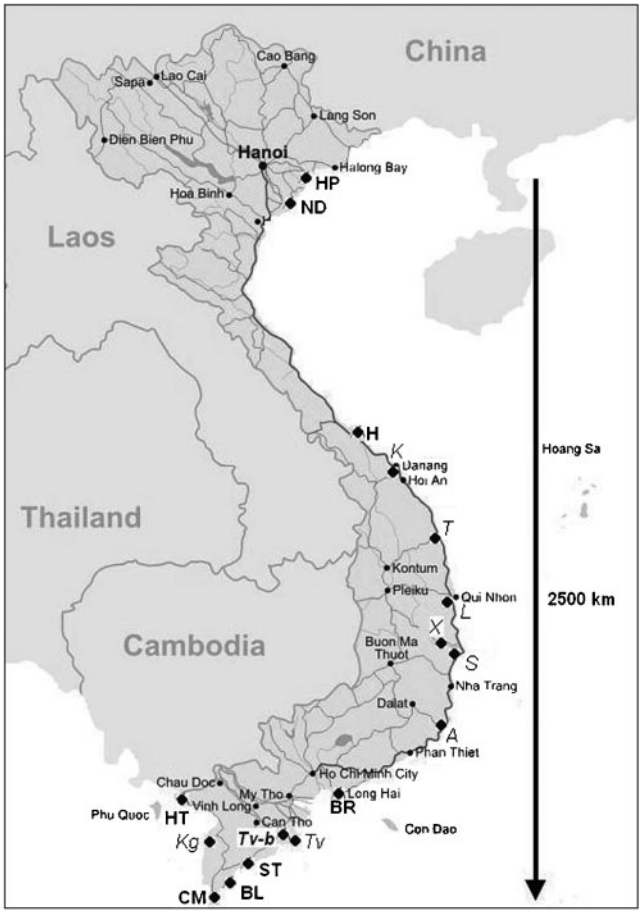

Fig. 1. Map of Vietnam, showing the geographical origins of the isolates used for this study, indicated by HP, ND, H, BR, Tv-b, ST, BL, CM and HT (in bold), according to Table 1. Isolates K, T, L, X, S, A, Tv and Kg, used in our previous study (Dieu *et al.*, 2004), are also in indicated (in italics).

Variable region ORF14/15

The TH-96-II isolate has an additional 6436 bp segment region in the ORF14/15 variable region compared with all other known WSSV isolates, and Marks *et al.* (2005) suggested that this genotype is representative of the common ancestor of WSSV in South-East Asia. The coordinates and size of the genomic deletions in ORF14/15 were determined for the new Vietnamese isolates using a similar approach as was used for the ORF23/24 variable region. TH-96-II was used as a reference sequence for determining the size of the deletion. A PCR with the VR14/15-screen primers (Supplementary Table S1; Fig. 3a) was performed. Almost all of the new WSSV-VN isolates (ND, HP, H, ST, BL, CM and Tv-b) showed an approximately 500 bp amplicon, similar in length to that reported previously for isolate K (Fig. 3b). Restriction enzyme analysis of the PCR products confirmed that these isolates have the same 6031 bp deletion as most WSSV-VN isolates analysed previously (Fig. 3a, c; Dieu *et al.*, 2004). However, the HT and BR isolates failed to give a PCR product with this primer set. Using the VR14/15-1 primers, an approximately 900 bp product was obtained for the HT isolate, whilst an approximately 750 bp product was obtained for the BR isolate (Supplementary Table S1). Cloning and sequencing of these PCR products showed

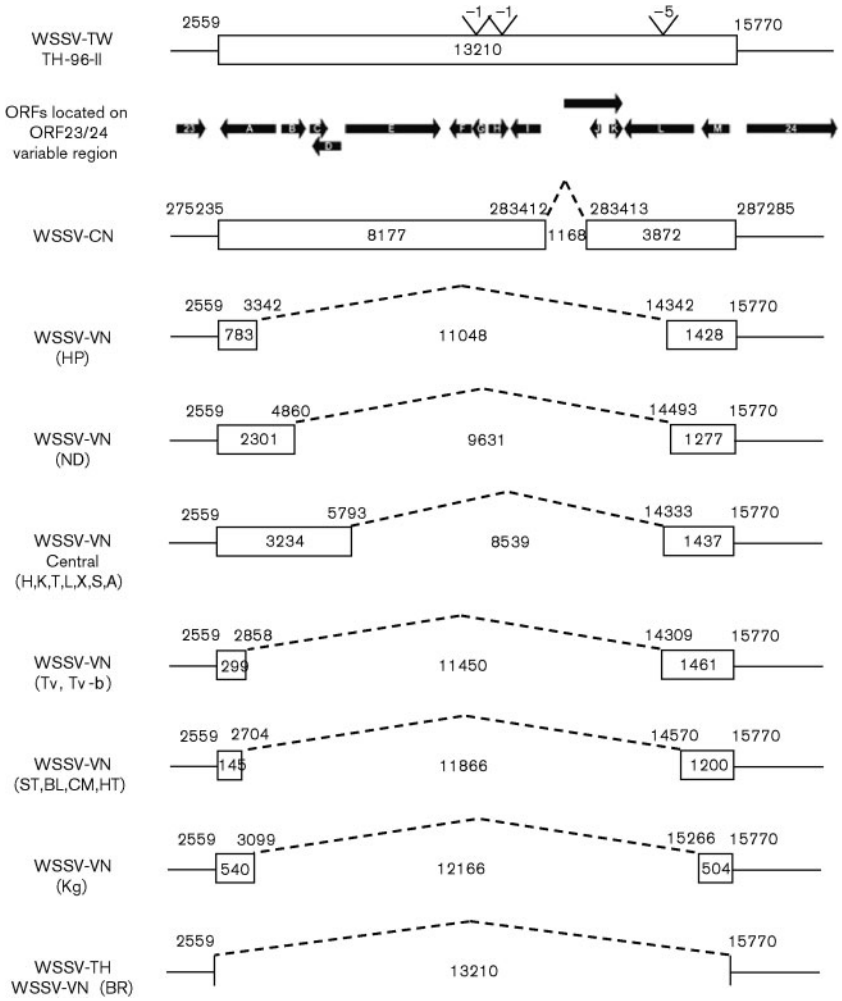

Fig. 2. Schematic representation of the ORF23/24 variable region of WSSV-TW, WSSV-TH-96-II, WSSV-CN, WSSV-TH and the VN isolates. The map numbers, indicated above each isolate, are in accordance with the numbers in GenBank for the genomic sequence of each isolate. The coordinates of the WSSV-VN isolates are according to the WSSV-TW annotation. The length of the fragments is indicated within boxes or sequences. The positions of the ORFs located in this region are indicated by closed arrows, which also represent the direction of transcription. ORFs are numbered in accordance with the numbering used by Marks *et al.* (2004) and Dieu *et al.* (2004).

that isolate HT had a genotype similar to that of WSSV-TW, whilst BR was similar to WSSV-TH (Fig. 3a).

VNTR loci ORF75, ORF94 and ORF125

The known WSSV VNTR loci were also analysed. ORF75 possesses two types of repeat units (RUs) with lengths of 102 and 45 bp. ORF94 and ORF125 have RUs with a single type of repeat sequence of 54 and 69 bp, respectively. These loci were analysed for the nine WSSV-VN isolates (Table 2) by PCR and sequencing. The number of RUs present in ORF94 ranges from four to 17, whereas those in ORF125 range from four to ten. The RU unit of 102 bp appeared to be present at a frequency ranging from one to four, whereas the 45 bp repeat is present between five and 14 times (Table 3). This variation in RUs is very similar to what has been found previously in other Vietnamese WSSV isolates (Dieu *et al.*, 2004; Hoa *et al.*, 2005). Overall, however, there do not seem to be any trends in these data (Table 2).

Statistical analysis

To test which loci are suitable as genetic markers for inferring WSSV spread on an intermediate scale, a runs test was performed on the number of repeat units (ORF75, ORF94 and ORF124) or the size of genomic deletion (ORF23/24 and ORF14/15). The ORF23/24 variable region was the only one giving a significant outcome for this test (Table 3). A significantly lower number of runs than expected by chance was found, indicating that geographically proximate isolates were related for this locus. For ORF14/15, ORF75, ORF94 and ORF125, the *P*-value was >0.010 and hence there was no evidence that deletion size (ORF14/15) and RU number (ORF75, ORF94, ORF125) were spatially structured.

We then performed a Jonckheere–Terpstra test to test for increasing or decreasing median trait value if the WSSV isolates from Vietnam were ordered sequentially from north to south. This test indicated that there was a significant increase in median deletion size in ORF23/24 if the samples were ordered along the Vietnamese coast from north to south (Table 3). For all other loci, there were no significant increases or decreases found with this test. The results from the runs test and Jonckheere–Terpstra test are therefore congruent: (i) the runs test indicates that only the ORF23/24 variable region is a suitable marker on an intermediate scale, and (ii) the Jonckheere–Terpstra test

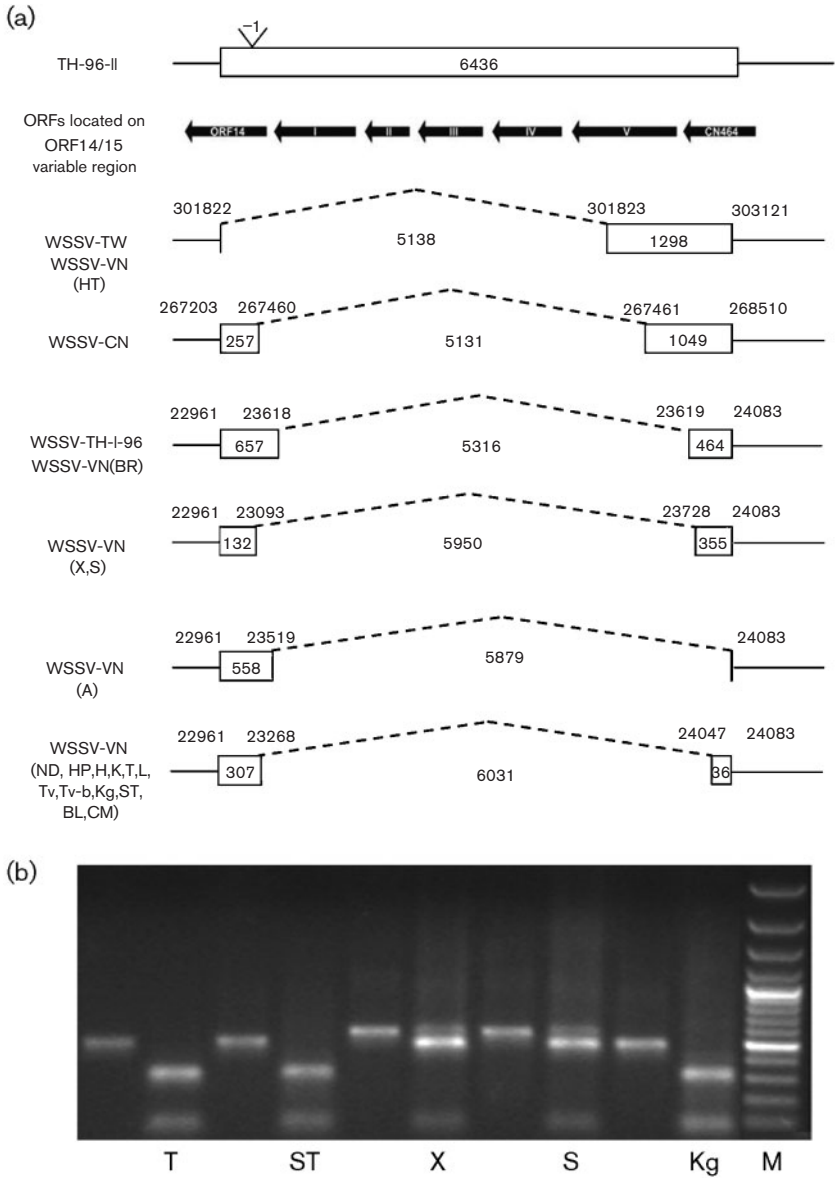

Fig. 3. (a) Schematic representation of the ORF14/15 variable region of the WSSV putative common ancestor (WSSV-TH-96-II), WSSV-TW, WSSV-CN, WSSV-TH and the VN isolates. The map numbers, indicated above each isolate, are in accordance with the numbers in GenBank for the genomic sequence of each isolate. The length of the fragments is indicated within boxes or sequences. (b) Restriction enzyme analysis using *Nco*I and *Pst*I of PCR products of WSSV-VN isolates with primer set VR14/15-screen (Supplementary Table S1).

indicates that there is a significant trend in trait value, but only for this locus.

DISCUSSION

The development of genetic markers for identifying and distinguishing WSSV isolates from different geographical origins is the first step in reconstructing, monitoring and ultimately controlling the spread of WSSV infection in shrimp farming. In a previous study, we showed that two variable regions with large deletions (ORF23/24 and ORF14/15) and three VNTR loci (ORF75, ORF94 and ORF125) could be used as molecular markers for epidemiological studies (Dieu *et al.*, 2004). Others have also employed one or more of these loci as markers to characterize WSSV isolates (Hoa *et al.*, 2005; Kiatpathomchai *et al.*, 2005; Marks *et al.*, 2005; Musthaq *et al.*, 2006; Waikhom *et al.*, 2006; Kang & Lu, 2007; Pradeep *et al.*, 2008a, b, 2009; Tan *et al.*, 2009), but a statistically supported model was not derived. Moreover, the degree of between-isolate variation for these loci appears to be very different. Therefore, a suitable WSSV marker locus should be selected, depending on the spatiotemporal scale to be considered in a study – be it population structure in and around shrimp ponds, or global virus spread.

We found that the ORF23/24 variable region was a suitable marker locus for determining patterns of WSSV spread at an intermediate spatiotemporal scale. Geographically proximate isolates had similarly sized deletions in this genomic region, and there was a clear overall spatial pattern: deletion size tended to become larger along the coast of Vietnam, from north to south (Table 3). On the other hand, the other WSSV variable loci were not suitable

Table 2. Number of RUs present within the non-homologous region unidirectional repeats of ORF75, ORF94 and ORF125 (van Hulten *et al.*, 2001)

WSSV isolates genotyped within the context of the current study are in bold; those genotyped by Dieu *et al.* (2004) are in plain text.

Region/WSSV isolate	No. RUs present in:		
	ORF75*	ORF94†	ORF125‡
North Vietnam			
HP	12 (9 and 3)	9	10
ND	7 (6 and 1)	4	§
Central Vietnam			
H	5 (3 and 2)	12	5
K	5 (3 and 2)	10	6
T	5 (3 and 2)	17	5
L	5 (3 and 2)	10	6
X	5 (3 and 2)	7	7
S	14 (10 and 4)	7	7
A	6 (4 and 2)	10	6
South Vietnam			
BR	5 (3 and 2)	8	7
Tv	||	||	||
Tv-b	6 (4 and 2)	10	9
ST	5 (3 and 2)	4	5
BL	6 (4 and 2)	§	9
CM	§	9	4
Kg	5 (3 and 2)	15	7
HT	5 (4 and 1)	11	6

*Length of RUs, 45 and 102 bp; WSSV-TH coordinates of total repeat, 107965–108675.
†Length of RUs, 54 bp; WSSV-TH coordinates of total repeat, 142744–143067.
‡Length of RUs, 69 bp; WSSV-TH coordinates of total repeat, 187899–188312.
§PCR gave no product.
||No data.

markers at this scale. For the ORF14/15 variable region, there may be too little between-isolate variation at an intermediate spatiotemporal scale. Although the runs test gave an insignificant outcome, 12 of 17 WSSV-VN isolates had the same 6030 bp deletion (Table 3). Conversely, for VNTR loci ORF75, ORF94 and ORF125, there appears to be too much between-isolate variation at this scale, as geographically proximate samples were unrelated (Table 3). Our data therefore suggest that the ORF14/15 variable region may be a suitable marker at larger scales (i.e. global spread), whereas VNTR loci may be suitable markers at smaller scales (i.e. between-farm spread). These conclusions are congruent with reported levels of ORF14/15 variation at large scales (Dieu *et al.*, 2004; Marks *et al.*, 2004; Pradeep *et al.*, 2008b), and VNTR variation at small (Pradeep *et al.*, 2008a) and very small (e.g. Hoa *et al.*, 2005) scales.

To date, single or multiple locus sequences of conserved and functional genes (Greiser-Wilke *et al.*, 2000; Uzcategui *et al.*, 2001; Eyer-Silva & Morgado, 2006), RFLP (Eda *et al.*, 2007) and amplified RFLP (ARFLP; Gouvea *et al.*, 1998; Sammels *et al.*, 1999; Hamano *et al.*, 2005) have been used as genetic markers for virus molecular epidemiology, depending on the level of between-isolate genomic variation. To our knowledge, we report here for the first time a systematic comparison between VNTR loci and variable genomic deletions as molecular markers to devise a model to explain the geographical spread of a DNA virus.

Our results indicate that a genomic deletion can be a suitable marker at an intermediate spatiotemporal scale, and probably also at large scales. The application of genomic-deletion markers is, however, probably both limited and transient, as it requires selection for the removal of redundant sequences from the virus genome. Such selection occurs only if a virus is introduced into specific novel environments – those allowing for adaptation in genome size – and for as long as size of the virus genome is evolving rapidly. In the case of WSSV, this means that whilst ORF23/24 is an excellent marker for studying spread early in the WSSV outbreak (i.e. until about 1998), it may be less useful for studying later WSSV spread (i.e. spread in the Americas). VNTRs are probably not useful markers on an intermediate spatiotemporal scale, regardless of whether the virus is adapting rapidly to a novel environment or not. This conclusion is more general because the occurrence of variation in VNTRs is not dependent on adaptation.

WSSV isolates from *P. monodon* in central and southern Vietnam showed deletions ranging from 8539 to 12 166 bp in the ORF23/24 variable region (Dieu *et al.*, 2004). The nine new isolates characterized in this study also had a deletion size within this range, and deletion size became progressively larger from central Vietnam to either the north or the south (Fig. 2; Table 3). Isolates from central Vietnam (K, T, L, H) were similar, and these isolates have the smallest ORF23/24 deletion size of all Vietnamese WSSV isolates (Fig. 2; Dieu *et al.*, 2004). We therefore propose that (i) the first introduction of WSSV in Vietnam was in central Vietnam and (ii) the virus then spread to the south and probably the north of Vietnam from this site (Fig. 4). This suggests that the spread of WSSV was concomitant with the spread of shrimp aquaculture in Vietnam, which was first introduced in central Vietnam and later in the south and the north of Vietnam (Nguyen, 2008). However, we cannot rule out the possibility that WSSV in northern Vietnam may have originated from China, which was reported to be a source of shrimp seed for northern Vietnam (MOFI, 2001). As more data on WSSV become available for this region (Tan *et al.*, 2009), it may become possible to evaluate and compare different spread models.

An alternative explanation of the patterns observed for the ORF23/24 locus (Fig. 2) is that they stem from environ-

Table 3. Statistical analysis of data for all five genomic loci

The runs test determines whether the number of runs in a sample is greater or lower than the number of runs expected if the outcomes for a trait are independent. A suitable genetic marker should give a significant *P*-value (indicated by an asterisk), indicating that RU or deletion-size values of consecutive geographical samples are not independent, but related. The Jonckheere–Terpstra test was used to determine whether RU or deletion-size values increase or decrease when the samples are ordered along the Vietnamese coast, from north to south. For all statistical tests, the significance threshold was adjusted ($\alpha' = 0.010$) because multiple comparisons are being made. Stand. *JT*, standard *JT* value.

Variable locus	No. cases	Runs test			Jonckheere–Terpstra test	
		Runs	*Z*-value	*P*-value	Stand. *JT*	*P*-value
ORF75	15	4	−1.139	0.255	−1.018	0.309
ORF94	15	6	−1.059	0.290	0.301	0.763
ORF125	15	10	0.556	0.578	0.051	0.959
ORF23/24	17	3	−3.002	0.003*	2.853	0.004*
ORF14/15	17	6	0.000	1.000	−1.290	0.197

**P*-value is significant.

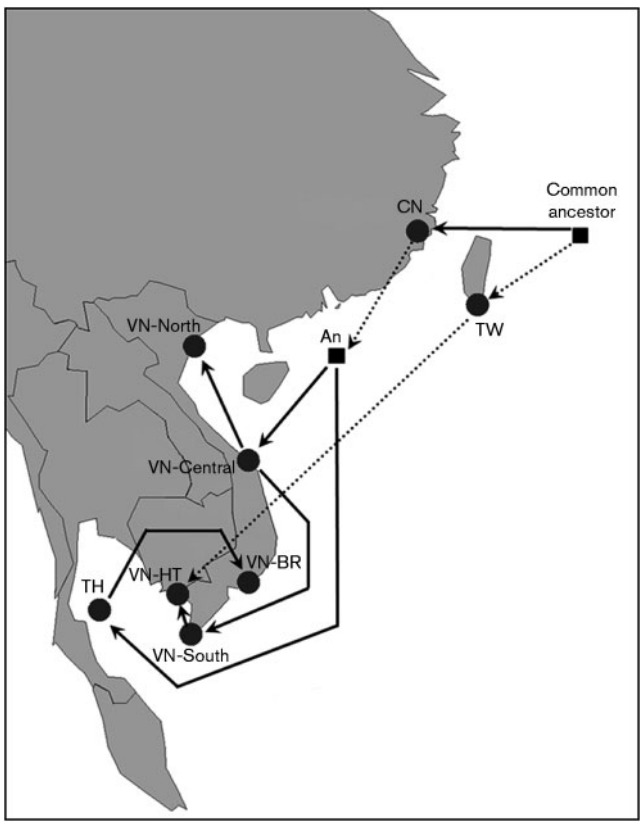

Fig. 4. Model of spread of WSSV in South-East Asia from Taiwan and China towards the west. ● represents identified isolates or groups of isolates, whilst ■ represents hypothetical isolates. An intermediate ancestor is abbreviated as An. Solid arrows represent deletions in the ORF23/24 region, whilst dotted arrows represent deletions in the ORF14/15 region.

mental differences along the Vietnamese coast. As the Vietnamese coastline extends further than 2500 km, there are differences in temperature, salinity, rainfall, vector species and their densities, and aquaculture practices. Central Vietnam, for example, is a relatively hot and dry region, and this may explain why certain ORF23/24 variants only appear there. However, small ORF14/15 deletion variants typical of Vietnam (X, S and A; Fig. 3) are only found in this region (HT and BR are not typical of Vietnam; the presence of these variants probably stems from import of post larvae, as discussed below). Although conditions in central Vietnam could result in selection for variants with smaller deletions at both loci, we think that the most parsimonious explanation is that these variants reflect the spread of WSSV.

We found two variants that are present as unique subgroups for Vietnam, represented as follows. (i) Isolate HT, from the south of Vietnam, has the same genotype in ORF14/15 as WSSV-TW, which suggests that this isolate may have been generated through recombination between directly introduced WSSV-TW and existing Vietnamese variants (Fig. 4). (ii) Isolate BR has the same genotype as WSSV-TH in both the ORF23/24 and ORF14/15 loci, suggesting that perhaps this WSSV variant was introduced from Thailand (Fig. 4). The presence of these variants in Vietnam suggests that human activities – such as the transportation of post larvae and broodstock – have contributed to the long-range spread of WSSV. However, these molecular data must be interpreted carefully, because there can be genetic diversity within WSSV populations (Hoa *et al.*, 2005; Pradeep *et al.*, 2008a; B. T. M. Dieu & J. M. Vlak, unpublished data) and the methods used here are expected to detect only the predominant genotype in an isolate.

We collected WSSV isolates at different geographical locations at a single time point (i.e. 2003/2004). However, the word 'spatiotemporal' is used to describe

the underlying spread process, to stress a temporal component implicit to our understanding of spread: WSSV molecular evolution during site-to-site transmission, and at each site where the virus has become established. The stability of WSSV genotypes at a geographical site will therefore influence how suitable these methods are for retrospective determination of the spread of WSSV. In other words, is the WSSV genotype(s) sampled at a location genetically representative of the WSSV strains first introduced into this area? Striking spatial patterns seen for WSSV isolates (Dieu *et al.*, 2004; Pradeep *et al.*, 2008b) suggest that this is the case, but solid empirical support – e.g. longitudinal studies of WSSV evolution – is missing. WSSV variable loci are considered key elements for understanding the rapid emergence and evolution of this rampant virus in shrimp culture, but a complete, experimentally supported framework would be valuable for interpreting marker data and knowing the limitations of the model.

METHODS

Infected-shrimp sampling. The origin of the WSSV-infected shrimp (*P. monodon*) analysed in this study is shown in Table 1. The shrimp were cleaned with 70 % ethanol and kept in 96 % ethanol during transportation to Can Tho University (Vietnam). After transportation, the ethanol was removed and the samples were stored at −20 °C until further processing.

Analysis of variable loci. DNA extracts of collected shrimp, primarily from gill tissue, were screened for the presence of WSSV with specific primers for *VP26*, as described by Dieu *et al.* (2004). PCR on the genomic variable loci of WSSV was performed with 250 ng DNA extract using *Taq* DNA polymerase (Promega). The specific primer sets, PCR conditions used and sizes of the PCR products are shown in Supplementary Table S1. PCR products were cloned, sequenced and analysed according to published procedures (Dieu *et al.*, 2004).

Statistical analysis. All statistical analysis was performed in SPSS 15.0 (SPSS Inc.). We considered the following quantitative traits for different loci: (i) the number of RUs for VNTR loci with one repeat type (ORF94 and ORF125); (ii) the total number of RUs, regardless of the identity of the repeat, for the VNTR locus with multiple repeat types (ORF75); or (iii) the size of the genomic deletion (ORF23/24 and ORF14/15 regions). WSSV isolates were given an ordinal code corresponding to their relative location along the Vietnamese coast, from north to south [e.g. VN-HP=1, VN-ND=2, (...), VN-BL=14, VN-CM=15, VN-Kg=16, VN-HT=17], and a 'runs test' (Wald & Wolfowitz, 1940) was then performed. A 'run' is a series of consecutive samples with a trait value greater than or less than the cut-off point, a threshold value for which we used the mean of a trait. This procedure tests whether the number of runs in a sample is greater than or lower than the number of runs expected if trait values are independent for each sample. If WSSV spread along the Vietnamese coast, a suitable genetic marker should give a significantly smaller numbers of runs than expected by chance, because geographically proximate isolates (i.e. from consecutive locations along the coast) are likely to have similar trait values.

A Jonckheere–Terpstra test (Bewick *et al.*, 2004) was used to determine whether median RU number or deletion size increased or decreased when the samples were ordered from north to south. We tested five variable loci, and a Šidák correction (Sokal & Rohlf, 1995) was therefore made to the significance threshold α for both the runs test and the Jonckheere–Terpstra test, such that the threshold P-value is $\alpha' = 1-(1-\alpha)^{1/n} = 1-(1-0.05)^{1/5} = 0.010$.

ACKNOWLEDGEMENTS

B. T. M. D. was supported by a WOTRO PhD scholarship (grant WB 83-207).

REFERENCES

Bewick, V., Cheek, L. & Ball, J. (2004). Statistics review 10: further nonparametric methods. *Crit Care* **8**, 196–199.

Corsin, F., Turnbull, J. F., Hao, N. V., Mohan, C. V., Phi, T. T., Phuoc, L. H., Tinh, N. T. N. & Morgan, K. L. (2001). Risk factors associated with white spot syndrome virus infection in a Vietnamese rice-shrimp farming system. *Dis Aquat Organ* **47**, 1–12.

Corsin, F., Phi, T. T., Phuoc, L. H., Tinh, N. T. N., Hao, N. V., Mohan, C. V., Turnbull, J. F. & Morgan, K. L. (2002). Problems and solutions with the design and execution of an epidemiological study of white spot disease in black tiger shrimp (*Penaeus monodon*) in Vietnam. *Prev Vet Med* **53**, 117–132.

Davis, C. L., Field, D., Metzgar, D., Saiz, R., Morin, P. A., Smith, I. L., Spector, S. A. & Wills, C. (1999). Numerous length polymorphisms at short tandem repeats in human cytomegalovirus. *J Virol* **73**, 6265–6270.

Dieu, B. T. M., Marks, H., Siebenga, J., Goldbach, R., Zuidema, D., Duong, T. P. & Vlak, J. M. (2004). Molecular epidemiology of white spot syndrome virus within Vietnam. *J Gen Virol* **85**, 3607–3618.

Eda, H., Ozawa, S., The Cooperation Group for HSV-1 RFLP Variant Study, Yoshino, K. & Yanagi, K. (2007). Contrasting geographic distribution profiles of the herpes simplex virus type 1 BgOL and BgKL variants in Japan suggest dispersion and replacement. *J Clin Microbiol* **45**, 771–782.

Eyer-Silva, W. A. & Morgado, M. G. (2006). Molecular epidemiology of HIV-1 infection in a small Brazilian county. Usefulness of envelope and polymerase sequences to epidemiologic studies. *J Acquir Immune Defic Syndr* **41**, 664–670.

Flegel, T. W. (1997). Major viral disease of the black tiger prawn (*Penaeus monodon*) in Thailand. *World J Microbiol Biotechnol* **13**, 433–442.

Gouvea, V., Hoke, C., Jr & Innis, B. L. (1998). Genotyping of hepatitis E virus in clinical specimens by restriction endonuclease analysis. *J Virol Methods* **70**, 71–78.

Greiser-Wilke, I., Fritzemeier, J., Koenen, F., Vanderhallen, H., Rutili, D., Mia, G. M. D., Romero, L., Rosell, R., Sanchez-Vizcaino, J. M. & Gabrie, A. S. (2000). Molecular epidemiology of a large classical swine fever epidemic in the European Union in 1997–1998. *Vet Microbiol* **77**, 17–27.

Hamano, M., Maeda, K., Kai, K., Mochizuki, M., Tohya, Y. & Akashi, H. (2005). A novel genetic marker to differentiate feline herpesvirus type 1 field isolates. *Vet Microbiol* **106**, 195–200.

Hoa, T. T. T., Hodgson, R. A. J., Oanh, D. T. H., Phuong, N. T., Preston, N. J. & Walker, P. J. (2005). Genotypic variations in tandem repeat DNA segments between ribonucleotide reductase subunit genes of white spot syndrome virus (WSSV) isolates from Vietnam. In *Diseases in Asian Aquaculture V*, pp. 339–351. Edited by P. J. Walker, R. G. Lester & M. G. Bondad-Reantaso. Manila, the Philippines: Fish Health Section, Asian Fisheries Society.

Kang, H. H. & Lu, C. P. (2007). Comparison of variable region genes of shrimp white spot syndrome virus (WSSV) in different areas in China. *Bing Du Xue Bao* **23**, 490–493 (in Chinese).

Kiatpathomchai, W., Taweetungtragoon, A., Jittivadhana, K., Wongteerasupaya, C., Boonsaeng, V. & Flegel, T. W. (2005). Target for standard Thai PCR assay identical in 12 white spot syndrome virus (WSSV) types that differ in DNA multiple repeat length. *J Virol Methods* **130**, 79–82.

Lan, Y., Lu, W. & Xu, X. (2002). Genomic instability of prawn white spot bacilliform virus (WSBV) and its association to virus virulence. *Virus Res* **90**, 269–274.

Lo, C.-F., Ho, C.-H., Peng, S.-E., Chen, C.-H., Hsu, H.-C., Chiu, Y.-L., Chang, C.-F., Liu, K.-F., Su, M.-S. & other authors (1996a). White spot syndrome baculovirus (WSBV) detected in cultured and captured shrimp, crabs and other arthropods. *Dis Aquat Organ* **27**, 215–225.

Lo, C.-F., Leu, J.-H., Ho, C.-H., Chen, C.-H., Peng, S.-E., Chen, Y.-T., Chou, C.-M., Yeh, P.-Y., Huang, C.-J. & other authors (1996b). Detection of baculovirus associated with white spot syndrome (WSBV) in penaeid shrimps using polymerase chain reaction. *Dis Aquat Organ* **25**, 133–141.

Lo, C.-F., Hsu, H.-C., Tsai, M.-F., Ho, C.-H., Peng, S.-E., Kou, G.-H. & Lightner, D.-V. (1999). Specific genomic DNA fragment analysis of different geographical clinical samples of shrimp white spot syndrome virus. *Dis Aquat Organ* **35**, 175–185.

Marks, H., Goldbach, R. W., Vlak, J. M. & van Hulten, M. C. W. (2004). Genetic variation among isolates of White spot syndrome virus. *Arch Virol* **149**, 673–697.

Marks, H., van Duijse, J. J. A., Zuidema, D., van Hulten, M. C. W. & Vlak, J. M. (2005). Fitness and virulence of an ancestral white spot syndrome virus isolate from shrimp. *Virus Res* **110**, 9–20.

MOFI (2001). *Annual Report of Fisheries Production in 2000 and Action Plan for the Year 2001* (in Vietnamese). Hanoi, Vietnam: Ministry of Fisheries.

Moon, C. H., Do, J. W., Cha, S. J., Yoon, W. J., Kim, S. B., Ko, M. S., Park, M. A., Kim, J. W., Sohn, S. K. & other authors (2003). Highly conserved sequences of three major virion proteins of a Korean isolate of white spot syndrome virus (WSSV). *Dis Aquat Organ* **53**, 11–13.

Musthaq, S. S., Sudhakaran, R., Ahmed, V. P. I., Balasubramanian, G. & Sahul Hameed, A. S. (2006). Variability in the tandem repetitive DNA sequences of white spot syndrome virus (WSSV) genome and suitability of VP28 gene to detect different isolates of WSSV from India. *Aquaculture* **256**, 34–41.

Nadala, E. C. B. & Loh, P. C. (1998). A comparative study of three different isolates of white spot virus. *Dis Aquat Organ* **33**, 231–234.

Nguyen, V. H. (2008). Vietnam's catfish and marine shrimp production: an example of growth and sustainability issues. *Panorama Acuicola Magazine* **13**, 46–55. http://www.panoramaacuicola.com/ediciones/pam_13-3/46-55.pdf

Pradeep, B., Shekar, M., Gudkovs, N., Karunasagar, I. & Karunasagar, I. (2008a). Genotyping of white spot syndrome virus prevalent in shrimp farms of India. *Dis Aquat Organ* **78**, 189–198.

Pradeep, B., Shekar, M., Karunasagar, I. & Karunasagar, I. (2008b). Characterization of variable genomic regions of Indian white spot syndrome virus. *Virology* **376**, 24–30.

Pradeep, B., Karunasagar, I. & Karunasagar, I. (2009). Fitness and virulence of different strains of white spot syndrome virus. *J Fish Dis* **32**, 801–805.

Sammels, L. M., Lindsay, M. D., Poidinger, M., Coelen, R. J. & Mackenzie, J. S. (1999). Geographic distribution and evolution of Sindbis virus in Australia. *J Gen Virol* **80**, 739–748.

Sokal, R. R. & Rohlf, F. J. (1995). *Biometry: the Principles and Practice of Statistics in Biological Research*. New York: W. H. Freeman.

Tan, Y., Xing, Y., Zhang, H., Feng, Y., Zhou, Y. & Shi, Z. L. (2009). Molecular detection of three shrimp viruses and genetic variation of white spot syndrome virus in Hainan Province, China, in 2007. *J Fish Dis* **32**, 777–784.

Uzcategui, N. Y., Camacho, D., Comach, G., Cuello de Uzcategui, R., Holmes, E. C. & Gould, E. A. (2001). Molecular epidemiology of dengue type 2 virus in Venezuela: evidence for in situ virus evolution and recombination. *J Gen Virol* **82**, 2945–2953.

van Hulten, M. C. W., Witteveldt, J., Peters, S., Kloosterboer, N., Tarchini, R., Fiers, M., Sandbrink, H., Lankhorst, R. K. & Vlak, J. M. (2001). The white spot syndrome virus DNA genome sequence. *Virology* **286**, 7–22.

Vlak, J. M., Bonami, J. R., Flegel, T. W., Kou, G. H., Lightner, D. V., Lo, C. F., Loh, P. C. & Walker, P. W. (2005). Family *Nimaviridae*. In *Virus Taxonomy: Eighth Report of the International Committee on Taxonomy of Viruses*, pp. 187–192. Edited by C. M. Fauquet, M. A. Mayo, J. Maniloff, U. Desselberger & L. A. Ball. San Diego, CA: Elsevier Academic Press.

Waikhom, G., John, K. R., George, M. R. & Jeyaseelan, M. J. P. (2006). Differential host passaging alters pathogenicity and induces genomic variation in white spot syndrome virus. *Aquaculture* **261**, 54–63.

Wald, A. & Wolfowitz, J. (1940). On a test whether two samples are from the same population. *Ann Math Stat* **11**, 147–162.

Wang, C.-H., Lo, C.-F., Leu, J.-H., Chou, C.-M., Yeh, P.-Y., Chou, H.-Y., Tung, M.-C., Chang, C.-F., Su, M.-S. & Kou, G.-H. (1995). Purification and genomic analysis of baculovirus associated with white spot syndrome (WSBV) of *Penaeus monodon*. *Dis Aquat Organ* **23**, 239–242.

Wang, Y. T., Liu, W., Seah, J. N., Lam, C. S., Xiang, J. H., Korzh, V. & Kwang, J. (2002). White spot syndrome virus (WSSV) infects specific hemocytes of the shrimp *Penaeus merguiensis*. *Dis Aquat Organ* **52**, 249–259.

Yang, F., He, J., Lin, X., Li, Q., Pan, D., Zhang, X. & Xu, X. (2001). Complete genome sequence of the shrimp white spot bacilliform virus. *J Virol* **75**, 11811–11820.