One would also expect that the future environment, which is predicted to be warmer and drier, may be still critical to its distribution and productivity. Thus efforts are required for its conservation.

1. Maikhuri, R. K., Nautiyal, M. C. and Khali, M. P., Lesser-known crops of food value in Garhwal Himalaya and a strategy to conserve them. *Plant Genet. Resour. Newsl.*, 1991, **86**, 33–36.
2. Sanz-Elorza, M., Dana, E. D., Gonzalez, A. and Sobrino, E., Changes in the high-mountain vegetation of the central Iberian Peninsula as a probable sign of global warming. *Ann. Bot.*, 2003, **92**, 273–280.
3. Thomas, C. D. *et al.*, Extinction risk from climate change. *Nature*, 2004, **427**, 145–148.
4. Berry, J. A. and Downton, W. J. S., Environmental regulation of photosynthesis. In *Photosynthesis, Vol. II, Development, Carbon Metabolism and Plant Productivity* (ed. Govindjee), Academic Press, New York, 1982, pp. 263–343.
5. Joshi, S. C. and Palni, L. M. S., Clonal variation in temperature response of photosynthesis in tea. *Plant Sci.*, 1998, **137**, 225–232.
6. Isebrands, J. G., Ceulemans, R. and Wiard, B., Genetic variation in photosynthetic traits among *Populus* clones in relation to yield. *Plant Physiol. Biochem.*, 1988, **26**, 427–437.
7. Berry, J. and Bjorkman, O., Photosynthetic response and adaptation to temperature in higher plants. *Annu. Rev. Plant Physiol.*, 1980, **31**, 491–543.
8. Schneider, S. H., The global warming debate heats up: An analysis and perspective. *Bull. Am. Meteorol. Soc.*, 1990, **71**, 1292–1304.
9. VonCaemmever, S. and Farquhar, G. D., Some relationships between the biochemistry of photosynthesis and the gas exchange of leaves. *Planta*, 1981, **153**, 376–387.
10. Mather, K., *Statistical Analysis in Biology*, Chapman and Hall, London, 1973.
11. Fisher, R. A., Bidinger, F., Syme, J. R. and Wall, P. C., Leaf photosynthesis, leaf permeability, crop growth and yield of short spring wheat genotypes under irrigation. *Crop Sci.*, 1981, **21**, 367–373.
12. Jones, H. G., *Plants and Microclimate: Quantitative Approach to Environmental Plant Physiology*, Cambridge University Press, Cambridge, 1992.
13. Lawlor, D. W., Photosynthesis, productivity and environment. *J. Exp. Bot.*, 1995, **46**, 1449–1461.
14. Rodriguez, Sanchez, Perez, P. and Martinez-Carrasco, R., Photosynthesis, carbohydrate levels and chlorophyll fluorescence estimated intercellular CO_2 in water stressed *Casuarina aquisetifolia* Forst and Forst. *Plant Cell Environ.*, 1999, **22**, 867–873.
15. Jones, H. G., Partitioning, stomatal and non-stomatal limitations to photosynthesis. *Plant Cell Environ.*, 1985, **8**, 95–104.
16. Ni, B.-R. and Pallardy, S. G., Stomatal and non-stomatal limitations to net photosynthesis in seedlings of woody angiosperms. *Plant Physiol.*, 1992, **99**, 1502–1508.
17. Alexander, J. D., Donnelly, J. R. and Shane, J. B., Photosynthetic and transpiration responses of red spruce understorey trees to light and temperature. *Tree Physiol.*, 1995, **15**, 393–398.
18. Rawson, H. M., Begg, J. R. and Woodward, R. G., The effect of atmospheric humidity on photosynthesis, transpiration and water use efficiency of leaves of several plant species. *Planta*, 1977, **134**, 5–10.
19. Cabuslay, G. S., Ito, O. and Alejar, A. A., Physiological evaluation of responses of rice (*Oryza sativa* L.) to water deficit. *Plant Sci.*, 2002, **163**, 815–827.
20. Zhang, J. W., Marshall, J. D. and Fins, L., Correlated population differences in dry matter accumulation, allocation and water use efficiency in three sympatric conifer species. *For. Sci.*, 1996, **42**, 242–249.
21. Miroslavov, E. A. and Kravkina, I. M., Comparative analysis of chloroplasts and mitochondria in leaf chlorenchyma from mountain plants grown at different altitudes. *Ann. Bot.*, 1991, **68**, 195–200.
22. Palit, P., Kundu, A., Mandan, R. K. and Sircan, S. M., Productivity of rice plant in relation to photosynthesis, photorespiration and translocation. *Indian J. Plant Physiol.*, 1979, **22**, 66–74.
23. Pearcy, R. W., Acclimation of photosynthetic and respiratory carbon dioxide exchange to growth temperature in *Atriplex lentiformis* (Torr.) Wats. *Plant Physiol.*, **59**, 795–799.

ACKNOWLEDGEMENTS. We thank the Director, G.B. Pant Institute of Himalayan Environment and Development, Almora for providing necessary facilities and encouragement.

Received 10 January 2005; accepted 30 May 2005

A computer-based identification of variable number tandem repeats in White spot syndrome virus genomes

Malathi Shekar, Indrani Karunasagar and I. Karunasagar*

Department of Fishery Microbiology, College of Fisheries, Mangalore 575 002, India

Three complete genome sequences of White spot syndrome virus (WSSV) from Thailand, Taiwan and China, were analysed for the presence of tandem repeats. Thirteen microsatellite and three minisatellite loci which could serve as potential molecular markers for studying and understanding the genetic and epidemiological relationship among WSSV strains were identified. Additionally, we report the presence of two new polymorphic megasatellite loci within the genomes.

Keywords: Microsatellites, minisatellites, megasatellites, variable number tandem repeats, White spot syndrome virus.

WHITE spot syndrome virus (WSSV), an important shrimp virus and one of the largest animal viruses is associated with shrimp mortalities, causing huge economic losses to cultured shrimp worldwide. WSSV has a large genome size; its size has been reported as 292,292 base pairs (bp)[1], 305,107 bp[2] and 307,287 bp[3] for Thailand, Taiwan and China isolates respectively. Since these three isolates show little variation across their genomes with much of their predicted proteins showing no sequence homology to any other viral genes or genes from other organisms[1,2,4,5], knowledge regarding the evolution, epidemiology or the pathogenic mechanism of this organism has been limited.

*For correspondence. (e-mail: mircen@sancharnet.in)

In the last few years, tandem repeats, i.e. DNA sequences occurring as multiple copies arranged tandemly at a single loci, are being increasingly recognized as informative markers for studying genotypic variations among strains[6,7] and in genomic evolution[8]. Tandem repeat loci exhibiting variability in their copy numbers are referred as variable number tandem repeats (VNTRs). VNTRs of microsatellites (with repeat unit tracts ranging from 1–6 bp) and minisatellites (repeat unit tracts of 7–100 bp) occur in both prokaryotic[9,10] as well as eukaryotic[11] genomes. Within genomes, VNTRs could be located either in the protein-coding or non-coding regions. Studies on the inter-individual variability in copy numbers of VNTR alleles have found applications in DNA fingerprinting in humans[12,13] as well as in other organisms[14]. In bacteria, in addition to studying genotypic variation[7,15,16], VNTRs serve as potential markers for the identification of pathogenic bacteria[17–19] and the virulence factors associated with their pathogenicity[20]. Studies with certain viral genomes have further shown that polymorphism exists among viruses[21] and has proven to be useful as markers in epidemiological[22] and virulence[23] studies.

In this study, we screened the complete genomes of WSSV obtained from GenBank for the presence of VNTR loci, which could facilitate the development of new markers for detection and characterization of WSSV strains.

The three complete WSSV genomes in GenBank (accession numbers AF369029 (Thailand), AF332093 (Taiwan), and AF440570 (China)) were analysed for the presence of tandem repeats by Tandem Repeats Finder (TRF) program developed by Benson and coworkers[24]. The program was run with parameters set to +2, −5, −3 (match, mismatch, indel) and minimum score adjusted to 30. Tandem repeats thus obtained for the three genomes were further compared for repeat size, copy number and genomic location. For comparative analysis, repeats obtained were compared for similarity regions and copy variability within these regions using BLAST and ClustalW programs. The presence of repeats within coding and non-coding regions was identified based on the annotation of the WSSV genomes in GenBank.

Figure 1 shows the output of analyses for the three WSSV genomes with the TRF program and is limited to the unit repeat size and to the total number of occurrences of these repeats within the genomes. Our analyses show that there is a relatively large number of repeats ranging between sizes 1 and 326 bp within WSSV genomes. Further, short repeats appear to be more abundant compared to larger repeats. As seen from Figure 1, there exists a high frequency of repeats of sizes 3, 9 and 12 bases within the genomes.

Computer survey also revealed the presence of 18 polymorphic VNTR loci, of varying repeat sizes in the WSSV genomes. Tables 1–3 summarize the repeat size, genomic position, copy number, repeat sequence, classification of repeats as perfect or compound based on whether they occur as a single unit or a combination of two or more repetitive units, presence within coding or non-coding regions and their associated open reading frames (ORFs) for micro, mini and megasatellite VNTRs within the three genomes.

Since the repeat sequence is determined based on a consensus pattern in the TRF program; the program output is not always similar to the actual repeat sequence. For example, the repeat sequence unit output using the TRF program was CTC with 14 copies for China (AF440570) isolate. However, on actual observation we see contiguous copies of TCC repeated 12 times with starting position at 230529 (Figure 2 *a*). Similarly; in Figure 2 *b*, instead of $(ACT)_{14}$, we have a compound repeat of $(ACA)_4$ ACC $(ACT)_4$ at position 14327 for Thailand genome (AF369029). Hence, in Tables 1–3 the genomic start position and the repeat unit sequence have been modified based on the observed repeated sequence in the prescribed region. This would enable us to more specifically design primers for amplification of a particular repeat region.

Almost all of microsatellite VNTRs observed were triplet repeats. Eleven out of these 13 microsatellite VNTRs were

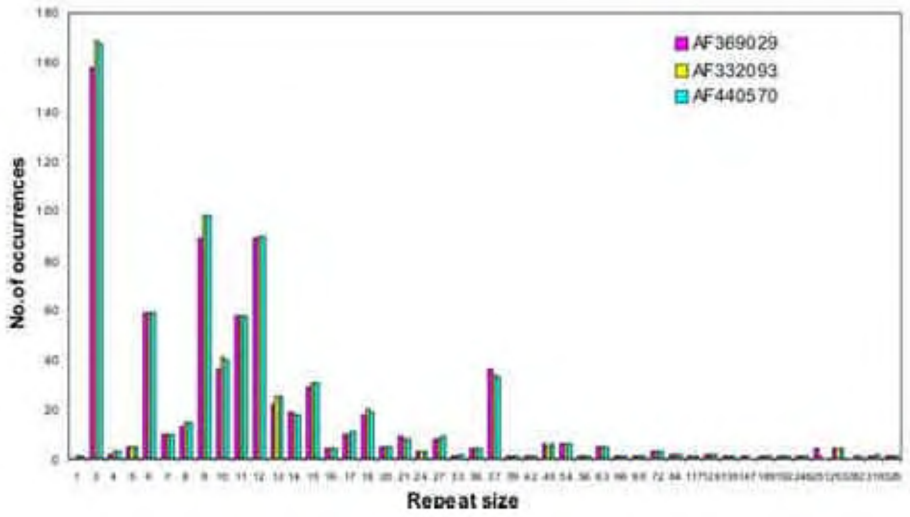

Figure 1. Comparative occurrence of various repeat sizes within the three WSSV genomes.

Table 1. Microsatellite copy variation within similar genomic loci of the three WSSV genomes

Repeat size	Genome/start position	Observed repeat sequence	Repeat classification	Coding/ noncoding	Associated ORF
3	TH-67689	(**AGA**)$_5$ (GGA)$_4$	Compound	coding	orf 42
	TW-18733	(**AGA**)$_6$ (GGA)$_4$	Compound	coding	wsv 037
	CN-52322	(**AGA**)$_6$ (GGA)$_4$	Compound	coding	wssv 094
3	TH-89487	(TGT)$_6$	Perfect	non-coding	–
	TW-40535	(TGT)$_7$	Perfect	non-coding	–
	CN-74121	(TGT)$_7$	Perfect	non-coding	–
3	TH-93164	(CAG)$_2$ (**TAG**)$_5$ (CAG)$_2$ (TAG)$_2$ (CAG)$_3$ (TAG)$_2$	Compound	coding	orf 65
	TW-44216	(CAG)$_2$ (**TAG**)$_4$ (CAG)$_2$ (TAG)$_2$ (CAG)$_3$ (TAG)$_2$	Compound	coding	wsv 091
	CN-77798	(CAG)$_2$ (**TAG**)$_5$ (CAG)$_2$ (TAG)$_2$ (CAG)$_3$ (TAG)$_2$	Compound	coding	wssv 149
3	TH-119135	(**GCT**)$_8$ (GCC)$_2$	Compound	coding	orf 84
	TW-70375	(**GCT**)$_5$ GCC (GCT)$_2$ GCC	Compound	coding	wsv 143
	CN-104284	(**GCT**)$_5$ GCC (GCT)$_2$ GCC	Compound	coding	wssv 198
3	TH-119219	(**GCT**)$_8$ (GCC)$_2$	Compound	coding	orf 84
	TW-70456	(**GCT**)$_5$ GCC (GCT)$_2$ GCC	Compound	coding	wsv 143
	CN-104365	(**GCT**)$_5$ GCC (GCT)$_2$ GCC	Compound	coding	wssv 198
3 and 6	TH-180799	AATGGA (GGA)$_2$ (AATGGA)$_4$	Compound	coding	orf 119
	TW-131838	AATGGA (**GGA**)$_2$ (**AATGGA**)$_2$ (GGA)$_2$ (AATGGA)$_4$	Compound	coding	wsv 238
	CN-165439	AATGGA (**GGA**)$_2$ (**AATGGA**)$_2$ (GGA)$_2$ (AATGGA)$_4$	Compound	coding	wssv 294
3	TH-185395	(C)$_3$ (T)$_5$ (C)$_3$ (**CTT**)$_{11}$	Compound	non-coding	–
	TW-136453	(C)$_3$ (T)$_5$ (C)$_3$ (**CTT**)$_{10}$	Compound	non-coding	–
	CN-170054	(C)$_3$ (T)$_5$ (C)$_3$ (**CTT**)$_9$	Compound	non-coding	–
3	TH-189654	(A)$_4$ T (**ACT**)$_7$	Compound	coding	orf 126
	TW-140847	(A)$_4$ T (**ACT**)$_6$	Compound	coding	wsv 252
	CN-174445	(A)$_4$ T (**ACT**)$_6$	Compound	coding	wssv 307
3	TH-200953	(TCA)$_4$ (**TCT**)$_5$ TCA (TCC)$_4$	Compound	coding	orf 134
	TW-152149	(TCA)$_4$ (**TCT**)$_4$ TCA (TCC)$_4$	Compound	coding	wsv 271
	CN-185741	(TCA)$_4$ (**TCT**)$_5$ TCA (TCC)$_4$	Compound	coding	wssv 326
3	TH-230529	(TCC)$_{10}$	Perfect	non-coding	–
	TW-181730	(TCC)$_{11}$	Perfect	non-coding	–
	CN-215323	(TCC)$_{12}$	Perfect	non-coding	–
3	TH-14327	(ACA)$_4$ ACC (**ACT**)$_7$	Compound	coding	orf 8
	TW-258569	(ACA)$_4$ ACC (**ACT**)$_7$	Compound	coding	wsv 446
	CN-293191	(ACA)$_4$ ACC (**ACT**)$_8$	Compound	coding	wssv 511
3 and 6	TH-18164	(CCT)$_5$ (**ACTCCT**)$_2$	Compound	non-coding	–
	TW-262406	(CCT)$_5$ (**ACTCCT**)$_2$	Compound	coding	wsv 451
	CN-267031	(CCT)$_5$ **ACTCCT**	Compound	coding	wssv 454
3	TH-231669	TTC (**CTC**)$_4$ TTC CTC	Compound	non-coding	–
	TW-182873	TTC (**CTC**)$_4$ TTC CTC	Compound	coding	wsv 318
	CN-216469	TTC (**CTC**)$_4$ **TTC** (**CTC**)$_3$ TTC CTC	Compound	coding	wssv 374

TH, Thailand; TW, Taiwan; CN, China isolates.
Compound repeats showing variation in their units are in boldface letters.

observed to occur as compound repeats showing copy number variation in one of their units. These have been marked bold in Table 1. We observed two large perfect tandem repetition units: a 253 and 326 bp repeated several times (Table 2), referred to here as megasatellites. While the former was located in a non-coding region, the latter was spread partially between a non-coding and a protein coding ORF located within the repeated region. Among the three minisatellite VNTRs, one was observed to be a compound repeat with a 45 bp repeat unit interspersed with a 57 bp nucleotide sequence. The 45 bp unit was a perfect repeat varying only in copy numbers (Table 2).

A similarity is observed among the three genomes with regard to the number of detected repeats and their occurrences within the genomes. The relatively large number of repeats existing within this viral genome suggests that the WSSV genome may be prone to mutational processes leading to insertion and deletion of nucleotides, and polymorphism, if any, arising due to these events could be useful in using these repeats as markers. This is akin to the earlier observations in several genomes studied. We looked further for polymorphism within the identified repeats and identified 13 microsatellites (unit repeat size < 10 bp), three minisatellites (unit repeat size < 100 bp) and two mega-

Table 2. Microsatellite copy variations within similar genomic loci of the three WSSV genomes

Repeat unit (bp)	Genome start position	(Repeat unit) copy numbers	Classification	Coding/ non-coding	Associated ORF
54	TH-142744	$(54)_6$	Perfect	Coding	orf 94
	TW-93475	$(54)_{12}$	Perfect	Coding	wsv 178
	CN-127388	$(54)_6$	Perfect	Coding	wssv 234
45 and 57	TH-107964	$(45)_2$ 57 $(45)_5$ 57 $(45)_3$ 57 (45)	Compound	Coding	orf 75
	TW-59013	$(45)_2$ 57 $(45)_5$ 57 $(45)_3$ 57 $(45)_3$ 57 $(45)_2$	Compound	Coding	wsv 128
	CN-92594	$(45)_2$ 57 $(45)_5$ 57 $(45)_4$ 57 $(45)_3$ 57 $(45)_5$ 60*$(45)_2$	Compound	Coding	wssv 183
69	TH-187894	$(69)_6$	Perfect	Coding	orf 125
	TW-138949	$(69)_8$	Perfect	Coding	wsv 249
	CN-172547	$(69)_6$	Perfect	Coding	wssv 304

TH, Thailand isolates; TW, Taiwan isolates; CN, China isolates.
*Extra copy of GAA present.

Table 3. Megasatellite copy variations within similar genomic loci of the three WSSV genomes

Repeat unit (bp)	Genome start position	(Repeat unit) copy numbers	Classification	Coding/non-coding	Associated ORF
253	TH-126432	$(253)_6$	Perfect	Non-coding	–
	TW-77669	$(253)_4$	Perfect	Non-coding	–
	CN-111575	$(253)_6$	Perfect	Non-coding	–
326	TH-28251	$(326)_6$	Perfect	Partially coding and non-coding	17
	TW-272678	$(326)_5$	Perfect	Partially coding and non-coding	473
	CN-2	$(326)_5$	Perfect	Partially coding and non-coding	532

TH, Thailand isolates; TW, Taiwan isolates; CN, China isolates.

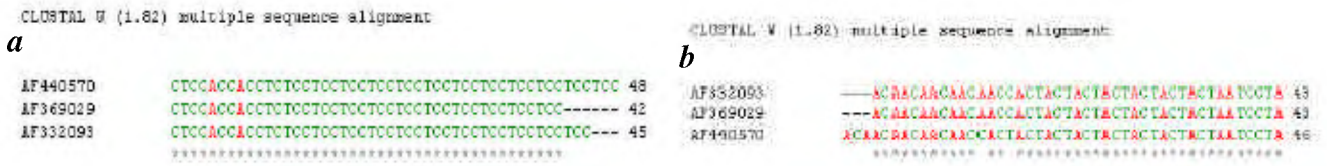

Figure 2. Multiple sequence alignment for microsatellite regions.

satellite (unit repeat size > 100 bp) polymorphic loci within the three WSSV genomes. Studies have shown that microsatellite VNTRs can bring about an alteration in gene expression in eukaryotes[25] and prokaryotes[7,26]. There is strong evidence to show that microsatellite copy variations can act as molecular switches, which play an important role in bacterial phase variation[19,27]. These could be located either within open reading frames, such that alteration in the length of the repeat alters the translational reading frame, or within promotors[28], where they affect the relative position of flanking promoter components and influence transcription. While triplet repeats are known to be associated with phase-mediated bacterial virulence, non-triplets within coding regions have the potential to radically modify protein and shift bacterial phenotypes. Phase-variation genes described as 'contingency genes' facilitate adaptation of bacterial population to changing environmental conditions[29] and are also recognized as an important mechanism for modulation of gene expression involved in bacterial virulence. Though similar mechanism has not been reported in viral genomes thus far, a study by Davis and coworkers[21] has demonstrated the presence of very short polymorphic repeats in human cytomegalovirus genome, suggesting a role of insertion–deletion events and the use of these repeats as moleculer markers. Similar studies with avian influenza virus, have shown variation in the length of trinucleotide repeats to be associated with increased virulence[23]. Most of the microsatellite repeats in this study were triplet repeats existing as compound repeats with variation in any one of their units. These triplet repeat arrays were much shorter than those frequently examined for studies of microsatellite variation. Identification of surrounding ORFs and the protein they encode may be one way of understanding the function of VNTRs. In this study, we tried to associate VNTR loci and their ORFs and found that most of the microsatellites identified were associated within probable protein-coding ORFs. It remains to be seen whether the variability associated with these has any implication as

contingency genes, which could be involved in the virulence and host adaptation in WSSV strains. This could possibly explain the vagaries associated with this viral disease in shrimps and its adaptation to different hosts.

Recently, a study by Wongteerasupaya and coworkers[30] has shown that different WSSV strains isolated from Thailand had a variable 54 bp VNTR locus and suggested its role as a marker for epidemiological studies of WSSV strains. In addition, we could identify two other minisatellite VNTRs; a 69 bp perfect repeat and a compound repeat of 45 and 57 bp (Table 2). Similar results were also observed by Marks and coworkers[31] using a different tool. Our observation also suggests that these three minisatellite variations have significant potential to be used as markers for studying VNTR diversity, which could help understand the genetic and epidemiological relationships between WSSV strains isolated from different geographical regions and from several different hosts. This is a rare feature of this virus.

While micro- and minisatellites are frequently being used as molecule marker tools, the role of higher tandem repeats such as megasatellites has not been determined in any genome except the human genome, where a few studies have suggested the association of these megasatellites with genes that are responsible for producing a large amount of transcripts such as ribosomal RNA[32] and U2 small nuclear RNA[33]. In WSSV, the exact role and nature of these repeats in their genomes require further study.

Understanding the molecular basis of VNTR variation will not only help in strain discrimination, but also on host adaptation and virulence mechanism of any organism. This study provides a substantial number of polymorphic loci in WSSV genomes. The minisatellite markers described here could help discriminate between WSSV strains in epidemiological and ecological studies. Further, the microsatellites described could provide a basis for determining phenotypic variations that can arise due to small differences in gene sequences and their role in the virulence of WSSV strains. Further, VNTR-based molecular biology experiments are required, which could significantly extend our understanding of the infection process in this important viral pathogen.

1. van Hulten M. C. W. *et al.*, The White spot syndrome virus DNA genome sequence. *Virology*, 2001, **286**, 7–22.
2. Yang, F., He, J., Lin, X., Li, Q., Pan, D., Zang, X. and Xu, X., Complete genome sequence of the shrimp White spot bacilliform virus. *J. Virol.*, 2001, **75**, 11811–11820.
3. Tsai, M. F. *et al.*, Transcriptional analysis of the ribonucleotide reductase genes of shrimp White spot syndrome virus. *Virology*, 2000, **277**, 92–99.
4. Moon, C. H. *et al.*, Highly conserved sequences of three major virion proteins of a Korean isolate of White spot syndrome virus (WSSV). *Dis. Aquat. Org.*, 2003 **53**, 11–13.
5. Wang, Q., Poulos, B. T. and Lightner, D. V., Protein analysis of geographic isolates of shrimp white spot syndrome virus. *Arch. Virol.*, 2000, **145**, 263–274.
6. Frothingham, R. and Meeker-O'Connell, W. A., Genetic diversity in the *Mycobacterium tuberculosis* complex based on variable numbers of tandem DNA repeats. *Microbiology*, 1998, **144**, 1189–1196.
7. van Belkum, A., Scherer, S., Van Leuwen, W., Willemse, D., Lock, A. and Verbrugh, H., Variable number of tandem repeats in clinical strains of *Haemophilus influenzae*. *Infect. Immunol.*, 1997, **65**, 5017–5027.
8. Tautz, D., Trick, M. and Dover, G., Cryptic simplicity in DNA is a major source of genetic variation. *Nature*, 1986, **322**, 652–656.
9. Gur-Arie, R., Cohen, C. J., Eitan, Y., Shelef, L., Hallerman, E. M. and Kashi, Y., Simple sequence repeats in *Escherichia coli*: Abundance, distribution, composition, and polymorphism. *Genet. Res.*, 2000, **62**, 62–71.
10. van Belkum, A., Scherer, S., van Alphen, L. and Verbrugh, H., Short-sequence DNA repeats in prokaryotic genomes. *Microbiol. Mol. Biol. Rev.*, 1998, **62**, 275–293.
11. Toth, G., Gaspari, Z. and Jurka, J., Microsatellites in different eukaryotic genomes: Survey and analysis. *Genet. Res.*, 2000, **10**, 967–981.
12. Jeffreys, A. J., Wilson, V. and Thein, S. L., Individual-specific 'fingerprints' of human DNA. *Nature*, 1985, **316**, 76–79.
13. Nakamura, Y. *et al.*, Variable number of tandem repeat (VNTR) markers for human gene mapping. *Science*, 1987, **235**, 1616–1622.
14. Le Fleche P. *et al.*, A tandem repeats database for bacterial genomes: Application to the genotyping of *Yersinia pestis* and *Bacillus anthracis*. *BMC Microbiol.*, 2001 **1**, 2.
15. Kim, W., Hong, Y. P., Yoo, J. H., Lee, W. B., Choi, C. S. and Chung, S. I., Genetic relationships of *Bacillus anthracis* and closely related species based on variable-number tandem repeat analysis and BOX–PCR genomic fingerprinting. *FEMS Microbiol. Lett.*, 2001 **207**, 21–27.
16. Vodopyanov, S. O. *et al.*, Variable Tandem Repeat *Vc*A of *Vibrio cholerae*. *Mol. Biol.*, 2002, **36**, 874–878.
17. Hood, D. W., Deadman, M. E., Jennings, M. P., Bisercic, M., Fleischmann, R. D., Venter, J. C. and Moxon, E. R., DNA repeats identify novel virulence genes in *Haemophilus influenzae*. *Proc. Natl. Acad. Sci. USA*, 1996, **93**, 11121–11125.
18. Klevytska, A. M., Price, L. B., Schupp, J. M., Worsham, P. L., Wong, J. and Keim, P., Identification and characterization of variable-number tandem repeats in the *Yersinia pestis* genome. *J. Clin. Microbiol.*, 2001, **39**, 3179–3185.
19. Peak, I. R. A., Jennings, M. P., Hood, D. W., Bisercic, M. and Moxon, E. R., Tetrameric repeat units associated with virulence factor phase variation in *Haemophilus* also occur in *Neisseria* spp. and *Moraxella catarrhalis*. *FEMS Microbiol. Lett.*, 1996, **137**, 109–114.
20. Saunders, N. J., Jeffries, A. C., Peden, J. F., Hood, D. W., Tettelin, H., Rappuoli, R. and Moxon, E. R., Repeat associated phase variable genes in the complete genome sequence of *Neisseria meningitidis* strain MC58. *Mol. Microbiol.*, 2000, **37**, 207–215.
21. Davis, C. L. *et al.*, Numerous length polymorphisms at short tandem repeats in human cytomegalovirus. *J. Virol.*, 1999, **73**, 6265–6270.
22. Falk, K., Gratama, J., Rowe, M., Zou, J., Khanim, F., Young, L., and Oosterveer, M., The role of repetitive DNA sequences in the size variation of Epstein–Barr virus (EBV) nuclear antigens and the identification of different EBV isolates using RFLP and PCR analysis. *J. Gen. Virol.*, 1995, **76**, 779–790.
23. Perdue, M. L., Garcia, M., Senne, D. and Fraire, M., Virulence-associated sequence duplication at the hemagglutinin cleavage site of avian influenza viruses. *Virus Res.*, 1997, **49**, 173–186.
24. Bensen, G., Tandem repeats finder: A program to analyse DNA sequences. *Nucleic Acids Res.*, 1999, **27**, 573–580.
25. Richards, R. I. and Sutherland, G. R., Dynamic mutation: possible mechanism and significance in human disease. *Trends Biochem. Sci.*, 1997, **22**, 432–436.
26. Adair, D. M., Worsham, P. L., Hill, K. K., Klevytska, A. M., Jackson, P. J., Friedlander, A. M. and Keim, P., Diversity in a variable number

tandem repeat from *Yersinia pestis*. *J. Clin. Microbiol.*, 2000, **38**, 1516–1519.
27. Weiser, J. N., Love, J. M. and Moxon, E. R., The molecular mechanism of phase variation of *Haemophilus influenzae* lipopolysaccharide. *Cell*, 1989, **59**, 657–665.
28. van Ham, S. M., van Alphen, L., Mooi, F. R. and van Putten, J. P. M., Phase variation of *Haemophilus influenzae* transcriptional control of two divergent genes through a variable combined promoter region. *Cell*, 1993, **73**, 1187–1196.
29. Moxon, E. R., Rainey, P. B., Nowak, M. A. and Lenski, R. E., Adaptive evolution of highly mutable loci in pathogenic bacteria. *Curr. Biol.*, 1995, **4**, 24–33.
30. Wongteerasupaya, C., Pungchai, P., Withyachumnarnkul, B., Boonsaeng, V., Panyim, S., Flegel, T. W. and Walker, P. J., High variation in repetitive DNA fragment length for White spot syndrome virus (WSSV) isolates in Thailand. *Dis. Aquat. Org.*, 2003, **54**, 253–257.
31. Marks, H., Goldbach, R. W., Vlak, J. M. and van Hulten, M. C. W., Genetic variation among isolates of White spot syndrome virus. *Arch. Virol.*, 2004, **149**, 673–697.
32. Wellauer, P. K. and David, I. B., Isolation and sequence organization of human ribosomal DNA. *J. Mol. Biol.*, 1979, **128**, 289–303.
33. van Arsdell, S. W. and Weiner, A. M., Human genes for U2 small nuclear RNA are tandemly repeated. *Mol. Cell. Biol.*, 1984, **4**, 492–499.

ACKNOWLEDGEMENTS. This work was supported by grants from Department of Biotechnology, New Delhi under BTIS programme and carried out at the Bioinformatics Centre, College of Fisheries, Mangalore.

Received 19 February 2005; revised accepted 31 May 2005

Assessing the role of *FecB* mutation in productivity of Indian sheep

V. C. Pardeshi[1], M. N. Sainani[1], J. F. Maddox[2], P. M. Ghalsasi[3], C. Nimbkar[3] and V. S. Gupta[1,*]

[1]National Chemical Laboratory, Division of Biochemical Sciences, Pune 411 008, India
[2]University of Melbourne, Centre for Animal Biotechnology, School of Veterinary Science, Victoria 3010, Australia
[3]Nimbkar Agricultural Research Institute, Animal Husbandry Division, Phaltan 415 523, India

DNA samples from four Indian sheep breeds, viz. Garole, Deccani, Bannur and Madras Red were screened by PCR–RFLP to determine the presence of *FecB* mutation in these breeds. The Garole was the only breed, which carried the *FecB* mutation. The *FecB* mutation was introgressed from Garole sheep into Deccani sheep and Bannur sheep, and the performance of the crossbred sheep is being monitored in subsequent generations. Approximately half of the first backcross ewes (progeny of *FecB* heterozygote F1 rams) was found to carry one copy of *FecB* mutation, as expected. The *FecB* PCR–RFLP test was found to be fast, accurate and useful as a tool for making breeding decisions.

Keywords: *FecB* (Booroola) gene, Garole sheep, PCR–RFLP test, prolificacy.

SHEEP occupy a special niche in the Indian agricultural production system and are important for the rural economy. They are efficient converters of otherwise unutilized poor-quality grass and crop residues into meat and skin. Profitable and sustainable sheep production is a high priority for India and genetic improvement of sheep for meat production is one of the important developmental priorities. Enhancing reproductive rate is a logical approach to improving economic efficiency of meat production. However, the litter size of almost all Indian sheep breeds, except the Garole, is low and is thus a major constraint to sheep-meat production. Introduction of a prolificacy gene into non-prolific sheep breeds having other desired traits may effectively increase the reproductive performance of local breeds. The concept of breeding for increased fecundity has been accepted and adapted in ruminant breeding[1]. However, the conventional method of selective breeding to increase the reproductive performance of local sheep breeds would be a slow process, since the heritability of litter size is typically low and there is not much variation to exploit in the non-prolific breeds.

The highly prolific 'Booroola', a strain of Merino sheep, developed in Australia was selected solely on the lambing performance of ewes[2]. Ovulation rate analysis of this breed further provided strong evidence for single gene control of the trait[3]. The Booroola fecundity gene, *FecB*, is a single autosomal mutation that shows additive effect for ovulation rate and partial dominance for litter size[4,5]. In early 2001, Wilson *et. al.*[6] identified the mutation that causes super prolificacy of Booroola, which was confirmed by groups in France and the UK[7,8]. The mutation was found in the bone morphogenetic receptor type 1B (*BMPR1B*) gene. It is hypothesized that this mutation might be reducing the signalling through the receptors of granulosa cells. The *BMPR1B* gene in humans is located in a region of chromosome 4 that contains homologues of genes that have been shown to be linked with the Booroola mutation[8]. Recently, it has been shown that the Garole possesses the same mutation as the Booroola[9], supporting the theory put forth by Turner[2] that the Bengal sheep that arrived in Australia in 1792 might be the probable source of *FecB* mutation in Booroola.

The Garole is the only known prolific sheep breed in India and is a native of Sunderban, the swampy delta region of the Ganga river in West Bengal[10]. This breed is mainly reared for meat production. However, it has low growth rate and body weight and poor survival in harsh semi-arid environment[11]. Therefore, three Indian sheep breeds, viz. Deccani, Bannur and Madras Red with more desirable

*For correspondence. (e-mail: vs.gupta@ncl.res.in)