Verschoor, E.J., Warren, K.S., Langenhuijzen, S., Heriyanto,, Swan, R.A. and Heeney, J.L. (2001) *Analysis of two genomic variants of orang-utan hepadnavirus and their relationship to other primate hepatitis B-like viruses.* Journal of General Virology, 82 (4). pp. 893-897.

Analysis of two genomic variants of orang-utan hepadnavirus and their relationship to other primate hepatitis B-like viruses

Ernst J. Verschoor[1], Kristin S. Warren[2,3], Susan Langenhuijzen[1], Heriyanto[4], Ralph A. Swan[2] and Jonathan L. Heeney[1]

Department of Virology, Biomedical Primate Research Centre (BPRC), PO Box 3306, 2280 GH Rijswijk, The Netherlands[1]

Division of Veterinary and Biomedical Sciences, Murdoch University, Western Australia[2]

Wanariset Orang-utan Reintroduction Centre, East Kalimantan, Indonesia[3]

Department of Animal Quarantine, East Kalimantan, Indonesia[4]

Author for correspondence: Ernst Verschoor. Fax +31 15 284 3986. e-mail verschoor@bprc.nl

Abstract

We recently described orang-utan hepadnavirus (OuHV) (Warren *et al.*, *Journal of Virology*, 73, 7860–7865, 1999). Phylogenetic analyses indicated that the various isolates of OuHV can be divided into two genomic variants. Two representatives from each genomic cluster were analysed both molecularly and phylogenetically. Their genome organization was highly similar to other hepadnaviruses of apes and humans. The complete genome sequences of the two OuHV types had an overall 5% sequence difference. Research on 25 seropositive Bornean orang-utans showed that, of the 19 animals infected with one variant, 12 originated from East Kalimantan. Phylogenetic analysis was performed using the full-length genomes of various primate hepadnaviruses. The tree topology revealed one cluster of Old World hepadnaviruses that is divided into two subclusters, one consisting of the ape viruses, and the other comprising the human genotypes A–E. These data suggest that the great apes and gibbons have been infected with a common ancestor hepadnavirus.

Hepatitis B virus (HBV) is a member of the family *Hepadnaviridae*, which consists of the avian hepatitis B-type viruses (avihepadnaviruses) and the orthohepadnaviruses, which infect mammals. In addition to the HBVs that infect humans, the orthohepadnaviruses include viruses which infect rodents such as ground squirrels (GSHV) (Marion *et al*., 1980) and woodchucks (WHV) (Summers *et al*., 1978), as well as hepatitis B-like viruses that infect nonhuman primates, like chimpanzees (ChHBV), gibbons (GiHBV), woolly monkeys (WMHBV), gorillas (GoHBV) and orang-utans (OuHV) (Grethe *et al*., 2000; Hu *et al*., 2000; Lanford *et al*., 1998, 2000; MacDonald *et al*., 2000; Norder *et al*., 1996; Takahashi *et al*., 2000; Vaudin *et al*., 1988; Warren *et al*., 1999).

Currently identified nonhuman primate hepadnaviruses are highly similar to the human HBVs in genome organization, and may cause acute or chronic hepatitis in their natural hosts (Lanford *et al*., 1998), or when passaged to other species (Norder *et al*., 1996). Previously, we described hepadnaviruses that occur naturally in Bornean orang-utans (*Pongo pygmaeus pygmaeus*), now designated orang-utan hepadnavirus (OuHV), by sequence analysis of the S gene (Warren *et al*., 1998, 1999). Here we describe the molecular and evolutionary analysis of full-length genomes of two variants of OuHV present in the Bornean orang-utan population.

Phylogenetic analysis of the small S gene of OuHV isolates from seven chronic carriers suggested that OuHV could be classified in two genomic groups (Warren *et al*., 1999). Viruses obtained from the orang-utans Doel, O'on, Papa and Mojo, and those infecting Romeo, Somad and Lisa, were found on two different branches in the OuHV cluster. Both branches were supported with high bootstrap values of 84 and 100%, respectively.

The division of HBV into genomic groups (A to F) has been based on the definition that the difference between the complete genomes of viruses from different genotypes must be 8% or more (Okamoto *et al*., 1988). To solve the question whether Bornean orang-utans are infected by two distinct genotypes of OuHV or merely by different genomic variants, we determined the sequence of the complete genome of two representative viruses, one from each OuHV subcluster (Somad and Papa).

The isolation of viral genomic DNA, PCR amplification of subgenomic fragments and sequence analysis were performed essentially as described previously (Warren *et al*., 1999). The primers used for PCR and sequence analysis of the OuHV genome and their positions on the chimpanzee HBV genome are given in Table 1. Sequencing of the cloned insert was performed using pUC/M13 forward and reverse sequencing primers, which bind on either side of the cloned insert, and with HBV-specific oligonucleotides. From each amplification reaction three clones were analysed.

The full-length genomes obtained from Somad and Papa were both 3182 nucleotides in length. Their genomic organization was similar to other hepadnaviruses described in humans and apes. Direct comparison of the two OuHV genomes revealed 95% similarity which, according to the HBV genotype definition, is too high to raise the different OuHV types to

the genotype level. Rather, we are dealing with genomic variants as have been described for GiHBV isolates also (Grethe *et al*., 2000).

Table 1. Primers used for amplification of subgenomic OuHV fragments and sequence analysis

Primer	Sequence (5′ to 3′)	Position*
OU 1	AACCCYRCMCCRAAYGCTCC	1095–1114
OU 2	CTGTWCARCATCTTGARTCC	2049–2068
OU 3	CCTTTACCCCGTTGCYCGGC	2422–2441
OU 4	GCATGGMGACCACCGTGAACGC	2888–2909
OU 5	CCTGGGTGGGHRDTAATTTRGAAG	209–232
OU 6	GGATTCAGCGCCGACGGGACG	2710–2730
OU 7	CAWGAGATGAYTAGGCAGAGG	3113–3130
OU 8	CGGAAGTGTTGATAAGATAG	413–432
OU 9	TCTGCGACGCGGCGATTGAGA	502–522
OU 10	GACWCAYAAGGTGGGRAAYTTTAC	562–585
OU 11	GTGACAGAAAGGTTCTGC	953–971
Hep B1	TTCAAGCCTCCAAGCTGTGCCTTGG	3145–3169

* Position on chimpanzee HBV isolate LSH (GenBank accession no. D00220). In addition to HBV/OuHV-specifc primers M13 forward and reverse sequencing primers were used to sequence the full-length OuHV genomes.

The Bornean orang-utan population consists of several geographically isolated and genetically different subpopulations (Groves *et al*., 1992 ; Rijksen & Meijaard, 1999; Warren *et al*., 2001; Zhi *et al*., 1996). Co-evolution of OuHV with their hosts may have caused their divergence into at least two genomic variants. To test this hypothesis we examined the geographical distribution of both variants in the Indonesian part of Borneo (Kalimantan) by using a combined PCR–restriction fragment length polymorphism (PCR–RFLP) assay to discriminate between infections with the genomic types (Fig. 1). Sera from 25 viraemic orang-utans were screened. The majority of OuHV infections (n=19) could be classified as belonging to the Somad cluster, while the remaining infections were caused by viruses closely related to the Papa isolate. All OuHV-positive animals from East Kalimantan (n=12) were viruses closely related to Somad, while the remaining animals from other geographical origins (n=13) showed an almost equal distribution over both clusters (7/6). These findings suggest a correlation with geographical origin or subpopulation of the host and the virus variants. Such a correlation has also been described by Grethe *et al*. (2000) for GiHBVs originating from Vietnam and Thailand. However, too few OuHV isolates from the various subpopulations have been examined to draw more definitive conclusions. Identification of additional viraemic orang-utans of known origin, together with development of a serological test capable of distinguishing between OuHV types, may be of assistance in further defining this issue in the future.

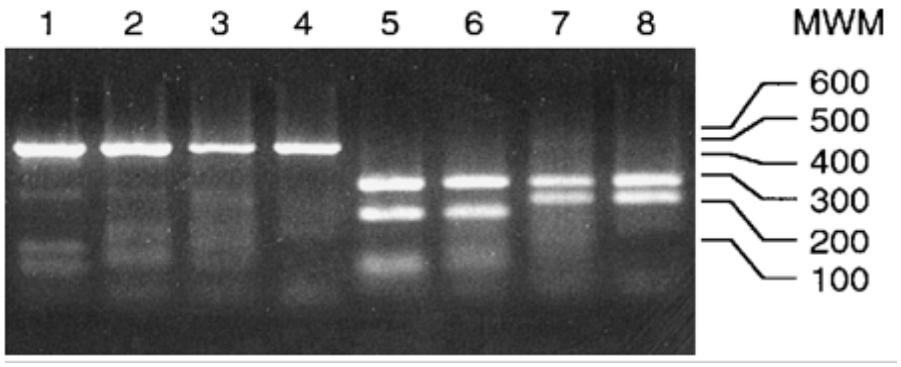

Fig. 1. Discrimination between OuHV genomic variants by using restriction enzyme analysis of pre-S region PCR fragments. A 458 bp fragment encompassing the 5′ end of the OuHV pre-S sequence was amplified using primers hepB-SF1 and hepB-SRout as described previously (Warren *et al*., 1999). The PCR fragments were purified from agarose gel with a QIAquick gel extraction kit for restriction enzyme digestion using the restriction enzyme *Bst*NI. Digestion of PCR fragments of viruses belonging to the Somad cluster generated fragments of 198 and 260 bp, while viruses belonging to the Papa cluster PCR yielded three fragments of 42, 156 and 260 bp, respectively. Lanes 1 to 4, undigested PCR fragments of Papa, O'on ('Papa cluster'), Somad and Romeo ('Somad cluster'), respectively. Lanes 5 to 8, as 1 to 4 but digested with *Bst*NI. A molecular mass marker is indicated on the right.

Recently, complete genomes of hepadnaviruses infecting gibbons, chimpanzees and a gorilla have become available (Grethe *et al*., 2000; Lanford *et al*., 2000; MacDonald *et al*., 2000; Takahashi *et al*., 2000). Sequence similarities between genomes of these viruses range from 89 to 97% (data not shown). We performed a phylogenetic analysis of 29 full-length genomic sequences of primate hepadnaviruses, including the OuHV genomes. Multiple sequence alignments were created using the Se-Al Sequence Alignment Editor program version 1.0 alpha 1 (Rambaut, 1995), and deletions were removed manually. Phylogenetic analysis of nucleotide sequences was performed using the PHYLIP software package, version 3.572 (Felsenstein, 1995), and PAUP 4.0b2a (Swofford, 1999). Distances between the aligned sequences were determined using the Kimura two-parameter algorithm, and the trees were visualized using NJplot (Perrière & Gouy, 1996) (Fig. 2).

A close phylogenetic relation exists between all hepadnaviruses isolated from the different ape species – chimpanzees, gorillas, gibbons and orang-utans. This relationship was well supported by a 76% bootstrap value. From the branch leading to the ape viruses, the African and Asian ape viruses split into two separate clusters. Both clusters are highly supported with a 100% bootstrap value. The human HBV genotypes A to E form another large cluster of hepadnaviruses. Analysis using maximum parsimony gave an essentially similar tree topology (data not shown). An exception was ch195, a chimpanzee virus that is found on the other main branch leading to the human HBV. On this branch ch195 clusters with the African genotype E viruses. However, the exact origin of this virus is unclear; it may be of human origin (Takahashi *et al*., 2000).

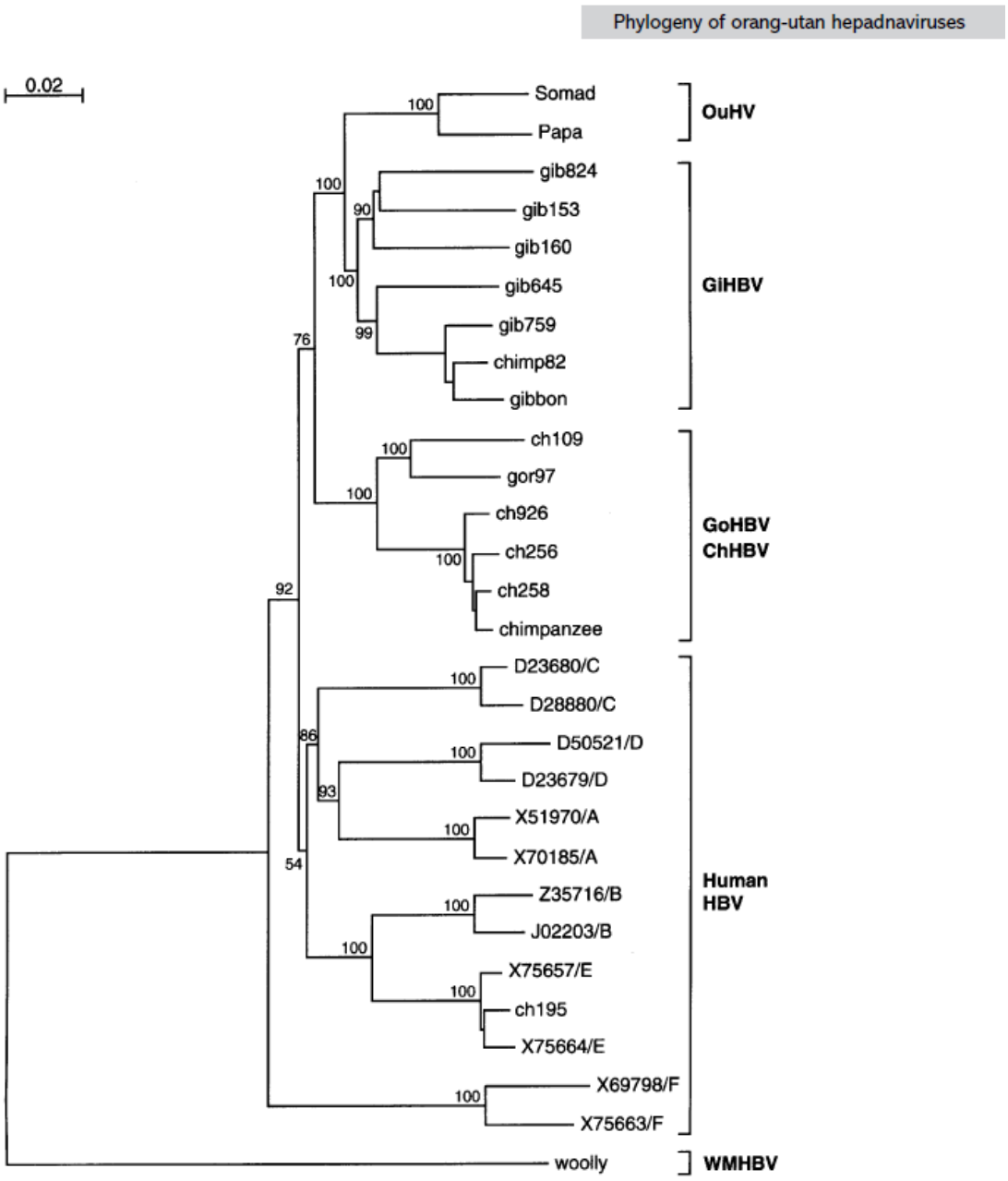

Fig. 2. Phylogenetic relationship of OuHV with other primate hepadnaviruses based on the analysis of full-length genomes. Distances were determined using the Kimura two-parameter method. Trees were constructed using the neighbour-joining method. The numbers at the nodes indicate the percentages probability after analysis of 1000 bootstrap re-samplings. The horizontal bar indicates number of base substitution per site. The nucleotide sequences of the Somad and Papa complete genomes have been deposited in the EMBL/GenBank data libraries, accession nos AF193863 and AF193864. EMBL/GenBank database accession nos are as follows: for GiHBV, U46935,AJ131568, AJ131569, AJ131571–AJ131573 and AJ131575; for ChHBV,D00220, AF222322, AF222323, AB032432 and AB032433; for GoHBV,AJ131567; and for WMHBV, AF046996. All human HBV are identified by their GenBank accession no./genotype identification.

Although the number of published genomes of ape hepadnaviruses is limited, certain conclusions may be drawn from our analysis. A distribution based on origin and species can be observed, consisting of the Asian and African ape hepadnaviruses. The former consist of two species-specific subclusters (orang-utan and gibbon viruses), both supported by high bootstrap values. The genomes of OuHV and GiHBV each differ by 8% or more from all other hepadnaviruses (data not shown); thus, based on the definition set by Okamoto *et al*. (1988), OuHV and GiHBV form separate genotypes. The viruses isolated from gorilla and chimpanzees form one African genotype (bootstrap support 100%) comprising the African cluster (93 to 99% sequence similarity). Interestingly, the single GoHBV sequence falls within the chimpanzee viruses and, unlike OuHV and GiHBV, does not form a separate genotype. Accumulation of more gorilla and chimpanzee virus sequences is necessary to further investigate the true relationship between GoHBV and ChHBV.

The ape viruses are closely related to the human HBV genotypes A to E (92% bootstrap support), yet they belong to different subclades which branch off from a single branch. This is indicative of a common ancestor virus from which the ape hepadnaviruses and the HBV genotypes A to E (the 'Old World HBV') have evolved, as has also been suggested by others (Grethe *et al*., 2000; Hu *et al*., 2000; MacDonald *et al*., 2000; Takahashi *et al*., 2000). A more recent cross-species transmission between humans and apes as proposed by Lanford *et al*. (1998, 2000) seems highly unlikely as this would have resulted in a mosaic branching pattern with human and ape viruses intermingled on the tree. In addition, the finding of closely related hepadnaviruses among ape species of Africa and Asia, as well as the cluster of species-specific hepadnaviruses within the Asian continent, strongly suggests an ancient hepadnavirus infection of the apes that may have been spread by multiple cross-species transmissions amongst these species. The evolution of different viral variants within one host species as described here for OuHV, and by Grethe *et al*. 2000 for GiHBV, also implies a much older infection event rather than a recent zoonotic event.

A question which remains is the relation of the HBV genotype F viruses and the virus carried by the South American woolly monkey with the other hepadnaviruses (Fig. 2). Several authors suggest that the F genotype viruses and the woolly monkey virus represent even older viruses (Arauz-Ruiz *et al*., 1997; Lanford *et al*., 1998; Norder *et al*., 1994, 1996). A more exhaustive search for hepadnavirus infections in nonhuman primates from all continents will be necessary to determine whether other primates harbour hepadnaviruses and at what prevalence, and how they differ from the ones already published. Only then it may be possible to test or verify our hypothesis and those put forward by others concerning the origin of primate hepadnaviruses and their evolutionary relationships.

Acknowledgements

We would like to thank the technicians at the Wanariset Orang-utan Reintroduction Project for their assistance in sample collection and the staff of the Wanariset I Samboja Research Station, especially Ir A. Susilo, Dr Ir W. Smits and Mr T. de Kam. Thanks also to the Indonesian Institute for Scientific Research for enabling research at the Centre. We gratefully acknowledge financial support provided by BHP Minerals, Melbourne, PT Kaltim Prima

Coal, CRA Foundation, PT Kelian Equatorial Mining and the Merck Foundation. The support and technical assistance of Nel Otting, Natasja de Groot and colleagues at the Biomedical Primate Research Centre and staff at the Division of Veterinary and Biomedical Sciences, Murdoch University, Australia, is greatly appreciated. The Department of Virology of the BPRC is committed to the conservation of nonhuman primates in their natural habitat.

E.J.V. and K.S.W. contributed equally to the data described in this article.

References

Arauz-Ruiz, P., Norder, H., Visona, K. A. & Magnius, L. O. (1997). Molecular epidemiology of hepatitis B virus in Central America reflected in the genetic variability of the small S gene. *Journal of Infectious Diseases* **176**,851-858.

Felsenstein, J. (1995). PHYLIP: Phylogeny Inference Package (version 3.572). Distributed by the author. University of Washington, Seattle, WA, USA.

Grethe, S., Heckel, J.-O., Rietschel, W. & Hufert, F. T. (2000). Molecular epidemiology of hepatitis B virus variants in non-human primates.*Journal of Virology* **74**, 5377-5381.

Groves, C. P., Westwood, C. & Shea, B. T. (1992). Unfinished business: mahalanobis and a clockwork orang. *Journal of Human Evolution* **22**, 327-340.

Hu, X., Margolis, H. S., Purcell, R. H., Ebert, J. & Robertson, B. H.(2000). Identification of hepatitis B virus indigenous to chimpanzees.*Proceedings of the National Academy of Sciences, USA* **97**, 1661-1664.

Lanford, R. E., Chavez, D., Brasky, K. M., Burns, R. B. & Rico-Hesse, R. (1998). Isolation of a hepadnavirus from the woolly monkey, a New World primate. *Proceedings of the National Academy of Sciences, USA* **95**, 5757-5761.

Lanford, R. E., Chavez, D., Rico-Hesse, R. & Mootnick, A. (2000). Hepadnavirus infection in captive gibbons. *Journal of Virology* **74**, 2955-2959.

MacDonald, D. M., Holmes, E. C., Lewis, J. C. & Simmonds, P.(2000). Detection of hepatitis B virus infection in wild-born chimpanzees (*Pan troglodytes verus*): phylogenetic relationships with human and other primate genotypes. *Journal of Virology* **74**, 4253-4257.

Marion, P. L., Oshiro, L. S., Regnery, D. C., Scullard, G. H. & Robinson, W. S. (1980). A virus in Beechey ground squirrels that is related to hepatitis B virus of humans. *Proceedings of the National Academy of Sciences, USA* **77**, 2941-2945.

Norder, H., Courouce, A. M. & Magnius, L. O. (1994). Complete genomes, phylogenetic relatedness, and structural proteins of six strains of the hepatitis B virus, four of which represent two new genotypes. *Virology* **198**,489-503.

Norder, H., Ebert, J. W., Fields, H. A., Mushahwar, I. K. & Magnius, L. O. (1996). Complete sequencing of a gibbon hepatitis B virus genome reveals a unique genotype distantly related to the chimpanzee hepatitis B virus.*Virology* **218**, 214-223.

Okamoto, H., Tsuda, F., Sakugawa, H., Sastrosoewignjo, R. I., Imai, M., Miyakawa, Y. & Mayumi, M. (1988). Typing hepatitis B virus by homology in nucleotide sequence: comparison of surface antigen subtypes.*Journal of General Virology* **69**, 2575-2583.

Perrière, G. & Gouy, M. (1996). WWW-Query: an on-line retrieval system for biological sequence banks. *Biochemie* **78**, 364-369.

Rambaut, A. (1995). SE-AL sequence alignment program V1.D1. Dept of Zoology, University of Oxford, Oxford, UK.

Rijksen, H. D. & Meijaard, E. (1999). *Our Vanishing Relative. The Status of Wild Orang-utans at the Close of the Twentieth Century*. Dordrecht: Kluwer Academic Publishers.

Summers, J., Smolec, J. M. & Snyder, R. (1978). A virus similar to human hepatitis B virus associated with hepatitis and hepatoma in woodchucks.*Proceedings of the National Academy of Sciences, USA* **75**, 4533-4537.

Swofford, D. L. (1999). PAUP*: Phylogenetic Analysis Using Parsimony (*and other methods), version 4.0b4, 4.0 edn. Sunderland, USA: Sinauer Associates.

Takahashi, K., Brotman, B., Usuda, S., Mishiro, S. & Prince, A. M.(2000). Full-genome sequence analyses of hepatitis B virus (HBV) strains recovered from chimpanzees infected in the wild: implications for an origin of HBV. *Virology* **267**, 58-64.

Vaudin, M., Wolstenholme, A. J., Tsiquaye, K. N., Zuckerman, A. J. & Harrison, T. J. (1988). The complete nucleotide sequence of the genome of a hepatitis B virus isolated from a naturally infected chimpanzee. *Journal of General Virology* **69**, 1383-1389.

Warren, K. S., Niphuis, H., Heriyanto, Verschoor, E. J., Swan, R. A. & Heeney, J. L. (1998). Seroprevalence of specific viral infections in confiscated orangutans (*Pongo pygmaeus*). *Journal of Medical Primatology* **27**, 33-37.

Warren, K. S., Heeney, J. L., Swan, R. A., Heriyanto & Verschoor, E. J.(1999). A new group of hepadnaviruses naturally infecting orangutans (*Pongo pygmaeus*). *Journal of Virology* **73**, 7860-7865.

Warren, K. S., Verschoor, E. J., Langenhuijzen, S., Heriyanto, Swan, R. A., Vigilant, L. & Heeney, J. L. (2001). Speciation and intra-subspecific variation of Bornean orangutans, *Pongo pygmaeus pygmaeus. Molecular Biology and Evolution* (in press).

Zhi, L., Karesh, W. B., Janczewski, D. N., Frazier-Taylor, H., Sajuthi, D., Gombek, F., Andau, M., Martenson, S. & O'Brien, S. J. (1996). Genomic differentiation among natural populations of orang-utan (*Pongo pygmaeus*). *Current Biology* **6**, 1326-1336.