

A preliminary analysis of the population genetics and molecular phylogenetics of *Onchocerca volvulus* (Nematoda: Filarioidea) using nuclear ribosomal second internal transcribed spacer sequences

Ramiro Morales-Hojas^{*/+}, Robert A Cheke^{*}, Rory J Post^{**}

Laboratório de Evolução Molecular, Instituto de Biologia Molecular e Celular, Universidade do Porto, Rua do Campo Alegre 823, 4150-180 Porto, Portugal ^{*}Natural Resources Institute, University of Greenwich at Medway, Kent, UK ^{**}Department of Entomology, The Natural History Museum, London, UK and Department of Infectious & Tropical Diseases, London School of Hygiene & Tropical Medicine, London, UK

Nuclear internal transcribed spacer 2 (ITS2) rDNA sequences were used for a molecular phylogenetics analysis of five Onchocerca species. The sister species of the human parasite O. volvulus was found to be the cattle parasite O. ochengi and not O. gibsoni, contrary to chromosomal evidence. The genetic differentiation of two African populations (representing the two African strains) and a Brazilian population of O. volvulus was also studied. Phylogenetic and network reconstruction did not show any clustering of ITS2 alleles on geographic or strain grounds. Furthermore, population genetics tests showed no indication of population differentiation but suggested gene flow among the three populations.

Key words: *Onchocerca volvulus* - onchocerciasis - strains - molecular phylogenetics - population differentiation

The parasite *Onchocerca volvulus* (Nematoda: Filarioidea) is the causative agent of onchocerciasis (river blindness), a major public health problem in tropical Africa and in isolated foci in Yemen, South and Central America. There is evidence for different pathogenicity between “strains” from the savanna and the forest of West Africa (Duke 1981) with reported levels of blindness among infected people in the savanna being higher than in the forest (up to a maxima of 15% compared to 2%, respectively) and in the forested areas severe ocular complications such as sclerosing keratitis are mostly absent, with dermatological symptoms predominant. There are also important differences in the competence of different vector species to support the different “strains” (Duke 1981), and geographic variation in morphology (Eichner & Renz 1990), isoenzymes (Cianchi et al. 1985), endosymbionts (Higazi et al. 2005), and O-150 repetitive DNA (Zimmerman et al. 1992) has been interpreted as correlating with forest/savanna “strain” designation. Furthermore, the O-150 repetitive DNA suggested that the New World populations of *O. volvulus* were derived from African savanna populations (Zimmerman et al. 1994). However, the morphological and isoenzymatic data were never followed-up, and subsequent data on the O-150 repetitive DNA has not given such a clear correlation in Africa (Fischer et al. 1996). The nuclear ITS rDNA sequence has been extensively used to define genetic markers for different species of nematodes (Powers et al. 1997), and the aims of this

study were to use the ITS2 to investigate the levels of genetic variation within *O. volvulus* and determine whether it is related to strain differentiation. This is information of importance for understanding the epidemiology of onchocerciasis and the design of control programmes. Thus, if there are normal levels of gene flow between populations (or strains), these populations are likely to be similar in their overall genetic characteristics (and response to control), genes for important characteristics (such as filaricide resistance) are likely to spread easily, and disease control at single locations or targeted at a single strain will be countered by immigration.

ITS2 rDNA sequences were already available from Brazilian and African forest samples of *O. volvulus* (Morales-Hojas et al. 2001, GenBank accession numbers AF228565-AF228570 and AF228572-AF228576). For samples of savanna *O. volvulus* (Kati district, Mali), *O. ochengi* (Bamako, Mali), *O. lienalis* (Shrewsbury, UK), *O. gutturosa* (Bamako, Mali) and *O. gibsoni* (Australia), amplification of the ITS2 was carried out using primers 5.8SOv-A 5-AGCGGTGGATCACTTGGCTC-3 (located in position 9-29 of the 5.8S gene) and NC2 5-TTAGT-TTCTTTTCCTCCGCT-3 (located 70 bp downstream of the 5 end of the 28s rRNA gene). The origin of this material is described in Morales-Hojas et al. (2006). Reactions were performed in 1x Buffer (Promega), 3 mM MgCl₂ (Promega), 200 μM of each dNTP (Promega), 0.2 μM of each primer, 1 unit of *Taq* DNA Polymerase (Promega) and 1.5 ml of the DNA extraction. Amplifications consisted of a first denaturation step at 94°C for 2 min, followed by 30 cycles of 30 s at 94°C, 30 s at 50°C (annealing) and 30 s at 72°C, with a final extension step of 5 min at 72°C. Amplification products were extracted directly from the gel using the GeneClean II kit (Anachem) and cloned using the TOPO TA Cloning Kit for Sequencing (Invitrogen). Positive clones were extracted from transformed bacteria using the S.N.A.P.

⁺ Corresponding author: rmhojas@ibmc.up.pt
Received 24 April 2007
Accepted 3 October 2007

MiniPrep kit (Invitrogen) and sent to the Advanced Biotechnology Centre, Imperial College School of Medicine, for sequencing in both directions, to correct for possible reading errors, using Big Dye v3.1 chemistry (Perkin Elmer/Applied Biosystems) and the universal primers T3 and T7. It is possible that the sequences include PCR-induced errors. Nevertheless, these would be difficult to correct for because the DNA used is a pool of genomes of the different individuals and thus amplification products are not expected to be homogeneous. Furthermore, the ribosomal genes are a family of tandemly repeated units that despite concerted evolution do not necessarily show complete homogenisation. However, the expected number of errors introduced by non-proofreading polymerases is very small, in the range of 1×10^{-4} to 10^{-5} errors per bp.

The new ITS2 sequences have been deposited in GenBank and have accession numbers DQ317639-DQ317666. Amplification products showed the expected size, ranging between 533 bp and 568 bp. The length of the ITS2 ranged between 318-319 bp in *O. lienalis*, 333-337 bp in *O. gutturosa*, 341-351 bp in *O. gibsoni*, 347-352 bp in *O. ochengi*, and 340-352 bp in *O. volvulus*.

Phylogenetic results using both Maximum Likelihood and Maximum Parsimony resulted in similar trees (Fig. 1). They supported the mitochondrial analyses that placed *O. ochengi* and *O. volvulus* as sister species and not *O. gibsoni* and *O. volvulus*, as suggested by karyotype evolution data (Post 2005). It is interesting to note that *O. ochengi* ITS2 alleles are paraphyletic with respect to *O. volvulus*, which is consistent with a recent cladogenesis of *O. volvulus* after the spread of domestic cattle in Africa (5000 to 2500 years ago) as discussed by Morales-Hojas et al. (2006). However, this is inconsistent with mitochondrial DNA analyses which resulted in a tree where *O. volvulus* and *O. ochengi* showed reciprocal monophyly (Morales-Hojas et al. 2006). This difference might be explained by the faster rate of coalescence of mitochondrial alleles (four times faster than that of the nuclear alleles; Ballard & Whitlock 2004). The evolutionary relationships between *O. volvulus* alleles were further explored with statistical parsimony networks (Templeton et al. 1992). The reconstructed network showed allele A (from the forest strain) as the most recent common ancestral allele (Fig. 2), and branches in a star-like shape with no geographical or "strain" clustering of sequences. Consistent with these results, population differentiation estimates F_{st} , K_{st}^* and S_{nn} [estimated with DnaSP 4.00 (Rozas et al. 2004) and Proseq 2.9 (Filatov 2002)] were all non-significant indicating no genetic differentiation in the ITS2 sequences between the three populations (Table I). Furthermore, the estimated number of migrant individuals per generation based on F_{st} values was higher than 1, which would be sufficient to prevent population differentiation.

The high haplotype diversity (H) observed, combined with the low sequence diversity (π) in the sample of *O. volvulus* (Table II), is often considered indicative of a population expansion after a reduced ancestral effective population size (Avice 2000). This would be consistent with a host switch from cattle to humans in which only a

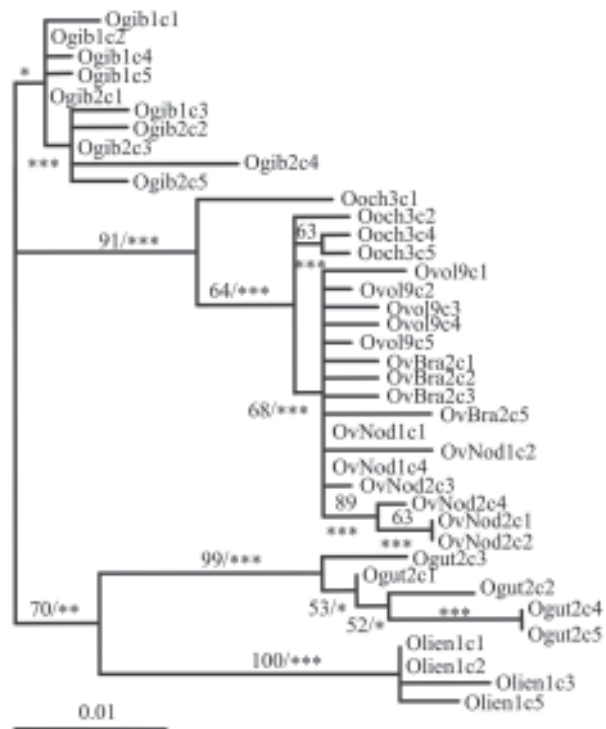


Fig. 1: maximum likelihood tree rooted using the midpoint rooting method obtained using internal transcribed spacer 2 rDNA sequences. The model found to best fit our data by Modeltest v3.7 (Posada & Crandall 1998) was the TVM+G model (nucleotide frequencies: A 0.3617, C 0.1444, G 0.1720 T 0.3219; substitution rate matrix: A-C 1.7337, A-G 7.3786, A-T 3.6829, C-G 1.1153, C-T 7.3786, and G-T 1; gamma distribution shape parameter 0.6225; proportion invariable sites 0). Values on branches indicate the bootstrap support and the asterisks the zero branch length test support (***) represents $p < 0.001$; ** $p < 0.01$; and * $p < 0.05$). Ogib stands for *Onchocerca gibsoni*, Ooch for *O. ochengi*, Ovol for savanna samples of *O. volvulus*, OvBra for *O. volvulus* of Brazilian origin, OvNod for the forest *O. volvulus*, Ogut for *O. gutturosa* and Olien for *O. lienalis*.

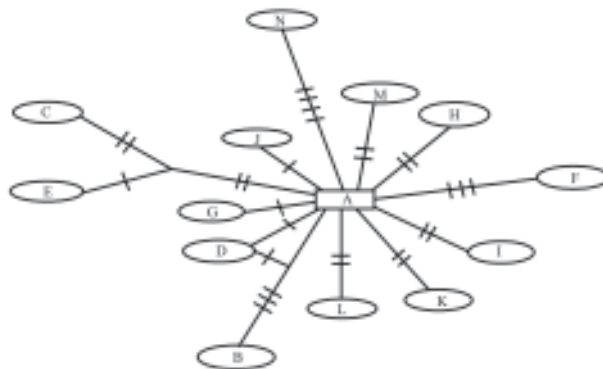


Fig. 2: statistical parsimony network showing the relationship between the internal transcribed spacer 2 alleles of *Onchocerca volvulus*. The network was obtained using Statistical Parsimony as implemented in TCS (Clement et al. 2000). Alignment gaps were treated as missing data. Transversal lines on branches represent the number of mutational events leading to each allele. The central squared box includes the most ancestral allele. Letters denote the different alleles of *O. volvulus*: alleles A to E from the forest sample; alleles F to J from the savanna sample; and alleles K to N of the Brazilian population.

TABLE I
Summary of population differentiation measures of *Onchocerca volvulus* samples at the rDNA.

	Fixed	Shared	<i>Fst</i>	<i>p</i>	<i>Kst</i>	<i>p</i>	<i>Snn</i>	<i>p</i>	Nm ^a
Forest - Savanna	0	0	0.0932	n.s.	0.0277	n.s.	0.5833	n.s.	2.43
Forest - Brazil	0	0	0.0840	n.s.	0.0363	n.s.	0.6363	n.s.	2.73
Savanna - Brazil	0	0	0.0000	n.s.	n.a.	n.s.	0.3703	n.s.	n.a.
3 populations	n.a.	n.a.	0.0564	n.s.	0.0314	n.s.	0.4375	n.s.	4.18

a: number of migrants based on *Fst* values; n.a.: not available; n.s.: not significant.

TABLE II
Summary of variability at the rDNA sequences in each *Onchocerca volvulus* population (forest, savanna and Brazil) and for the combination of the three (total).

Parameters	Forest	Savanna	Brazil	Total
N	7	5	4	16
S	10	9	9	27
h	5	5	4	14
H	0.9047	1	1	0.9833
π	0.0068	0.0064	0.0080	0.0073

N: sequence number; S: number of segregating sites; h: number of alleles; H: allele diversity; π : average number of pairwise nucleotide differences per base pair.

few individuals were involved, thus resulting in a bottleneck. Following the host switch, an increase in population size in the new (human) host must have occurred to give the current millions of infected people.

Although a small number of errors may have been introduced in the sequences due to non-proofreading polymerase mistakes (around 1×10^{-4} to 10^{-5} per bp), the main consequence of these would be to increase the genetic and haplotype diversities. This could affect the interpretation that after a bottleneck the species population has increased its effective size. Nevertheless, despite this possibility of an artificial increase of the diversity the estimated levels of genetic diversity remain low suggesting that the PCR-introduced mutations are few, and not affecting the historical demographic interpretation. Furthermore, this is otherwise supported by the evolutionary history of the parasite, with a recent host-switch and an increase in population size alongside its human host. Other potential artifact of these PCR mistakes could be to increase the number of nucleotide differences between populations. Despite this, still non-significant levels of genetic differentiation between populations were obtained with this marker. Thus, the possible PCR-induced errors introduced in the sequences appear to have little consequences in the interpretation of the results.

In conclusion, the rDNA ITS2 shows no intraspecific structure in *O. volvulus* and the suggested gene flow between the African populations is consistent with what would be expected as a result of the known annual migrations of the West African vector species (Boakye et al. 1998). However, further samples and additional (more

sensitive) nuclear and mitochondrial genes should be investigated in order to corroborate the results presented in this study. This will be of great importance for control strategies. Poor response to the microfilaricidal drug Ivermectin has been reported in some populations of *O. volvulus* (e.g. Awadzi et al. 2004). Results presented here could suggest that resistance genes would spread freely among strains and populations in Africa if resistance to Ivermectin resulted in a re-emergence of parasite transmission at significant levels. Thus, strategies for onchocerciasis control should take into account the fact that populations may not be isolated and that "strain" delimitation is not necessarily indicative of genetic isolation and generalised genetic differentiation.

REFERENCES

- Avice JC 2000. *Phylogeography: The History and Formation of Species*, Harvard University Press, Cambridge, Massachusetts, 447 pp.
- Awadzi K, Boakye DA, Edwards G, Opoku NO, Attah SK, Osei-Atweneboana MY, Lazdins-Helds JK, Ardrey AE, Addy ET, Quartey BT, Ahmed K, Boatman BA, Soumbey-Alley EW 2004. An investigation of persistent microfilaridemia despite multiple treatments with ivermectin, in two onchocerciasis-endemic foci in Ghana. *Ann Trop Med Parasitol* 98: 231-149.
- Ballard JO, Whitlock MC 2004. The incomplete natural history of mitochondria. *Mol Ecol* 13: 729-744.
- Boakye DA, Back C, Fiasorgbor GK, Sib APP, Coulibaly Y 1998. Sibling species distributions of the *Simulium damnosum* complex in the West African Onchocerciasis control Programme area during the decade 1984-93, following intensive larviciding since 1974. *Med Vet Entomol* 12: 345-358.
- Cianchi R, Karam M, Henry MC, Villani F, Kumlien S, Bullini L 1985. Preliminary data on the genetic differentiation of *Onchocerca volvulus* in Africa (Nematoda: Filarioidea). *Acta Trop* 42: 341-351.
- Clement M, Posada D, Crandall K 2000. TCS: a computer program to estimate gene genealogies. *Mol Ecol* 9: 1657-1660.
- Duke BOL 1981. Geographical aspects of onchocerciasis. *Ann Soc Belg Med Trop* 61: 179-186.
- Eichner M, Renz A 1990. Differential length of *Onchocerca volvulus* infective larvae from the Cameroon rainforest and savanna. *Trop Med Parasitol* 41: 29-32.
- Filetov DA 2002. ProSeq: A software for preparation and evolutionary analysis of DNA sequence data sets. *Mol Ecol Notes* 2: 621-624.
- Fischer P, Bamuhiga J, Kilian AHD, Büttner DW 1996. Strain

- differentiation of *Onchocerca volvulus* from Uganda using DNA probes. *Parasitology* 112: 401-408.
- Higazi TB, Filiano A, Katholi CR, Dadzie Y, Remme JH, Unnasch TR 2005. *Wolbachia* endosymbiont levels in severe and mild strains of *Onchocerca volvulus*. *Mol Biochem Parasitol* 141: 109-112.
- Morales-Hojas R, Cheke RA, Post RJ 2006. Molecular systematics of five *Onchocerca* species (Nematoda: Filarioidea) including the human parasite, *O. volvulus*, suggest sympatric speciation. *J Helminthol* 80: 281-290.
- Morales-Hojas R, Post RJ, Shelley AJ, Maia-Herzog M, Coscarón S, Cheke RA 2001. Characterisation of nuclear ribosomal DNA sequences from *Onchocerca volvulus* and *Mansonella ozzardi* (Nematoda: Filarioidea) and development of a PCR-based method for their detection in skin biopsies. *Int J Parasitol* 31: 169-177.
- Posada D, Crandall KA 1998. MODELTEST: testing the model of DNA substitution. *Bioinformatics* 14: 817-818.
- Post RJ 2005. The chromosomes of the Filariae. *Filaria J* 4: 10.
- Powers TO, Todd TC, Burnell AM, Murray PCB, Fleming CC, Szalanski AL, Adams BA, Harris TS 1997. The rDNA internal transcribed spacer region as a taxonomic marker for nematodes. *J Nematol* 29: 441-450.
- Rozas J, Sánchez-DelBarrio JC, Messeguer X, Rozas R 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* 19: 2496-2497.
- Templeton AR, Crandall KA, Sing CF 1992. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics* 132: 619-633.
- Zimmerman PA, Dadzie KY, De Sole G, Remme J, Soumbeay Alley E, Unnasch TR 1992. *Onchocerca volvulus* DNA probe classification correlates with epidemiological patterns of blindness. *J Infect Dis* 165: 964-968.
- Zimmerman PA, Katholi CR, Wooten MC, Lang-Unnasch N, Unnasch TR 1994. Recent evolutionary history of American *Onchocerca volvulus*, based on analysis of a tandemly repeated DNA sequence family. *Mol Biol Evol* 11: 384-392.