

New molecular identifiers for *Simulium limbatum* and *Simulium incrustatum* s.l. and the detection of genetic substructure with potential implications for onchocerciasis epidemiology in the Amazonia focus of Brazil

Priscila A. Conceição^a, James L. Crainey^a, Tatiana P. Almeida^a,
Anthony J. Shelley^b, Sergio L.B. Luz^{a,*}

^a Ecologia de Doenças Transmissíveis na Amazônia, Instituto Leônidas e Maria Deane – Fiocruz Amazônia Rua Terezina, 476. Adrianópolis, CEP: 69.057-070 Manaus, Amazonas, Brazil

^b Department of Entomology, The Natural History Museum, Cromwell Road, London SW7 5BD, UK

ARTICLE INFO

Article history:

Received 29 November 2012

Received in revised form 13 March 2013

Accepted 22 March 2013

Available online xxx

Keywords:

Simulium

S. limbatum

S. incrustatum

DNA barcoding

Taxonomy

Systematics

Cytochrome oxidase 1 (CO1)

Amazonia focus

Onchocerciasis

ABSTRACT

The Amazonia onchocerciasis focus of southern Venezuela and northern Brazil is the larger of the two remaining Latin American onchocerciasis foci where disease transmission still occurs and is often regarded as the most challenging of all the Latin American foci to eliminate onchocerciasis. The site is home to a population of over 20,000 semi-nomadic, hunter-gatherer Yanomami people and is made-up of a mosaic of rainforest and savannah ecologies, which are influenced by the area's undulating terrain and rich geological diversity. At least six blackfly vectors have been implicated in onchocerciasis transmission in this focus; however, because of the difficulty in their routine identification the relative importance of each has been obscured. *Simulium limbatum* and *Simulium incrustatum* s.l. have both been recorded as vectors in the Amazonia focus, but they are difficult to discriminate morphologically and thus the ecological range of these species, and indeed the presence of *S. limbatum* in the Amazonia focus at all, have remained controversial. In the work described here, we report 15 *S. incrustatum* s.l. CO1 sequences and 27 *S. limbatum* sequences obtained from field-caught adult female blackflies collected from forest and savannah localities, inside and just outside the Amazonia focus. Phylogenetic analysis with the sequences generated in this study, showed that both the *S. limbatum* and the *S. incrustatum* s.l. CO1 sequences obtained (even from specimens living in sympatry) all fell into discrete species-specific bootstrap-supported monophyletic groups and thus confirmed the utility of the CO1 gene for identifying both these species inside the Amazonia focus. As the *S. limbatum*-exclusive cluster included CO1 sequences obtained from forest-caught and morphologically identified specimens these results provide the clearest evidence yet of the presence of *S. limbatum* inside the Amazonia focus. The question, however, of whether *S. limbatum* is actually a vector in the focus still remains unanswered as the data presented here also suggest that *S. limbatum* found in the savannahs adjacent to, but outside the Amazonia focus (and which represent the only *S. limbatum* population to be unambiguously incriminated as a host of *Onchocerca volvulus*), are genetically distinct from those living inside the focus. These findings highlight the need for a clearer picture of the vector taxonomy inside the Amazonia onchocerciasis focus.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

Onchocerciasis (more commonly known as river blindness) is a debilitating neglected tropical disease that is estimated to affect, directly, around 37 million people worldwide and to have 120 million people living at risk of being infected by it (Basáñez et al., 2006;

Cheke and Garms, 2013). The disease causes a variety of severe skin conditions, eliciting symptoms including depigmentation, lichenification and intense itching and blindness. It is only present in Africa, the Arabian Peninsula and Latin America. In the 1990s it was estimated that 250,000 people were living with irreversible blindness caused by onchocerciasis infections (1995 WHO report in Post et al., 2011), although more recent estimates suggest that the number of people presently living with blindness caused by the disease could be over 35% higher (Vos et al., 2012). The disease is caused by the filarial worm *Onchocerca volvulus*, which is transmitted by simuliid blackflies.

* Corresponding author. Tel.: +55 92 3656 0363; fax: +55 92 3656 0363.
E-mail address: sergioluz@amazonia.fiocruz.br (S.L.B. Luz).

Following the success of the Onchocerciasis Elimination Program for the Americas, there are now just two foci in Latin America where onchocerciasis transmission is still thought to occur, the Amazonia focus and the Northeast Venezuelan focus (Gustavsen et al., 2011). The larger of these is the Amazonia focus, a vast area spanning large parts of Brazil's Amazonas and Roraima states and the adjacent areas of Venezuela. As well as its size, vegetation and geology, which make logistics at the site extremely difficult, the semi-nomadic lifestyle of the >20,000 Yanomami people who live in the Amazonia focus also contributes to making it, arguably, the more difficult of the remaining Latin American foci for onchocerciasis eradication (Botto et al., 2011; Coelho et al., 1998; Gustavsen et al., 2011; Shelley et al., 1997). It has thus been argued that a more targeted approach may be necessary for the successful elimination of onchocerciasis in the Amazonia focus (Botto et al., 2011).

The site is a mosaic of dense forest and open savannah, both of which contain a wide variety of ecologies influenced by altitude, geology and climate and each supporting a unique simuliid fauna. At least six simuliid species have been incriminated in the transmission of onchocerciasis within the Amazonia focus: *Simulium exiguum* s.l., *Simulium guianense* s.l., *Simulium oyapockense* s.l., *S. roraimense* s.l., *Simulium incrustatum* s.l. and *Simulium limbatum*. Of these, *S. guianense* s.l. is the most important vector and is found in the mountainous forested areas and the plateau upland savannah areas where onchocerciasis is hyperendemic (Botto et al., 2011; Shelley et al., 1997). *S. oyapockense* s.l. is the principal vector in the lowland hypoendemic areas of the site, and is generally regarded not to be of great epidemiological importance (Botto et al., 2011; Shelley et al., 1997). *S. incrustatum* s.l. has been recorded in both forested highland and forested lowland sites within the Brazilian side of the focus and in the hyperendemic and mesodendemic areas on the Venezuelan side of the focus (Botto et al., 2011; Shelley et al., 1997). Recent feeding experiments with this vector have suggested that it may play a more important role in transmission than had been previously thought (Botto et al., 2011; Grillet et al., 2008); however its importance has been obscured by difficulties in discriminating it from *S. limbatum*. While *S. limbatum* has been recorded in savannahs immediately adjacent to the focus and has been shown to be a competent host in experimental infection experiments, it has hitherto never been recorded in the Brazilian part of the focus (Shelley et al., 2010, 1997). On the Venezuelan side of the focus, *O. volvulus* infections of *S. limbatum* have been reported, but in subsequent descriptions these infected simuliids have been described variously as *S. incrustatum* or *S. limbatum* (Vivas-Martinez et al., 1998), *S. incrustatum*/*S. limbatum* (Basáñez et al., 2009) and *S. incrustatum* (Grillet et al., 2008). Questions concerning the distribution and relative importance of these two species for onchocerciasis dynamics of the Amazonia focus thus remain. Indeed, even whether *S. limbatum* is an active vector within the Amazonia focus at all is still uncertain.

Although *S. incrustatum* s.l. and *S. limbatum* can be distinguished reliably by the scutal markings, these characters can easily be misinterpreted if the light source is not correctly positioned as detailed in Shelley et al. (2010, 1997) and thus morphological identifications are unsuitable for reliable routine identifications. Molecular identifications using DNA barcoding (PCR and DNA sequencing of targeted taxonomic identifier genes), in contrast: can be easily done through a commercial service; do not require high-levels of specialist training, and can be readily used to identify all life-stages and sexes of a target organism. DNA barcodes thus offer the potential to provide a low-skill alternative to onchocerciasis vector identification in Latin America. Although many genes have been used for DNA barcoding, the mitochondrial cytochrome oxidase 1 (CO1) gene has emerged as among the most reliable for insect species identification and has already been proven to have a rate

of evolution suitable for resolving a variety of simuliid taxonomy problems (Day et al., 2008, 2010; Hernández-Triana et al., 2012).

In this study we have collected *S. limbatum* specimens from a number of sites in the Roraima state savannahs, which are immediately adjacent to the Amazonia focus and represent the closest known site where a gold standard *S. limbatum* population has been recorded (Shelley et al., 1997). (A gold standard *S. limbatum* population of *S. limbatum* being defined here as a population which has had representatives verified as such by direct comparison with type specimens.) We have then compared these gold standard *S. limbatum* specimens morphologically to material collected in the present work from forested areas inside the Amazonia focus in Brazil where the species had not been previously recorded. CO1-based barcoding has subsequently been used to make unambiguous confirmation of the existence of *S. limbatum* inside the Amazonia onchocerciasis focus and to demonstrate that *S. incrustatum* s.l. and *S. limbatum* collected within the focus can be reliably distinguished using this technique.

2. Methods

2.1. Field study-site selection, blackfly field collection and morphological identifications

Fig. 1 shows a map of the selected field sites. Savannah and forest areas are labelled appropriately, as are the field collection localities. Adult female blackflies were collected from the eight localities indicated, by human landing catches in which blackflies were allowed to land but not to bite volunteers, using an approach similar in detail to those described in Shelley et al. (2000, 1982). The coordinates of the sites where these collections were made are provided in Table 1. The collection zone was selected because it spans areas of both savannah and forest ecologies and includes localities both within and adjacent to the Amazonia onchocerciasis focus. Equally important, the site was chosen because it spanned an area where both unambiguously identified *S. incrustatum* s.l. and *S. limbatum* have previously been collected (Shelley et al., 1997) and an area very close to where a contested report of *S. limbatum* has been made (Basáñez et al., 1988). Anthropophilic females were collected in 100% ethanol and identified to morphospecies using descriptions in Shelley et al. (1997, 2010).

2.2. DNA extraction, PCR amplification and sequencing

DNA was extracted from individual blackfly specimens using QIAGEN DNeasy blood and tissue kits and eluted in 200 µl of TE (10 mM Tris–Cl; 0.5 mM EDTA; pH 9.0). These DNA extracts were then used in separate 50 µl PCR reaction mixes containing: 5 µl of DNA extract; 5 µl of 10× Invitrogen PCR buffer; 0.2 µl of Invitrogen taq DNA polymerase (at 5 u/µl); 2 µl of 50 mM MgCl₂; 1 µl of dNTPs, at a concentration of 10 mM each and 2 µl of each of the forward and reverse primers, from 10 mM stocks. The PCR primers (LC1049: 5'-GGTCAACAAATCATAAAGATATTGG-3' and HC02198: 5'-TAAACTTCAGGGTGACCAAAAAATCA-3') used for CO1 gene amplification were the same as those described by Hebert et al. (2003) as were the PCR thermocycling conditions. PCR amplicons were individually cleaned or gel extracted following the appropriate protocol provided with the Promega SV Gel and PCR clean-up system kit used for this purpose.

Cleaned-up PCR products were Sanger sequenced in both directions using the same primers that were used to amplify them. Briefly, following optical density quantification, PCR products were prepared for sequencing using an ABI BigDye terminator v3.1 kit and protocol. Some of the DNA sequencing for this study was performed using an ABI 3130 DNA genetic analyser situated at the Fundação de Dermatologia e Venearologia, Alfredo da Matta,

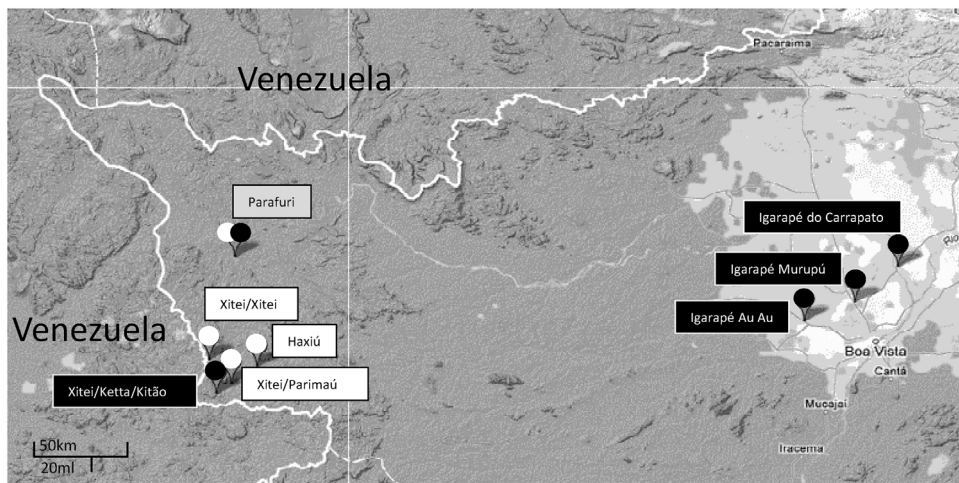


Fig. 1. The location of the eight Brazilian blackfly collection sites used in this study and are listed in Table 1. The map image is an annotated screen-shot taken from google maps which has been converted to grey scale. Light areas around Boa vista represent areas of natural savannah, dark areas indicate primary rain forest. The forested collection sites are all within the Amazonia focus and the savannah sites are all outside of it, with the interface of the two corresponding roughly with the focus's perimeter. Black spots and box-labels indicate sites where *S. limbatum* specimens were collected and white spots and box-labels where *S. incrustatum* specimens were collected; grey box-labels are used to indicate where the species were found in sympatry. A fuller picture of the terrain and vegetation around the collection sites and the whole Amazonia focus can be browsed with the following KML file: XXXXXXXXXXXXX.

Table 1
The site coordinates, species level morphological identifications and sequence accession numbers of each of the CO1 sequences obtained from each of 42 simuliid specimens used in this study. Identical sequences have identical accession numbers and are only discriminated by a suffixed roman numeral.

Collection site	Site co-ordinates	Ecology	Species	Sequence accession
Igarapé do Carrapato, Boa Vista – Roraima	2°58'39.46" N - 60°42'42.16" W	Savannah	<i>Simulium limbatum</i>	HF9484456
			<i>Simulium limbatum</i>	HF9484454iii
			<i>Simulium limbatum</i>	HF9484455i
			<i>Simulium limbatum</i>	HF9484455ii
			<i>Simulium limbatum</i>	HF9484455iii
			<i>Simulium limbatum</i>	HF9484455iv
			<i>Simulium limbatum</i>	HF9484455v
			<i>Simulium limbatum</i>	HF9484455vii
Igarapé Murupú, Boa Vista – Roraima	3°1'11.06" N - 60°46'34.96" W	Savannah	<i>Simulium limbatum</i>	HF9484455x
			<i>Simulium limbatum</i>	HF9484454ii
			<i>Simulium limbatum</i>	HF9484454iv
			<i>Simulium limbatum</i>	HF9484455vi
			<i>Simulium limbatum</i>	HF9484455viii
Igarapé Au Au Boa Vista – Roraima	2°56'24.48" N - 61°3'2.71" W	Savannah	<i>Simulium limbatum</i>	HF9484455viii
			<i>Simulium limbatum</i>	HF9484455vii
			<i>Simulium limbatum</i>	HF9484454i
			<i>Simulium limbatum</i>	HF9484453i
			<i>Simulium limbatum</i>	HF9484453ii
			<i>Simulium limbatum</i>	HF9484453iii
			<i>Simulium limbatum</i>	HF9484453iv
			<i>Simulium limbatum</i>	HF9484453v
Xitei/Ketta/Kitão, Boa Vista – Roraima	2°31'42.9N - 63°55'92" W	Forest	<i>Simulium limbatum</i>	HF9484453vi
			<i>Simulium limbatum</i>	HF9484453vii
			<i>Simulium limbatum</i>	HF9484449
			<i>Simulium limbatum</i>	HF9484451
Xitei/Xitei, Boa Vista – Roraima	2°36'428 N - 63°52'411" W	Forest	<i>Simulium limbatum</i>	HF9484448
			<i>Simulium limbatum</i>	HF9484450
			<i>Simulium limbatum</i>	HF9484455
Xitei/Parimaú, Boa Vista – Roraima	2°35'0.89N - 63°52'6.01" W	Forest	<i>Simulium incrustatum</i>	HF9484445iii
			<i>Simulium incrustatum</i>	HF9484445viii
			<i>Simulium incrustatum</i>	HF9484446iii
			<i>Simulium incrustatum</i>	HF9484446v
			<i>Simulium incrustatum</i>	HF9484447i
Haxiú, Boa Vista – Roraima	2°40'26.38" N - 63°44'7.65" W	Forest	<i>Simulium incrustatum</i>	HF9484447ii
			<i>Simulium incrustatum</i>	HF9484445ii
			<i>Simulium incrustatum</i>	HF9484446iv
			<i>Simulium incrustatum</i>	HF9484446vi
			<i>Simulium incrustatum</i>	HF9484446vii
Parafuri, Boa Vista – Roraima	3°17'0.96" N - 63°50'52.98" W	Forest	<i>Simulium incrustatum</i>	HF9484447iii
			<i>Simulium incrustatum</i>	HF9484445i
			<i>Simulium incrustatum</i>	HF9484446i
			<i>Simulium incrustatum</i>	HF9484446ii
			<i>Simulium incrustatum</i>	HF9484447iv
			<i>Simulium limbatum</i>	HF94844452

Manaus (Amazonas), the rest was done by the Brazilian barcoding of life data (BOLD) research group (in Rio de Janeiro).

2.3. DNA sequence editing, alignment and phylogenetic analysis

Consensus sequences were prepared from matching forward and reverse CO1 sequence trace files using DNA star and then were aligned and edited using freely available software as described previously (Crainey et al., 2009, 2010). BLASTn searches, comparing the acquired sequences to all non-redundant sequence deposits, were used to confirm that all of the obtained sequences were from simuliids and the targeted CO1 gene. Conceptual translation of sequences was carried out using the European Molecular Biology Laboratory's on-line conceptual translation tool 'transeq'. Edited sequence alignments were used to infer phylogenetic trees, using three different methodologies: (1) the Neighbour joining method; (2) Maximum Likelihood method and (3) the Maximum parsimony method. These methodologies were implemented using the PYHLIP software suite, as described previously (Crainey et al., 2009, 2010).

3. Results

3.1. The generation of mitochondrial cytochrome 1 (CO1) DNA sequence from the known *O. volvulus* hosts: *S. incrustatum* s.l. and *S. limbatum*

PCR amplifiable DNA was successfully extracted from a total of 42 specimens, caught in eight different localities (see Table 1). These collection sites were from two different ecological environments: five collection sites we classified as forest and three were classified as savannah (see Fig. 1 and Table 1). Morphological identification of the collected specimens confirmed that 15 of the collected flies were *S. incrustatum* s.l. and that 27 were *S. limbatum* (the provenance of the specimens is provided in Table 1). Good-quality cytochrome oxidase (CO1) gene sequences (both forward and reverse) were obtained from all 42 of these DNA extracts and were used to generate consensus sequences. Unique consensus sequences have been deposited in the sequence databases of the European Molecular Biology Laboratory (EMBL) and the US National Center of Biotechnology Information (NCBI); the sequences have been assigned the following accession numbers: HF948445–HF948456.

3.2. CO1 barcoding is a novel tool for discriminating Amazonia focus *S. limbatum* and *S. incrustatum*

Databases, Blastn, searches with each of the 42 consensus nucleotide sequences confirmed that their closest database relatives were all from the targeted CO1 DNA sequence region and that the closest sequences (for which source species identification was available) were all derived from the genus *Simulium*. Conceptual translation of the sequences, using the invertebrate mitochondrial universal code, confirmed the sequences all coded for uninterrupted open reading frame amino acid sequence and are therefore unlikely to derive from a non-target nuclear mitochondrial pseudogenes (numts) origin. All of the obtained sequences were therefore judged to also be from the targeted *Simulium* specimens and thus included in a more detailed analysis.

In a preliminary analysis, the 42 sequences generated in this study were aligned to a set of 872 unique simuliid CO1 sequences, derived from a diverse range of simuliid species. Neighbour joining analysis with these sequences showed that they all fell into a single bootstrap-supported monophyletic group that excluded all the other 872 simuliid background sequences. It also showed that their closest 27 previously published sequences fell into a bootstrap-supported sister group. More computationally demanding heuristic

phylogenetic tree building methods were thus carried out on a subset of sequences from the initial 914 sequence dataset, including only the study-generated sequences and their 27 closest relatives. An alignment of these sequences, spanning 519 nucleotide positions, has been provided as supplementary material.

Fig. 2 shows a representative Maximum Likelihood tree built from this alignment. The tree clearly shows that all the CO1 sequences generated from this study form two distinct groups that coincide with the morphologically identified species: all the *S. limbatum* sequences are within a *S. limbatum*-exclusive bootstrap-supported monophyletic group and all the *S. incrustatum* s.l. are all within a *S. incrustatum*-exclusive bootstrap-supported monophyletic group. Because both the *S. limbatum* and the *S. incrustatum* s.l. specimens came from numerous collections (see Table 1), these results support the value of CO1 barcoding as a tool for discriminating between the two species and because of the substantial genetic distance between the two groups, which is never less than 10% (55 nucleotide differences across 519 positions), the results suggest the CO1 gene's utility for identification might be extended to all the onchocerciasis vectors of the Amazon focus.

3.3. CO1 barcodes reveal hidden genetic diversity and provide evidence for a distinct *S. limbatum* forest ecotype

While no clear morphological differences were detected among the 27 *S. limbatum* specimens collected for this study, CO1 DNA sequence analysis showed a clear difference between those collected from the forest and those collected in the savannah. As well as the species groupings referred to above, Fig. 2 also shows that all the *S. limbatum* specimens that were collected from the savannah (at three geographically distinct localities – see Table 1) fall into a single bootstrap-supported monophyletic group, nested within the *S. limbatum* species group. The clustering of these sequences supports the existence of a distinct *S. limbatum* ecotype or subspecies. The existence of such a group is also supported by the nucleotide variation found even among samples collected at the same site as in-variant sequences could easily be explained by a sampling bias generated from collecting closely related samples. The observed variation found within the nucleotide sequences, however, supports a more ancient relationship that has been around long enough to show signs of nucleotide-level selection.

Among the 22 *S. limbatum* specimens collected in the savannah, just four of the 519 nucleotide positions can be seen to vary; with each variant nucleotide position supported by just one sequence each varying by one position from the consensus. These single nucleotide polymorphisms (SNPs) are supported on both strands of the sequence runs and conceptual translation of them shows that three out of four code for silent (synonymous) substitutions, which is suggestive of selection having operated on this group of sequences and thus that the collected *S. limbatum* are from an ecotype with an ancient origin. The specimens collected in the forest, however, show much higher levels of diversity, relative to the specimens collected in the savannah, suggesting an even more ancient origin and thus implying the savannah ecotype may have evolved from a forest ancestor. None of the five CO1 sequences obtained from *S. limbatum* specimens collected in the forest were seen to be identical; with variation between sequences ranging from one to 13 nucleotide positions (0.19–2.5%), across the same 519 nucleotide positions.

3.4. Genetically indistinct *S. incrustatum* found at multiple Amazon forest sites

Much lower levels of genetic variation were found among the tested *S. incrustatum* s.l. specimens than was found for *S. limbatum*

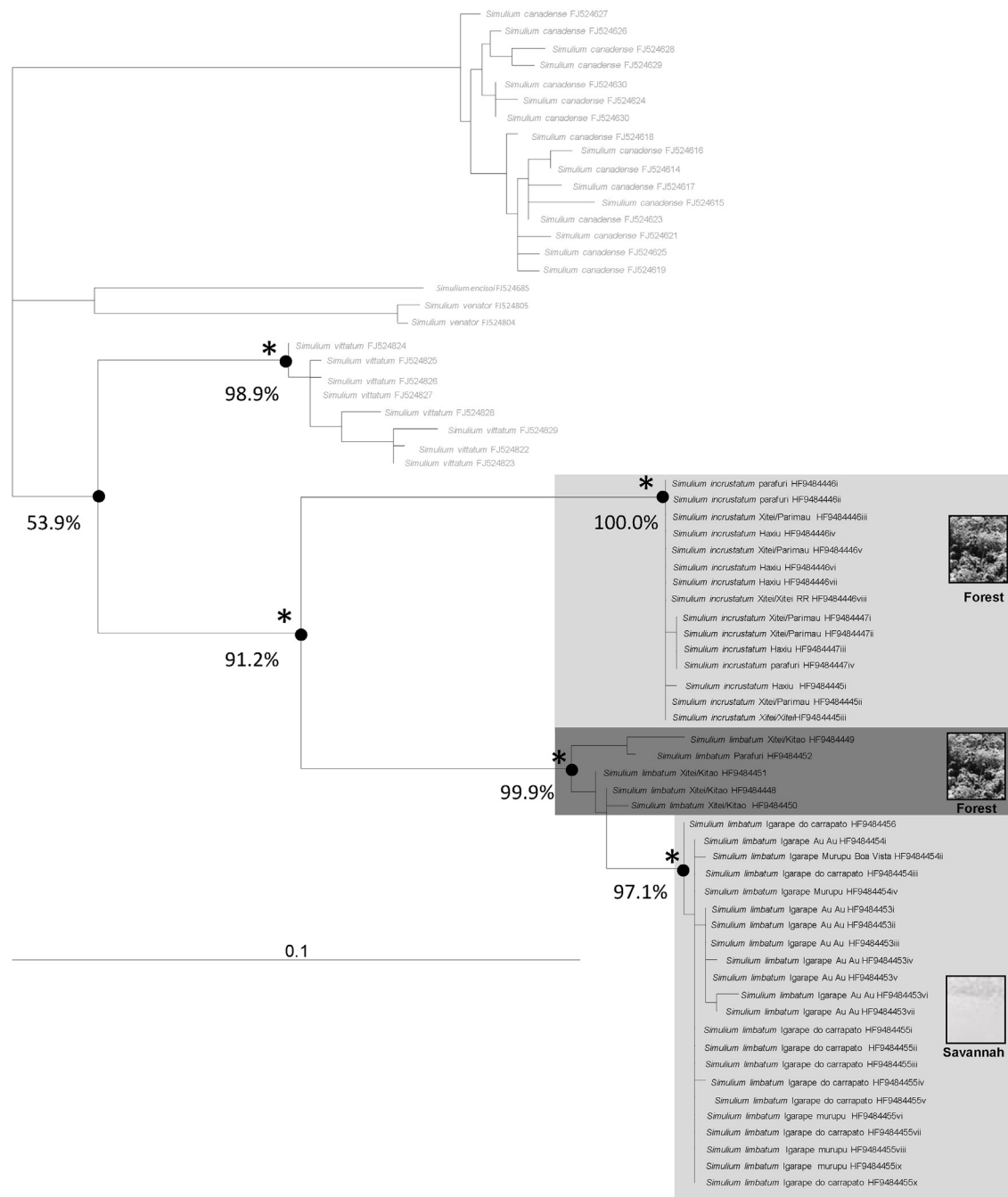


Fig. 2. A representative ML tree constructed from a 519 nucleotide CO1 sequence alignment (provided in the supplementary material). Branch labels indicate the species and the location that the sequences were obtained from. A nucleotide database accession number follows each of these labels so that the corresponding sequences, which each label represents, maybe easily accessed. Ecology specific sequence clusterings are highlighted with background shading and labelling. Bootstrap supported nodes discussed in the text are indicated with their percentage values. Those which are also supported with >50% support from both NJ and MP analysis are indicated with an asterisk.

(see Fig. 2). While the 27 *S. limbatum* sequences were observed to vary at 21 nucleotide positions and by up to 19 nucleotide positions (3.67%), the 15 *S. incrustatum* s.l. sequences were seen to vary at just two nucleotide positions (with one change being silent and one coding for a conservative amino acid change); sequences were never seen to vary by more than 0.39%. These results thus suggest *S. incrustatum* s.l. is more recently derived than *S. limbatum*. No clear distinction between CO1 sequence haplotypes from *S. incrustatum* s.l. specimens collected at different forest sites was detected. Consistent with the morphological data, the sequence data provided no support for the existence of savannah-dwelling *S. incrustatum* s.l.

4. Discussion

4.1. Confirmation of *S. limbatum* in the Amazonia onchocerciasis focus and the utility of CO1 barcoding for vector identification and for onchocerciasis control

The fact that in all the phylogenetic identification methods that we employed in this study, the CO1 sequences obtained from morphologically identified *S. limbatum* and *S. incrustatum* fell into highly bootstrap supported species-specific monophyletic groups supports the species status of the two and shows that

the portion of the CO1 gene sequence we targeted can be used to distinguish between specimens of these species found within the Amazonia onchocerciasis focus. The high levels of between-species nucleotide diversity detected, of around 10%, is consistent with what has been observed in the related simuliid subgenus *Trichodagmia* (Hernández-Triana et al., 2012), but is higher than has been used to resolve other closely related Simuliidae species (Day et al., 2008; Hamada et al., 2010; Ilmonen et al., 2009) and thus supports the possibility that CO1 barcoding might be used to distinguish between these two species across their entire geographic range and might also be targeted for the identification and discrimination of other vectors within the Amazonia focus. This notion is also supported by the fact that the specimens collected for this study came from sites where the species are living in sympatry (see Fig. 1). The clear genetic distinction between all of the forest collected *S. limbatum* and *S. incrustatum* CO1 sequences thus suggests there is no introgression between these species and that the observed differences are species differences and are not simply a consequence of geographical isolation. More geographic sampling will be required to identify the existence of species-diagnostic single nucleotide polymorphisms (SNPs), which could be used to develop simpler single-step species-diagnostic PCR reactions similar to what has been developed and continues to play an important role in the identification and control of members of the *Anopheles gambiae* complex (Mwangangi et al., 2013; Scott et al., 1993), but even without such developments the sequences described here can still provide a useful aid for vector identification and incrimination.

The depositing of *S. limbatum* and *S. incrustatum* reference CO1-sequences in public databases will very likely provide a robust tool for *S. incrustatum* and *S. limbatum* identification throughout their geographic range and certainly assist with taxonomic identifications of Amazonia vectors. A clear, specific, example of how this novel technique could have an immediate impact on onchocerciasis vector monitoring comes from a published report of *S. incrustatum* distribution. While almost all accounts of *S. incrustatum* come from collections made in the forest environment, the Brazilian Simuliid reference laboratory 'Laboratório de Simulídeos e Onco-cercose' (laboratory for Simuliids and Onchocerciasis) has recorded *S. incrustatum* at a savannah site (Igarapé, Murupu, near Boa Vista) (Marchon-Silva et al., 2007), where specimens collected and used for this study and all seven previous specimen collections (dating from 1979 to 1986) have found exclusively *S. limbatum* (Shelley et al., 1997). Whether this discrepancy is a consequence of simple misidentifications or a consequence of a genuine sudden, and temporary, vector faunal shift, is not clear, however, what is clear is that it suggests all future *S. incrustatum* savannah reports should be supported with CO1-barcode identifications and highlights the urgent need for reliable vector identification tools.

Perhaps more importantly, however, the sequences reported here may also help (in the future) to resolve the vector status of *S. limbatum* and help to produce a clearer picture of its distribution and of the distribution of *S. incrustatum*. While there have been records of infected *S. limbatum* on the Venezuelan side of the focus (Basáñez et al., 1988), there is some doubt as to whether these specimens were *S. incrustatum* or *S. limbatum* (Shelley et al., 1997, 2010). To our knowledge, no *S. limbatum* collected from the Amazonia focus has ever been compared directly with type material from Guyana. The closest localities to the Amazonia focus where *S. limbatum* populations have had their species status confirmed in this way are in the savannah regions of Roraima state (Shelley et al., 1997). As some of the savannah material used in this study came from these sites and as our morphological and molecular identification of *S. limbatum* found inside the focus have been made by direct comparison with this gold standard material, our report here of forest dwelling *S. limbatum* on the Brazilian side

of the Amazonia focus is not only the first of its kind, but is also the most certain account of the species inside of focus, as well.

4.2. Hidden and potentially epidemiologically important diversity within *S. limbatum*

The tight clustering of all the savannah-caught *S. limbatum* specimens (nested within the species clade) is inconsistent with forest and savannah caught specimen belonging to a single continuous interbreeding population and suggests that they belong to a distinct ecotype or subspecies of *S. limbatum*. Further geographical sampling will be required to determine whether these specimens are part of a subspecies or ecotype or whether the savannah-caught specimens surveyed here are simply a genetically isolated population. As the only unambiguous incrimination of *S. limbatum* as a host of onchocerciasis (Shelley et al., 1987) comes from blackflies taken from a savannah population outside the focus, and which in our the molecular analysis presented here can be seen to be distinct from those found inside the focus, the vectorial status of this species remains uncertain. The present evidence suggests that one of the vectors of *O. volvulus* in the Venezuelan part of the Amazonia focus (Basáñez et al., 1988) may be *S. limbatum*, but this requires confirmation.

Accurate cytospecies identifications have often played a key role in understanding the epidemiology of onchocerciasis in Africa and for its affective control (Garms et al., 2009; Traoré et al., 2009); it, therefore, may be that future disease control in the Amazonia onchocerciasis focus can benefit similarly from a higher resolution of vector taxonomy. Further sampling will also be needed to resolve the genetic structure of the forest-dwelling *S. limbatum* in Brazil. While we have only managed to obtain five forest-dwelling specimens for this analysis these specimens are more diverse (at the molecular level) than their more diversely sampled savannah-dwelling relatives. The relationships observed between the forest-dwelling *S. limbatum* specimens also appear to be more complex than the relationships among the savannah-dwelling specimens. Some of the forest-collected specimens are more closely related to the savannah-collected specimens than they are to other forest-collected specimens and thus the forest-collected specimens do not form a discrete monophyletic group, as do the savannah-collected specimens. It is, therefore, possible that vector competencies may even vary within the forest population of *S. limbatum*.

4.3. Hidden genetic diversity provides insights into the origin and evolution of *S. limbatum*

Although the CO1 sequence diversity data used for this study derives from mitochondria DNA and can therefore be easily be distorted by homogenising sweeps such as those driven by the common insect reproductive parasite *Wolbachia* (which have been found in simuliids [Crainey et al., 2010]), there is, presently, no evidence to suggest that *S. limbatum* mitochondrial diversity has been affected in this way. Thus the much higher levels of genetic diversity found among the much smaller sampled population of forest-dwelling *S. limbatum* is, presently at least, most parsimoniously explained by the *S. limbatum* savannah-dwelling behaviour having evolved from a forest-dwelling ancestral state. Of course, establishing this will not only require more geographic sampling (to prove the existence of a *S. limbatum* ecotype), it will also require some nuclear gene sampling.

5. Conclusions

The results presented here demonstrate that *S. limbatum* and *S. incrustatum* from the Amazonia focus can be reliably discriminated

using CO1 DNA barcoding. More generally, the results also support the utility of CO1 barcoding for both more reliable and higher resolution vector identifications within the Amazonia Onchocerciasis focus and provide some insight into the species evolution.

Acknowledgements

This study was conducted in partnership with the Brazilian barcoding of life data (BOLD) system. We would like to thank Claudete Schuertz and her team of DSEI-Yanomami/SESAI for all of their help in this study and especially for their help with blackfly collections made for this study, which were carried out for a study entitled: 'Avaliação do risco potencial de estabelecimento de focos secundários de transmissão de *Onchocerca volvulus* na amazônia', which received ethical clearance certificate CAAE-0001/2006 from the 'comité de ética em pesquisa'. We would also like to thank four anonymous referees for their comments, which have helped to improve this manuscript. This paper is contribution number 15 of the Research Programme on Infectious Disease Ecology in the Amazon (RP-IDEA) of the Instituto Leônidas & Maria Deane – Fiocruz Amazônia.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.actatropica.2013.03.012>. These data include Google maps of the most important areas described in this article.

References

- Basáñez, M.G., Churcher, T.S., Grillet, M.E., 2009. *Onchocerca-Simulium* interactions and the population and evolutionary biology of *Onchocerca volvulus*. *Adv. Parasitol.* 68 (68), 263–313.
- Basáñez, M.G., Pion, S.D.S., Churcher, T.S., Breitling, L.P., Little, M.P., Boussinesq, M., 2006. River blindness: a success story under threat? *PLoS Med.* 3, 1454–1460.
- Basáñez, M.G., Yarzabel, L., Takaoka, H., Suzuki, H., Noda, S., Tada, I., 1988. The vectorial role of several blackfly species (Diptera: Simuliidae) in relation to human onchocerciasis in the Sierra Parima and Upper Orinoco regions of Venezuela. *Ann. Trop. Med. Parasitol.* 82, 597–611.
- Botto, C., Vilamizar, N.J., Jock, Z., Cortez, J., Escakona, M., Grillet, M.E., Basanez, M.G., 2011. Landscape epidemiology of human Onchocerciasis in Southern Venezuela. In: Nriagu, J. (Ed.), *Encyclopaedia of Environmental Health*. Elsevier, Oxford, pp. 366–379.
- Cheke, R.A., Garms, R., 2013. Indices of onchocerciasis transmission by different members of the *Simulium damnosum* complex conflict with the paradigm of forest and savanna parasite strains. *Acta Trop.* 125, 43–52.
- Coelho, G.E., Vieira, J.B.F., Garcia-Zapata, M.T.A., Schuertz, J.C.M., 1998. Identificação de áreas de estratificação epidemiológica no foco de oncocercose na região Yanomami, Roraima, Brasil. *Cadernos de Saúde Pública* 14, 109–118.
- Crainey, J.L., Wilson, M.D., Post, R.J., 2009. An 18S ribosomal DNA barcode for the study of *Isomeris lairdi*, a parasite of the blackfly *Simulium damnosum* s.l. *Med. Vet. Entomol.* 23, 238–244.
- Crainey, J.L., Wilson, M.D., Post, R.J., 2010. Phylogenetically distinct *Wolbachia* gene and pseudogene sequences obtained from the African onchocerciasis vector *Simulium squamosum*. *Int. J. Parasitol.* 40, 569–578.
- Day, J.C., Goodall, T.I., Post, R.J., 2008. Confirmation of the species status of the blackfly *Simulium galeratum* in Britain using molecular taxonomy. *Med. Vet. Entomol.* 22, 55–61.
- Day, J.C., Mustapha, M., Post, R.J., 2010. The subgenus *Eusimulium* (Diptera: Simuliidae) in Britain. *Aquatic Insects* 32, 281–292.
- Garms, R., Lakwo, T.L., Ndyomugenyi, R., Kipp, W., Rubaale, T., Tukesiga, E., Katamanywa, J., Post, R.J., Amazigo, U.V., 2009. The elimination of the vector *Simulium neavei* from the Itwara onchocerciasis focus in Uganda by ground larviciding. *Acta Trop.* 111, 203–210.
- Grillet, M.-E., Villamizar, N.J., Frontado, H.L., Cortez, J., Escalona, M., Botto, C., Basanez, M.-G., 2008. Vector competence of *Simulium oyapockense* s.l. and *S. incrustatum* for *Onchocerca volvulus*: Implications for ivermectin-based control in the Amazonia focus of human onchocerciasis, a multi-vector-host system. *Acta Trop.* 107, 80–89.
- Gustavsen, K., Hopkins, A., Sauerbrey, M., 2011. Onchocerciasis in the Americas: from arrival to (near) elimination. *Parasites Vectors*, 4.
- Hamada, N., Pepinelli, M., Mattos-Gloria, A., Luz, S.L.B., 2010. A new black fly species from Brazil, closely related to *Simulium guianense* Wise (Diptera, Simuliidae), revealed by morphology and DNA barcoding. *Zootaxa*, 22–36.
- Hebert, P.D.N., Cywinska, A., Ball, S.L., DeWaard, J.R., 2003. Biological identifications through DNA barcodes. *Proc. Roy. Soc. Lond. Ser. B: Biol. Sci.* 270, 313–321.
- Hernández-Triana, L.M., Crainey, J.L., Hall, A., Faith, F., Mackenzie-Dodds, J., Shelley, A.J., Zhou, X., Post, R.J., Gregory, T.R., Hebert, P.D.N., 2012. DNA barcodes reveal cryptic genetic diversity within the blackfly subgenus *Trichodagma* Enderlein (Diptera: Simuliidae: Simulium) and related taxa in the New World. *Zootaxa* 3514, 43–69.
- Ilmonen, J., Adler, P.H., Malmqvist, B., Cywinska, A., 2009. The *Simulium venum* group (Diptera: Simuliidae) in Europe: multiple character sets for assessing species status. *Zool. J. Linnean Soc.* 156, 847–863.
- Marchon-Silva, V., Caer, J.C., Post, R.J., Maia-Herzog, M., Fernandes, O., 2007. Detection of *Onchocerca volvulus* (Nematoda: Onchocercidae) infection in vectors from Amazonian Brazil following mass Mectizan™ distribution. *Mem. Inst. Oswaldo Cruz* 102, 197–202.
- Mwangangi, J.M., Mbogo, C.M., Orindi, B.O., Muturi, E.J., Midega, J.T., Nzovu, J., Gatakala, H., Githure, J., Borgemeister, C., Keating, J., Beier, J.C., 2013. Shifts in malaria vector species composition and transmission dynamics along the Kenyan coast over the past 20 years. *Malar. J.*, 12.
- Post, R.J., Onyenwe, E., Somiari, S.A.E., Mafuyai, H.B., Crainey, J.L., Ubachukwu, P.O., 2011. A guide to the *Simulium damnosum* complex (Diptera: Simuliidae) in Nigeria, with a cytotoxicomic key for the identification of the sibling species. *Ann. Trop. Med. Parasitol.* 105, 277–297.
- Scott, J.A., Brogdon, W.G., Collins, F.H., 1993. Identification of single specimens of the *Anopheles gambiae* complex by the polymerase chain-reaction. *Am. J. Trop. Med. Hyg.* 49, 520–529.
- Shelley, A.J., Dias, A., Moraes, M.A.P., Procunier, W.S., 1987. The status of *Simulium oyapockense* and *Simulium limbatum* as vectors of human onchocerciasis in Brazilian Amazonia. *Med. Vet. Entomol.* 1, 219–234.
- Shelley, A.J., Hernández, L.M., Maia-Herzog, M., Luna Dias, A.P.A., Garritano, P.R., 2010. The blackflies (Diptera: Simuliidae) of Brazil. In: Jorge, R., Arias, Sergei Golovach, Karl Matthias Wantzen, Eduardo Domínguez, Peter Adler (Eds.), *Aquatic Biodiversity in Latin America*, vol. 6. PENSOFT, Sofia-Moscow, 821 pp.
- Shelley, A.J., Lowry, C.W., Maia-Herzog, M., Luna Dias, A.P.A., Moraes, M.A.P., 1997. Biosystematic studies on the Simuliidae (Diptera) of the Amazonia Onchocerciasis Focus. *Bull. Br. Mus. (Natural History) Entomol.* 66, 1–121.
- Shelley, A.J., Maia-Herzog, M., Lowry, C.A., Luna Dias, A.P.A., Garritano, P.R., Shelley, A., Camargo, M., Carter, H.G., 2000. The Simuliidae (Diptera) of the secondary onchocerciasis focus at Minaçu in central Brazil. *Bull. Nat. Hist. Mus.* 69, 171–221.
- Shelley, A.J., Pinger, R.R., Moraes, M.A.P., 1982. The taxonomy, biology and medical importance of *Simulium amazonicum* Goeldi (Diptera: Simuliidae), with a review of related species. *Bull. Br. Mus. (Nat. Hist.)* 44, 1–29.
- Traoré, S., Wilson, M.D., Sima, A., Barro, T., Diallo, A., Ake, A., Coulibaly, S., Cheke, R.A., Meyer, R.R.F., Mas, J., McCall, P.J., Post, R.J., Zoure, H., Noma, M., Yameogo, L., Seketeli, A.V., Amazigo, U.V., 2009. The elimination of the onchocerciasis vector from the island of Bioko as a result of larviciding by the WHO African Programme for Onchocerciasis Control. *Acta Trop.* 111, 211–218.
- Vivas-Martinez, S., Basáñez, M.G., Grillet, M.E., Weiss, H., Botto, C., Garcia, M., Vilamizar, N.J., Chavasse, D.C., 1998. Onchocerciasis in the Amazonia focus of southern Venezuela: altitude and blackfly species composition as predictors of endemicity to select communities for ivermectin control programmes. *Trans. R. Soc. Trop. Med. Hyg.* 92, 613–620.
- Vos, T., Flaxman, A.D., Naghavi, M., Lozano, R., Michaud, C., Ezzati, M., Shibuya, K., Salomon, J.A., Abdalla, S., Aboyans, V., Abraham, J., Ackerman, I., Aggarwal, R., Ahn, S.Y., Ali, M.K., Alvarado, M., Anderson, H.R., Anderson, L.M., Andrews, K.G., Atkinson, C., Baddour, L.M., Bahalim, A.N., Barker-Collo, S., Barrero, L.H., Bartels, D.H., Basanez, M.-G., Baxter, A., Bell, M.L., Benjamin, E.J., Bennett, D., Bernabe, E., Bhalla, K., Bhandari, B., Bikbov, B., Bin Abdulhak, A., Birbeck, G., Black, J.A., Blencowe, H., Bllore, J.D., Blyth, F., Bolliger, I., Bonaventure, A., Boufous, S.A., Bourne, R., Boussinesq, M., Braithwaite, T., Brayne, C., Bridgett, L., Brooker, S., Brooks, P., Brugh, T.S., Bryan-Hancock, C., Bucello, C., Buchbinder, R., Buckle, G.R., Budke, C.M., Burch, M., Burney, P., Burstein, R., Calabria, B., Campbell, B., Canter, C.E., Carabin, H., Carapetis, J., Carmona, L., Cella, C., Charlson, F., Chen, H., Cheng, A.T.-A., Chou, D., Chugh, S.S., Coffeng, L.E., Colan, S.D., Colquhoun, S., Colson, K.E., Condon, J., Connor, M.D., Cooper, L.T., Corriere, M., Cortinovis, M., de Vaccaro, K.C., Couser, W., Cowie, B.C., Criqui, M.H., Cross, M., Dabhadkar, K.C., Dahiya, M., Dahodwala, N., Damsere-Derry, J., Danaei, G., Davis, A., De Leo, D., Degenhardt, L., Dellavalle, R., Delossantos, A., Denenberg, J., Derrett, S., Des Jarlais, D.C., Dharmaratne, S.D., Dherani, M., Diaz-Torne, C., Dolk, H., Dorsey, E.R., Driscoll, T., Duber, H., Ebel, B., Edmond, K., Elbaz, A., Ali, S.E., Erskine, H., Erwin, P.J., Espindola, P., Ewoigbokhan, S.E., Farzadfar, F., Feigin, V., Felson, D.T., Ferri, A., Ferri, C.P., Fevre, E.M., Finucane, M.M., Flaxman, S., Flood, L., Foreman, K., Forouzanfar, M.H., Fowkes, F.G.R., Franklin, R., Fransen, M., Freeman, M.K., Gabbe, B.J., Gabriel, S.E., Gakidou, E., Ganatra, H.A., Garcia, B., Gaspari, F., Gillum, R.F., Gmel, G., Gosselin, R., Grainger, R., Groeger, J., Guillemin, F., Gunnell, D., Gupta, R., Haagsma, J., Hagan, H., Halasa, Y.A., Hall, W., Haring, D., Maria Haro, J., Harrison, J.E., Havmoeller, R., Hay, R.J., Higashi, H., Hill, C., Hoen, B., Hoffman, H., Hotez, P.J., Hoy, D., Huang, J.J., Ibeanusi, S.E., Jacobsen, K.H., James, S.L., Jarvis, D., Jasrasaria, R., Jayaraman, S., Johns, N., Jonas, J.B., Karthikeyan, G., Kassebaum, N., Kawakami, N., Keren, A., Khoo, J.-P., King, C.H., Knowlton, L.M., Kobusingye, O., Koranteng, A., Krishnamurthi, R., Laloo, R., Laslett, L.L., Lathlean, T., Leasher, J.L., Lee, Y.Y., Leigh, J., Lim, S.S., Limb, E., Lin, J.K., Lipnick, M., Lipshultz, S.E., Liu, W., Loane, M., Ohno, S.L., Lyons, R., Ma, J., Mabweijano, J., MacIntyre, M.F., Malekzadeh, R., Mallinger, L., Manivannan, S., Marcenes, W., March, L., Margolis, D.J., Marks, G.B., Marks, R., Matsumori, A., Matzopoulos, R., Mayosi, B.M., McAnulty, J.H., McDermott, M.M., McGill, N., McGrath, J., Elena Medina-Mora, M., Meltzer, M., Mensah, G.A., Merriman, T.R., Meyer, A.-C., Miglioli, V., Miller, M.,

- Miller, T.R., Mitchell, P.B., Mocumbi, A.O., Moffitt, T.E., Mokdad, A.A., Monasta, L., Montico, M., Moradi-Lakeh, M., Moran, A., Morawska, L., Mori, R., Murdoch, M.E., Mwaniki, M.K., Naidoo, K., Nair, M.N., Naldi, L., Narayan, K.M.V., Nelson, P.K., Nelson, R.G., Nevitt, M.C., Newton, C.R., Nolte, S., Norman, P., Norman, R., O'Donnell, M., O'Hanlon, S., Olives, C., Omer, S.B., Ortblad, K., Osborne, R., Ozgediz, D., Page, A., Pahari, B., Pandian, J.D., Rivero, A.P., Patten, S.B., Pearce, N., Perez Padilla, R., Perez-Ruiz, F., Perico, N., Pesudovs, K., Phillips, D., Phillips, M.R., Pierce, K., Pion, S., Polanczyk, G.V., Polinder, S., Pope III, C.A., Popova, S., Porrini, E., Pourmalek, F., Prince, M., Pullan, R.L., Ramaiah, K.D., Ranganathan, D., Razavi, H., Regan, M., Rehm, J.T., Rein, D.B., Remuzzi, G., Richardson, K., Rivara, F.P., Roberts, T., Robinson, C., De Leon, F.R., Ronfani, L., Room, R., Rosenfeld, L.C., Rushton, L., Sacco, R.L., Saha, S., Sampson, U., Sanchez-Riera, L., Sanman, E., Schwebel, D.C., Scott, J.G., Segui-Gomez, M., Shahraz, S., Shepard, D.S., Shin, H., Shivakoti, R., Singh, D., Singh, G.M., Singh, J.A., Singleton, J., Sleet, D.A., Sliwa, K., Smith, E., Smith, J.L., Stapelberg, N.J.C., Steer, A., Steiner, T., Stolk, W.A., Stovner, L.J., Sudfeld, C., Syed, S., Tamburlini, G., Tavakkoli, M., Taylor, H.R., Taylor, J.A., Taylor, W.J., Thomas, B., Thomson, W.M., Thurston, G.D., Tleyjeh, I.M., Tonelli, M., Towbin, J.R.A., Truelsen, T., Tsilimbaris, M.K., Ubeda, C., Undurraga, E.A., van der Werf, M.J., van Os, J., Vavilala, M.S., Venketasubramanian, N., Wang, M., Wang, W., Watt, K., Weatherall, D.J., Weinstock, M.A., Weintraub, R., Weisskopf, M.G., Weissman, M.M., White, R.A., Whiteford, H., Wiersma, S.T., Wilkinson, J.D., Williams, H.C., Williams, S.R.M., Witt, E., Wolfe, F., Woolf, A.D., Wulf, S., Yeh, P.-H., Zaidi, A.K.M., Zheng, Z.-J., Zonies, D., Lopez, A.D., Murray, C.J.L., 2012. [Years lived with disability \(YLDs\) for 1160 sequelae of 289 diseases and injuries 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 380, 2163–2196.](#)