Mito-nuclear genomics challenges the theory of clonality in *Trypanosoma congolense*. Reply to Tibayrenc and Ayala.

Frederik Van den Broeck, Lisse J.M. Tavernier, Lieve Vermeiren, Jean-Claude Dujardin, Jan Van Den Abbeele

Abstract

We recently published the first genomic diversity study of *Trypanosoma congolense*, a major etiological agent of Animal African Trypanosomiasis. We demonstrated striking levels of SNP and indel diversity in the Eastern province of Zambia as a consequence of hybridization between divergent trypanosome lineages. We concluded that these and earlier findings in *T. congolense* challenge the Predominant Clonal Evolution (PCE) model. In a recent Comment, Tibayrenc and Ayala claim that there are many features in *T. congolense* supporting their theory of clonality. While we can follow the reasoning of the authors, we also identify major limitations in their theory and interpretations that resulted in incorrect conclusions. First, we argue that each *T. congolense* subgroup should be analyzed independently as they may represent different (sub)species rather than 'near-clades'. Second, the authors neglect major findings of two robust population genetic studies on Savannah *T. congolense* that provide clear evidence of frequent recombination. Third, we reveal additional events of introgressive hybridization in *T. congolense* by analyzing the maxicircle coding region using Next Generation Sequencing analyses. Finally, we pinpoint two important misinterpretations by the authors and show that there are no spatially and temporally widespread clones in *T. congolense*. We stand by our earlier conclusions that the clonal framework is unlikely to accurately model the population structure of *T. congolense.* Other theoretical frameworks such as Maynard Smith's epidemic model may better represent the complex ancestry seen in *T. congolense*, where clones delimited in space and time seem to arise against a background of recombination.

Keywords: semi-clonal model, epidemic model, mito-nuclear discordance, Bioinformatics/Phyloinformatics, Hybridization, Parasitology

In their Comment, Tibayrenc and Ayala raised concerns regarding misquoting their theory of Predominant Clonal Evolution (PCE) in pathogens because we stated the following: *"Although this model does not exclude the occurrence of hybridization, it does imply restrained recombination that is scarce enough to not break a pattern of clonal population structure (Tibayrenc and Ayala, 2012). Such a population structure is generally characterized by genetically homogenous parasite populations containing a few predominant genotypes that propagated mitotically."* We agree with the authors' point (iii) and recognize that we should have used their broader definition of clonality by including reproductive modes other than mitotic propagation, such as selfing and homogamy (Tibayrenc and Ayala, 2012). However, we only partially agree with their argumentation in point (v) that "*PCE is by no means characterized by the existence of genetically homogeneous parasite populations"*, but by (a) strong linkage disequilibrium (LD), (b) propagation of (nearly) identical genotypes and (c) near-clading. On the one hand, we believe their argument is misleading because propagation of (nearly) identical genotypes with associated markers (high LD) results in relatively homogenous (sub-)populations. For instance, whole genome analyses of 75 isolates of the asexually evolving *T. brucei gambiense* in Côte d'Ivoire, Cameroon and Guinea demonstrated *"an extremely low level of intra-group diversity (e.g. the two most distantly related isolates differed only at 435 SNP loci)"* (Weir et al., 2016). On the other hand, we do agree that a low level of intra-group diversity is not necessarily a consequence of clonal evolution because recent diversification and genetic drift could have similar genetic consequences (also in sexual species). Unfortunately, by focusing on our restricted definition of clonality, Tibayrenc and Ayala (2018) almost entirely ignored our and previous observations of frequent recombination in *T. congolense* and argue that our data supports PCE in *T. congolense*. Here, we list the four main limitations of their appraisal.

1. Genetic subdivision does not imply clonal evolution

Our genomic data confirmed previous observations based on early enzyme analyses (e.g. Young and Godfrey, 1983) that Savannah and Forest parasites represent two discrete lineages (Tihon et al., 2017). Tibayrenc and Ayala (2018) believe this observation supports PCE and they equate these lineages to near-clades, namely a phylogenetic line that is clouded by occasional recombination. However, we believe that it makes no sense to explore PCE at the level of the entire *T. congolense* species as many (unexplored) factors other than clonality could lead to reproductive isolation. There is currently no evidence from the field for genetic exchange between any of the *T. congolense* subgroups, suggesting that the subgroups represent clades (not near-clades). Indeed, earlier observations based on congopain gene diversity showed that the genetic distances between the *T. congolense* subgroups (Forest, Savannah and Kilifi) are larger than the divergence between *T. simiae* and *T. godfreyi*, suggesting that the *T. congolense* subgroups may represent different species (Rodriguez et al., 2014). Hence, the significant divergence between *T. congolense* Forest and Savannah supports the biological species concept (see also in Gibson, 2007), namely a species concept based on restricted gene flow, in which genes are exchanged by recombination within but not between species. This reproductive isolation between Savannah and Forest parasites may be caused by various factors such as the occupation of different ecological/geographical niches (Gibson, 2007, 2008), implying infection of different tsetse fly species. Similarly, at a lower evolutionary scale (i.e. within the Savannah subgroup), genetic subdivision may act as (partial) reproductive barriers (see section 4). Therefore, we believe that it only makes sense to explore recombination frequency within populations that are un-subdivided in space and time (Ramirez and Llewelyn, 2014, 2015).

2. Evidence for frequent – not infrequent – recombination

Contrary to our conclusions, Tibayrenc and Ayala (2018) stated that "*the genomic data by Tihon et al. (2017) as well as data from previous authors support the presence in this species of PCE features*". We are not sure which *"previous authors*" the authors are referring to (see also concerns raised by Ramirez and Llewellyn (2015) and Rougeron et al. (2015) on the practice of self- and mis-citation by Tibayrenc and Ayala). Perhaps the authors were referring to the work of Simo et al. (2013) or Rodriguez et al. (2014), but these studies found evidence of clonality in Forest *T. congolense* while the Savannah subgroup was the main subject in our study (see also section 1). Importantly, note that the results of Rodriguez and colleagues suggested that recombination may be an important process in generating a diverse repertoire of the congopain genes in the Savannah subgroup (Rodriguez et al., 2014). To our knowledge, genetic studies that provide robust data/tools to test PCE in Savannah *T. congolense* are Morrison et al. (2009) and Tihon et al. (2017). Indeed, these are the only studies that (i) explored recombination within un-subdivided populations in Savannah *T. congolense* (see also section 1) and (ii) provided enough power to detect recombination, either through the analyses of a large number of samples (Morrison et al., 2009) or markers (Tihon et al., 2017). We believe it is important to shortly iterate the main findings of these studies as they challenge PCE:

- Morrison and colleagues typed 84 Savannah *T. congolense* isolates using seven microsatellite markers. Samples were collected within a small temporal (2006-2007) and spatial scale (within 50 km radius in the Central River District of The Gambia). Population genetic analyses revealed four subpopulations (their Figure 2) exhibiting no or limited linkage disequilibrium (LD) (their table 2) as estimated using the Index of Association (I_A) test (Morrison et al., 2009). While LD can occur for reasons other than a lack of random mating (see Box 2 in Holzmuller et al., 2010 and section 2 in Rougeron et al., 2015), Tibayrenc and Ayala (2012) strongly emphasize LD for inferring clonality in a species and they accept the I_A test as a reliable measure to do so. Morrison and colleagues identified one subpopulation exhibiting linkage equilibrium ($I_A = 0.006$; $p = 0.45$) indicative of random mating and the other subpopulations demonstrating limited LD ($I_A = 0.067$-0.095; $p = 0.01$). Unfortunately, this important finding is not taken into consideration in the Comment of Tibayrenc and Ayala (2018).
- Tihon and colleagues used whole-genome sequencing to characterize the type (SNPs, indels, CNVs) and extent of genetic diversity in 52 natural isolates of Savannah *T. congolense* from various regions in sub-Saharan Africa (Tihon et al., 2017). While providing preliminary insights into the distribution of genetic diversity across Africa, they focused their study on one region in particular, namely the Eastern province of Zambia. This is because (i) it was the only geographic region for which they sequenced a sufficient number (N=26) of samples, allowing detailed analyses and (ii) the fact that hybridization occurred between divergent *T. congolense* lineages provided sufficient phylogenetic power to robustly summarize the ancestry in these isolates (their Figure 5). Their results revealed a striking mosaic of ancestry from multiple genetic backgrounds.

Tihon and colleagues concluded their work by stating that *"together with earlier findings (Morrison et al., 2009), our observations suggest that hybridization in T. congolense may be more widespread and persistent than previously anticipated"* (Tihon et al., 2017). While Tibayrenc and Ayala (2018) acknowledged these findings, they disregard them within their analysis and argue in their point (iv) that *"the model includes occasional bouts of genetic exchange"*. However, the evidence for *T. congolense* does not indicate a singular or occasional event of hybridization, and therefore cannot be considered as paradigmatic cases of PCE. Therefore, in our view, the two empirical studies mentioned above that warrant any conclusion regarding PCE in Savannah *T. congolense* provide robust evidence of extensive mating in

populations from The Gambia and Zambia (Morrison et al., 2009; Tihon et al., 2017), indicating that recombination may be a relatively prevalent and universal process in this species.

3. Biogeographic mito-nuclear discordance reveals additional events of hybridization

Tihon et al. (2017) focused their study on the discovery of hybridization in Zambia for reasons we have mentioned in section 2. However, their results indicated the presence of recombination in other populations as well. The co-ancestry matrix resulting from the ChromoPainter analyses (Lawson et al., 2012) revealed many signals of asymmetric gene flow with varying degrees of shared haplotype segments among the sequenced *T. congolense* isolates (their Figure 4b). These findings show that most *T. congolense* genomes have a fragmented ancestry that is impossible to contain in a single phylogenetic tree. This challenges PCE because populations dominated by clonality should reveal the same phylogenetic history across the genome, a property known as congruence (Maiden, 2006).

We will describe one illustrative example. PCA and fineSTRUCTURE analyses based on nuclear SNPs revealed that isolates from Ethiopia and Uganda (Eastern Africa) largely clustered together with isolates from The Gambia, Cameroon, Mali, Togo and Burkina Faso (Western/Central Africa) (Figures 3 and 4a in Tihon et al., 2017). However, ChromoPainter analyses revealed a much more complex underlying ancestry, and indicated that isolates from Ethiopia and Uganda also contain a minor fraction of nuclear haplotype segments similar to those found in isolates from Zambia (Southern Africa) (Figure 4b in Tihon et al., 2017). This could possibly be the result of introgressive hybridization, whereby hybridization followed by repeated backcrossing resulted in the introduction of *T. congolense* genes from Zambia into the gene pool of *T. congolense* from Uganda and Ethiopia. Recurrent cycles of backcrossing typically result in a complex mixture of parental genes at the nuclear level, which makes it challenging to depict past hybridization events.

One approach that could help to formally demonstrate past recombination events is the study of biogeographic discordance between mitochondrial and nuclear DNA (e.g. Messenger et al., 2012). Until now, this was not possible in *T. congolense* because of the lack of mitochondrial sequence data. Therefore, in order to be able to include the mitochondrial DNA sequence diversity in our genetic analysis, we generated the first complete and annotated maxicircle reference genome for *T. congolense*. In short, the kinetoplast DNA of *T. congolense* IL3000 was separated from the nuclear DNA (Pérez-Morga & Englund, 1993), prepared for sequencing using the TruSeq nano DNA library preparation kit and sequenced using 500 cycles (2x250) on an Illumina MiSeq platform. The maxicircle was *de novo* assembled with SPAdes (Bankevich et al., 2012), elongated with NOVOplasty (Dierckxsens et al. 2017), chopped in two and circularized by assembling overlapping ends with CAP3 (Huang and Madan, 1999) (Figure 1a). The final assembly was annotated with RATT (Otto et al., 2011) using the *T. lewisi* annotation file (Lin et al., 2015) (Figure 1b). The *T. congolense* sequence data (Tihon et al. 2017) was aligned against this novel maxicircle reference using SMALT (https://www.sanger.ac.uk/science/tools/smalt-0) and SNPs were called with SAMTOOLS mpileup and bcftools (Li et al., 2009), including only reads with a minimum mapping quality of 20. The resulting 55 SNPs (with a minimum phred-scaled quality of 60) were concatenated and used to reconstruct a maxicircle haplotype network in PopART (Leigh and Bryant, 2015). Note that this does not reflect the true SNP diversity across the entire maxicircle coding region as the coverage was too low in some maxicircle regions. The resulting haplotype network based on maxicircle SNPs revealed that isolates from Ethiopia and Uganda had a mitochondrial sequence similar to those found in Zambia (Figure 1c), while their nuclear genome was similar to those found in The Gambia, Mali, Togo, Burkina Faso and Cameroon (Figure 1d). This is a typical pattern of biogeographic mito-nuclear discordance as the result of hybridization and

introgression. Although e.g. incomplete lineage sorting could also lead to mito-nuclear discordance, this would not result in such a drastically different mitochondrial genome as observed here (Toews and Brelsford, 2012). Similar observations have been found in other African trypanosomes, with genome analyses identifying genomic introgression between East African and West African *T. brucei* subspecies in Uganda (Goodhead et al., 2013), and also in American trypanosomes (Messenger et al., 2012, 2015; Ramirez et al., 2012).
Altogether, our observations at the nuclear and mitochondrial level clearly evidence introgressive hybridization *in T. congolense*, illustrated by the isolates from Uganda and Ethiopia. This has several implications with respect to the theory of clonality. First, we have now additional indications that recombination is universal in *T. congolense*, as it has been described in populations from West, Central, East and Southern Africa (namely The Gambia, Zambia, Ethiopia and Uganda; see also section 2). Second, introgressive hybridization can only happen through extensive backcrossing, which is often a long-term process, and which can only be explained by the existence of prevalent recombination, or in other words, it cannot be explained by a model that assumes predominant clonality.

4. Pairs of (near-)identical clones against a background of recombination: the epidemic/semi-clonal model of evolution

Tibayrenc and Ayala (2018) advanced several statements that were based on misinterpretations of our data, leading to incorrect conclusions regarding clonality in *T. congolense.*
First, the authors claim that the genetic subdivision of Savannah *T congolense* into TC1 and TC2 clusters represents a typical Russian doll pattern at a lower evolutionary scale that cannot be explained by geographical separation. This is incorrect: our results clearly suggested that *T. congolense* population structure strongly follows a geographic pattern (Figure 4a in Tihon et al. 2017). To further illustrate this, we assessed patterns of isolation-by-distance within the TC1 cluster (this was not possible for the TC2 cluster that only contained three isolates). This was done using a Mantel test in the R package vegan (R Development Core Team, 2015; Oksanen et al., 2013) between a matrix of pairwise log-transformed geographical distances and a matrix of pairwise log-transformed genetic distances. The Mantel test revealed a significant increase of genetic distance with geographical distance between *T. congolense* strains ($r = 0.63$; $p < 0.001$), suggesting the presence of physical barriers to long-distance gene flow (i.e. dispersal limitation) (Orsini et al., 2013). These results indicate that genetic subdivision is mainly explained by geographical separation (i.e. a Wahlund effect), presenting a physical obstacle to recombination and invalidating the Russian doll feature of PCE.
Second, the authors claim that we found 14 *T. congolense* strains that are virtually identical and widespread in space and time, evidencing PCE. This is incorrect: we have clearly stated that we found seven pairs of near-identical *T. congolense* isolates, and six of them were isolated within a single geographic region and during the same year (Table 3 in Tihon et al., 2017). For instance, within Cameroon we found two pairs of near-identical isolates/clones from 2005 (KONT2-133 w/ KONT2-155 and KONT1-129 w/ BEKA), and while the isolates of each pair showed no fixed SNP differences, the pairs themselves showed 3,528-3,545 fixed SNP differences. When we compare these clonal pairs from Cameroon to – for instance – the three clonal pairs found in Zambia, we identified a total of 15,597-18,453 fixed SNP differences. These examples clearly show that the 14 *T. congolense* strains listed in Table 3 in Tihon et al. (2017) are not identical. Moreover, the clonal pairs were restricted in time and space – most likely due to dispersal limitation (see above) – and appear against a background of recombination. To illustrate this, recall that we identified three clonal pairs in the Eastern Province of Zambia (restricted in space) and each pair was found in the same year, either in 1996 or 2003 (restricted in time) (Table 3 in Tihon et al. 2017). One clonal pair were F1-like hybrids and the two others were backcrossed hybrids (Figure 5 in Tihon et al. 2017), suggesting

that these clonal genotypes appeared following recombination. This is also illustrated for the other *T. congolense* populations in the co-ancestry matrix in Figure 4b of Tihon et al. (2017), where clonal pairs (black boxes) appear within populations showing a fragmented ancestry. These observations suggest that the population structure of *T. congolense* probably follows an epidemic/semi-clonal model of evolution (Maynard Smith et al., 1993; Maiden, 2006). This model assumes that there is relatively frequent recombination within all members of a given population, but that occasionally a highly successful individual arises and increases rapidly in frequency to produce an epidemic clone with a relatively short lifetime (see Figure 1c in Maynard Smith et al., 1993). Note that such epidemic clones may be oversampled if they arise from virulent organisms (Maynard Smith, 1991), possibly masking the true frequency of recombination in *T. congolense*.

Conclusions

We believe that the analyses by Tibayrenc and Ayala (2018) were too limited to claim clonality in *T. congolense.* Here we have listed several limitations in their Comment and advanced novel insights to update this complex debate. We stand by our earlier conclusions that recombination in *T. congolense* may be much more prevalent and universal than previously anticipated (Tihon et al., 2017), contrasting strongly with the asexual evolution of *T. b. gambiense* Type 1 (Weir et al., 2016). We argue that there is not 'one theory to rule them all' and that African trypanosomes – and parasitic protozoa in general – may use a variety of reproductive strategies ranging between clonal reproduction and sexual recombination, depending on the species, the population and the epidemiological setting examined. Within this context, novel analytical tools are needed to model the population structure of organisms that deviate from the purely phylogenetic (developed for haploid clonal organisms) and population genetic (developed for diploid sexual organisms) framework. As outlined here, (i) focusing on populations that are undivided in space and time, (ii) guaranteeing sufficiently large numbers of isolates avoiding oversampling of virulent strains, (iii) using next-generation sequencing tools to study discordant patterns of evolution (mito-nuclear, but also between windows of the nuclear genome) and (iv) advancing/optimizing novel analytic tools such as ChromoPainter (e.g. Yahara et al., 2014) will contribute valuable new insights into the recombination history of parasitic protozoa.

Acknowledgments

We thank Achim Schnaufer and Sinclair Cooper for helpful advice with the extraction and sequencing of the kinetoplast DNA. Frederik Van den Broeck (ITM-AAP KINUGEN) and Lisse Tavernier (ITM-PPP) were supported by the Department of Economy, Science and Innovation in Flanders. This work, including sequencing, was supported by a Research Grant (1528117N) of the Research Foundation Flanders (FWO) awarded to Frederik Van den Broeck.

Data accessibility

The sequence data (ERR2681942) and annotated assembly (ERZ667256) is available in EBI under the accession number PRJEB27608 (http://www.ebi.ac.uk/ena/data/view/PRJEB27608).

References

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A., Dvorkin, M., Kulikov, A. S., … Pevzner, P.A. (2012). SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *Journal of Computational Biology*, 19, 455-77.

Carver, T. J., Rutherford, K. M., Berriman, M., Rajandream, M. A., Barrell, B. G., & Parkhill, J. (2005). ACT: the Artemis Comparison Tool. *Bioinformatics*, 21, 3422-3.

Dierckxsens, N., Mardulyn, P., & Smits, G. (2017). NOVOPlasty: *de novo* assembly of organelle genomes from whole genome data. *Nucleic Acids Research*, 45, e18.

Gibson, W. (2007). Resolution of the species problem in African trypanosomes. *International Journal of Parasitology*, 37, 829-838.

Gibson, W. (2008). Molecular epidemiology of African trypanosomiasis: the contributions of David George Godfrey OBE to the biochemical characterization of trypanosomes. *Parasite*, 15, 233-6.

Goodhead, I., Capewell, P., Bailey, J. W., Beament, T., Chance, M., Kay, S., . . . Hall, N. (2013). Whole-genome sequencing of *Trypanosoma brucei* reveals introgression between subspecies that is associated with virulence. *mBio*, 4, e00197–13.

Holzmuller, P., Herder, S., Cuny, G., & De Meeûs, T. (2010). From clonal to sexual: A step in *T. congolense* evolution? *Trends in Parasitology*, 26, 56–60.

Huang, X., & Madan, A. (1999). CAP3: A DNA Sequence Assembly Program. *Genome Research*, 9, 868-877.

Lawson, D. J., Hellenthal, G., Myers, S., & Falush, D. (2012). Inference of population structure using dense haplotype data. *PLoS Genetics*, 8, e1002453.

Leigh, J. W., & Bryant, D. (2015). PopART: Full-feature software for haplotype network construction. *Methods in Ecology and Evolution*, 6, 1110-1116.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R. and 1000 Genome Project Data Processing Subgroup (2009). The Sequence alignment/map (SAM) format and SAMtools. *Bioinformatics*, 25, 2078-9.

Lin, R.-H., Lai, D.-H., Zheng, L.-L., Wu, J., Lukes, J., Hide, G., & Lun, Z.-R. (2016). *Parasite and Vectors*, 13, 9-15.

Maiden, M. C. J. (2006). Multilocus Sequence Typing of Bacteria. *Annu. Rev. Microbiol*. 2006. 60:561–88.

Maynard Smith, J., Dowson, C. G., & Spratt, B. G. (1991). Localized sex in bacteria. *Nature*, 349, 29-31.

Maynard Smith, J., Smith, N. H., O'Rourke, M., & Spratt, B. G. (1993). How clonal are bacteria? *Proceedings of the National Academy of Sciences of the United States of America*, 90, 4384-88.

Messenger, L. A., Llewellyn, M. S., Bhattacharyya, T., Franzén, O., Lewis, M. D., Ramirez, J. D., Carrasco, H. J., Andersson, B., & Miles, M. A. (2012). Multiple mitochondrial introgression events and heteroplasmy in *Trypanosoma cruzi* revealed by maxicircle MLST and next generation sequencing. *PLoS Neglected Tropical Diseases*, 6, e1584.

Messenger, L. A., Garcia, L., Vanhove, M., Huaranca, C., Bustamante, M., Torrico, M., Torrico, F., Miles, M.A., & Llewellyn, M.S. (2015). Ecological host fitting of *Trypanosoma cruzi* TcI in Bolivia: mosaic population structure, hybridization and a role for humans in Andean parasite dispersal. *Molecular Ecology*, 24, 2406-2422.

Morrison, L. J., Tweedie, A., Black, A., Pinchbeck, G. L., Christley, R. M., Schoenefeld, A., . . . Tait, A. (2009). Discovery of mating in the major African livestock pathogen *Trypanosoma congolense*. *PLoS One*, 4, e5564.

Oksanen, J. F., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D. … Wagner, H. (2017). vegan: Community Ecology Package. R package version 2.4-4. https://CRAN.R-project.org/package=vegan

Orsini, L., Vanoverbeke, J., Swillen, I., Mergeay, J., & De Meester, L. (2013). Drivers of population genetic differentiation in the wild: isolation by dispersal limitation, isolation by adaptation and isolation by colonization. *Molecular Ecology*, 22, 5583-5999.

Otto, T., Dillon, G. P., Degrave, W. S., & Berriman, M. (2011). RATT: Rapid Annotation Transfer Tool. *Nucleic Acids Research*, 39, e57.

Pérez-Morga, D., & Englund, P. T. (1993). The structure of replicating kinetoplast DNA networks. *Journal of Cell Biology*, 123, 1069–79.

Ramirez, J.D., Guhl, F., Messenger, L.A., Lewis, M.D., Montilla, M., Cucunuba, Z., Miles, M.A., Llewellyn, M.S. (2012). Contemporary cryptic sexuality in *Trypanosoma cruzi*. *Molecular Ecology*, 21, 4216-4226.

Ramirez, J. D., & Llewellyn, M. S. (2014). Reproductive clonality in protozoan pathogens-truth or artefact? *Molecular Ecology*, 23, 4195–4202.

Ramirez, J. D., & Llewellyn, M. S. (2015). Response to Tibayrenc and Ayala: Reproductive clonality in protozoan pathogens – Truth or artefact? *Molecular Ecology*, 24, 5782–5784.

R Development Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, 1, 409.

Rodrigues, A. C., Ortiz, P. A., Costa-Martins, A. G., Neves, L., Garcia, H. A., Alves, J. M., . . . Teixeira, M. M. (2014). Congopain genes diverged to become specific to Savannah, Forest and Kilifi subgroups of *Trypanosoma congolense*, and are valuable for diagnosis, genotyping and phylogenetic inferences. *Infection, Genetics and Evolution*, 23, 20–31.

Rougeron, V., De Meeûs, T., Banuls, A.-L. (2015). Response to Tibayrenc et al.: can recombination in *Leishmania* parasites be so rare? *Trends in Parasitology*, 31, 280-281.

Simo, G., Sobgwi, P. F., Njitchouang, G. R., Njiokou, F., Kuiate, J. R., Cuny, G., & Asonganyi, T. (2013). Identification and genetic characterization of *Trypanosoma congolense* in domestic animals of Fontem in the South-West region of Cameroon. *Infection, Genetics and Evolution*, 18, 66–73.

Tibayrenc, M., & Ayala, F. J. (2012). Reproductive clonality of pathogens: A perspective on pathogenic viruses, bacteria, fungi, and parasitic protozoa. *Proceedings of the National Academy of Sciences of the United States of America*, 109, E3305–E3313.

Tibayrenc, M., & Ayala, F. J. (2018). Hybridization in *Trypanosoma congolense* does not challenge the predominant clonal evolution model. A comment on Tihon et al., 2017, Mol. Ecol. *Molecular Ecology*, this issue.

Tihon, E., Imamura, H., Dujardin J.-C., Van Den Abbeele J., & Van den Broeck F. (2017). Discovery and genomic analyses of hybridization between divergent lineages of *Trypanosoma congolense*, causative agent of African Animal Trypanosomiasis. *Molecular Ecology*, 26, 6524-6538.

Toews, D. P. L, & Brelsford, A. (2012). The biogeography of mitochondrial and nuclear discordance in animals. *Molecular Ecology*, 21, 3907-3930.

Weir, W., Capewell, P., Foth, B., Clucas, C., Pountain, A., Steketee, P., . . . MacLeod, A. (2016). Population genomics reveals the origin and asexual evolution of human infective trypanosomes. *Elife*, 5, e11473.

Yahara, K., Didelot, X., Azim Ansari, M., Sheppard, S.K., & Falush, D. (2014). Efficient inference of recombination hot regions in bacterial genomes. *Molecular Biology and Evolution*, 31, 1593-1605.

Young, C. J., & Godfrey, D. G. (1983). Enzyme polymorphism and the distribution of *Trypanosoma congolense* isolates. *Annals of Tropical Medicine and Parasitology*, 77, 467–481.

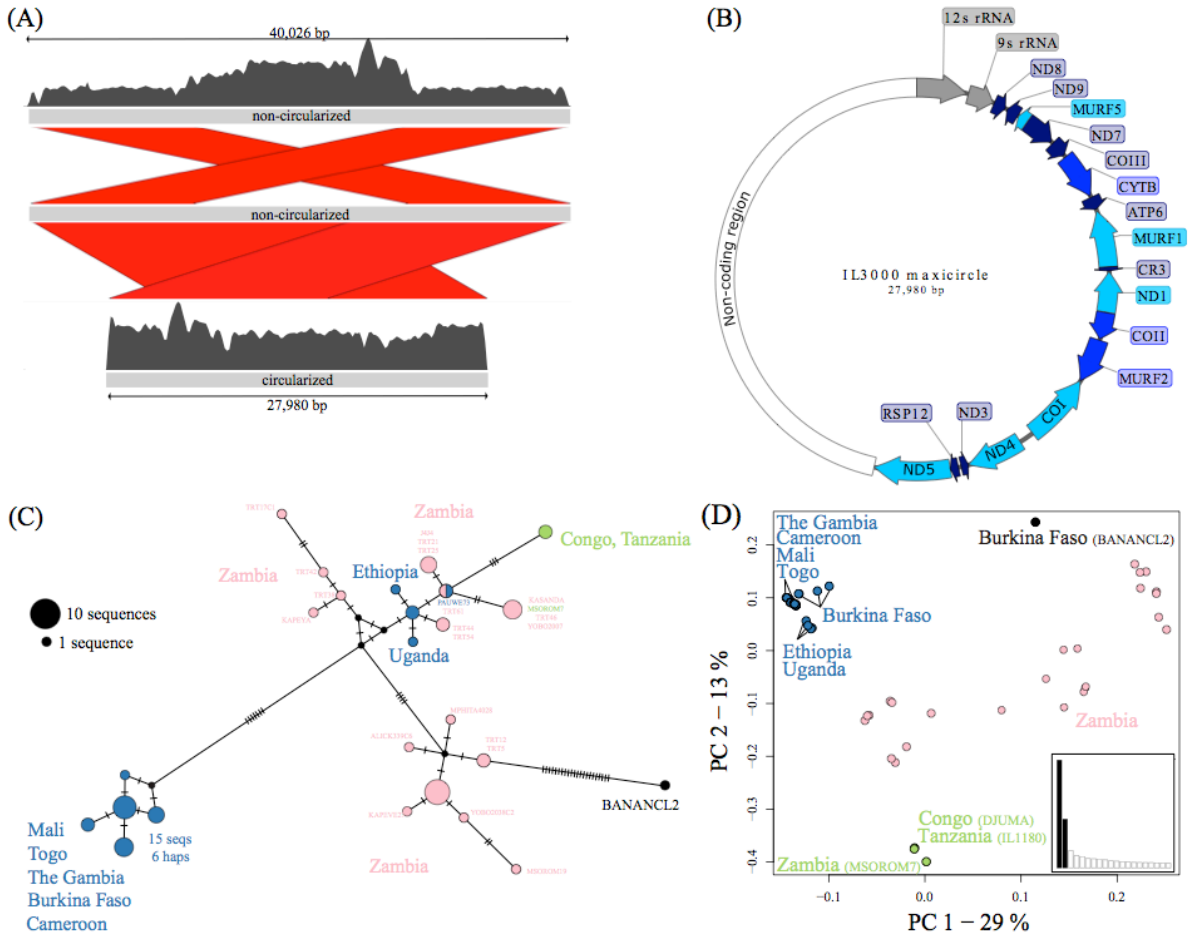

Figure 1. Mito-nuclear discordance in Savannah *T. congolense.* **(A)** Figure from ACT (Carver et al., 2005) visualizing the BLASTN comparisons (red bands) between the non-circularized and circularized maxicircle sequences (= part of the mitochondrial genome). These comparisons show the matching ends of the non-circularized maxicircle, as well as the overlapping parts between the non-circularized and circularized maxicircle. Coverage plots (black) show the properly paired 250bp reads aligned with an average read depth of 1,424 and 2,042 to the non-circularized and circularized maxicircles respectively. **(B)** Visualization from SnapGene (http://www.snapgene.com) representing the fully assembled and annotated maxicircle sequence containing a repetitive non-coding region, 2 ribosomal genes and 17 protein coding genes that are never (light blue), partially (blue) or completely (dark blue) edited. **(C)** Haplotype network based on maxicircle SNPs and statistical parsimony. Each pie diagram represents a haplotype (i.e. unique sequence) and hashes represent haplotypes that were either not sampled or went extinct and can thus be regarded as mutational steps. The sizes of the pie diagrams reflect the number of sequences that represent the respective haplotypes. **(D)** PCA plot based on nuclear SNPs, adopted from Tihon et al. (2017). The colors of the pie diagrams in panels C and D are: blue depicting the TC1 cluster (Mali, Togo, The Gambia, Burkina Faso, Cameroon, Ethiopia, Uganda), green the TC2 cluster (Congo, Tanzania) and pink the isolates from Zambia. Note that isolate BANANCL2 contains a divergent maxicircle haplotype (panel C), and this isolate was identified as a hybrid resulting from the outcrossing between a TC1-like parent and a unknown divergent parent (Figure 4d in Tihon et al., 2017).