Return of the Mitochondrial DNA

A *Fusarium oxysporum* story

Scan for more information

Balázs Brankvics[1,2], Peter van Dam[3], Martijn Rep[3], G. Sybren de Hoog[1,2], Theo A. J. van der Lee[4], Cees Waalwijk[4] & Anne D. van Diepeningen[1,4]

[1]Westerdijk Fungal Biodiversity Institute (former CBS-KNAW), Royal Netherlands Academy of Arts and Sciences (KNAW), Utrecht, The Netherlands; [2]Institute of Biodiversity and Ecosystem Dynamics, University of Amsterdam, Amsterdam, The Netherlands; [3]Swammerdam Institute for Life Sciences, University of Amsterdam, Amsterdam, The Netherlands; [4]Wageningen Plant Research,Wageningen University and Research Centre, Wageningen, The Netherlands

Introduction

In the 1990s, the barcoding of life project was started, aiming to use a single sequence to identify animals, fungi and plants to the species level. The first marker that was proposed was a mitochondrial marker: the mitochondrially encoded cytochrome oxidase I (*cox1* in filamentous fungi).

However, amplifying the *cox1* sequence proved problematic in many fungal groups, because the frequent insertions of introns into the region made universal primer design difficult. Hence, the mitochondrial marker was abandoned and the barcoding community chose the ITS region as official barcode for fungi. Unfortunately, the ITS barcode proved to have insufficient resolution in many closely related species. Thus, multi-locus analysis became the new standard, most of which included at least one mitochondrial marker. There has been no consensus on which mitochondrial loci to include.

With next generation sequencing and new assembly tools it is possible to assemble the complete mitochondrial genome of isolates, which provide all the benefits that are associated with using mitochondrial markers. In addition, using the complete mitochondrial genome offers better resolution for phylogenetic analyses and with sufficient sampling it can be placed in the context of previous works done on mitochondrial barcoding markers.

Demonstration: How the mitochondrial genome revealed recombination[1]

Two parts of the mitochondrial genome

The mitochondrial genome of *Fusarium* spp. can be divided into two parts: the conserved part (shows complete synteny between species) and the large variable (LV) region (Fig. 1). From NGS data, we have assembled, annotated and compared the mitochondrial genomes of 61 strains of the *Fusarium oxysporum* species complex (FOSC) together with 1 *F. commune* and 2 *F. proliferatum* strains as outgroup.

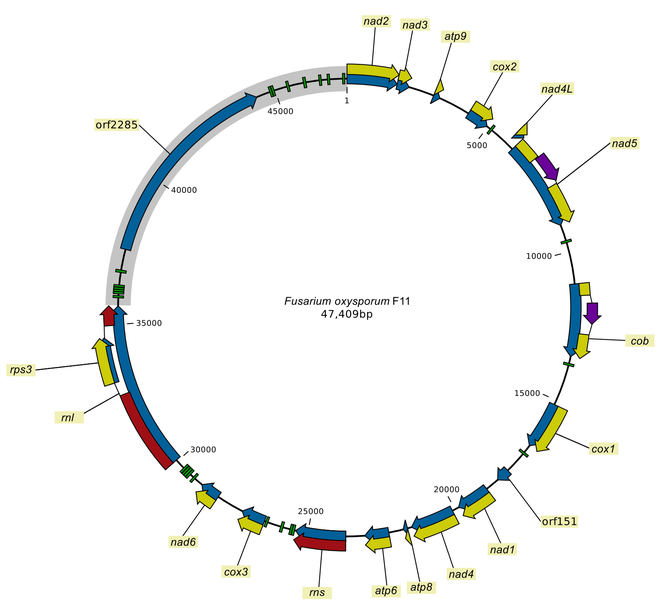

Figure 1: Mitochondrial genome of *Fusarium oxysporum* f. sp. *cumini* strain F11. Green blocks: tRNA coding genes, blue arrows: genes, yellow arrows: protein coding sequences, red arrows: rDNA coding sequence, purple arrows: intron encoded homing endonuclease genes, gray segment: large variable (LV) region with orf2285 (LV-uORF).

Non-homologous variants of the LV region

Three different variants of the LV region (Fig. 2) were found within the FOSC. Variant 1 (Fig. 2a) is homologous to the LV region found in other *Fusarium* spp. The three variants contain at least 13 tRNA genes, their order shows partial synteny. However, the variants have highly divergent sequences, which is demonstrated by the fact that BLASTN is unable to identify synteny blocks between the variants.

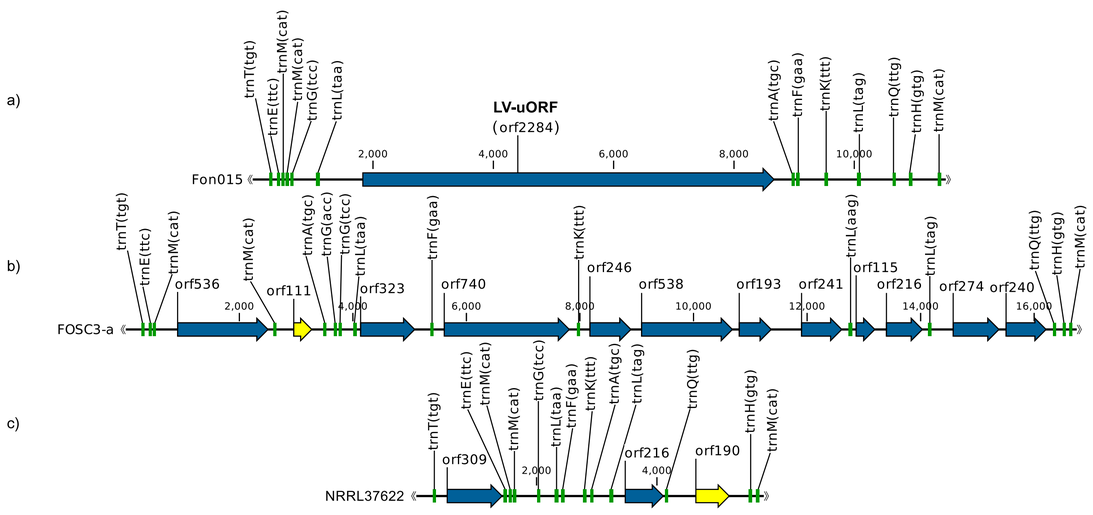

Figure 2: The three variants of the large variable region. a) Variant 1 represented by *F. oxysporum* strain Fon015, **b)** variant 2 represented by *F. oxysporum* strain FOSC3-a and **c)** variant 3 represented by *F. oxysporum* strain NRRL37622. Green blocks: tRNA coding genes, blue arrows: ORFs, yellow arrows: ORFs that are not present in all representatives of the given variant.

Both variants 1 & 2 are not clade specific

The strains of the FOSC used in this study could be grouped into three clades which were recognized as phylogenetic species based on genealogical concordance. Variant 1 was present in all three clades, variant 2 was present in clades 2 & 3 and variant 3 was restricted to clade 2. Possible hypotheses for the distribution of variants 1 & 2 are as follows:

H_1: Variant 2 emerged in either clade 1 or 2 and the complete mitochondrial genome was transferred to a strain of the other clade, then the genome spread in the population without recombination.

H_2: Variant 2 emerged in either clade 1 or 2 and the complete mitochondrial genome was transferred to a strain of the other clade, then the genome spread in the population by recombination.

H_3: Variant 2 emerged in the ancestor of clades 1 & 2 and was maintained within both lineages during the separation of the two phylogenetic species without recombination.

H_4: Variant 2 emerged in the ancestor of clades 1 & 2 and was maintained within both lineages during the separation of the two phylogenetic species by recombination.

Co-evolution of the conserved part and the LV region variants

The conserved part of the mitochondrial genome of the strains supports the separation of the three clades (Fig. 3: Conserved mt region). Thus, the conserved part of the mitogenomes of strains belonging to the same clade are more similar irrespective of which variant the strains have, this rules out H_1 & H_3. Both variants 1 & 2 support the separation of the two clades (Fig. 3), this also rules out H_1 & H_3. If H_2 were correct, we would expect that one of the clades is paraphyletic both for the tree based on variant 2 and the conserved part, but there is no sign of this in the data. Therefore, we conclude that H_4 is correct: variant 2 emerged in the ancestor of clades 1 & 2 and was maintained within both lineages during the separation of the two phylogenetic species by recombination.

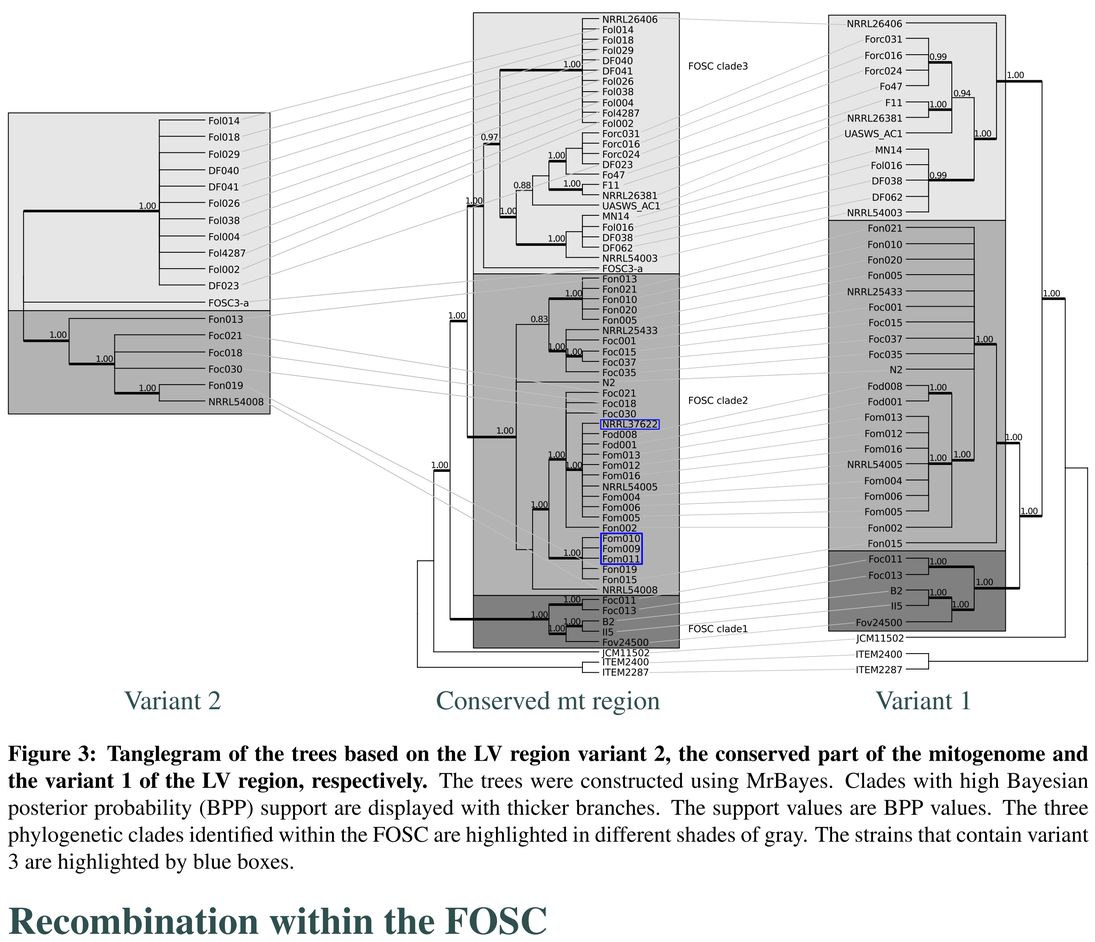

Figure 3: Tanglegram of the trees based on the LV region variant 2, the conserved part of the mitogenome and the variant 1 of the LV region, respectively. The trees were constructed using MrBayes. Clades with high Bayesian posterior probability (BPP) support are displayed with thicker branches. The support values are BPP values. The three phylogenetic clades identified within the FOSC are highlighted in different shades of gray. The strains that contain variant 3 are highlighted by blue boxes.

Recombination within the FOSC

Variant 2 emerged in the common ancestor of clades 2 & 3 of the FOSC and was maintained within both lineages during the separation of the two phylogenetic species. This shows that mitochondrial recombination is going on within the FOSC.

Conclusions

- Mitochondrial genomes can be efficiently assembled from NGS data
- It is feasible to analyze the mitogenome of a large number of strains
- Complete mitogenomes offer sufficient information to delineate even closely related species
- A detailed analysis of the mitogenomes may offer new insights into the biology of the organism

Funding

The investigations are supported by the Division for Earth and Life Sciences (ALW) with financial aid from the Netherlands Organization for Scientific Research (NWO) under grant number 833.13.006. This work was further supported by the Horizon programme (project 93512007) of the Netherlands Genomics Initiative (NGI) through a grant to Martijn Rep.

Reference

[1] Brankovics, B., van Dam, P., Rep, M., de Hoog, G. S., van der Lee, T. A. J., Waalwijk, C. and van Diepeningen, A. D. Mitochondrial genomes reveal recombination in the presumed asexual *Fusarium oxysporum* species complex. (Under review)