UPPSALA
UNIVERSITET

Phylogenetic distribution and role of Spok homologs in the genus Fusarium

Armin
Rassooli Tilehnovi

Degree project in biology, Bachelor of science, 2023
Examensarbete i biologi 15 hp till kandidatexamen, 2023
Biology Education Centre and Biology Education Center and Department of Organismal Biology, Uppsala University
Supervisor: Aaron Vogan

Abstract

Spok genes are found in the fungus *Podospora Anserina* and code for meiotic drivers. They specifically code for a toxin that kill spores not inheriting the gene, and therefore increasing their own transmission rate. Homologs of said genes also exist in the pathogenic fungi genus *Fusarium.* However, whether they act as meiotic drivers or not is unknown. The point of this study is to analyze the phylogenetic distribution of these homologs and see their effect on *Fusarium.*

This was done by analyzing publicly available genomes of different *Fusarium* strains and finding as many homologs as possible, and generating a phylogenetic tree. The phylogenetic tree showed that the genes have likely been horizontally transferred among different individuals. Some strains had a lot of homologs, where many were located on their accessory chromosomes that, that are related to their pathogenicity. There were some evidence, that the genes code for meiotic drivers or proteins causing similar effects, but more research will have to be done for a more conclusive answer.

Background

Understanding the evolutionary mechanisms among fungi is crucial for understanding how they spread, which is of special importance when studying pathogenic strains. Genetic elements important for the evolution of organisms can be part of intragenomic conflicts. Genes can act as selfish genes that compete with other genes in the genome, instead of working together. A common type of selfish gene are meiotic drivers. They increase their frequency during meiosis, often reducing the frequency of other genetic elements (Courret *et al.* 2019).

Spore killers

A selfish element used in some fungi are spore killers. Due to separation of alleles that occur during meiosis where different offspring inherit different alleles, some haploid offspring will have haplotypes containing the selfish gene and will kill the offspring with the alternative haplotype not containing the selfish gene. Half of the offspring will inherit the spore killer and kill the other half. The killing of spores can occur easily since all *ascomycete* fungi can package spores together in ascus, that contain different meiotic products, that can then interact with each other (Vogan *et al.* 2022).

Two mechanisms for spore killing have been discovered. One common mechanism is the antidote/poison mechanism. This mechanism works by coding for a toxin that kills other spores, and coding for an antidote that saves the spores killing, by binding the toxins, and working as inhibitors (Vogan *et al.* 2022). These mechanisms can be coded by selfish genes, or complexes of genes. These genes are then usually located in regions with suppressed recombination to be

assured they are inherited together, such as on the sex chromosome. In the case of poison/antidote mechanism, one gene can code for different proteins that are poison or antidote, or code for a protein with different domains responsible for antidote and poison. These genes can also be very widespread and spread quickly in populations, despite having proven to have a deleterious effect and potentially killing half the offspring. This is due to several factors making the effect less negative. First there can be a fitness advantage as the fungi can use resources from dead spores to grow more spores. The fungi can also modify their reproduction from sexual to asexual, to reduce the negative effect. They have also evolved resistance mechanisms (Vogan *et al*. 2022).

Spok genes

A group of genes confirmed to code for spore killing mechanisms are *Spok* genes. These are selfish genes discovered in the fungi *Podospora Anseria*. They code for single proteins with different domains responsible for both antidotes and poisons. Individual *Spok* genes have proved to work as meiotic driver independently (Vogan *et al*. 2019).

There are four *Spok* genes described; *Spok1*, *Spok2*, *Spok3* and *Spok4*. *Spok 3* and 4 are believed to have the closest physical relationship to each other and are found in a region on the genome called the *Spok* block (size of the region is 74-167 kb). They exist in different frequencies, where *Spok1* and *Spok2* exist in highest frequency, and are found in pretty much all strains, compared to the *Spok block* (Vogan *et al*. 2021).

The *Spok* block in *P. Anserina* can relocate itself through translocation and is therefore found on different locations on chromosomes 1,3 and 5 of different strains. They are only found once in the genome. The movement occurs by having tyrosine-recombinase –mobilized DNA transposon called *The Enterprise*, coded by its *Kirc* genes. Transposable elements are also selfish elements, and it is believed that the Enterprise, has parasitized the *Spok* block and can move it around and then spread even further by using the meiotic driver abilities of the block, acting as a superparasite. *Spok1* and *Spok2* are also believed to have transposable elements surrounding them, shifting them around on the chromosome, possibly though another transposition mechanism (Vogan *et al*. 2021). The *Spok* block seems to be associated with higher killing frequency; 50-90% killing efficacy compared to 40% efficacy measured on *Spok2* genes. The killing efficacy can increase depending on what genomic location it is in, such as being higher in centromeric regions. The block was shown in some instances to be deleterious, but not in other cases, and there does not seem to be a correlation between higher killing frequency and lower fitness. A hypothesis is that the insertion of the *Spok block* might affect recombination and can therefore, depending on where it's inserted, have a bigger effect on the organism (Vogan *et al*. 2021).

Thanks to transposition, the Spok genes can most likely transfer from species to species through horizontal gene transfer. The characteristics of transposons have also been shown to be effective for creating the DNA intermediates in their procedure stable enough to survive being transferred between organisms (Wells & Feschotte 2020). This would explain why some of the *Spok* genes exist in diverged lineages of *Podospora*, despite not being very divergent themselves in their structure. A possible origin for some *Spok* genes is therefore hybridization events. The gene can also spread through duplication. *Spok3* and *Spok4* are found in duplicated regions and are believed to possibly be paralogs, although an alternative explanation does exist (Vogan *et al*. 2019).

Phylogenetics of *Spok* homologs

Homologs of *Spok* genes with a similar gene structure, have been found in other ascomycetes. They are very frequent and very diverged. While many of them have not been confirmed spore killers, many have diverged, which could indicate meiotic drivers in fact are drivers that have increased their transmission *(*Vogan *et al.* 2019).

The phylogeny of *Spok* homologs is distributed into two clades. Clade 1 has fewer copies, and strains who only contain one copy of the homologs. Clade 2 contains *Spok*, *Spok2*, Spok3 and *Spok4*. Clade 2 is also where you find homologs belonging to strains, who have a lot of copies (Figure 7, Vogan *et al* 2019). Meaning clade 2 has more polymorphic variation and more likely to possess meiotic drivers. The proteins in the homologs of clade 2 seem to diversify in a similar manner to *Spok2* genes (Vogan *et al*. 2019).

The tree of the gene homologs does not correspond closely to the species tree. (Figure 7, Vogan *et al* 2019). This reinforces the idea that these types of genes can move from species to species through horizontal gene transfer, as that would lead to gene tree discordance. To support this, homologs to the *Kirc* genes, which were responsible for the transposition in *Spok* genes have also been found in some of the other species (Vogan *et al*. 2021).

The unusual phylogeny can also be explained due to patchy distributions, where some meiotic drivers have gone extinct in certain lineages due to their negative fitness. This would further explain the presence of pseudogenes (Grogner et al. 2014). Many of the homologs were most likely pseudogenes, due to having frameshift mutations and premature stop codons (Vogan *et al*. 2019).

Fuspoks

One genus of fungi where said homologs have been discovered is *Fusarium. Fusarium* is a genus of plant pathogens from the *Nectriacae* family, responsible for many issues such as maize ear and stalk rot, and ruin many crops through its mycotoxins. Analysis of different genes such as rRNA and *B-tubulin* showed consistently that the phylogeny has over a thousand species divided into seven clades, with high bootstrap values (the same observation is seen when generating the tree several times) (Watanabe *et al.* 2011).

The *Spok* homologs discovered in *Fusarium* [*Fuspoks*] seem to be associated with strain specific chromosomes (accessory chromosomes) that in turn are correlated with the fungi's ability to spread diseases (Vogan et al. 2019). Many of the *Spok* homologs found in *Fusarium* were respectively found in clade 1 and clade 2 previously described. Most have no proven spore killing abilities, with the expectation of one homolog found in *F.Vanettenii* (old name is *Nectria haemotocca),* after experiments where *F.Vanetenii* mated with *Podospora Anserina,* and the resulting offspring had empty ascari (Grognet *et al.* 2014). Many strains of *Fusarium Oxysporum* had homologs belonging to the second clade, with more than one copy of the gene indicating possible spore killing. Even a strain from clade 1 (*F. Oxysporum f.sp pisi* (Fop)) had multiple copies of the gene, in contrast to most strains of clade 1. Another strain of *Oxysporum* had 4 copies of its gene which is very similar to the homologs of clade 2.

Mutated *Spok3* showed killing abilities in vegetative tissue. Both toxins and antidote, produced by *Spok3* have been observed in vegetative cells. This would correlate well with how killing could operate in *Fusarium*, since many of them have no sexual cycle. A vegetative cell produces the toxin and antidote that exists in the cytoplasm. The proteins will then exist in both new cells that are formed during mitosis, and then its speculated that the toxin kills the cell not containing the *Fuspok*. However due to the discovery of alleles related to sexual reproduction, such a cycle might still possibly exist in species where it has not been observed (Vogan *et al.* 2019). Transforming of strains of F. *Graminearum* with the *Spok1* gene proved that *Spok* genes caused spore killing, indicating that *Fuspoks* could have similar effects in natural Fusarium strains (Gardiner *et al.* 2020).

Aim

The aim is to study phylogenetic distribution of *Fuspoks* in *Fusarium.* This is done to discover if there is a connection to the accessory chromosomes. Its also done to find out if the *Fuspoks*, just like their homologs in *P. Ansperina* act as meiotic drivers, or if they code for either spore or vegetative killing. This will give a better understanding of the evolutionary mechanisms in *Fusarium*, which can later be used in future applications when dealing with the fungi.

Method

Downloading sequences

61 genomes of different *Fusarium* strains that were de-novo assembled were downloaded from NCBI, and later used (some other genomes not following the criteria had to be removed later in the analysis). The genomes were compiled into a big file, that was later used to build a database using BLAST software (version 2.12.0). A Tblastn (version 2.12.0) was done later with parameters: e-value 0.001, outfmt 6 and word-size 6. The *Spok3* protein sequence (MK521590) was "blasted" against the genomes.

Building a database with the gene homologs

All hits from the Tblastn that belonged to the same gene, were patched using a custom made script in python. The nucleotide sequences with entire haplotypes for all hits were then retrieved with another script. All of the genes were compiled into a new file with their sequences in fasta format.

Some of the sequences had to be reverse complemented. This was done by firstly gathering all sequences from the Tblastn table were read in the reverse direction. The full haplotypes for these sequences were retrieved, reverse complemented, and later replaced the sequences in the compiled file with similar sequence id.

Filtrating out pseudogenes

After retrieving all the F*uspok* genes, the pseudogenes had to be removed. This had to be done since nonfunctioning genes could create long branches which could cause issues when generating a phylogenetic tree. The nucleotide sequences of the genes were therefore translated into protein sequences, and all sequences with premature stop codons were identified and removed.

Alignment

The sequences were then aligned using MACSE_V2.07 (default settings). Sequences of *Spok2* (CDP29201.1), *Spok3* (MK52188) and *Spok4* (MK521590) were also added to the alignment file before aligning, to be used as reference in the phylogenetic tree after. They were trimmed so that the coding regions were used.

Generating the tree

After aligning, any further pseudogenes with frameshift mutations were discovered and removed. Using Iqtree version 2.07 , a phylogenetic tree using aligned sequences was generated. Iqtree automatically chose the best fit substation model (GTR+I+G) and generated the tree, with 1000 ultrafast bootstrap replicates. The tree was later visualized using ITOL (Se figure 3 and 4). A

second tree using highly trimmed sequences was also generated with similar settings, to see if the trimming would change the tree.

Analyze Data

The data was later analyzed using custom made python scripts, that counted the *Fuspoks* (non-pseudogenes and pseudogenes) for each strain, and categorizing the strains based on species.

Result

Finding the *Fuspoks*

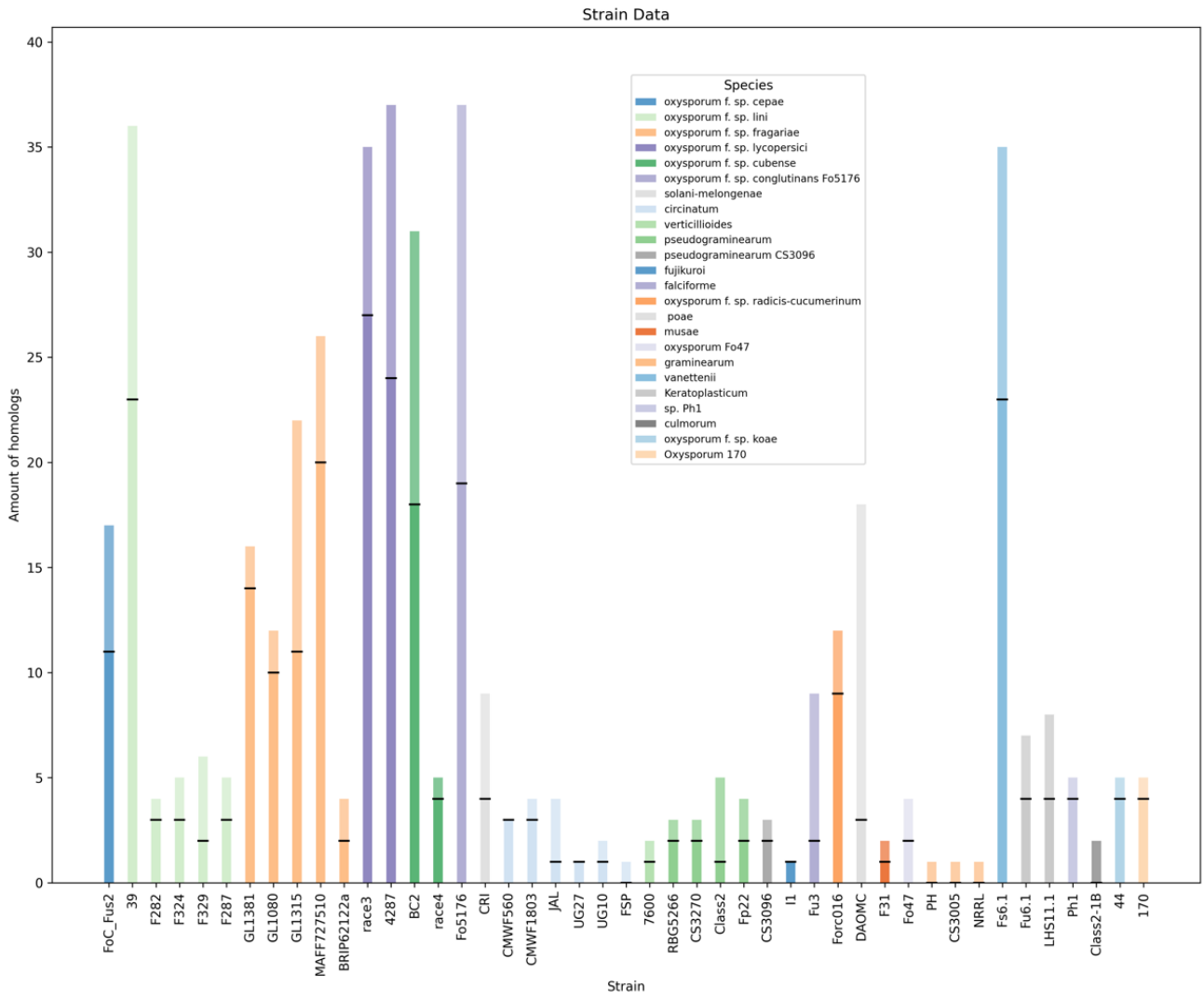
"Blasting" the *Spok3* protein sequence against the database of *Fusarium* genomes, resulted in 539 hits. After stitching up all hits that belonged to the same sequence, 456 sequences were left, whose haplotype was retrieved. 239 sequences were read in the wrong direction and had to be reverse complemented.

Removing Pseudogenes and badly sequenced genes

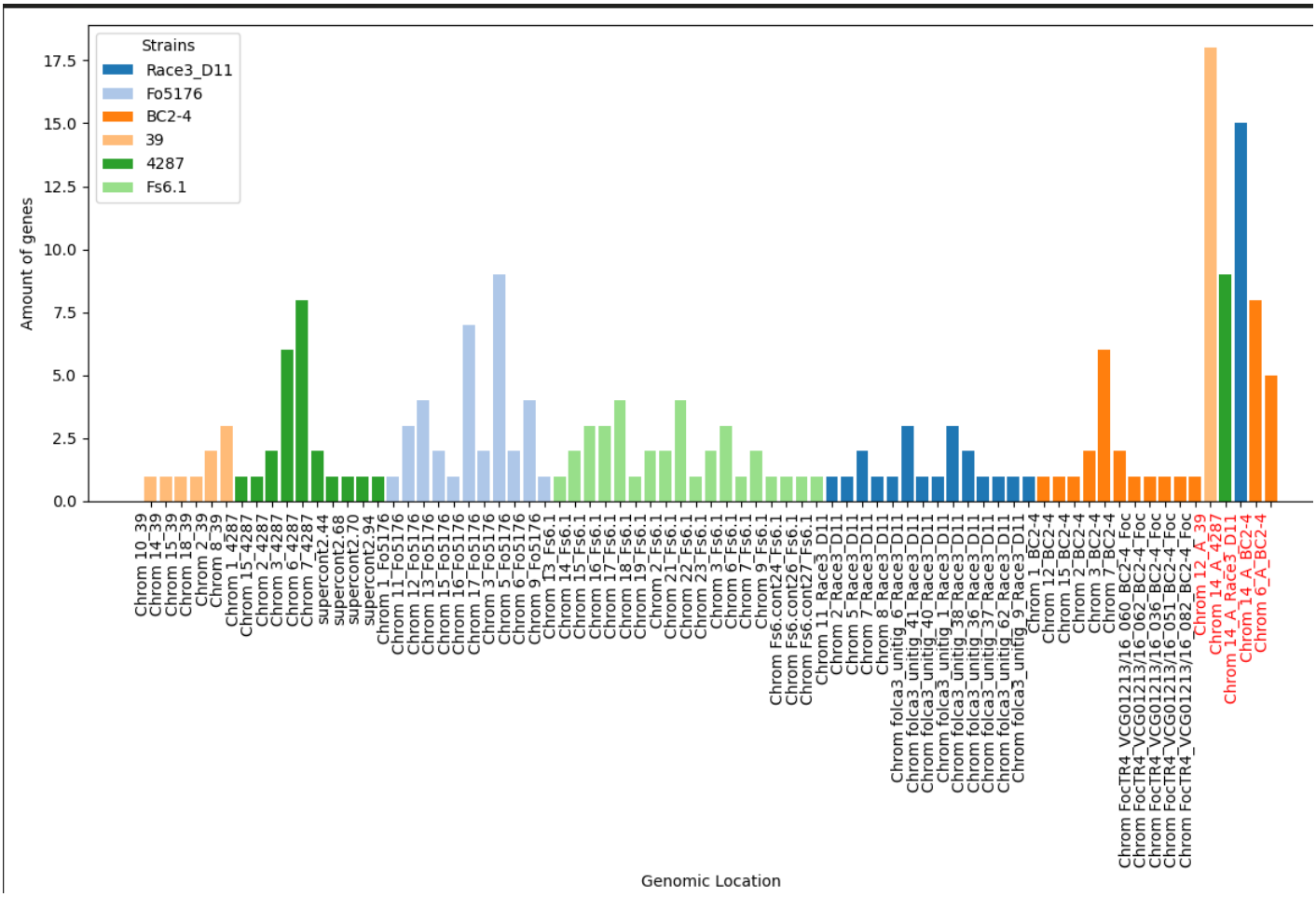
Pseudogenes were identified by finding sequences with premature stop codons and looking at the alignment and discovering frame shift mutations. We had 273 *Fuspoks*, and 183 pseudogenes.

Total amount of *Fuspoks* in different strains

Six strains contained a very large amount of *Fuspoks* (figure 1). *F. Oxysporum f. sp. Lycopersici strain 4287* and. *F. Lycopersici Strain* race3_D11 had many on chromosome 14. Chromosome 14 in Lycopersici is an accessory chromosome related to disease (Ma *et al.* 2010). *F. f. sp. Cubense race BC2* had a lot, that were mainly chromosomes 14, 6 and 5 as well. 14 and 6 has shown to be accessory in *Cubense* too (Guo *et al.* 2014). One strain of *F. Oxysporum* f. sp. lini strain 39 compared to the rest of that species had a lot of Fuspoks, mainly on its chromosome 12 and 13, and chromosome 12 is associated with disease (Samsonova *et al.* 2021). *F. Oxysporum f.sp. Conglutians Fo176* had a similar amount *Fuspoks*, but they were located on several different chromosomes; 16, 11 and 3 (figure 2).



(Figure 1. Total amount of Fuspoks in each strain. All strains of similar color next to each are of the same species. The lighter shaded parts of the bars after the black line are total amount of pseudogenes. If no black bar is nos found, only pseudogenes exist)



(Figure 2: Amount of Fupoks on different genomic locations in six of the strains. Red color is an accessory chromosome)

Phylogenetic tree

Using maximum likelihood models, a phylogenetic tree with all 273 *Fuspoks* that were putatively functional, together with 3 *Spok* genes were generated. All clades, with the expectation of a few specified lineages, were highly supported. There are four big clades, and a few small divergent branches. (Full tree in appendix)

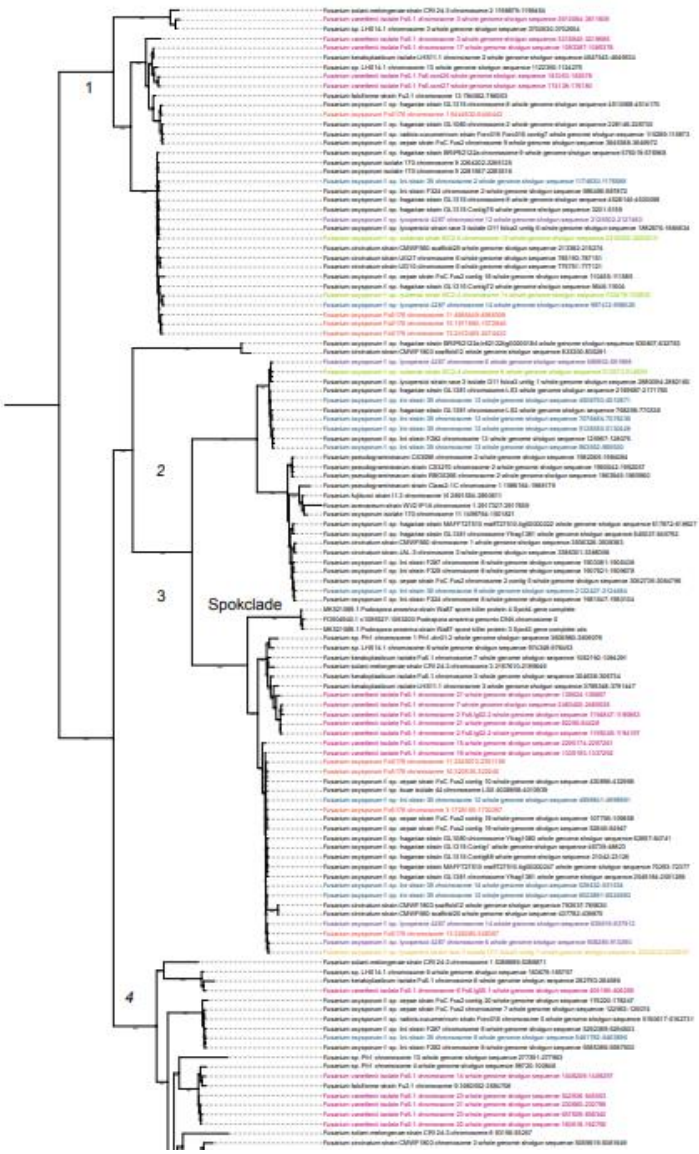
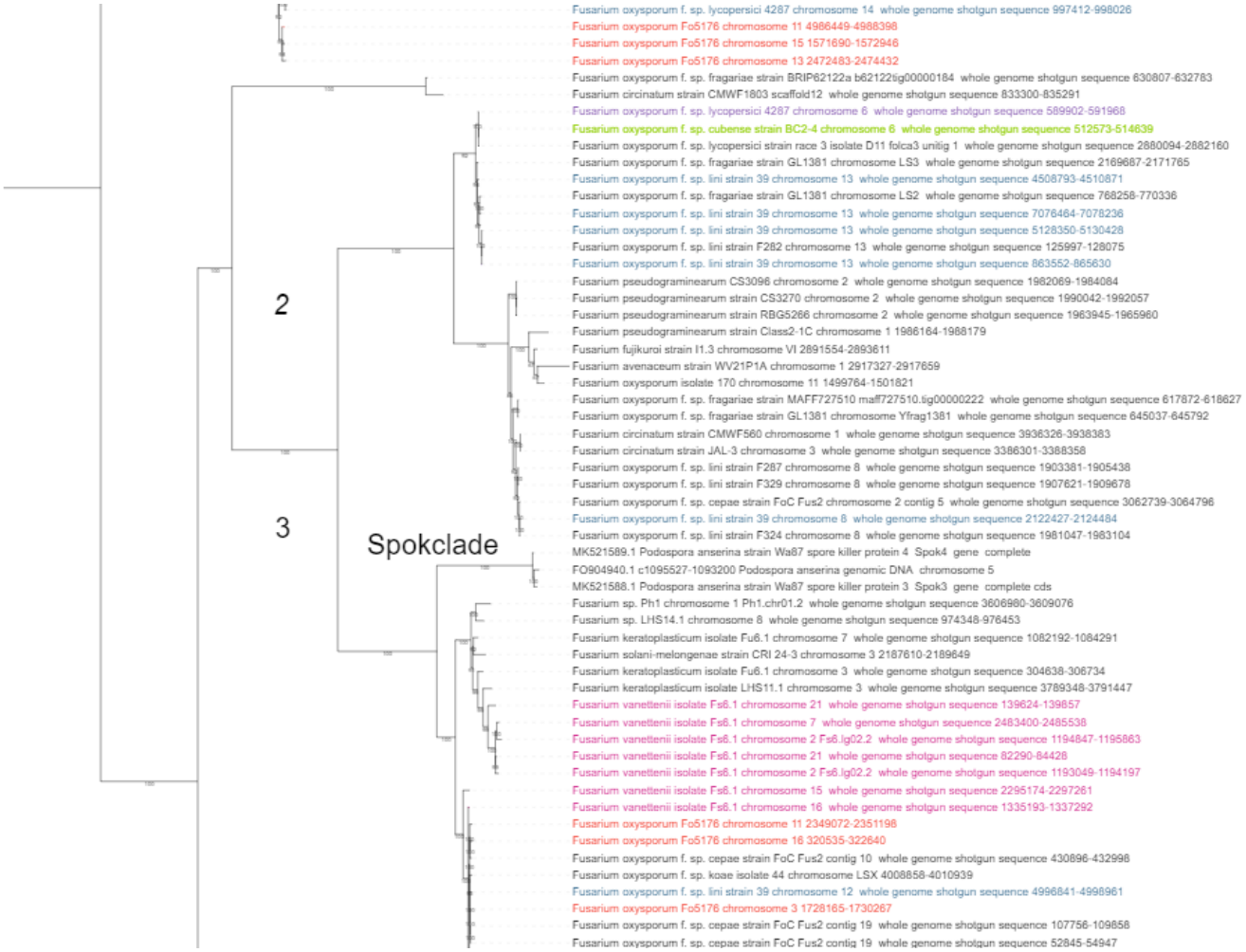


Figure 3. (All 4 clades of the phylogenetic tree)

Strain BC2-4: Green, strain 39: blue, strain 4287: purple, strain Fs6.pink: strain Fo5176; red, strain race3: D11: yellow color)



(Figure 4. Clade 3 and Clade 4 of Phylogenic tree of Fuspoks,

Strain BC2-4: Green, strain 39: blue, strain 4287: purple, strain Fs6.pink: strain Fo5176; red, strain race3: D11: yellow color)

Discussion

Firstly, by looking at the phylogeny, the 4 clades described here correlate well with the phylogenetic tree described in previous literature. Clades 2 and 3 described in figure 4 correspond to clades 1 and 2 (Vogan *et al*. 2019). There is clear discordance between the species tree and the gene tree. This can firstly be seen by the fact that several of the genes from the same strain or species do not cluster together, creating several clearly divided clusters. There are many instances in this gene tree where individual genes from different species are closer apart than individual genes from their own species. There are even instances where different genes of the same strain, but different genomic locations are closer to genes from other strains than each other. This further supports that horizontal gene transfer of the *Fuspoks* is most likely occurring between different individuals. This pattern could be observed, if there was a loss of some of the

genes in certain lineages. However had that been the case, a larger difference in divergence between the genes would have been expected.

Six strain had a lot of *Fuspoks.* The fact that many *Oxysporum* strains have a lot of *Fuspoks* correlates well with previous literature (Vogan *et al* 2019). However, this could very likely be due to the fact that strains of said species have more sequenced data. It was not only *Oxysporum* strains that had many though. *F. Vanetti Fs6.1* had a range of them spread among different chromosomes. These were also very close to the *Spok* genes in the tree. The F*uspoks* found on these strains are most likely also horizontally transferred due to existing everywhere randomly among several of the clades in the phylogeny similarly as all the other genes. Some of them even existed among the most divergent clades. Interestingly, far more copies and strains with many copies exist in clade 3 compared to clade 2, which is similar to previous results (Vogan *et al.* 2019). This makes sense since clade 3 is where the *Spok* genes exist. While they are widespread, in many cases one can find several copies next to each other, indicating said genes have a common origin. Many are most likely orthologs, and those on the same genomic location are possibly paralogs. This is especially true for several of the copies of *Vanetti Fs6.1* copies in clade 3.

Many of these strains had the *Fuspoks* on known accessory chromosomes. Their ability to perform Horizontal gene transfer would correlate well with this. Horizontal chromosome transfer has been shown to move accessory chromosomes to transfer pathogenic genes than can then diversify among lineages. This has specifically been documented in *F. Oxysporum* (Henry *et al.* 2021). It is speculated that the *Fuspoks* can spread in a similar manner. Since the strains have a lot of copies as well, the species (not counting *F. Vanettenii*) have a lot of polymorphic variation, which could also indicate meiotic drive. The horizontal gene transfer would also make sense if the genes were meiotic drivers. Just like the Enterprise moved the *Spok* genes, possible transposable elements could move these drivers, and transposable elements would be more inclined to do horizontal gene transfer. However, there is no big difference in divergence among lineages, which could have been expected if some of the lineages were meiotic drivers and others not.

The strains with a lot of *Fuspoks* had a lot of pseudogenes. A lot of lost pseudogenes could potentially indicate lost meiotic drivers (Vogan *et al.* 2019). Some strains only had pseudogenes such as *F. Gramineraum PH* and *F. Circinatum FSP*, and naturally those genes would not have any function, which is supported by the fact that there was little variation in number of homologs the species those strains belong to.

The big issue with the possibility of these genes coding for meiotic drivers in all of the strains in *F. Oxysporum* is the lack of evidence of a sexual cycle in that species. (McTaggart *et al.* 2021).

However, as explained in the section on *Fuspoks* in the background, possibilities of vegetative killing still exists. This could cause similar effects as spore killing (Vogan *et al.* 2019). *F.Vanettenii* does actually have a sexual cycle (Xie *et al.* 2022), and had shown spore killing abilities when having Spok1 inserted in them (Grogner *et al.* 2014). So despite the lack of polymorphic variation, more research on *F.Vanettenii's* possible meiotic drive abilities should definitely be done.

Future research to verify if the genes code for spore or vegetative killing is necessary. This could possibly be done with site-directed mutagenesis experiments or using Crispr to remove the gene, and then analyzing the difference in phenotype. Further research looking into *Kirc* homologs or any other similar gene responsible for transposition close to these genes would also help to discover the potential function of them. Studying meiotic drivers that are not *Spok* genes in *Fusarium* could also give more insight into how spore killings works in the genus, as other loci in some species such *as F. Monoliliforme* have shown to be responsible for spore killing (Kathariou & Spieth 1982).

One issue with analyzing these sequences, is that a lot of them still had gaps in their coding regions. This can cause issues during the alignment and generate a less accurate tree. Removing a lot of gaps did not significantly change the phylogenetic tree though, so the effect is most likely not profound. However future studies verifying the validity of phylogeny would be necessary. This would be extra important when trying to understand the dynamics of the possible horizontal gene transfer better, and for knowing exactly how the genes have transferred in the tree.

Conclusion

In conclusion, analyzing the phylogenetic tree of the *Fuspoks* showed a clear difference between gene distribution and species distribution, indicating these genes are horizontally transferred. It was clear that six strains contained a very high amount of *Fuspoks*, and that there is most likely a correlation between those genes on these strains and accessory chromosomes. The results also highlight the possibility that these genes, just like the *Spok* genes code for meiotic drivers, or for some type of killing abilities. Future research to verify the function of the genes, especially on those six strains must be done, to get a much better understanding of the genes and their effect on *Fusarium*.

References

Courret C, Chang C-H, Wei KH-C, Montchamp-Moreau C, Larracuente AM. 2019. Meiotic drive mechanisms: lessons from Drosophila. Proceedings of the Royal Society B: Biological Sciences 286: 20191430.

Gardiner DM, Rusu A, Barrett L, Hunter GC, Kazan K. 2020. Can natural gene drives be part of future fungal pathogen control strategies in plants? New Phytologist 228: 1431–1439.

Grognet P, Lalucque H, Malagnac F, Silar P. 2014. Genes That Bias Mendelian Segregation. PLoS Genetics 10: e1004387.

Guo L, Han L, Yang L, Zeng H, Fan D, Zhu Y, Feng Y, Wang G, Peng C, Jiang X, Zhou D, Ni P, Liang C, Liu L, Wang J, Mao C, Fang X, Peng M, Huang J. 2014. Genome and Transcriptome Analysis of the Fungal Pathogen Fusarium oxysporum f. sp. cubense Causing Banana Vascular Wilt Disease. PLoS ONE 9: e95543.

Henry PM, Pincot DDA, Jenner BN, Borrero C, Aviles M, Nam M-H, Epstein L, Knapp SJ, Gordon TR. 2021. Horizontal chromosome transfer and independent evolution drive diversification in Fusarium oxysporum f. sp. fragariae. New Phytologist 230: 327–340.

Kathariou S, Spieth PT. 1982. SPORE KILLER POLYMORPHISM IN FUSARIUM MONILIFORME. Genetics 102: 19–24.

Ma L-J, van der Does HC, Borkovich KA, Coleman JJ, Daboussi M-J, Di Pietro A, Dufresne M, Freitag M, Grabherr M, Henrissat B, Houterman PM, Kang S, Shim W-B, Woloshuk C, Xie X, Xu J-R, Antoniw J, Baker SE, Bluhm BH, Breakspear A, Brown DW, Butchko RAE, Chapman S, Coulson R, Coutinho PM, Danchin EGJ, Diener A, Gale LR, Gardiner DM, Goff S, Hammond-Kosack KE, Hilburn K, Hua-Van A, Jonkers W, Kazan K, Kodira CD, Koehrsen M, Kumar L, Lee Y-H, Li L, Manners JM, Miranda-Saavedra D, Mukherjee M, Park G, Park J, Park S-Y, Proctor RH, Regev A, Ruiz-Roldan MC, Sain D, Sakthikumar S, Sykes S, Schwartz DC, Turgeon BG, Wapinski I, Yoder O, Young S, Zeng Q, Zhou S, Galagan J, Cuomo CA, Kistler HC, Rep M. 2010. Comparative genomics reveals mobile pathogenicity chromosomes in Fusarium. Nature 464: 367–373.

McTaggart AR, James TY, Shivas RG, Drenth A, Wingfield BD, Summerell BA, Duong TA. 2021. Population genomics reveals historical and ongoing recombination in the Fusarium oxysporum species complex. Studies in Mycology 99: 100132.

Samsonova A, Kanapin A, Bankin M, Logachev A, Gretsova M, Rozhmina T, Samsonova M. 2021. A Genomic Blueprint of Flax Fungal Parasite

Fusarium oxysporum f. sp. lini. International Journal of Molecular Sciences 22: 2665.

Vogan AA, Ament-Velásquez SL, Bastiaans E, Wallerman O, Saupe SJ, Suh A, Johannesson H. 2021. The Enterprise, a massive transposon carrying Spok meiotic drive genes. Genome Research, doi 10.1101/gr.267609.120.

Vogan AA, Ament-Velásquez SL, Granger-Farbos A, Svedberg J, Bastiaans E, Debets AJ, Coustou V, Yvanne H, Clavé C, Saupe SJ, Johannesson H. 2019. Combinations of Spok genes create multiple meiotic drivers in Podospora. eLife 8: e46454.

Vogan AA, Martinossi-Allibert I, Ament-Velásquez SL, Svedberg J, Johannesson H. 2022. The spore killers, fungal meiotic driver elements. Mycologia 114: 1–23.

Watanabe M, Yonezawa T, Lee K, Kumagai S, Sugita-Konishi Y, Goto K, Hara-Kudo Y. 2011. Molecular phylogeny of the higher and lower taxonomy of the Fusarium genus and differences in the evolutionary histories of multiple genes. BMC evolutionary biology 11: 322.

Wells JN, Feschotte C. 2020. A Field Guide to Eukaryotic Transposable Elements. Annual review of genetics 54: 539–561.

Xie S-Y, Ma T, Zhao N, Zhang X, Fang B, Huang L. 2022. Whole-Genome Sequencing and Comparative Genome Analysis of Fusarium solani-melongenae Causing Fusarium Root and Stem Rot in Sweetpotatoes. Microbiology Spectrum 10: e00683-22.

Appendix

Full tree:

Genome

Fusarium species: https://www.ncbi.nlm.nih.gov/data-hub/genome/?taxon=5506&assembly_level=2%3A3

Spok3:
MK52188

https://www.ncbi.nlm.nih.gov/nuccore/MK521588

Spok4:
https://www.ncbi.nlm.nih.gov/nuccore/MK521589

MK52189

Spok2:

https://www.ncbi.nlm.nih.gov/protein/CDP29201.1/

CDP29201.1

Strain information

Strains/genomes used from the different species

(first number is functional Fusspoks, and second is fusspoks)

Fusarium oxysporum f. sp. cepae:

FoC_Fus2: 11 6

Fusarium oxysporum f. sp. lini:

39: 23 13

F282: 3 1

F324: 3 2

F329: 2 4

F287: 3 2

Fusarium oxysporum f. sp. fragariae:

GL1381: 14 2

GL1080: 10 2

GL1315: 11 11

MAFF727510: 20 6

BRIP62122a: 2 2

Fusarium oxysporum f. sp. lycopersici:

race3 : 27 8

4287: 24 13

Fusarium oxysporum f. sp. cubense:

BC2: 18 13

race4: 4 1

Fusarium oxysporum f. sp. conglutinans Fo5176:

Fo5176: 19 18

Fusarium solani-melongenae:

CRI: 4 5

Fusarium circinatum:

CMWF560: 3 0

CMWF1803: 3 1

JAL: 1 3

UG27: 1 0

UG10: 1 1

FSP: 0 1

CMWF2633: 0 0

Fusarium verticillioides:

7600: 1 1

BRIP53590: 0 0
BRIP53263: 0 0
RBG5266: 0 0

HN2117187: 0 0

Fusarium pseudograminearum:

RBG5266: 2 1

CS3270: 2 1

Class2: 1 4

Fp22: 2 2

Fusarium pseudograminearum CS3096:

CS3096: 2 1

Fusarium fujikuroi:

I1: 1 0

IMI 58289: 0 0
CSV1: 0 0

Augusto2: 0 0

Fusarium falciforme:

Fu3: 2 7

Fusarium oxysporum f. sp. radicis-cucumerinum:

Forc016: 9 3

Fusarium poae:

DAOMC: 3 15

Fusarium musae:

F31: 1 1

Fusarium oxysporum Fo47:

Fo47: 2 2

Fusarium venenatum:

A3: 0 1

Fusarium avenaceum:

WV21P1A: 1 0

Fusarium graminearum:

PH: 0 1

CS3005: 0 1

NRRL: 0 1

FG-12: 0 0

Fusarium vanettenii:

Fu6.1: 4 3

Fusarium Keratoplasticum:

Fu6.1: 4 3

LHS11.1: 4 4

Fusarium sp. Ph1:

Ph1: 4 1

Fusarium culmorum

Class2-1B: 0 2

Fusarium venenatum

A3/5: 0 1

Fusarium oxysporum:

170: 4 1

Fusarium oxysporum f. sp. koae:

44: 4 1

Fusarium asiaticum

KCTC 16664: 0 0

Fusarium culmorum 5516
5516: 0 0

Software used

Iqtree2: http://www.iqtree.org/

Itol: https://itol.embl.de/

TblastnN: https://blast.ncbi.nlm.nih.gov/doc/blast-help/downloadblastdata.html

Macse: https://bioweb.supagro.inra.fr/macse/

Everything was done using custom made scripts in Ubuntu/Linux:
https://github.com/Armin-RT/Guide-to-phylogenetic-Tree

Filtration of sequences

One gene was found to have a frameshift mutation during alignment:
Fusarium_oxysporum_f._sp._fragariae_strain_GL1315_Contig34,_whole_genome_shotgun_sequence_7610-8086) was also found to have a frameshift mutation during the alignment.

During the analysis, genes from a few genomes had to be removed. There were also 2 duplications (*from Fusarium graminearum strain PH-1*) So I recounted how many genes in each step that existed without these. (using custom made script)

This is the original count

702 hits: (539)
602 stitched (456)

Functioning Fuspoks: 369 (273)

Pseudogenes: 233 (183) -with the GL1315 one counted

4 strains

- Fusarium oxysporum f. sp. cubense isolate VCG0125
- Fusarium oxysporum f. sp. cubense isolate VCG0120
- Fusarium oxysporum f. sp. cubense strain TC1-1
- Fusarium oxysporum f. sp. cubense race 1 isolate VCG01220

2 duplications removed:
Fusarium graminearum genome assembly, chromosome: II_4155879-4157538

Fusarium graminearum chromosome 2, complete genome_4156656-4158354

Full Phylogenetic Tree

Tree scale: 10