

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ

VITÓRIA LUISA CAVALHEIRO

**ANÁLISE GERAL DO PANGENOMA DE *Setosphaeria spp.* COM ENFOQUE NOS
GENES DE PATOGENICIDADE**

DOIS VIZINHOS

2023

VITÓRIA LUISA CAVALHEIRO

**ANÁLISE GERAL DO PANGENOMA DE *Setosphaeria spp.* COM ENFOQUE NOS
GENES DE PATOGENICIDADE**

**GENERAL ANALYSIS OF THE PANGENOME OF *Setosphaeria spp.* FOCUSING
ON THE PATHOGENICITY GENES**

Trabalho de conclusão de curso de Especialização apresentado como requisito para obtenção do título de Especialização em Biologia Molecular – Habilitação Bioinformática da Universidade Tecnológica Federal do Paraná (UTFPR).

Orientador(a): Profa. Dra Juliana Morini Küpper Cardoso Persegui.

Coorientador(a): Profa. Dra Naiana Cristine Gabiatti.

DOIS VIZINHOS

2023



[4.0 Internacional](https://creativecommons.org/licenses/by-nc-nd/4.0/)

Esta licença permite download e compartilhamento do trabalho desde que sejam atribuídos créditos ao(s) autor(es), sem a possibilidade de alterá-lo ou utilizá-lo para fins comerciais. Conteúdos elaborados por terceiros, citados e referenciados nesta obra não são cobertos pela licença.

VITÓRIA LUISA CAVALHEIRO

**ANÁLISE GERAL DO PANGENOMA DE *Setosphaeria spp.* COM ENFOQUE NOS
GENES DE PATOGENICIDADE**

Trabalho de conclusão de curso de Especialização apresentado como requisito para obtenção do título de Especialização em Biologia Molecular – Habilitação Bioinformática da Universidade Tecnológica Federal do Paraná (UTFPR).

Data de aprovação: 16/fevereiro/2023

Deborah Catharine de Assis Leite
Doutorado
Universidade Tecnológica Federal do Paraná - Campus Dois Vizinhos

Fábio Antônio Antonelo
Mestrado
Universidade Tecnológica Federal do Paraná - Campus Dois Vizinhos

Nédia de Castilhos Ghisi
Doutorado
Universidade Tecnológica Federal do Paraná - Campus Dois Vizinhos

DOIS VIZINHOS

2023

Dedico este trabalho à minha família,
pois trilharam o caminho mais difícil para mim.
Ao meu companheiro,
por mais motivos do que eu consigo descrever.

AGRADECIMENTOS

Nesta página gostaria de agradecer a todos que contribuíram de alguma forma para realização deste trabalho.

Agradeço à minha orientadora, Prof^a Dr^a Juliana Morini Küpper Cardoso Perseguini, por toda a ajuda durante a trajetória.

Ao laboratório de Análises Biológicas e Biologia Molecular (BioMol) da Universidade Tecnológica Federal do Paraná - Câmpus Dois Vizinhos.

Aos meus colegas de trabalho, que se tornaram amigos.

Aos meus amigos, que estiveram comigo em todos os momentos importantes dos últimos anos, e estarão para toda a vida.

Aos demais professores do curso, pela cooperação e conhecimento compartilhado.

A minha família que trilhou o caminho mais difícil para me proporcionar tudo ao que tive acesso.

Ao meu noivo, por sempre acreditar em mim, até quando eu mesma deixo de fazer.

A todos os que por algum motivo contribuíram para a realização desta pesquisa.

Cada um de nós é uma colcha de retalhos
daqueles que nos amaram, daqueles que
acreditaram em nosso futuro, daqueles que
nos mostraram empatia e bondade ou nos
disseram a verdade mesmo quando não era
fácil de ouvir.

(SWIFT, 2022)

RESUMO

Conhecido amplamente como patógeno responsável pela helmintosporiose em diversas culturas de plantas, *Setosphaeria turcica* é responsável por perdas econômicas de até 40% em diferentes plantios. Apesar do gênero *Setosphaeria* ser tradicionalmente composto por fitopatógenos, relatos recentes apontam para a possibilidade de patogenicidade em gatos e em humanos. Dentre os objetivos deste trabalho estão a análise pangênômica e a compreensão dos fatores de virulência e patogenicidade associados a este organismo. As sequências foram obtidas da base de dados do *National Center for Biotechnology Information*, para anotação dos genomas foi utilizado o programa genSAS (V.6.0), uma plataforma que fornece um *pipeline* para anotação estrutural e funcional do genoma inteiro para eucariotos e procariotos com uma variedade de ferramentas para mascaramento de repetição, previsão de modelos de genes e outros recursos. Os programas FastANI e OrthoVenn2 (V.2) foram utilizados para gerar os dados e as imagens relacionadas à análise pangênômica. Foram analisadas cinco diferentes cepas de *S. rostratum* e *S. turcica*, dentre as quais apenas três puderam ser adicionadas ao *pipeline* de anotação funcional, pois não possuíam os requisitos de menos de 25.000 sequências totais e mais de 50% das sequências contendo mais de 2.500 bases. Os resultados demonstraram que apenas a cepa do organismo de *S. rostratum* teve cerca de 6% dos seus clusters associados com processos metabólicos de lipídio e 17% para o processamento de macromoléculas, frequentemente associadas ao mecanismo de penetração no hospedeiro. A mesma cepa apresentou sete clusters relacionados à patogênese, que quando investigados demonstraram ser proteínas potencialmente associados ao citocromo p450, tioredoxina redutase e proteínas similares às de outros fitopatógenos. Genes codificadores dessas proteínas estão atrelados a produção de micotoxinas da classe fumosinas, cuja ingestão prolongada pode causar carcinogênese de esôfago em humanos. O presente estudo evidencia a falta de informações e a importância de novos sequenciamentos de diferentes linhagens desse organismo, bem como a importância da investigação dos diferentes genes para infecção de diferentes hospedeiros.

Palavras-chave: helmintosporiose; fitopatógenos; Bioinformática.

ABSTRACT

Widely known as the pathogen responsible for helminthosporiasis in several plant crops, *Setosphaeria turcica* is responsible for economic losses of up to 40% in different plantations. Although the genus *Setosphaeria* is traditionally composed of phytopathogens, recent reports point to the possibility of pathogenicity in cats and humans. Among the objectives of this work are the pangenomic analysis and the understanding of the virulence and pathogenicity factors associated with this organism. The sequences were obtained from the National Center for Biotechnology Information database, for genome annotation the genSAS program (V.6.0) was used, a platform that provides a pipeline for structural and functional annotation of the entire genome for eukaryotes and prokaryotes with a variety of tools for repeat masking, gene model prediction, and other features. FastANI and OrthoVenn2 (V.2) programs were used to generate data and images related to the pangenomic analysis. Five different strains of *S. rostratum* and *S. turcica* were analyzed, among which only three could be added to the functional annotation pipeline, as they did not have the requirements of less than 25,000 total sequences and more than 50% of the sequences containing more than 2,500 bases. The results showed that only the strain of the *S. rostratum* organism had about 6% of its clusters associated with lipid metabolic processes and 17% for the processing of macromolecules, often associated with the host penetration mechanism. The same strain showed seven clusters related to pathogenesis, which when investigated proved to be proteins potentially associated with cytochrome p450, thioredoxin reductase and proteins similar to those of other phytopathogens. Genes encoding these proteins are linked to the production of mycotoxins of the fumonisin class, whose prolonged ingestion can cause esophageal carcinogenesis in humans. The present study highlights the lack of information and the importance of new sequencing of different strains of this organism, as well as the importance of investigating the different genes for infection in different hosts.

Keywords: northern corn leaf blight; phytopathogens; bioinformatics.

SUMÁRIO

1	INTRODUÇÃO	13
1.1	Objetivos	14
1.1.1	Objetivo Geral	14
1.1.2	Objetivos Específicos	14
2	DESENVOLVIMENTO	15
2.1	Revisão de literatura	15
2.2	Material e métodos	17
2.3	Resultados	19
2.4	Discussão	26
3	CONCLUSÃO	29
	REFERÊNCIAS	30

1 INTRODUÇÃO

Classificado como um fungo ascomiceto heterotático, o organismo *Sethosphaeria turcica* é amplamente conhecido por ser o agente causador da helmintosporiose em plantas de sorgo, milho (Zhang, 2012) e, por vezes, alguns relatos em plantas de feijão. Periodicamente epidemias desta doença são relatadas, que resultam em grandes perdas na produtividade de milho (Raymundo, Hooker, 1981; Perkins, Pedersen, 1987), principalmente em áreas com temperatura moderada e alta umidade (Carson, 1999).

Descrito pela primeira vez em 1876 com o nome de *Helminthosporium turcicum*, esse organismo atualmente possui nomes distintos para suas fases de reprodução, *S. turcica* para a forma sexuada, e *Exserohilum turcicum* para a forma assexuada (Caligiorno *et al.*, 1999), mas que essencialmente se referem ao mesmo organismo. Estima-se que casos avançados dessa doença geram perdas em plantações de milho, entre 30% até 50% do rendimento no plantio, caso a doença se estabeleça em uma planta mais jovem (De Rossi *et al.*, 2022)

Tradicionalmente, esse gênero é composto apenas por fitopatógenos e infectam seus hospedeiros com alta especificidade (Tang *et al.*, 2015), entretanto, relatos recentes por Amaya *et al.* 2016 observaram doenças cutâneas, infecções sistêmicas e locais em pacientes imunossuprimidos e imunocompetentes causadas por estes patógenos.

Análises pangenômicas são análises do repertório de genes observados em uma determinada espécie ou grupo taxonômico, esse tipo de estudo pode auxiliar a compreender e investigar os diferentes mecanismos de patogenicidade e fatores de virulência, avaliando de forma completa variabilidades genéticas existentes conferidas a população de uma única região, por exemplo, o que não seria possível caso fosse utilizado apenas um genoma de referência (Rouli *et al.*, 2015). Este trabalho visa compreender parte dos fatores que tornam os organismos da espécie *Sethosphaeria* tão complexos, e uma classe eficiente de patógenos capazes de infectar diversos hospedeiros .

1.1 Objetivos

1.1.1 Objetivo Geral

Anotar, classificar, analisar e compreender os mecanismos de patogenicidade do gênero *Setosphaeria*.

1.1.2 Objetivos Específicos

- Anotar os genomas das cepas disponíveis;
- Analisar o pangenoma dos organismos do gênero *Setosphaeria*;
- Descrever os genes de patogenicidade e fatores de virulência envolvidos na infecção provocada pelos organismos do gênero *Setosphaeria*, de acordo com as informações disponíveis.

2 DESENVOLVIMENTO

2.1 Revisão de literatura

O organismo do gênero *Sethosphaeria* possuem classificação taxonômica pertencentes à subclasse *Loculoascomycetidae*, ordem *Pleosporales*, é um ascomiceto, ou seja, pertencente ao Filo *Ascomycota*, e, portanto, classificado como fungo. Por ser um ascomiceto, o *S. turcica* produz seus esporos em esporângios específicos que são denominados de ascos. Este fungo pode ser chamado também de heterotático, pois hiberna em restos de plantas hospedeiras como micélio dormente ou como clamidósporos no solo (Leach *et al.*, 1977; Galiano-Carneiro & Miedaner, 2017).

Este fitopatógeno é o agente causador da doença “*Northern Corn Leaf Blight*” (NCLB), que em português é chamada de helmintosporiose. Esta doença foi primeiramente identificada na Itália em 1876 (Mitra, M., 1923), sendo que a propagação da doença nas lavouras ocorre principalmente devido a dispersão dos conídios através do ar, da chuva ou pelo vento (De Rossi & Reis, 2014). Temperaturas entre 15 e 25°C, períodos de orvalho de pelo menos 4 horas e umidade relativa entre 90% - 100% acabam por favorecer essas infecções (Levy & Cohen, 1983; Bentolila *et al.*, 1991; Ogliari *et al.*, 2005; Dong *et al.*, 2008).

Após a infecção, o micélio fúngico penetra a cutícula e a epiderme das folhas (Setyawan *et al.*, 2016). As hifas crescem intracelularmente no mesófilo, e se alastram para os feixes vasculares, penetrando no xilema da planta (Jennings & Ullstrup, 1957). Uma vez em contato com o xilema, esse fungo secreta uma toxina denominada de HT (proveniente do organismo *H. turcicum*). A toxina HT é composta por compostos solúveis em água que contém baixo peso molecular, de modo que inibem a síntese de clorofila e são, portanto, fitotóxicos (Bashan *et al.*, 1995; Li *et al.*, 2016).

Quanto a sua classificação, *S. turcica* quase sempre é dividido de acordo com sua virulência para um determinado hospedeiro ou grupo de hospedeiros, muitas delas conseguindo infectar apenas alguns hospedeiros específicos (Tang *et al.*, 2015). Além disso, essa espécie possui nomenclaturas distintas para sua forma sexuada

(*Setosphaeria turcica*) e assexuada (*Exserohilum turcicum*); no entanto, muitos estudos limitam a pesquisa à interação patógeno e hospedeiro, sem levar em conta as análises moleculares disponíveis para abranger a diversidade genética dos fungos (Caligiome *et al.* 1999).

Tradicionalmente, este gênero é composto apenas por fitopatógenos, como mencionado previamente, embora de acordo com Amaya *et al.* (2020), existam registros recentes de diferentes espécies de *Exserohilum* causando doenças cutâneas, infecções sistêmicas e locais tanto em pacientes imunossuprimidos quanto em aqueles imunocompetentes.

A reprodução do Filo ao qual pertencem esses organismos pode ser de forma sexuada ou assexuada. A reprodução assexuada dos fungos ascomicetos acontece através da formação de esporos formados por mitose. Já a reprodução sexuada deste grupo pode ser bastante variável e pode acontecer pelos processos de contato ou copulação de gametângios, por somatogamia e/ou espermatização. Essa forma de reprodução delimita o grupo de forma muito precisa, e confere ao grupo sua característica mais marcante, a formação dos ascos (Santos; Junior, 2015)

Os ascos são estruturas em formato de saco formadas de hifas entrelaçadas e bastante compactadas, denominadas de ascoma. É dentro dessa estrutura altamente especializada e com diferentes classificações que acontecem os processos de cariogamia, meiose e maturação interna dos esporos. (Maia; Carvalho, 2010).

Em estudo utilizando diversos marcadores moleculares do tipo RAPD (*Random Amplified Polymorphic DNA*) em mais de 250 isolados de *S. turcica* de diferentes regiões do planeta, Borchardt *et al.* (1998) verificaram que em países tropicais como México, Quênia e a porção mais ao sul da China, essas populações possuem alta diversidade genética indicando quase que exclusivamente a reprodução sexuada destes fungos, enquanto que em regiões mais temperadas, como Europa e a porção norte da China, apresentam uma diversidade genética muito menor, indicando que a recombinação genética seja menor nessas temperaturas.

As análises pangenômicas, unindo o conhecimento das regiões conservadas ao restante do genoma, muitas vezes dito como “acessório” foi algo que originalmente foi proposto em modelos de procariotos como uma ferramenta para observar genes e suas variabilidades intraespecíficas (McCarthy & Fitzpatrick, 2019). No entanto, esse

modelo de análise e *pipeline* de processamento de dados já foi comprovadamente validado para diferentes organismos eucariotos (Parfrey *et al.*, 2008).

2.2 Material e Métodos

Neste trabalho foi utilizada a base de dados NCBI (*National Center of Biotechnology Information*), na categoria NCBI Genome, onde foi realizada a pesquisa utilizando os termos “*Setosphaeria turcica*” (forma sexuada) e “*Exserohilum turcicum*” (forma assexuada). Foram encontradas sequências de apenas duas linhagens (*E. turcica* Et28A, de referência, e *E. turcica* f. sp. sorghi GD003) e, por este motivo, ampliamos a busca para o gênero *Exserohilum* inteiro.

Os genomas encontrados foram anotados com o auxílio do programa GenSas (V. 6.0), uma plataforma que fornece um *pipeline* para anotação estrutural e funcional do genoma inteiro para eucariotos e procariotas. Os usuários podem fazer *upload* de sequências de genoma e selecionar uma variedade de ferramentas para mascaramento de repetição, previsão de modelos de genes e outros recursos estruturais, bem como ferramentas de anotação funcional (Humann *et al.*, 2019).

Dentro do *pipeline*, o programa PRINSEQ foi utilizado já no momento do *input* dos dados em formato FASTA. Este programa pode ser utilizado para filtrar, reformatar ou ajustar dados de sequenciamento genômico ou metagenômico (Schmieder & Edwards, 2011). Foram encontradas 5 diferentes sequências, sendo duas da espécie *S. turcica* e três da espécie *S. rostratum*, onde duas das sequências de *S. rostratum* foram identificadas como sendo de baixa qualidade por este programa foram descartadas da análise.

Nas etapas seguintes foram utilizados quatro programas para predição de genes: AUGUSTUS para predição de genes em sequências genômicas eucarióticas (Stanke & Morgenstern, 2005; Stanke *et al.*, 2004); EvidenceModeler (também conhecido como EVM), que combina previsões de genes *ab initio*, ou seja, que utiliza apenas aos arquivos das sequências de nucleotídeos como dados de entrada, e alinhamentos de proteínas e transcritos em estruturas de genes de consenso ponderadas (Haas *et al.*, 2008); GeneMarkES, que usa treinamento não supervisionado de novos genomas para realizar a previsão de genes *ab initio* (Lomsadze *et al.*, 2005; Ter-Hovhannisyan *et al.*, 2008); o *software* GlimmerM, um

localizador de genes derivado do Glimmer, desenvolvido especificamente para eucariontes (Salzberg *et al*, 1999). Ao fim da execução de todos esses programas, foi implementada uma análise utilizando o software BUSCO (V. 5) como controle de qualidade para a predição de genes (Simão *et al*, 2015), onde as predições com maior *score* eram selecionadas como conjuntos de genes oficiais daquela cepa. O programa BUSCO fornece medidas quantitativas para a avaliação da montagem do genoma, conjunto de genes e completude do transcriptoma, com base nas expectativas evolutivas da quantidade de genes ortólogos de cópia única quase universais selecionados do OrthoDB (V. 10) (Simão *et al*, 2015).

Enfim, as sequências passaram pela fase de anotação funcional, utilizando o banco de dados NCBI RefSeq (O'Leary *et al*, 2016) e os programas: BLAST (*Basic Local Alignment Tool*), que localiza regiões de similaridade entre sequências dentro de um banco de dados entre nucleotídeos (BLASTn) ou de proteínas (BLASTp); DIAMOND, que realiza a mesma comparação que o BLASTp, mas com uma velocidade otimizada (Camacho *et al*, 2009); os programas InterProScan e Pfam, que realizam comparação entre sequências, sendo que o Pfam utiliza alinhamentos de sequências e modelos ocultos de Markov (HMM's) (Jones *et al*, 2015; Finn *et al*, 2016). Em seguida, mais uma etapa de controle de qualidade foi adicionada utilizando o software BUSCO (V. 5) para avaliar a qualidade das anotações. As sequências não tiveram quaisquer alterações manuais realizadas, e apenas as anotações funcionais realizadas pelo programa BLAST foram utilizadas para as demais etapas.

Na sequência, foi utilizado o programa OrthoVenn2 (Xu, 2019), que realiza a análise de clusters ortólogos de genomas completos e realiza análises comparativas entre eles, com o objetivo de avaliar o pangenoma de *S. turcica*. Por fim, para a avaliação da distância filogenética entre as cepas, as sequências originais dos genomas foram submetidas ao software fastANI (JAIN *et al*, 2018), que compara os genomas par a par para calcular a identidade nucleotídica média e construir uma matriz de similaridade, na qual o valor 1 significa 100% de identidade e o valor 0 significa 0% de identidade. O pipeline de análises simplificado pode ser consultado na Figura 1.

Figura 1. Pipeline de Análise Pangenômica do gênero *Sethoespheria*.

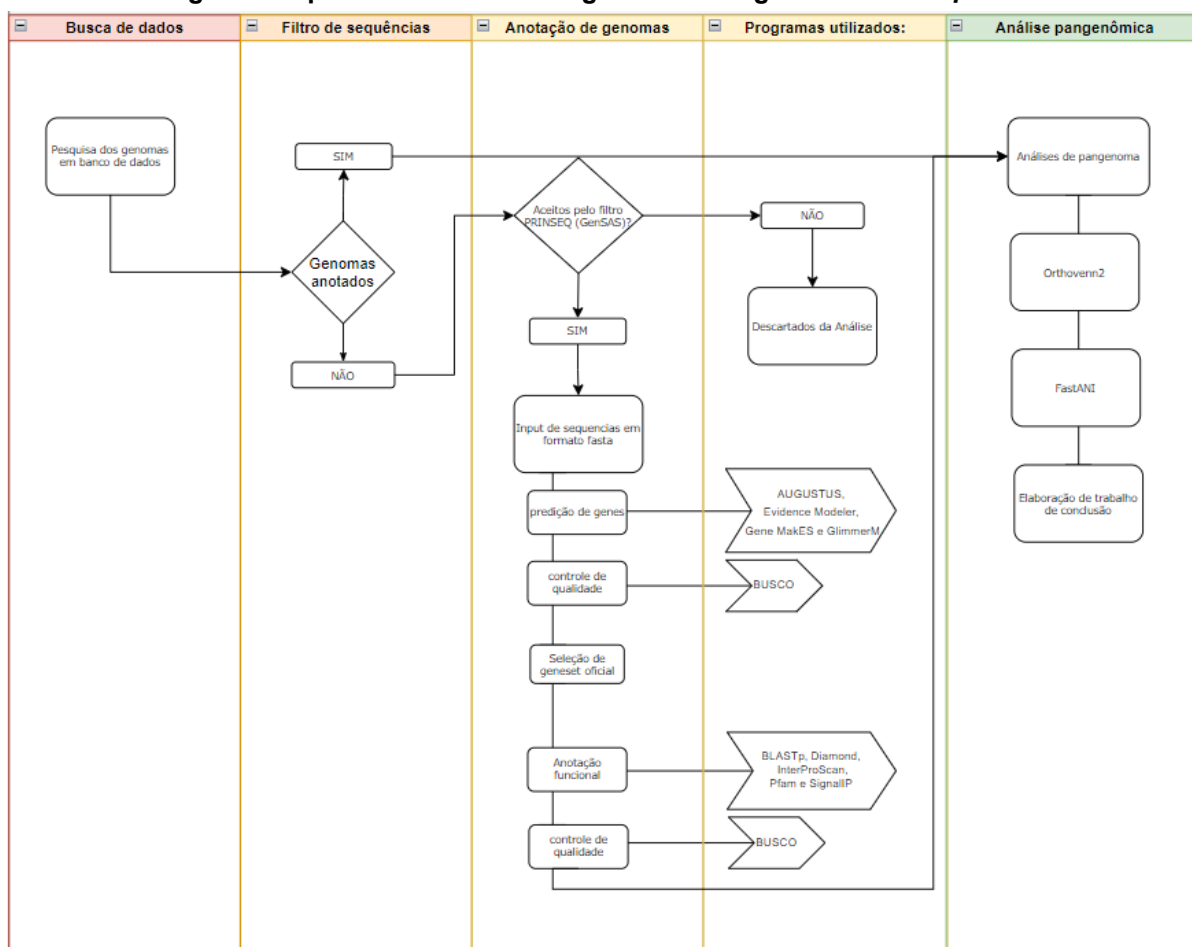


Figura gerada na ferramenta diagrams.net. Fonte: A AUTORA, 2023.

2.3 Resultados

Em uma busca utilizando o banco de dados Genome, do *GenBank*, foram encontradas duas montagens de genoma completo da linhagem Et28A de *Exserohilum turcica*, sendo que apenas uma delas está finalizada (BioProject: PRJNA638013). As demais cepas encontradas e consideradas para a análise pangenoma estão descritas na Tabela 1:

Tabela 1. Sequências de cepas avaliadas para a análise pangenômica de *Sethosphaeria*.

espécie	cepa	n° de contigs	software GenSAS
<i>S. rostratum</i>	ER1	29	não aceito
<i>S. rostratum</i>	SR-KPL1	769	aceito
<i>S. rostratum</i>	ZM170581	667	não aceito
<i>S. turcica</i>	Et28A	407	aceito
<i>S. turcica</i>	GD003	22	aceito

Fonte: A AUTORA, 2023.

Após avaliação para garantir a não ocorrência de repetições das sequências, iniciou-se o processo o *pipeline* descrito na Figura 1, dentro do servidor GenSAS. Contudo, apenas as cepas *S. rostratum* SR_KPL1, *S. turcica* Et28A e *S. turcica sorghi* GD003 tiveram a qualidade de sequenciamento suficiente para dar continuidade ao processamento de dados. De acordo com o manual do usuário do software PRINSEQ, os arquivos de montagem devem conter menos de 25.000 sequências totais e mais de 50% das sequências devem ter mais de 2.500 bases de comprimento (Schmieder & Edwards, 2011). Os resultados para a predição de genes encontrada por cada um dos programas utilizados e a qualidade final da anotação funcional podem ser consultados na Tabela 2.

Tabela 2. Resultados das predições de genes obtidas pelos diferentes softwares

Programa	Resultado	Média	Cepas		
			<i>S. turcica</i>		
			Et28A	SR-KPL1	GD003
AUGUSTUS	predição de genes	12.059	12.262	11.504	12.411
GeneMarkES	predição de genes	13.067	13.610	11.847	13.743
GlimmerM	predição de genes	7.782	7.549	8.447	7.349
EvidenceModeler	sequência consenso	11.735	11.739	11.576	11.891
BUSCO	qualidade da anotação funcional	-	99%	99,10%	97,40%

Fonte: A AUTORA, 2023..

Para a análise de valor de identidade nucleotídica média entre as diferentes espécies, as cepas *S. rostratum* SR_KPL1 e *S. turcica* Et28A tiveram identidade igual à 84,3, 99,5 para *S.turcica* Et28A e *S. turcica sorghi* GD003 e, por fim, as cepas *S.*

rostratum SR_KPL *S. turcica sorghi* GD003 apresentaram uma identidade nucleotídica média de 84,3. A Figura 2 tem um diagrama de Venn representando numericamente as proteínas comuns e individuais de cada cepa utilizada.

Figura 2. Diagrama de Venn comparando o número de proteínas entre diferentes espécies do gênero *Sethosphaeria*.

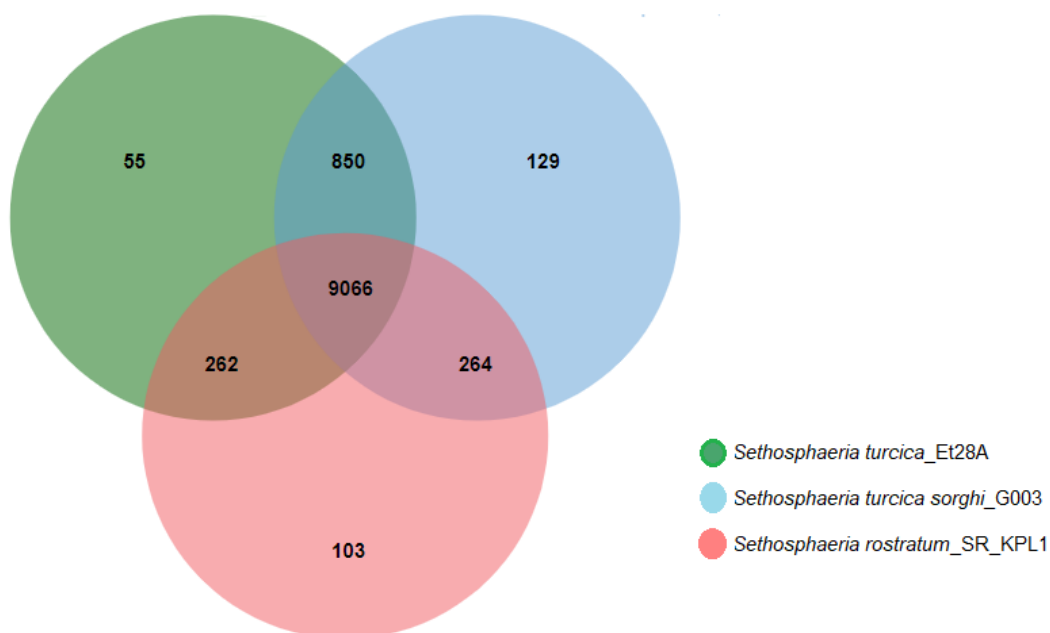
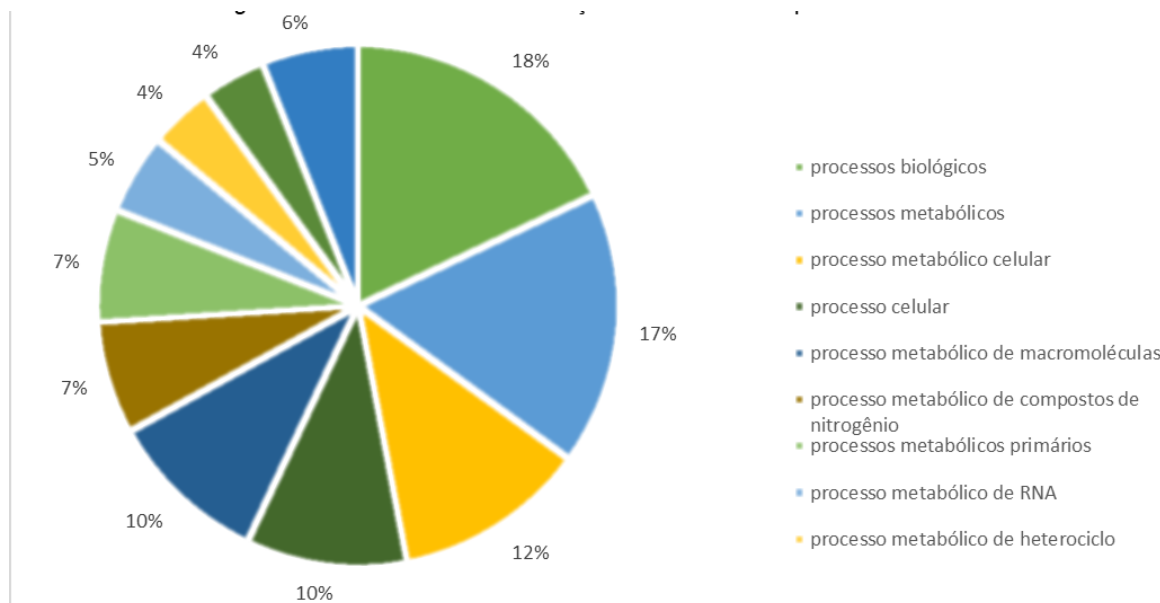


Imagem gerada no programa OrthoVenn2. Fonte: A AUTORA, 2023.

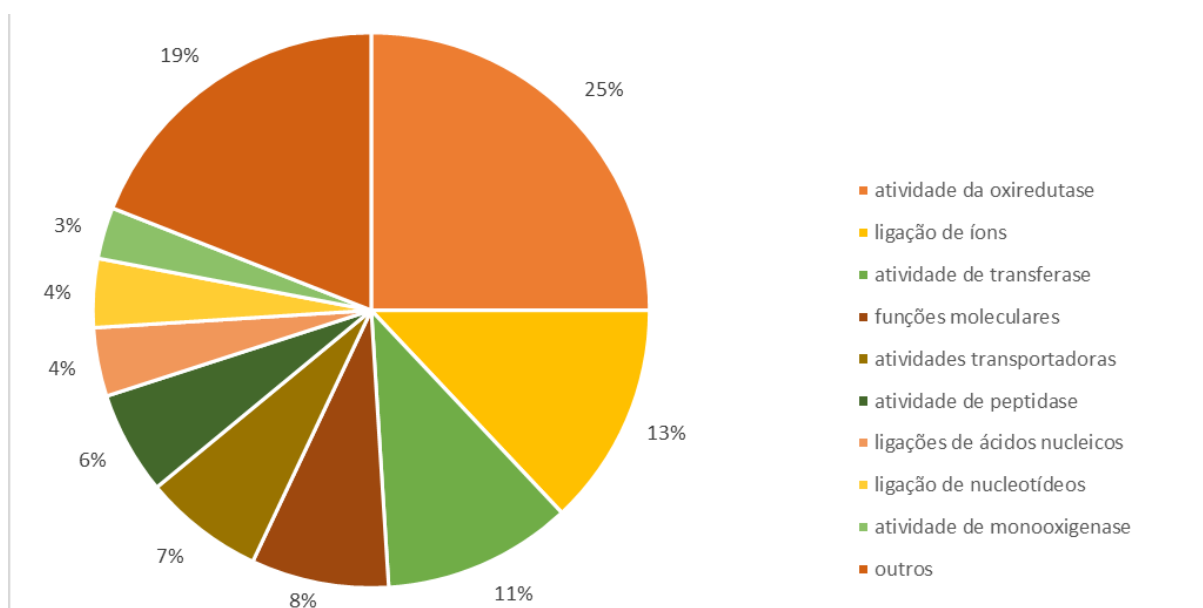
Nas Figuras 3 a 5 são demonstradas as diferentes identificações de funcionalidades para o genoma central, ou seja, contendo as proteínas comuns a todas as cepas analisadas, definidos de acordo com a classificação do OrthoVenn2.

Figura 3. Porcentagem de proteínas do genoma central envolvidas em processos biológicos de acordo com a classificação do GO realizada pelo OrthoVenn2.



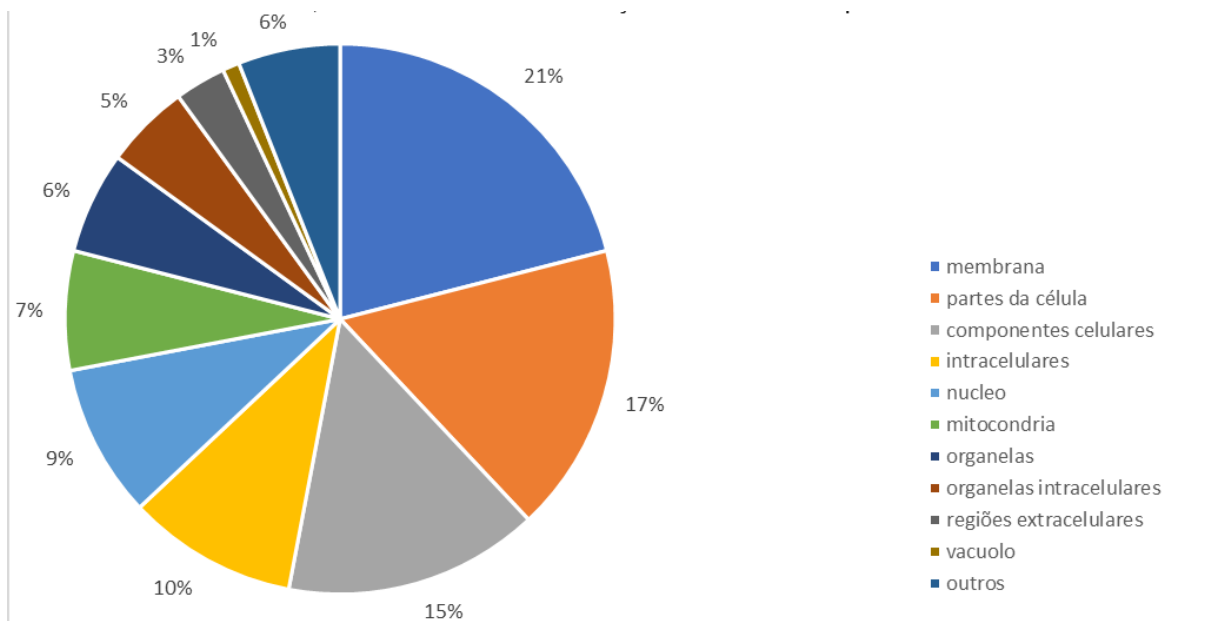
Fonte: A AUTORA, 2023.

Figura 4. Porcentagem de proteínas do genoma central associadas à funções moleculares, de acordo com a classificação de GO, realizada pelo OrthoVenn2.



Fonte: A AUTORA, 2023.

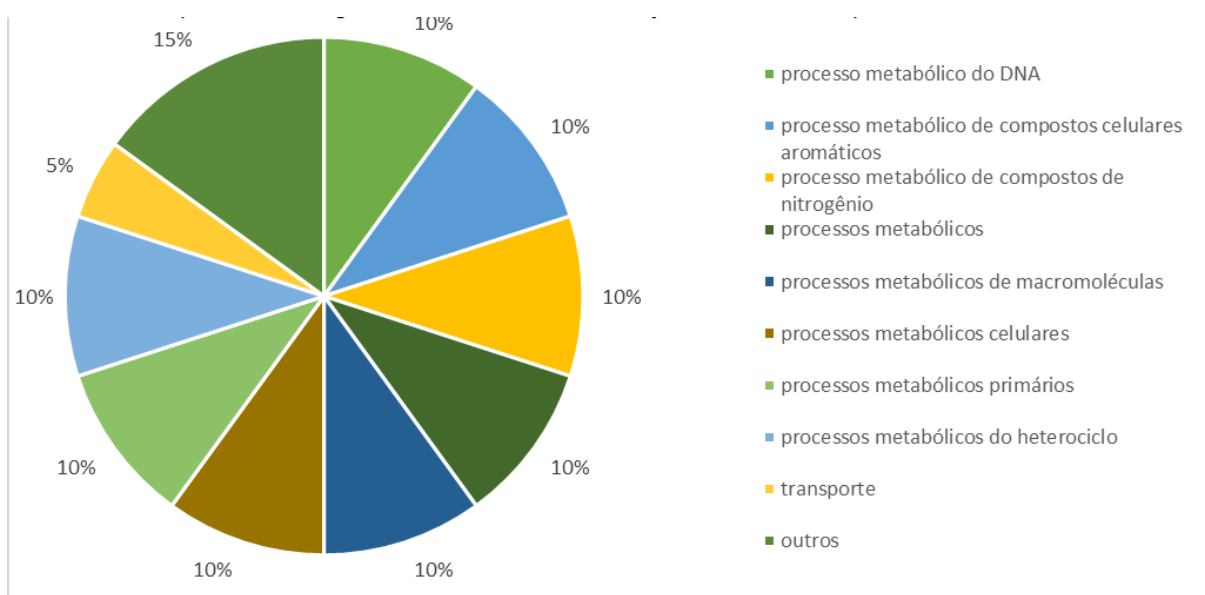
Figura 5. Porcentagem de proteínas do genoma central associadas a componentes celulares, de acordo com a classificação de GO realizada pelo OrthoVenn2.



Fonte: A AUTORA, 2023.

Nas Figuras 6 a 10 estão apresentados os resultados para as classificações quanto às proteínas pertencentes ao genoma periférico, ou seja, individuais à cada cepa analisada. Quanto à classificação de proteínas associadas a componentes celulares, a cepa *S. turcica sorghi* (GD003) a identificação realizada pelo OrthoVenn2 apenas apresentou a classificação de região extracelular, e não obteve resultados para a classificação de proteínas associadas à funções moleculares.

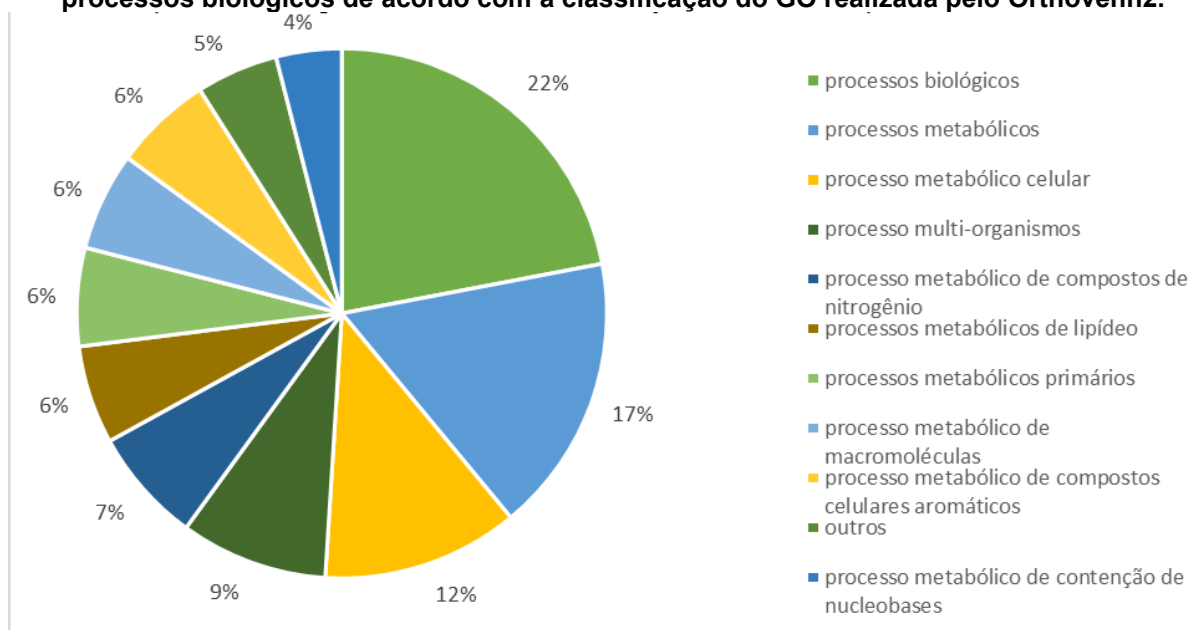
Figura 6. Porcentagem de proteínas do genoma de *S. turcica sorghi* (GD003) envolvidas nos processos biológicos de acordo com a classificação de GO realizada pelo OrthoVenn2.



Fonte: A AUTORA, 2023.

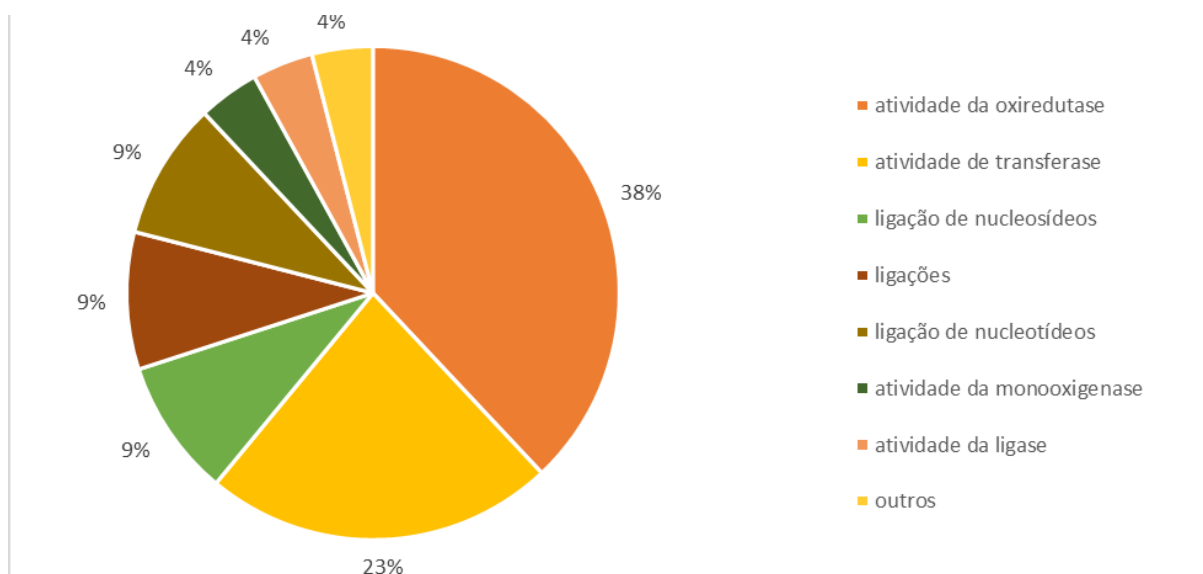
As Figuras 7 a 9 apresentam os dados de classificação do GO realizada pelo OrthoVenn2 para as proteínas associadas a determinados processos biológicos, funções moleculares e componentes celulares do genoma periférico correspondente à cepa de *S. rostratum* (SR_KPL1).

Figura 7. Porcentagem de proteínas do genoma de *S. rostratum* (SR_KPL1) envolvidas em processos biológicos de acordo com a classificação do GO realizada pelo OrthoVenn2.



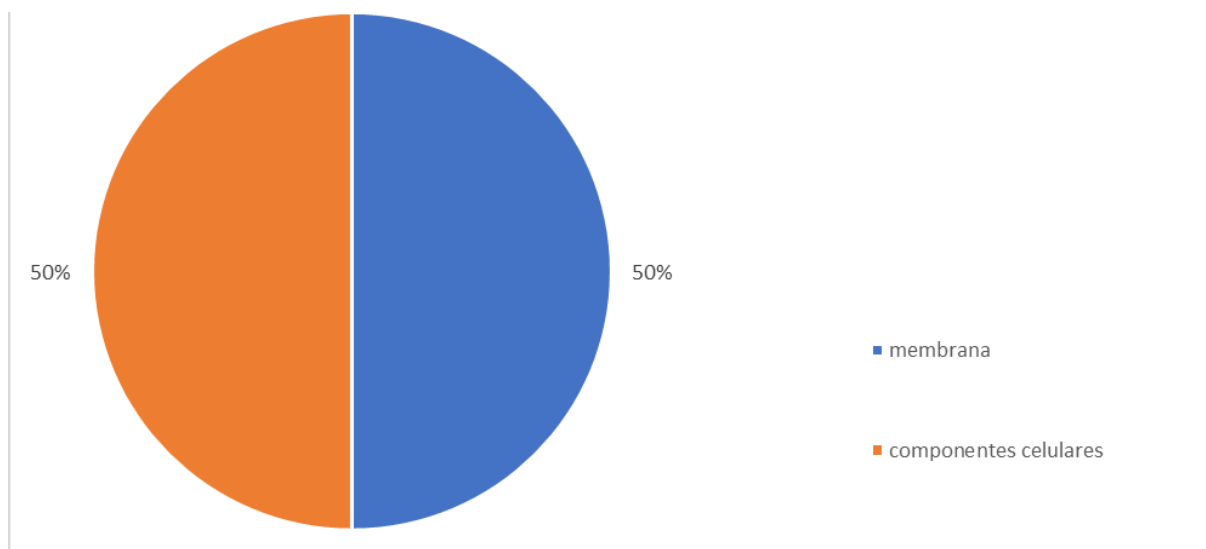
Fonte: A AUTORA, 2023.

Figura 8. Porcentagem de proteínas do genoma de *S. rostratum* (SR_KPL1) associadas a funções moleculares de acordo com a classificação do GO realizada pelo OrthoVenn2.



Fonte: A AUTORA, 2023.

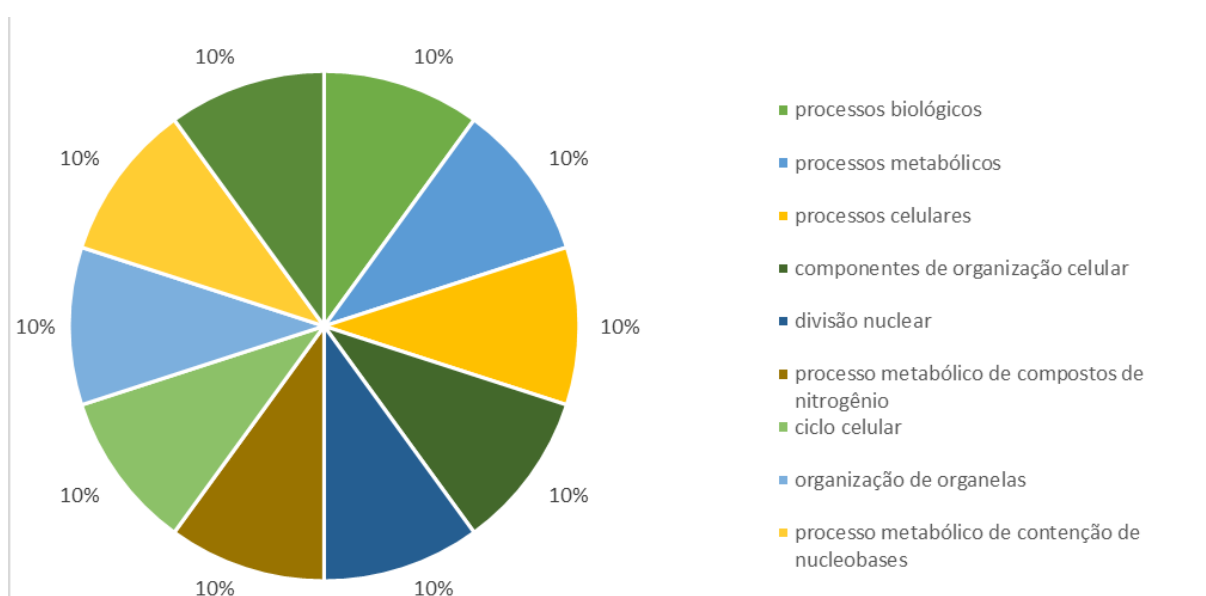
Figura 9. Porcentagem de proteínas do genoma de *S. rostratum* (SR_KPL1) associadas a componentes celulares, de acordo com a classificação do GO realizada pelo OrthoVenn2.



Fonte: A AUTORA, 2023.

Na Figura 10, estão apresentados os resultados para as classificações quanto às proteínas pertencentes ao genoma periférico da cepa de *S. turcica* (Et28A). Quanto à classificação de proteínas associadas a componentes celulares, a cepa Et28A não apresentou resultados, sendo a única associação para função molecular dentre as classificações realizadas pelo OrthoVenn2, foi a de atividade da hidrolase.

Figura 10. Porcentagem de proteínas do genoma de *S. turcica* (Et28A) envolvidas nos processos biológicos de acordo com a classificação de GO realizada pelo OrthoVenn2.



Fonte: A AUTORA, 2023.

No genoma central analisado, foram encontradas 80 proteínas associadas com patogênese de acordo com a classificação do GO realizada pelo OrthoVenn2, enquanto que no genoma periférico, apenas a cepa de *S. rostratum* analisada teve ocorrências de proteínas associadas à patogenicidade de acordo com a mesma classificação. As sete ocorrências encontradas foram investigadas quanto a sua origem e os resultados estão apresentados na Tabela 3.

Tabela 3. Relação de proteínas de *S.rostratum* relacionadas à patogênese com suas correspondentes de maior score de acordo com a maior identidade e similaridade identificados pelo BLASTp

cluster	código da proteína	nome	query cover	identidade	código NCBI
cluster 232	00g075360	relacionado com a monooxigenase do citocromo P450 (lovA) [<i>Ramularia collo-cygni</i>]	100%	87,99%	XP_023628074.1
cluster 232	00g009610	citocromo P450 [<i>Phaeosphaeriaceae</i> sp. PMI808]	100%	87,71%	KAH8703001.1
cluster 232	00g096090	proteína hipotética AJ80_05567 [<i>Polytolypa hytricus</i> UAMH7299]	98%	64,16%	PGH15383.1
cluster 233	00g009510	tioredoxina redutase [<i>Phaeosphaeriaceae</i> sp. PMI808]	99%	80,83%	KAH8723049.1
cluster 233	00g075450	relacionado com tioredoxina redutase [<i>Ramularia collo-cygni</i>]	96%	81,96%	XP_023628083.1
cluster 233	00g096100	tioredoxina redutase glit [<i>Phaeosphaeriaceae</i> sp. PMI808]	100%	85,48%	KAH8723042.1
cluster 9231	00g108620	EntF, proteína do módulo de peptídeo sintetase não ribossomal [<i>Pyrenophora tritici-repentis</i>]	99%	62,77%	KAF7572929.1
cluster 9231	00g104680	HC-toxina sintetase [<i>Pyrenophora tritici-repentis</i> Pt-1C-BFP]	100%	64,26%	XP_001940764.1
cluster 9248	00g047670	proteína hipotética GQ44DRAFT_693644 [<i>Phaeosphaeriaceae</i> sp. PMI808]	99%	90,57%	KAH8701275.1
cluster 9248	00g003400	proteína hipotética HBI10_107890 [<i>Parastagonospora nodorum</i>]	100%	74,88%	KAH4000130.1
cluster 9249	00g075370	proteína não caracterizada C7974DRAFT_292919, parcial [<i>Boeremia exigua</i>]	91%	79,75%	XP_045992328.1
cluster 9249	00g009530	proteína hipotética GQ44DRAFT_831630 [<i>Phaeosphaeriaceae</i> sp. PMI808]	86%	81,05%	KAH8703008.1
cluster 9269	00g112320	EntF, proteína do módulo de peptídeo sintetase não ribossomal [<i>Pyrenophora tritici-repentis</i>]	82%	76,92%	KAF7572928.1
cluster 9269	00g102630	EntF, proteína do módulo de peptídeo sintetase não ribossomal [<i>Pyrenophora tritici-repentis</i>]	100%	72,91%	KAF7572928.1
cluster 9288	00g022800	proteína não caracterizada ALTATR162_LOCUS9826 [<i>Alternaria atra</i>]	98%	57,25%	XP_043173397.1
cluster 9288	00g004710	proteína hipotética J4E81_010817 [<i>Alternaria</i> sp. BMP 2799]	100%	75,69%	KAI4677792.1

Fonte: A AUTORA, 2023.

2.4 Discussão

O organismo *Sethosphaeria turcica* pertencente à família *Ascomycota* é principalmente conhecido por ser o agente causador da helmintosporiose (Zhang *et al.*, 2012). Periodicamente, essa espécie causa prejuízos na cadeia produtora de milho, relacionadas a epidemias dentro de plantações localizadas em regiões com altas temperaturas e umidade elevada (Raymundo & Hooker, 1981., Perkins & Pedersen, 1987; Carson, 1999), sendo desta forma, uma espécie com que desperta o interesse quanto ao seu entendimento.

Ainda que *S. turcica* tenha sua atenção justificada, ao observar os dados apresentados na Tabela 1, em pesquisa no NCBI genome foram encontradas apenas seis diferentes sequências do organismo *Sethosphaeria*, e dentre elas, apenas três delas foram aceitas pelo *pipeline* de anotação funcional. Ou seja, além dos poucos dados disponíveis, 50% deles não tinham os requisitos adequados para participar do *pipeline* de análises.

Dentre os dados analisados na Tabela 2, podemos observar os resultados para a predição de genes para cada cepa de acordo com os diferentes programas utilizados. Os três programas apresentaram dados concordantes entre si, e até mesmo dados aproximados entre as diferentes cepas dos mesmos organismos. Naturalmente, a cepa da espécie *S. rostratum* apresentou uma divergência das outras duas de *S. turcica*, se tratando de diferentes espécies. Na mesma tabela, é possível observar também que o programa de GlimmerM apresentou uma contagem de genes muito abaixo dos demais, isso pode ser explicado por conta dos modelos de treinamento que o programa recebeu, contendo apenas os organismos *Arabidopsis thaliana*, *Oryza sativa*, e *Plasmodium falciparum* (Salzberg *et al.*, 1999). O treinamento que o programa recebe é intrínseco ao seu software e não pode ser modificado, como nenhum destes organismos citados se assemelha especificamente ao objeto de estudo neste trabalho, esses resultados foram descartados nas demais etapas deste projeto.

Nas Figura 2 e Tabela 3, estão apresentados o Diagrama de Venn com o número de clusters comuns e individuais a cada cepa de *Sethosphaeria*, onde o genoma central apresentou 28906 proteínas e 9066 *clusters*. Quanto às classificações realizadas, representadas graficamente nas Figuras 3, 6, 7 e 10 é possível observar

as porcentagens dos *clusters* envolvidos nos processos biológicos do genoma central e individual de cada cepa. Destes resultados, podemos destacar a Figura 7, que apresenta 6% de clusters identificados para os processos metabólicos de lipídio e 17% para o processamento de macromoléculas.

Uma característica comum a algumas espécies de Ascomicetos fitopatogênicos é a dependência da digestão de corpos lipídicos armazenados no citoplasma para conseguir realizar a penetração no hospedeiro (Mims et al., 1995; Schadeck; Leite; Bucchi, 1998). A digestão dessas moléculas supre a demanda de substratos para os processos de síntese e auxilia ao fornecer a energia necessária para a germinação, além de produzir glicerol e melanina, necessários para obtenção da pressão de turgor, essencial à infecção (Weete, 1984; Thimes; Weber; Talbot, 2000).

Ao investigar a patogenicidade dessas espécies, foi possível identificar 80 proteínas associadas à patogênese dentro do genoma central, no entanto, ao consultar o código de classificação dentro do GO, o sistema informava que essa se tratava de um código de classificação obsoleto que atualmente contempla duas subdivisões, podendo significar: 1. Modulação processos de um hospedeiro através de um simbiote; 2. Indução de morte celular programada no hospedeiro realizada por um simbiote.

Nas análises para investigar de forma mais profunda a patogenicidade individual de cada linhagens, apenas a cepa proveniente da espécie de *S. rostratum* teve clusters associadas à patogênese, de acordo com a classificação do GO realizada pelo programa OrthoVenn2. Isso representa um achado interessante, pois muitos dos genes de patogenicidade são comuns ao gênero, no entanto, somente a linhagem Et28A tem diferentes genes associados diante dessa classificação. Para o melhor entendimento deste repertório genético, seria necessária a comparação apenas entre diferentes linhagens contendo amostras apenas da espécie de *S. rostratum*.

Com 16 proteínas em 7 diferentes clusters, os resultados organizados na Tabela 5, demonstram uma frequência de quatro principais proteínas, das quais: citocromo p450 e relacionadas, proteínas relacionadas a tioredoxina redutase, proteína do módulo de peptídeo sintetase de outros ascomicetos como *Ramularia collo-cygni*, *Phaeosphaeriaceae sp.* e *Pyrenophora tritici-repentis*.

Genes codificadores para essas proteínas estão fortemente atrelados à produção de micotoxinas denominadas fumosinas, usualmente encontradas no milho e seus derivados, associadas a diversas doenças em animais (Lorenzzeti, Villas-Boas, Garcia, 2006). Comprovadamente a ingestão de grãos contaminados com fumonisinas pode trazer sintomas alérgicos ou ainda, causar carcinogênese de esôfago em humanos se ingeridos por longos períodos de tempo (Norred, 1993; Pitt, 2000; Marasas, 2001). Ainda que muitas dessas proteínas tenham apresentado um baixo percentual de identidade, isso pode fornecer informações quanto a outros organismos com mecanismos de virulência similares e também como contê-los.

3 CONCLUSÃO

O presente trabalho se propôs a analisar o pangenoma de *Setosphaeria turcica*, visando descrever genes de patogenicidade e fatores de virulência envolvidos na infecção provocada pelo agente causador da helmintosporiose. Mesmo causando grande impacto nos plantios, os dados disponíveis em relação a esta espécie são insuficientes para estudos genômicos mais aprofundados, capazes de desvendar os mecanismos de infecção deste organismo. Na pesquisa realizada, apenas os genomas de seis cepas de duas espécies distintas de *Setosphaeria* foram encontradas, dos quais apenas três possuíam qualidade suficiente e puderam ser utilizados para o desenvolvimento do trabalho.

Vale ressaltar que a maior parte do pangenoma do gênero *Setosphaeria* é formado pelo o genoma central, mesmo realizando a comparação entre diferentes espécies, o que evidencia uma baixa variabilidade intra-cepa e intra-específica. Quanto aos genes de patogenicidade, é importante destacar a presença de genes de função ainda desconhecida exclusivos à cepa de *S. rostratum* (Et28A) analisada, que podem ser de importância para a infecção e o desenvolvimento da doença.

É necessário ampliar o sequenciamento de linhagens distintas que sejam provenientes de diferentes hospedeiros, pois além de possibilitar a compreensão das variações genéticas que as tornam capazes de infectar diferentes cultivares, nos ajudariam a ter um maior vislumbre dos mecanismos de infecção desenvolvidos para a recém adquirida capacidade de infectar mamíferos.

REFERÊNCIAS

- Andrews, S. (2010). FastQC: a quality control tool for high throughput sequence data. Available online at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>.
- Amaya, Paula Pena et al. Feohifomicosis por *Exserohilum rostratum* en un paciente pediátrico con leucemia linfoblástica aguda postransplante de médula ósea. *Rev. argent. microbiol.* [online]. 2020, vol.52, n.3, pp.91-100. ISSN 0325-7541.
- Bashan, B., Levy, R. S., Cojocar, M., and Levy, Y. (1995). Purification and structural determination of a phytotoxic substance from *Exserohilum turcicum*. *Physiol. Mol. Plant Pathol.* 47, 225–235. doi: 10.1006/pmpp.1995.1054.
- Bashan, B., and Levy, Y. (1992). Differential response of sweet corn cultivars to phytotoxic water soluble compounds from culture filtrates of *Exserohilum turcicum*. *Plant Dis.* 76, 451–454. doi: 10.1094/PD-76-0451.
- Bentolila, S., Guitton, C., Bouvet, N., Sailland, A., Nykaza, S., and Freyssinet, G. (1991). Identification of an RFLP marker tightly linked to the Ht1 gene in maize. *Theor. Appl. Genet.* 82, 393–398. doi: 10.1007/BF00588588.
- Borchardt DS, Welz HG, Geiger HH. Genetic Structure of *Setosphaeria turcica* Populations in Tropical and Temperate Climates. *Phytopathology.* 1998 Apr;88(4):322-9. doi: 10.1094/PHYTO.1998.88.4.322. PMID: 18944955.
- Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., Madden, T.L. (2009) BLAST plus: architecture and applications. *BMC Bioinformatics* 10. doi:10.1186/1471-2105-10-421.
- Caligiome, R. B., Resende, M. A., Diasneto, E., Oliveira, S. C., and Azevedo, V. 1999. Dematiaceous fungal pathogens: analysis of ribosomal DNA gene polymorphism by polymerase chain reaction - restriction fragment length polymorphism. *Mycoses* 42:609-614.
- Carson, M. L.; Compendium of corn diseases D.G. White, St. Paul (Eds.), *Helminthosporium leaf spots and blights*, The American Phytopathology Society Press, Minnesota (1999), pp. 15-24.
- Chen X, Zhang Y, Zhang Z, et al. PGAweb: A Web Server for Bacterial Pan-Genome Analysis. *Front Microbiol.* 2018;9:1910. Published 2018 Aug 21. doi:10.3389/fmicb.2018.01910.
- Dai Y, Gan L, Lan C, Lu X, Yang X, Gao Z. Genetic Differentiation and Mixed Reproductive Strategies in the Northern Corn Leaf Blight Pathogen *Setosphaeria turcica* From Sweet Corn in Fujian Province, China. *Front Microbiol.* 2021 May 26;12:632575. doi: 10.3389/fmicb.2021.632575. PMID: 34122358; PMCID: PMC8187859.
- De Rossi, R. L., and Reis, E. M. (2014). Semi-selective culture medium for *Exserohilum turcicum* isolation from corn seeds. *Summa Phytopathol.* 40, 163–167. doi: 10.1590/0100-5405/1925.

De Rossi, R. L.; Guerra, F.A.; Plazas, M. C. et al. Crop damage, economic losses, and the economic damage threshold for northern corn leaf blight. *Crop Protection*. Volume 154, 2022. 105901, ISSN 0261-2194. <https://doi.org/10.1016/j.cropro.2021.105901>.

Dong, J.; Fan, Y.; Gui, X.. Geographic Distribution and Genetic Analysis of Physiological Races Of *Setosphaeria turcica* in Northern China. **Am J Agric Biol Sci**, v. 3, n. 1, p. 389-398, 2008.

Finn, R.D., Coggill, P., Eberhardt, R.Y., Eddy, S.R., Mistry, J., Mitchell, A.L., Potter, S.C., Punta, M., Qureshi, M., Sangrador-Vegas, A., Salazar, G.A., Tate, J., Bateman, A. (2016). The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res* 44 (D1):D279-D285. doi:10.1093/nar/gkv1344.

Galiano-Carneiro AL and Miedaner T (2017) Genetics of Resistance and Pathogenicity in the Maize/*Setosphaeria turcica* Pathosystem and Implications for Breeding. *Front. Plant Sci.* 8:1490. doi: 10.3389/fpls.2017.01490.

Haas, B.J., Salzberg, S.L., Zhu, W., Pertea, M., Allen, J.E., Orvis, J., White, O., Buell, C.R., Wortman, J.R. (2008). Automated eukaryotic gene structure annotation using EVIDENCEModeler and the program to assemble spliced alignments. *Genome Biol* 9 (1). doi:10.1186/Gb-2008-9-1-R7.

Hammerschmidt, R., Metraux, J. P., and Van Loon, L. C. 2001. Inducing resistance: A summary of papers presented at the first international symposium on induced resistance to plant diseases, Corfu, May 2000. *Eur. J. Plant Pathol.* 107:1-6.

Humann, J. L., Lee, T., Ficklin, S., Main, D. "Structural and Functional Annotation of Eukaryotic Genomes with GenSAS." *Gene Prediction: Methods and Protocols*. Ed. Martin Kollmar. New York: Humana Press, 2019. 29-51. DOI: 10.1007/978-1-4939-9173-0_3.

Jones, P., Binns, D., Chang, H.Y., Fraser, M., Li, W.Z., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G., Pesseat, S., Quinn, A.F., Sangrador-Vegas, A., Scheremetjew, M., Yong, S.Y., Lopez, R., Hunter, S. (2014). InterProScan 5: genome-scale protein function classification. *Bioinformatics* 30 (9):1236-1240. doi:10.1093/bioinformatics/btu031.

Jennings, P. R., and Ullstrup, A. J. (1957). A histological study of three *Helminthosporium* leaf blight of corn. *Phytopathology* 47, 707–714.

Leach, C. M., Fullerton, R. A., and Young, K. (1977). Northern leaf blight of maize in New Zealand: relationship of *Drechslera turcia* airspora to factors influencing sporulation, conidium development, and chlamydospore formation. *Phytopathology* 67, 629–636. doi: 10.1094/Phyto-67-629

Leggett, R.M.; Ramirez-Gonzalez, R.H.; Clavijo, B.J.; Waite, D.; Davey, R.P. Sequencing quality assessment tools to enable data-driven informatics for high throughput genomics. *Front Genet.* 2013 Dec 17;4:288. doi: 10.3389/fgene.2013.00288. PMID: 24381581; PMCID: PMC3865868.

Levy, Y., and Cohen, Y. (1983). Differential effect of light on spore germination of *Exserohilum turcicum* on corn leaves and corn leaf impressions. *Phytopathology* 73, 249–252. doi: 10.1094/Phyto-73-249.

Li, P., Gong, X., Jia, H., Fan, Y., Zhang, Y., Cao, Z., et al. (2016). MAP kinase gene STK1 is required for hyphal, conidial, and appressorial development, toxin biosynthesis, pathogenicity, and hypertonic stress response in the plant pathogenic fungus *Setosphaeria turcica*. *J. Integr. Agric.* 15, 2786–2794. doi: 10.1016/S2095-3119(16)61472-7.

Lipps, P. and Mills, D. Northern corn leaf blight. Extension FactSheet. Ohio State University Extension. <https://ohioline.osu.edu/factsheet/plpath-cer-10>. Acesso em 20 Abril 2022.

Lomsadze, A., Ter-Hovhannisyan, V., Chernoff, Y.O., Borodovsky, M. (2005) Gene identification in novel eukaryotic genomes by self-training algorithm. *Nucleic Acids Res* 33 (20):6494-6506. doi:10.1093/nar/gki937.

Lorenzetti, R. S.; Vilas-Boas, L. A.; Garcia, J. E. / UNOPAR Cient., Ciênc. Biol. Saúde, Londrina, v. 8, n. 1, p. 63-70, out. 2006.

Marasas, W.F. Discovery and Occurrence of the Fumonisin: A historical perspective. *Environmental Health Perspectives*. Research Triangle Park, v. 109, p. 239-243, 2001.

McCarthy, C., & Fitzpatrick, D. A. (2019). Pan-genome analyses of model fungal species. *Microbial genomics*, 5(2), e000243. <https://doi.org/10.1099/mgen.0.000243>.

Mims, C. W. Using electron microscopy to study plant pathogenic fungi. *Mycologia*, v. 83, p. 1-19. 1991.

Mitra M., 1923. Helminthosporium spp. on cereals and sugarcane in India. Part I.(Diseases of Zea mays and Sorghum vulgare caused by species of Helminthosporium.). *Bot. Ser.* 6, pp: 219-242.

Norred, W.P. Fumonisin-mycotoxins produced by *Fusarium moniliforme*. *Journal of Toxicology and Environmental Health*, London, v. 38, suppl. Mar, p. 309- 328, 1993.

Ogliari, J. B., Guimarães, M. A., Geraldi, I. O., and Camargo, L. E. A. (2005). New resistance genes in the *Zea mays* - *Exserohilum turcicum* pathosystem. *Genet. Mol. Biol.* 28, 435–439. doi: 10.1590/S1415-47572005000300017.

O'Leary, N.A., Wright, M.W., Brister, J.R. et al, (2016). Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res* 44 (D1):D733-745. doi:10.1093/nar/gkv1189.

OrthoVenn2: a web server for whole-genome comparison and annotation of orthologous clusters across multiple species. *Nucleic Acids Res.* 2019.

Page, A.; Cummins, C.; Hunt, M.; Wong, V.; Reuter, S.; Holden, M., Roary: rapid large-scale prokaryote pan genome analysis, *Bioinformatics*, Volume 31, Issue 22, 15 November 2015, Pages 3691–3693, <https://doi.org/10.1093/bioinformatics/btv421>.

Parfrey L.W.; Lahr D.J.G.; Katz L.A.; The dynamic nature of eukaryotic genomes. *Mol Biol Evol.* 2008;25:787–794. doi: 10.1093/molbev/msn032.

Perkins, J.M.; Pedersen, W. L. Disease development and yield losses associated with northern leaf blight on corn. *Plant Disease*, 71 (1987), pp. 940-943.

Pitt, J. I. Toxigenic Fungi: Which Are Important? *Medical Mycology*, Oxford, v. 38, suppl. 1, p. 17-22, 2000.

Raymundo, A. D.; Hooker A. L., Measuring the relationship between northern corn leaf blight and yield losses. *Plant Disease*, 65 (4) (1981), pp. 325-327.

Rouli L, Merhej V, Fournier PE, Raoult D. The bacterial pangenome as a new tool for analysing pathogenic bacteria. *New Microbes New Infect.* 2015;7:72–85. doi: 10.1016/j.nmni.2015.06.005.

Salzberg, S.L., Pertea, M., Delcher, A.L., Gardner, M.J., Tettelin, H. (1999) Interpolated Markov models for eukaryotic gene finding. *Genomics* 59 (1):24-31. doi:10.1006/geno.1999.5854.

Santos, E.R.D.; Junior, P.A.H; *Sistemática Vegetal I: Fungos*. 2015.

Schadeck, R. J. G.; Buchi, D. F.; Leite, B. Ultrastructural aspects of *Colletotrichum graminicola* conidium germination, appressorium formation and penetration on cellophane membranes: focus on lipid reserves. *J. Submicrosc. Cytol. Pathol.*, v. 30, p. 555-561, 1998b.

Schmieder R. & Edwards, R.. Quality control and preprocessing of metagenomic datasets. *Bioinformatics* 2011, 27:863-864. [PMID: 21278185].

Setyawan, B., Suliansyah, I., Anwar, A., and Swasti, E. (2016). Resistance of eleven new hybrid maize genotypes to turicum leaf blight (*Exserohilum turicum*). *Biodiversitas* 17, 604–608. doi: 10.13057/biodiv/d170230.

Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., & Zdobnov, E. M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*, 31(19), 3210-3212.

Stanke, M., Morgenstern, B. AUGUSTUS: a web server for gene prediction in eukaryotes that allows user-defined constraints. *Nucleic Acids Res.* 2005; 33:W465-W467. doi:10.1093/Nar/Gki458.

Stanke, M., Steinkamp, R., Waack, S., Morgenstern, B. AUGUSTUS: a web server for gene finding in eukaryotes. 2004. *Nucleic Acids Res* 32:W309-W312. doi:10.1093/Nar/Gkh379.

Swift, T. 2022. Disponível em: <https://www.youtube.com/watch?v=XQEHXuCmi5I>

Tang, L., Gao, Z. G., Yao, Y., and Liu, X. 2015. Identification and genetic diversity of formae speciales of *Setosphaeria turcica* in China. *Plant Dis.* 99:482-487.

Ter-Hovhannisyan, V., Lomsadze, A., Chernoff, Y.O., Borodovsky, M. (2008) Gene prediction in novel fungal genomes using an ab initio algorithm with unsupervised training. *Genome Res* 18 (12):1979-1990. doi:10.1101/gr.081612.108.

Thines, E.; Weber, R. W. S.; Talbot, N. J. MAP kinase and protein kinase A – Dependent mobilization during of triacylglycerol and glycogen during appressorium turgor generation by *Magnapothe grisea* . *Plant Cell*, v. 12, p. 1703-1718, 2000.

Vanderplank, J. E. (1984). *Disease Resistance in Plants*. London: Academic Press, Inc.

Wathaneeyawech, S., Sirithunya, P., and Smitamana, P. (2015b). Study of the host range of northern corn leaf blight disease and effect of *Exserohilum turcicum* toxin on sweet corn. *J. Agric. Technol.* 11, 953–963.

Weete, J. D. Spore lipid surface wax composition; metabolism during germination In: Infection process of fungi. 1984.

Xu, L.; Dong, Z.; Fang, L.; Luo, Y.; Wei, Z.; Guo, H. et al., OrthoVenn2: a web server for whole-genome comparison and annotation of orthologous clusters across multiple species, *Nucleic Acids Research*, Volume 47, Issue W1, 02 July 2019, Pages W52–W58, <https://doi.org/10.1093/nar/gkz333>.

Zhang, S.; Hao, Z.; Wang, L. et al. StRas2 regulates morphogenesis, conidiation and appressorium development in *Setosphaeria turcica*. *Microbiological Research*, Volume 167, Issue 8, 2012. Pages 478-486. ISSN 0944-5013.