Projet de Recherche Fictif

UNIVERSITÉ DE MONTPELLIER

# An epigenetic mechanism for mutation rate modulation in the fungal wheat pathogen *Zymoseptoria tritici*

**MCKEON Conor**

**M2 Translational Medicinal Chemistry**

**Spring, 2022**

## Abstract

DNA double strand breaks (DSBs) are repaired by two distinct pathways: homology-directed repair (HDR) and nonhomologous end joining (NHEJ). Several factors contribute to the choice between these repair pathways, including cell-cycle stage, the complexity of the lesion, and as suggested by recent studies, the chromatin context where a DSB occurs. Notably, these pathways tend to introduce errors with different frequencies, and result in different mutational signatures. Modulation of the chromatin environment might therefore influence the choice of DSB repair pathway, which in turn might influence the type and rate of a mutation at a given locus. The underlying idea of this proposal is that organisms could potentially leverage chromatin properties such as to favor mutations in regions in which a high mutation rate would be advantageous, for example in genes that have important roles in virulence and niche colonization, while minimizing mutation rate at loci in which mutations are most deleterious, such as in essential genes.

The fungal wheat-pathogen *Zymoseptoria tritici* is emerging as a remarkable model organism for investigations on the genetic basis of environmental adaptation due to its remarkable genome plasticity. This plasticity was proposed to drive its rapid adaptation to fungicides and wheat defense mechanisms. In addition to 13 "core" chromosomes found in all isolates, the *Z. tritici* genome has up to 8 additional "accessory" chromosomes, which reportedly have specialized roles in virulence and niche colonization, and have significantly increased mutation rates. These accessory chromosomes are greatly enriched in the repressive histone mark H3K27Me3, which has been shown to increase the mutation rate in *Z. tritici*.

In this study I aim to investigate whether the distinct chromatin environment of accessory chromosomes dictates whether DSBs are repaired via the error-prone NHEJ pathway or the more-reliable HDR pathway, thereby actively modulating mutation rates at specific genomic locations in *Z. tritici*. A CRISPR-Cas9-based approach will be used to create DSBs at target loci on core and accessory chromosomes followed by quantification of the choice of repair pathway employed.

## Introduction (400)

Evolutionary theory has been dominated by the idea that mutations occur randomly with respect to their consequences. Yet, the mutation rate varies along a genome, such that genetic diversity is lower in functionally important regions, such as coding sequences [1]. A central paradigm in biology is that these differences occur as a result of natural selection. However, recent discoveries suggest that organisms can actively modulate the frequency of variants across their genomes [2]. Specific factors and mechanisms would maintain low mutation rates at loci where mutations are most deleterious, while stimulating mutation rates where they confer an immediate adaptative advantage. One proposed way in which organisms might control the mutation rate across the genome is through the regulation of chromatin structure and function. Interestingly, a recent large-scale analysis of *de-novo* mutations in *Arabidopsis thaliana* demonstrates that the *de novo* mutation frequency is reduced by half inside gene bodies and by two-thirds in essential genes [2]. This research demonstrated that epigenomic features explain over 90% of variance in the genome-wide pattern of mutation bias.

*Zymoseptoria tritici*, (formerly *Mycosphaerella graminicola*) which causes the major fungal wheat disease septoria tritici blotch, has emerged as a powerful model system for the study of evolutionary trends associated with speciation and host-plant adaptation [3,4]. The *Z. tritici* genome is contained on 13 “core” chromosomes, found in all isolates, and up to 8 dispensable, “accessory” chromosomes (Figure 1) [5,6]. While the core chromosomes have been described as primarily made up of euchromatic regions rich in housekeeping genes, the accessory chromosomes have a lower gene density and are comprised mainly of heterochromatic regions [6,7]. It has been suggested that accessory chromosomes may be conducive for rapid adaptive evolution of virulence-related genes, perhaps due to fewer selective constraints on sequences. Indeed, genes located on the accessory chromosomes appear to evolve considerably faster compared to genes on the core chromosomes [8].

A recent mutation accumulation study was conducted in *Z. tritici* to investigate the impact of epigenetic modifications on mitotic mutation rates [8]. In addition to observing an increased mutation rate in accessory chromosomes, it was observed that depletion of the repressive histone mark H3K27me3 resulted in decreases of variants, demonstrating a link between chromatin structure and the mutation rate. Together, these observations underscore an incomplete understanding of the forces driving evolution, and the extent to which epigenomic features have been harnessed by organisms to maximize their environmental adaption.

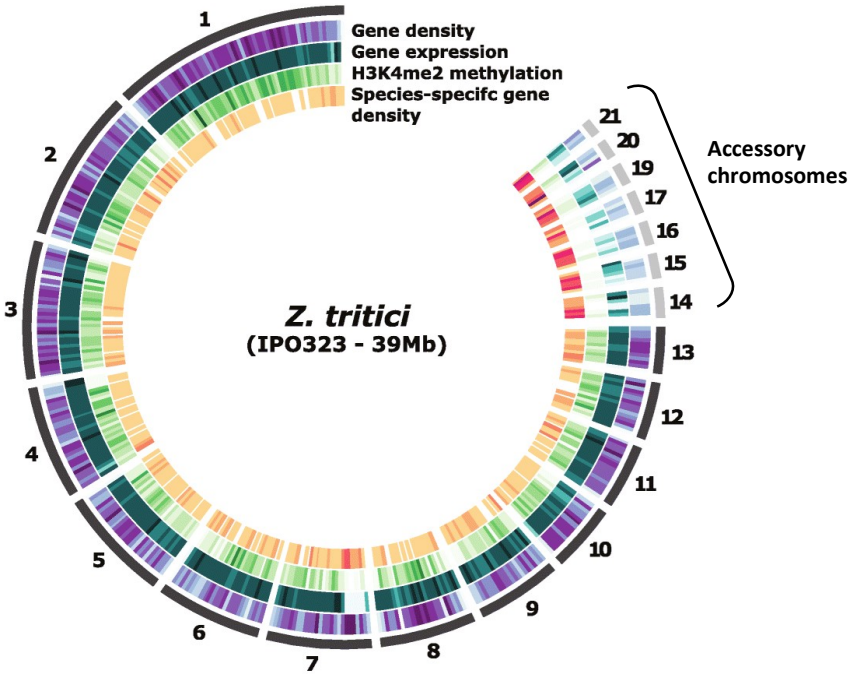


Figure 1: Genome architecture of the reference *Z. tritici* genome IPO323 [6].

## Research project for the period of training + specific aims

The question I pose is "why do accessory chromosomes have a higher mutation rate in *Z. tritici*?". I propose that there is an epigenetic mechanism controlling the mutation rate. This hypothesis is based on the evidence that:

- in *A. thaliana*, epigenetic features explain over 90% of the mutation rate variation [1],
- in *Z. tritici*, accessory chromosomes have a higher mutation rate than core chromosomes and are linked to virulence and niche/host adaption, for which a more rapid evolution would presumably be beneficial [8],
- in *Z. tritici,* accessory chromosomes are significantly enriched in H3K27Me3 which promotes an increase in mutation rate [7].

However, directly investigating the mutation rate is labor intensive and given the rarity of mutations, one can only examine factors that affect mutation rates on a global scale, rather than at specific loci. This difficulty prompted me to search for other ways to test this hypothesis. Given that the chromatin environment controls how DNA can interacts with proteins, one way in which the epigenome might influence mutation rates is by influencing the repair mechanism of DNA double stranded breaks to increase the likelihood of mutations in certain chromatin contexts [9,10]. The two main DSB repair pathways, non-homologous end joining (NHEJ) and homology-directed repair (HDR), have different mutation frequencies and outcomes and thus could conceivably be employed to modulate the mutation rate to the benefit of the organism [11]. Here I propose to

investigate whether histone modifications influence the choice of DSB repair pathway in *Z. tritici* and to establish a causal role of epigenetic factors in the modulation of the mutation rate.

The research project will consist in implementing a CRISPR-Cas9 based system to introduce double strand breaks at specific target loci within either accessory or core chromosomes, and determine which predominant pathway promotes their repair. For this, I will investigate whether the repair resulted in incorporation of a homologous template via HDR or introduction of nucleotide insertions or deletions (INDELS) via NHEJ. Two approaches, described in detail below, will be employed. The first is a sequencing-based approach to characterize the repair pathway choice at endogenous loci while the second will utilize fluorescence reporter-constructs to elicit different phenotypes depending on the repair pathway utilized.

## Significance, importance, novelty in the field

The aim of this research is to advance our understanding of the influence of epigenetic marks on DNA repair and on the rate of spontaneous mutation. Appreciating the unique genomic landscape of Z. tritici led me to wonder what is responsible for the increased mutation rate at accessory chromosomes. Any such information would provide a very important step towards understanding *Z tritici*'s rapid evolution of resistance to pesticides and perhaps towards development of more resilient pesticides or wheat cultivars.

However, given the distinct chromatin environment of accessory chromosomes, their roles in adaptive evolution, and the mounting evidence that chromatin features have a role in shaping the mutation rate, it became evident that identifying the factors responsible for the increased mutation frequency at accessory chromosomes could have much broader implications in evolutionary biology and epigenetic research, potentially demonstrating existence of a novel evolutionary mechanism utilizing histone marks to modulate the mutation rate to provide an adaptory benefit to an organism.

If this analysis finds that chromatin environments modulate the mutation rate by influencing the choice of DSB repair mechanism, it will not necessarily imply that the system is actively tailored to produce this outcome. Further studies would be needed to clarify whether the increased mutation rate as a result of a repressive chromatin environment is actively sought by the cell to boost adaptive evolution in regions with less selection constraints, or simply a fortunate side-effect of an increased surveillance-and-repair mechanism reportedly at play in actively transcribed regions [12].

## Development of the project, methods, and procedures

### Establishment of a Z. tritici lab strain

The finished Z. tritici genome for the reference WT strain IPO323 was published in 2011 and is among the most complete reference genomes available [13]. This strain is available from The Westerdijk Fungal Biodiversity Institute. Extensive protocols are available detailing efficient and

reliable culturing conditions. All culturing will be done at 18°C to maintain Z tritici in yeast-like morphology.

**Epigenetic characterization**

To confirm the epigenetic environment of the target loci, histones with the modification of interest will be immunoprecipitated by adding antibodies to purified, soluble chromatin, following formaldehyde fixation, cell lysis and chromatin shearing [14]. qPCR with a variety of primers of target regions will be performed on DNA isolated from immunoprecipitated chromatin, mock control chromatin or total chromatin.

**Genetic editing**

This analysis will rely on the establishment of a functional CRISPR based system to introduce double-strand breaks at target regions. Codon-optimization of foreign DNA is necessary for functional expression in fungal systems [15]. On the basis of the codon frequency of Z. tritici, we will optimize the Cas9 gene of *Streptococcus pyogenes* and the sequence of the SV40 nuclear localization signal to maximize the likelihood of functional Cas9 expression. The codon-optimized Cas9 will be placed under the control of an RNA polymerase II promoter, and inserted into a plasmid vector alongside a hygromycin selectable marker, flanked by regions of homology for the desired genome incorporation site. This plasmid will be integrated into the genome at the "soft-landing" sd1 site using the protocol of by Kilaru et al. employing the *Agrobacterium tumefaciens* mediated transformation method which has been well characterized in *Z. tritici* [16,17]. Given the multiplicity of targets, sgRNA designed using a 20bp sequence of the target gene PCR amplified into a DNA template containing the constant tracrRNA sequence, will be transcribed *in vitro* (ThermoFisher GeneArt™ Precision gRNA Synthesis Kit) and delivered via using a polyethylene glycol-mediated protoplast transfection procedure, alongside homologous templates for the target region, with homology arms ranging from 200bp to 1Kbp [18].

**Investigating DSB repair pathway**

Two approaches are proposed with each having their own advantages and disadvantages. The first possibility is to target the Cas9 to create a break in an endogenous locus to which a homologous template will be provided. Sanger sequencing of a PCR amplified locus subcloned into vector will enable a direct readout of the frequency with which the locus is unmodified, incorporated small INDELs, or incorporated a homologous template, and thus the tendency to repair the break by NHEJ or HDR. This approach has the advantage that multiple loci can be easily and simultaneously targeted by designing a suitable multiplexed sgRNA vector, however the interpretation of the results might be complicated by the fact that insertion of a large reporter gene in the homologous template may inherently lower the efficiency of HDR vs NHEJ and bias the quantification of repair pathway choice by Next Generation Sequencing (NGS) [19].

I will therefore implement a second approach, which utilizes a reporter construct, which is integrated into the genome on core and accessory chromosomes. Targeting of the Cas9-sgRNA to create a DSB in the reporter construct will enable a phenotypic readout depending on the repair

pathway utilized. Introduction of INDELs as a result of repair via the error-prone NHEJ machinery will likely lead to a shift in the reading frame resulting in a truncated or non-functional protein and loss of the fluorescent signal. In contrast, if homology-directed repair is employed to incorporate a homologous template containing a RFP cassette in place of the GFP reporter, a shift in the fluorescent wavelength will demonstrate a repair via the HDR pathway. The difficulty in this approach is effective incorporation of the reporter construct at several loci, while maintaining the epigenetic environment and expression of the reporter gene. Repair frequencies will be quantified by flow cytometry.

**Perspective**

Establishment of a functional CRISPR-Cas9 system to introduce DSBs at target loci, and a successful implementation of one or both of the techniques outlined to observe the frequency of each repair mechanism will enable us to determine whether modulation of DSB repair mechanism is in part responsible for the increasingly evident correlation between chromatin state and the mutation rate. A successful analysis, providing evidence for such a role of chromatin, will prompt further investigation and an attempt to demonstrate causality by repeating the analysis with global deletion of H3K27Me3 via *Kmt1* knockout as done by Habig et al [8]. To eliminate potential secondary effects as a result of global deletion of a histone writer protein, a local "epigenetic editing" approach could be envisaged to locally modify the epigenetic environment by targeting a non-endonucleolytic dead-Cas9, fused to the H3K27me3 demethylase, to the target site. This novel use of the CRISPR machinery is gaining popularity, providing unprecedented ability to investigate epigenetic phenomenon at a local scale. [20]

## Facilities available and justification of the budget

This project will require readily-available media and widely-used reagents for conventional microbiology (cell-line maintenance, transformation), biochemistry and molecular biology experiments (ChIP-qPCR, DNA extraction, PCR). In terms of facilities and equipment, this project will require laboratory space equipped to handle BSL-1 biological material, with incubators and sterile laminar-flow hoods for aseptic cell-culture. PCR and qPCR machines will be indispensable throughout this project. Coordination with external platforms will be required for flow-cytometry analyses and targeted deep-sequencing. The open-source bioinformatics program, R equipped with free tools such as Fastq-join, and Cas-analyzer will be used to analyze in-house the resulting NGS data for quality-control and repair pathway quantification.

| Application | Cost |
|---|---|
| Establishment and culture of *Z. tritici*. | €2,000 |
| Establishment of CRISPR-Cas9-sgRNA system in *Z. tritici.* | €5,000 |
| Biochemistry and molecular biology (ChIP, qPCR, DNA extraction, PCR) | €15,000 |
| Targeted deep sequencing and data analysis | €15,000 |
| Flow cytometry | €1,000 |
| Total | €38,000 |

# Bibliography


1. Joshi, M., Kapopoulou, A., & Laurent, S. (2021). Impact of Genetic Variation in Gene Regulatory Sequences: A Population Genomics Perspective. *Frontiers in Genetics*, *12*. https://www.frontiersin.org/article/10.3389/fgene.2021.660899

2. Monroe, J. G. et al. (2022). Mutation bias reflects natural selection in *Arabidopsis thaliana*. *Nature*, *602*(7895), 101–105. https://doi.org/10.1038/s41586-021-04269-6

3. Marshall, R. (2011). Analysis of two in planta expressed LysM effector homologs from the fungus *Mycosphaerella graminicola* reveals novel functional properties and varying contributions to virulence on wheat. *Plant Physiology*, *156*(2), 756–769. https://doi.org/10.1104/pp.111.176347

4. Poppe, S. et al. (2015). Rapidly Evolving Genes Are Key Players in Host Specialization and Virulence of the Fungal Wheat Pathogen *Zymoseptoria tritici (Mycosphaerella graminicola*). *PLOS Pathogens*, *11*(7), e1005055-. https://doi.org/10.1371/journal.ppat.1005055

5. Steinberg, G. (2015). Cell biology of *Zymoseptoria tritici*: Pathogen cell organization and wheat infection. *Fungal Genetics and Biology*, *79*, 17–23. https://doi.org/https://doi.org/10.1016/j.fgb.2015.04.002

6. Feurtey, A. et al. (2020). Genome compartmentalization predates species divergence in the plant pathogen genus Zymoseptoria. *BMC Genomics*, *21*(1), 588. https://doi.org/10.1186/s12864-020-06871-w

7. Schotanus, K. (2015). Histone modifications rather than the novel regional centromeres of *Zymoseptoria tritici* distinguish core and accessory chromosomes. *Epigenetics & Chromatin*, *8*(1), 41. https://doi.org/10.1186/s13072-015-0033-5

8. Habig, M. et al. (2021). Epigenetic modifications affect the rate of spontaneous mutations in a pathogenic fungus. *Nature Communications*, *12*(1), 5869. https://doi.org/10.1038/s41467-021-26108-y

9. Clouaire, T. et al. (2018). Comprehensive Mapping of Histone Modifications at DNA Double-Strand Breaks Deciphers Repair Pathway Chromatin Signatures. *Molecular Cell*, *72*(2), 250-262.e6. https://doi.org/10.1016/j.molcel.2018.08.020

10. Miyaoka, Y. et al. (2016). Systematic quantification of HDR and NHEJ reveals effects of locus, nuclease, and cell type on genome-editing. *Scientific Reports*, *6*(1), 23549. https://doi.org/10.1038/srep23549

11. Lobinska, G., Pilpel, Y., & Ram, Y. (2021). Epigenetically Inherited Mutation Rates Predicted to Maximize Adaptation Rates. *BioRxiv*, 2021.07.14.452333. https://doi.org/10.1101/2021.07.14.452333

12. Aymard, F. et al. (2014). Transcriptionally active chromatin recruits homologous recombination at DNA double-strand breaks. *Nature Structural & Molecular Biology*, *21*(4), 366–374. https://doi.org/10.1038/nsmb.2796

13. Goodwin, S. B. et al. (2011). Finished Genome of the Fungal Wheat Pathogen *Mycosphaerella graminicola* Reveals Dispensome Structure, Chromosome Plasticity, and Stealth Pathogenesis. *PLOS Genetics*, *7*(6), e1002070-. https://doi.org/10.1371/journal.pgen.1002070

14. Soyer, J. L. et al. (2015). Chromatin analyses of *Zymoseptoria tritici*: Methods for chromatin immunoprecipitation followed by high-throughput sequencing (ChIP-seq). *Fungal Genetics and Biology*, *79*, 63–70. https://doi.org/https://doi.org/10.1016/j.fgb.2015.03.006

15. Song, R. et al. (2019). CRISPR/Cas9 genome editing technology in filamentous fungi: progress and perspective. *Applied Microbiology and Biotechnology*, *103*(17), 6919–6932. https://doi.org/10.1007/s00253-019-10007-w

16. Kilaru, S. et al. (2015). A gene locus for targeted ectopic gene integration in *Zymoseptoria tritici*. *Fungal Genetics and Biology*, *79*, 118–124. https://doi.org/https://doi.org/10.1016/j.fgb.2015.03.018

17. Zwiers, L.-H., & de Waard, M. A. (2001). Efficient Agrobacterium tumefaciens-mediated gene disruption in the phytopathogen *Mycosphaerella graminicola. Current Genetics*, *39*(5), 388–393. https://doi.org/10.1007/s002940100216

18. Payne, A. C., Grosjean-Cournoyer, M.-C., & Hollomon, D. W. (1998). Transformation of the phytopathogen *Mycosphaerella graminicola.* to carbendazim and hygromycin B resistance. *Current Genetics*, *34*(2), 100–104. https://doi.org/10.1007/s002940050372

19. Song, B. et al. (2021). Analysis of NHEJ-Based DNA Repair after CRISPR-Mediated DNA Cleavage. *International Journal of Molecular Sciences*, *22*(12), 6397. https://doi.org/10.3390/ijms22126397

20. Brocken, D.J.W. et al. (2018). dCas9: A Versatile Tool for Epigenome Editing. Curr. Issues Mol. Biol. 26, 15-32. https://doi.org/10.21775/cimb.026.015