Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Spatio-Temporal Repeat-Induced Point Mutation (RIP) Dynamics in Global Parastagonospora nodorum Pan-genomes

Mohitul Hossain

mdmohitul.hossain@curtin.edu.au

Curtin University https://orcid.org/0000-0002-6833-2234

Darcy Jones

Curtin University

Naomi Gray

Curtin University

Angela Williams

Curtin University

Huyen Phan

Curtin University https://orcid.org/0000-0002-8971-2308

James Hane

Curtin University https://orcid.org/0000-0002-7651-0977

Article

Keywords: Repeat-Induced Point Mutation (RIP), Parastagonospora nodorum, Pan-genome, Crop fungal diseases, Spatio-temporal dynamics, Effector diversification, Phylogeography, Climate-associated population shift

Posted Date: November 13th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11486955/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: There is **NO** Competing Interest.

Title: Spatio-Temporal Repeat-Induced Point Mutation (RIP) Dynamics in Global *Parastagonospora nodorum* Pan-genomes

Mohitul Hossain[1*], Darcy Jones[1], Naomi Gray[1], Angela Williams[1], Huyen Phan[1], James Hane[1]

[1]Centre for Crop and Disease Management (CCDM), Curtin University, Kent Street, Bentley, Western Australia 6102, Australia

*Corresponding author: Mohitul Hossain; email: mdmohitul.hossain@curtin.edu.au

Abstract:

Repeat-induced point mutation (RIP) is a fungal genome defence mechanism that generates extensive C→T-biased mutations, yet its spatial and temporal dynamics across pathogen populations remain poorly understood. We analysed whole-genome variation in a global pan-genome of 649 *Parastagonospora nodorum* isolates, including 269 from the Western Australian (WA) wheatbelt and 380 from international populations. Populations showed strong geographic differentiation. WA and United States (US) populations had the highest overall RIP activity, largely confined to non-genic regions, whereas European and South African populations showed greater non-synonymous RIP enrichment, indicating selection for functional diversification via RIP leakage into coding regions. RIP signatures accumulated earlier in global populations, particularly the US, whereas WA populations showed a more recent but rapid increase in high-RIP regions. Chromosomes harbouring major effector loci exhibited elevated RIP signatures, particularly around ToxA (Chr4) and Tox1 (Chr10) in WA populations, suggesting RIP accumulation may mark locally important virulence-associated regions. Although rainfall

gradients and sampling decade did not explain population-wide effector composition, phylogeographic analyses revealed evidence consistent with climate-associated pathotype movement across the wheatbelt. Spatio-temporal RIP dynamics thus generate distinct genomic signatures that can help identify adaptive effector regions and the evolutionary drivers of virulence in fungal crop pathogens.

Keywords: Repeat-Induced Point Mutation (RIP), *Parastagonospora nodorum*, Pan-genome, Crop fungal diseases, Spatio-temporal dynamics, Effector diversification, Phylogeography, Climate-associated population shift

1. Introduction

Septoria nodorum blotch (SNB) of wheat (*Triticum aestivum*) is caused by haploid necrotrophic fungal pathogen *Parastagonospora nodorum*, a globally important disease that reduces grain quality and yield [1,2]. In major wheat-growing regions such as Western Australia, SNB typically causes 5–15% yield loss, with severe epidemics exceeding 30% [3,4]. Disease development is governed by complex, often additive or epistatic interactions between secreted necrotrophic effectors (NEs), and the interactions of NEs with cognate host susceptibility genes [5-7]. *P. nodorum* populations are highly diverse and predominantly reproduce sexually, with long-distance dispersal via infected seed, crop residues, and wind-borne ascospores, and secondary spread through splash-dispersed pycnidiospores [8]. The combination of frequent recombination, high standing genetic variation, and strong agricultural selection pressure confers substantial adaptive potential [9-11],

making *P. nodorum* a useful model for investigating genome evolution, population structure, and repeat-associated mutational processes in necrotrophic plant pathogens.

Transposable elements (TEs) are major drivers of genome size variation and genome architecture across eukaryotes and constitute the primary source of non-coding DNA expansion [12]. In fungi, TE content varies widely [13-15], from near absence in compact yeast genomes to 18.6% of the genome in *Zymoseptoria tritici* [16] to more than 70% of the genome *Blumeria graminis (*64% to 90%) [15]. TE mobilization imposes opposing evolutionary pressures: while insertions can be deleterious and destabilize genome integrity, TEs also generate genetic variation (i.e. duplications) that may facilitate adaptation [17,18]. Many fungal plant pathogens exhibit genome regions enriched in TEs that evolve at accelerated rates through sequence rearrangements, with elevated polymorphism in TE-proximal genes relative to conserved regions [19-21].

Repeat-induced point mutation (RIP) is a fungal-specific genome defense mechanism against TEs, that targets duplicated DNA during sexual reproduction, introducing C:G to T:A transitions primarily in CpA/TpG contexts across the phylum Ascomycota [14,22-24]. First described in *Neurospora crassa*, RIP generates the highest known mutation rates outside viruses, reaching 3.38 $\times 10^{-6}$ per bp per generation—several orders of magnitude higher than mitotic mutation rates in the same species [25]. Two mechanistically distinct RIP pathways have been characterized in *N. crassa*: a Rid1-dependent pathway acting directly on repetitive sequences and a Dim2/Dim5-dependent pathway targeting single-copy regions flanking repeats. Although RIP signatures are widespread among ascomycetes, the underlying machinery shows a patchy phylogenetic distribution, with repeated loss of key components (Dim5, Rid1, Dim2) across the phylum [26-28].

Importantly, RIP activity can extend beyond repeats into adjacent non-repetitive regions, generating mutational gradients ("RIP leakage") that affect nearby genes [29]. In plant pathogenic fungi, such leakage has significant evolutionary consequences, often driving the rapid diversification of effector genes located in repeat-rich regions and the emergence of compartmentalized genomes with contrasting repeat and gene content, a phenomenon well-documented in *P. nodorum*, *Leptosphaeria maculans*, and *Neurospora crassa* [9,20,30]. A very recent synthesis argues that the molecular machinery underlying the fungal RIP "hypermutator" phenotype has deeper evolutionary roots shared with other eukaryotic lineages, rather than being wholly fungal-specific [31].

Population-level analyses highlight the complexity of genome evolution in *P. nodorum*, shaped by the intertwined dynamics of transposable elements (TEs) and repeat-induced point mutation (RIP). This interplay can affect virulence, as key pathogenicity genes, including necrotrophic effectors, are frequently located near repetitive elements [9,11,19]. Although TEs account for only ~6% of the *P. nodorum* genome, proximity to repeats strongly correlates with gene presence–absence polymorphisms and is associated with horizontal transfer events, such as the transfer of the *ToxA* effector locus between wheat pathogen species [32,33], illustrating the dual role of repeats in promoting both genome instability and adaptive innovation. These processes operate within a "two-speed" genome architecture, in which polymorphisms and gene gains or losses are unevenly distributed within repeat-rich, G:C-depleted DNA regions [11]. Recent pangenomic analyses revealed that RIP-like mutations are concentrated on accessory chromosome 23 (AC23) and in the subtelomeric regions of core chromosomes, where secreted effector candidates show accelerated evolution [9,10]. Broader pangenome analysis revealed RIP-like SNPs occur frequently across the

whole pangenome, and that RIP-like mutations "leak" from TE-rich sequences into neighbouring non-repetitive DNA regions, including effector loci where they account for over half of non-synonymous amino acid changes [9]. This high potential for RIP-driven diversification is counter-balanced by strong purifying selection against non-synonymous SNPs within most gene coding regions, maintaining essential gene functions and isolate viability while enabling rapid adaptation at the genome's more dynamic margins.

Pangenomic surveys across *P. nodorum* populations from Iran, Europe, Africa, Australia, and the USA revealed strong geographic structuring, with Iranian populations displaying the highest genetic diversity, consistent with a longer co-evolutionary history during early wheat domestication in the Fertile Crescent [34,35]. In North America, population-scale resequencing of 197 U.S. isolates showed the effect of geographically structured selection on effector loci, by identifying two distinct (North-South) pathogen populations that respond differently to local wheat varieties based on host susceptibility [33]. Similarly, the necessity for high-resolution, population-level surveillance is further underscored by regional studies from Western Canada, where the discovery of *SnToxA* Haplotype 21 (H21), alongside the high prevalence of the H5 haplotype, demonstrated that local adaptations can lead to distinct pathogenic profiles that challenge traditional resistance breeding strategies [36]. In Western Australia, surveys spanning 44 years documented temporal shifts in population structure associated with changes in wheat cultivar deployment, particularly following widespread adoption of SnToxA-insensitive cultivars after 2013, with emergent clusters reported to be more aggressive despite increased host resistance [37].

Despite increasing population-level studies of *P. nodorum*, the spatial and temporal patterns of RIP signatures - and how differences in reproductive mode, demography, and selection influence their accumulation - remain incompletely resolved. RIP-like mutation signatures have been systematically assessed across the fungal kingdom, demonstrating that canonical C→T–biased mutational patterns are largely restricted to Ascomycetes and closely associated with RIP-related genes [38]. However, this study did not explore how RIP varies within species across agro-ecological zones or through time, and while it provided critical insight into the role of RIP in shaping local genomic diversity and effector adaptation, it was geographically restricted and did not address broader spatial or temporal dynamics. Consequently, how RIP signatures vary across global *P. nodorum* populations, how these patterns have changed over time, and whether regional differences in RIP activity reflect distinct evolutionary or epidemiological histories remain unknown. Addressing these gaps is essential for understanding RIP's contribution to effector-mediated virulence, long-term pathogen adaptation, and genome evolution at a global scale.

Results

Phylogeny and population structure of the global *Parastagospora nodorum* pangenome indicate similarities and differences between major regions that may be associated with climate

We analysed global genome sequencing data for 648 *P. nodorum* isolates and identified 4,778,479 high-confidence SNP markers with a minor allele frequency of 5%. RIP-like C:G→T:A substitutions (C→T and G→A) represented 37.9% of all SNPs and constituted the most frequent

mutation class in the genome. SNPeff annotation showed that 1,001,505 SNPs (21.0%) resulted in non-synonymous amino-acid changes, whereas 1,799,286 (37.7%) were synonymous, and 26,039 (0.5%) introduced premature stop codons, indicating that only a subset of mutations are predicted to directly alter protein function. To determine how this genome-wide variation is structured across global populations, we next examined phylogenetic relationships and population structure among isolates.

The Western Australian (WA) *P. nodorum* population exhibited substantial diversity compared with other major global regions (Figure 2). Although this region was initially subdivided into hot, warm and cold sub-groups (North to South) (Supplementary Table 1B), these latitudinal and climatic sub-groupings did not appear to be grouped phylogenetically. The WA isolates were phylogenetically most similar to some (but not all) Iranian isolates and to South African isolates and were distinct from isolates from the USA and Europe. Most US isolates belonged to two distinct groups – one from warmer regions (Southern) and another from colder regions (Northern), as reported in a previous study [19]. US isolates from warm regions were most similar to isolates from Europe, and US isolates from cold regions were most similar to (the non-WA-like) Iranian isolates. The two mating-type loci (MAT1-1 and MAT1-2) were observed in all regions across the global pangenome. Within WA, the MAT1-1 and MAT1-2 loci were dispersed throughout the population at nearly equal ratios (Figure 2), with the exception of a few isolates that were phylogenetically extremely similar and restricted to a small geographic area.

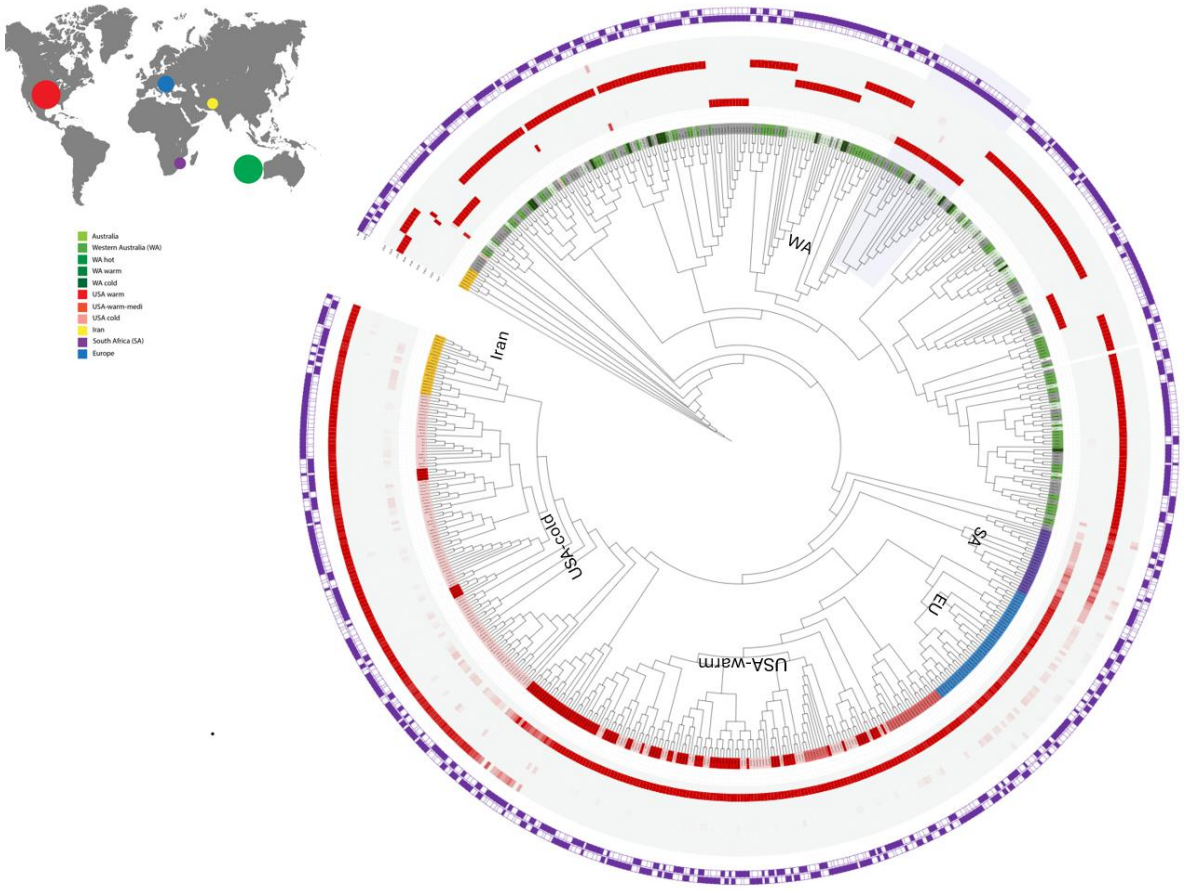


Figure 2: Global phylogenetic tree of *Parastagonospora nodorum* [clades color-coded by region], showing the diversity of mating type loci [purple] and population sub-group structure [teal] across 5 major regions: the USA, Europe, Iran, South Africa, and Western Australia (WA). The USA was sub-divided into warm and cold regions (South to North), and WA was sub-divided into hot, warm and cold regions (North to South).

The distribution of RIP mutation signatures across the world indicates variable selection pressures between different *P. nodorum* chromosomes, with widespread selection against non-synonymous RIP mutations that is contrasted by hotspots of RIP-diversification

To quantify RIP accumulation across the pangenome, we calculated a normalised metric termed 'RIPness', representing the proportion of RIP-like C→T and G→A transitions within non-overlapping 10 kb genomic windows relative to all C→N and G→N mutations, allowing comparison across regions of variable GC content. Comparisons of whole-genome RIPness across the major global regions showed that isolates from the WA and US regions had higher proportions of RIP-affected regions. Across all SNPs, Western Australian (WA) and US populations exhibited the highest number of high-RIP windows (>0.6 RIPness per 10 kb), significantly exceeding those in Europe, South Africa (SA), and Iran (Figure 3). This may suggest that the WA and US populations are 'hyper-adaptive' relative to the rest of the world. However, this scenario shifts when considering SNPs within protein-coding genes and those that lead to non-synonymous amino acid changes. For non-synonymous SNPs, Europe and SA displayed the highest values, while WA and US showed the lowest. These regional differences indicate that the high whole-genome RIPness of WA and the US is concentrated in non-genic compartments (e.g. TEs, intergenic regions) that lead to protein and functional changes less often, whereas Europe and SA exhibited stronger RIPness, leading to non-synonymous changes (Figure 3).

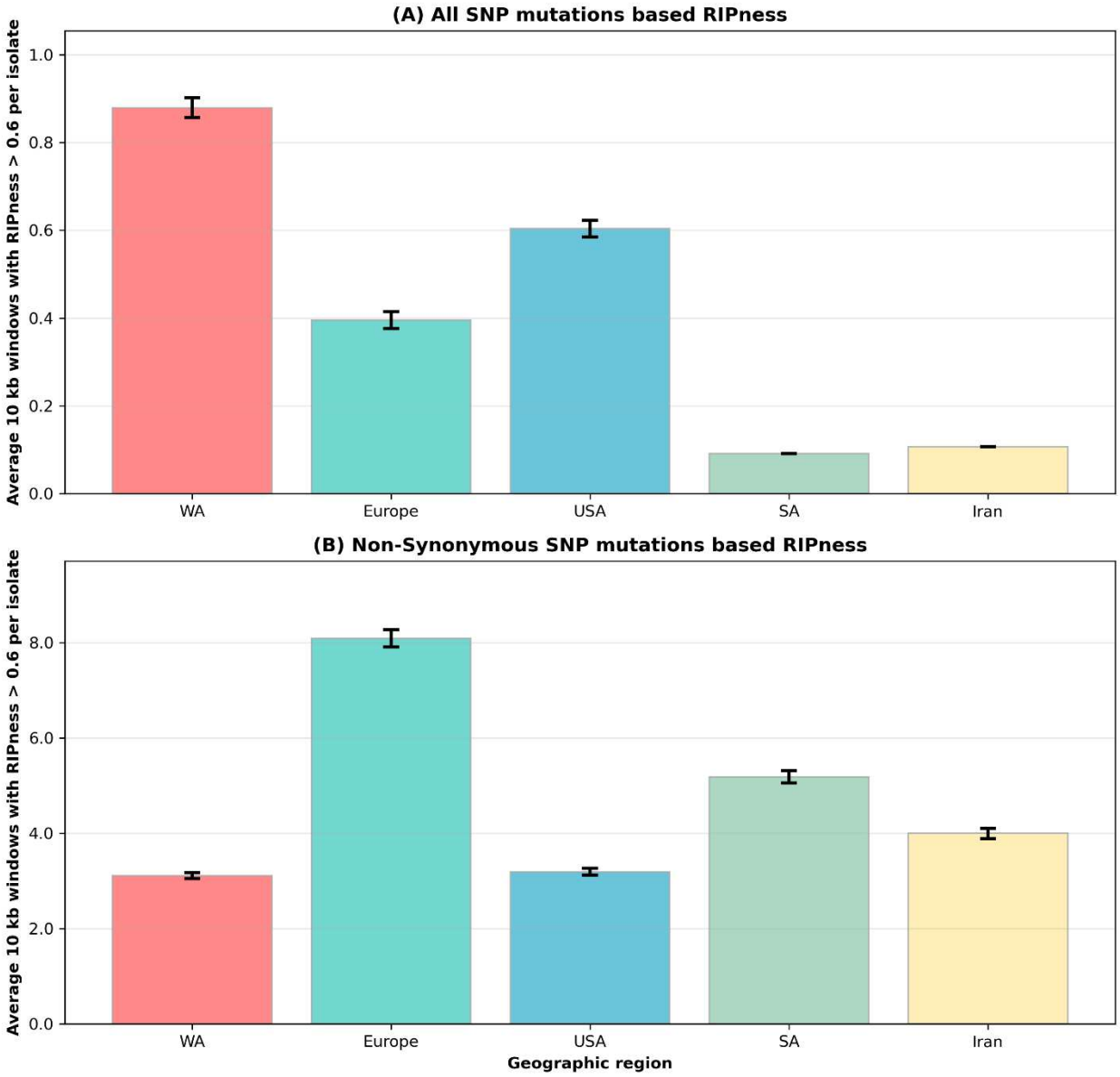


Figure 3. Comparison of whole-genome RIPness across six sub-populations of *Parastagonospora nodorum*, from Western Australia (WA), Europe, the USA, South Africa (SA) and Iran. Bar plots indicate the whole-genome average RIPness calculated from all 10kb genomic windows with RIPness > 0.6, and averaged for all isolates belonging to each region. Data is presented for two categories of SNP mutation: (A) All SNPs and (B) Non-synonymous SNPs. Error bars represent standard deviation.

In addition to these population-level patterns at the whole-genome level, we also observed chromosome-specific RIP dynamics that further highlighted regional differences (Figure 4; Supp Figure 2). Consistent with population-level trends, chromosome-level analysis of all SNPs revealed more high-RIP windows in WA and US populations, whereas only a few chromosomes showed elevated RIP in SA and Iranian populations. In contrast, for non-synonymous SNPs, European populations exhibited higher RIPness across most chromosomes relative to other regions (Figure 4; Supp Figure 2). Examination of chromosomes harbouring known effector loci—Chr4 (*ToxA*), Chr10 (*Tox1*), Chr11 (*Tox3*), and Chr12 (*Tox267*) - based on alignment to the SN15 (WA) reference genome - revealed elevated RIPness in Europe and SA relative to WA and US populations (Supp. Figure 2), suggesting stronger non-synonymous RIP enrichment around effector loci in these regions.

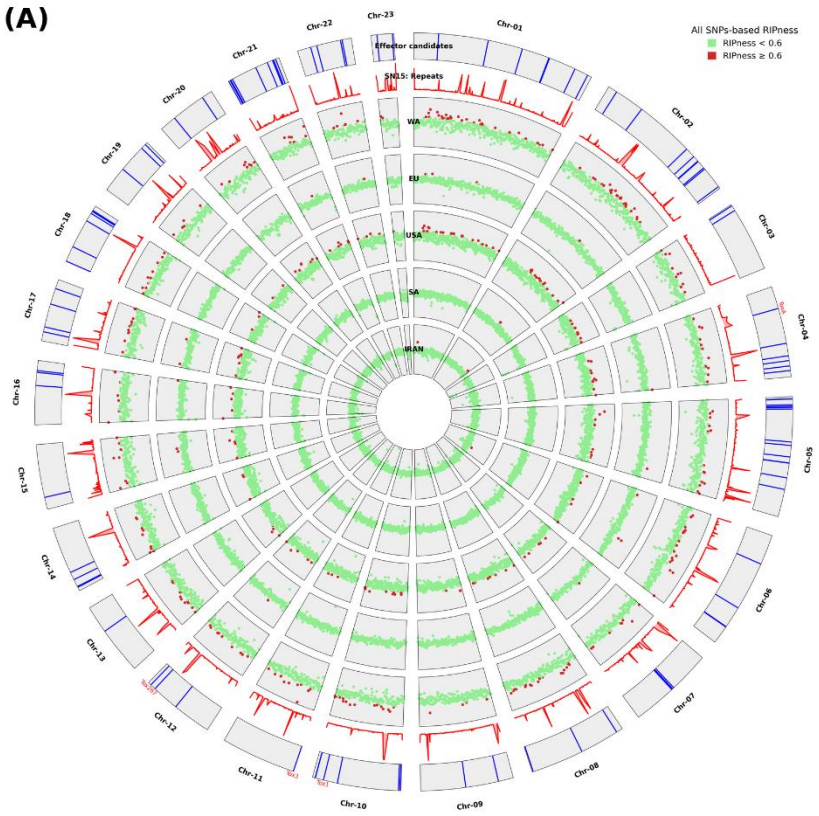
Most Candidate Secreted Effector Proteins (CSEPs) were associated with genomic windows exhibiting elevated RIPness values (Figure 5). However, when considering all SNPs, genomic windows containing known effector and CSEP genes showed no consistent or pronounced difference in RIPness compared to non-effector windows. In contrast, when restricted to non-synonymous SNPs, effector/CSEP-containing windows generally displayed higher RIPness than non-effector windows across most chromosomes, with few exceptions (Figure 5). Among Sn15 chromosomes carrying known effector loci—Chr4 (*ToxA*), Chr10 (*Tox1*), Chr11 (*Tox3*), and Chr12 (*Tox267*)—non-synonymous SNPs exhibited significantly elevated RIPness within effector windows on Chr4 and Chr10, a moderate increase on Chr11, and only a weak signal on Chr12. Figure 4 and 5 further highlights how effector-chromosome importance varies by region, with specific effector-chromosomes having higher RIPness in different sub-populations. This is

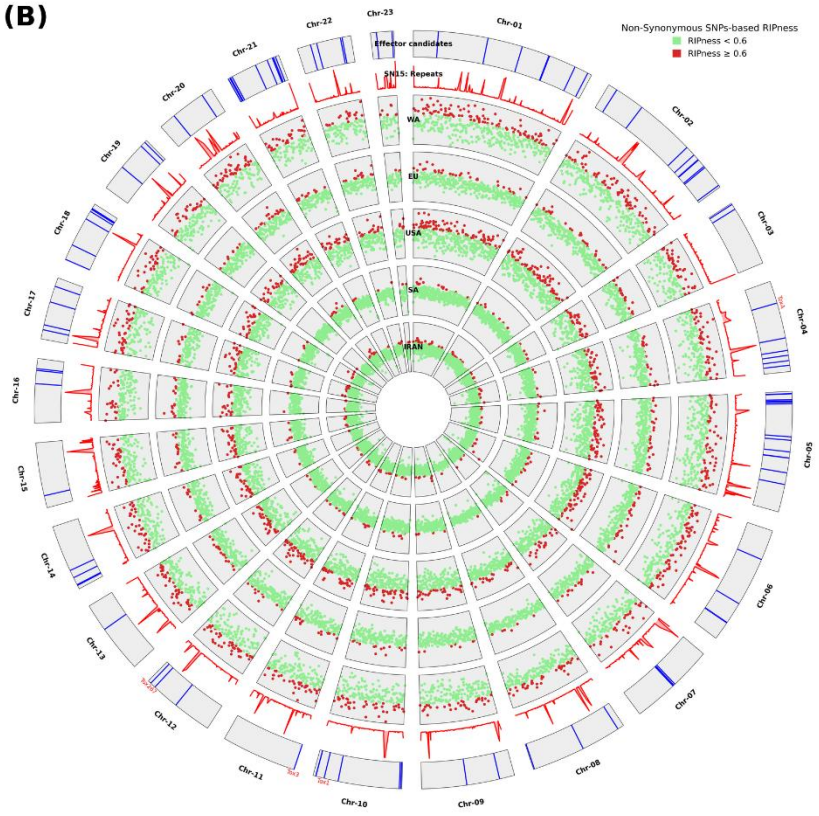
214 particularly evident in WA populations, where regions of Chr4 (*ToxA*) and Chr10 (*Tox1*) around
215 their respective effector loci showed elevated non-synonymous RIP mutations. Together, these
216 spatial patterns of RIP accumulation across global populations raised the question of how RIP
217 signatures have changed over time.

218

219

220


(A)
All SNPs-based RIPness
RIPness < 0.6
RIPness ≥ 0.6
Effector candidates
SIX15: Repeats
WA
EU
USA
SA
IRAN
Chr-01
Chr-02
Chr-03
Chr-04
Chr-05
Chr-06
Chr-07
Chr-08
Chr-09
Chr-10
Chr-11
Chr-12
Chr-13
Chr-14
Chr-15
Chr-16
Chr-17
Chr-18
Chr-19
Chr-20
Chr-21
Chr-22
Chr-23


(B)
Non-Synonymous SNPs-based RIPness
RIPness < 0.6
RIPness ≥ 0.6
Effector candidates
SIX15: Repeats
WA
EU
USA
SA
IRAN
Chr-01
Chr-02
Chr-03
Chr-04
Chr-05
Chr-06
Chr-07
Chr-08
Chr-09
Chr-10
Chr-11
Chr-12
Chr-13
Chr-14
Chr-15
Chr-16
Chr-17
Chr-18
Chr-19
Chr-20
Chr-21
Chr-22
Chr-23

221

Figure 4. Circular genome plot showing genomic distribution of RIPness values (RIPness > 0.6 per isolate) across all chromosomes of *P. nodorum* based on (A) all SNPs and (B) non-synonymous SNPs. Concentric tracks represent populations from Western Australia (WA), Europe (EU), USA, South Africa (SA), and Iran. RIPness scores (10kb) are displayed as coloured scatter points, where green indicates lower RIPness and red indicates higher RIPness. The outermost blue tracks show the positions of effector candidate genes in WA isolates, while the red lines in the outer ring indicate repeat-rich genomic regions of the SN15 reference genome.

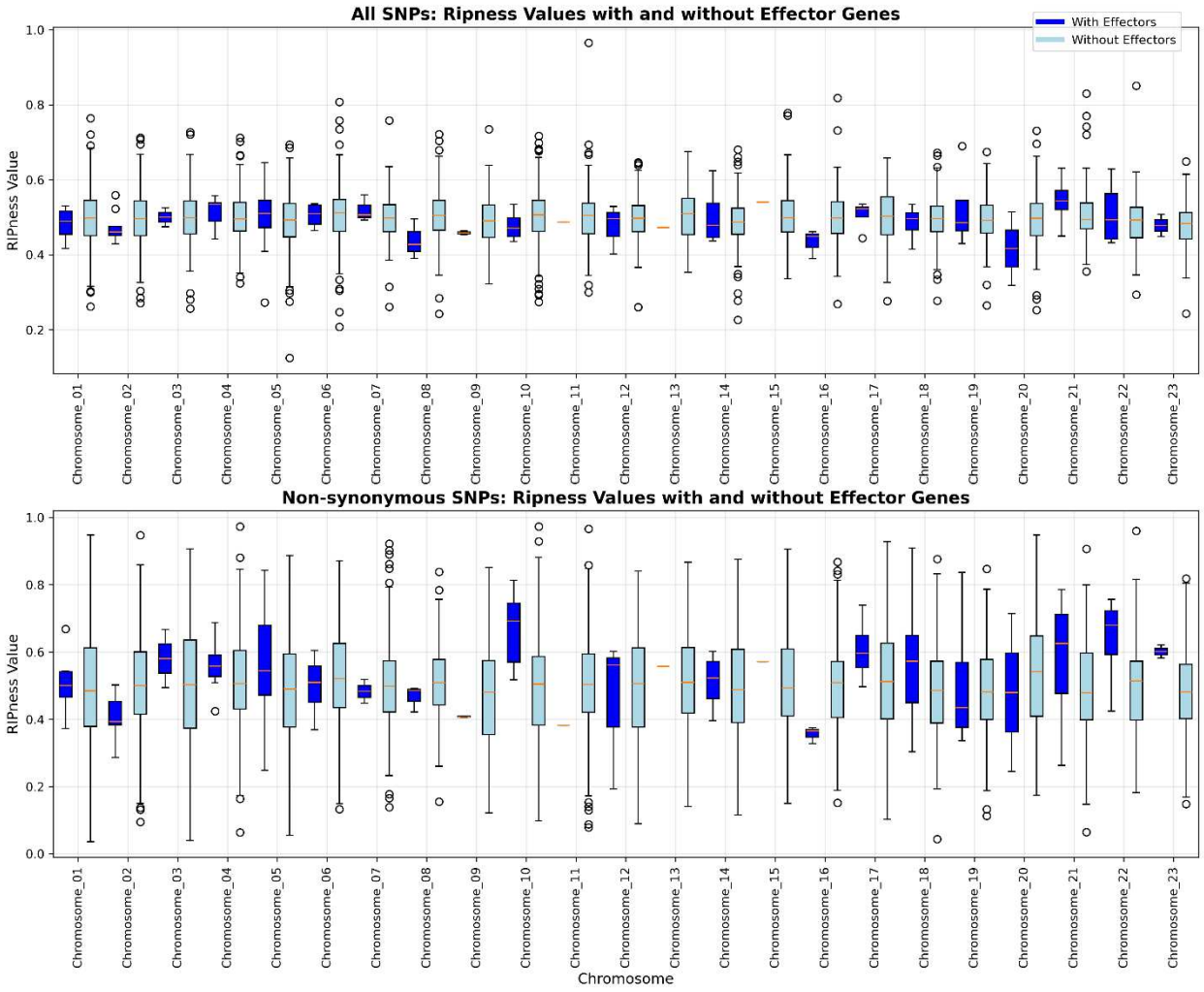


Figure 5: WA-specific (state-level) RIPness values for 10kb genomic windows containing known effector genes or CSEPs (blue) versus windows without effector genes (light blue) across all chromosomes of *P. nodorum*. Data is presented for two SNP categories: (A) All SNPs and (B) Non-synonymous SNPs.

RIP Temporal Dynamics and Chromosomal Heterogeneity across state-level (WA) and global-level *P. nodorum* populations

Analysis of genome-wide RIPness revealed a pronounced temporal signal, with a progressive increase in RIP accumulation across sampling decades at both global and WA scale. At the global level, the trend became apparent in the 1990s, with an average of >0.5 high-RIP windows per isolate (>0.6 RIPness per 10 kb window, based on all SNPs), whereas WA populations exhibited lower values (<0.2) during the same period and reached comparable levels (>0.5) only in the 2000s. Global populations, and USA populations in particular, therefore reached comparably high RIP levels earlier than WA populations. However, over the most recent decade, WA populations exhibited substantially higher RIPness (~1.0) than the global average (~0.7) (Figure 6).

Among all chromosomes, Chr6, Chr11, Chr15, and Chr21 exhibited RIP activity signatures across all 5 decades at a global scale, and at the state level (WA), Chr11, Chr15, and Chr21 varied across 4 decades. These were the only chromosomes showing historic RIP patterns (before 1990s), while the majority exhibited recent RIP signatures (Figure 6). Examination of chromosomes harbouring known effector loci—Chr4 (*ToxA*), Chr10 (*Tox1*), Chr11 (*Tox3*), and Chr12 (*Tox267*)—revealed elevated RIP metrics specifically on Chr11 across multiple decades, extending back to the 1960s (Supp Figure 3 and 4), while the *ToxA*-bearing Chr4 showed detectable RIP signatures beginning in the 1980s. Notably, Chr4 exhibited a clear increase in RIPness over time at both the global and WA scales, indicating increasing selection pressure, whereas Chr11 displayed a contrasting (decreasing) pattern at the global scale, but an increasing trend in WA populations. In contrast,

Chr10 and Chr12 showed RIP signatures largely confined to recent decades. Among the non-(known) effector-containing chromosomes, Chr2, Chr6, Chr15, and Chr21 exhibited historical RIP signatures dating back to the 1960s but showed declining trends globally, except for Chr2. At the WA scale, these chromosomes displayed more variable patterns. The remaining chromosomes predominantly exhibited recent RIP activity, with no evidence of long-term selective retention. To determine whether these temporal RIP dynamics correspond to shifts in pathotype distribution across the WA wheat belt, we conducted phylogeographic analysis of WA isolates in relation to geographic location, sampling year, and local climate data.

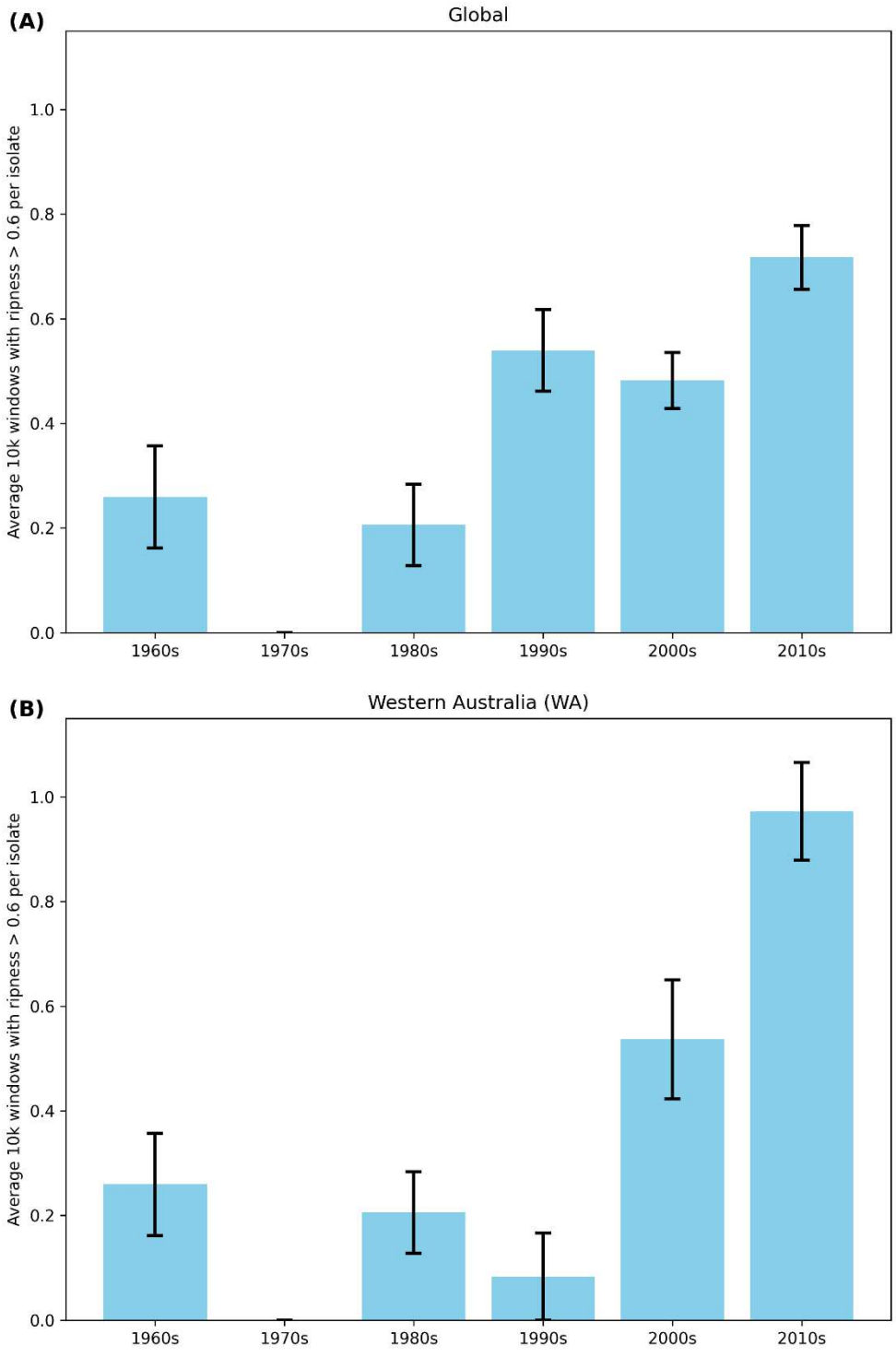


Figure 6: Temporal change in the average number of high-RIPness (> 0.6, based on all SNPs) 10 kb windows per isolate across decades in (A) Global and (B) WA *P. nodorum* population.

State-level phylogeography analysis suggests that pathotype migrations may be associated with changes to local climate

To investigate how local climate may influence the distribution of *P. nodorum* isolates and their pathotypes, we performed phylogenetic and population-structure analyses, mating-type profiling, and effector/CSEP profiling. Prediction of the structure of the *P. nodorum* population of Western Australia (WA) formed four distinct groups, where groups 1 and 4 were formed from two highly phylogenetically similar sets of isolates from Geraldton. These groups, with few exceptions, were dominated by a single mating type (although different between groups 1 and 4). In further analysis of the WA population, these two groups were excluded, leaving the main WA population divided into two major groups (2 and 3). Considering only groups 2 and 3, the mating type loci were mixed in approximately equal proportions, indicating potential for sexual recombination across the WA region.

The inferred phylogeny of the WA *P. nodorum* population can be roughly divided into two major clades: upper and lower (Figure 7). The upper clade was dominated by Group 2 isolates, which are diverse in both sampling year and geographic location. The lower clade was mostly composed of Group 3 isolates and was also restricted to very recent sampling years. Despite this overall division, isolates from several geographic locations were interspersed throughout the phylogeny rather than forming single, location-specific clusters, suggesting ongoing gene flow between regions. Phylogenetically similar isolates typically had identical profiles of known effectors and predicted CSEPs (Figure 7); however, diversity in effector profiles did not appear to differ noticeably between the upper and lower groups. Considering the 2-year average rainfall values at the sampling

295 locations, the lower/group 3 subset exhibited a narrower rainfall range than the upper/group 2
296 subset. To further assess whether observed phylogenetic groupings and climate factors affect
297 effector profiles, we performed PERMANOVA on Jensen–Shannon divergences in effector and
298 CSEP protein-isoform compositions across 162 WA isolates with known locations, spanning
299 approximately 50 years and a rainfall gradient of ~577.4 mm. This analysis detected no statistically
300 significant effects of mean annual rainfall ($p = 0.070$) or temporal grouping ($p = 0.181$) on overall
301 effector/CSEP diversity, indicating that broad precipitation or time-based gradients do not explain
302 population-wide patterns of effector composition in this system (Supp Figure 5). However, case-
303 by-case phylogeographic analysis identified several examples of pathotype co-occurrence across
304 different geographic locations and sampling periods (Supp Figure 7).

305

306 The *ad hoc* collection of isolates that comprised the WA *P. nodorum* pangenome (Figure 7)
307 revealed a small number of examples of small groups of isolates that appeared to belong to the
308 same pathotype (i.e. highly similar phylogenetically and identical in effector profile) yet were
309 sampled across different locations and years (note: Northam and South Perth were excluded as
310 these isolates may have been sampled from field trials, with true locations unknown). Isolates
311 DPIRD12 and DPIRD27 from Moora in 1969 were highly similar to isolate DPIRD42 sampled
312 from Badgingarra in 1980. DPIRD14 and DPIRD28 from Welbungin in 1969 were highly similar
313 to DPIRD111 from Badgingarra in 1991. Lastly, DPIRD40 from Badgingarra (1980) was highly
314 similar to WAC13447 from Geraldton 2011 and WAC13418 from Geraldton 1996. These
315 examples represent geographic dispersal from inland to coastal regions (Figure 1).

316

317

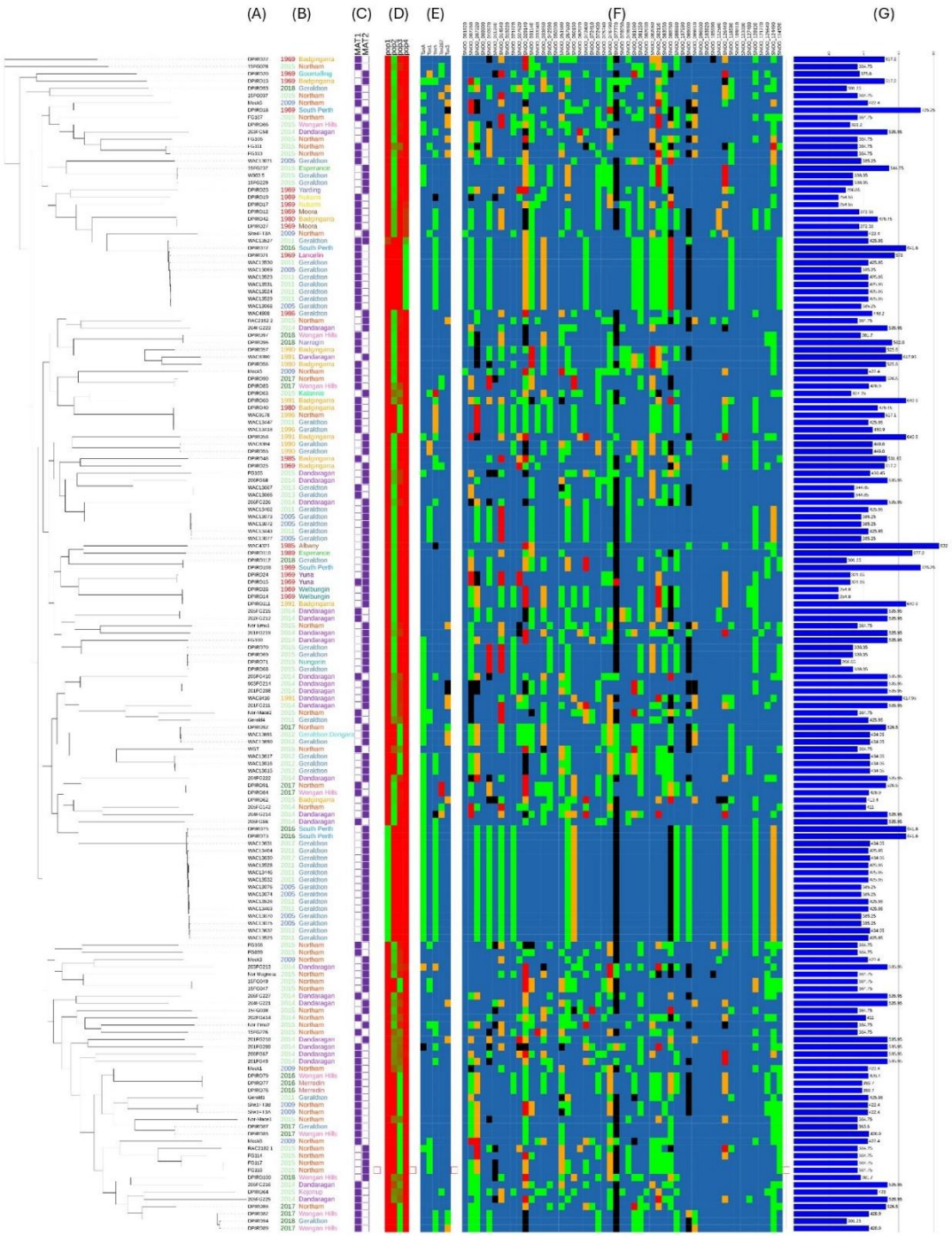

(A)
(B)
(C)
(D)
(E)
(F)
(G)
MAT1
MAT2
pop1
pop2
pop3
pop4

Figure 7. Phylogenetic distribution of WA *P. nodorum* isolates showing (A) phylogenetic tree constructed using IQ-TREE2 from genome-wide high-confidence SNP markers, together with (B) year and location of collection, (C) mating type, (D) population structure, (E) known effectors, (F) top candidate secreted effector proteins (CSEPs), and (G) average annual rainfall (2-year mean, mm) for each region. The primary purpose of the figure is to illustrate the distribution of effector diversity and associated metadata across the phylogeny.

Within the constraints of the limited data available to this study, we also predicted *P. nodorum* pathotypes present in WA based on either a) the profiles of 5 known effectors, or b) the top 20 CSEPs and observed changes to their distribution across time and space (Supplementary Table 3). Considering pathotypes constructed from the 5 known effectors, one frequent pathotype (PT_001) was observed in 13 isolates and was present primarily in Dandaragan from at least 1991 to 2014, was observed in Badgingarra in 1990, and may have moved to Geraldton from 2013 to 2018. A moderately frequent pathotype (PT_003) was detected in 6 isolates, first in Badgingarra in 1980, and then in Geraldton from 2005 to 2011. Some less frequent pathotypes (3 isolates) such as PT_004, PT_006 and PT_007 were first detected inland (Badgingarra and Dandaragan) and later detected in Geraldton. Other pathotypes (e.g. PT_011 and PT_012) were observed in 14 and 5 isolates, respectively, but were only detected in Geraldton from 2005-2012 and 2005-2011. Considering pathotypes constructed from the top 20 CSEPs, most were only detected in Geraldton (2005-2012). PT_001 was detected first in Badgingarra in 1980, and later in Geraldton from 1996-2011, consistent with the pattern observed in the known effector-based pathotype.

Discussion

In this study, we investigated long-term spatial and temporal RIP dynamics across a global *Parastagonospora nodorum* pan-genome of 648 isolates across five major regions - Western Australia (Australia), Europe, South Africa (SA), Iran and the USA. Our analyses revealed that RIP signatures are geographically and temporally dynamic, and enriched at effector-associated loci, with distinct regional patterns pointing to divergent evolutionary histories and selection pressures. Phylogenetic analysis revealed WA isolates were most closely related to South African and some Iranian isolates, while US and European isolates formed largely distinct groups, with US isolates further partitioned into northern (cold) and southern (warm) sub-groups, consistent with Pereira et al. [19]. The similarity of the SA population to some WA isolates may be due to pathogen adaptations to similar latitudes, which may indicate similar climatic conditions, or alternatively may suggest a migration history linked to the export of wheat lines adapted to similar climate zones [39]. Within WA, population structure analysis identified four distinct groups, and two groups comprised highly phylogenetically similar isolates from Geraldton, dominated by a single mating type and restricted to a small geographic area, suggesting predominantly asexual or homothallic reproduction within these lineages, if these isolates were not co-located with isolates possessing alternate mating types. However, the presence of both mating type loci (MAT1-1 and MAT1-2) across all regions suggests widespread potential for sexual recombination [40,41] and ongoing RIP activity [42]. Together, these patterns show how gene flow, founder effects and regional reproductive strategy shape the standing genomic variation on which both climate- and host-driven selection, including cultivar turnover, can act, making *P. nodorum* a tractable system for studying how pathogen genomes record past selective pressures [39].

RIP randomly introduces C→T mutations and has been shown to be the predominant type of sequence mutation (~80%), with the potential to occur at any position across the 23 chromosomes of the *P. nodorum* genome [9]. 'RIP leakage' from interspersed repeats may occur several kilobases away from the RIP-targeted repeat site, although the maximum effective distance of leakage is unknown [9]. In this study, we have used RIPness as a normalised ratio estimating the cumulative effects of RIP across chromosomes - as opposed to quantifying total mutation rate, or mutation density, as these raw frequency metrics may vary considerably across DNA regions with different G:C contents. We assume that RIP and RIP-leakage occur ubiquitously across the genome, but over successive generations, may be enriched or depleted in some genome regions due to the effects of selection, as RIP is likely to generate many non-viable isolates through psuedogenisation [9,14]. Thus, RIPness reflects how much RIP is retained, or allowed to occur, after selection pressures have acted upon a population. High RIPness indicates either low selection pressure or a potential advantage from diversification of a genomic region, whereas low RIPness reflects a high selection pressure to conserve current functions, as would be expected for most genomic regions important for pathogen fitness.

We observed substantial variability in RIP accumulation patterns across the global *P. nodorum* population, with distinct regional signatures emerging at both the whole-genome and chromosomal levels. Specifically, WA and US populations exhibited the highest whole-genome RIPness based on all SNPs, yet showed the lowest proportion of non-synonymous RIPness. These regional differences may indicate that the high whole-genome RIPness of WA and the US is concentrated in non-genic compartments (e.g. TEs, intergenic regions) but leads to protein and functional changes less often. This pattern is consistent with a "two-speed" genome architecture [11,43], in which

RIP is effectively partitioned away from coding sequences, preserving protein function while allowing rapid diversification. Whereas Europe and SA exhibited lower overall RIPness but stronger non-synonymous RIPness, indicating that RIP leakage into coding regions is proportionally greater in these populations, potentially reflecting relatively more relaxed or diversified environments that are more likely to lead to changes in protein function (Figure 3) or a different history of host–pathogen co-evolution alongside locally deployed wheat cultivars [44]. The coevolution of regional *P. nodorum* populations with their respective regional host cultivars reflects the complex history of the diverse routes of host domestication, highlighted by recent wheat genome studies [44,45]**.** The elevated genome-wide RIPness concentrated in non-genic regions of WA and US populations may indicate that these populations are under some form of selection pressure related to crop management in those regions that causes higher diversity to arise without systematically disrupting protein-coding sequences. It is also worth noting that RIP effectiveness is not necessarily a fixed property of a lineage: kingdom-wide phylogenomic analysis shows RIP signatures can vary markedly even between closely related taxa [38] and in *Zymoseptoria tritici* the recent loss of *Dim2* in some lineages has substantially reduced C→T mutation in repeats during mitotic growth [46]. Because classical RIP occurs during the sexual cycle [24,47], whereas DIM-2-associated RIP-like mutation can also accumulate mitotically [46], some of the regional differences in RIPness reported here could therefore reflect variation in RIP input, through differences in the frequency of sexual reproduction [48] or in RIP pathway activity (e.g. allelic variation in *RID* or *DIM-2* homologues; [28,49]), in addition to selection acting on the mutations it generates.

Spatio-temporal RIP analyses revealed that RIP, though widespread throughout the pangenome, is not uniformly distributed but is consistently enriched within specific regions at the population

level. At finer resolution, specific chromosomal regions harbouring known or putative effector loci showed elevated non-synonymous RIPness compared to non-effector loci [9,50], whereas no consistent or pronounced difference was observed in overall SNP-RIPness between effector and non-effector windows. These results suggest that effector regions do not receive more RIP overall, but rather retain a greater proportion of RIP-induced amino-acid changes, consistent with relaxed purifying selection or diversifying selection acting on effector genes [51,52]. Although RIP mutations accumulate across the genome, their disproportionate recurrence within particular chromosomes or chromosomal regions, including effector-enriched regions, is unlikely to reflect stochastic noise alone, and more plausibly reflects the combined effects of genome architecture (e.g. proximity to repeats and heterochromatin) and differential retention under selection [16,20,53-55]. Regionally, this may indicate that some populations experience selection pressures favouring the retention of RIP-induced changes in chromosomes harbouring locally adaptive effectors. This is particularly evident in WA populations, where regions of Chr4 (containing *ToxA*) and Chr10 (containing *Tox1*) showed elevated non-synonymous RIP mutations around their respective effector loci. Elevated non-synonymous RIPness may thus serve as a genomic signature of the relative importance of effector loci between regional populations and may also flag novel genomic regions harbouring unknown effectors for functional characterisation [56].

Beyond these spatial patterns, RIP signatures also show a pronounced temporal dimension that further distinguishes regional populations. Extensive RIP mutations accumulated predominantly in more recent populations rather than being uniformly distributed across the historical dataset of both global and WA scales, though the timing and magnitude of the increase differ between the regions. Owing to the large number of isolates from the USA and WA, these regions are likely to

436 contribute most to the observed global temporal trend. However, the enrichment of high-RIP
437 windows in recent decades coincides with periods of cultivar turnover and changing host-
438 associated selection in WA *P. nodorum* populations [57], although a direct link between these factors
439 and RIP accumulation remains untested. At the chromosome level, RIPness is observed to increase
440 or decrease over time differentially across chromosomes, and it is important to distinguish for
441 different chromosomes whether they possess known effectors or other biological roles. The long-
442 term RIP signals on Chr4 (*ToxA*) and Chr11 (*Tox3*) reflect ancient RIP activity with elevated C→T
443 rates; balancing selection likely maintains functional effector polymorphisms despite ongoing RIP
444 leakage [9]. The contrasting pattern observed for Chr11—decreasing in the global dataset but
445 increasing in Western Australian (WA) populations—may reflect region-specific selection acting
446 on or near the SnTox3 locus, consistent with evidence for geographically structured selection and
447 population-specific diversification of effector-associated genomic regions in *P. nodorum* [33]. In
448 contrast, the RIP signatures of Chr10 and Chr12 being largely confined to recent decades,
449 alongside the predominantly recent activity of most non-effector chromosomes, suggest transient
450 mutational processes rather than sustained adaptive evolution for these chromosomes. To assess
451 whether these genomic patterns are influenced by environmental factors, we examined the
452 geographic and climatic context of the WA isolates.
453
454 The WA wheatbelt has a Mediterranean climate with hot, dry summers and cool, wet winters,
455 although the locations sampled in this study range from coastal sites with higher rainfall and milder
456 temperatures to inland sites with hotter, more arid summers. *P. nodorum* sporulates in spring and
457 early summer [58], when easterly to north-easterly winds prevail across most of WA, countered in
458 coastal areas by afternoon southerly to south-westerly winds [59]. The region's climate has also

experienced recent and dramatic changes. All sampled locations, past and present isolates (e.g. Badgingarra, Dandaragan, Geraldton, Northam and Esperance; Supp Figure 7) have been drier and hotter since 2000, consistent with DPIRD analyses showing that drier and hotter rainfall zones shifted westward to south-westward between 2000–2024 relative to 1910–1999 [60]. Climate shifts can drive the migration or emergence of pathogens and their more aggressive strains [61,62], but this has not yet been examined for *P. nodorum*. Recent work in other fungal pathosystems suggests that precipitation and humidity, rather than temperature alone, often drive climate-linked adaptive divergence, with populations facing the largest projected moisture deficits showing the greatest genomic offset, that is, the largest mismatch between current genotypes and those predicted to suit future conditions [63].

When we tested whether rainfall gradients and sampling period explained effector-profile diversity across WA isolates, we found no significant relationship at the population level or among temporal groupings. However, case-by-case phylogeographic analysis identified three instances of possible climate-associated pathotype movement (Supp Figure 7). Two were westward (Moora-to-Badgingarra and Welbungin-to-Badgingarra) and one north-westward (Badgingarra-to-Geraldton) (Figure 1), each broadly consistent with prevailing spring winds and the direction of recent rainfall-zone shifts [59,60]. Such movement could reflect direct climatic adaptation, or indirect selection through cultivars deployed in response to changing local climate. However, these examples should be interpreted with caution, as only a few sites were sampled across multiple time periods; therefore, more extensive spatio-temporal sampling would be needed to confirm these movements.

The Badgingarra-to-Geraldton case is the most informative, because both locations were sampled before and after 2000. Genetic diversity was higher at Badgingarra, which lies within a wheatbelt rainfall-transition zone (Supp Figure 6), whereas the lower diversity at Geraldton may reflect localised clonal expansion or dominance by a few lineages. Post-2000 Badgingarra isolates are increasingly clustered within Geraldton-associated clades. Pathotype profiles based on both known effectors and top-ranked CSEPs also showed multiple pathotypes detected first at Badgingarra and subsequently at Geraldton. This directional shift is consistent with the prevailing southerly to south-easterly coastal winds, which blow north-northwest along the coast during the sporulation season[59] (Figure 1). However, before 2000, although spores may have been dispersed from Badgingarra to Geraldton, the larger climatic differences between the two sites may have prevented Badgingarra-derived pathotypes from establishing. After 2000, rainfall declined much more at Badgingarra (~−130 mm) than at Geraldton (~−50 mm), bringing conditions at the two sites closer together and potentially favouring establishment. In genomic-offset terms, Badgingarra's steeper rainfall decline implies a greater mismatch between its historically adapted pathotypes and its new climate envelope, consistent with genotypes moving from the more climate-mismatched site towards the comparatively stable one [63], although genomic offset was not estimated directly here. Although not conclusive, this trend is suggestive of drought-associated selection of either pathogens or host cultivars, with directional gene flow of pathotypes along climatic gradients, consistent with prevailing southerly winds in the Geraldton region during spring and early summer [62]. Detecting such shifts reliably will require regular, spatially dense sampling, as has been used to resolve population structure and inter-regional migration in wheat stripe rust [64].

Taken together with temporal RIP dynamics, these results suggest that climate change, cultivar deployment, and effector evolution are coupled components of pathogen adaptation [62]. Under elevated RIP activities, effector loci can diversify rapidly and generate virulence gains on resistant cultivars, as shown in *Z. tritici* [65,66]. In *L. maculans*, RIP-rich, TE-associated regions similarly promote high variation in effector genes, enabling both the emergence of new virulence alleles and the loss of avirulence alleles that have become detrimental under resistance gene selection [20]. Our *P. nodorum* pan-genome shows comparable patterns, with effector neighbourhoods retaining a greater proportion of RIP-induced amino-acid changes, suggesting potential for resistance breakdown when particular cultivars are used over long periods [67], and distinct RIP profiles of regional populations likely reflect differing histories of co-evolution with regionally deployed cultivars [44]. The increasing number of RIP signatures in WA populations over time is of particular interest. If they reflect elevated mutational input rather than retention alone, they could accelerate effector diversification and adaptive responses to resistant cultivars, heightening the risk of resistance erosion and outbreak emergence (Figure 8). This would be consistent with the broader idea that regulated, elevated mutagenesis ("hypermutation") can itself be an adaptive feature of fungal pathogens under selection, rather than simply an incidental by-product of genome defence [68]. In this context, resistance breakdown may reflect structured evolutionary dynamics shaped by interacting climatic and genomic forces, rather than purely stochastic events, as reported for climate- and management-driven structuring in other plant pathosystems [69].

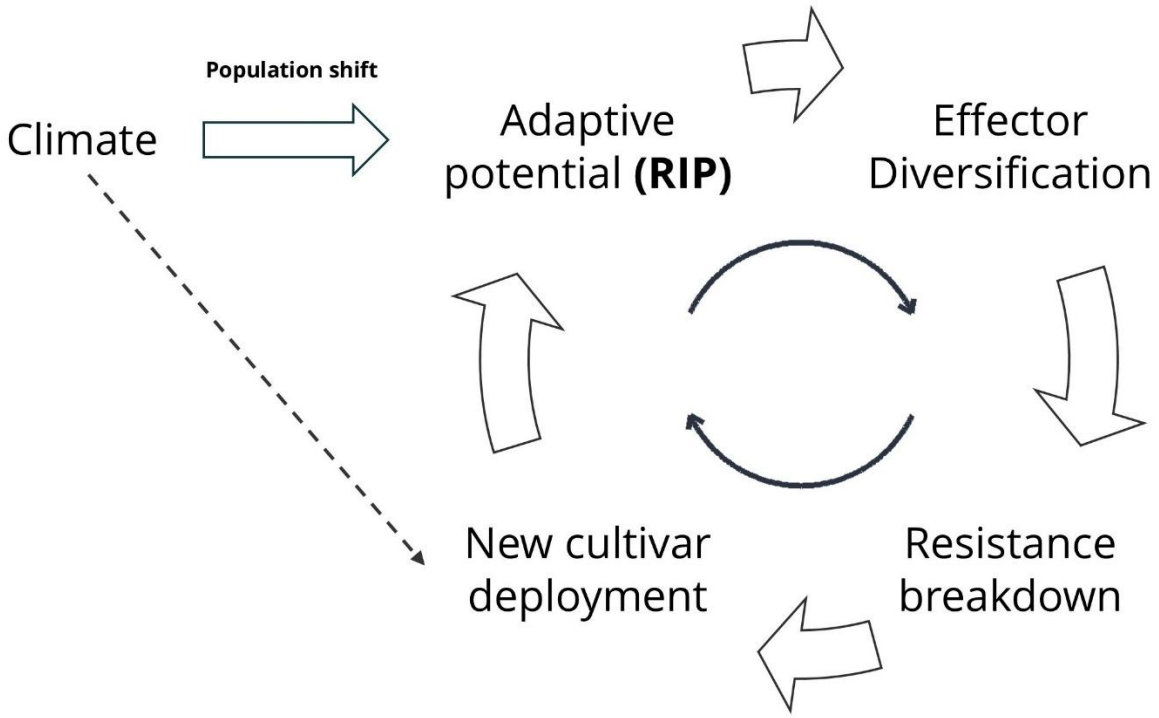


Figure 8: Conceptual framework of RIP-mediated adaptive dynamics linking climate-driven population shifts, effector diversification and cultivar deployment. Climate may act directly on pathogen populations through population shifts (solid arrow) and indirectly by influencing cultivar deployment (dashed arrow). RIP-associated genomic diversity supports effector diversification, which can lead to resistance breakdown; the deployment of new cultivars then imposes renewed selection, completing the cycle.

Conclusions

This global analysis shows that RIP signatures in *P. nodorum* are not uniformly distributed but are spatially and temporally structured, varying among regions, chromosomes and sampling decades. Although effector-associated regions did not accumulate more RIP overall, they retained a greater proportion of RIP-induced amino-acid changes, most clearly around *ToxA* and *Tox1* in WA, consistent with region-specific selection acting on effector loci. RIP signatures have increased in

recent decades at both global and WA scales, a trend that may coincide with a changing climate and shifting cultivar deployment. In Western Australia, we find preliminary phylogeographic evidence, most clearly the Badgingarra-to-Geraldton lineage shift, consistent with climate-associated pathotype movement along a rainfall gradient. Confirming a causal role for climate will require denser, more balanced spatio-temporal sampling than was available here. Nevertheless, our results suggest that tracking RIP signatures at effector loci, alongside climate and cultivar-deployment data, could provide an early indicator of emerging pathogen variants and help anticipate resistance breakdown before it occurs, rather than responding to it after the fact.

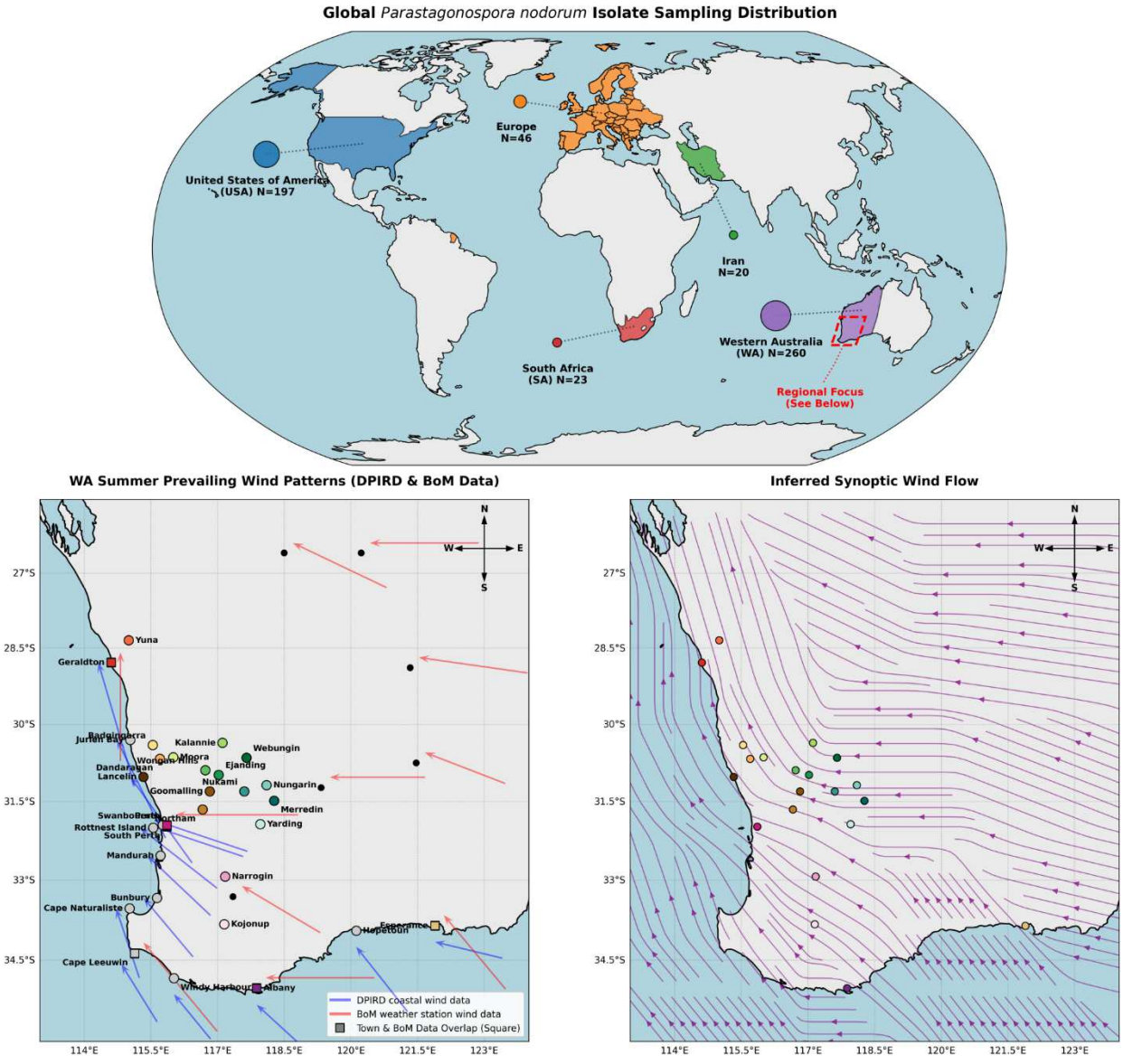
Methods

Sampling, gathering and generation of pangenome data for a global population of *Parastagonospora nodorum*

This study analyzed 649 *Parastagonospora nodorum* isolates, comprising 269 from the Western Australian (WA) wheat-belt and 380 international isolates. (i) WA isolates: Of the 269 WA isolates, sequences from 156 previously published *P. nodorum* isolates were obtained from NCBI BioProject PRJNA398070. These were sequenced at the Australian Genome Research Facility (Melbourne, Australia) using Illumina HiSeq2500 with TruSeq PCR-free libraries, generating 125 bp paired-end reads and 600 bp insert sizes [9]. An additional 104 newly sequenced WA isolates were generated using Illumina HiSeq2500 with TruSeq PCR-free libraries (150 bp paired-end reads, 350 bp insert size). For WA-specific analyses, only isolates with confirmed WA locations

and collection years were included. (ii) International isolates: 380 isolates originated from Switzerland (CH, n=43), South Africa (SA, n=22), Iran (IR, n=32), USA (US, n=261), and other Australian isolates (n=22). All 197 US isolates had paired-end sequencing (100–150 bp read length). Sequencing data were accessed from NCBI BioProject PRJNA398070, PRJNA606320, PRJNA476481 and PRJNA1130627.

The seqeunce reads of all isolates were trimmed using CutAdapt v2.1 [70] to remove TruSeq universal adapters with parameters '-q30' and '--minimum-length=50'. The specified adapters trimmed were *'AAGTCGGAGGCCAAGCGGTCTTAGGAAGACAA'* and *'AAGTCGGATCGTAGCCATGTCGTTCTGTGAGCCAAGGAGTTG'*. Quality control statistics were generated using FastQC v0.11.8 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and summarised using MultiQC v1.14 [71]. The trimmed reads were merged and gap filled by bbmerge v38.90 [72] included in BBTools suite. Potential sample contaminants were searched against a database (PlusPFP) constructed from all NCBI Refeq standard (bacterial, viral, human, plasmid) plus Refeq protozoa, fungi and plant using Kraken2 v-2.1.2 [73].



578

579

580 **Figure 1. Geographic distribution of *Parastagonospora nodorum* isolates and sampling**
581 **locations.** Global sampling distribution of *P. nodorum* isolates is shown on the left, with circle
582 size proportional to the number of isolates collected from each region: Western Australia, USA,
583 Europe, South Africa, and Iran. The right panel presents a detailed map of south-western Western

584 Australia showing individual sampling locations with names and corresponding isolate counts (N),
585 along with spring morning wind roses [59].
586
587 Whole-genomes of newly sampled isolates (excluding reference isolates) were assembled *de novo*
588 with SPAdes v3.12.0 [74] (--only-assembler, --cov-cutoff auto) and MitoZ v3.6 [75] (all, --
589 genetic_code 4, --clade Chordata, --thread_number 8, --assembler spades, --requiring_taxa
590 Chordata, with --insert_size and --fastq_read_length adjusted according to the input data).
591 Genome assemblies generated with SPAdes were filtered to remove sequences matching MitoZ
592 mtDNA contigs using BLASTn [76] (≥95% mtDNA coverage). Filtered SPAdes and MitoZ
593 sequences were combined into a final assembly and annotated as "nuclear" or "mtDNA"
594 respectively. Quast v5.0.2 [77] and OcculterCut v1.1 were used to assess genome assembly quality
595 and find AT-rich regions in fungal genomes [14](Supplementary table 1A).
596
597 **Identifying genome mutations across the global *P. nodorum* pangenome**
598
599 Cleaned sequence reads were aligned to the *P. nodorum* Sn15 reference genome [NCBI
600 Assembly: GCA_016801405.1] [10] using BWA-MEM v0.7.17 [78] with default parameters. GATK
601 v4.2.6.1 was used for short variant calling (SNPs and indels)[79]. HaplotypeCaller was first run in
602 GVCF mode for the reads of each isolate independently, after which GenotypeGVCFs was used
603 to perform joint genotyping of all isolates. Raw SNPs and indels were hard-filtered separately
604 using the VariantFiltration tool with the following parameters: SNPs: QD < 2, QUAL < 30, SOR
605 > 3, FS > 60, MQ < 40, MQRankSum < -12.5, ReadPosRankSum < -8 and Indels: QD < 2, QUAL
606 < 30, FS > 200, ReadPosRankSum < -20 [53,80]. Additional filtering for bi-allelic SNPs was

performed to remove sites with >5% missing data per site, LD >0.9, and retaining 1 high-quality SNP per 5 kb window. Filtered variants were annotated using SNPEff v5.0 [81], versus a recent re-annotation of the *P. nodorum* Sn15 reference genome from [10].

Repeat-induced point (RIP)-like mutations were quantified using a new genome-wide metric called the "RIPness index", which represents the normalised proportion of RIP-like transitions within fixed genomic windows. This allowed the relative impact of RIP mutation to be compared between sequences with variable G:C content, e.g. G:C-equilibrated regions could be roughly compared to AT-rich regions. 'RIPness' was calculated by quantifying the normalised frequency of C→T (and G→A) transition mutations within non-overlapping 10 kb windows relative to other C→N mutations**:** i.e. RIPness = (C→T + G→A) / (C→N + G→N)**.** Although C→T and G→A transitions can arise through multiple mutational processes, they represent the predominant and well-established mutational signatures associated with RIP in fungi. Therefore, the RIPness index was designed to measure the relative abundance of RIP-like signatures across genomic regions rather than to attribute every individual transition event exclusively to RIP activity. C/T and G/A variable sites were determined from GVCF data across all isolates in the pangenome (not reference-based), and were only counted in the RIPness numerator (C→T + G→A) if they contained a base that was the potential result of RIP mutation i.e. T or A bases. Windows lacking C/T or G/A variable sites were excluded, as they would have no detectable RIP-associated transitions. The RIPness index was calculated both genome-wide and for specific genomic contexts (i.e., regions containing effector genes) using all SNPs or non-synonymous SNPs. While RIPness provides a comparative measure of RIP-like mutation signatures, it cannot distinguish individual RIP events from other processes that may generate C→T/G→A transitions.

Determining the Phylogenetic and Population Structure of the *P. nodorum* global pangenome:

Filtered SNPs were converted from VCF format into the NEXUS using 'vcf2phylip' v2.8 [82] to construct a phylogenetic tree using IQTree2 v2.1.2 [83]. IQ-Tree2 was used to automatically test substitution models (-m TEST+ASC) for the SNP alignment (-st DNA). The final tree was constructed using the best-fitting model selected by IQ-Tree2, with 1000 ultrafast bootstrap iterations (-bb 1000) and 1000 replicates for the approximate likelihood ratio test (-alrt 1000). Phylogenetic trees were visualised along with other genome-derived metadata (see below) via iTOL v5 [84].

FastSTRUCTURE version [85], which is based on a Bayesian clustering approach, was run on the entire SNPs alignment to infer population structure. Before running fastSTRUCTURE, VCF files were converted into the PLINK bed format using PLINK v1.90b7. For k = 1 to 12, twelve independent runs were performed with default parameters. The fastSTRUCTURE function "chooseK.py" was used to determine the optimal number of subpopulations that maximise the marginal likelihood. After determining the best value of K, the population structure of *P. nodorum* was inferred using FastSTRUCTURE. The population structure was visualised using the distruct.py script from FastSTRUCTURE [85].

Prediction of global pan-genome gene content and candidate secreted effector-like proteins (CSEPs):

We employed the reference-guided gene prediction software MetaEuk v0.021 [86] to identify putative genes in our *P. nodorum* isolates. A representative gene ortholog dataset, termed "SNOO-set," was previously generated from *P. nodorum* pangenome data from a small set of international isolates and extensive isolates from the Western Australian (WA) region [9]. This was used as a representative proteome reference for input to MetaEuk, to predict gene content in all isolates across the global pangenome. We confirmed the applicability of the WA-based SNOO-set to the new set of global *P. nodorum* pangenome data by constructing a rarefaction curve (Supplementary Fig. 1). This rarefaction curve demonstrated that the number of genes captured by the SNOO-set plateaued rapidly after all WA isolates were considered, which further flattened with the addition of European (EU), Iranian and US isolates, indicating that representation of the global pan-proteome by this dataset was near complete and validating its use in our study. Additionally, we utilised MetaEuk separately to screen the population for datasets of specialised proteins, including mating types and known effector genes.

The Predector v1.2.7 [87] pipeline was used to predict functional annotations and candidate secreted effector-like proteins (CSEPs), processing the translated gene annotation outputs from MetaEuk versus the representative ortholog sequences of the SNOO-set. Predector combines the outputs of multiple tools including SignalP (different versions) [88,89], DeepLoc version 1.0 [90], TargetP version 2.0 [91], Phobius version 1.01 [92], TMHMM version 2.0c [93,94], EffectorP versions 1, 2 and 3 [95,96], ApoplastP version 1.0.1 [97], LOCALIZER version [98], HMMER version 3.3 [99] which was used for homology searches against the dbCAN version 8 database [100], and MMSeqs2 version 11.e1a1c [101] to perform sequence matches against PHI-base version 4.9 [102]. The Predector pipeline generated a

summarised list of ranked effector candidates with a range of effector properties and their homologies to various databases, including Pfam [103], SwissProt [104], InterPro [105], dbCAN [106], and PHIbase [102], as well as homology to the Predector dataset of known effectors [87]. CSEPs were predicted by the following criteria: a Predector score of ≥1, the presence of ≥2 cysteines, and a total residue count of <300 (Supplementary Table 2). The selected criteria were based on the common characteristics of known fungal CSEPs, which are typically small, cysteine-rich secreted proteins with positive effector prediction scores.

A region-focussed analysis was performed for isolates sampled from WA, where the protein isoform profiles of known effectors were analysed to test for rainfall and year-based structuring using PERMANOVA on Jensen-Shannon divergence (JSD) distances [107,108]. The profile of each isolate comprised proportions of 118 top ranked protein families that were classified into six categories: none (no variants), var1 (reference only), var2/var3/var4plus (single derived variants), or multiple (highest diversity). Pairwise JSD matrices quantified compositional differences, with PERMANOVA assessing rainfall (continuous/categorical via tertiles: low/medium/high) and collection year effects (1969–2018, *n*=164 isolates). All analyses were implemented in Python 3.8 or later using scikit-learn (v1.4.0)(https://scikit-learn.org/stable/), scikit-optimize (v0.9.0)(https://scikit-optimize.github.io/stable/), NumPy (v1.19.0) [109], and pandas (v1.3.0)(http://pandas.pydata.org/pandas-docs/stable/whatsnew/v3.0.1.html). Key functions included jensenshannondivergence, createvariantcompositionprofiles, and permanovatest. Custom Python implementations partitioned JSD matrix variance via permutational ANOVA, yielding pseudo-F statistics, R^2 (variance explained), and p-values without normality assumptions. Models

tested included: (1) rainfall continuous, (2) rainfall categorical, (3) year continuous, and (4) year decades; principal coordinates analysis (PCoA) visualized ordinations.

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and materials

All sequencing data used in this study are publicly available from the National Center for Biotechnology Information (NCBI) under the BioProject accessions PRJNA398070, PRJNA606320, PRJNA476481 and PRJNA1130627. All processed datasets and associated metadata supporting the findings of this study are available in the Supplementary Information and/or via Figshare at https://doi.org/10.6084/m9.figshare.34065522.

Competing interests

Not applicable

Funding

This project was supported by the Grains Research and Development Corporation (GRDC) and Curtin University as part of the co-investment in the Centre for Crop and Disease Management (CCDM) phase III project: "Advanced Bioinformatics Approaches" [CUR00023]. M. Hossain and N. Gray were supported by Research Training Program (RTP) grants provided by the Australian Government Department of Education. Additional support for M. Hossain was provided by a GRDC Research Scholarship [CUR2301-006RSX], with support from the WA Agricultural Research Collaboration (WAARC).

Authors' contribution

MH and JH led the conceptual development and design the study. DJ, NG, AW and H.P. generated some of the data used in this chapter. MH performed all analyses and MH and JH generated all figures and tables. MH, HP and JH wrote and edited the manuscript. All authors provided comments and suggestions and approved the manuscript.

Acknowledgements

Not applicable

References

1. Eyal, Z., Center, I.M. & Wheat, I. *The Septoria Diseases of Wheat: Concepts and Methods of Disease Management*, 56 (CIMMYT, 1987).

2. Solomon, P.S., Lowe, R.G.T., Tan, K.-C., Waters, O.D.C. & Oliver, R.P. Stagonospora nodorum: cause of stagonospora nodorum blotch of wheat. *Molecular Plant Pathology* **7**, 147–156 (2006).
3. Bhathal, J.S., Loughman, R. & Speijers, J. Yield Reduction in Wheat in Relation to Leaf Disease From Yellow (tan) Spot and Septoria Nodorum Blotch. *European Journal of Plant Pathology* **109**, 435–443 (2003).
4. Brennan, J.P. & Murray, G.M. Economic importance of wheat diseases in Australia. *Economic importance of wheat diseases in Australia* (1998).
5. Friesen, T.L., Faris, J.D., Solomon, P.S. & Oliver, R.P. Host-specific toxins: effectors of necrotrophic pathogenicity. *Cell Microbiol* **10**, 1421–8 (2008).
6. Friesen, T.L., Meinhardt Sw Fau - Faris, J.D. & Faris, J.D. The Stagonospora nodorum-wheat pathosystem involves multiple proteinaceous host-selective toxins and corresponding host sensitivity genes that interact in an inverse gene-for-gene manner. *The Plant journal* **51**, 681–692 (2007).
7. Phan, H.T. *et al.* Differential effector gene expression underpins epistasis in a plant fungal disease. *Plant J* **87**, 343–54 (2016).
8. Downie, R.C. *et al.* Septoria Nodorum Blotch of Wheat: Disease Management and Resistance Breeding in the Face of Shifting Disease Dynamics and a Changing Environment. *Phytopathology*® **111**, 906–920 (2021).
9. Jones, D.A.B. *et al.* Repeat-induced point mutations driving Parastagonospora nodorum genomic diversity are balanced by selection against non-synonymous mutations. *Communications Biology* **7**, 1614 (2024).
10. Bertazzoni, S., Jones, D.A.B., Phan, H.T., Tan, K.-C. & Hane, J.K. Chromosome-level genome assembly and manually-curated proteome of model necrotroph Parastagonospora nodorum Sn15 reveals a genome-wide trove of candidate effector homologs, and redundancy of virulence-related functions within an accessory chromosome. *BMC Genomics* **22**, 382 (2021).
11. Syme, R.A. *et al.* Pan-Parastagonospora Comparative Genome Analysis—Effector Prediction and Genome Evolution. *Genome Biology and Evolution* **10**, 2443–2457 (2018).
12. Almojil, D. *et al.* The Structural, Functional and Evolutionary Impact of Transposable Elements in Eukaryotes. *Genes* **12**, 918 (2021).
13. Castanera, R. *et al.* Transposable Elements versus the Fungal Genome: Impact on Whole-Genome Architecture and Transcriptional Profiles. *PLoS Genetics* **12**, e1006108 (2016).
14. Testa, A.C., Oliver, R.P. & Hane, J.K. OcculterCut: A Comprehensive Survey of AT-Rich Regions in Fungal Genomes. *Genome Biology and Evolution* **8**, 2044–2064 (2016).
15. Frantzeskakis, L. *et al.* Signatures of host specialization and a recent transposable element burst in the dynamic one-speed genome of the fungal barley powdery mildew pathogen. *BMC Genomics* **19**, 381 (2018).
16. Komluski, J., Habig, M. & Stukenbrock, E.H. Repeat-Induced Point Mutation and Gene Conversion Coinciding with Heterochromatin Shape the Genome of a Plant-Pathogenic Fungus. *mBio* **14**, e0329022 (2023).
17. Razali, N.M., Cheah, B.H. & Nadarajah, K. Transposable Elements Adaptive Role in Genome Plasticity, Pathogenicity and Evolution in Fungal Phytopathogens. *International Journal of Molecular Sciences* **20**(2019).
18. Charlesworth, B. & Langley, C.H. THE POPULATION GENETICS OF DROSOPHILA TRANSPOSABLE ELEMENTS. *Annual Review of Genetics* **23**, 251–287 (1989).
19. Pereira, D., Oggenfuss, U., McDonald, B.A. & Croll, D. Population genomics of transposable element activation in the highly repressive genome of an agricultural pathogen. *Microbial Genomics* **7**, 000540 (2021).

20. Rouxel, T. *et al.* Effector diversification within compartments of the Leptosphaeria maculans genome affected by Repeat-Induced Point mutations. *Nature communications* **2**(2011).
21. Grandaubert, J. *et al.* Transposable element-assisted evolution and adaptation to host plant within the Leptosphaeria maculans-Leptosphaeria biglobosa species complex of fungal pathogens. *BMC Genomics* **15**(2014).
22. Hane, J.K. & Oliver, R.P. RIPCAL: a tool for alignment-based analysis of repeat-induced point mutations in fungal genomic sequences. *BMC Bioinformatics* **9**, 478 (2008).
23. Fudal, I. *et al.* Repeat-Induced Point Mutation (RIP) as an Alternative Mechanism of Evolution Toward Virulence in Leptosphaeria maculans. *Molecular Plant-Microbe Interactions®* **22**, 932–941 (2009).
24. Hane, J.K., Williams, A.H., Taranto, A.P., Solomon, P.S. & Oliver, R.P. Repeat-Induced Point Mutation: A Fungal-Specific, Endogenous Mutagenesis Process. in *Genetic Transformation Systems in Fungi, Volume 2* (eds. Van Den Berg, M.A. & Maruthachalam, K.) 55–68 (Springer International Publishing, Cham, 2015).
25. Wang, L. *et al.* Repeat-induced point mutation in Neurospora crassa causes the highest known mutation rate and mutational burden of any cellular life. *Genome biology* **21**, 142 (2020).
26. Gladyshev, E. & Kleckner, N. DNA sequence homology induces cytosine-to-thymine mutation by a heterochromatin-related pathway in Neurospora. *Nature Genetics* **49**, 887–894 (2017).
27. Kouzminova, E. & Selker, E.U. dim-2 encodes a DNA methyltransferase responsible for all known cytosine methylation in Neurospora. *The EMBO Journal* **20**, 4309–4323 (2001).
28. Freitag, M., Williams Rl Fau - Kothe, G.O., Kothe Go Fau - Selker, E.U. & Selker, E.U. A cytosine methyltransferase homologue is essential for repeat-induced point mutation in Neurospora crassa. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 8802–8807 (2002).
29. Van de Wouw, A.P. *et al.* Evolution of linked avirulence effectors in Leptosphaeria maculans is affected by genomic environment and exposure to resistance genes in host plants. *PLoS Pathog* **6**, e1001180 (2010).
30. Van de Wouw, A.P., Elliott, C.E., Popa, K.M. & Idnurm, A. Analysis of Repeat Induced Point (RIP) Mutations in Leptosphaeria maculans Indicates Variability in the RIP Process Between Fungal Species. *Genetics* **211**, 89–104 (2019).
31. Malagnac, F., Gladyshev, E., Croll, D. & Badet, T. The fungal RIP hypermutator mechanism has deep eukaryotic roots. *Trends Genet* **42**, 793–796 (2026).
32. Hane, J.K. *et al.* Dothideomycete plant interactions illuminated by genome sequencing and EST analysis of the wheat pathogen Stagonospora nodorum. *Plant Cell* **19**, 3347–68 (2007).
33. Richards, J.K. *et al.* Local adaptation drives the diversification of effectors in the fungal wheat pathogen Parastagonospora nodorum in the United States. *PLOS Genetics* **15**, e1008223 (2019).
34. Ghaderi, F., Sharifnabi, B., Javan-Nikkhah, M., Brunner, P.C. & McDonald, B.A. SnToxA, SnTox1, and SnTox3 originated in Parastagonospora nodorum in the Fertile Crescent. *Plant Pathology* **69**, 1482–1491 (2020).
35. Choupannejad, R. *et al.* Interaction of wheat-Parastagonospora nodorum isolates: Exploring host susceptibility and fungal virulence. *Physiological and Molecular Plant Pathology* **130**, 102235 (2024).
36. Hafez, M. *et al.* Parastagonospora nodorum and Related Species in Western Canada: Genetic Variability and Effector Genes. *Phytopathology®* **110**, 1946–1958 (2020).

37. Trainor, G., Zaicou-Kunesch, C., Curry, J., Shackley, B. & Nicol, D. *2019 Wheat Variety Sowing Guide for Western Australia*, 44 (Department of Primary Industries and Regional Development, 2018).
38. Badet, T. & Croll, D. Phylogenomic signatures of repeat-induced point mutations across the fungal kingdom. *PLOS Biology* **23**, e3003433 (2025).
39. Feurtey, A. *et al.* A thousand-genome panel retraces the global spread and adaptation of a major fungal crop pathogen. *Nat Commun* **14**, 1059 (2023).
40. Milgroom, M.G. Recombination and the Multilocus Structure of Fungal Populations. *Annual Review of Phytopathology* **34**, 457–477 (1996).
41. Sun, S. *et al.* Fungal sexual reproduction and mating-type loci. *Curr Biol* **35**, R496–R503 (2025).
42. Graïa, F. *et al.* Genome quality control: RIP (repeat-induced point mutation) comes to Podospora. *Molecular Microbiology* **40**, 586–595 (2001).
43. Torres, D.E., Oggenfuss, U., Croll, D. & Seidl, M.F. Genome evolution in fungal plant pathogens: looking beyond the two-speed genome model. *Fungal Biology Reviews* **34**, 136–143 (2020).
44. White, J. *et al.* The genetic diversity of UK, US and Australian cultivars of Triticum aestivum measured by DArT markers and considered by genome. *Theor Appl Genet* **116**, 439–53 (2008).
45. Jiao, C. *et al.* Pan-genome bridges wheat structural variations with habitat and breeding. *Nature* **637**, 384–393 (2025).
46. Moller, M. *et al.* Recent loss of the Dim2 DNA methyltransferase decreases mutation rate in repeats and changes evolutionary trajectory in a fungal pathogen. *PLoS Genet* **17**, e1009448 (2021).
47. Selker, E.U. Premeiotic instability of repeated sequences in Neurospora crassa. *Annu Rev Genet* **24**, 579–613 (1990).
48. Lin, M., Ficke, A., Cockram, J. & Lillemo, M. Genetic Structure of the Norwegian Parastagonospora nodorum Population. *Frontiers in Microbiology* **Volume 11 - 2020**(2020).
49. Gladyshev, E. & Kleckner, N. DNA sequence homology induces cytosine-to-thymine mutation by a heterochromatin-related pathway in Neurospora. *Nat Genet* **49**, 887–894 (2017).
50. van Wyk, S. *et al.* Repeat-Induced Point Mutations Drive Divergence between Fusarium circinatum and Its Close Relatives. *Pathogens* **8**, 298 (2019).
51. Stukenbrock, E.H. *et al.* The making of a new pathogen: insights from comparative population genomics of the domesticated wheat pathogen Mycosphaerella graminicola and its wild sister species. *Genome Res* **21**, 2157–66 (2011).
52. Brunner, P.C., Torriani, S.F.F., Croll, D., Stukenbrock, E.H. & McDonald, B.A. Coevolution and Life Cycle Specialization of Plant Cell Wall Degrading Enzymes in a Hemibiotrophic Pathogen. *Molecular Biology and Evolution* **30**, 1337–1347 (2013).
53. Jones, D.A.B. *et al.* Repeat-induced point mutations driving Parastagonospora nodorum genomic diversity are balanced by selection against non-synonymous mutations. *Commun Biol* **7**, 1614 (2024).
54. Syme, R.A. *et al.* Pan-Parastagonospora Comparative Genome Analysis-Effector Prediction and Genome Evolution. *Genome Biol Evol* **10**, 2443–2457 (2018).
55. Oggenfuss, U. & Croll, D. Recent transposable element bursts are associated with the proximity to genes in a fungal plant pathogen. *PLoS Pathog* **19**, e1011130 (2023).
56. Lorrain, C., Feurtey, A., Moller, M., Haueisen, J. & Stukenbrock, E. Dynamics of transposable elements in recently diverged fungal pathogens: lineage-specific transposable element content and efficiency of genome defenses. *G3 (Bethesda)* **11**, jkab068 (2021).

57. Phan, H.T.T. *et al.* Low Amplitude Boom-and-Bust Cycles Define the Septoria Nodorum Blotch Interaction. *Front Plant Sci* **10**, 1785 (2019).
58. Bathgate, J.A. & Loughman, R. Ascospores are a source of inoculum of Phaeosphaeria nodorum, P. avenaria f. sp. avenaria and Mycosphaerella graminicola in Western Australia. *Australasian Plant Pathology* **30**, 317–322 (2001).
59. Australian Government Bureau of Meteorology. Wind maps for Australia. Australian Government Bureau of Meteorology (BoM). (2011).
60. DPIRD. South-west Western Australia average annual rainfall zones 2000–2024 compared to 1910–1999. Department of Primary Industries and Regional Development (DPIRD), Western Australia. (2025).
61. Lahlali, R. *et al.* Effects of climate change on plant pathogens and host-pathogen interactions. *Crop and Environment* **3**, 159–170 (2024).
62. Singh, B.K. *et al.* Climate change impacts on plant pathogens, food security and paths forward. *Nature Reviews Microbiology* **21**, 640–656 (2023).
63. Hessenauer, P., Feau, N., Heinzelmann, R. & Hamelin, R.C. Genomic Exploration of Climate-driven Evolution and Evolutionary Convergence in Forest Pathogens. *Genome Biology and Evolution* **17**, evaf069 (2025).
64. Li, Y. *et al.* Genomic analysis, trajectory tracking, and field surveys reveal sources and long-distance dispersal routes of wheat stripe rust pathogen in China. *Plant Commun* **4**, 100563 (2023).
65. Sampaio, A.M. *et al.* Diversification, loss, and virulence gains of the major effector AvrStb6 during continental spread of the wheat pathogen Zymoseptoria tritici. *PLoS Pathog* **21**, e1012983 (2025).
66. Sampaio, A.M. *et al.* Diversification, loss, and virulence gains of the major effector AvrStb6 during continental spread of the wheat pathogen Zymoseptoria tritici.
67. Wilson, H. *et al.* Integrated Disease Management, Adaptation and Genomics of Fungal Plant Pathogens in Cropping Systems. *Plant Pathology* **74**, 1445–1469 (2025).
68. Baril, T. & Croll, D. Hypermutator fungal pathogens: from threat to meltdown. *Trends Microbiol* **34**, 42–49 (2026).
69. Ostos, E. *et al.* Effect of Cultivar Resistance and Soil Management on Spatial-Temporal Development of Verticillium Wilt of Olive: A Long-Term Study. *Front Plant Sci* **11**, 584496 (2020).
70. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**, 10–12 (2011).
71. Ewels, P., Magnusson, M., Lundin, S. & Kaller, M. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **32**, 3047–8 (2016).
72. Bushnell, B., Rood, J. & Singer, E. BBMerge – Accurate paired shotgun read merging via overlap. *PLOS ONE* **12**, e0185056 (2017).
73. Wood, D.E., Lu, J. & Langmead, B. Improved metagenomic analysis with Kraken 2. *Genome Biol* **20**, 257 (2019).
74. Prjibelski, A., Antipov, D., Meleshko, D., Lapidus, A. & Korobeynikov, A. Using SPAdes De Novo Assembler. *Curr Protoc Bioinformatics* **70**, e102 (2020).
75. Meng, G., Li, Y., Yang, C. & Liu, S. MitoZ: a toolkit for animal mitochondrial genome assembly, annotation and visualization. *Nucleic Acids Res* **47**, e63 (2019).
76. Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. Basic local alignment search tool. *J Mol Biol* **215**, 403–10 (1990).
77. Gurevich, A., Saveliev, V., Vyahhi, N. & Tesler, G. QUAST: quality assessment tool for genome assemblies. *Bioinformatics* **29**, 1072–5 (2013).
78. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–60 (2009).

79. McKenna, A. *et al.* The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* **20**, 1297–303 (2010).
80. Kumar, P. *et al.* Genetic structuring and estimation of reproductive adults in Onchocerca volvulus: A genome-wide analysis across hosts and regions. *PLoS Negl Trop Dis* **19**, e0013221 (2025).
81. Janssen, K. *et al.* SNPeffect 5.0: large-scale structural phenotyping of protein coding variants extracted from next-generation sequencing data using AlphaFold models. *BMC Bioinformatics* **24**, 287 (2023).
82. Ortiz, E.M. vcf2phylip v2.0: convert a VCF matrix into several matrix formats for phylogenetic analysis. *Zenodo* (2019).
83. Minh, B.Q. *et al.* IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Mol Biol Evol* **37**, 1530–1534 (2020).
84. Letunic, I. & Bork, P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Research* **49**, W293–W296 (2021).
85. Raj, A., Stephens, M. & Pritchard, J.K. fastSTRUCTURE: Variational Inference of Population Structure in Large SNP Data Sets. *Genetics* **197**, 573–589 (2014).
86. Levy Karin, E., Mirdita, M. & Soding, J. MetaEuk-sensitive, high-throughput gene discovery, and annotation for large-scale eukaryotic metagenomics. *Microbiome* **8**, 48 (2020).
87. Jones, D.A.B. *et al.* An automated and combinative method for the predictive ranking of candidate effector proteins of fungal plant pathogens. *Scientific Reports* **11**, 19731 (2021).
88. Petersen, T.N., Brunak, S., von Heijne, G. & Nielsen, H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature Methods* **8**, 785–786 (2011).
89. Teufel, F. *et al.* SignalP 6.0 predicts all five types of signal peptides using protein language models. *Nature Biotechnology* **40**, 1023–1025 (2022).
90. Odum, M.T. *et al.* DeepLoc 2.1: multi-label membrane protein type prediction using protein language models. *Nucleic Acids Res* **52**, W215–W220 (2024).
91. Emanuelsson, O., Nielsen H Fau - Brunak, S., Brunak S Fau - von Heijne, G. & von Heijne, G. Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *Journal of molecular biology* **300**, 1005–1016 (2000).
92. Käll, L., Krogh, A. & Sonnhammer, E.L.L. A Combined Transmembrane Topology and Signal Peptide Prediction Method. *Journal of Molecular Biology* **338**, 1027–1036 (2004).
93. Krogh, A., Larsson, B., von Heijne, G. & Sonnhammer, E.L. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* **305**, 567–80 (2001).
94. Sonnhammer, E.L., von Heijne, G. & Krogh, A. A hidden Markov model for predicting transmembrane helices in protein sequences. *Proc Int Conf Intell Syst Mol Biol* **6**, 175–82 (1998).
95. Sperschneider, J., Dodds, P.N., Gardiner, D.M., Singh, K.B. & Taylor, J.M. Improved prediction of fungal effector proteins from secretomes with EffectorP 2.0. *Molecular Plant Pathology* **19**, 2094–2110 (2018).
96. Sperschneider, J. *et al.* EffectorP: predicting fungal effector proteins from secretomes using machine learning. *New Phytol* **210**, 743–61 (2016).
97. Sperschneider, J., Dodds, P.N., Singh, K.B. & Taylor, J.M. ApoplastP: prediction of effectors and plant proteins in the apoplast using machine learning. *New Phytologist* **217**, 1764–1778 (2018).
98. Sperschneider, J. *et al.* LOCALIZER: subcellular localization prediction of both plant and effector proteins in the plant cell. *Scientific reports* **7**(2017).
99. Potter, S.C. *et al.* HMMER web server: 2018 update. *Nucleic Acids Research* **46**, W200–W204 (2018).

100. Zheng, J. *et al.* dbCAN-seq update: CAZyme gene clusters and substrates in microbiomes. *Nucleic Acids Research* **51**, D557–D563 (2023).
101. Steinegger, M. & Söding, J. MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nature Biotechnology* **35**, 1026–1028 (2017).
102. Urban, M. *et al.* PHI-base: the pathogen–host interactions database. *Nucleic Acids Research* **48**, D613–D620 (2020).
103. Finn, R.D. *et al.* Pfam: the protein families database. *Nucleic acids research* **42**(2014).
104. Bairoch, A. & Apweiler, R. The SWISS-PROT Protein Sequence Data Bank and Its New Supplement TREMBL. *Nucleic Acids Research* **24**, 21–25 (1996).
105. Apweiler, R. *et al.* The InterPro database, an integrated documentation resource for protein families, domains and functional sites. *Nucleic Acids Res* **29**, 37–40 (2001).
106. Zheng, J. *et al.* dbCAN3: automated carbohydrate-active enzyme and substrate annotation. *Nucleic Acids Research* **51**, W115–W121 (2023).
107. Tang, Z.Z., Chen, G. & Alekseyenko, A.V. PERMANOVA-S: association test for microbial community composition that accommodates confounders and multiple distances. *Bioinformatics* **32**, 2618–2625 (2016).
108. Pascual-Garcia, A. & Bell, T. Community-level signatures of ecological succession in natural bacterial communities. *Nat Commun* **11**, 2386 (2020).
109. Harris, C.R. *et al.* Array programming with NumPy. *Nature* **585**, 357–362 (2020).

Supplementary Tables:

Available via Figshare at https://doi.org/10.6084/m9.figshare.34065522

Supplementary Table 1: (A) Sequencing quality, genome assembly metrics, repeat content, and GC distribution profiles for *P. nodorum* isolates and (B) geographic and climatic data of Western Australia isolates

Supplementary Table 2: Predector output, including functional annotations and effector-likelihood scores, used to prioritise candidate effector proteins for downstream analysis

Supplementary Table 3: Pathotyping based on known effectors and top 20 effector candidates

Supplementary Figures:

Available via Figshare at https://doi.org/10.6084/m9.figshare.34065522

Supplementary Figure 1: Rarefaction curve, demonstrating the saturation of gene content capture without our comprehensive set of isolates. The plateau indicates that the current number of isolates are sufficient to represent the majority of potential gene content within a global pan-genomic population study.

Supp Figure 2 (Chromosome-specific RIPness by region): chromosome specific average number of 10 kb genomic windows with RIPness > 0.6 per isolate across five geographic regions: Western Australia (WA, red), Europe (teal), United States (USA, blue), South Africa (SA, light green), and Iran (yellow). Data are presented for two SNP mutation categories: (A) All SNPs and (B) Non-synonymous SNPs.

Supp Figure 3: Comparison of RIPness across decades from the 1960s to the 2010s across the 23 chromosomes of *P. nodorum*, for the (A) global pangenome and (B) the Western Australian (WA) sub-set (right). Circular genome plots show RIPness scores across based on all SNPs within 10 kb windows, with concentric rings representing sequential decades from the 1960s (innermost) to the 2010s (outermost). Scatter points indicate SNP-level RIPness scores, coloured from green (low RIPness) to red (high RIPness).

Supp Figure 4: RIPness across the (A) Global and (B) Western Australia *Parastagonospora nodorum* populations: Chromosome-specific average 10kb genomic windows with RIPness score

(based on all SNPs) > 0.6 per isolate across five collection decades in WA *Parastagonospora nodorum* isolates.

Supplementary Figure 5. Principal Coordinates Analysis (PCoA) of isolate-level protein variant composition based on Jensen–Shannon divergence of variant-class distributions. Left panel: PCoA coloured by annual rainfall (mm) and Right panel: PCoA coloured by Years. Each point represents a single isolate. Overall clustering reflects similarities in the relative proportions of protein variant classes across isolates, rather than protein identity or sequence length.

Supplementary Figure 6: Average annual rainfall zones in south-west Western Australia (2000-2024) compared to 1910-1999 baseline (DPIRD, 2025 [https://library.dpird.wa.gov.au/gis_climate/4/]), with study location indicated (red circle).

Supplementary Figure 7: Average annual rainfall and max temperature before and after 2000 at five different locations, where we have old and new isolates

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementaryfigures.docx