

# Pangenome of *Pyrenophora tritici-repentis* reveals high plasticity and the association of virulence factors with mobile elements

Ryan Gourlie<sup>1,2</sup>, Megan McDonald<sup>3</sup>, Mohamed Hafez<sup>1</sup>, Rodrigo Ortega-Polo<sup>1</sup>, Stephen Strelkov<sup>4</sup> and Reem Aboukhaddour<sup>1\*</sup>

<sup>1</sup>AAFC-LRDC, Cereal Pathology, Lethbridge, AB

<sup>2</sup>University of Lethbridge, Department of Biosciences, Lethbridge, AB

<sup>3</sup>University of Birmingham, School of Biosciences, United Kingdom,

<sup>4</sup>Department of Agricultural, University of Alberta, Edmonton, AB



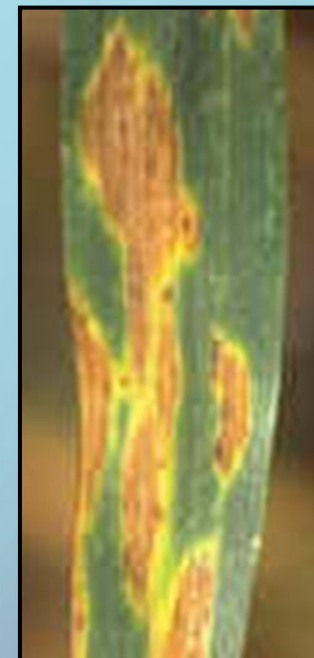
Agriculture and  
Agri-Food Canada

University of  
Lethbridge



# Tan spot of wheat

- *Pyrenophora tritici-repentis* (Ptr)
- Tan Spot is a foliar disease of wheat
- Worldwide occurrence, causes ~5% global losses<sup>1</sup>
  - Among top diseases of wheat
- Stubble borne with recent emergence as wheat pathogen in the 1940s via HGT<sup>2,3</sup>
- Ptr is a model for necrotrophic pathogens, pathogen emergence, and the inverse gene-for-gene model



<sup>1</sup>Savary et al. 2019. Nat. Ecol. Evol. 3, 430-439

<sup>2</sup>Friesen et al. 2006. Nat. Genet. 38, 953-956

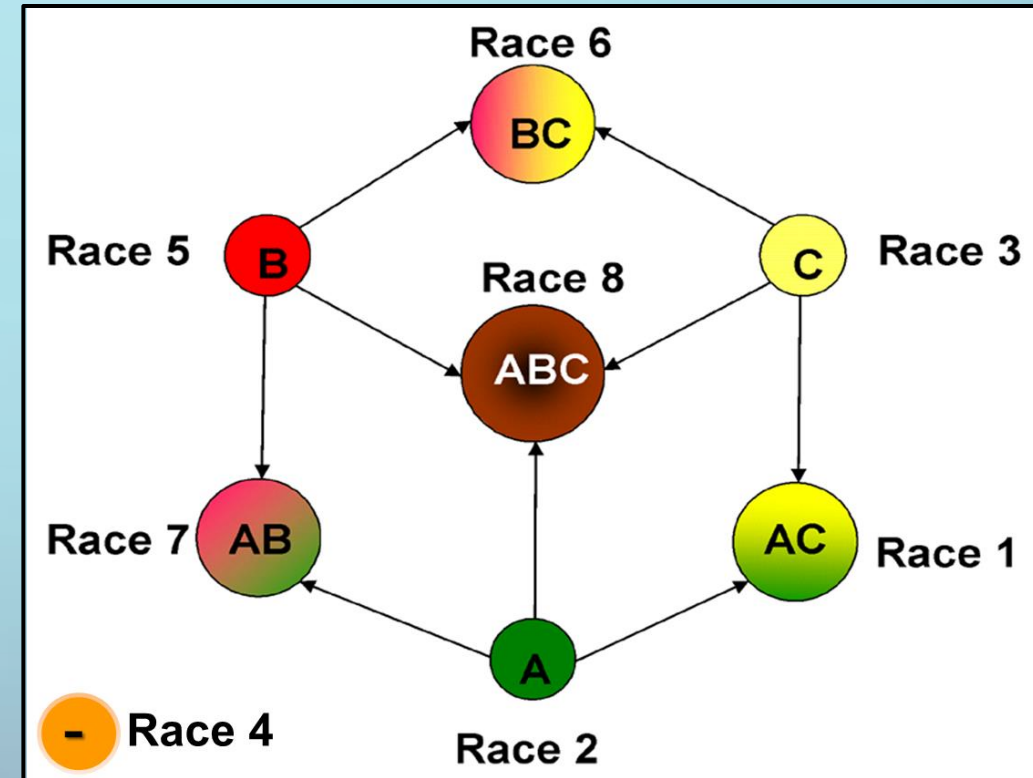
<sup>3</sup>McDonald et al. 2018. Mbio. 10, e01515-19



# Ptr and its necrotrophic effectors

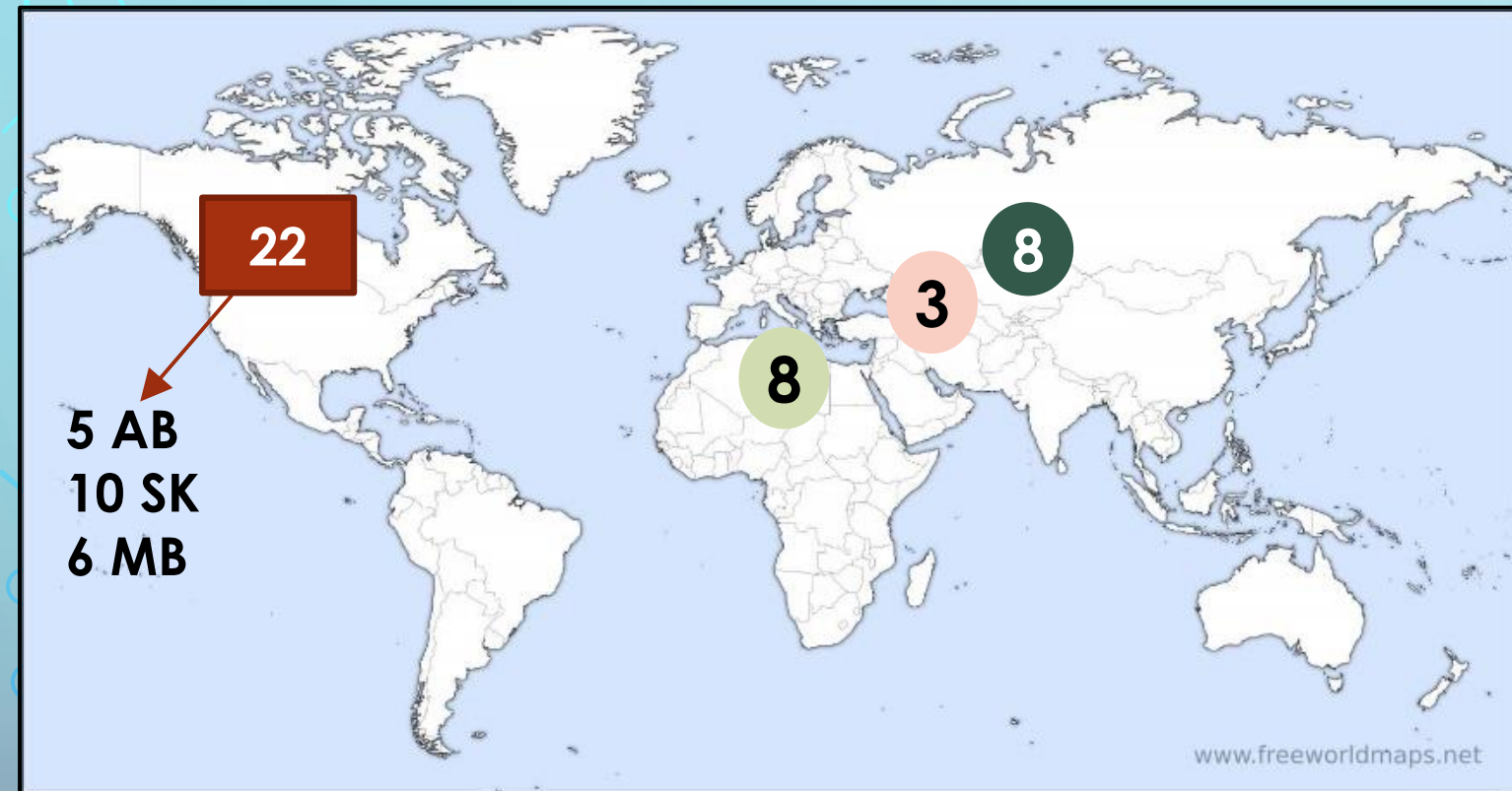
- Produces three necrotrophic effectors (previously host selective toxins)

<i>ToxA</i>	<i>ToxB</i>	<i>ToxC</i>
Necrosis	Chlorosis B	Chlorosis C
		



Strelkov and Lamari, 2003. CJPP 25, 339-349

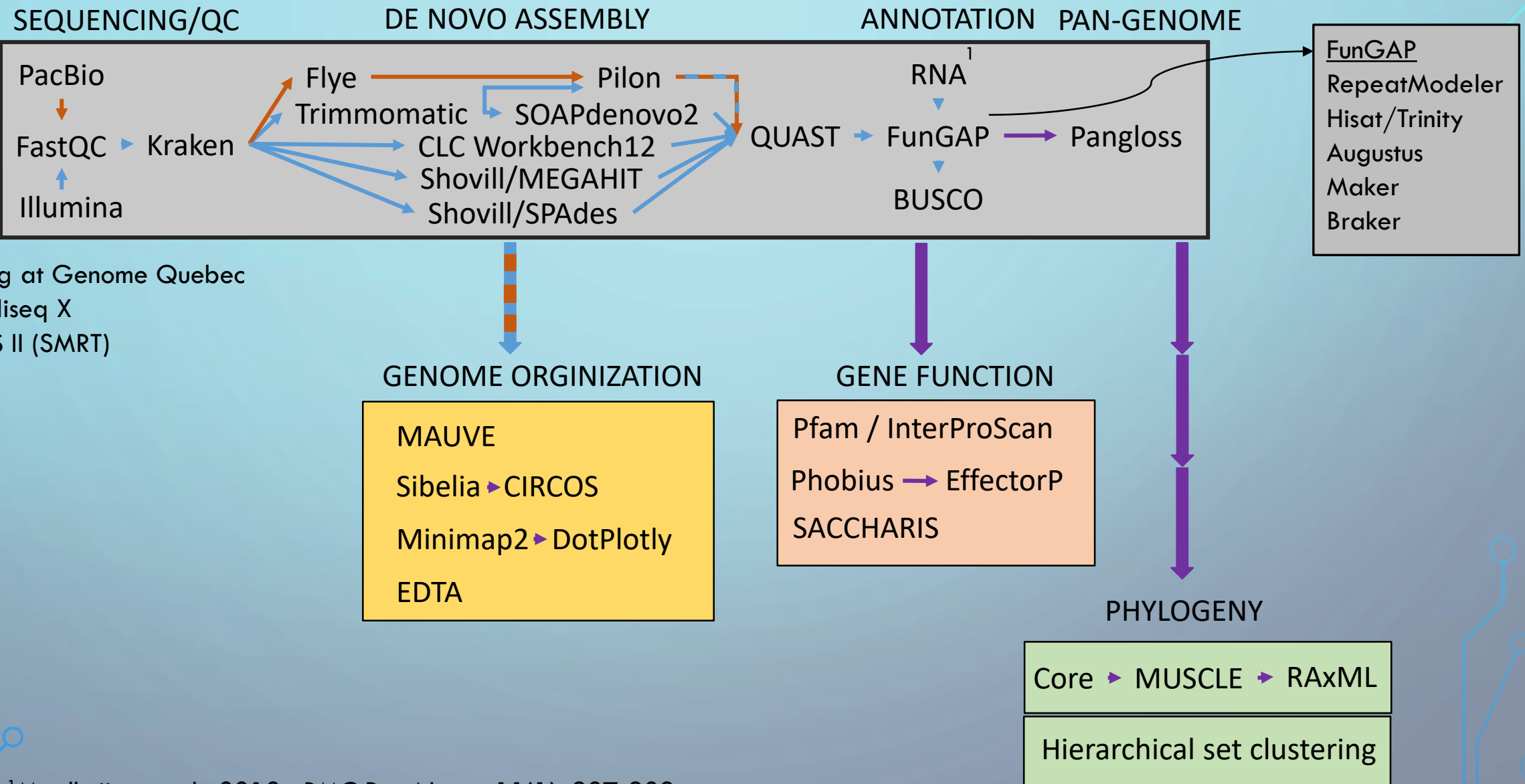
# Sequenced isolate details



Dates range from 1990 to 2017

Race	ToxA	ToxB	ToxC	Number of isolates sequenced	
1	+	-	+	10	
2	+	-	-	6	
3*	-	-	+	6	
4	-	-	-	3	
5	-	+	-	7	
6	-	+	+	1	
7	+	+	-	3	
8*	+	+	+	3	
novel	-	+	-	1	
				Total	40

# Current genomics pipeline



<sup>1</sup>Moolhuijzen et al., 2018. BMC Res. Notes, 11(1), 907-909

# Assemblies

- Consistent high quality assemblies for all short-read sequenced isolates (BUSCO >99%)
- Average Ptr genome size:  $34.8 \pm 2.1$  Mb
- Average Ascomycota<sup>1</sup>: 36.9 Mb
- Non-pathogen Ascomycota<sup>2</sup>: 34.8 Mb
- Plant pathogenic Ascomycota<sup>2</sup>: 39.4 Mb
- Largest: G9-4 (race 4) 36.97 Mb
- Smallest: T128-1 (novel) 34.12 Mb
- Average gene count: 13,071
- Ascomycota average<sup>1</sup>: 11,129
- Long read assemblies (Flye+Pilon)
  - I73-1 (ToxA, B, C): 39.9 Mb
  - D308 (ToxC): 39.7 Mb
  - Primarily due to transposons and repetitive elements

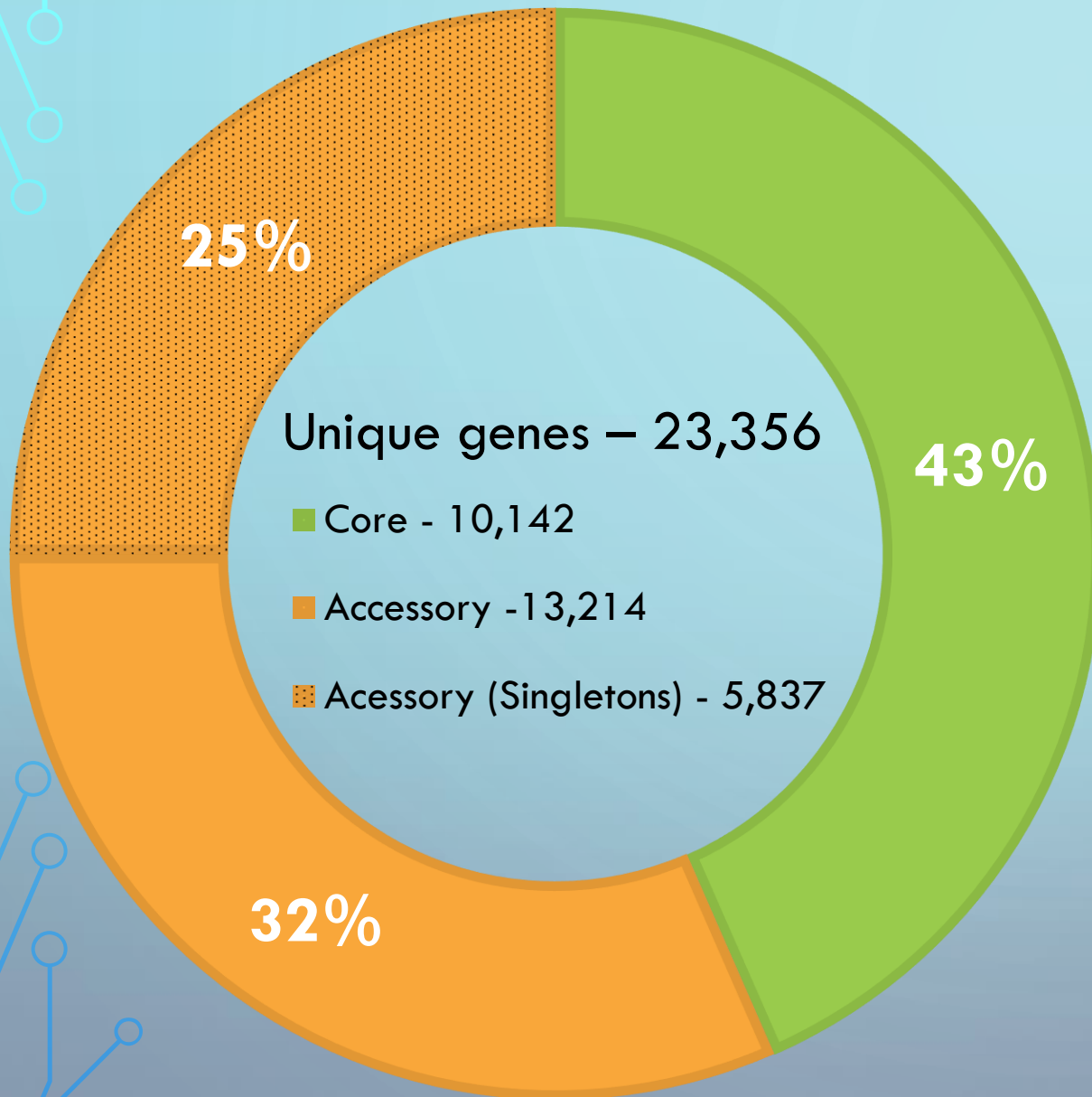
<sup>1</sup>Biol Proced Online 17, 8 (2015)

<sup>2</sup>IMA Fungus 8, 1–15 (2017)

Isolate	Race	HST	Year	Location	Size (MB)	Contigs	N50	Genes
ASC1	1	AC	1990	Manitoba	34.78	6,495	65,481	13,124
I33-1	1	AC	2001	Azerbaijan	35.06	6,850	75,237	13,055
L3-1	1	AC	2016	Alberta	34.93	6,666	77,181	13,115
L4-1	1	AC	2016	Alberta	34.74	6,361	72,358	13,063
SW20-7	1	AC	2016	Saskatchewan	35.00	6,568	76,189	13,116
SW2-1	1	AC	2016	Saskatchewan	35.11	6,789	74,336	13,116
SW21-1	1	AC	2016	Saskatchewan	34.62	6,282	74,274	12,965
SW21-7	1	AC	2016	Saskatchewan	34.97	6,741	74,663	13,126
SW21-8	1	AC	2016	Saskatchewan	34.96	6,631	75,023	13,073
SW7-5	1	AC	2016	Saskatchewan	35.65	6,701	73,682	13,490
86-124	2	A	1990	Manitoba	34.90	6,832	60,693	13,209
AB88-2	2	A	2010	Alberta	34.83	6,465	75,256	13,126
L2-1	2	A	2016	Alberta	34.96	6,923	72,851	13,130
SW1-2	2	A	2016	Saskatchewan	35.08	6,599	71,303	13,365
SW15-1	2	A	2016	Saskatchewan	34.83	6,272	78,811	13,201
T132-2	2	A	2017	Tunisia	34.41	6,472	57,094	12,935
331-2	3	C	2001	Manitoba	34.44	6,828	55,623	12,909
D308	3	C	1990	Manitoba	34.33	6,809	58,536	12,826
I72-1	3	C	2001	Syria	34.35	6,734	58,599	13,011
I72-7	3	C	2001	Syria	34.35	6,696	58,913	12,901
SC29-1	3	C	1999	Saskatchewan	34.19	6,464	58,971	12,951
SW21-5	3	C	2016	Saskatchewan	34.66	6,619	63,491	13,029
90-2	4	absent	2016	Alberta/Saskatchewan	35.22	3,818	225,924	12,909
G9-4	4	absent	2016	Alberta	36.97	8,035	78,161	13,148
T126-1	4	absent	2017	Tunisia	34.15	6,373	62,353	12,837
92-171-R5	5	B	1997	Canada	36.81	14,647	45,051	13,393
Alg3-24	5	B	1995	Algeria	34.30	6,098	73,593	12,900
Alg4x-1	5	B	1995	Algeria	35.71	8,072	70,965	13,193
I17-2	5	B	2001	Azerbaijan	34.25	6,555	62,487	12,820
I34-5	5	B	2001	Azerbaijan	34.29	6,315	61,895	12,841
I35-56	5	B	2001	Azerbaijan	34.24	6,516	62,616	12,918
I36-1	5	B	2001	Azerbaijan	34.36	6,467	62,439	12,881
AlgH1	6	BC	1995	Algeria	34.74	6,902	61,661	13,159
AZ35-5	7	AB	2001	Azerbaijan	35.30	7,165	72,088	13,239
T176-2	7	AB	2017	Tunisia	34.70	6,896	57,095	13,141
T181-1	7	AB	2017	Tunisia	34.78	6,718	57,473	13,583
I34-1	8	ABC	2001	Azerbaijan	34.38	6,405	64,478	13,036
I35-18	8	ABC	2001	Azerbaijan	34.85	6,698	65,336	13,071
I73-1	8	ABC	2001	Syria	34.62	6,619	63,370	12,941
T128-1	atypical	B	2017	Tunisia	34.12	6,095	59,072	13,002

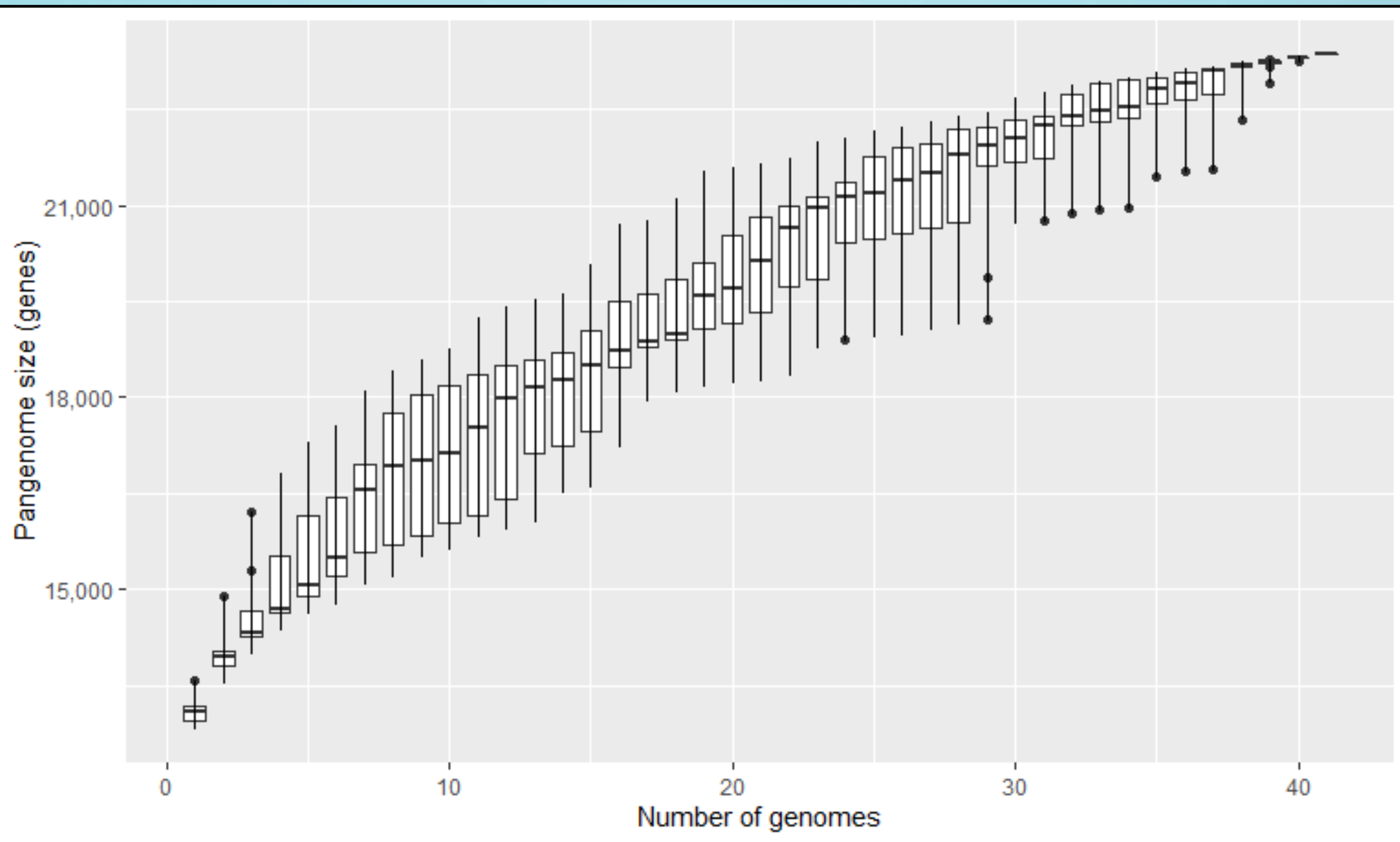


# Pangenome of *Pyrenophora tritici-repentis*



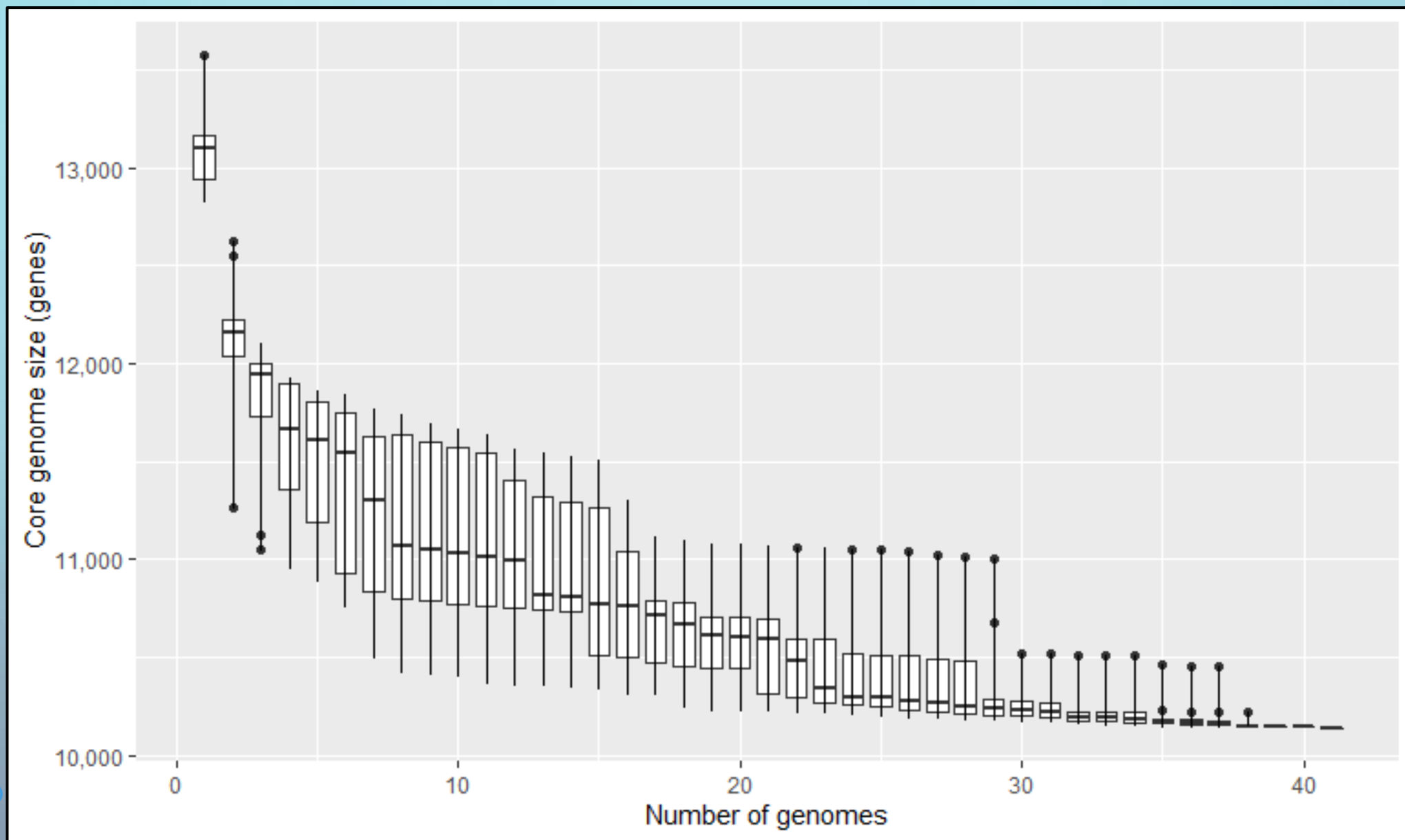
- Core = genes present in all isolates
- Accessory = present in some isolates
- Singletons = present in one isolate
- Core genome (43%)
  - 69% had domains in Pfam database
- Accessory (57%)
  - 28% had domains in Pfam database
- Large accessory genome and very large singleton count
- More genes ~ more functions ~ higher adaptability
- Huge portion of singletons are from race 4 non-pathogenic isolates and a divergent race 5 (56% of singletons)

# Pangenome of *Pyrenophora tritici-repentis*

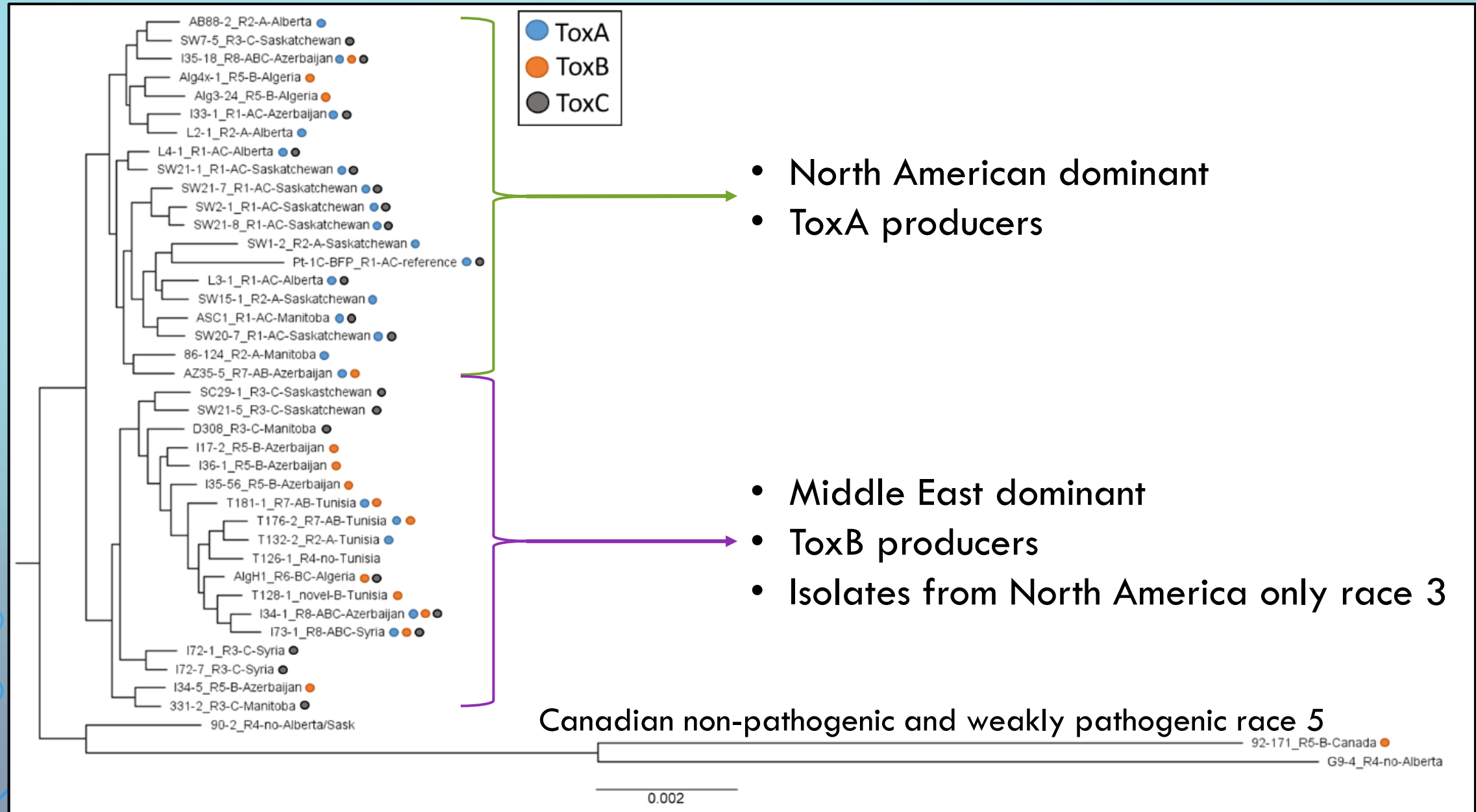




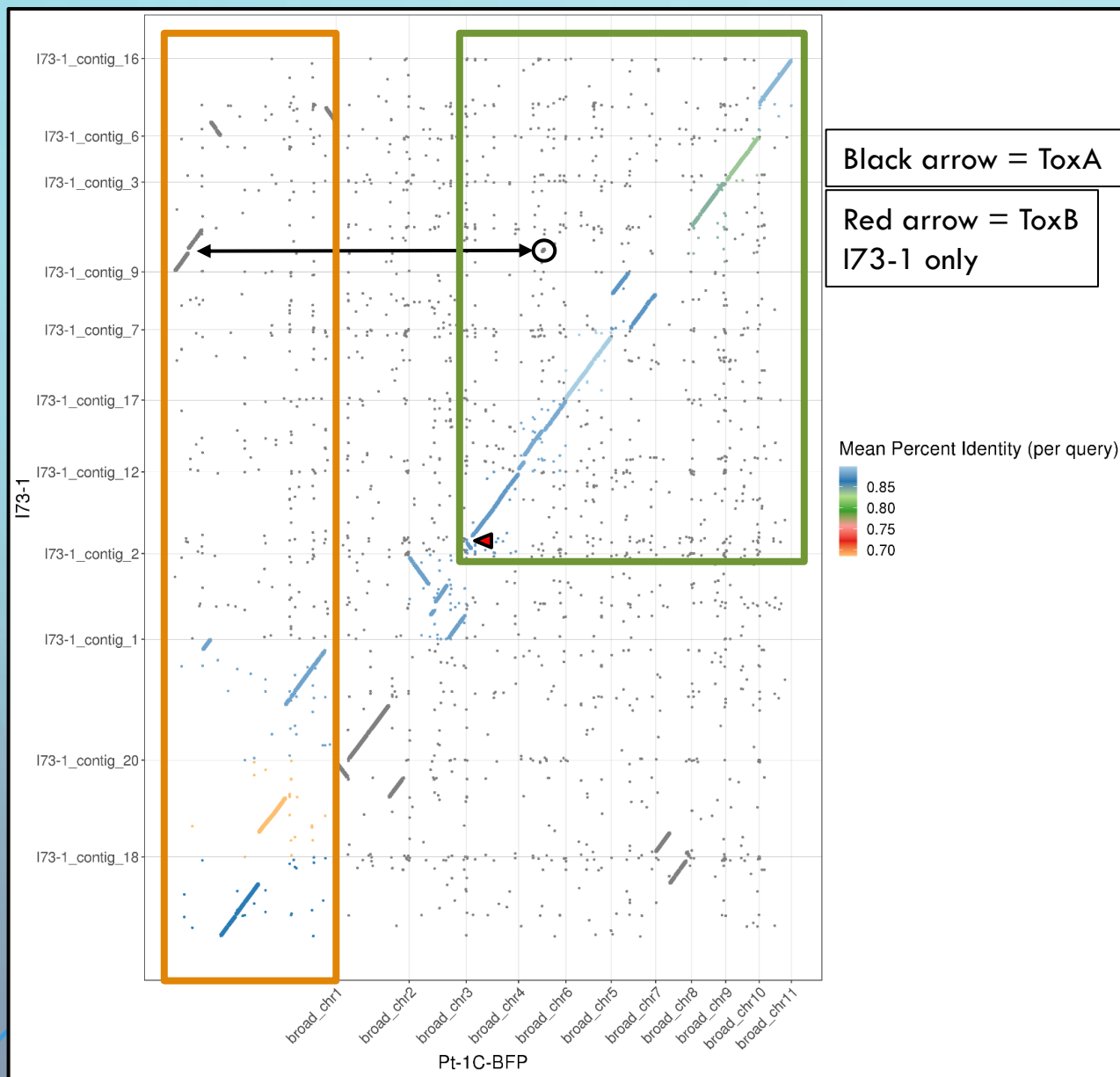
# Pangenome of *Pyrenophora tritici-repentis*



# Core gene Maximum-Likelihood phylogeny of *Ptr*

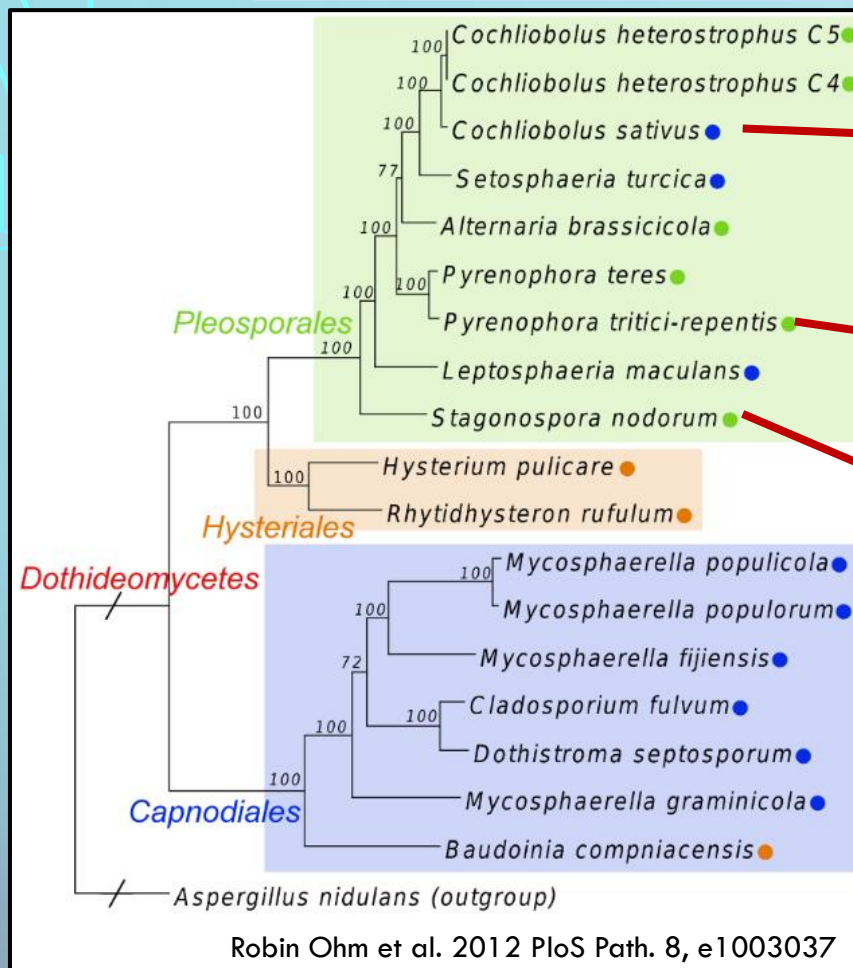


# Chromosomal rearrangements

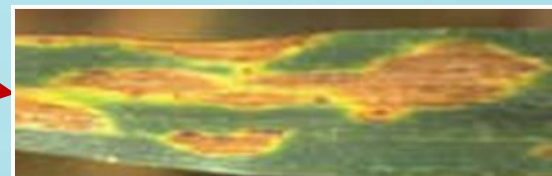


- Dotplot full-genome alignment of reference isolate Pt-1C-BFP race 1 (A, C) and I73-1 race 8 (all effectors)
- Chromosomes 3, 4, 5, 6, 7, and 10 largely intact between race 1 and 8 (green box)
- Major fragmentation of chromosome 1 with sections present in 5 contigs (orange box)
- Few other rearrangements and large inversions in other chromosomes
- Translocation of *ToxA* within Ptr
- Absence of *ToxB* region in reference

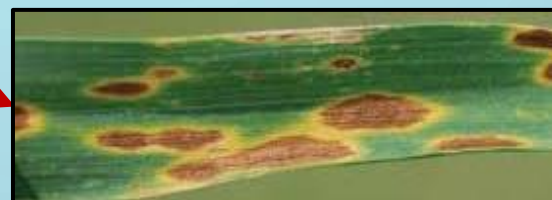
# ToxA is present in other species



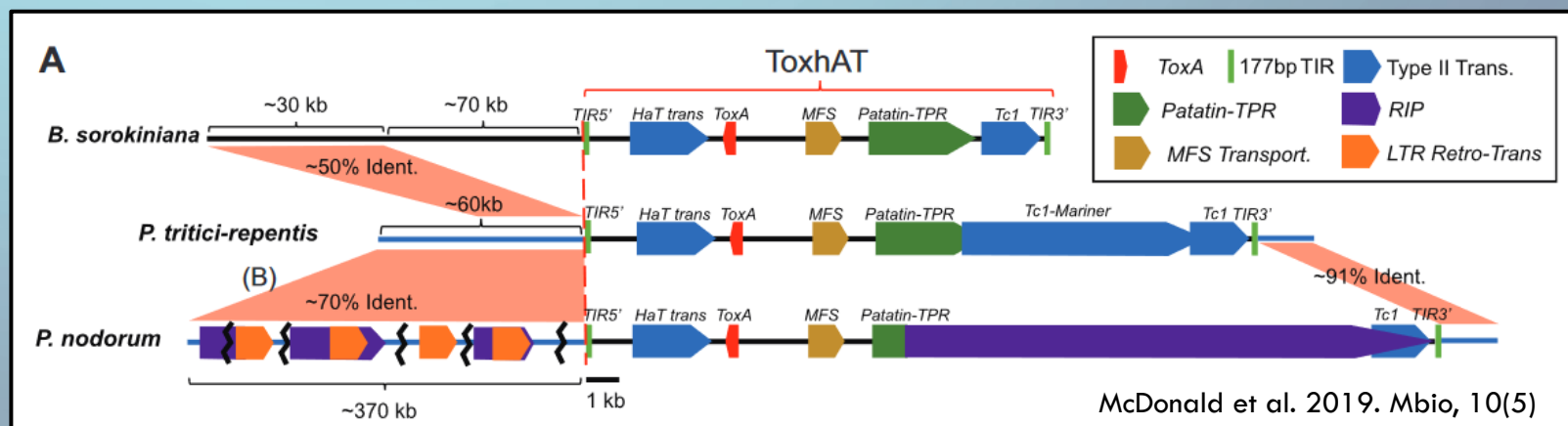
Spot blot



Tan spot

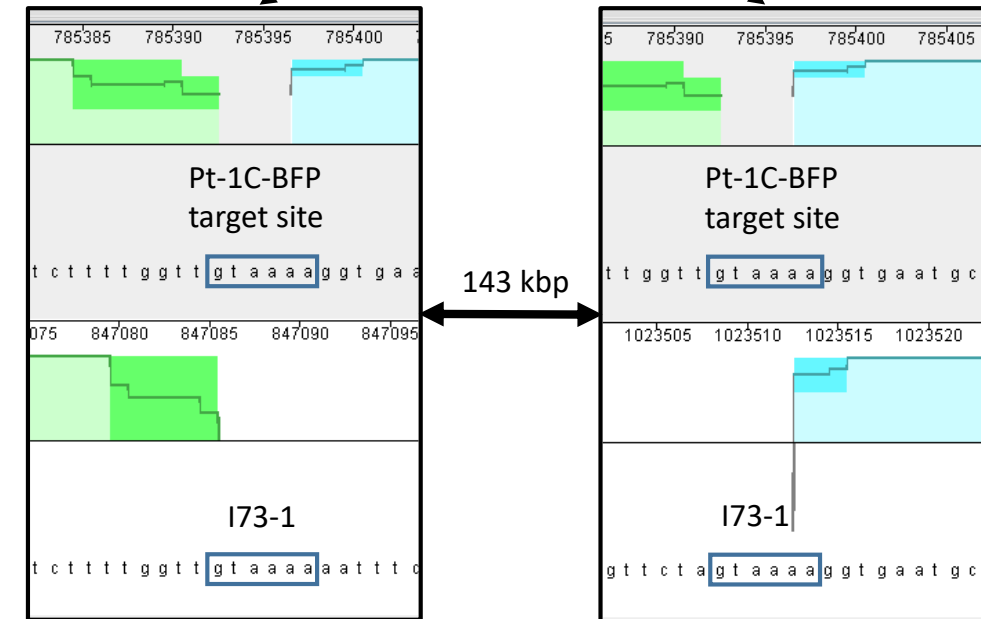
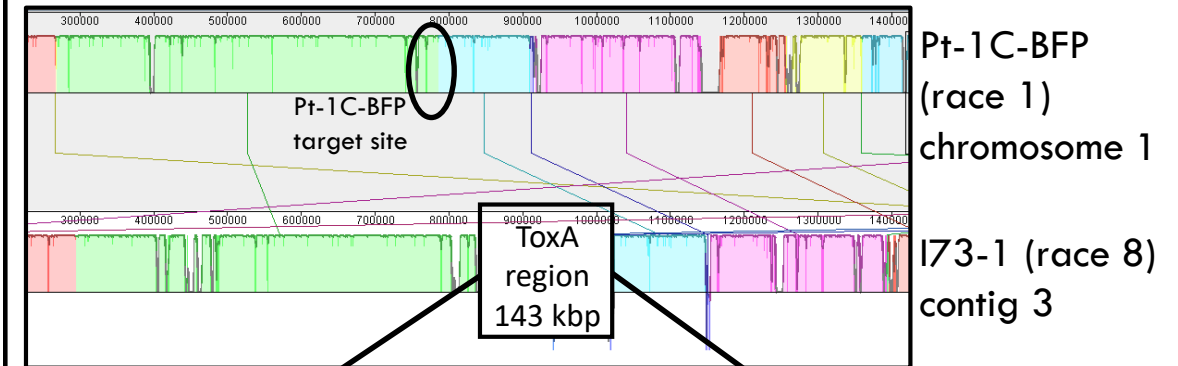
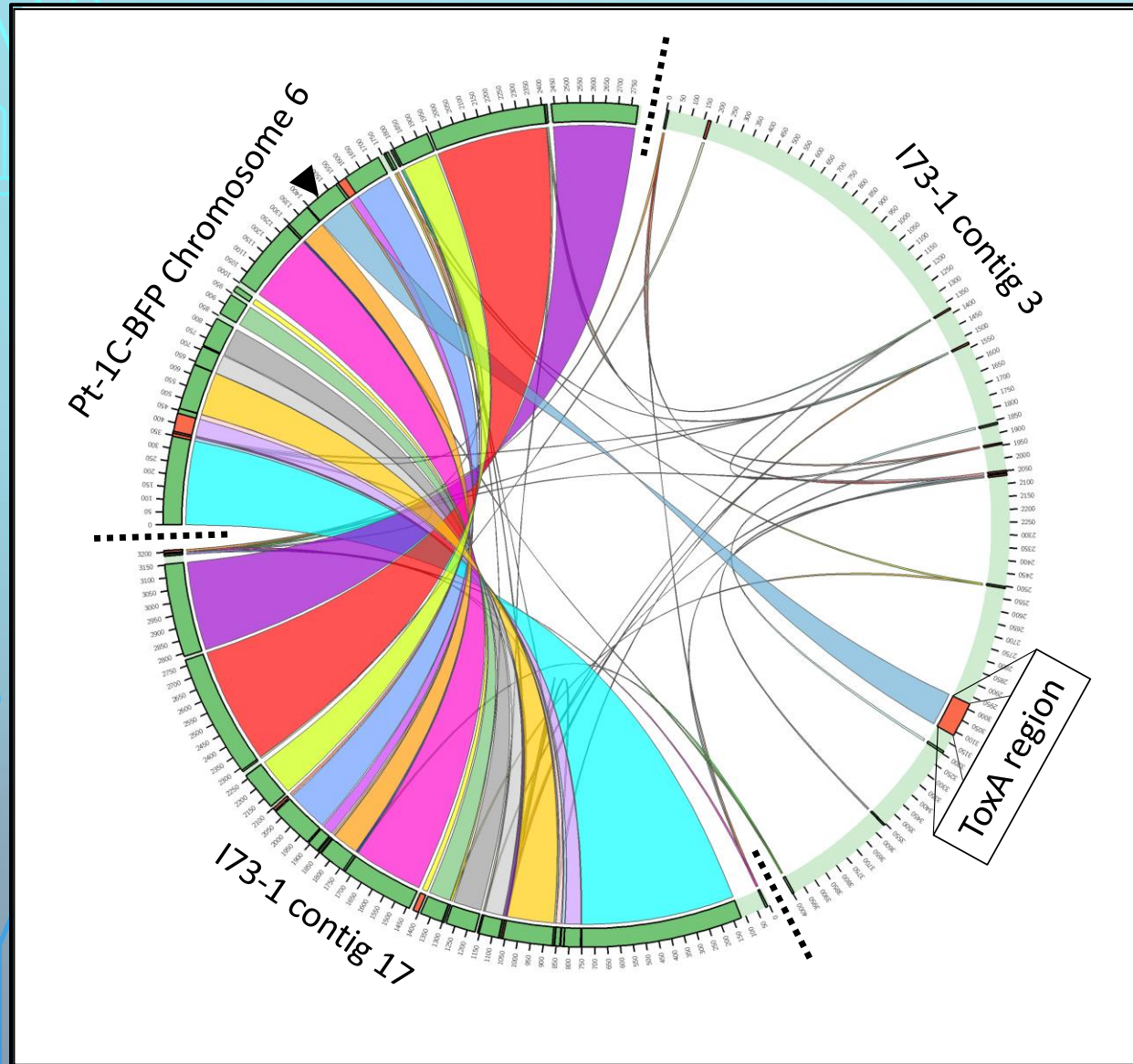


Septoria nodorum



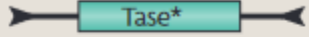
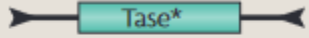
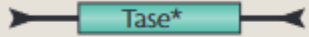
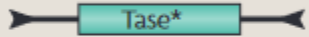
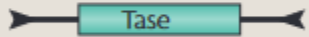
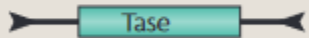


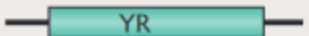


# Intra-specific translocation of *ToxA* via massive transposon



**Short Direct Repeats**

# Intra-specific translocation of *ToxA* via massive transposon

Class II (DNA transposons) - Subclass 1				Wicker et al. 2007. Nat. Rev. Genet. 8, 973-982		
TIR	<i>Tc1-Mariner</i>			TA	DTT	P,M,F,O
	<i>hAT</i>			8	DTA	P,M,F,O
	<i>Mutator</i>			9-11	DTM	P,M,F,O
	<i>Merlin</i>			8-9	DTE	M,O
	<i>Transib</i>			5	DTR	M,F
	<i>P</i>			8	DTP	P,M
	<i>PiggyBac</i>			TTAA	DTB	M,O
	<i>PIF-Harbinger</i>			3	DTH	P,M,F,O
	<i>CACTA</i>			2-3	DTC	P,M,F
Crypton	Crypton			0	DYC	F



		Phyre2	
98.2	11	<b>PDB header:</b> recombination <b>Chain:</b> A: <b>PDB Molecule:</b> tyrosine recombinase xerh; <b>PDB title:</b> crystal structure of xerh site-specific recombinase bound to 2 palindromic DNA substrate: post-cleavage complex	

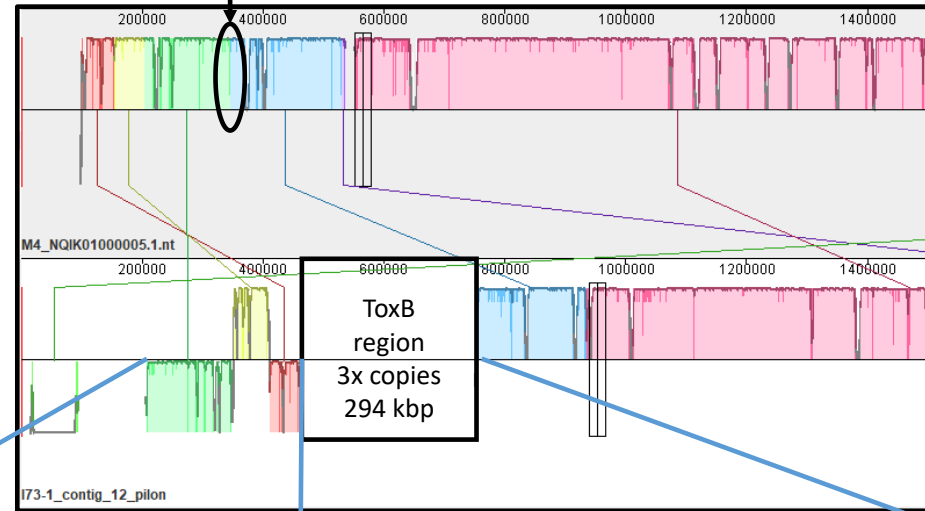
Protein folding prediction of gene\_09853  
located near 5' end of putative crypton

# Multiple copies of *ToxB* cluster in a massive transposon

M4 (race 1)  
contig NIK5

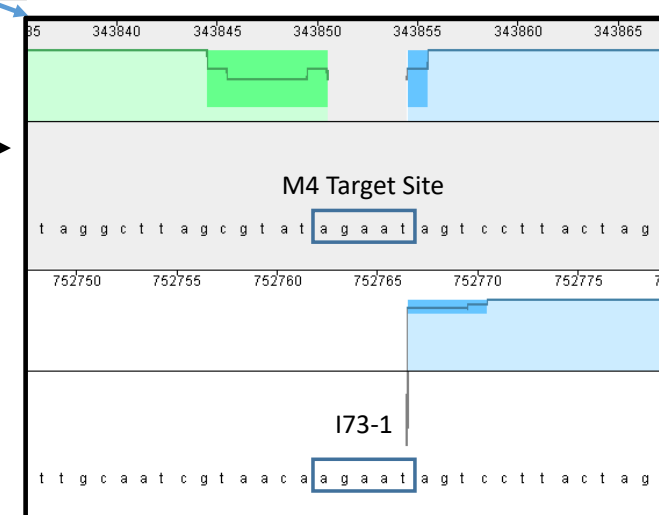
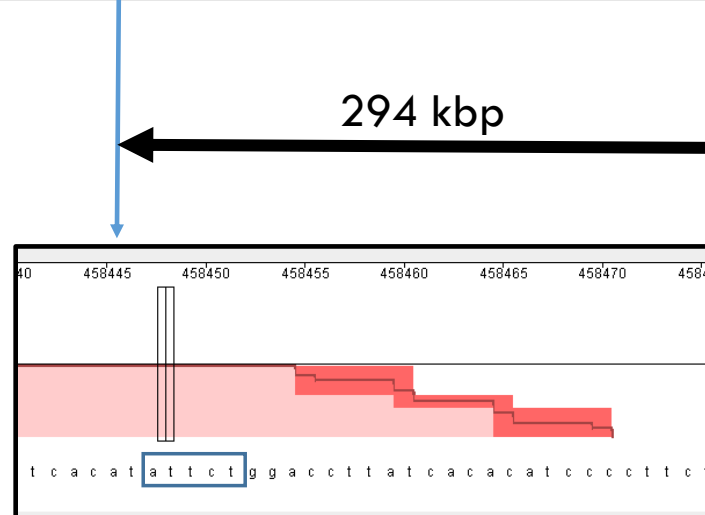
I73-1 (race 8)  
contig 12

M4 Target site



- *ToxB* present on essential chromosome

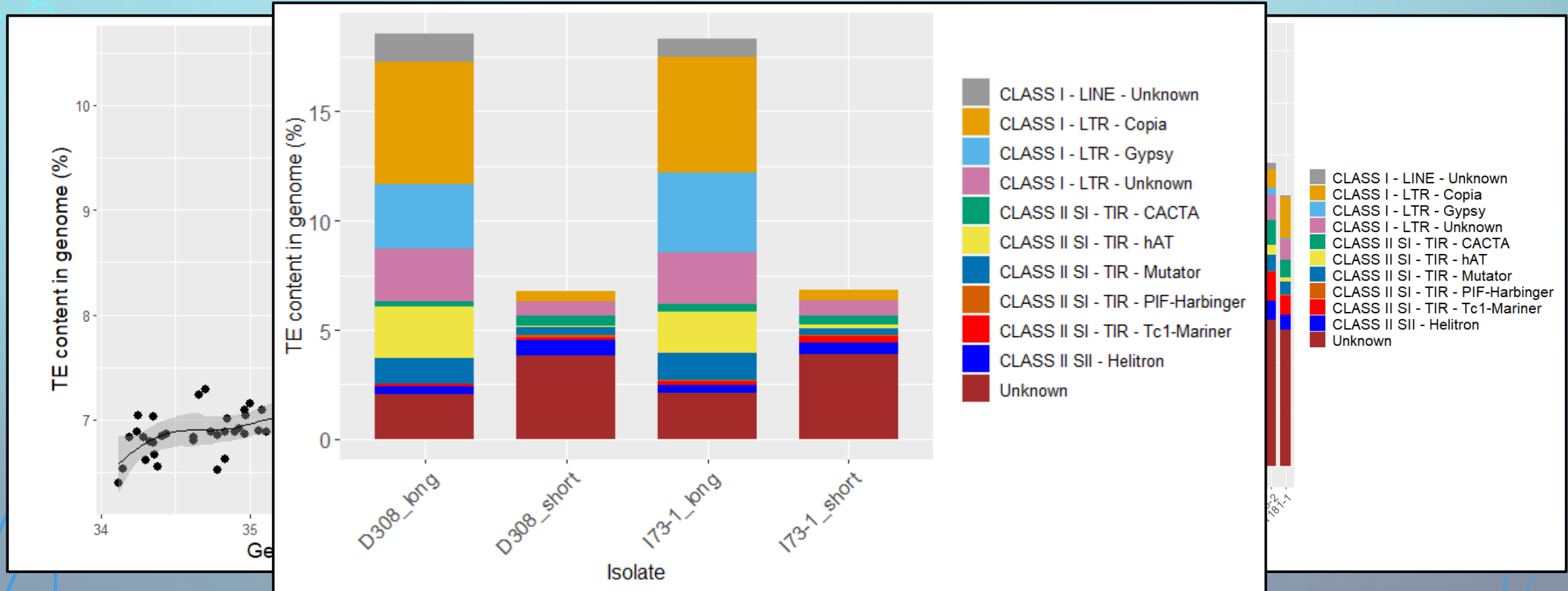
- D308 (race 3) also has transposon but it contains inactive *tox**b*



Terminal Inverted Repeat(s)

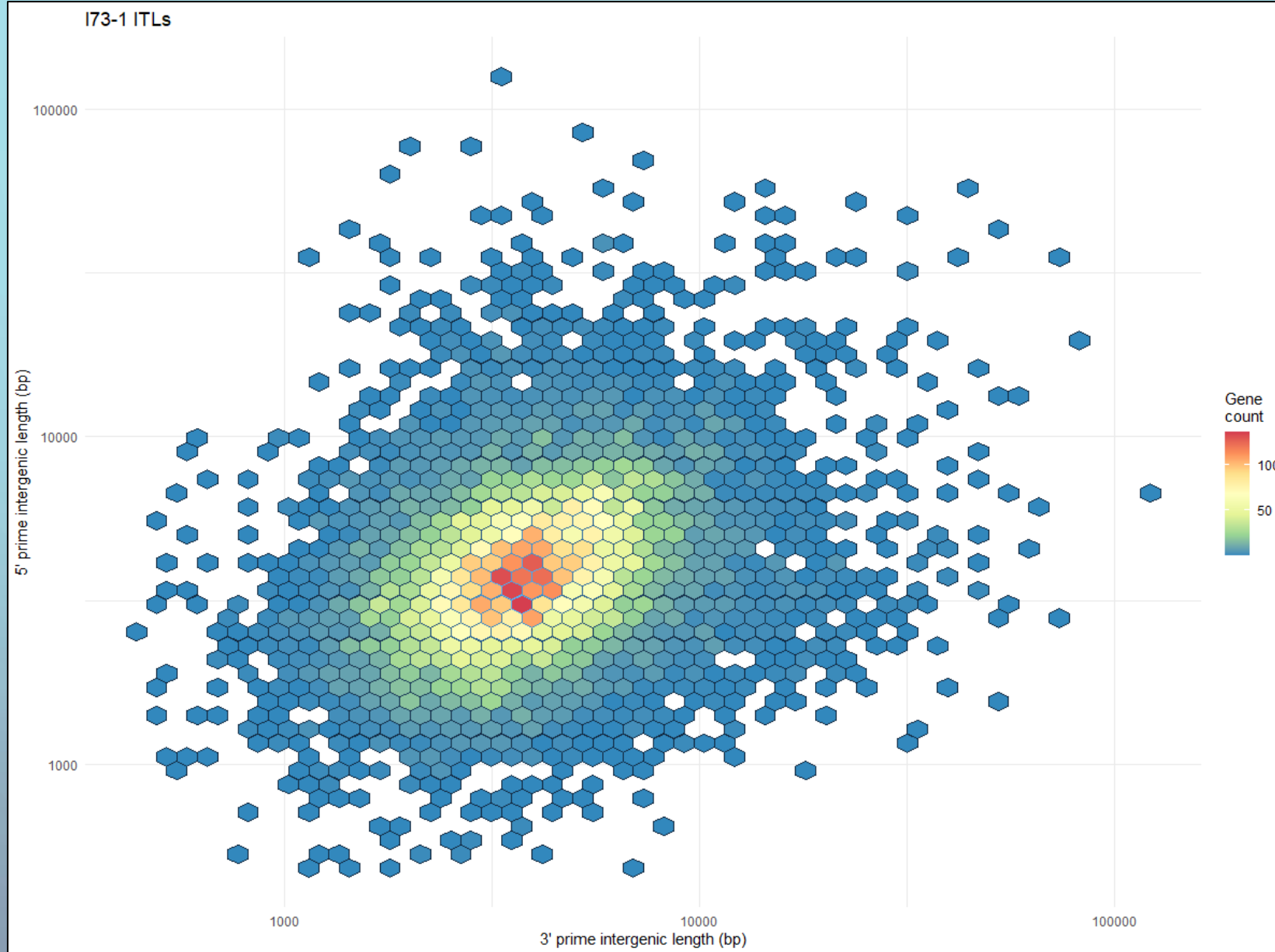
# Transposable element content

- Genome expansion in some isolates is driven by TE invasion
- Long read assemblies captured >150% more transposons





# Ptr likely has 'one-compartment' genome



## Main results/conclusions to date

- New high quality short-read assemblies representing all races
- First long-read assemblies of races 3 (ToxC) and 8 (ToxA, B, C)
- Ptr has an open, 'one-compartment' genome and is highly adaptable
  - Huge accessory gene count
  - Large numbers of gene gains and losses
  - Distinct gene sets between pathogen vs non-pathogen (race 4)
- Large structural reorganizations between races
- Distinct phylogenetic clustering by ability to produce certain effectors
- Confirmation of *ToxA* translocation within Ptr species
- *ToxhAT* is nested within a larger mobile element, likely a crypton
- Evidence that *ToxB* is present on massive transposon
- *ToxB* is located as multiple copies on essential chromosome
- High TE content contributes to both genome expansion and the movement of virulence factors

## On going analysis and future work

- Current priority is pangenome manuscript
- Sequence more Ptr isolates with long-read
- Sequence related *Pyrenophora* species with ToxB homologs
- Analysis of *ToxB* regions ~ How does *ToxB* (and/or its homologs) replicate and/or move?
- Identify the nature of ToxC via genome wide association study
- Further characterize genome structure between races/isolates
  - Requires more long-read seq
  - Non-pathogenic (race 4) vs. pathogenic isolates
  - Related species to Ptr
  - Confirm one vs. two compartment genome (AT/GC content?)

### *Related non-thesis work in our lab*

- Novel effector identification (Mohamed Hafez)
- Fungicide sensitivity and mutations (Mouldi Zid)

# ACKNOWLEDGEMENTS

- Dr. Reem Aboukhaddour
- Dr. Megan McDonald
- Dr. Dmytro Yevtushenko
- Dr. Rodrigo Ortega-Polo
- Dr. Mohamed Hafez
- Dr. Stephen Strelkov
- Dr. Wade Abbott and Kristin Low
- Dr. Fouad Dyaaf
- High Performance Computing Biocluster Team
- Funders and Supporters



UNIVERSITY OF  
BIRMINGHAM



# Pipeline references

Kraken2: Wood et al., 2019. *Genome Biology* 20  
SPAdes: Bankevich et al., 2012. *Journal of Computation Biology*, 19(5), 455-477  
Shovill: Seemann, 2019. [github.com/tseemann/shovill](https://github.com/tseemann/shovill)  
MEGAHIT: Li et al., 2015. *Bioinformatics*, 31(10), 1674-1676  
SOAPdenovo2: Luo et al., 2012. *Gigascience*, 1(1), 18  
Flye: Lin et al., 2016. *Proceedings of the National Academy of Sciences*, 113(52), E8396-E8405  
Pylon: Walker et al., 2014. *PloS One*, 9(11), e112963  
BUSCO: Simão et al., 2015. *Bioinformatics* 31(19), 3210-3212  
Fungap: Min et al., 2017. *Bioinformatics* 33(18), 2936-2937  
RNA: Moolhuijzen et al., 2018. *BMC Research Notes*, 11(1), 907-909  
BUSCO: Simão et al., 2015. *Bioinformatics* 31(19), 3210-3212  
Pangloss: McCarthy & Fitzpatrick, 2019. *Genes* 10(7), 521  
Reference isolate: Manning et al., 2013. *G3* 3(1), 41-63  
Hierarchical Sets: Pedersen 2016. [github.com/thomasp85/hierarchicalSets](https://github.com/thomasp85/hierarchicalSets)  
Phobius: Käll et al., 2004. *Journal of Molecular Biology* 338(5), 1027-1036  
EffectorP: Sperschneider et al., 2018. *Molecular Plant Pathology*.  
MUSCLE: Edgar, 2004. *Nucleic Acid Research* 32(5), 1792-1797  
RAxML: Stamatakis, 2014. *Bioinformatics* 30(9), 1312-1313  
Mauve: Darling, et al., 2004. *Genome Research* 14(7), 1394-1403  
EDTA: Ou et al., 2019. *Genome Biology* 20(1), 1-18  
Sibelia: Minkin et al., 2013. *Int. Workshop on Algorithms in Bioinformatics*, 215-229  
CIRCOS: Krzywinski et al., 2009. *Genome Research* 19(9), 1639-1645  
DotPlotly: Poorten, 2018. <https://github.com/tpoorten/dotPlotly>  
Minimap2: Li, 2018. *Bioinformatics*, 34(18), 3094-3100  
Phyre2: Kelley et al., 2015. *Nature Protocols*, 10(6), 845-858

