

# WHOLE GENOME SEQUENCE OF *Puccinia striiformis* f. sp. *tritici* AND GENOME SIZE COMPARISONS WITH *P. graminis* f. sp. *tritici* and *P. triticina*

Christina CUOMO<sup>1</sup>, Sarah YOUNG<sup>1</sup>, Meinan WANG<sup>2</sup>, Chuntao YIN<sup>2</sup>, Scot HULBERT<sup>2</sup>, and Xianming CHEN<sup>2,3</sup>

<sup>1</sup> Broad Institute of the Massachusetts Institute of Technology and Harvard University, Cambridge, MA, USA; <sup>2</sup> Department of Plant Pathology, Washington State University; and <sup>3</sup> USDA-ARS Wheat Genetics, Quality, Physiology, and Disease Research Unit, Pullman, WA, USA



## Abstract

*Puccinia striiformis* f. sp. *tritici* (*Pst*) causes stripe rust, an economically important disease of wheat worldwide. Whole genome sequencing of the fungus was done using genomic DNA from urediniospores of race PST-78, representing a group of races that have become predominant in the United States since 2000. The sequence assembly was generated using multiple assemblers from a combination of illumina and 454 data which had 81x coverage of the genome. The assembly size is 117.31 Mb, including 79.31 Mb for the total length of 17,295 contigs arranged into 9,715 scaffolds. The assembled *Pst* genome size is considerably bigger than that (88.64 Mb) of the stem rust pathogen (*P. graminis* f. sp. *tritici*), but smaller than that (135.34 Mb) of the leaf rust pathogen (*P. triticina*). The *Pst* total contig length is slightly smaller than that (81.92 Mb) of the stem rust pathogen, and smaller than that (106.57 Mb) of the leaf rust pathogen. The *Pst* genome contains 19,542 genes much more than that of *P. graminis* or *P. triticina*. Sequencing various cDNA libraries from PST-78 are underway for gene confirmation. Four additional *Pst* isolates are being sequenced to identify genome variations among races of the wheat stripe rust pathogen. The assembled *Pst* sequence is useful for understanding the genome and genes of the pathogen. It provides a whole genome database for developing molecular markers to study population structures of the pathogen and interactions between the pathogen and host, and also for comparative studies of important *Puccinia* and other fungal pathogens.

## Introduction

*Puccinia striiformis* f. sp. *tritici* (*Pst*) causes stripe rust, an important diseases of wheat worldwide. It has become increasingly important in the U.S. where epidemics since 2000 have caused yield losses of over 246 million bushels in spite of millions of dollars spent on fungicides. *Pst* populations change their virulence amazingly rapidly by mechanisms that are largely unknown. *Pst* is recalcitrant to classical genetic manipulations because it is an obligate biotroph producing essentially by asexual reproduction; our best chance of understanding its genetic nature and interaction with wheat lie in comparing genomic and transcriptome sequences of different isolates and comparisons to related *Puccinia* species.

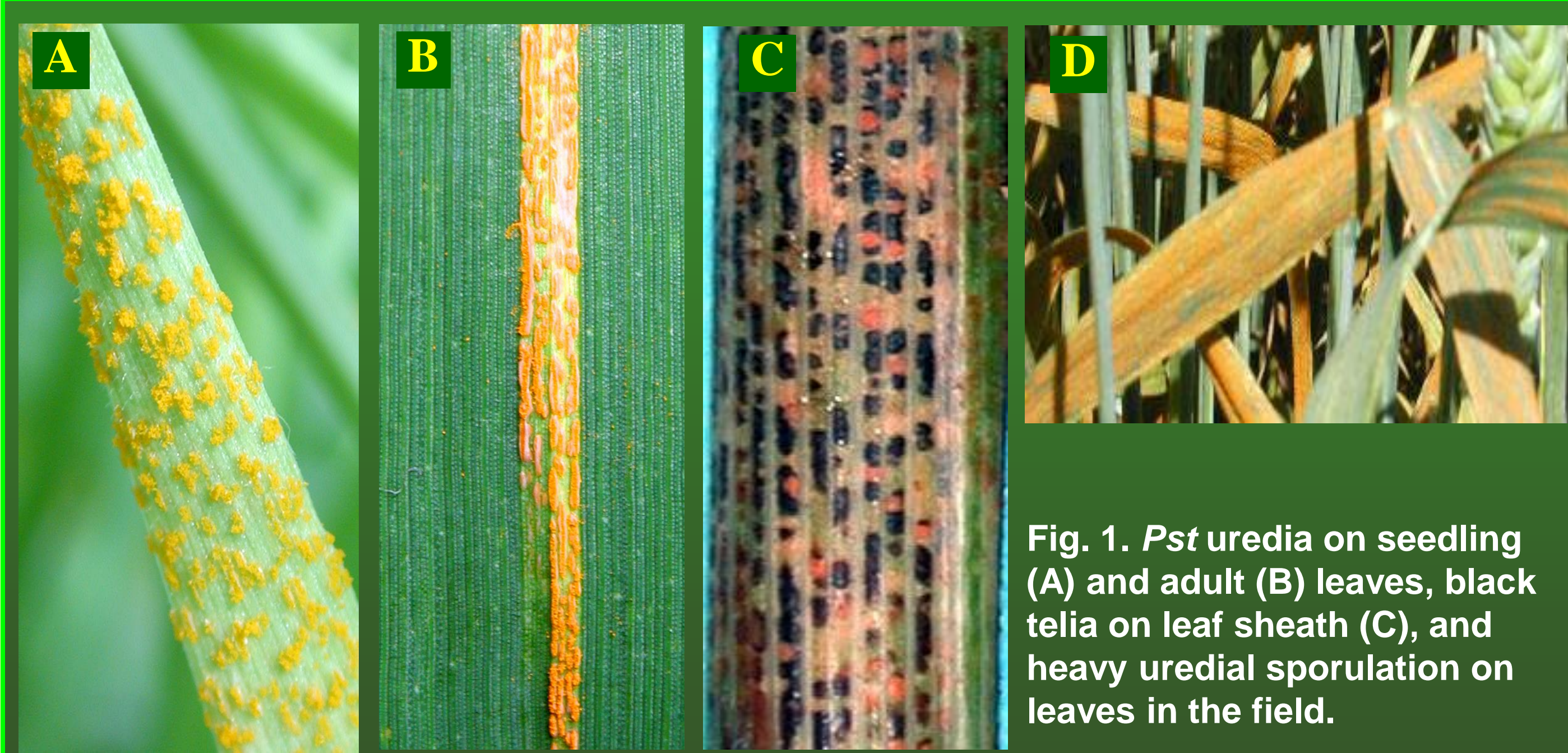


Fig. 1. *Pst* uredia on seedling (A) and adult (B) leaves, black telia on leaf sheath (C), and heavy uredial sporulation on leaves in the field.

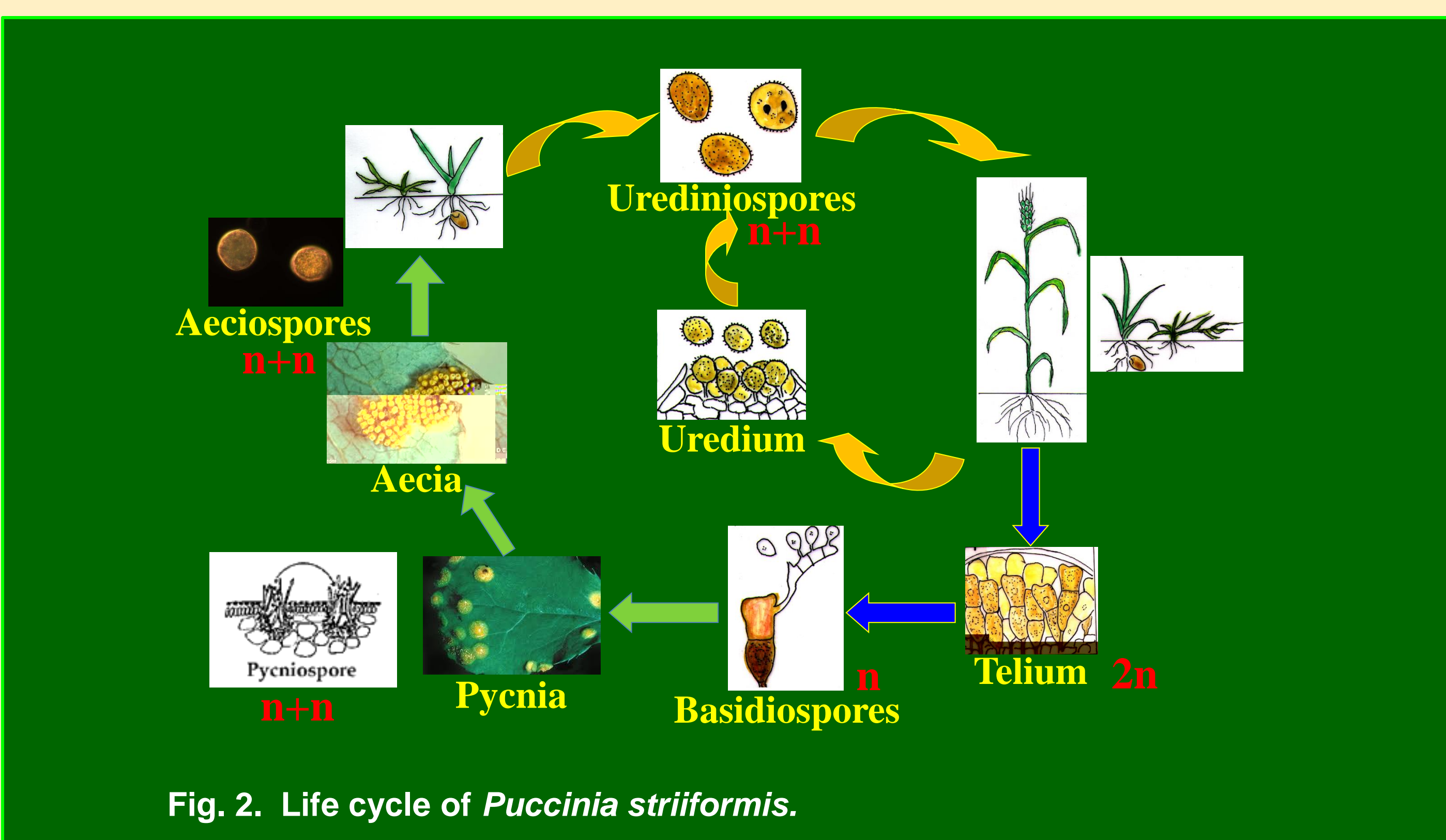


Fig. 2. Life cycle of *Puccinia striiformis*.

## Project Goals

- 1) Produce an assembled whole genome sequence of *Pst* from 15X coverage of 454 fragment reads, 30X coverage from 454 paired-end reads, and 10X coverage of Fosmid-end reads.
- 2) Sequence ESTs from several developmental stages and annotate the sequence/genes using both automated analyses of genomic sequence and directed annotations using ESTs.
- 3) Produce 15X genome coverage for four additional isolates for discovering polymorphisms using Illumina sequencing.
- 4) Make the data available to the public through a series of linked web sites that offer the sequence data, gene predictions, and preliminary in-silico expression data as well as tools to search and analyze the data.
- 5) Compare genomes and genes of *Pst* with those of *Pgt* (stem rust) and *Pt* (leaf rust).
- 6) Train young scientists in bioinformatics and genomics and enhance the biological interest of underserved minority youngsters.

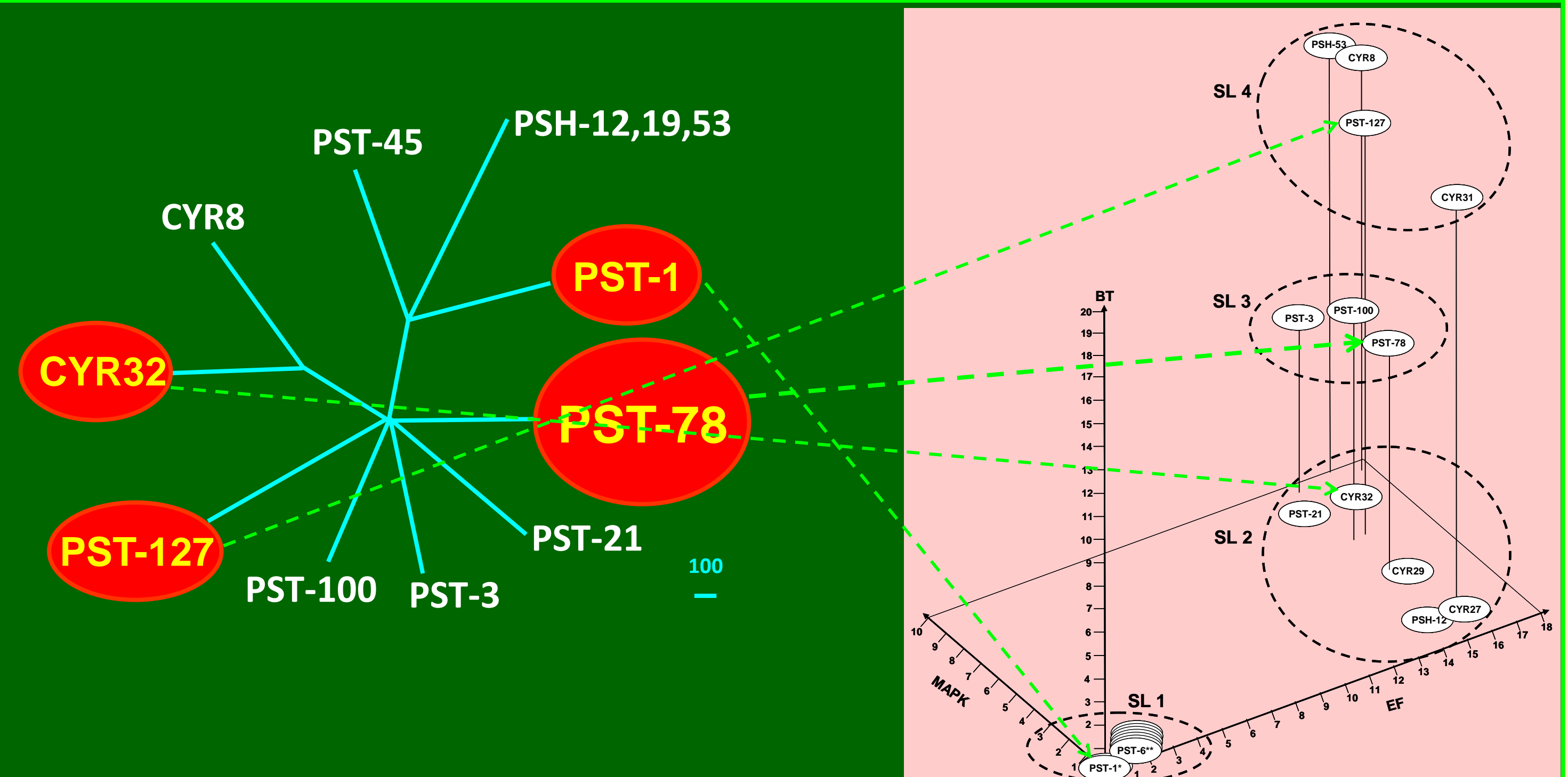


Fig. 3. Genetic relationships of races PST-78, PST-1, PST-127, and CYR32 and their relationships to other races based on a total of 10,750 bp sequences from seven genes, including a hypothetical conidiation protein, TATA binding protein, serine/threonine receptor kinase, hypothetical glucan 1,3-beta glucosidase, beta 3- tubulin (BT), elongation factor (EF), and MAP kinase (MAPK) (left) and based on BT, EF, and MAPK (right) (Liu et al. 2012). PST-78 (2K04-Yr9) is the isolate used for complete sequencing. PST-1, PST-127, CYR32 (an Asian isolate) and UK94/519 (an European isolate, not in the figure) are being sequenced using standard Illumina

Table 1. Genome size, GC content, and numbers of genes, tRNAs, and rRNAs of *Puccinia striiformis* f. sp. *tritici* in comparison with *P. triticina* and *P. graminis*

Item	<i>P. striiformis</i>	<i>P. triticina</i>		<i>P. graminis</i>
Isolate	PST-78	BBBD Race 1 (V1)	BBBD Race 1 (V2)	CRL 75-36-700-3
Accession	AJJV01000000	ADAS01000000	ADAS02000000	AAWC01000000
Coverage	81.00X	16.00X	31.00X	7.88X
Genome size (Mb)	117.31	162.95	135.34	88.64
GC (%)	44.43	46.34	46.72	43.35
No. of genes	19,542	11,638	14,878	15,800
No. of tRNAs	666	1,484	767	501
No. of rRNAs	4	72	47	13



Fig. 4. Comparison of the beta-tubulin proteins among *Puccinia striiformis* (PS) (CQM\_05340), *P. triticina* (PT) (PGTG\_09142), and *P. graminis* (PG) (PTTG\_02943).

## Conclusions

- The whole genome sequencing of *Puccinia striiformis* is completed by sequencing genomic DNA fragments covering 81X of the fungal genome using a US isolate.
- The sequenced genome of *P. striiformis* has a total of 117.31 Mb. The genome size is considerably bigger than that of *P. graminis*, but smaller than that of *P. triticina*.
- The *P. striiformis* genome contains 19,542 genes, more than that of *P. graminis* or *P. triticina*.
- The sequence database of *P. striiformis* is useful to study particular genes among various isolates and also for comparison with other *Puccinia* species to study rust biology.

[http://www.broadinstitute.org/annotation/genome/puccinia\\_group.9/MultiHome.html](http://www.broadinstitute.org/annotation/genome/puccinia_group.9/MultiHome.html)

## Reference

Liu, B., Chen, X. M., and Kang, Z. S. 2012. Gene sequencing reveals heterokaryotic variations and evolutionary mechanisms in *Puccinia striiformis*. *Open Journal of Genomics* 1:1.