

RESEARCH ARTICLE

Open Access

# Conserved loci of leaf and stem rust fungi of wheat share synteny interrupted by lineage-specific influx of repeat elements

John P Fellers<sup>1\*</sup>, Bahram M Soltani<sup>2</sup>, Myron Bruce<sup>1</sup>, Rob Linning<sup>2</sup>, Christina A Cuomo<sup>3</sup>, Les J Szabo<sup>4</sup> and Guus Bakkeren<sup>2</sup>

## Abstract

**Background:** Wheat leaf rust (*Puccinia triticina* Eriks; *Pt*) and stem rust fungi (*P. graminis* f.sp. *tritici*; *Pgt*) are significant economic pathogens having similar host ranges and life cycles, but different alternate hosts. The *Pt* genome, currently estimated at 135 Mb, is significantly larger than *Pgt*, at 88 Mb, but the reason for the expansion is unknown. Three genomic loci of *Pt* conserved proteins were characterized to gain insight into gene content, genome complexity and expansion.

**Results:** A bacterial artificial chromosome (BAC) library was made from *P. triticina* race 1, BBBB and probed with *Pt* homologs of genes encoding two predicted *Pgt* secreted effectors and a DNA marker mapping to a region of avirulence. Three BACs, 103 Kb, 112 Kb, and 166 Kb, were sequenced, assembled, and open reading frames were identified. Orthologous genes were identified in *Pgt* and local conservation of gene order (microsynteny) was observed. Pairwise protein identities ranged from 26 to 99%. One *Pt* BAC, containing a RAD18 ortholog, shares syntenic regions with two *Pgt* scaffolds, which could represent both haplotypes of *Pgt*. Gene sequence is diverged between the species as well as within the two haplotypes. In all three BAC clones, gene order is locally conserved, however, gene shuffling has occurred relative to *Pgt*. These regions are further diverged by differing insertion loci of LTR-retrotransposon, *Gypsy*, *Copia*, *Mutator*, and *Harbinger* mobile elements. Uncharacterized *Pt* open reading frames were also found; these proteins are high in lysine and similar to multiple proteins in *Pgt*.

**Conclusions:** The three *Pt* loci are conserved in gene order, with a range of gene sequence divergence. Conservation of predicted haustoria expressed secreted protein genes between *Pt* and *Pgt* is extended to the more distant poplar rust, *Melampsora larici-populina*. The loci also reveal that genome expansion in *Pt* is in part due to higher occurrence of repeat-elements in this species.

**Keywords:** Wheat leaf rust, BAC construction, Synteny

## Background

Plants and pathogens are in a constant struggle as each co-evolves to adapt to genomic changes. Plant genomes are adapting to different modes of infection by pathogens while pathogens are evolving different avenues to circumvent defense systems of their respective hosts. Rust fungi are among the most economically important pathogens, yet are part of elusive host-pathogen systems. The order

Pucciniales (formerly Uredinales or Urediniomycetes) contains over 7,000 different species from 100 genera [1]. Adding to the complexity, individual cereal crops can be infected by several rust fungi adapted to the specific crop.

Cereal rust fungi are obligate biotrophs and have alternate hosts where sexual recombination takes place, allowing for diversification of the population [2]. The life cycle of cereal rust fungi begins with a urediniospore landing on a leaf surface and germinating in the presence of adequate humidity. A germ tube emerges and moves towards a stomate via a thigmotrophic response and probable chemical clues [3] where an appressorium will

\* Correspondence: john.fellers@ars.usda.gov

<sup>1</sup>USDA-ARS, Hard Winter Wheat Genetics Research Unit, Department of Plant Pathology, Manhattan, KS 66506, USA

Full list of author information is available at the end of the article

form. A hypha grows inside the substomatal space until a mesophyll cell is encountered. The fungus will penetrate the cell wall and produce a haustorium by invagination of the plasma membrane [4,5]. At each stage of infection, the fungus is postulated to secrete effectors to inhibit cell defenses and reprogram cells to redirect nutrients. Though some candidate effectors are shared among the rust fungi, most are specific to their host and include transcription factors, zinc finger proteins, small secreted proteins and cysteine-rich proteins [6]. Certain classes of effectors, such as ones modulating host immunity, are believed to rapidly change to overcome resistance, however, the mechanisms generating this variation are not known. In several studied pathogens, certain classes of predicted effectors are found in variable and highly mutagenic regions of the genome. Mobile elements induced mutations in effectors in *Phytophthora* [7], *Magnaporthe* [8], and *Leptosphaeria* [9] while *Fusarium oxysporum* has a specialized chromosome with effectors [10,11]. Effectors can be clustered in the genome (*Ustilago*; [12]) including at telomeres (*Fusarium*, [13]; *Magnaporthe*; [14]). Avirulence genes from the flax rust fungus, *Melampsora lini* are all small secreted proteins [15,16]. Currently, two effectors have been identified in urediniospores of *Puccinia graminis* f.sp. *tritici* (*Pgt*) that induce the *in vivo* phosphorylation and degradation of the barley resistance protein, RPG1 [17].

Sequencing technology has made significant advancements in recent years. Complete genomes of more species, including fungi, are being sequenced. Comprehensive catalogs of genes can be generated, annotated, and comparisons made to other genomes. Core sets of genes needed for function, adaptations for life cycle, and host specificity can now be found. Comparisons of several obligate fungal plant parasites have identified common losses of genes involved in nitrate and sulfur metabolism [6,18]. *Melampsora larici-populina* (*Mlp*) and *Pgt* have approximately 8,000 orthologous genes which could be suggested as a core set needed for biotrophism. However, 74% and 84% of the secreted proteins, respectively, are lineage specific [6] suggesting proteins that are needed for the individual life cycle. Corn pathogens, *U. maydis* and *S. reilianum* are also closely related and share 71% of effector genes in so-called divergence clusters. However, 10% are *U. maydis* specific while 19% are specific to *S. reilianum* [19].

*Puccinia triticina* (*Pt*) is the causal agent of wheat leaf rust and new races emerge each year aided by a crop monoculture placing a strong selection pressure on the pathogen. Genetic variation is generally believed to increase through sexual recombination to generate new allele combinations. Two related wheat rust fungi, *Pgt* and *P. striiformis* f.sp. *tritici* (*Pst*), causing stripe rust, have a sexual cycle on North American *Berberis* spp. and

have a greater race variability where the alternate host is present [20]. The *Pt* aeciospore stage is on *Thalictrum* spp. and *Isopyrum*, found mainly in the Mediterranean region, however, other *Thalictrum* species present in North America can support a reduced level of infection [21] but are generally resistant to *Pt* [22,23]. Populations are essentially asexual; supported by the lack of recombination found in numerous North American races [24-28]. A parasexual cycle may exist allowing recombination since germ tube fusion, nuclear migration, and bridging structures between nuclei have been observed in *Pt* [29].

The obligate biotrophic nature of cereal rusts makes experimental manipulation difficult, however, genomics provides a means of studying evolution and gene function. We set out to understand the genome variation of two rust fungi at three regions. A *Pt* bacterial artificial chromosome (BAC) library was made and clones were identified using three probes that would isolate regions of predicted secreted proteins and avirulence. Sequenced DNA regions of *Pt* were compared to syntenic regions in two rust species with complete genome sequences, *Pgt*, and *Mlp* [6], and evaluated for genomic conservation, expansion and mutations.

## Results

### BAC library construction

Urediniospores harbor two haploid nuclei with an estimated total genome complexity for *Pt* of approximately 135 Mb, based on comparative DNA fluorescence (L. Szabo, unpublished) and the current total size of the genome assembly ([http://www.broadinstitute.org/annotation/genome/puccinia\\_group/GenomeStats.html](http://www.broadinstitute.org/annotation/genome/puccinia_group/GenomeStats.html)). The generated *P. triticina* BAC library contained 15,360 clones arrayed in 384 well plates with an average insert size of 105 kb representing an estimated 10 to 12 genome equivalents. A single-copy probe identified nine positive clones on high density filters, and assuming fragments were randomly cloned during library construction, this is in agreement with the estimated genome coverage.

### BAC clone selection, sequencing and characterization

Three genomic regions were targeted for comparison. Previous work had mapped a *Pgt* RAD18 homolog in a genomic region harboring an avirulence gene [30]. Using *Pgt*RAD18 as a reference, PT0313.J16.C21 (GenBank accession number GR497566) was identified from a *Pt* EST database using TBLASTN (E value = e-107; [31,32]) and used as a probe. Nine positive BAC clones were found and clone 1F16 (Pt1F16) was selected for sequencing because of its longer length and the centralized location of PtRAD18 within the BAC clone. Sequences from Pt1F16 were assembled into two contiguous sequences of 39,219 and 63,874 bp, totaling 103,093 bp (GenBank JX489506). The GC content of these sequences was 47%. Subclones

were generated spanning the gap for orientating and ordering of the two contigs. However, due to a region of 60 near-perfect 46 bp repeats of ACCAGCCCGCCGAGAG GAAGCCCTCTCGGCGAGCTGGTGTGTGTAT, the gap could not be closed. FGENESH, with gene models from the *Puccinia* group genome project ([http://www.broadinstitute.org/annotation/genome/puccinia\\_group/](http://www.broadinstitute.org/annotation/genome/puccinia_group/)), predicted 30 open reading frames (ORFs) ranging from 210 to 4,077 bp in length.

In a functional screen of *Uromyces viciae-fabae*, secreted peptide effector protein UF5 was related to the flax rust *Melampsora lini* haustorial-expressed secreted protein HESP-379 [16]. A *Pgt* genome search revealed several predicted secreted protein homologs in close proximity, suggesting the presence of small clusters of predicted secreted proteins. UF5 (Genbank ES608162) aligned with two predicted secreted proteins, PGTG\_03708 (E score  $5 e^{-33}$ ) and PGTG\_03709 (E score 0), both transcribed and located 513 bp apart on the *Pgt* contig. Using these *Pgt* sequences, PtContig18 (Genbank accession HP451841) and PtContig7347 (Genbank accession HP458556) were identified by a BLASTN *Pt* EST database search. A PCR product from the cDNA clone, *Pt* EST PT0061b.D10.TB that aligned to Contig18 (GenBank accession EC400508), was used as a probe to identify *Pt* BAC PtHSP02. Sequencing of this BAC resulted in four assembled contigs. Gaps could be spanned and thus the contigs could be ordered and oriented. Sizes of the contigs in bp were 16,991, 30,055, 5,014, and 60,277 for a total of 112,337 bp (GenBank JX489507). Gaps were present in regions of repeated DNA and could not be assembled. GC content was 46.3% and FGENESH predicted 31 ORFs in the contig ranging from 174 bp to 7,167 bp in length. The smaller ORFs were generally within repeated elements.

The bean rust effector UfHSP42c Uf011 (GenBank ES608167; [33]) matched three predicted protein sequences in *Pgt*, PGTG\_17547 (E Score 0), PGTG\_17548 (E Score  $1 e^{-21}$ ) and PGTG\_17549 (E Score  $1 e^{-4}$ ). UfHSP42c matched five *Pt* ESTs, including clone PT0131d.B10.BR (GenBank accession EC414978) from which probes were derived to identify *Pt* BAC clone HSP04. Sequencing of HSP04 produced two contiguous sequences of 9,276 bp and 157,027 bp for a total of 166,303 bp (GenBank JX489508). GC content was 46.3% and 61 ORFs were predicted ranging from 120 bp to 5,214 bp in length.

#### BAC annotation

The predicted ORFs from each BAC clone were aligned using BLASTN to the *Pgt* genome, *Pgt* predicted transcripts and *Pt* ESTs, and using BLASTX, to the *Pgt*, *Mlp*, and *U. maydis* (*Um*) predicted proteomes (Table 1). Pt1F16 had nine ORFs with synteny in *Pgt*. Identity across the protein sequences ranged from 37-87% in

these alignments and putative annotations could be assigned to five of the proteins. Pt1F16-4 contained many gaps when compared to PGTG\_13013. Proteins Pt1F16-5, 6, 7, 8 and 9 aligned with two proteins each from *Pgt*. Pt1F16-7 aligned with PgtRAD18, which has one copy in each of the *Pgt* haplotype genomes. All but one homolog could also be found in *Mlp* and four were represented in *Um* (Table 1).

Nine predicted proteins in PtHSP02 were confirmed through EST sequence alignment [32] and a putative function could be assigned to eight of them. Alignment identity ranged from 30-100% in PtHSP02. Eight homologs could be found in both *Mlp* and *Um* in PtHSP02. The most highly conserved protein is PtHSP02-6, a G-protein  $\beta$ -subunit containing a conserved WD-40 repeat motif. The first 343 amino acids were 100% identical to PGTG\_03727 and 99% to *Mlp* accession GL883091 (Table 1). Conversely, PtHSP02-3 was only 30% identical to PGTG\_3706 and had no homologs in the other two fungi. PtHSP02-4 and PtHSP02-5 aligned with *Mlp* HESP-379, the haustorial expressed predicted secreted protein homolog from *M. lini*, and a homolog was found for each in *Pgt* (Table 1, Figure 1). Two insertions/deletions were found in PtHSP02-4 and PGTG\_3708 (Figure 1A). PtHSP02-5 and PGTG\_3709 aligned to homologs from *M. lini*, *Mlp*, *M. medusae deltoidis*, and *U. maydis*. The N-terminal half of the protein was conserved between *Puccinia* and *Melampsora* (Figure 1B). There appeared to be 48 genus-specific amino acid changes across the protein. *Um* was the most diverged with only a few conserved motifs.

Fourteen predicted proteins were identified in PtHSP04 and could be supported through EST sequence alignment. Every protein had a homolog in *Pgt* with protein identities ranging from 26-95% (Table 1); nine could be assigned a putative function. Eight PtHSP04 proteins had homologs in *Mlp* and five in *Um*. PtHSP04-1, 5, and 14 appeared to be unique to *Pt* with little homology to *Pgt*. The predicted transcripts of PtHSP04-6, 7, 8 and 9 aligned to a single EST of *P. striiformis* predicted to encode a secreted protein (ADA54575; [34]) at scores of  $4 e^{-5}$ ,  $2 e^{-8}$ ,  $6 e^{-48}$ , and  $3 e^{-9}$ , respectively. PtHSP04-6 and 7 aligned both to PGTG\_17549, though revealing 26 and 60% identity, respectively. The predicted HSP04-7 ORF is 1,095 bp in length and contains a 3' in-frame repeat of nine nucleotides, GG(C/T) AC(T/A) AC(T/A), translating to 30, three amino acid repeats of Gly-Thr-Thr. Without the repeat, PtHSP04-7 is a homolog to PGTG\_17549, while PtHSP04-6 is unique to *Pt*. PtHSP04-8 and 9 are responsible for the homology to Uf-HSP42c and isolation of the BAC clone (Figure 2, Table 1). They are very highly identical except for the C-terminal 18 amino acids, where PtHSP04-9 has a five amino acid deletion and only four identities (Figure 2). Each aligned to PGTG\_17547 and PGTG\_17548, adjacent proteins which themselves are 100% identical. PtHSP04-8

**Table 1 Gene features in three *Puccinia triticina* BAC clones and their alignment to other sequenced Basidiomycetes, *P. graminis tritici* (*Pgt*), *Melampsora larici-populina* (*Mlp*), and *Ustilago maydis* (*Um*)**

<i>P. triticina</i> ORF	<i>Pgt</i> Gencxxe Feature*	<i>Pgt</i> protein identities	<i>Mlp</i> gene feature**	<i>Um</i> gene feature	BLASTX Annotation
1 F16-1	PGTG_12990	54%	35.90208/11	-	Similar to Uhrf1
1 F16-2	PGTG_13012	60%	104.11397	Um_00786	Hypothetical protein
1 F16-3	PGTG_13013	63%	74. 94858	-	Similar to esterase
1 F16-4	PGTG_13016	37%	-	-	Predicted protein
	PGTG_18731	64%			
1 F16-5	PGTG_13018	67%	49.39112	Um_02725	Similar to molybdopterin synthase sulpherylase
	PGTG_18732	68%			
1 F16-6	PGTG_13021	68%	74.73436	-	Hypothetical protein
	PGTG_18735	64%			
1 F16-7	PGTG_13023	56%	74.94864	Um_05085	RAD18
	PGTG_18741	56%			
1 F16-8	PGTG_13024	64%	74.000024	-	cystein rich SCP-like extracellular protein
	PGTG_18744	65%			
1 F16-9	PGTG_13026	87%	74.50754/28052	Um_00594	Similar to pyruvate dehydrogenase complex
	PGTG_18746	86%			
HSP02-1	PGTG_3730/1	79%	22.87674	Um_00736	Conserved protein
HSP02-2	PGTG_6672	48%	68.40205	Um_04270	Aspartyl-tRNA synthetase
HSP02-3	PGTG_3706	30%	-	-	-
HSP02-4	PGTG_3708	69%	2.70587	Um_01555	<i>Mlp</i> HESP-379
HSP02-5	PGTG_3709	83%	2.70587	Um_01555	<i>Mlp</i> HESP-379
HSP02-6	PGTG_3727	100%	2.76428	Um_00703	G-protein beta subunit
HSP02-7	PGTG_3728	68%	2.115002	Um_03486	Nucleotide-binding protein 2
HSP02-8	PGTG_3729	82%	2.46419	Um_05743	Pre-mRNA splicing factor ATP- dependent RNA helicase PRP16
HSP02-9	PGTG_3730	87%	10.115914	Um_04551	Similar to cyclin <i>Ctk2</i>
HSP04-1	PGTG_16978	36%	-	-	Predicted protein
HSP04-2	PGTG_16976	73%	27.88522	Um_00639	Nucleoporin-like
HSP04-3	PGTG_10949	95%	47.72927	Um_02479	60S ribosomal protein
HSP04-4	PGTG_02586	33%	27.88520	-	Heat shock protein 90
HSP04-5	PGTG_14539	41%	-	-	Predicted protein
HSP04-6	PGTG_17549	26%	-	-	Predicted secreted protein
HSP04-7	PGTG_17549	60%	-	-	Predicted secreted protein
HSP04-8	PGTG_17547/8	76%	16.85997	-	<i>Uf</i> -HSP42c
HSP04-9	PGTG_17547/8	71%	16.85997	-	<i>Uf</i> -HSP42c
HSP04-10	PGTG_05205	82%	22.48630	-	Integral membrane protein
HSP04-11	PGTG_17545	52%	-	Um_00662	Predicted protein
HSP04-12	PGTG_17544	83%	23.87824	Um_03820	Vacuolar sorting protein PEP5
HSP04-13	PGTG_17543	62%	23.72100	Um_02189	Hypothetical protein
HSP04-14	PGTG_17537/8	38%	-	-	Predicted protein

\* *Pgt* gene features indicate gene nUmbere assigned by the Broad institute during assembly ([http://www.broadinstitute.org/annotation/genome/puccinia\\_group/](http://www.broadinstitute.org/annotation/genome/puccinia_group/) verified September 28, 2012). Two nUmbere indicate different scaffolds, as highlighted in Figure 1.

\*\* *Mlp* gene features are indicated by scaffold nUmbere; gene nUmbere as assigned by the Joint Genome Initiative at the Department of Energy (<http://genomeportal.jgi-psf.org/Mellp1/Mellp1.home.html> verified September 28, 2012).

**A**

Pt HSP02-4 1 MQFFTAISGLVALSSGVTGKEIAVYD T LSVTSPSGSSYWVQFSTNTIAWA 50  
 PGTG\_3708 1 MQFYTALLGLVALSSGVT - - - - - A L S I T S P S G S S Y W V Q F S T N T I A - - 40

Pt HSP02-4 51 LSSGDPTSVS IQIVNPTNKQFNGPFSIAEYVPAAKM SYDVTNVQLIVGDG 100  
 PGTG\_3708 41 - - - - - C I Q I I N P T N K Q F N G P F S I A E Y V P A A K K S Y D V T N V Q L I I A D G 81

Pt HSP02-4 101 YQVQFVSP LNNTQVYLATSAPFSVKVSGTPEAQVTYAPGTHDSAS - - NSS 147  
 PGTG\_3708 82 YE VQFVN PLNNT E V Y A T S A P F S V K T S G T A P A Q V T Y A P G T H E S S N S S N N S S 131

Pt HSP02-4 148 SNSTSSNSSTSHPDINGNSFKNNGTANANAKSA SVRAAAVPSMAVFAIG 197  
 PGTG\_3708 132 S N A N T S S K N S T M N A D I N G N S F K N S G N A N T N A K S S S V R A A A M P S I V A L A L G 181

Pt HSP02-4 198 SLMALIA 204  
 PGTG\_3708 182 SLCALVA 188

**B**

Pt HSP02-5 1 MRFLNLFGLA GLTASAAAFSTINSP - SASSYWVQFA T N T I A W S N A P N D S P 49  
 PGTG\_3709 1 MRS L N L L I S L A G L I A S A A A F T I N S P - S A S S Y W V Q F A T N T I A W T N A P N D S P 49  
 M. lini HESP-379aa 1 MRTSA LLTGLVSLAAS SLAFTVNGP - SASSYWVQFS T N T I S W T Q G S G D P P 49  
 M. medusae deltoideis HESP-379 1 MRTST LLTGLXSLAAS SLAFTVNGP - SASSYWVQFS T N T I S W S T G S G D P P 49  
 M. laricis-populina 70587 1 MRTSA FLTGLVSLAAS SLAFTVNGP - S A A S Y W V Q F S T N T I S W S Q G S G D P P 49  
 U. maydis UM01555 1 MQIAAA I T A L A S V Y G L S Q A F E I T F P N S N G G Y W V T N Y T N T L N W K A N S S D P S 50

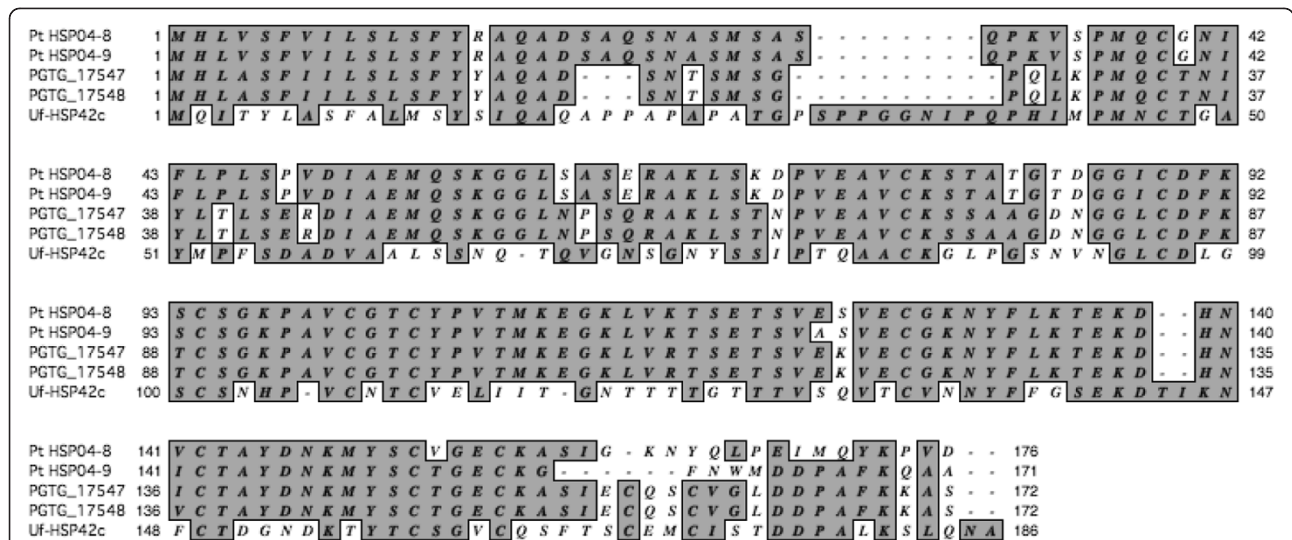
Pt HSP02-5 50 QVTLQIVNSNHTLLNGVFSIAEYVPAALEAYTVTNVTLVVDGYQVQMVN 99  
 PGTG\_3709 50 QVTLQIVNSNHTLLNGVFSIAEYVPAALEAYTVTNVTLVVDGYI VQMVN 99  
 M. lini HESP-379aa 50 VVSLQIINP SKNLLNGVFSIAEYVKA SLESYTVTNVTLVVDGYQVQMVN 99  
 M. medusae deltoideis HESP-379 50 VVTLQIVNSNRNLLNGVFSIAEYVKA SLESYTVTNVTLVVDGYQVQMVN 99  
 M. laricis-populina 70587 50 VVSLQIVNP NHNLLNGVFSIAEYVKA SLESYTVTNVTLVVDGYQVQMVN 99  
 U. maydis UM01555 51 FFSVQLLNANSSQLNGNFQIGNAL LTANGTAQIF - I D R I P S G T Y S L L F V N 99

Pt HSP02-5 100 PAN - - - - S L V Y A T S S P F S V K S S D T P P A P G S N I S P P S S L S G M G G T N G D M R 145  
 PGTG\_3709 100 PAN - - - - S S Q V Y A S S S P F S V K P S D T P A A P G S G P G A P S S L S G L P G T S N D M R 145  
 M. lini HESP-379aa 100 STN - - - - A T Q V Y A S S G A F S V K D A N T T A A P V A T S S G S G S - N G S G S A T G - T N 143  
 M. medusae deltoideis HESP-379 100 STN - - - - A T Q V Y A S S G T F S V K D A T T T P A P V A S T S N G S G - N G S G N G S G T T A 144  
 M. laricis-populina 70587 100 STN - - - - A T Q V Y A S S G T F S V K D A N T T A A P V A S T S G S G S A N G S G N G S G S T T 145  
 U. maydis UM01555 100 S S N Y E L D R P Q V Y Y T S S S F E V K P N G T Q P A E V T A N T I N A D P - - - - - N S S A T T 143

Pt HSP02-5 146 NMTGMSGMPG SATN GRPGNTTGFNNNSN NNGTANSNARS DA SSVYLF SFA 195  
 PGTG\_3709 146 NMTGMSGMPGN NMSG R P G N T T G F S N N A N - - G T A N S N A R S D A S I V S P I S F A 193  
 M. lini HESP-379aa 144 GN S A N G S S M T S N T T S T S S G T R T A G S G A G - N S T A G S N A K S S A T I S N P - A M G 191  
 M. medusae deltoideis HESP-379 145 GTNGNSG SMT SNA T - T A T G A R T L G S G A G - N S T A G T N A K S S A T M S S P - A M G 191  
 M. laricis-populina 70587 146 GTNGNSG S T T S N T T S T S S G A R T A G S G A G - N S T A G S N A K S S A T I L S P - A M G 193  
 U. maydis UM01555 144 A A T G S A T A T L S S A T G S S T I S A R P S N G S S N - - - - S G S S M I Q G S T V A V V G S L Y 189

Pt HSP02-5 196 TWSIVAS AIGFVTL 209  
 PGTG\_3709 194 VLSVVA SAIGL A TL 207  
 M. lini HESP-379aa 192 LLTVS LAAIGFASL 205  
 M. medusae deltoideis HESP-379 192 LLTVS LAA MGFVSL 205  
 M. laricis-populina 70587 194 LFTVSLAAIGFASL 207  
 U. maydis UM01555 190 SAVLGLGAGLLV - - 201

**Figure 1 ClustalW alignments of secreted proteins coded on BAC PtHSP02. A) PtHSP02-4 aligned to a Pgt homolog. B) PtHSP02-5 aligned to homologs from Pgt, Melampsora lini, M. medusae deltoideis, Mlp and U. maydis.**



**Figure 2** ClustalW alignments of two predicted secreted proteins coded on BAC clone PtHSP04. PtHSP04-8 and 9 are aligned to homologs from *Pgt* and *Uromyces fabae*.

and 9 are 76% and 71% identical to PGTG\_17547, respectively (Figure 2, Table 1).

### Repetitive elements and repeated sequences

Each BAC was evaluated for repeat elements by using REPBASE against *Pgt*, *Pt* and *Pst* genomes. Complete and incomplete terminal inverted repeats (TIR), LTRs, *Copia*, *Gypsy*, *Mariner*, *Mutator*, *Harbinger*, *Helitron*, hAT, and DNA transposons were found. (Additional file 1: Tables S1 and S2). Major insertions are represented in Figure 1. *Copia* elements were found inserted within *Gypsy* elements in Pt1F16 and PtHSP02. PtHSP02 and PtHSP04 also had localization of LTRs.

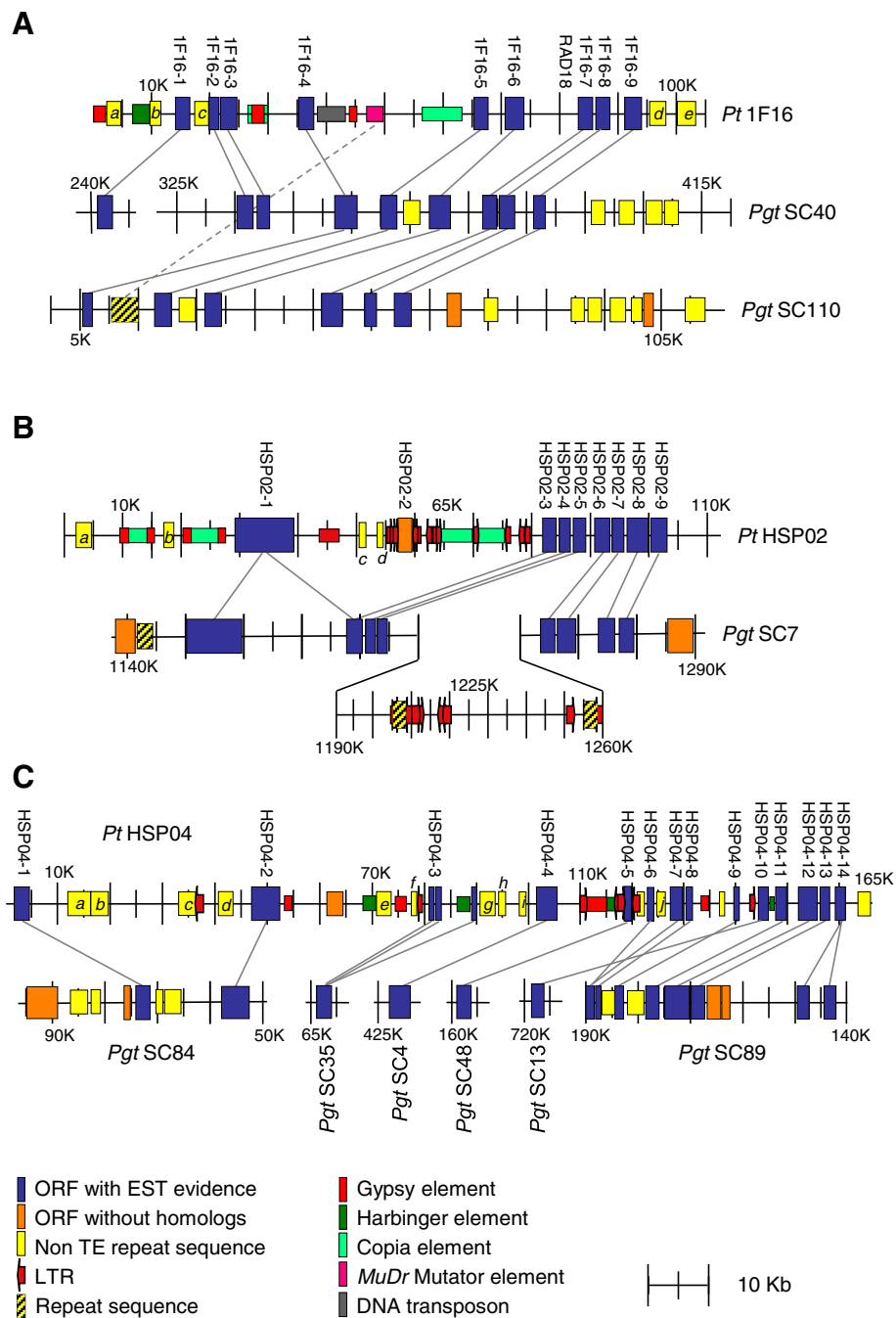
### Synten

To investigate whether the high number of candidate orthologs with *Pgt* maintained the same gene order, the *Pt* BAC sequences were aligned to the available *Pgt* contig sequences. Figure 3 graphically represents the location along each BAC clone of *Pt* ORFs with EST sequence or protein homology support. The majority of Pt1F16 aligned to the 325,000 bp to 415,000 bp region of *Pgt* scaffold (SC) 40 but also to the 5,000 to 65,000 bp region of PgtSC110. PgtSC40 and PgtSC110 could either represent the two *Pgt* haplotypes or a duplication of this region in the genome. Overall, gene order was maintained in both scaffolds. As previously noted, eight of the Pt1F16 ORFs aligned to homologs in *Pgt* but Pt1F16-1 to 3 were found only on PgtSC40 (Table 1, Figure 3A). Pt1F16-1 aligned to PGTG\_12990 85 kb upstream in SC40 of PGTG\_13012 whereas Pt1F16-2 and 3 were similarly spaced as their counterparts on this *Pgt* SC. Between Pt1F16-4 and 5, four retrotransposons were found, of

which one was similar to a retroelement in PgtSC110. No mobile elements were found in this region on PgtSC40. PtRAD18 (Pt1F16-7) is a single ORF while Pt1F16-8 aligned to an ORF corresponding a cysteine rich SCP family protein in both SCs of *Pgt*.

PtHSP02 aligned to a single scaffold, PgtSC7 (Figure 3B). A second haplotype was not detected as the *Pgt* assembly represents most loci with a single sequence [6]. Nine *Pt* ORFs could be aligned to homologs on PgtSC7 (position 1,135,000 to 1,280,000). As with the other BAC clones, gene order was generally maintained. However, PtHSP02-1 and PtHSP02-2 were found embedded between retroelements and LTRs. While PtHSP02-1 aligned to two fragments on PgtSC7, PtHSP02-2 was 48% homologous to a gene on PgtSC15 elsewhere in the genome. The remaining genes in PtHSP02 were in the same order as on PgtSC7, except a large insertion of approximately 70 kb of DNA, including sequence similar to mobile elements, was found between PGTG\_03709 and PGTG\_03708 on PgtSC7. Additional PgtSC7 DNA insertions were evident within this gene cluster whereas the *Pt* homologs were packed in a tighter arrangement. Across this region, a higher number of retrotransposon elements were found on PtHSP02 (Additional file 1: Table S1).

PtHSP04 aligned to at least six regions within the *Pgt* genome and represents the least syntenic sequence amongst the three BAC clones (Figure 3C). PtHSP04-1 and 2 were found on PgtSC84, however, there were several repeat elements within both the *Pt* and *Pgt* regions. PtHSP04-3 appeared to be a fragmented ORF because a single homologous ORF was found on PgtSC35. PtHSP04-4 and 5 were found on two separate scaffolds, PgtSC4 and PgtSC48, respectively. PtHSP-6, 7,



**Figure 3** Graphical representation of three BAC clones from *P. tritricina* Pt1F16 (A), PtHSP02 (B) and PtHSP04 (C) and their synteny to super contigs (SC) of *P. graminis tritici* (*Pgt*). Lines connect homologs between genomes.

8, and 9 have homologs on *Pgt* SC89 in the same order and similar gene distance (conserved micro-synteny). PtHSP04-10, flanked by an LTR and *Harbinger* element, does not have a homolog on *Pgt*SC89, but on *Pgt*SC13. Microsynteny of PtHSP04-11, 12, and 13 to *Pgt* is maintained. PtHSP04-14 is a single copy gene in *Pt* but is repeated in *Pgt*. between BAC positions 60,000 and 125,000 there are a high number repeat elements.

One of the most interesting sets of sequences were *Pt* ORFs for which numerous homologous copies were found in the *Pgt* genome but were not classified as typical mobile elements (identified by small letters in Figure 3; Table 2). Twenty of these ORFs had repeats in the *Pgt* genome numbering from 19 to 474. Table 2 lists the conserved amino acid domains, if present, in each of the ORFs and the percent identity, which, ranged from

**Table 2 Detailed analysis of non-transposable element, repeated sequences in three *P. triticina* BAC clones**

ORF	Size bp	<i>Pgt</i> repeats	Conserved domain	% Ident	Exp*	Dominant peptide	Notes
1F16-a	757	35	194-443	51	no	9.2% Lys	216-252-40.3% CI Winged helix DNA/RNA binding
1F16-b	458	98	97-458	40	no	11.8% Lys	Highly helical
1F16-c	462	98	all	40	no	13.2% Lys	6-91-60.3% CI ubiquitin ligase
1F-16-d	489	44	344-485	48	no	10.8% Lys 13.0% Ser	3 alpha helices and 7 beta sheets in conserved domain
1F16-f	651	80	51-245	74	yes	9.2% Lys	38-163-38.7% CI oxidoreductase
HSP02-a	658	19	all	31	no	9.7% Lys	
HSP02-b	299	80	35-94	68	no	9.4% Lys 11.7% Ser	32-73-80% CI metal binding protein 145-256-80% CI protease hydrolase inhibitor
HSP02-c	252	50	33-127	52	yes	11.1% Lys 10.7% Glu	33-95 53% CI DNA binding domain
HSP02	243	35	all		yes	11.9% Ala 13.6% Ser 11.5% Thr	Alignment in <i>Pgt</i> are to DNA, not protein
HSP04-a	952	74	256-470	48	no	9.9% Ala	
HSP04-b	484	76	300-454	74	no	none	
HSP04-c	442	69	all	34	yes	9.0% Lys	
HSP04-d	679	81	10-351	68	no	9.3% Lys	11 alpha helices in conserved region
HSP04-e	420	131	55-386	55	yes	9.5% ala	216-388-51% identical in Mlp
HSP04-f	262	474	20-189	44	no	11.5% ser	EST sequence hits
HSP04-g	543	96	all	69	yes	11.4% Lys 9.0% ser	Highly helical
HSP04-h	262	66	199-258	65	yes	12.6% Lys	P-loop NTPase/DEAD box
HSP04-i	326	20	127-242	34	yes	12.8% Lys 11.0% Glu	Highly helical
HSP04-j	309	69	191-303	35	no	9.7% Lys 10.7% Ser	2-191-99.7% CI Recombinant DNA binding/ RAD54 like

Each was evaluated for number of repeated homologs in the *Pgt* genome, region of conservation, alignment identity, expression in infected tissue, the dominant peptides, and their best detected fit to a protein structural model.

\* Exp - expression in infected tissue 6 dpi.

\*\* PHYRE 2.0 confidence interval (CI) of protein segment matching a structural model with the listed function [35].



34-74%. Each ORF was compared to an RNAseq cDNA library of *Pt*-infected leaf tissue (Fellers and Bruce, unpublished) and nine aligned to the experimental cDNA sequences. The predicted proteins were analyzed for peptide content and most had an abundance of Lys, which is suggestive of helical structures. Each of the proteins was also compared to the PHYRE 2.0 structural data base [35] resulting in seven that revealed regions that aligned, with confidence, to known structures. The first 191 peptides of PtHSP04-j had a structure similar to RAD54, with 99.7% confidence. Of note, PtHSP04-e was expressed and was 51% identical to a protein in *Mlp*.

### Discussion

This study was performed to look at three regions of the *Pt* genome that were hypothesized to be under selection pressure because of the presence of putative secreted proteins or loci associated with avirulence. To begin with, gene order is conserved between the *Pt* BACs and *Pgt*. However, there is a wide range of protein conservation. A previous comparison of ESTs of *Pt* and *Pgt* found a similar level of variation in sequence, but only 40% of the *Pt* EST unigenes had orthologs in *Pgt* [32]. Many genes were likely missing in the unigene set because of the difficulty of sampling other *Pt* life stages to sufficient depth, affecting the percentage. Nevertheless, within the BAC clones, many protein identities were supported by ESTs and similar sequence variation was present [32]. Some proteins were highly conserved between the two wheat rust fungi and had homologs in *Mlp* and *Um* [36].

The three genes used for identifying the BACs were of most interest, in particular, the amount of variation within the sequence. PgtRAD18 had been associated with an avirulence locus in *Pgt* [30]. PtRAD18 protein length is relatively similar but the sequence has diverged from the PgtRAD18 with only 56% identity. Structurally, PtRAD18 is still closely associated with a predicted secreted protein. *Pt* has two genes similar to HESP-379 from *M. lini* [16]. Two indels in PtHSP02-4 suggest a recombination event or splicing difference evolved since the two species diverged, while the sequence differences in the C-terminus of PtHSP02-5 suggest that this region could be very variable. PtHSP04 contained a four-gene locus predicted to code for secreted proteins. Two of them are unique while two are recently duplicated paralogs. Secreted proteins are believed to be most variable amongst fungal proteins because they are under the highest selection pressure to avoid recognition by the host [16,19,37]. At least with these examples, It can be said that sequence variation, recombination, and duplication are driving the changes in these proteins.

Numerous fungal genomes have recently been generated, analyzed, and published. Now comparisons can be

made to find core gene families associated with specific life styles and cycles. In an extensive comparison, Duplessis et al. [6] identified core conserved genes needed for biotrophic life in both rust species. It appears that PtHSP02-6 may be one of those genes. PtHSP02-6 aligns with a G-protein beta subunit (GPBS) and no peptide differences were found between *Pt* and *Pgt*. Furthermore, there is little difference between *Pt* and *Mlp* suggesting that this protein is under strong purifying selection in rusts. Yet, the genes flanking PtHSP02-6 are relatively conserved indicating strong selection and the importance of this gene. In *Verticillium dahliae*, mutations in GPBS had reduced virulence, increased microsclerotia and conidiation and decreased ethylene production [38]. GPBS is also involved in similar functions in *F. oxysporum* [39]. In *M. grisea*, GPBS mutants could not form appressorium, and hyphae could not penetrate and grow in rice leaves [40]. The authors also showed that by over expressing GPBS in the fungus, appressorium could form on a hydrophilic surfaces suggesting that GPBS is necessary for control of surface recognition, growth and appressorium formation [40]. Surface recognition and appressorium formation are the key to rust fungal establishment. This suggests that PtHSP02-6 is indispensable for the biotrophic lifecycle and could be a regulating link in pathogenicity.

A strong correlation between genome size and repetitive element content has been found for many fungal genomes. Genome expansion is significant between *Pt* and *Pgt*, even though they are both closely related and are both dikaryotic. The assembled genome for *Pgt* is 89 Mb [6] while *Pt* is currently estimated to be 135 Mb (Broad Institute). The sequence analysis of the three BAC clones gives some indication on why the *Pt* genome may be larger than the *Pgt* genome. Pt1F16 had the least mobile element complexity, but had *Gypsy* elements within *Copia* elements, as did PtHSP02. PtHSP02 also harbored numerous TEs and LTRs in the region between PtHSP02-1 and 3. Meanwhile, PtHSP04 contains more non-TE repeat ORFs, its homologous genes are scattered across *Pgt* scaffolds, and its sequence reveals recombination and/or transposition events disrupting syntenic genes. There is also evidence of gene movement by active elements. PtHSP02-2 was directly flanked by LTRs and was not found in PgtSC7, PtHSP04-5 was also flanked by LTRs and could be found in PgtSC48, and PtHSP04-10 only had a single LTR flanking it, but was flanked on the opposite side by a partial *Harbinger* element. It is possible that since these regions are in repetitive sequence there are assembly errors in *Pgt*, however, each *Pgt* homolog are in high confidence scaffolds.

Most surprising are the non-transposable element, repeated sequences found in the *Pt* BACs (Table 2). Each

had homologs throughout the *Pgt* genome. Most had conserved domains that were maintained, while flanking sequences were greatly diverged. Many were high in Lys suggesting a helix protein structure. Some are expressed, based on the presence of an aligning EST, and have homologs in *Mlp*, suggesting an importance. The helical nature of these proteins would suggest their involvement as nucleotide binding elements. *Pt* has five different spore types in its lifecycle involving two different hosts requiring a significant level of cell modifications and cell types. Sequences like these have not been described before and could represent undiscovered elements in the disease cycle.

This work has shown significant genome synteny between two closely related wheat rust fungi. Gene sequences confirmed previous findings of the existence of EST sequence variation between *Pt* and *Pgt*. Various levels of homologies are present, but many of the genes are diverging in a manner that is species specific [32]. Both genomes have a significant amount of mobile elements. Some TE copies are conserved between the two species suggesting ancestral insertion. The insertion of TE sequences helps explain genome expansion, and their insertion near secreted protein genes may alter their regulation or cause their duplication and spread or deletion. Most surprising was the presence of small predicted non-TE genes with numerous homologs in *Pgt*. As many of the small repeated sequences are highly helical in predicted structure, one could suggest they are involved in DNA binding and regulation. Further work is needed to determine when they are expressed and at what stage of the life cycle. When analysis of the *Pt* and *Pst* genomes has been concluded, it can be determined if the repeated nature of these predicted genes is maintained within the wheat rust fungi.

## Methods

### *Pt* BAC library

Total genomic DNA for the BAC library construction was isolated from *P. triticina* (*Pt*) Race1, BBBB [41] urediniospores collected from susceptible wheat (*Triticum aestivum* L.) cultivar Thatcher. Spores were increased on plants spray-inoculated with a urediniospore suspension in light mineral oil (Soltrol 170 isoparaffin, Conoco-Phillips Chemical Co, Borger TX). The oil was allowed to evaporate for 30 min, then plants were moved to a dark dew chamber at 20°C and 100% relative humidity for 24 hrs for urediniospore germination and appressorium formation. Plants were grown in a growth chamber under 16-hour day at 20°C. After 10 days, urediniospores were collected and germinated by densely dusting them over sterile water in dishes for 8 hrs using a volatile nonanol solution (1.56 µl nonanol (Sigma-Aldrich, St. Louis MO), 1 ml acetone, 19 ml of ddH<sub>2</sub>O) spotted on filter paper which was suspended in the

lids to stimulate urediniospore germination under crowded conditions. The BAC library was constructed by BioS&T (Montreal, Quebec, Canada; www.biost.com). In brief, nuclei were isolated from collected germinated urediniospores and embedded in 1% low melting point agarose plugs. Total genomic DNA embedded in the plugs was partially digested with *Hind*III, separated by electrophoresis by pulse field gel electrophoresis, and the 100–200 kb region was isolated. After electro-elution and dialysis, the DNA fragments were cloned into the *Hind*III site of BAC vector pIndogoBAC5 (Epicenter Technologies, Madison, WI) and propagated in *E. coli* DH10B (Life Technologies, Grand Island, NY).

### BAC clone selection and sequencing

The resulting BAC library of 15,360 individual clones was arrayed on nylon membranes. After colony lysis, DNA was bound to the membranes using standard procedures [41]. BAC filters were probed to identify clones for sequencing. Several candidate fragments were selected as probes. The *Sfi*I insert from a *Pt* cDNA clone, PT0313.J16.C21 (GenBank accession GR497566; [32]) was labeled with αP<sup>32</sup>-dCTP using a random primer labeling kit (GE Healthcare, Pittsburg, PA). Positive BAC clones were verified by PCR using primers Forward 5'-AGCTCTTCAC ACGATTCC and Reverse 5'-ATCTTGGCATTGAGC ATC. The second probe, SP02, was amplified from *Pt* cDNA clone PT0061b.D10.TB (GenBank accession EC400508) by PCR using primers Forward 5'-CTTTCTA GACCTAGGCAACTTAACAC and Reverse 5'-GCGCC ATGGACTAGTTGAAGAGGGA. The third probe, *SP04* was amplified from cDNA clone, PT0131d.B10.BR (GenBank accession EC414978) using PCR primers Forward 5'-CACGAGGGGAACCGATGGGGGT and Reverse 5'-TGGGTTGGTAAACTATTAATGTGCAC. Southern hybridizations were as described [41]

Selected BAC clones were sent as a stab culture to the Genome Center at Washington University, St. Louis, MO. BAC clones were cultured, subcloned, shot gun sequenced, and assembled (Washington University Genome Center, St. Louis, MO). Gene calls were made using FGENESH with gene models specific to *Puccinia* (<http://linux1.softberry.com/berry.phtml>); BAC clone gene predictions were compared to *Pgt*, *Mlp* and *Um* genomic resources ([http://www.broadinstitute.org/annotation/genome/puccinia\\_group/Blast.html](http://www.broadinstitute.org/annotation/genome/puccinia_group/Blast.html); <http://genomeportal.jgi-psf.org/Mellp1/Mellp1.home.html> verified May 7, 2012 and [http://www.broadinstitute.org/annotation/genome/ustilago\\_maydis](http://www.broadinstitute.org/annotation/genome/ustilago_maydis)) using the BLASTN and BLASTX algorithms with settings of E value = 1e<sup>-3</sup>, Matrix = BLOSUM62, and gapped alignment. Repeats were identified using fungaldB of RepBase 17.04 [42], containing the repeats of *Pt*, *Pgt* and *Pst*. Long terminal repeats (LTR) were determined by LTR\_Finder [43] (e.g., red arrows in Figure 3).

*Mention of a trademark of a proprietary product does not constitute a guarantee of warranty of the product by the United States Department of Agriculture, and does not imply its approval to the exclusion of other products that may also be suitable. USDA is an equal opportunity provider and employer.*

## Additional file

**Additional file 1: Table S1.** Repeat elements found in three *P. triticina* BAC clones. BAC sequences were compared to RepBASE, a database containing characterized repeat elements from *P. triticina* (Pt), *P. graminis tritici* (Pgt), and *P. striiformis tritici* (Pst). Repeats are listed by position in the respective Pt BAC clone, DNA strand, and the species specific element. **Table S2.** Description of matching repeats, type of element and which rust fungus they are from.

## Abbreviations

BAC: Bacterial artificial chromosome; Pt: *Puccinia triticina*; Pgt: *P. graminis tritici*; Pst: *P. striiformis f.sp. tritici*; Mlp: *Melampsora larici-populina*; bp: Base pair; HESP: Haustoria expressed secreted protein.

## Competing interests

The authors declare that they have no competing interests.

## Authors' contributions

GB and JPF conceived the study. GB provided the DNA and constructed the BAC library. BMS and RL designed the primers and isolated the BACs. JPF sequenced the BACs, analyzed the sequence and made gene calls along with MB. CAC analyzed the sequence for repetitive sequence. LS provided the probes for RAD18. JPF, GB, and CAC wrote the manuscript. All authors read and approved the final manuscript.

## Acknowledgements

The authors would like to thank Drs. Robert Bowden and Randy Dinkens for their comments in the preparation of the manuscript. JPF is supported by USDA-ARS CRIS 5430-21000-005-00.

## Author details

<sup>1</sup>USDA-ARS, Hard Winter Wheat Genetics Research Unit, Department of Plant Pathology, Manhattan, KS 66506, USA. <sup>2</sup>Agriculture & Agri-Food Canada, Pacific Agri-Food Research Centre, Summerland, BC V0H 1Z0, Canada. <sup>3</sup>Broad Institute, 7 Cambridge Center, Cambridge, MA 02142, UK. <sup>4</sup>USDA-ARS, Cereal Disease Laboratory, 1551 Lindig, St. Paul, MN 55108, UK.

Received: 28 September 2012 Accepted: 11 January 2013

Published: 29 January 2013

## References

- Cummins GB, Hiratsuka Y: *Illustrated genera of rust fungi*. 3rd edition. St Paul: American Phytopathological Society; 2003:225.
- Liu JQ, Kolmer JA: Molecular and virulence diversity and linkage disequilibria in asexual and sexual populations of the wheat leaf rust fungus, *puccinia recondita*. *Genome* 1998, **41**:832–840.
- Uppalapati S, Ishiga Y, Doraiswamy V, Bedair M, Mittal S, Chen J, Nakashima J, Tang Y, Tadege M, Ratet P, Chen R, Schultheiss H, Mysore KS: Loss of abaxial leaf epicuticular wax in *Medicago truncatula* *irg1/palm1* mutants results in reduced spore differentiation of anthracnose and nonhoast rust pathogens. *Plant Cell* 2012, **24**(1):353–370.
- Harder JH: Developmental ultrastructure of hyphae and spores. In *The cereal rusts, Vol 1*. Edited by Bushnell WR, Roelfs AP. Orlando, Florida: Academic; 1984:333–373.
- Bolton MD, Kolmer JA, Garvin DF: Wheat leaf rust caused by *Puccinia triticina*. *Mol Plant Path* 2008, **9**(5):563–575.
- Duplessis S, Cuomo C, Lin Y-C, Aerts A, Tisserant E, Veneault-Fourrey C, Joly DL, Hacquard S, Amselem J, Cantarel BL, Chiu R, Coutinho PM, Feaue N, Field M, Frey P, Gelhaye E, Goldberg J, Grabherr MG, Kodira CD, Kohler A, Kües U, Lindquist EA, Lucas SM, Mago R, Mauceli E, Morin E, Murat C, Pangilinan JL, Park R, Pearson M, Quesneville H, Rouhier N, Sakthikumar S, Salamov AA, Schmutz J, Selles B, Shapiro H, Tanguay P, Tuskan GA, Henrissat B, Van de Peer Y, Rouzéc P, Ellis JG, Dodds PN, Scheinh JE, Zhong S, Hameline RC, Grigoriev IV, Szabo LS, Martin F: Obligate biotrophy features unraveled by the genomic analysis of rust fungi. *Proc Nat Acad Sci* 2011, **108**:9166–9171.
- Kasuga T, Kozanitas M, Bui M, Hüberli D, Rizzo DM, Garbelotto M: Phenotypic diversification is associated with host-induced transposon derepression in the sudden oak death pathogen *Phytophthora ramorum*. *PLoS One* 2012, **7**(4):e34728.
- Kang S, Lebrun MH, Farrall L, Valent B: Gain of virulence caused by insertion of a *Pot3* transposon in a *Magnaporthe grisea* avirulence gene. *MPMI* 2001, **14**:671–674.
- Gout L, Fudal I, Kuhn M, Blaise F, Eckert M, Cattolico L, Balesdent M-H, Rouxel T: Lost in the middle of nowhere: the *AvLm1* avirulence gene of the Dithideomycete *Leptosphaeria maculans*. *Molec Microb* 2006, **60**(1):67–80.
- Ma L, van der Does C, Borkovich KA, Coleman JJ, Daboussi M, et al: Comparative genomics reveals mobile pathogenicity chromosomes in *Fusarium*. *Nature* 2010, **464**:367–373.
- Rep M, Kistler HC: The genomic organization of plant pathogenicity in *Fusarium* species. *Curr Opin in Plant Biol* 2011, **13**(4):420–426.
- Kämper J, Kahmann R, Bölker M, Ma LJ, Brefort T, Saville BJ, Banuett F, Kronstad JW, Gold SE, Müller O, Perlin MH, Wösten HA, de Vries R, Ruiz-Herrera J, Reynaga-Peña CG, Snelaelar K, McCann M, Pérez-Martín J, Feldbrügge M, Basse CW, Steinberg G, Ibeas J, Holloman W, Guzman P, Farman M, Stajich JE, Sentandreu R, González-Prieto JM, Kennell JC, Molina L, Schirawski J, Mendoza-Mendoza A, Greilinger D, Münch K, Rössel N, Scherer M, Vranes M, Ladendorff O, Vincon V, Fuchs U, Sandrock B, Meng S, Ho EC, Cahill MJ, Boyce KJ, Klose J, Klosterman SJ, Deelstra HJ, Ortiz-Castellanos L, Li W, Sanchez-Alonso P, Schreier PH, Häuser-Hahn I, Vaupel M, Koopmann E, Friedrich G, Voss H, Schlüter T, Margolis J, Platt D, Swimmer C, Gnirke A, Chen F, Vysotskaia V, Mannhaupt G, Güldener U, Münsterkötter M, Haase D, Oesterheld M, Mewes HW, Mauceli EW, DeCaprio D, Wade CM, Butler J, Young S, Jaffe DB, Calvo S, Nusbaum C, Galagan J, Birren BW: Insights from the genome of the biotrophic fungal plant pathogen *Ustilago maydis*. *Nature* 2006, **444**(7115):97–101.
- Cuomo C, Güldener U, Xu JR, Trail F, Turgeon BG, Di Pietro A, Walton JD, Ma L, Baker SE, Rep M, Adam G, Antoniw J, Baldwin T, Calvo S, Chang Y, DeCaprio D, Gale LR, Gnerre S, Goswami RS, Hammond-Kosack K, Harris LJ, Hilburn K, Kennell JC, Kroken S, Magnuson JK, Mannhaupt G, Mauceli E, Mewes H, Mitterbauer R, Muehlbauer G, Münsterkötter M, Nelson D, O'Donnell K, Ouellet T, Qi W, Quesneville H, Roncero MIG, Seong K, Tetko IV, Urban M, Waalwijk C, Ward TJ, Ao J, Birren BW, Kistler HC: The *Fusarium graminearum* genome reveals a link between localized polymorphism and pathogen specialization. *Science* 2007, **317**(5843):1400–1402.
- Dean RA, Talbot NJ, Ebbole DJ, Farman ML, Mitchell TK, Orbach MJ, Thon M, Kulkarni R, Xu JR, Pan H, Read ND, Lee YH, Carbone I, Brown D, Oh YY, Donofrio N, Jeong JS, Soanes DM, Djonovic S, Kolomiets E, Rehmeier C, Li W, Harding M, Kim S, Lebrun MH, Bohnert H, Coughlan S, Butler J, Calvo S, Ma LJ, Nicol R, Purcell S, Nusbaum C, Galagan JE, Birren BW: The genome sequence of the rice blast fungus *Magnaporthe grisea*. *Nature* 2005, **434**(7036):980–986.
- Dodds PN, Lawrence GJ, Catanzariti AM, Teh T, Wang CIA, Ayliffe MA, Kobe B, Ellis JG: Direct protein interaction underlies gene-for-gene specificity and coevolution of the flax resistance genes and flax rust avirulence genes. *Proc Nat Acad Sci USA* 2006, **103**:8888–8893.
- Catanzariti AM, Dodds PN, Lawrence GJ, Ayliffe MA, Ellis JG: Haustorially expressed secreted proteins from flax rust are highly enriched for avirulence elicitors. *Plant Cell* 2006, **18**:243–256.
- Nirmala J, Drader T, Lawrence PK, Yin C, Hulbert S, Steber CM, Steffenson BJ, Szabo LJ, von Wettstein D, Kleinhofs A: Concerted action of two avirulent spore effectors activates *Reaction to Puccinia graminis 1* (*Rpg1*)-mediated cereal stem rust resistance. *Proc Nat Acad Sci USA* 2011, **108**(35):14676–14681.
- Spanu PD, Abbott JC, Amselem J, Burgis TA, Soanes DM, Stüber K, Ver Loren Van Themaat E, Brown JK, Butcher SA, Gurr SJ, Lebrun MH, Ridout CJ, Schulze-Lefert P, Talbot NJ, Ahmadinejad N, Ametz C, Barton GR, Benjida M, Bidzinski P, Bindschedler LV, Both M, Brewer MT, Cadle-Davidson L, Cadle-Davidson MM, Collemare J, Cramer R, Frenkel O, Godfrey D, Harriman J, Hoede C, King BC, Klages S, Kleemann J, Knoll D, Koti PS, Kreplak J, López-Ruiz FJ, Lu X, Maekawa T, Mahanil S, Micali C, Milgroom MG, Montana G, Noir S, O'Connell RJ, Oberhaensli S, Parlange F, Pedersen C, Quesneville H,

- Reinhardt R, Rott M, Sacristán S, Schmidt SM, Schön M, Skamnioti P, Sommer H, Stephens A, Takahara H, Thordal-Christensen H, Vigouroux M, Wessling R, Wicker T, Panstruga R: **Genome expansion and gene loss in powdery mildew fungi reveal tradeoffs in extreme parasitism.** *Science* 2010, **330**(6010):1543–15436.
19. Schirawski J, Mannhaupt G, Münch K, Brefort T, Schipper K, Doehlemann G, Stasio MD, Rössel N, Mendoza-Mendoza A, Pester D, Müller O, Winterberg B, Meyer E, Ghareeb H, Wollenbert T, Münsterkötter M, Wong P, Walter M, Stukenbrock E, Güldener U, Kahmann R: **Pathogenicity determinants in smut fungi revealed by genome comparison.** *Science* 2010, **330**:1546–1548.
20. Jin Y, Szabo LJ, Carson M: **Century-old mystery of *Puccinia striiformis* life history solved with the identification of *Berberis* as an alternate host.** *Phytopathology* 2010, **100**:432–435.
21. Saari EE, Young HC, Kernkamp MF: **Infection of North American *Thalictrum* ssp. with *Puccinia recondita* f.sp. *tritici*.** *Phytopathology* 1968, **58**:939–943.
22. Chester KS: *The Nature And Prevention Of The Cereal Rusts As Exemplified In The Leaf Rust Of Wheat.* Waltham Mass: Chronica Botanica Co; 1946.
23. Jackson HS, Mains EB: **Aecial stage of the orange leaf rust of wheat, *Puccinia triticina* Eriks.** *Journal of Agr Res* 1921, **22**:151–172.
24. Goyeau H, Halkett F, Zapater M, Carlier J, Lannou C: **Clonality and host selection in the wheat pathogenic fungus *Puccinia triticina*.** *Fungal Genetics and Biol* 2007, **44**(6):474–483.
25. Kolmer JA, Liu JQ, Sies M: **Virulence and molecular polymorphism in *Puccinia recondita* fsp *tritici* in Canada.** *Phytopathology* 1995, **85**:276–285.
26. Kolmer JA, Ordoñez ME: **Genetic differentiation of *Puccinia triticina* populations in central Asia and the Caucasus.** *Phytopathology* 2007, **97**:1141–1149.
27. Ordoñez ME, Kolmer JA: **Differentiation of molecular genotypes and virulence phenotypes of *Puccinia triticina* from common wheat in North America.** *Phytopathology* 2009, **99**:750–758.
28. Ordoñez ME, German SE, Kolmer JA: **Genetic differentiation within the *Puccinia triticina* population in South America and comparison with the North American population suggests common ancestry and intercontinental migration.** *Phytopathology* 2010, **100**:376–383.
29. Wang X, McCallum B: **Fusion body formation, germ tube anastomosis, and nuclear migration during the germination of urediniospores of the wheat leaf rust fungus, *Puccinia triticina*.** *Phytopathology* 2009, **99**(12):1355–1364.
30. Zambino PJ, Kubelik AR, Szabo LJ: **Gene action and linkage of avirulence genes to DNA markers in the rust fungus *Puccinia graminis*.** *Phytopathology* 2000, **90**:819–826.
31. Hu G, Linning R, McCallum B, Banks T, Cloutier S, Butterfield Y, Liu J, Kirkpatrick R, Stott J, Yang G, Smailus D, Jones S, Marra M, Schein J, Bakkeren G: **Generation of a wheat leaf rust, *Puccinia triticina*, EST database from stage-specific cDNA libraries.** *Mol Plant Pathol* 2007, **8**:451–467.
32. Xu J, Linning R, Fellers J, Dickinson M, Zhu W, Antonov I, Joly D, Donaldson ME, Eilam T, Anikster Y, Banks T, Munro S, Mayo M, Wynhoven B, Ali J, Moore R, McCallum B, Borodovsky M, Saville B, Bakkeren G: **Gene discovery in EST sequences from the wheat leaf rust fungus *Puccinia triticina* sexual spores, asexual spores and haustoria, compared to other rust and corn smut fungi.** *BMC Genomics* 2011, **12**:161.
33. Link TI, Voegelé RT: **Secreted proteins of *Uromyces fabae*: similarities and stage specificity.** *Mol Plant Pathology* 2008, **9**(1):59–66.
34. Dong YL, Yin CT, Hulbert S, Chen XM, Kang ZS: **Cloning and expression analysis of three secreted protein genes from wheat stripe rust fungus *Puccinia striiformis* f. sp. *tritici*.** *J Microbiol Biotechnol* 2011, **27**(5):1261–1265.
35. Kelley LA, Sternberg MJE: **Protein structure prediction on the web: a case study using the Phyre server.** *Nat Protoc* 2009, **4**:363–371.
36. Oberhaensli S, Parlange F, Buchmann JP, Jenny FH, Abbott JC, Burgis TA, Spanu PD, Keller B, Wicker T: **Comparative sequence analysis of wheat and barley powdery mildew fungi reveals colinearity, dates of divergence and indicates host-pathogen co-evolution.** *Fungal Gen Biol* 2011, **48**:327–334.
37. Laurie JD, Ali S, Linning R, Mannhaupt G, Wong P, Güldener U, Münsterkötter M, Link TI, Voegelé RT: **Secreted proteins of *Uromyces fabae*: similarities and stage specificity.** *Mol Plant Pathol* 2008, **9**:59–66.
38. Tzima AK, Paplomatas EJ, Tsitsigiannis DI, Kang S: **The G protein  $\beta$  subunit controls virulence and multiple growth- and development -related traits in *Verticillium dahlia*.** *Fungal Genet Biol* 2012, **49**(4):271–283.
39. Delgado-Jarana J, Martínez-Rocha AL, Roldán-Rodríguez R, Roncero MI, Di Pietro A: ***Fusarium oxysporum* G-protein beta subunit Fgb1 regulates hyphal growth, development, and virulence through multiple signalling pathways.** *Fungal Genet Biol* 2005, **42**(1):61–72.
40. Nishimura M, Park G, Xu JR: **The G-beta subunit MGB1 is involved in regulating multiple steps of infection-related morphogenesis in *Magnaporthe grisea*.** *Mol Microbiol* 2003, **50**(1):231–243.
41. Sambrook J, Russell DW: *Molecular Cloning: A Laboratory Manual.* 3rd edition. Cold Spring Harbor: NY: Cold Spring Harbor Laboratory Press; 2001.
42. Jurka J, Kapitonov W, Pavlicek A, Klonowski P, Kohany O, Walichiewicz J: **Repbase Update, a database of eukaryotic repetitive elements.** *Cytogenetic and Genome Res* 2005, **110**:462–467.
43. Xu Z, Wang H: **LTR\_FINDER: an efficient tool for the prediction of full-length LTR retrotransposons.** *Nucleic Acids Res* 2007, **35**:W265–W268.

doi:10.1186/1471-2164-14-60

**Cite this article as:** Fellers et al.: Conserved loci of leaf and stem rust fungi of wheat share synteny interrupted by lineage-specific influx of repeat elements. *BMC Genomics* 2013 **14**:60.

**Submit your next manuscript to BioMed Central and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
www.biomedcentral.com/submit

