

TECHNICAL NOTE

Open Access

# Development of a rapid multiplex SSR genotyping method to study populations of the fungal plant pathogen *Zymoseptoria tritici*

Angélique Gautier<sup>1†</sup>, Thierry C Marcel<sup>1†</sup>, Johann Confais<sup>1</sup>, Charles Crane<sup>2</sup>, Gert Kema<sup>3</sup>, Frédéric Suffert<sup>1</sup> and Anne-Sophie Walker<sup>1\*</sup>

## Abstract

**Background:** *Zymoseptoria tritici* is a hemibiotrophic ascomycete fungus causing leaf blotch of wheat that often decreases yield severely. Populations of the fungus are known to be highly diverse and poorly differentiated from each other. However, a genotyping tool is needed to address further questions in large collections of isolates, regarding regional population structure, adaptation to anthropogenic selective pressures, and dynamics of the recently discovered accessory chromosomes. This procedure is limited by costly and time-consuming simplex PCR genotyping. Recent development of genomic approaches and of larger sets of SSRs enabled the optimization of microsatellite multiplexing.

**Findings:** We report here a reliable protocol to amplify 24 SSRs organized in three multiplex panels, and covering all *Z. tritici* chromosomes. We also propose an automatic allele assignment procedure, which allows scoring alleles in a repeatable manner across studies and laboratories. All together, these tools enabled us to characterize local and worldwide populations and to calculate diversity indexes consistent with results reported in the literature.

**Conclusion:** This easy-to-use, accurate, repeatable, economical, and faster technical strategy can provide useful genetic information for evolutionary inferences concerning *Z. tritici* populations. Moreover, it will facilitate the comparison of studies from different scientific groups.

**Keywords:** Microsatellite, Single sequence repeat (SSR), Genotyping, Multiplex, Population genetics, Diversity, *Mycosphaerella graminicola*, Phytopathogenic fungus

## Background

*Zymoseptoria tritici* (syn. *Mycosphaerella graminicola*, anamorph *Septoria tritici*) [1] is an ascomycete fungus that causes Septoria leaf blotch, the most important foliar disease of winter wheat in Northern Europe. Moreover, this species has adapted to fungicides (especially from the azole and strobilurin families), which have been used intensively since the 1970s [2,3]. Intensive research has been undertaken to understand the biology of this fungus and the mechanisms of its pathogenicity, with the objective to develop durable management tools [4]. The *Z. tritici* genome has been fully sequenced and annotated, which revealed a

high chromosome plasticity [5]. The sequenced genome contained 21 chromosomes, eight of them being facultatively gained or lost during meiosis in some isolates and constituting the dispensome, which accounts for up to approximately 12% of the genome size [5,6]. It was recently demonstrated that these accessory chromosomes originated mainly from ancient core chromosomes through a degeneration process that included breakage-fusion-bridge cycles, nondisjunction and mutational decay of duplicated sequences [7].

*Z. tritici* populations have been shown to be highly diverse on scales going from leaf to continents [8]. Wholesale gene flow and frequent recombination are supposed to occur at a global level [9]. Nevertheless, population differentiation has been detected in Iran, western USA, Germany and France [9-13], *i.e.* several of these areas

\* Correspondence: walker@versailles.inra.fr

†Equal contributors

<sup>1</sup>UR 1290 BIOGER-CPP, INRA, BP01, Avenue Lucien Brétignières, F-78850 Thiverval-Grignon, France

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

were not included in the previous worldwide surveys. Additional large collections of isolates need to be analyzed to confirm and refine the population differentiation observed in those specific areas, and to identify the factors that influence the differentiation (for example, anthropogenic selection pressure or host/cultivar adaptation; seasonal adaptation; history of geographic spreads). Those questions should be answered in parallel on the so-called core genome (chromosomes 1 to 13) and the accessory (or dispensable) genome (chromosomes 14 to 21), which are characterized by strikingly different organizations [6,7] and rates of evolution [14] likely to result in diverging dynamics between the two compartments. Such a compartmentalized genome has been termed a “2-speed genome” [7].

Several marker sets are available to study the genetic structure of *Z. tritici* populations, ranging from sequence of mating-type genes [15-17], random amplified polymorphic DNA (RAPD) [18], restriction fragment length polymorphism (RFLP) [8] to amplified fragment length polymorphism (AFLP) [11,19]. A limited set of microsatellite/simple sequence repeat (SSR) markers was developed in the late 1990s [20] and used for field studies [10,21]. It was later complemented by a larger set of 99 new SSRs, derived from an EST database [22], and used at least in one study [23]. SSRs are of particular value for population studies, since they are co-dominant, easy to use, and highly polymorphic markers. As diversity in *Z. tritici* is unanimously considered high, a reduced panel of adapted SSRs should suffice to analyze population genetic structure [17]. Ideally, such markers should be genetically independent, *i.e.*, reside on different chromosomes, and selectively neutral to avoid bias in subsequent population genetic analyses [24]. Moreover, multiplex PCR, *i.e.*, the simultaneous amplification of several markers within a single reaction, is a good way to improve the SSR genotyping efficiency and to considerably reduce cost and labour [25].

Here, we optimize and complete a method that enables the amplification of 24 highly informative SSRs from *Z. tritici*, multiplexed into three independent panels.

## Findings

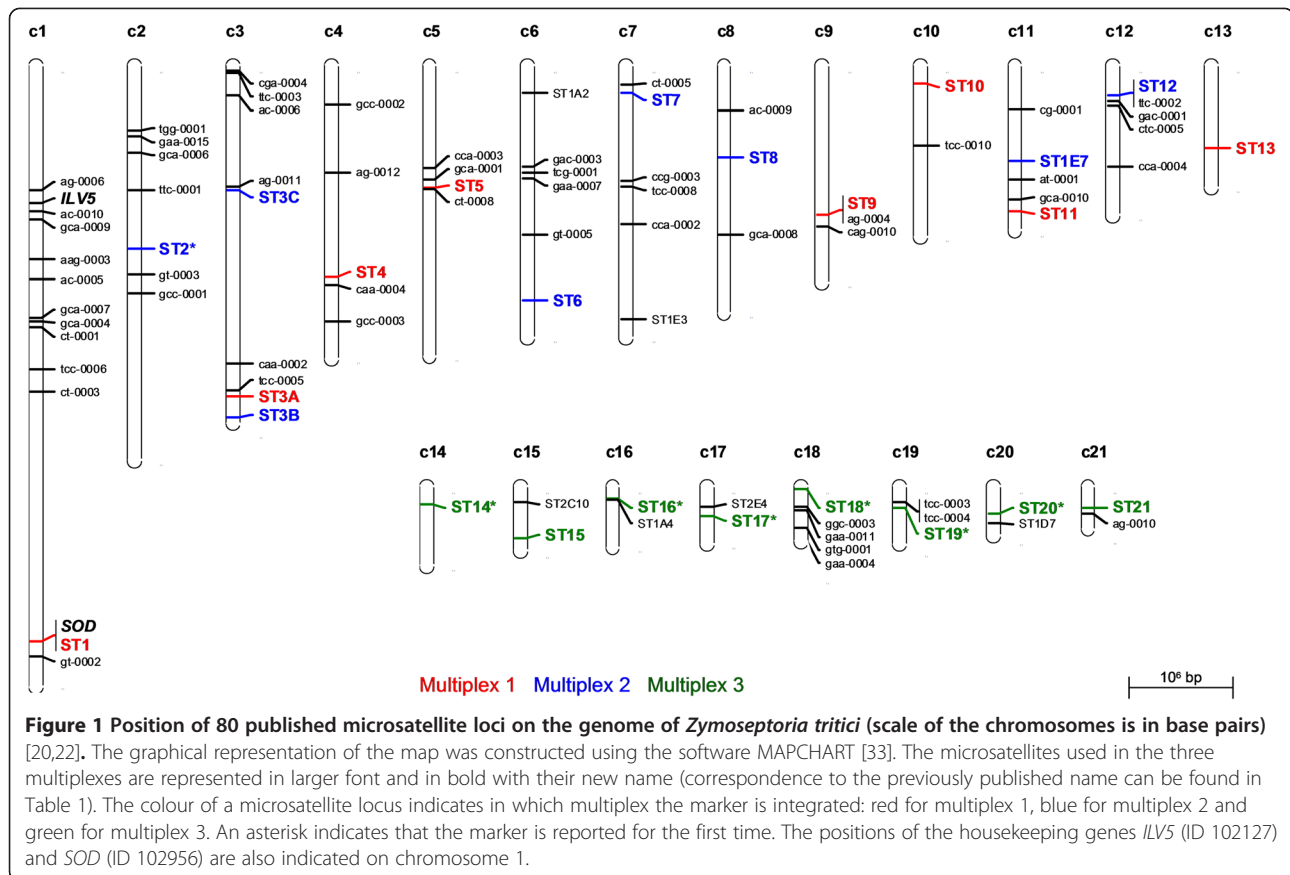
### Development of three efficient and multiplexed panels of SSRs for population studies

To ensure relative independence of markers, and to detect the variably present accessory chromosomes, at least one SSR was chosen per chromosome. Therefore, the 9 [20] and 99 [22] previously developed SSRs were mapped *in silico* by blast-searching their primers against the *Z. tritici* genome at the JGI Genome Portal website (<http://genome.jgi-psf.org/>) [5]. The chromosome position of 7 and 73 SSRs was finally determined in each marker set, respectively (Figure 1). Blast of the primers from the

remaining markers yielded non-specific or no result (Table 1). Interestingly, the position of 21 from the 23 SSRs genetically mapped by Goodwin *et al.* [22] was confirmed while the last 2 could not be positioned *in silico*.

Additionally, in order to increase the number of available SSRs from which to choose markers to constitute the panels, microsatellites were mined from the fasta-formatted genome of strain IPO-323 using the same pipeline of Perl scripts and C++ program that has been used to obtain microsatellites from EST sequences of higher plants [26]. This pipeline output primer sequences, motifs and repeat counts of simple and compound microsatellites, primer annealing temperatures, and primer locations in the genome. It searched for motifs of one to four nucleotides, with respective minimum repeat counts of 12, 6, 4, and 3. Up to 10 percent of mismatching positions were allowed, and the maximum distance between consecutive repeat blocks was five nucleotides in compound microsatellites. Primers were chosen to amplify products 100 to 500 nucleotides long, and to anneal at  $60 \pm 1.2^\circ\text{C}$ , with preference to annealing temperatures in the interval  $60.0 \pm 0.2^\circ\text{C}$ . In addition to the assembled genome of IPO-323, 86,776,768 short (45-base) reads were available for strain IPO-94269 (unpublished data). Reads with sufficiently high quality scores were mapped to the genome of IPO-323 by blastn, and a Perl-collated consensus of these reads provided an approximate assembly of the genome of IPO-94269. IPO-94269 lacks the accessory chromosomes 18 and 20, present in IPO-323 [5]. Otherwise, it was assumed for the sake of tractability that the two strains differed only in substitutions, and not in indels. Another Perl script blast-aligned the microsatellite primers from above to the IPO-323 and IPO-94269 genomes, and selected the primer pairs that hit each genome exactly once. For chromosomes 18 and 20, we selected the primer pairs that hit the genome of IPO-323 exactly once. Together, these were the 3781 “specific” microsatellites that are expected to produce single products from both genomes. Details on these SSRs are available in Additional file 1. A subset of 119 microsatellites having at least 50, 17, 10, or 8 repeats respectively for motifs of 1, 2, 3, or 4 nucleotides is also available in Additional file 2.

Thereafter, markers were chosen according to (1) their chromosomal location, (2) their polymorphism, (3) their allele size range, (4) their specificity towards related species and (5) the annealing temperature of their primers. For those reasons, candidate markers were selected in priority from the published sets because their polymorphism and species specificity had previously been assessed in a number of isolates [16,21,23]. SSR markers from core chromosomes were assembled in multiplexes 1 and 2 (8 markers per multiplex), and SSR markers from accessory chromosomes were combined in multiplex 3. The forward primers were labeled with fluorescent dyes (blue, green or



black from Sigma-Aldrich®) in a way that the same color was given only to markers with non-overlapping range of allele sizes. Candidate markers were individually tested on eight representative *Z. tritici* DNAs from diverse origins, and on DNAs from fungal species closely related to *Z. tritici* and of agronomical importance [27], including three *Zymoseptoria passerinii*, one *Stagonospora nodorum* and one *Mycosphaerella fijiensis*. When markers were clearly amplified, were specific to *Z. tritici* and were found polymorphic between isolates, they were assembled in multiplexes and tested on the same DNA samples. According to the band patterns observed on high resolution agarose gel, multiplexes were optimized by means of marker exchange until all markers within a multiplex produced the expected amplification pattern at the consensus annealing temperature of 57°C, with the exception of multiplex 3 for which null alleles were allowed. The marker ST1E7 [20] was included in multiplex 2 to allow comparison with previously published datasets [10]. Finally, primers amplifying a 108 bp fragment from the housekeeping gene *ILV5*, keto-acid reductoisomerase (Chromosome 1, Protein ID 102127), and a 96 bp fragment from the housekeeping gene *SOD*, superoxide dismutase (Chromosome 1, Protein ID 102956), were introduced in the multiplexes to serve as amplification

controls. This is particularly important for multiplex 3, since isolates without any accessory chromosome may be detected in field populations.

Once the three panels were optimized on the limited number of samples, the procedure was automatized for high-throughput routine work. Fresh sporidia, collected from 5 day-old mono-pycnidia cultures on malt-yeast agar MYA (20 g.l<sup>-1</sup> malt extract, 5 g.l<sup>-1</sup> yeast extract and 12.5 g.l<sup>-1</sup> agar), were distributed in 96-well plates of the following extraction kit and ground for 90 s at 50 rps with a tungsten bead in 50 µl lysis (AP-1) buffer (65°C) with a Retsch-MM300 grinder. DNA extraction was carried-out using the Qiagen DNA 'Biorobot 3000' extraction robot and the Qiagen DNeasy Plant Maxi Kit, following the manufacturer's protocol. After extraction, DNAs were quantified on a Nanodrop spectrophotometer (Thermo Scientific®), adjusted to 15 ng.µl<sup>-1</sup> and stored at -20°C until use. 10 µl PCR reactions were performed using the Type-it microsatellite kit (Qiagen), specifically designed for multiplexed PCRs. Each reaction contained 2 µl deionized water, 1 µl Q-solution, 1 µl of the primer mix (containing 2 µM of each SSR/housekeeping gene primers), 5 µl of the Type-it mix and 1 µl DNA. Panel composition is detailed in Table 1. PCR conditions were identical for all multiplexes, with preheating

**Table 1 Composition of three microsatellite PCR multiplexes developed for *Zymoseptoria tritici* population genetic studies**

Multiplex <sup>a</sup>	Chrom.	Locus <sup>b</sup>	Locus in literature <sup>c</sup>	Repeat motif of IPO-323 <sup>d</sup>	Primer sequences (5' -> 3')	Fluores-cent dye <sup>e</sup>	Size range (bp)	Protein ID <sup>f</sup>	Na <sup>g</sup>	Ne <sup>h</sup>	Fna <sup>i</sup>	Gene diversity (H <sub>E</sub> ) <sup>j</sup>	
												Total pop. (n = 166)	Total pop. (n = 166)
<b>Multiplex 1 (core chromosomes)</b>	1	ST1	GAC0002	(GAC) <sub>6</sub>	F: AATCGACCCCTT CCTCAAC R: GGGGGAGAGGC ATAGTCTTG	Black	192-222	33894	3	1.4	0.6%	0.305	0.308
	3	ST3A	AG0003	(AG) <sub>15</sub>	F: ACTTGGGGAGGT GTTGTGAG R: ACGAATTGTTCATCCAGCG	Black	226-258	103950	16	5.8	2.4%	0.785	0.825
	4	ST4	AC0002	(AC) <sub>7</sub>	F: TGAACATCAACC TCACACGC R: AGAAGAGGACG ACCCAGAG	Green	182-206	71387	7	2.0	0.0%	0.479	0.504
	5	ST5	GGC0001	(GGC) <sub>8</sub>	F: GATACCAAGGT GGCCAAGG R: CACGTTGGGAG TGTCGAAG	Green	232-256	104484	6	1.4	0.0%	0.114	0.345
	9	ST9	CT0004	(CT) <sub>6</sub>	F: CACCTCACTCCT CAATTCCG R: GAAAGGTTGGT GTCGTGTC	Green	336-348	62567	6	1.7	3.0%	0.404	0.413
	10	ST10	CAA0003	(CAA) <sub>7</sub>	F: TCCGTCATCAAC AACACCAG R: TGGCCGTAGAA CTGCTGAG	Blue	139-160	62856	6	2.7	0%	0.624	0.631
	13	ST13	AG0009	(AG) <sub>10</sub> (GGCA) <sub>3</sub>	F: GACTCCATTAC CTGTGGCG R: TGTGAAGGACA CGCAAAGAG	Blue	192-200	106529	5	2.4	0%	0.566	0.587
	11	ST11	GCA0002	(GCA) <sub>6</sub>	F: GTGTGAAACGA AAAGCGGAG R: TACTGGGTATCG AGATCGGC	Blue	293-296	-	2	1.1	17.5%	0.071	0.087
<b>Multiplex 2 (core chromosomes)</b>	8	ST8	CT0007	(CT) <sub>6</sub>	F: TGCAGGGCATT TAATTGAGG R: TCCATCCATTTA GGCTCGTC	Black	180-188	61457	5	1.6	40.4%	0.322	0.404

**Table 1 Composition of three microsatellite PCR multiplexes developed for *Zymoseptoria tritici* population genetic studies (Continued)**

	12	ST12	TCC0002	(TCC) <sub>6</sub>	F: GAATCCACCTCT TCCTTGCC  R: AGGAGGATATC AAGGCCAG	Black	226-232	111480	3	1.1	5.4%	0.059	0.137
	3	ST3B	CAA0005	(CAA) <sub>8</sub>	F: AAGAATCCCAC CACCCAAAC  R: CACACGGCTCC TTTGACAC	Black	263-299	-	10	2.9	0.6%	0.663	0.648
	6	ST6	TCC0009	(TCC) <sub>8</sub>	F: TCAATTGCCAAT AATTCGGG  R: AGACGAGGCAG TTGGTTGAG	Green	161-179	73474	6	1.5	3.6%	0.242	0.342
	3	ST3C	GCA0003	(GCA) <sub>7</sub>	F: TCCTATCAACTC CCGAGACG  R: CCGCACGTAGG AATTTTCAG	Green	229-253	108534	5	2.2	1.2%	0.580	0.533
	2	ST2	Chr02-140	(GCT) <sub>12</sub>	F: ACACCAAAGAA GGATCCACG  R: GCCGGAGGTCT ATCAGTTTG	Green	338-365	108062	9	2.9	0%	0.439	0.689
	11	ST1E7	ST1E7	(CGG) <sub>3</sub>	F: GATCTCGAGCA GGGCCGAAGT  R: TCACACGCTGGT CTGTGAATC	Blue	86-98	-	4	1.5	20.0%	0.204	0.373
	7	ST7	AC0001	(AC) <sub>21</sub>	F: CACCACACCGT CGTTCAAG  R: CGTAAGTTGGTG GAGATGGG	Blue	171-227	105049	15	1.9	0.6%	0.463	0.487
<b>Multiplex 3 (accessory chromosomes)</b>	15	ST15	AC0007	(AC) <sub>15</sub>	F: TGCTCGCAAGAC ATAAAACG  R: CTCTTAGCATTG GTCGGTGG	Black	92-152	106622	15	8.0	16.9%	0.858	0.885
	21	ST21	GGA0001	(GGA) <sub>6</sub>	F: GTACGACACGG GCTATGGAG  R: GGCGATGACGA TGAAACC	Black	149-161	111806	5	1.8	32.5%	0.632	0.640
	18	ST18	Chr18-007*	(GGA) <sub>5</sub>	F: AGAGGCAATGG TGGGTGAT	Black	245-272	111768	2	1.2	71.1%	0.411	0.430

**Table 1 Composition of three microsatellite PCR multiplexes developed for *Zymoseptoria tritici* population genetic studies (Continued)**

19	ST19	Chr19-031	(TCG) <sub>6</sub>	R: TCCTCCTCGTCC GACAATAC F: CTACTGTATTC CCGGGGGT	Black	351-363	97944	5	2.3	12.6%	0.700	0.631
20	ST20	Chr20-066*	(GAA) <sub>22</sub>	R: ACCCCTCTCTCC TTCTCTCG F: CTCCTCCACATC CATCCAAC	Green	229-289	-	19	13.9	10.8%	0.891	0.914
16	ST16	Chr16-016	(GAA) <sub>6</sub>	R: GACGGAGGGGG AGAGGTAT F: TTCACGGTATCA CAGACCCA	Green	496-538	97708	7	1.2	13.3%	0.369	0.364
17	ST17	Chr17-040*	(AGA) <sub>5</sub>	R: GGGTTGTCCAAG CTGTTGTT F: ACAAGAGGCGG AAGACTGAA	Blue	236-239	97821	2	1.0	18.7%	0.367	0.311
14	ST14	Chr14-023	(TCC) <sub>20</sub> (TCG) <sub>5</sub>	R: TCATGCGTCGTA TTCTTGGA F: CAAACAGCCAGA TCCACCTT R: GTCGTGGTCGGA GAGAGAAG	Blue	361-529	111734	36	12.1	13.8%	0.898	0.920

<sup>a</sup>Primers for the housekeeping gene SOD are added in multiplexes 1 and 3 (green fluorescent dye; 96 bp amplicon; F:CTACCGATGCTCCAGCTGAGA and R:CACCACCTTGCCATCCA), and primers from the housekeeping gene ILV5 are added in multiplexes 1 and 2 (black fluorescent dye; 108 bp amplicon; F:GCTCTATCGCCATCTCCAGG and R:TCCCTCTCGAAGGTGGTCTTGT).

<sup>b</sup>Proposed name of microsatellite loci starting with the fungus initials and followed by the chromosome number on which the microsatellite is located (except for ST1E7).

<sup>c</sup>Loci name found in the literature [20,22] or in the Additional file 1 (primers for Chr18-007, Chr20-066 and Chr17-040 have been modified from the original design to fit the multiplex range size requirements).

<sup>d</sup>Repeat motif of IPO-323 as published in [20,22], with exception of ST7 and ST13, for which the motif was found slightly different after sequencing.

<sup>e</sup>Primers were purchased from Sigma Aldrich (France). They are WellRed oligos, licensed by Beckman Coulter (USA), and bearing fluorochromes at their 5' end: D1: red, D2: black, D3: green, D4: blue.

<sup>f</sup>Identity number of the protein corresponding to SSR loci identified within expressed sequences (Filtered Gene Models from Mgraminicolav2 at JGI).

<sup>g</sup>Number of alleles identified in the total population (excluding null alleles).

<sup>h</sup>Effective number of alleles (excluding null alleles).

<sup>i</sup>Frequency of missing data or null alleles (no amplification) in the total population.

<sup>j</sup>Unbiased gene diversity  $H_E$  calculated as the probability that two individuals taken at random have different genotypes (excluding null alleles) [31].

at 95°C for 5 mn, followed by 30 cycles of 95°C for 30 s, 57°C for 90 s and 72°C for 30 s, with a final extension step of 60°C for 30 mn, using a MJ research PTC-200 thermocycler. The PCR products were run on a 2% agarose gel to check for amplification products. Sequencing reactions were prepared as follows: for multiplexes 1 and 2, 0.8 µl PCR product, 40 µl Sample Loading Solution (SLS) and 0.5 µl 400 bp size standard (red dye); the mix for multiplex 3 was the same, with the exception that 0.5 µl 600 bp size standard was used instead of the previous one. Fragments were separated using a Beckman Coulter CEQ-8000 DNA Analyzer. Results were visualized using the CEQ-8000 Genetic Analysis System Software (Beckman Coulter).

#### Automation of allele scoring

We automated the allele scoring procedure using the binning analysis function of the CEQ-8000 Genetic Analysis System Software (Beckman Coulter) on the 166 genotyped samples (see below for description of populations). Such automation helps on one hand to solve problems related to the size assignment of ambiguous peaks produced on the chromatogram by Taq polymerase stuttering or fluorescent dyes reflection, and provides on the other hand a reproducible scoring method across readers and laboratories. In short, the software first provides for each marker a distribution of peak height and approximate size relative to the size-standard ladder. Each marker was then manually delimited ('binned') for each allele according to its size variability and a minimal peak height, in order to eliminate false positive or aberrant peaks and to minimize bin width. Binned alleles were finally assigned a definitive name (generally derived from the mean allele size) and allele size ranges were established (Table 1). The generated binning files could be used to annotate allele peaks from new data in any laboratory using a Beckman Coulter DNA Analyzer, and are described in Additional file 3 and provided as Additional file 4 or upon request to the authors. The alleles' sizes of IPO-323 and other 100 reference isolates are provided as examples of this procedure (Additional file 3) to help further assignment by other operators. Resequencing data of all these reference isolates will be soon released on the JGI Genome Portal website (<http://genome.jgi-psf.org/>). This procedure assures reliable and accurate allele scoring, even when different or inexperienced operators manipulate the data, increasing the reproducibility between distinct datasets generated over time. This procedure can also be progressively improved, since new alleles of the same markers and also completely new markers can be implemented in the "binning files".

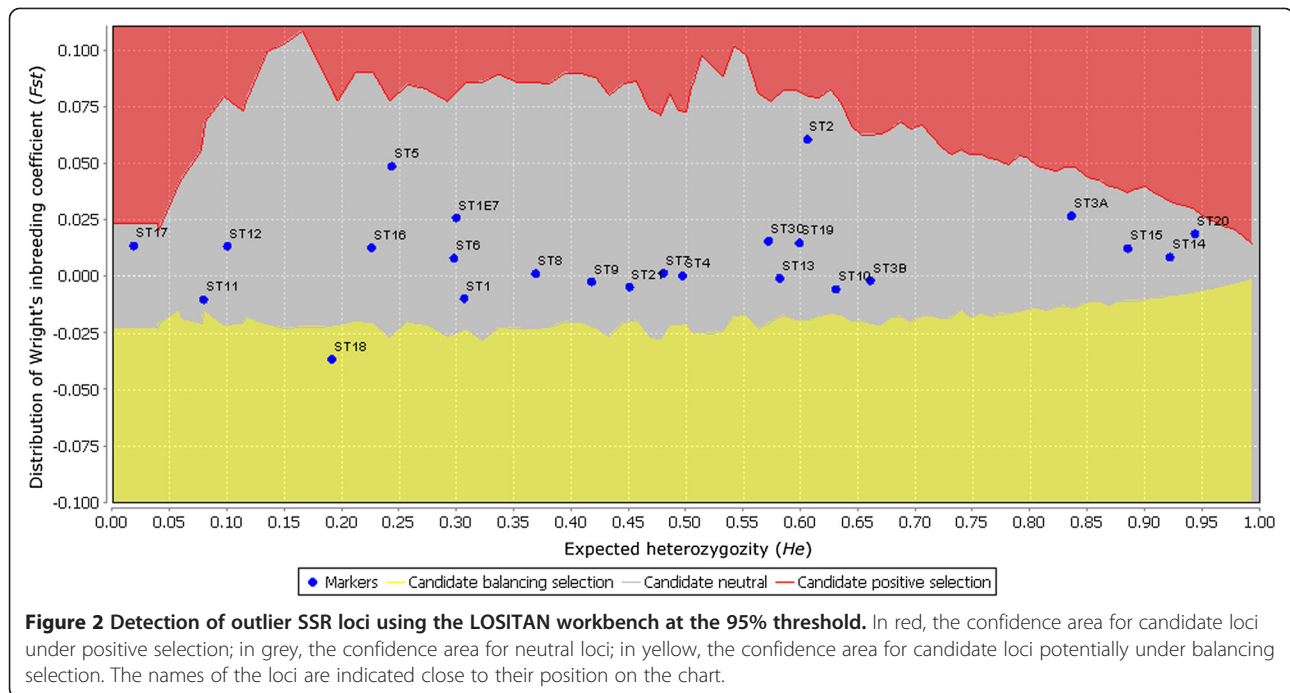
#### Checking the quality of the multiplexed SSR panels

Frequency parameters were calculated for the 166 individuals (see below for the description of the populations) with the software GENALEX 6.4 [28]. For each marker,

the number of SSR repeats present in the isolate IPO-323 was determined according to its genome sequence and used to infer the number of repeats from all the other alleles. In a few cases, fewer than 0 repeats were deduced for alleles present in never more than one to three individuals; such alleles were coded as missing alleles and were not considered further. Excluding those, a total of 204 alleles were detected. The number of alleles per SSR ranged from 2 to 16 for panels 1 and 2 on the core chromosomes (Table 1). Markers on the accessory chromosomes 14 and 20 had particularly numerous alleles with 36 and 20 detected alleles, respectively. The 3 markers ST11 in panel 1, and ST8 and ST1E7 in panel 2 could not be amplified in a large number of isolates probably because of the presence of null alleles (Table 1). The same observation is true for all markers from panel 3 on the accessory chromosomes with 11 to 71% null amplifications, on chromosomes 20 and 18, respectively. This result is supported by other studies, which revealed that chromosome 18 is the accessory chromosome the most often missing in field isolates of *Z. tritici* [7]. We also sequenced the PCR product of ST18 from 3 isolates carrying the 245 bp allele (IPO-02166, IPO-98032, IPO-95054) and 2 isolates carrying the 272 bp allele (IPO-98038, IPO-95052) (Table 1) and found that this marker was not polymorphic at its microsatellite motif (GGA)<sub>5</sub>. The allele size difference at ST18 is caused by an insertion of 27 bp in the larger allele.

The eight accessory chromosomes exhibit strong differentiation between strains, in number but also in length of chromosomes [7]. Indeed, the frequency in presence or absence of chromosomal segments varies strongly along each accessory chromosome. For example, chromosome 14 is present in more than 80% of the strains while it harbors a substantial insertion present in less than 20% of these same strains [7]. Except for chromosome 18, the SSR markers from panel 3 on accessory chromosomes are in conserved segments of their respective chromosome, which make them reliable indicators of the presence or absence of each accessory chromosome. The presence or absence of chromosome 18 is particularly difficult to predict due to its high differentiation between strains from different geographical origins [7].

Non-neutral loci may bias population genetic studies [24]. To identify possible outlier loci in the panels, the 24 SSR markers constituting the three multiplex panels were subjected to analysis with the selection detection workbench LOSITAN [29]. LOSITAN detects outlier loci from the expected distribution of Wright's inbreeding coefficient  $F_{st}$  versus the expected heterozygosity (or gene diversity)  $H_e$  under an island model of migration with neutral markers. A neutral mean  $F_{st}$  was computed after excluding loci for which selection had been detected initially. The mean observed  $F_{st}$  value was 0.005, while individual marker  $F_{st}$  values ranged from -0.036 to 0.061. Only one

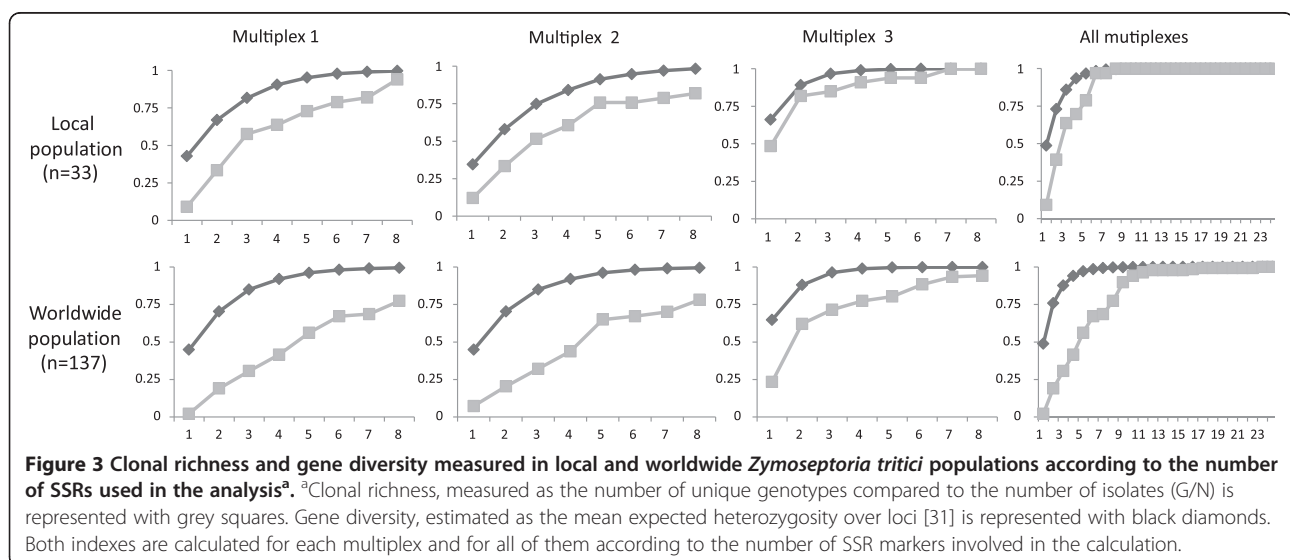


out of the 24 SSR markers, ST18, was declared an outlier, since it appeared in the confidence area for loci potentially under balancing selection (Figure 2). Either directional selection or genotyping errors can generate outlier effects. The very high proportion (71%) of null alleles for ST18 is the most likely explanation for this marker being detected as an outlier. The other markers are considered as neutral.

#### Validation of the multiplexed SSR panels: diversity measurement in two populations

To validate the SSR multiplexing procedure, the automated allele assignment, and the value of the selected

SSR markers for population genetic studies, two distinct fungal populations were genotyped following the proposed protocol. Isolates belonging to a first, global, population ( $n = 133$ ) were collected worldwide over 25 years; they covered 28 countries from the five continents, including different areas in France, and were retrieved from laboratory collections. Isolates belonging to a second, local, population ( $n = 33$ ) were collected during the same crop season 2009–2010 in a single experimental field plot in Grignon, France ( $48^{\circ}51'N$ ,  $1^{\circ}58'E$ ; [30]). Samples were genotyped and analyzed as previously described.





For further index calculations, missing data were coded as missing alleles for multiplexes 1 and 2, and as null alleles for multiplex 3. Gene diversity, estimated as the mean expected heterozygosity  $H_e$  over the loci [31], and clonal richness ( $G/N$ ; number of unique genotypes compared to the number of individuals) were plotted against the number of loci used to calculate these indexes in the three multiplexes and for all of them (Figure 3), using the MULTILOCUS program [32]. As expected, gene diversity was similarly high in the two populations, and exceeded 0.95 after 4 loci were included in calculation for multiplexes 1 and 2, and after two loci were included for multiplex 3, probably because null alleles are considered as distinct alleles for markers located on the dispensome. The 24 SSR loci present in the three panels are then sufficient to saturate the whole gene diversity, at a local and a larger scale. The observed ranges for gene diversity, for single loci (Table 1) or for genotypes (Figure 3), are similar or even greater compared to what has been reported in the literature for various *Z. tritici* populations, either with SSR, RAPD, AFLP or RFLP markers [9-13,18]. The 24 SSR loci are also sufficient to detect possible clonal individuals as only 11 and 23 markers were necessary to distinguish all haplotypes in the local and world populations, respectively (Figure 3); in practice 164 haplotypes were recognized among the 166 isolates. Such a low proportion of clonal individuals ( $G/N$ ) corroborates several studies pointing out the importance of sexual recombination in field populations of *Z. tritici* [9,13,17,18,21].

## Conclusion

The proposed multiplex PCR genotyping method, together with the automatic allele assignment procedure, constitutes a reliable tool for population genetics studies on *Z. tritici*. Moreover, these ready-to-use methods enable important savings in terms of money, time, reproducibility and accuracy, and produce results in agreement with others previously published about *Z. tritici* populations. As the chosen SSRs are highly polymorphic and describe most of the population diversity, the three multiplex panels could be used "à la carte" to address specific questions. For example, the use of a single panel (either multiplex 1 or 2) could suffice to describe diversity in local/reduced populations, whereas both panels representing the core chromosomes (multiplexes 1 and 2 together) would be useful to study larger unknown populations. Multiplex 3 is of great interest to acquire complementary information on the dispensome dynamics and evolution [7] of the studied populations, but users should keep in mind that false negative results are always possible due to the variation in chromosomal content between strains. Multiplex 3 should also be used with care for population genetics analyses, as the great number of missing alleles may bias the results. The three markers with null alleles from the multiplexes 1 and 2 (i.e. ST11, ST8,

ST1E7) may be as well omitted from certain analyses for the same reason. These sets of optimized markers will be of great relevance to the *Z. tritici* scientific community, as they provide a method that can be applied in any laboratory and will allow facile comparison of studies from various locations, scales, and with different aims. The method is currently being used successfully in two laboratories to genotype different European populations consisting of several thousand individuals.

## Additional files

**Additional file 1: Characteristics of 3781 "specific" microsatellites identified in silico from *Zymoseptoria tritici* genome.** Primer pairs are proposed for each microsatellite, which hit each genome of strains IPO-323 and IPO-94269 exactly once.

**Additional file 2: Subset of 119 microsatellites with highly repeated motifs.**

**Additional file 3: Alleles computed in the automated allele scoring procedure and frequency of alleles recorded among 166 individual strains.** Strains were selected from a local French population and from a worldwide collection. The multilocus genotypes of 101 strains is provided as reference.

**Additional file 4: Binning files, to be used for automatic allele scoring.** Files to use with the CEQ-8000 Genetic Analysis System Software (Beckman Coulter) to provide automatic assignment of alleles for each SSR marker.

## Competing interest

The authors declare that they have no competing interest.

## Authors' contributions

AG designed and optimized the multiplexes. JC genotyped the populations. TM localized the SSRs on the genome, designed primers and selected candidate markers. CC identified the new SSRs from the genome sequence. GK and FS provided the worldwide and local populations, respectively. ASW led the project and developed allele automatic assignment. ASW and TM analyzed the data and wrote the manuscript. All authors read and approved the final manuscript.

## Author details

<sup>1</sup>UR 1290 BIOGER-CPP, INRA, BP01, Avenue Lucien Brétignières, F-78850 Thiverval-Grignon, France. <sup>2</sup>USDA-ARS, Crop Production and Pest Control Research Unit, Department of Botany and Plant Pathology, Purdue University, 915 West State Street, West Lafayette, IN 47907-2054, USA. <sup>3</sup>Plant Research International, Biointeractions and Plant Health, P.O. Box 16, 6700 AA Wageningen, The Netherlands.

Received: 16 May 2013 Accepted: 30 May 2014

Published: 18 June 2014

## References

1. Quaadvlieg W, Kema GHJ, Groenewald JZ, Verkley GJM, Seifbarghi S, Razavi M, Gohari AM, Mehrabi R, Crous PW: *Zymoseptoria* gen. nov.: a new genus to accommodate *Septoria*-like species occurring on graminicolous hosts. *Persoonia* 2011, **26**:57-69.
2. Stammler G, Carstensen M, Koch A, Semar M, Strobel D, Schlehuber S: Frequency of different CYP51-haplotypes of *Mycosphaerella graminicola* and their impact on epoxiconazole-sensitivity and -field efficacy. *Crop Prot* 2008, **27**:1448-1456.
3. Leroux P, Walker A: Multiple mechanisms account for resistance to sterol 14 $\alpha$ -demethylation inhibitors in field isolates of *Mycosphaerella graminicola*. *Pest Manag Sci* 2011, **67**:44-59.
4. Orton ES, Deller S, Brown JKM: *Mycosphaerella graminicola*: from genomics to disease control. *Mol Plant Pathol* 2011, **12**:413-424.
5. Goodwin SB, Ben M'Barek S, Dhillon B, Wittenberg AHJ, Crane CF, Hane JK, Foster AJ, Van der Lee TAJ, Grimwood J, Aerts A, Antoniw J, Bailey A, Bluhm B,

- Bowler J, Bristow J, van der Burgt A, Canto-Canche B, Churchill ACL, Conde-Ferra ez L, Cools HJ, Coutinho PM, Csukai M, Dehal P, De Wit P, Donzelli B, van de Geest HC, van Ham RCHJ, Hammond-Kosack KE, Henrissat B, Kilian A, et al: **Finished genome of the fungal wheat pathogen *Mycosphaerella graminicola* reveals dispense structure, chromosome plasticity, and stealth pathogenesis.** *PLoS Genet* 2011, **7**(6):e1002070.
6. Wittenberg AHJ, van der Lee TAJ, Ben M'Barek S, Ware SB, Goodwin SB, Kilian A, Visser RGF, Kema GHJ, Schouten HJ: **Meiosis drives extraordinary genome plasticity in the haploid fungal plant pathogen *Mycosphaerella graminicola*.** *PLoS One* 2009, **4**(6):e5863.
  7. Croll D, Zala M, McDonald BA: **Breakage-fusion-bridge cycles and large insertions contribute to the rapid evolution of accessory chromosomes in a fungal pathogen.** *PLoS Genet* 2013, **9**(6):e1003567.
  8. Linde CC, Zhan J, McDonald BA: **Population structure of *Mycosphaerella graminicola*: from lesions to continents.** *Phytopathology* 2002, **92**:946–955.
  9. Zhan J, Pettway RE, McDonald BA: **The global genetic structure of the wheat pathogen *Mycosphaerella graminicola* is characterized by high nuclear diversity, low mitochondrial diversity, regular recombination, and gene flow.** *Fungal Genet Biol* 2003, **38**:286–297.
  10. El Chartouni L, Tisserant B, Siah A, Duyme F, Leducq JB, Deweer C, Fichter-Roisin C, Sanssene J, Durand R, Halama P, Reignault P: **Genetic diversity and population structure in French populations of *Mycosphaerella graminicola*.** *Mycologia* 2011, **103**:764–774.
  11. Kabbage M, Leslie JF, Hulbert SH, Bockus WW: **Comparison of natural populations of *Mycosphaerella graminicola* from single fields in Kansas and California.** *Physiol Mol Plant Pathol* 2009, **74**:55–59.
  12. Abrinbana M, Mozafari J, Shams-bakhsh M, Mehrabi R: **Genetic structure of *Mycosphaerella graminicola* populations in Iran.** *Plant Pathol* 2010, **59**:829–838.
  13. Schnieder F, Koch G, Jung C, Verreet JA: **Genotypic diversity of the wheat leaf blotch pathogen *Mycosphaerella graminicola* (anamorph) *Septoria tritici* in Germany.** *Eur J Plant Pathol* 2001, **107**:285–290.
  14. Stukenbrock EH, Jorgensen FG, Zala M, Hansen TT, McDonald BA, Schierup MH: **Whole-genome and chromosome evolution associated with host adaptation and speciation of the wheat pathogen *Mycosphaerella graminicola*.** *PLoS Genet* 2010, **6**(12):e1001189.
  15. Waalwijk C, Mendes O, Verstappen ECP, de Waard MA, Kema GHJ: **Isolation and characterization of the mating-type idiomorphs from the wheat septoria leaf blotch fungus *Mycosphaerella graminicola*.** *Fungal Genet Biol* 2002, **35**:277–286.
  16. Zhan J, Kema GHJ, Waalwijk C, McDonald BA: **Distribution of mating type alleles in the wheat pathogen *Mycosphaerella graminicola* over spatial scales from lesions to continents.** *Fungal Genet Biol* 2002, **36**:128–136.
  17. Gurung S, Goodwin SB, Kabbage M, Bockus WW, Adhikari TB: **Genetic differentiation at microsatellite loci among populations of *Mycosphaerella graminicola* from California, Indiana, Kansas, and North Dakota.** *Phytopathology* 2011, **101**:1251–1259.
  18. Razavi M, Hughes GR: **Molecular variability of *Mycosphaerella graminicola* as detected by RAPD markers.** *J Phytopathol* 2004, **152**:543–548.
  19. Schnieder F, Koch G, Jung C, Verreet JA: **The application of molecular markers for genetic characterization of *Septoria tritici* populations.** *Z Pflanzenk Pflanzens-J Plant Dis Prot* 1998, **105**:452–461.
  20. Owen PG, Pei M, Karp A, Royle DJ, Edwards KJ: **Isolation and characterization of microsatellite loci in the wheat pathogen *Mycosphaerella graminicola*.** *Mol Ecol* 1998, **7**:1611–1612.
  21. Razavi M, Hughes GR: **Microsatellite markers provide evidence for sexual reproduction of *Mycosphaerella graminicola* in Saskatchewan.** *Genome* 2004, **47**:789–794.
  22. Goodwin SB, van der Lee TAJ, Cavaletto JR, Hekkert BTL, Crane CF, Kema GHJ: **Identification and genetic mapping of highly polymorphic microsatellite loci from an EST database of the septoria tritici blotch pathogen *Mycosphaerella graminicola*.** *Fungal Genet Biol* 2007, **44**:398–414.
  23. Boukef S, McDonald BA, Yahyaoui A, Rezgui S, Brunner PC: **Frequency of mutations associated with fungicide resistance and population structure of *Mycosphaerella graminicola* in Tunisia.** *Eur J Plant Pathol* 2012, **132**:111–122.
  24. Negrini R, D'Andrea M, Crepaldi P, Colli L, Nicoloso L, Guastella AM, Sechi T, Bordonaro S, Ajmone-Marsan P, Pilla F, Econogene C: **Effect of microsatellite outliers on the genetic structure of eight Italian goat breeds.** *Small Rumin Res* 2012, **103**:99–107.
  25. Hayden MJ, Nguyen TM, Waterman A, Chalmers KJ: **Multiplex-ready PCR: a new method for multiplexed SSR and SNP genotyping.** *BMC Genomics* 2008, **9**:80.
  26. Crane C: **Patterned sequence in the transcriptome of vascular plants.** *BMC Genomics* 2007, **8**:173.
  27. Goodwin SB, Zismann VL: **Phylogenetic analyses of the ITS region of ribosomal DNA reveal that *Septoria passerinii* from barley is closely related to the wheat pathogen *Mycosphaerella graminicola*.** *Mycologia* 2001, **93**:934–946.
  28. Peakall R, Smouse PE: **GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research.** *Mol Ecol Notes* 2006, **6**:288–295.
  29. Antao T, Lopes A, Lopes RJ, Beja-Pereira A, Luikart G: **LOSITAN: a workbench to detect molecular adaptation based on a F(st)-outlier method.** *BMC Bioinformatics* 2008, **9**:323.
  30. Suffert F, Sache I: **Relative importance of different types of inoculum to the establishment of *Mycosphaerella graminicola* in wheat crops in north-west Europe.** *Plant Pathol* 2011, **60**:878–889.
  31. Nei M: **Estimation of average heterozygosity and genetic distance from a small number of individuals.** *Nature* 1978, **89**:583–590.
  32. Agapow PM, Burt A: **Indices of multilocus linkage disequilibrium.** *Mol Ecol Notes* 2001, **1**:101–102.
  33. Voorrips RE: **MapChart: software for the graphical presentation of linkage maps and QTLs.** *J Hered* 2002, **93**:77–78.

doi:10.1186/1756-0500-7-373

**Cite this article as:** Gautier et al.: Development of a rapid multiplex SSR genotyping method to study populations of the fungal plant pathogen *Zymoseptoria tritici*. *BMC Research Notes* 2014 **7**:373.

**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

