A highly multiplexed assay to monitor virulence, fungicide resistance and gene flow in the fungal wheat pathogen *Zymoseptoria tritici*

Hadjer Bellah[1], Gwilherm Gazeau[2], Sandrine Gélisse[2], Reda Amezrou[2], Thierry C. Marcel[2,*], Daniel Croll[1,*]

[1] Laboratory of Evolutionary Genetics, Institute of Biology, University of Neuchâtel, CH-2000, Neuchâtel, Switzerland

[2] Université Paris Saclay, INRAE, UR BIOGER, 78850 Thiverval-Grignon, France

[*] Corresponding authors: thierry.marcel@inrae.fr, daniel.croll@unine.ch

Abstract

Crop pathogens pose severe risks to global food production due to the rapid rise of resistance to pesticides and host resistance breakdowns. Predicting future risks requires monitoring tools to identify changes in the genetic composition of pathogen populations. Here we report the design of a microfluidics-based amplicon sequencing assay to multiplex 798 loci targeting virulence and fungicide resistance genes, and randomly selected genome-wide markers for the fungal pathogen *Zymoseptoria tritici*. We optimized the primer design by integrating polymorphism data from 632 genomes of the same species. To test the performance of the assay, we genotyped 192 samples in two replicates. The fungus causes one of the most devastating diseases on wheat showing rapid adaptation to fungicides and host resistance. Analysis of the short-read sequence data generated by the assay showed a fairly stable success rate across samples to amplify a large number of loci. The performance was consistent between samples originating from pure genomic DNA as well as material extracted directly from infected wheat leaves. In samples with mixed genotypes, we found that the assay recovers variations in allele frequencies. We explored also the potential of the amplicon assay to recover transposable element insertion polymorphism relevant for fungicide resistance. As a proof-of-concept, we show that the assay recovers the pathogen population structure across French wheat fields. Genomic monitoring of crop pathogens contributes to more sustainable crop protection and yields.

Introduction

Approximately 30 percent of all crop diseases are caused by fungi (Jain et al. 2019). Plant pathogenic fungi affect crops at various life cycle stages and plant tissues, including seeds, root and leaf development, and inflorescence (Gonzalez et al. 2011; Serrato-Diaz et al. 2014; He et al. 2019; Zhang et al. 2020). Yield reductions by pathogenic fungi cause food insecurity and economic losses (Horbach et al. 2011; Avery et al. 2019). Crop protection is primarily achieved through the application of a variety of fungicides and resistance breeding (Nelson et al. 2018; Corkley et al. 2022). However, fungal pathogens have evolved resistance to all major fungicides currently in use (Mohd Assaad et al. 2016). In addition, efforts to breed resistant crop varieties have repeatedly been defeated by rapid evolutionary change in pathogen populations allowing to circumvent resistance mechanisms (Nelson et al. 2018). Predicting future breakdowns in fungicide efficacy and crop resistance remains challenging. Fungicide resistance is monitored across the European continent by analyzing mutations in known target genes related to the fungicide mode of action (Brent and Hollomon 1998; Klappach et al. 2020). However, the rise of pathogen strains defeating crop resistance is not comprehensively monitored. Notable exceptions include the screening of rust fungi (Park et al. 2011; Olivera Firpo et al. 2017; Cook et al. 2021; Fontyn et al. 2022). To reduce damage caused by plant pathogens, a timely and accurate detection of both fungicide resistance mutations and mutations associated with the defeat of crop resistance is essential.

Fungal plant pathogen populations that evolved resistance to specific fungicides harbor numerous mutations in or nearby the genes encoding the targets of the chemical compounds (Deising et al. 2008; Bowyer et al. 2012; Mohd Assaad et al. 2016; Omrane et al. 2017). Similarly, pathogen populations virulent on previously resistant crop varieties have often mutated or deleted a specific set {Citation}of genes that encode proteins recognized by the plant immune system (Kashiwa et al. 2016; Hartmann et al. 2017; Hartmann 2022; Olivera et al. 2022). Fungicide resistance has traditionally been detected using in vitro fungicide sensitivity assays (Saville et al. 2015; Samils et al. 2021). Such analyses require the

isolation and culturing of individual fungal strains that can then be tested for growth on media containing different fungicide concentrations. The fungicide dose that effectively inhibits growth by 50% is determined for comparison among samples (*i.e.*, EC50) (Saville et al. 2015; Wang et al. 2017). The method is laborious and limited to fungal species that can be cultured in absence of the host. With advances in molecular techniques, a number of genetic screening methods have been developed including Sanger sequencing, TaqMan assays based on fluorescently-tagged, allele-specific probes (Schleinitz et al. 2011). In general, such screening approaches are labor-intensive and have low potential for multiplexing large numbers of individual loci. Virulence surveillance of fungal plant pathogens has been implemented for the wheat rust pathogen *Puccinia graminis f. sp. tritici* based on simple sequence repeat (SSR) markers (Zhong et al. 2009; Visser et al. 2011) and used to differentiate the virulent Ug99 race lineages (Jin et al. 2008; Visser et al. 2011). However, these SSR makers have been less useful in distinguishing differentiation within the Ug99 race (Singh et al. 2015). The tomato wilt pathogen *Fusarium oxysporum f. sp. lycopersici* (*Fol*) was monitored using loop-mediated isothermal amplification (LAMP) to distinguish the pathogenic race 3 from other races based on point mutations in the secreted in xylem 3 (SIX3) gene characteristic of race 3. However, LAMP assays can be expensive given costs of individual probes (Ayukawa et al. 2017). Therefore, there is a need for cost effective, sensitive, and quantitative genotyping tools that are able to identify specific mutations associated with pathogen emergence to improve early detection and effective disease management.

The advent of next generation sequencing (NGS) approaches has removed a series of limitations in pathogen monitoring. The most general application of NGS techniques is whole genome sequencing (WGS) that can be used to detect single nucleotide polymorphisms (SNPs) and structural variation (Cao et al. 2016).

Applications of WGS have contributed to the mapping and characterization of virulence and resistance factors primarily through genome-wide association mapping (Mohd□Assaad et al. 2016; Zhong et al. 2017; Pereira et al. 2020; Hartmann 2022). Low-cost, high-throughput methods based on NGS include

reduced representation sequencing genotyping methods such as restriction-site-associated DNA sequencing (RAD-seq) and Genotyping-by-Sequencing (GBS), both methods rely on restriction enzymes to reduce genome size and complexity and exploring SNPs adjacent to restriction enzyme sites (Baird et al. 2008; Elshire et al. 2011). However, such genotyping approaches assess only mutations near restriction enzyme cut sites. Applications in fungal pathogens mainly include assessments of recombination rates and mapping of quantitative traits in experimental crosses (Croll et al. 2015; Talas and McDonald 2015; Lendenmann et al. 2016; Stewart et al. 2018; Wyka et al. 2022). The analysis of individual regions involved in fungicide resistance has been improved by the recent development of a PacBio long-read sequencing assay based on the multiplex amplification of target genes in fungal wheat pathogen *Zymoseptoria tritici*. The main advantage is the ability to generate long-reads capturing significant haplotype information of individual strains revealing a series of alterations conferring increased resistance in response to different commercial fungicides. However, due to the varying amplicon sizes generated by this assay two separate multiplex PCRs were required to separate shorter and longer amplicons (Samils et al. 2021). High degrees of multiplexing for amplicons and samples were recently achieved using two parallel approaches for animal and plant species. Genotyping-in-thousands by sequencing (GT-seq) is based on multiplex PCR targeted amplicon sequencing to simultaneously genotype thousands of loci and hundreds of samples in a single Illumina sequencing run (Campbell et al., 2015). A limitation of this approach is the extended time required for its development (about ~4 months according to (Meek and Larson 2019). One challenge to overcome is imbalanced amplification of individual loci and samples. Such bias can be reduced by the use of microfluidics assays that physically separate sets of amplicons and samples (Kessler et al. 2022).

The fungal pathogen *Z. tritici* causes one of the economically most important wheat diseases called Septoria tritici blotch (STB) (Fones and Gurr 2015). The pathogen has emerged at the onset of wheat domestication in the Middle East (Stukenbrock et al. 2007) and has since spread to all wheat-producing areas of the world (Zhan et al. 2003). Populations have evolved resistance to all commercially used

fungicides and repeatedly across continents (Hartmann 2022). Major routes to resistances included the rise of mutations in genes encoding the targets of the fungicide, in particular in *CYP51* encoding the target of azoles (Mohd Assaad et al. 2016; Corkley et al. 2022). Furthermore, upregulation of the transporter gene *MFS1* due to the insertion of transposable elements in the promoter region contributed to azole resistance (Omrane et al. 2017). The rise of succinate dehydrogenase inhibitor (SDHI) resistance mutations are the most recent of the observed gains in resistance (Fungicide Resistance Action Committee, FRAC, 2021). In parallel to the rapid evolution to resist fungicides, *Z. tritici* has also surmounted most known resistance factors segregating among wheat cultivars (Brown et al. 2015). Association mapping in *Z. tritici* has recently revealed specific mutations underlying the gain of virulence on previously resistant wheat cultivars including cultivars carrying the resistance gene *Stb6* and others (Zhong et al. 2017; Meile et al. 2018; Singh et al. 2021). Recently, Amezrou et al. (*preprint available shortly*) identified an additional 58 candidate virulence genes based on association mapping on 12 wheat differential cultivars. The genes linked to gains of virulence are typically referred to as effector genes and show rapid evolutionary change in populations of *Z. tritici* (Hartmann et al. 2017; Meile et al. 2018; Singh et al. 2021). Gene flow among *Z. tritici* populations is leading to significant weak differentiation at the continental scale and high local diversity (Linde et al. 2002; Zhan et al. 2003; Singh et al. 2021). Monitoring of fungicide resistance mutations is mainly achieved through the sequencing of target genes including the recent development of long-read sequencing assays (Samils et al. 2021). A joint monitoring of virulence mutations and genetic diversity is lacking though.

Here, we report the design and validation of a microfluidics based multiplex targeted amplicon sequencing assay that allows the simultaneous monitoring of mutations in fungicide resistance genes and effector genes associated with a wide range of host resistance factors. In addition, we enable the monitoring of hundreds of equally spaced polymorphisms along chromosomes to identify recent changes in the genetic composition of pathogen populations. We validate the performance of the assay using

replication, sensitivity analyses to low input DNA, mixed samples as well as the performance on DNA directly obtained from infected wheat leaves.

Results

Marker design based on whole-genome sequenced individuals across species

We used whole-genome sequencing datasets of 632 *Z. tritici* isolates collected in Oceania (Australia, New Zealand), the United States, Switzerland, France, and Israel to identify segregating SNPs and improve the design of a total of 798 amplicons. Genome-wide polymorphism was used to mask sites to avoid primer mismatches and amplification drop-outs (Fig. 1A). We designed 25 amplicons across genes associated with fungicide targets and resistance including *CYP51*, *alternative oxidase* (*AOX*), *beta-tubulin* (*TUB1*) and *SDH1-4* genes including *ZtSDHC3* as well as *Cytochrome b (CYTB)* gene (See Supplementary Table S2). For each gene, we prioritized amplicons covering non-synonymous substitution. Due to the complexity of the transposable element insertion polymorphism in the promoter region of the transporter gene *MFS1*, we designed multiple amplicons matching known sequence variants near three insertion sites (Omrane et al. 2017) (See Supplementary Table S2). For loci associated with virulence on diverse cultivars, we retained a set of 67 amplicons successfully passing primer design (See Supplementary Table S2). We also randomly selected SNPs at ~50 kb distances including a SNP in the previously selected fungicide resistance gene *cytochrome b (CYTB)* to monitor the genetic make-up of populations for a total of 691 designed amplicons across all chromosomes (See Supplementary Table S2).

Assessment of loci quality across the targeted sequencing assay

We performed targeted sequencing of all 798 loci based on the Fluidigm Juno system in a single run using microfluidics (Fig. 1B). The 192 samples included four sets of pure DNA from different isolates mixed in equal proportions, ten samples including each DNA of the same three isolates in different proportions, and 178 samples constituted from extracted leaf material from different wheat fields across France mostly (*i.e.* n = 172), Belgium, Ireland and the United Kingdom. The complete set of samples was replicated once for the amplification and Illumina sequencing step. The total sequencing output over both replicates was 2,418,905,407 read pairs and 338.89 Gb. For 31 samples, the amplification and Illumina

sequencing procedures failed in either one of the two replicates of each sample, therefore the failed replicates were eliminated. Across a replicate run (*i.e.* FC2), samples produced between 5,976-173,049,611 read pairs with numbers broadly consistent between the two replicate runs (Fig. 2A). We found that the mapping rate against the reference genome ranged from 96.93-100% among all replicates of samples (Fig. 2B).

To assess the faithful amplification of individual loci, we first focused on the four samples with mixtures of pure fungal DNA of 26 to 30 isolates. Combining the two replicates, we used eight samples to evaluate sequencing read coverage across the 782 amplicons designed outside of the *MFS1* region. We found 17 loci with a read depth of 0. The highest read depth was 1,779,927 for an effector locus on chromosome 1. For the set of genome-wide, equally spaced amplicons on core chromosomes, we retained the locus if the read counts were between 20,000 and 100,000 in the set of reference samples (Fig. 2C). We considered this read count range to reflect the loci consistently amplifying across samples and not showing evidence for duplications. With this filter, we discarded 149 loci falling outside of the read count range (Fig. 2D). For randomly selected markers on accessory chromosomes, we expected lower amplification success because not all isolates of the species carry the locus. We retained loci with a read count between 10,000 and 100,000 in the set of reference samples leading to the rejection of 21 loci (Fig. 2C-D). For randomly selected mitochondrial markers, we found read counts ranging from 203,502 to 1,372,965 in the set of reference samples reflecting the high copy number of mitochondria compared to the nuclear genome. All 12 randomly selected mitochondrial loci were kept. For effector loci, the number of mapped reads ranged from 502 to 1,779,927 reads indicating significant variation in the amplification success and possibly copy number (Fig. 2C). We retained all 67 designed amplicons due to the general interest in polymorphism at such loci (Fig. 2D). For resistance gene loci, the number of mapped reads ranged from 2,986-1,372,965 reads (Fig. 2C). As for effector gene loci, all 24 designed amplicons were retained (Fig. 2D). In addition, we retained the amplicon for the mitochondrial resistance locus of *CYTB* with a read count of 1,372,965. In summary, we retained 521 high-quality loci representing 75% of the randomly

selected markers designed for genetic structure analyses, as well as all 67 effector and 24 fungicide resistance loci (Fig. 2D).

Reproducibility among replicate assays and recovery of allele frequencies

To assess the reproducibility of the sequencing assay, we repeated the amplification and sequencing procedure two times. We found that the number of read pairs recovered for each sample were positively correlated between replicates (r = 0.78, p-value < 0.0001; Fig. 3A). We also found a positive correlation in the mapping rate of reads recovered from the same samples (r = 0.85, p-value < 0.0001; Fig. 3B). To investigate effects on allele frequencies assessed for mixed samples, we compared the pooled DNA of population 41 sample. We used allele frequencies estimated from read depth for the reference and alternative allele at SNP loci. Reference allele frequencies at 201 SNP loci calculated in both replicates of each sample were highly correlated (r = 0.89, p-value < 0.0001) with outliers corresponding to poorly covered loci in either one of the two replicates of the same sample (Fig. 4A). Furthermore, we analyzed allele frequencies in 10 samples constituted from a mix of pure DNA from the same three isolates in different proportions. We focused on the effector locus *AvrStb6* on chromosome 5 with known differences among the three isolates. We found that the allele frequency variation at the SNP position 69,631 among samples with different isolate DNA mixtures was reflecting differences in the proportion of DNA mixtures (Fig. 4B). We analyzed correlations between the proportion of DNA of a specific isolate and allele frequencies (Fig. 4C; Supplementary Figure S2). As expected, based on the known genotypes, increasing proportions of DNA from isolate INRA10-FS1006 tended to be correlated with higher alternative allele frequencies (Fig. 4C). Increasing proportions of DNA from isolate INRA10-FS1022 tended to be correlated with lower alternative allele frequencies (Supplementary Figure S2B). However, neither correlation was significant.

Amplicons for the promoter region of MFS1

The amplicons designed for the promoter region of *MFS1* are matching known haplotypes differing in their insertion of transposable element sequences. Due to the sequence complexity, we chose to first

cluster sequencing reads into individual amplicons instead of directly mapping reads to a *MFS1* haplotype. Analyzing the 10 samples with different DNA mixtures of three isolates including replicates, we identified 10 sequence clusters with at least 22 reads (lowest number observed in sample G3). We used BLAST to retrieve the subset (n =10) of the clustered sequences matching the *MFS1* promoter region. The sequences matched positions from 1-4946 bp (for sample G9) on the consensus *MFS1* sequence with all being upstream of the coding sequence as expected (Fig. 5A). We did not recover any amplicon matching forward and/or reverse primer positions based on the amplicon design (Fig. 5B). However, all amplicons did not match the expected amplicon length most likely due to the complexity of the underlying sequence. Furthermore, the pooled amplification of multiple primer pairs matching the promoter region has likely produced chimeric amplicons in some contexts. We used the retrieved amplicons matching the promoter region to form clusters of near identical BLAST matches based on alignment length and positions. We identified 10 well supported amplicon clusters showing variation in abundance among the analyzed samples (Supplementary Table S3).

Genetic differentiation in French and European wheat field populations

We used the 158 wheat leaf samples infected by *Z. tritici* collected from fields across France with at least five genotyped samples per location and additional samples from Belgium, Ireland and the United Kingdom to assess the genetic structure using the genome-wide marker set. Based on a principal component analysis of 85 genome-wide SNPs, we found that the majority of the European samples grouped homogeneously (Fig. 6A). Focusing on the genetic differentiation among French regions (n = 82 genome-wide SNPs for the French populations only), we found some modest differentiation of genotypes from wheat fields in Midi-Pyrénées and Champagne (Fig. 6B). However, the overall differentiation of the field samples was low with the first and second principal component explaining only ~4%.

Discussion

We developed a microfluidics-based amplicon sequencing assay combining the advantages of high-throughput sequencing and multiplex PCR. We assessed the performance of 798 loci to reliably and sensitively genotype randomly selected genome-wide markers, as well as virulence and fungicide resistance genes in a diverse set of *Z. tritici* samples. We show that a large portion of the designed markers can be amplified consistently across samples and used to monitor the emergence of relevant mutations. The set of genome-wide markers provides means to assess the genetic structure of the pathogen directly from field collected wheat leaves.

Within-species polymorphism can lead to amplification failures due to mismatching primers. We considered this issue particularly relevant for the wheat pathogen *Z. tritici* as the species harbors genetically highly diverse populations within single fields (Singh et al. 2021). As expected, we detected a high number of SNPs in regions intended for amplicon design leading to the rejection of amplicon candidates prior to the genotyping stage. Furthermore, we noticed targeted regions with weak amplification success. The poor performance of some primer pairs is most likely explained by a combination of factors. First, we ignored low-frequency SNPs at the masking stage to be able to proceed to the amplicon design for more loci. Second, our species-wide genomic survey of SNPs may have missed polymorphisms present in the assayed samples. The filtering thresholds can be adjusted and more genome sequencing datasets could be included in future amplicon design efforts. Despite some failed attempts at amplifying individual loci, we obtained high degrees of sequencing read coverage for most loci. Most samples yielded hundreds to thousands of reads for each locus. Such deep coverage across the amplicon assay provides a detailed picture of genotypic diversity particularly for mixed samples directly obtained from infected leaves. A major limitation with the multiplexed amplicon sequencing assay is the shortness of the amplified sequence (~200 bp). The short amplicon length ensures a high degree of multiplexing by providing stable amplification across the entire assay. However, longer amplicons would

be needed to recover entire sequences (*i.e.* haplotypes) of the azole resistance locus *CYP51* or several effector genes of interest. A potential solution would be to design overlapping amplicons to cover an entire locus. However, this approach was unsuccessful *e.g.* for the effector gene *AvrStb6* providing no sufficiently conserved sections inside or adjacent to the coding sequence for an overlapping amplicon design.

A versatile pathogen genotyping assay should perform well with low-input pure fungal DNA as well as mixed samples containing DNA both from multiple strain genotypes and the host (*i.e.* wheat plants). We find that the assay replicated well across all tested sample types both in terms of the number of recovered reads per sample as well as the proportion of reads that could be mapped to the *Z. tritici* reference genome. We found that in mixed samples (*i.e.* containing more than one genotype), the assay reproduces well the allele frequencies across the two independent genotyping runs. We also assessed the ability of the assay to recover allele frequencies at a specific locus (*i.e.* the effector gene *AvrStb6*) using three-isolate mixtures of different proportions. We found a generally positive correlation of the proportion of a specific isolate in the DNA mixture and the allele associated with the isolate, however the recovery of allele frequencies was not well reproduced across all tested DNA mixtures. Hence, the assay requires further validation if within-sample allele frequencies need to be recovered with high accuracy. The genotyping of TE insertions in the *MFS1* promoter region was not conclusive. The overlapping amplicons and very high levels of sequence polymorphism prevented a clear assignment of amplicons to TE insertion genotypes, but our data opens up a path for a more comprehensive design strategy to capture inserted sequences. Overall, the assay performed well in recovering *Z. tritici* DNA across hundreds of loci even from wheat samples obtained directly from the field.

The microfluidics-based multiplex PCR targeted amplicon sequencing recovered successful genotypes at hundreds of targeted loci. The number of recovered loci for targeted amplicon sequencing remains below untargeted approaches such as RAD-seq and GBS. Untargeted reduced-representation approaches provide

however only genome-wide information on genetic differentiation. This may be informative *e.g.* for virulence profiles in clonal pathogens (Aoun et al. 2020), however this approach is unsuitable to recover genotypes at specific loci. Targeted amplification such as the microfluidics based multiplex PCR performs also well in mixed samples. RAD-seq and GBS are unlikely to perform well if substantial proportions of plant DNA are present, because large plant genomes will typically contain many more restriction sites compared to the fungal genome. Limitations in amplicon length can be overcome using long-read sequencing as developed to monitor fungicide resistance loci in *Z. tritici* (Samils et al. 2021). Long-read sequencing may also help to overcome issues with amplifying the highly polymorphic promoter region of *MFS1*. Long amplicons could capture the entire promoter region instead of focusing on individual insertion points. Ultimately, a combination of different approaches performing highly-multiplexed short reads sequencing and separate long-read sequencing for the most complex loci will be required.

The developed microfluidics-based targeted amplicon assay allows a cost-effective and reproducible monitoring of hundreds of loci to track virulence and fungicide resistance evolution in field populations. The integration of genome-wide markers greatly enhances the information provided, compared to monitoring focused on adaptive loci alone, by providing crucial information about patterns of gene flow. Our study revealed only weak differentiation across Western European countries and among French regions consistent with high levels of gene flow and genetic diversity. Knowledge of genetic structure can help identify recent movements of the pathogen due to natural or human-mediated dispersal. The rapid rise in resistance of *Z. tritici* populations after the application of fungicides can more effectively be monitored due to the large number of loci that can be assayed simultaneously. Furthermore, tracking mutations at effector loci opens new opportunities to track adaptation to different wheat cultivars across regions. With the availability of whole genome sequencing data for an increasing number of crop pathogens, the targeted amplicons could be expanded to simultaneously or separately genotype other major pathogens including rusts to improve the surveillance and management of crop diseases globally.

Methods

Isolates used for the design and validation of the assay

We retrieved genome sequencing information from 632 *Z. tritici* isolates collected across the global distribution range of wheat. Isolates included six different populations with a sample size of 29-178. A total of 88 isolates were collected in Australia including Tasmania in 2001 and 2015 (Hartmann et al. 2017). Additional isolates from Oceania included 75 isolates collected in New Zealand in 2013 and 2015 (Feurtey et al. 2022). A total of 154 isolates were collected in Oregon, USA, in 1990 and 2015 (Hartmann et al. 2017). 178 isolates were in wheat fields near Zurich in Switzerland in 1999 and 2016 (Hartmann et al. 2017; Singh et al. 2021) and 29 isolates were isolated in the Nahal Oz region in Israel in 1992 (Hartmann et al. 2017). Finally, 108 isolates were retrieved from a panel of French isolates (Zhong et al. 2017). We assessed the performance of the microfluidics assay using different sets of samples collected from wheat fields in Europe. Four samples included equimolar DNA mixtures of 26 to 30 isolates obtained by culturing single spore isolates from field-collected wheat leaves. Three single spore isolates identified as INRA10-FS1006, INRA10-FS1022 and IPO-09455 were collected in 2009 and 2010 in the Ile-de-France region and were used to create DNA mixtures in ten different proportions. Finally, a total of 178 samples were obtained by extracting DNA directly from infected wheat leaves collected in different regions of France, Belgium, Ireland and United Kingdom. See Supplementary Table S1 for details on all included samples.

SNP calling and identification of polymorphisms for the amplicon design

We performed read alignment and SNP discovery for the generated genomic datasets, as previously described (Hartmann et al. 2017; Zhong et al. 2017). In summary, we trimmed raw Illumina reads using Trimmomatic v. 0.38 (Bolger et al. 2014) and mapped retained reads to the reference genome IPO323 (Goodwin et al. 2011) using bowtie v2.3.5 (Langmead and Salzberg 2012). We used the Genome Analysis Toolkit (GATK) v4.0.1 (Van der Auwera et al. 2013) including the HaplotypeCaller tool to

identify candidate SNPs. We filtered for a set of high-quality polymorphisms using the GATK VariantFiltration tool and vcftools v.0.1.15 (Danecek et al. 2011). A more extensive description of the filtering procedures and validations are available (Hartmann et al. 2018).

Polymorphism selection for neutral markers, effector and fungicide resistance genes

Effector candidate genes were retrieved from GWAS focused to identify candidate effectors interacting with major wheat resistance genes (Amezrou et al., unpublished, Hartmann et al. 2017, Meile et al. 2018). We included 65 candidate effector genes showing a significant association for symptom development on at least one wheat cultivar. We designed at least one amplicon overlapping the most significantly associated SNP in each of the effector genes. If a significantly associated SNP could not be reproduced in the worldwide isolate collection, a random nearby SNP (within ~200 bp) was selected as the target for the amplicon design. If a different SNP was selected, we filtered for SNPs with a minor allele count of 5 and a minimal genotyping rate of 80%. For the effector gene *AvrStb6,* we designed two additional amplicons to cover polymorphism in the coding sequence. To monitor fungicide resistance gene mutations, we covered 25 genes related to fungicide resistance in *Z. tritici* populations including the mitochondrial gene *CYTB* associated with resistance to strobilurin (QoI), the nuclear genes beta tubulin 1 (carbendazim resistance), *CYP51* (azole resistance), *AOX* (alternate oxidase inhibitor resistance), as well as *SDH1, SDH2, SDH3* and *SDH4* (SDHI resistance). The amplicons covered known resistance mutations if known. If no mutation was documented in *Z. tritici,* the amplicon covered randomly selected SNPs in the coding sequence. Similar to the procedure for effector loci, if a known SNP associated with fungicide resistance could not be recovered, we selected a SNP within ~100 bp (minor allele count of 3, minimum genotyping rate 80%). The broader inclusion of polymorphisms for filtering was possible due to the generally lower degree of detected variants in resistance genes. We defined an additional amplicon to target the paralog of *SDH3* (*ZtSDHC3*; Steinhauer et al., 2019). For this, we analyzed the paralog sequence discovered in the pangenome of *Z. tritici* (Badet et al. 2020). Multidrug fungicide resistance in *Z. tritici* is mediated by transposable element (TE) insertions in the promotor region of the transporter *MFS1*. We designed 16

amplicons covering three previously reported TE insertions and haplotypes (Omrane et al. 2017). The amplicons were designed to either amplify if an insertion was present or not. Amplicons were designed on a consensus sequence of previously described haplotypes (Omrane et al. 2017). In addition to polymorphisms related to virulence and fungicide resistance, we randomly selected equally spaced polymorphisms along all 21 chromosomes to capture neutral population structure. For this, we selected 691 SNPs with a minor allele frequency of 5% and a minimal genotyping rate of 80%. SNPs were selected at a distance of 50 kb (if available) using the *--thin* option in *vcftools*. In conclusion, a total of 798 amplicons were designed for virulence, fungicide resistance as well as gene flow tracking across *Z.tritici* populations. See Supplementary Table S2 for details on all selected effector and fungicide resistance genes as well as whole genome neutral markers.

Amplicon design

For genome-wide markers and markers in virulence and fungicide resistance genes (except *ZtSDHC3* and MFS1), we produced a 401 bp fasta sequence from the reference genome centered on the SNP to target. We identified the SNP using the IUPAC code marked with parenthesis according to the design instructions. To improve genotyping success across a broad range of *Z. tritici* genotypes, we masked known polymorphic sites near the targeted SNP to prevent the accidental primer design in known polymorphic regions. We used bcftools v1.9 (Li 2011) to mask non-target sites showing evidence for polymorphisms in the panel of 632 analyzed isolates using the -I option of the consensus command and re-wrote sequences with samtools v1.9 (Li et al. 2009). For resistance and effector genes, we used a minor allele count of 3 and 5, respectively to consider the polymorphism for masking. For genome-wide markers, we used a minor allele frequency cut-off of 5%. If the resulting sequence contained more than 10% masked sites, the amplicon was not considered further. Additional sequences were excluded by Fluidigm Inc. if the sequences failed to yield adequate primer candidates. If the initial amplicon design had failed, we repeated the procedure for effector loci but relaxed the filter to consider only SNPs with a minor allele count of $\geq$25.

DNA extractions and microfluidics assay

DNA extractions to test the microfluidics assay were performed using the following procedures. For pure cultures and directly from infected wheat leaves using DNeasy® Plant Mini Kit (Qiagen, Hilden, Germany). DNA was quantified using a Qubit 2.0 fluorometer (Thermo Fisher, Waltham, Massachusetts, USA). We followed the Fluidigm Inc. (San Francisco, California, USA) Juno™ targeted amplicon sequencing protocol according to the manufacturer's protocol. As input DNA, we used the following 1.5-200 ng of total amount (See Supplementary Table S1). We performed the entire microfluidics procedure twice independently on different Juno LP 192.24 integrated fluidic circuits plate (IFC). Libraries were prepared following the manufacturer's protocol. Target amplicons were generated for each sample and pools of primers using PCR on a specialized thermocycler (Juno system; Fluidigm). Illumina sequencing was performed in paired-end mode to generate 100 bp reads on the NovaSeq™6000 platform at Integragen Inc. (Evry, France) and produced 363.89 Gb of raw sequencing data for both independent chips combined.

Amplicon sequence data analyses

We used Trimmomatic v0.38 (Bolger et al. 2014) with the following parameters: LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:36. Due to the short amplicon length compared to the read lengths, we used FLASH v1.2.11 (Magoč and Salzberg 2011) to merge forward and reverse reads per pair into single pseudo-reads. Finally, pseudo-reads were aligned to the IPO323 reference genome using bowtie2 v2.3.5 (Goodwin et al. 2011; Langmead and Salzberg 2012). We assessed individual read counts at each analysis step using MultiQC v.1.7 (Ewels et al. 2016). After individual genotyping using the GATK HaplotypeCaller tool, we performed multi-sample genotype calling using CombineGVCFs and GenotypeGVCFs (Van der Auwera and O'Connor 2020). Variant sites were removed if these met the following conditions: QD < 5, MQ < 20, -2 > ReadPosRankSum > 2, -2 > MQRankSum > 2, -2 >

BaseQRankSum > 2. For amplicons targeting the promoter region of *MFS1*, we first used seqtk (Shen et al. 2016) to subsample 10.000.000 reads from large merged paired-end reads FASTQ files and we then performed a clustering analysis of Illumina reads to obtain read sets originating from the same locus. We used CD-HIT-EST (Fu et al. 2012) with an identity threshold set to 100% to cluster sequencing reads. For each cluster, the representative sequence identified by CD-HIT-EST was aligned to the *MFS1* promoter consensus sequence using BLASTn 2.12.0 (Altschul et al. 1990). Only BLASTn best hits with a bit score above 100 and identity > 90% were kept. To identify clusters of nearly identical hits based on position and identity, we performed *k*-means clustering with the R packages {factoextra}(Kassambara and Mundt 2020), {clustertend} (Wright et al. 2022), {cluster} (Maechler et al. 2012), {NbClust} (Charrad et al. 2014). For each sample, we identified the optimal number of clusters (K = 1–10) by performing a silhouette analysis (Rousseeuw 1987).

Data visualization and population genetic analyses

Data analyses were performed using R 4.0.4 (R Core Team 2021). The R packages included in {tidyverse} (Wickham et al. 2019) were used for summarizing and plotting coverage across loci, visualizing retained SNPs, the outcomes of different filtering stages and genotyping. We used bcftools v1.9 (Li 2011) to calculate allele frequencies at SNP loci. The allele frequency correlation between both flow cells chips was analyzed with the R package {report} (Makowski et al. 2021) and visualized using {ggpubr} (Kassambara 2020) and {ggplot2} (Wickham 2016). To analyze genetic diversity and population structure, we performed a principal component analysis (PCA) using the R packages {vcfR} (Knaus and Grünwald 2017), {adegenet}(Jombart and Ahmed 2011), {ade4} (Dray and Dufour 2007) and {ggplot2} (Wickham 2016). For population analyses, we focused only on the second replicate (flow cell) and genome-wide SNPs were filtered for a minor allele frequency of 0.05 and allowing for 20% missing data (--max-missing 0.8).

Data availability: Raw sequencing data is available on the NCBI Sequence Read Archive (SRA) under BioProject PRJNA847707.

Acknowledgements: We thank Anne-Sophie Walker for providing infected leaf samples that were used to develop the assay. We are grateful for the sequence alignment shared by Sabine Fillinger. The microfluidics assay was conducted on the genotyping platform GENTYANE at INRAE Clermont-Ferrand (https://gentyane.clermont.inrae.fr/), with the help of Rachel Fourdin and Lydia Jaffrelo.

Funding: HB was supported by the Swiss State Secretariat for Education, Research and Innovation (SERI) through a Swiss Government Excellence Scholarship. Funding was also awarded by the French Fund to support Plant Breeding (FSOV 2018 S-Div*R*) to TM and DC. INRAE BIOGER benefits from the support of Saclay Plant Sciences-SPS (ANR-17-EUR-0007).

References

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J., 1990. Basic local alignment search tool. *Journal of molecular biology* [online], 215 (3), 403–410. Available from: https://doi.org/10.1016/S0022-2836(05)80360-2.

Aoun, M., Kolmer, J. A., Breiland, M., Richards, J., Brueggeman, R. S., Szabo, L. J. and Acevedo, M., 2020. Genotyping-by-sequencing for the study of genetic diversity in Puccinia triticina. *Plant Disease* [online], 104 (3), 752–760. Available from: https://doi.org/10.1094/PDIS-09-19-1890-RE.

Avery, S. V., Singleton, I., Magan, N. and Goldman, G. H., 2019. The fungal threat to global food security. *Fungal biology* [online], 123 (8), 555–557. Available from: https://doi.org/10.1016/j.funbio.2019.03.006.

Ayukawa, Y., Hanyuda, S., Fujita, N., Komatsu, K. and Arie, T., 2017. Novel loop-mediated isothermal amplification (LAMP) assay with a universal QProbe can detect SNPs determining races in plant pathogenic fungi. *Scientific reports* [online], 7 (1), 1–9. Available from: https://doi.org/10.1038/s41598-017-04084-y.

Badet, T., Oggenfuss, U., Abraham, L., McDonald, B. A. and Croll, D., 2020. A 19-isolate reference-quality global pangenome for the fungal wheat pathogen Zymoseptoria tritici. *BMC biology* [online], 18 (1), 1–18. Available from: https://doi.org/10.3929/ethz-b-000400871.

Baird, N. A., Etter, P. D., Atwood, T. S., Currey, M. C., Shiver, A. L., Lewis, Z. A., Selker, E. U., Cresko, W. A. and Johnson, E. A., 2008. Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PloS one* [online], 3 (10), e3376. Available from: https://doi.org/10.1371/journal.pone.0003376.

Bolger, A. M., Lohse, M. and Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* [online], 30 (15), 2114–2120. Available from: https://doi.org/10.1093/bioinformatics/btu170.

Bowyer, P., Mosquera, J., Anderson, M., Birch, M., Bromley, M. and Denning, D. W., 2012. Identification of novel genes conferring altered azole susceptibility in Aspergillus fumigatus. *FEMS microbiology letters* [online], 332 (1), 10–19. Available from: https://doi.org/10.1111/j.1574-6968.2012.02575.x.

Brent, K. J. and Hollomon, D. W., 1998. *Fungicide resistance: the assessment of risk (Vol. 2)*. Global Crop Protection Federation Brussels, Belgium.

Brown, J. K., Chartrain, L., Lasserre-Zuber, P. and Saintenac, C., 2015. Genetics of resistance to Zymoseptoria tritici and applications to wheat breeding. *Fungal Genetics and Biology* [online], 79, 33–41. Available from: https://doi.org/10.1016/j.fgb.2015.04.017.

Cao, K., Zhou, Z., Wang, Q., Guo, J., Zhao, P., Zhu, G., Fang, W., Chen, C., Wang, X., Wang, X., Tian, Z. and Wang, L., 2016. Genome-wide association study of 12 agronomic traits in peach. *Nature Communications* [online], 7 (1), 13246. Available from: https://doi.org/10.1038/ncomms13246.

Charrad, M., Ghazzali, N., Boiteau, V. and Niknafs, A., 2014. {NbClust}: An {R} Package for Determining the Relevant Number of Clusters in a Data Set. *Journal of Statistical Software* [online], 61 (6), 1--36. Available from: https://doi.org/10.18637/jss.v061.i06.

Cook, N. M., Chng, S., Woodman, T. L., Warren, R., Oliver, R. P. and Saunders, D. G., 2021. High frequency of fungicide resistance□associated mutations in the wheat yellow rust

pathogen Puccinia striiformis f. sp. tritici. *Pest Management Science* [online], 77 (7), 3358–3371. Available from: https://doi.org/10.1002/ps.6380.

Corkley, I., Fraaije, B. and Hawkins, N., 2022. Fungicide resistance management: Maximizing the effective life of plant protection products. *Plant Pathology* [online], 71 (1), 150–169. Available from: https://doi.org/10.1111/ppa.13467.

Croll, D., Lendenmann, M. H., Stewart, E. and McDonald, B. A., 2015. The impact of recombination hotspots on genome evolution of a fungal plant pathogen. *Genetics* [online], 201 (3), 1213–1228. Available from: https://doi.org/10.1534/genetics.115.185280.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T. and Sherry, S. T., 2011. The variant call format and VCFtools. *Bioinformatics* [online], 27 (15), 2156–2158. Available from: https://doi.org/10.1093/bioinformatics/btr330.

Deising, H. B., Reimann, S. and Pascholati, S. F., 2008. Mechanisms and significance of fungicide resistance. *Brazilian Journal of Microbiology* [online], 39 (2), 286–295. Available from: https://doi.org/10.1590/S1517-838220080002000017.

Dray, S. and Dufour, A.-B., 2007. The ade4 Package: Implementing the Duality Diagram for Ecologists. *Journal of Statistical Software* [online], 22 (4), 1–20. Available from: https://doi.org/10.18637/jss.v022.i04 [Accessed 6 Mar 2022].

Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S. and Mitchell, S. E., 2011. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PloS one* [online], 6 (5), e19379. Available from: https://doi.org/10.1371/journal.pone.0019379.

Ewels, P., Magnusson, M., Lundin, S. and Käller, M., 2016. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics (Oxford, England)* [online], 32 (19), 3047–3048. Available from: https://doi.org/10.1093/bioinformatics/btw354.

Feurtey, A., McDonald, M. C., Milgate, A., Solomon, P., Warren, R., Puccetti, G., Scalliet, G., Gout, L., Marcel, T., Suffert, F., Alassimone, J., Grigoriev, I., Goodwin, S. B., Genissel, A., Lebrun, M.-H., Kema, G., McDonald, B. A. and Croll, D., 2022. A thousand-genome panel retraces the global spread and climatic adaptation of a major fungal crop pathogen. *BioRxiv (DOI pending)*.

Fones, H. and Gurr, S., 2015. The impact of Septoria tritici Blotch disease on wheat: An EU perspective. *Fungal genetics and biology* [online], 79, 3–7. Available from: https://doi.org/10.1016/j.fgb.2015.04.004.

Fontyn, C., Zippert, A.-C., Delestre, G., Marcel, T. C., Suffert, F. and Goyeau, H., 2022. Is virulence phenotype evolution driven exclusively by Lr gene deployment in French Puccinia triticina populations? *Plant Pathology* [online], 00, 1–14. Available from: https://doi.org/10.1111/ppa.13599.

Fu, L., Niu, B., Zhu, Z., Wu, S. and Li, W., 2012. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* [online], 28 (23), 3150–3152. Available from: https://doi.org/10.1093/bioinformatics/bts565.

Gonzalez, M., Pujol, M., METRAUX, J., GONZALEZ□GARCIA, V., Bolton, M. D. and BORRÁS□HIDALGO, O., 2011. Tobacco leaf spot and root rot caused by Rhizoctonia solani Kühn. *Molecular plant pathology* [online], 12 (3), 209–216. Available from: https://doi.org/10.1111/j.1364-3703.2010.00664.x.

Goodwin, S. B., Ben M'Barek, S., Dhillon, B., Wittenberg, A. H., Crane, C. F., Hane, J. K., Foster, A. J., Van der Lee, T. A., Grimwood, J. and Aerts, A., 2011. Finished genome of the fungal wheat pathogen Mycosphaerella graminicola reveals dispensome structure, chromosome plasticity, and stealth pathogenesis. *PLoS genetics* [online], 7 (6), e1002070. Available from: https://doi.org/10.1371/journal.pgen.1002070.

Hartmann, F. E., 2022. Using structural variants to understand the ecological and evolutionary dynamics of fungal plant pathogens. *New Phytologist* [online], 234 (1), 43–49. Available from: https://doi.org/10.1111/nph.17907.

Hartmann, F. E., McDonald, B. A. and Croll, D., 2018. Genome□wide evidence for divergent selection between populations of a major agricultural pathogen. *Molecular ecology* [online], 27 (12), 2725–2741. Available from: https://doi.org/10.1111/mec.14711.

Hartmann, F. E., Sánchez-Vallet, A., McDonald, B. A. and Croll, D., 2017. A fungal wheat pathogen evolved host specialization by extensive chromosomal rearrangements. *The ISME journal* [online], 11 (5), 1189–1204. Available from: https://doi.org/10.1038/ismej.2016.196.

He, C., Zhang, Y., Zhou, W., Guo, Q., Bai, B., Shen, S. and Huang, G., 2019. Study on stripe rust (Puccinia striiformis) effect on grain filling and seed morphology building of special winter wheat germplasm Huixianhong. *Plos one* [online], 14 (5), e0215066. Available from: https://doi.org/10.1371/journal.pone.0215066.

Horbach, R., Navarro-Quesada, A. R., Knogge, W. and Deising, H. B., 2011. When and how to kill a plant cell: infection strategies of plant pathogenic fungi. *Journal of plant physiology* [online], 168 (1), 51–62. Available from: https://doi.org/10.1016/j.jplph.2010.06.014.

Jain, A., Sarsaiya, S., Wu, Q., Lu, Y. and Shi, J., 2019. A review of plant leaf fungal diseases and its environment speciation. *Bioengineered* [online], 10 (1), 409–424. Available from: https://doi.org/10.1080/21655979.2019.1649520.

Jin, Y., Szabo, L., Pretorius, Z., Singh, R., Ward, R. and Fetch Jr, T., 2008. Detection of virulence to resistance gene Sr24 within race TTKS of Puccinia graminis f. sp. tritici. *Plant Disease* [online], 92 (6), 923–926. Available from: https://doi.org/10.1094/PDIS-92-6-0923.

Jombart, T. and Ahmed, I., 2011. adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics (Oxford, England)* [online], 27 (21), 3070–3071. Available from: https://doi.org/10.1093/bioinformatics/btr521.

Kashiwa, T., Suzuki, T., Sato, A., Akai, K., Teraoka, T., Komatsu, K. and Arie, T., 2016. A new biotype of Fusarium oxysporum f. sp. lycopersici race 2 emerged by a transposon-driven mutation of avirulence gene AVR1. *FEMS Microbiology Letters* [online], 363 (14), fnw132. Available from: https://doi.org/10.1093/femsle/fnw132.

Kassambara, A., 2020. ggpubr: 'ggplot2' Based Publication Ready Plots. [online]. Available from: https://CRAN.R-project.org/package=ggpubr.

Kassambara, A. and Mundt, F., 2020. actoextra: Extract and Visualize the Results of Multivariate Data Analyses. [online]. Available from: https://CRAN.R-project.org/package=factoextra.

Kessler, C., Brambilla, A., Waldvogel, D., Camenisch, G., Biebach, I., Leigh, D. M., Grossen, C. and Croll, D., 2022. A robust sequencing assay of a thousand amplicons for the high□throughput population monitoring of Alpine ibex immunogenetics. *Molecular Ecology Resources* [online], 22 (1), 66–85. Available from: https://doi.org/10.1111/1755-0998.13452.

Klappach, B. K., Stammler, G., Bryson, R., Semar, M., Mehl, B. A., Steiger, D., Derpmann, J., Genet, C. J.-L., MBoup, M. and Sierotzki, S. H., 2020. Succinate Dehydrogenase Inhibitor (SDHI) Working Group. *In*: [online]. Meeting on January (Vol. 21, p. 22), 22. Available from: https://www.frac.info/docs/default-source/working-groups/sdhi-fungicides/sdhi-meeting-minutes/minutes-of-the-2020-sdhi-meeting-21-22th-of-january-2020-with-recommendations-for-2020.pdf?sfvrsn=5918499a_2.

Knaus, B. J. and Grünwald, N. J., 2017. vcfr: a package to manipulate and visualize variant call format data in R. *Molecular ecology resources* [online], 17 (1), 44–53. Available from: https://doi.org/10.1111/1755-0998.12549.

Langmead, B. and Salzberg, S. L., 2012. Fast gapped-read alignment with Bowtie 2. *Nature methods* [online], 9 (4), 357–359. Available from: https://doi.org/1038/nmeth.1923.

Lendenmann, M. H., Croll, D., Palma-Guerrero, J., Stewart, E. L. and McDonald, B. A., 2016. QTL mapping of temperature sensitivity reveals candidate genes for thermal adaptation and growth morphology in the plant pathogenic fungus Zymoseptoria tritici. *Heredity* [online], 116 (4), 384–394. Available from: https://doi.org/10.1038/hdy.2015.111.

Li, H., 2011. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* [online], 27 (21), 2987–2993. Available from: https://doi.org/10.1093/bioinformatics/btr509.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G. and Durbin, R., 2009. 1000 genome project data processing subgroup. The sequence alignment/map format and SAMtools. *Bioinformatics* [online], 25 (16), 2078–2079. Available from: https://doi.org/10.1093/bioinformatics/btp352.

Linde, C. C., Zhan, J. and McDonald, B. A., 2002. Population structure of Mycosphaerella graminicola: from lesions to continents. *Phytopathology* [online], 92 (9), 946–955. Available from: https://doi.org/10.1094/PHYTO.2002.92.9.946.

Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M. and Hornik, K., 2012. Cluster: cluster analysis basics and extensions. *R package version* [online], 1 (2), 56. Available from: https://CRAN.R-project.org/package=cluster.

Magoč, T. and Salzberg, S. L., 2011. FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* [online], 27 (21), 2957–2963. Available from: https://doi.org/10.1093/bioinformatics/btr507.

Makowski, D., Ben-Shachar, M. S., Patil, I. and Lüdecke, D., 2021. Automated Results Reporting as a Practical Tool to Improve Reproducibility and Methodological Best Practices Adoption. *CRAN* [online]. Available from: https://github.com/easystats/report.

Meek, M. H. and Larson, W. A., 2019. The future is now: Amplicon sequencing and sequence capture usher in the conservation genomics era. *Molecular Ecology Resources* [online], 19, 795–803. Available from: https://doi.org/10.1111/1755-0998.12998.

Meile, L., Croll, D., Brunner, P. C., Plissonneau, C., Hartmann, F. E., McDonald, B. A. and Sánchez□Vallet, A., 2018. A fungal avirulence factor encoded in a highly plastic genomic region triggers partial resistance to septoria tritici blotch. *New Phytologist* [online], 219 (3), 1048–1061. Available from: https://doi.org/10.1111/nph.15180.

Mohd□Assaad, N., McDonald, B. A. and Croll, D., 2016. Multilocus resistance evolution to azole fungicides in fungal plant pathogen populations. *Molecular ecology* [online], 25 (24), 6124–6142. Available from: https://doi.org/10.1111/mec.13916.

Nelson, R., Wiesner-Hanks, T., Wisser, R. and Balint-Kurti, P., 2018. Navigating complexity to breed disease-resistant crops. *Nature Reviews Genetics* [online], 19 (1), 21–33. Available from: https://doi.org/10.1038/nrg.2017.82.

Olivera Firpo, P., Newcomb, M., Flath, K., Sommerfeldt□Impe, N., Szabo, L., Carter, M., Luster, D. and Jin, Y., 2017. Characterization of Puccinia graminis f. sp. tritici isolates derived from an unusual wheat stem rust outbreak in Germany in 2013. *Plant Pathology* [online], 66 (8), 1258–1266. Available from: https://doi.org/10.1111/ppa.12674.

Olivera, P. D., Villegas, D., Cantero□Martínez, C., Szabo, L. J., Rouse, M. N., Luster, D. G., Bartaula, R., Lopes, M. S. and Jin, Y., 2022. A unique race of the wheat stem rust pathogen with virulence on Sr31 identified in Spain and reaction of wheat and durum cultivars to this race. *Plant Pathology* [online], 71 (4), 873–889. Available from: https://doi.org/10.1111/ppa.13530.

Omrane, S., Audéon, C., Ignace, A., Duplaix, C., Aouini, L., Kema, G., Walker, A.-S. and Fillinger, S., 2017. Plasticity of the MFS1 promoter leads to multidrug resistance in the wheat pathogen Zymoseptoria tritici. *MSphere* [online], 2 (5), e00393-17. Available from: https://doi.org/10.1128/mSphere.00393-17.

Park, R., Fetch, T., Hodson, D., Jin, Y., Nazari, K., Prashar, M. and Pretorius, Z., 2011. International surveillance of wheat rust pathogens: progress and challenges. *Euphytica* [online], 179 (1), 109–117. Available from: https://doi.org/10.1007/s10681-011-0375-4.

Pereira, D., McDonald, B. A. and Croll, D., 2020. The genetic architecture of emerging fungicide resistance in populations of a global wheat pathogen. *Genome biology and evolution* [online], 12 (12), 2231–2244. Available from: https://doi.org/10.1093/gbe/evaa203.

R Core Team, 2021. R: A Language and Environment for Statistical Computing.

Rousseeuw, P. J., 1987. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of computational and applied mathematics* [online], 20, 53–65. Available from: https://doi.org/10.1016/0377-0427(87)90125-7.

Samils, B., Andersson, B., Edin, E., Elfstrand, M., Rönneburg, T., Bucur, D., Hutton, F., Heick, T. M., Hellin, P. and Kildea, S., 2021. Development of a PacBio long-read sequencing assay for high throughput detection of fungicide resistance in Zymoseptoria tritici. *Frontiers in microbiology* [online], 12, 692845. Available from: https://10.3389/fmicb.2021.692845.

Saville, A., Graham, K., Grünwald, N. J., Myers, K., Fry, W. E. and Ristaino, J. B., 2015. Fungicide sensitivity of US genotypes of Phytophthora infestans to six oomycete-targeted compounds. *Plant Disease* [online], 99 (5), 659–666. Available from: https://doi.org/10.1094/PDIS-05-14-0452-RE.

Schleinitz, D., DiStefano, J. K. and Kovacs, P., 2011. Targeted SNP genotyping using the TaqMan® assay. *In*: *Disease gene identification* [online]. Springer, 77–87. Available from: https://doi.org/10.1007/978-1-61737-954-3_6.

Serrato-Diaz, L., Rivera-Vargas, L. and French-Monar, R., 2014. First report of Neofusicoccum mangiferae causing rachis necrosis and Inflorescence blight of mango (Mangifera indica) in Puerto Rico. *Plant Disease* [online], 98 (4), 570–570. Available from: https://doi.org/10.1094/PDIS-08-13-0878-PDN.

Shen, W., Le, S., Li, Y. and Hu, F., 2016. SeqKit: A Cross-Platform and Ultrafast Toolkit for FASTA/Q File Manipulation. *PLOS ONE* [online], 11 (10), 1–10. Available from: https://doi.org/10.1371/journal.pone.0163962.

Singh, N. K., Karisto, P. and Croll, D., 2021. Population-level deep sequencing reveals the interplay of clonal and sexual reproduction in the fungal wheat pathogen Zymoseptoria tritici. *Microbial genomics* [online], 7 (10). Available from: https://doi.org/10.1099/mgen.0.000678.

Singh, R. P., Hodson, D. P., Jin, Y., Lagudah, E. S., Ayliffe, M. A., Bhavani, S., Rouse, M. N., Pretorius, Z. A., Szabo, L. J. and Huerta-Espino, J., 2015. Emergence and spread of new races of wheat stem rust fungus: continued threat to food security and prospects of genetic control. *Phytopathology* [online], 105 (7), 872–884. Available from: https://doi.org/10.1094/PHYTO-01-15-0030-FI.

Steinhauer, D., Salat, M., Frey, R., Mosbach, A., Luksch, T., Balmer, D., Hansen, R., Widdison, S., Logan, G. and Dietrich, R. A., 2019. A dispensable paralog of succinate dehydrogenase subunit C mediates standing resistance towards a subclass of SDHI fungicides in Zymoseptoria tritici. *PLoS pathogens* [online], 15 (12), e1007780. Available from: https://doi.org/10.1371/journal.ppat.1007780.

Stewart, E. L., Croll, D., Lendenmann, M. H., Sanchez□Vallet, A., Hartmann, F. E., Palma□Guerrero, J., Ma, X. and McDonald, B. A., 2018. Quantitative trait locus mapping reveals complex genetic architecture of quantitative virulence in the wheat pathogen Zymoseptoria tritici. *Molecular plant pathology* [online], 19 (1), 201–216. Available from: https://doi.org/10.1111/mpp.12515.

Stukenbrock, E. H., Banke, S., Javan-Nikkhah, M. and McDonald, B. A., 2007. Origin and domestication of the fungal wheat pathogen Mycosphaerella graminicola via sympatric speciation. *Molecular biology and evolution* [online], 24 (2), 398–411. Available from: https://doi.org/10.1093/molbev/msl169.

Talas, F. and McDonald, B. A., 2015. Genome-wide analysis of Fusarium graminearum field populations reveals hotspots of recombination. *BMC genomics* [online], 16 (1), 1–12. Available from: https://doi.org/10.1186/s12864-015-2166-0.

Van der Auwera, G. A., Carneiro, M. O., Hartl, C., Poplin, R., Del Angel, G., Levy□Moonshine, A., Jordan, T., Shakir, K., Roazen, D. and Thibault, J., 2013. From FastQ data to high□confidence variant calls: the genome analysis toolkit best practices pipeline. *Current protocols in bioinformatics* [online], 43 (1), 11–10. Available from: https://doi.org/10.1002/0471250953.bi1110s43.

Van der Auwera, G. A. and O'Connor, B. D., 2020. *Genomics in the cloud: using Docker, GATK, and WDL in Terra* [online]. O'Reilly Media. Available from: https://play.google.com/store/books/details?id=vsXaDwAAQBAJ.

Visser, B., Herselman, L., Park, R. F., Karaoglu, H., Bender, C. M. and Pretorius, Z. A., 2011. Characterization of two new Puccinia graminis f. sp. tritici races within the Ug99 lineage in South Africa. *Euphytica* [online], 179 (1), 119–127. Available from: https://doi.org/10.1007/s10681-010-0269-x.

Wang, J., Bradley, C. A., Stenzel, O., Pedersen, D. K., Reuter-Carlson, U. and Chilvers, M. I., 2017. Baseline sensitivity of Fusarium virguliforme to fluopyram fungicide. *Plant disease* [online], 101 (4), 576–582. Available from: 1https://doi.org/0.1094/PDIS-09-16-1250-RE.

Wickham, H., 2016. *ggplot2: elegant graphics for data analysis* [online]. springer. Springer-Verlag New York. Available from: https://ggplot2.tidyverse.org.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L. and Hester, J., 2019. Welcome to the Tidyverse. *Journal of open*

source software [online], 4 (43), 1686. Available from: https://doi.org/10.21105/joss.01686.

Wright, K., YiLan, L. and RuTong, Z., 2022. clustertend: Check the Clustering Tendency. [online]. Available from: https://CRAN.R-project.org/package=clustertend.

Wyka, S., Mondo, S., Liu, M., Nalam, V. and Broders, K., 2022. A large accessory genome and high recombination rates may influence global distribution and broad host range of the fungal plant pathogen Claviceps purpurea. *PloS one* [online], 17 (2), e0263496. Available from: https://doi.org/10.1371/journal.pone.0263496.

Zhan, J., Pettway, R. E. and McDonald, B. A., 2003. 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 Genetics and Biology* [online], 38 (3), 286–297. Available from: https://doi.org/10.1016/s1087-1845(02)00538-8.

Zhang, X., Wang, R., Ning, H., Li, W., Bai, Y. and Li, Y., 2020. Evaluation and management of fungal-infected carrot seeds. *Scientific Reports* [online], 10 (1), 1–8. Available from: https://doi.org/10.1038/s41598-020-67907-5.

Zhong, S., Leng, Y., Friesen, T., Faris, J. and Szabo, L., 2009. Development and characterization of expressed sequence tag-derived microsatellite markers for the wheat stem rust fungus Puccinia graminis f. sp. tritici. *Phytopathology* [online], 99 (3), 282–289. Available from: https://doi.org/10.1094/PHYTO-99-3-0282.

Zhong, Z., Marcel, T. C., Hartmann, F. E., Ma, X., Plissonneau, C., Zala, M., Ducasse, A., Confais, J., Compain, J. and Lapalu, N., 2017. A small secreted protein in Zymoseptoria tritici is responsible for avirulence on wheat cultivars carrying the Stb6 resistance gene. *New Phytologist* [online], 214 (2), 619–631. Available from: https://doi.org/10.1111/nph.14434.

Figures

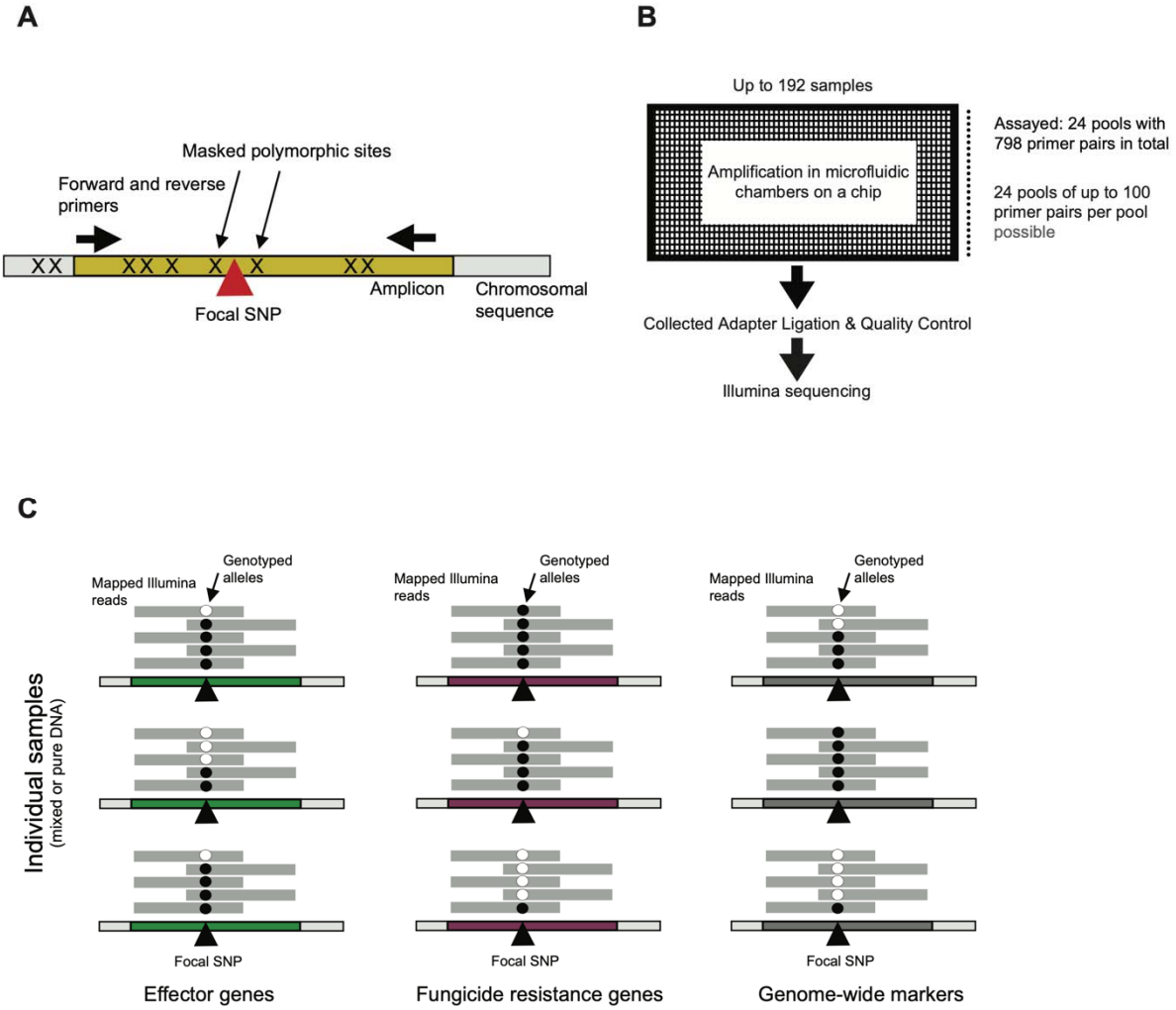
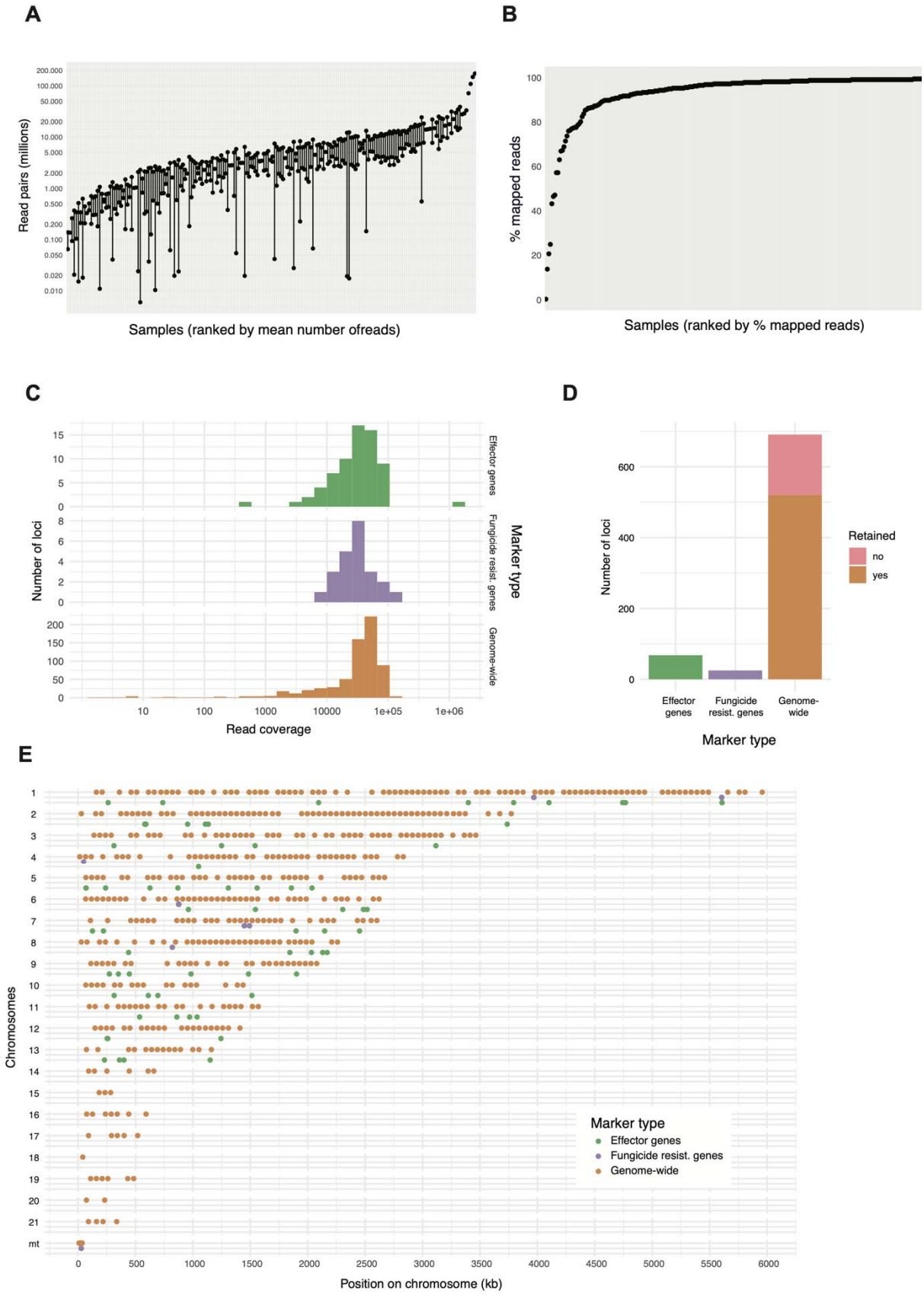


Figure 1: Schematic overview of the targeted amplicon assay design. A) Design of individual amplicons (~200 bp) with primers designed to not overlap known polymorphic sites. B) Schematic overview of the microfluidic chambers of a Fluidigm Juno chip accommodating up to 192 samples and 24 pools of primers (each up to 100 primer pairs). Following amplification in microfluidic wells, barcoded products are pooled and finalized for Illumina sequencing. C) Genotypes of individual samples (pure or mixed individuals) are assessed by analyzing mapped reads at each locus in the genome. Markers were designed for three different categories including effector genes, genes encoding targets of fungicides and genome-wide evenly spaced markers.


A
Read pairs (millions)
Samples (ranked by mean number ofreads)
B
% mapped reads
Samples (ranked by % mapped reads)
C
Number of loci
Effector genes
Fungicide resist. genes
Genome-wide
Marker type
Read coverage
D
Number of loci
Retained
no
yes
Effector genes
Fungicide resist. genes
Genome-wide
Marker type
E
Chromosomes
Marker type
Effector genes
Fungicide resist. genes
Genome-wide
Position on chromosome (kb)

776

Figure 2: Sequencing data recovered for the amplicon assay and loci assessment. A) Read pairs recovered per sample and replicate. Each sample was amplified and sequenced two times (two different microfluidic flow cells). B) Ranking of percent mapped reads to the reference genome per sample (including both replicates if available). C) Number of reads mapped per locus for the three different categories of markers. The read numbers correspond to the total obtained from four pooled samples performed in replicates. D) Summary of loci retained after read number filtering. Only genome-wide markers were removed if they failed filtering criteria. E) Overview of retained markers per category across the 21 chromosomes and mitochondrion.

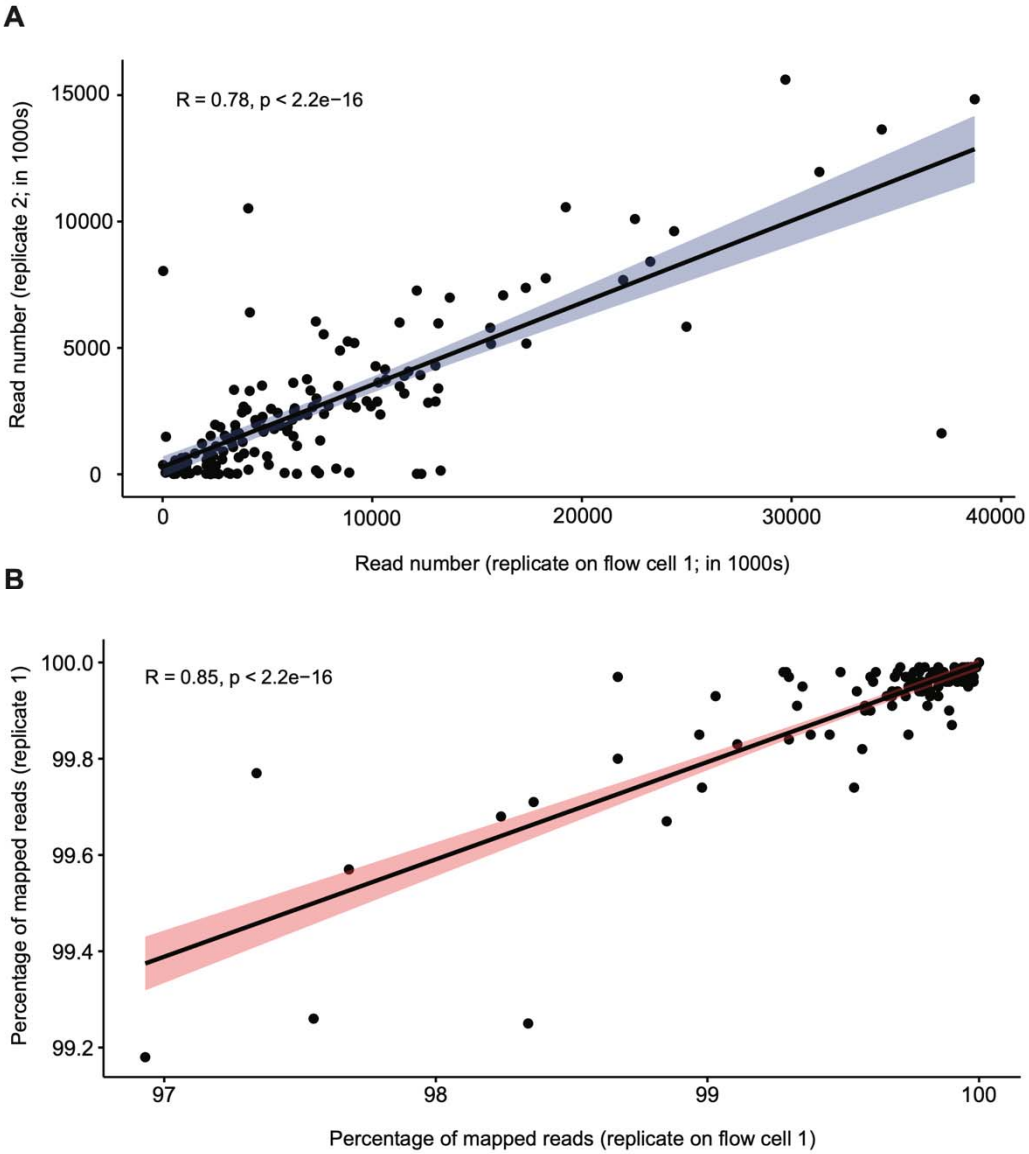


Figure 3: Consistency between replicate runs of the amplicon assay. A) Read numbers per sample and B) percentage of reads mapped to the reference genome.

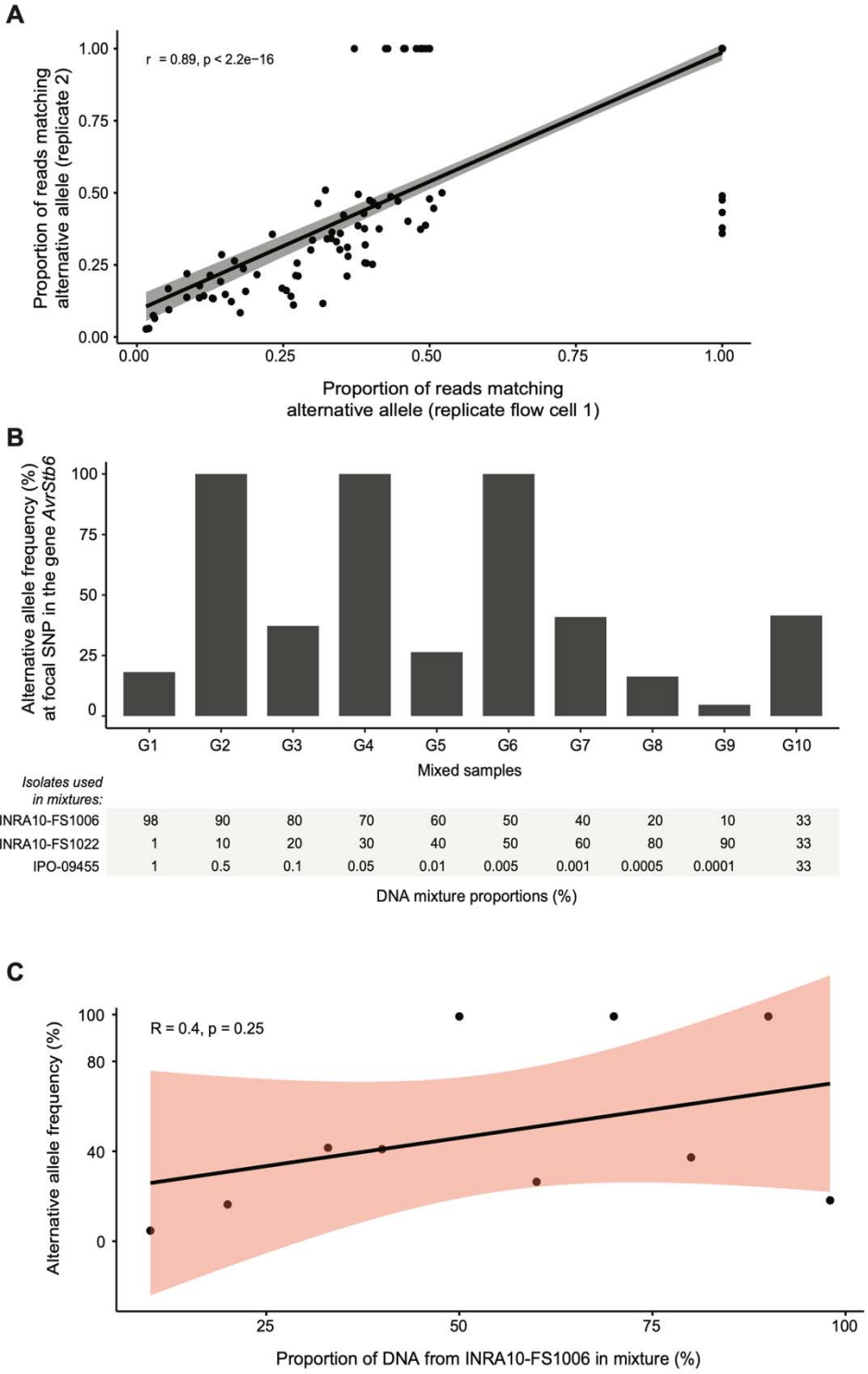


Figure 4: Evaluation of mixed sample analyses. A) Comparison of alternative allele frequencies within samples between the two replicate runs for each sample. B) Comparison of alternative allele frequencies within DNA mixture samples containing different proportions of three isolates. C) Correlation of proportion of the isolate INRA10-FS1006 compared to the alternative allele frequency.

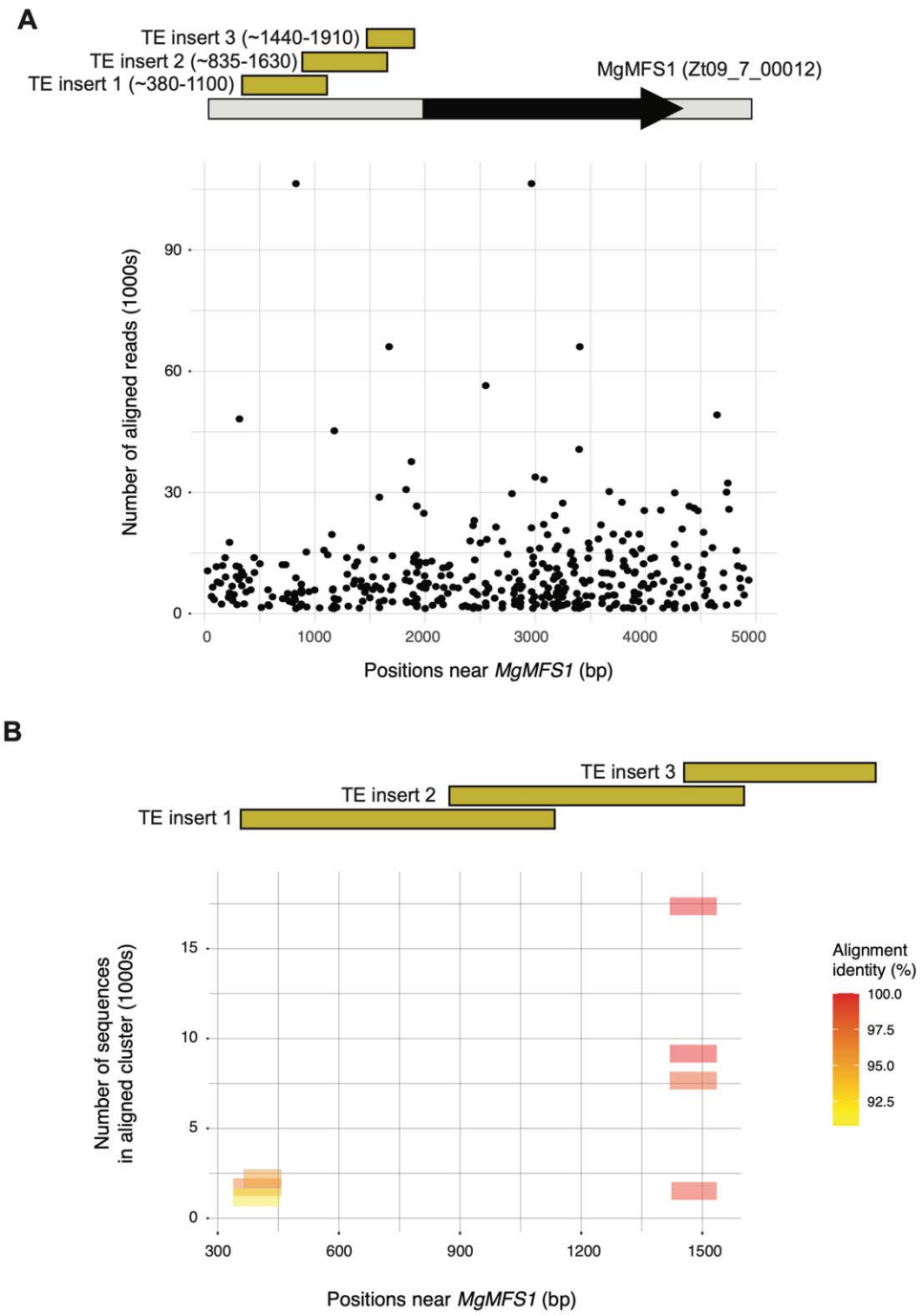


Figure 5: Analyses of amplicons designed on polymorphic transposable element insertions upstream of the multidrug transporter gene *MgMFS1*. A) Overview of the location of amplicons designed for each of three transposable element insertion site (1-3). Multiple amplicons were designed for each insertion site. The aligned reads are shown for positions near the coding sequence of *MgMFS1* for sample G9 (only positions with >10 reads mapping are shown). B) After read clustering for sample G9, consensus sequences were blasted against positions near the coding sequence of *MgMFS1*. The horizontal bars indicate the extent of a BLASTn alignment with colors indicating the percent identity of the

alignment. The vertical position indicates the number of sequences that were clustered for the aligned consensus sequence.

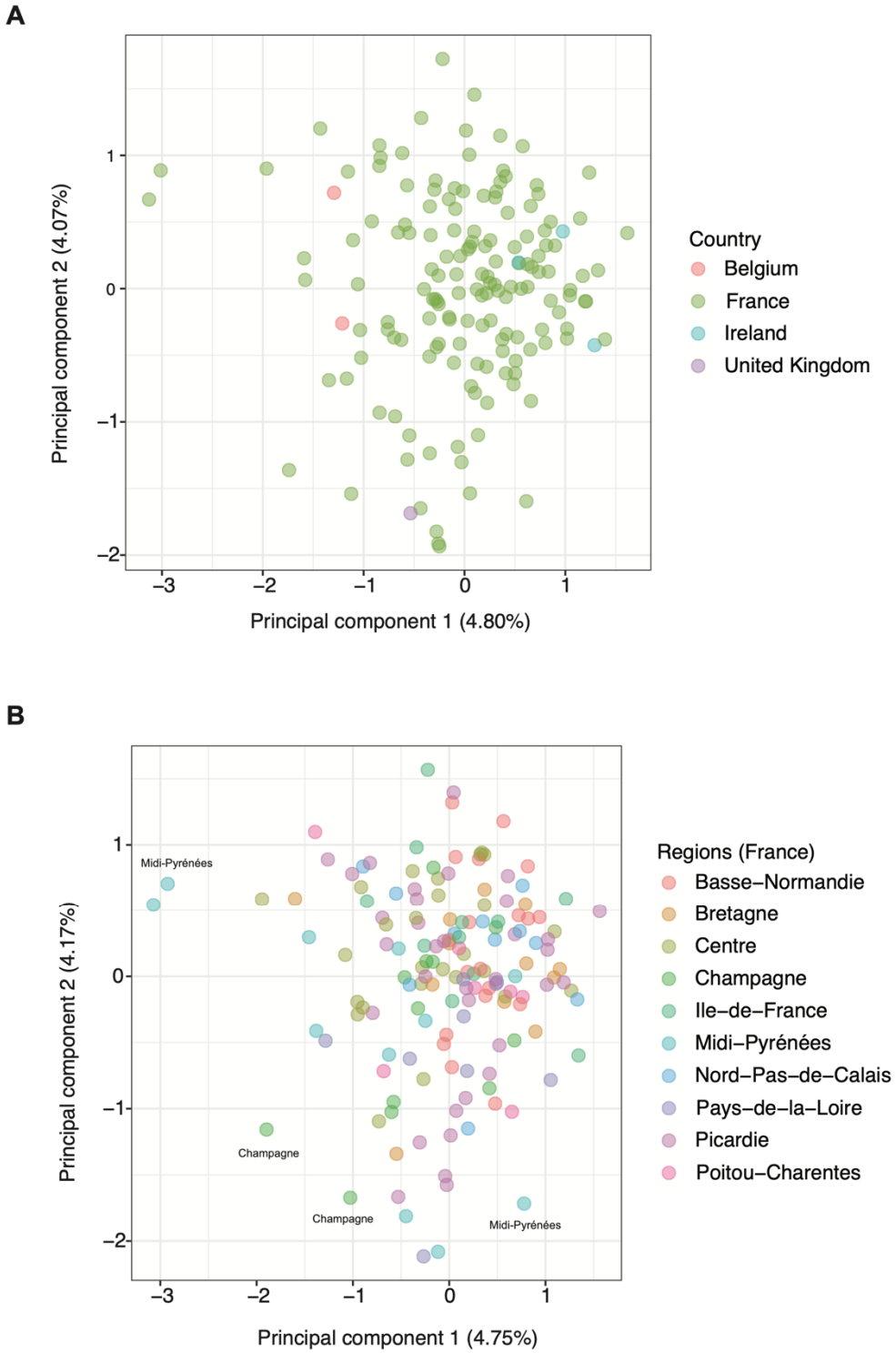
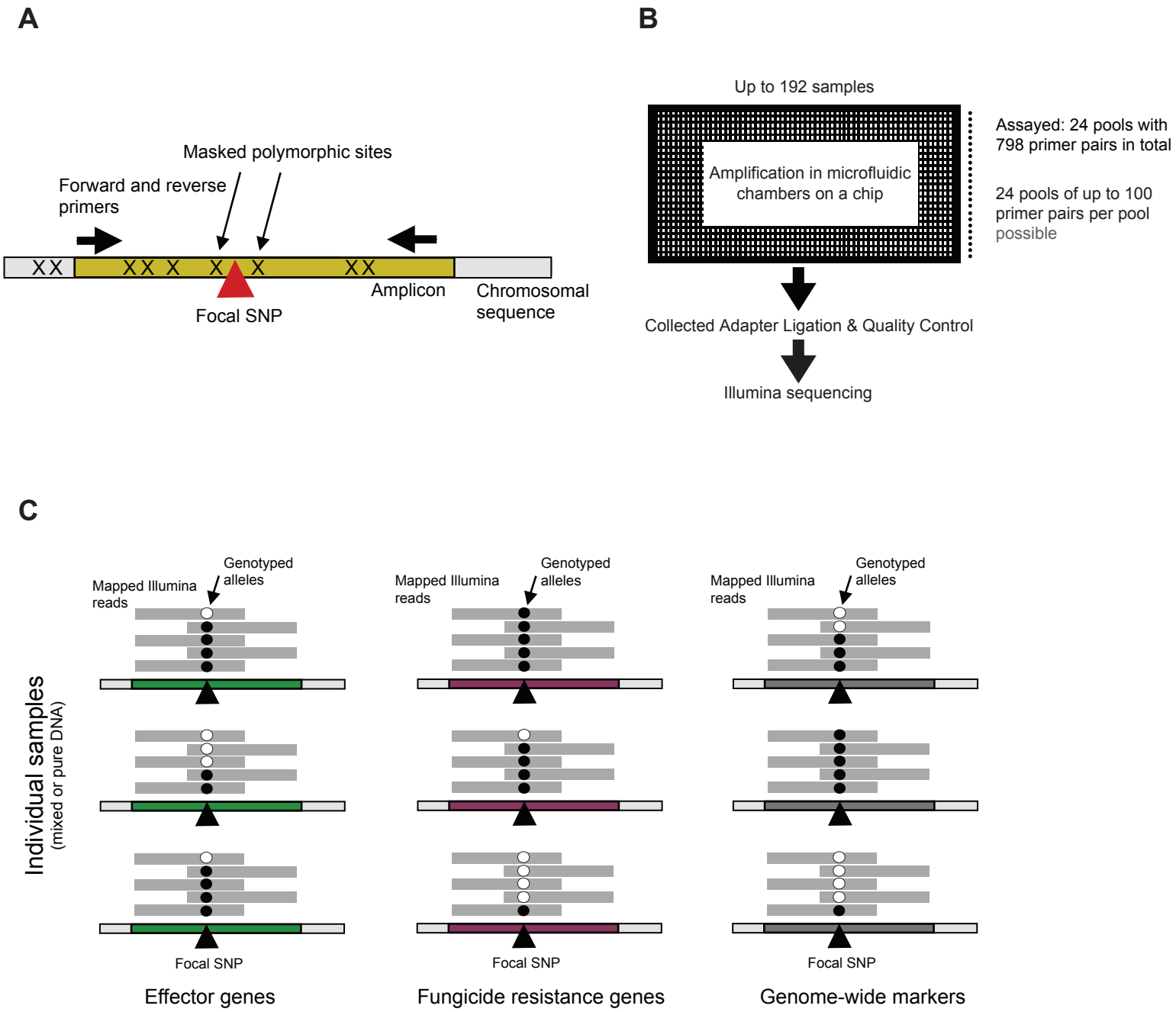
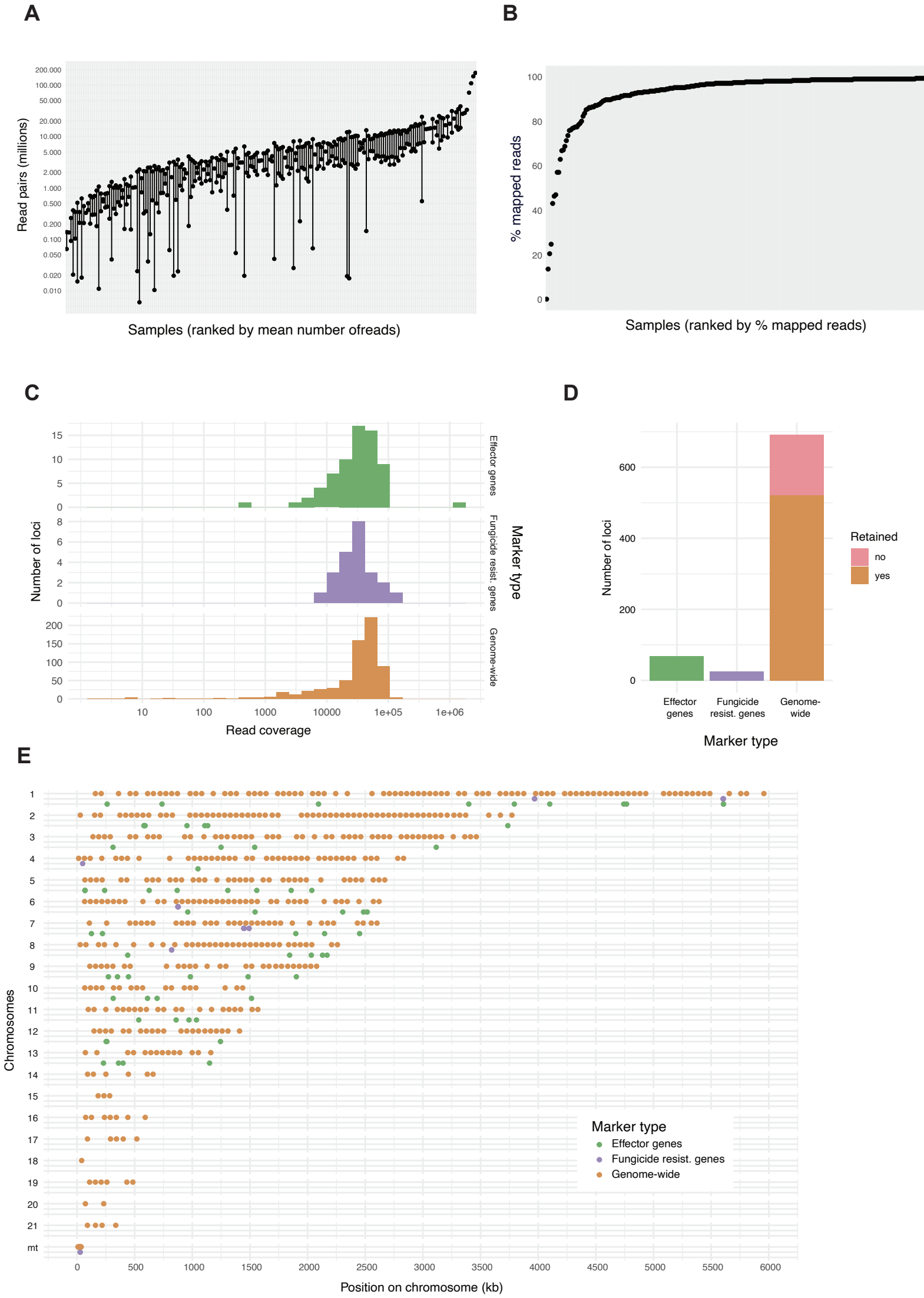
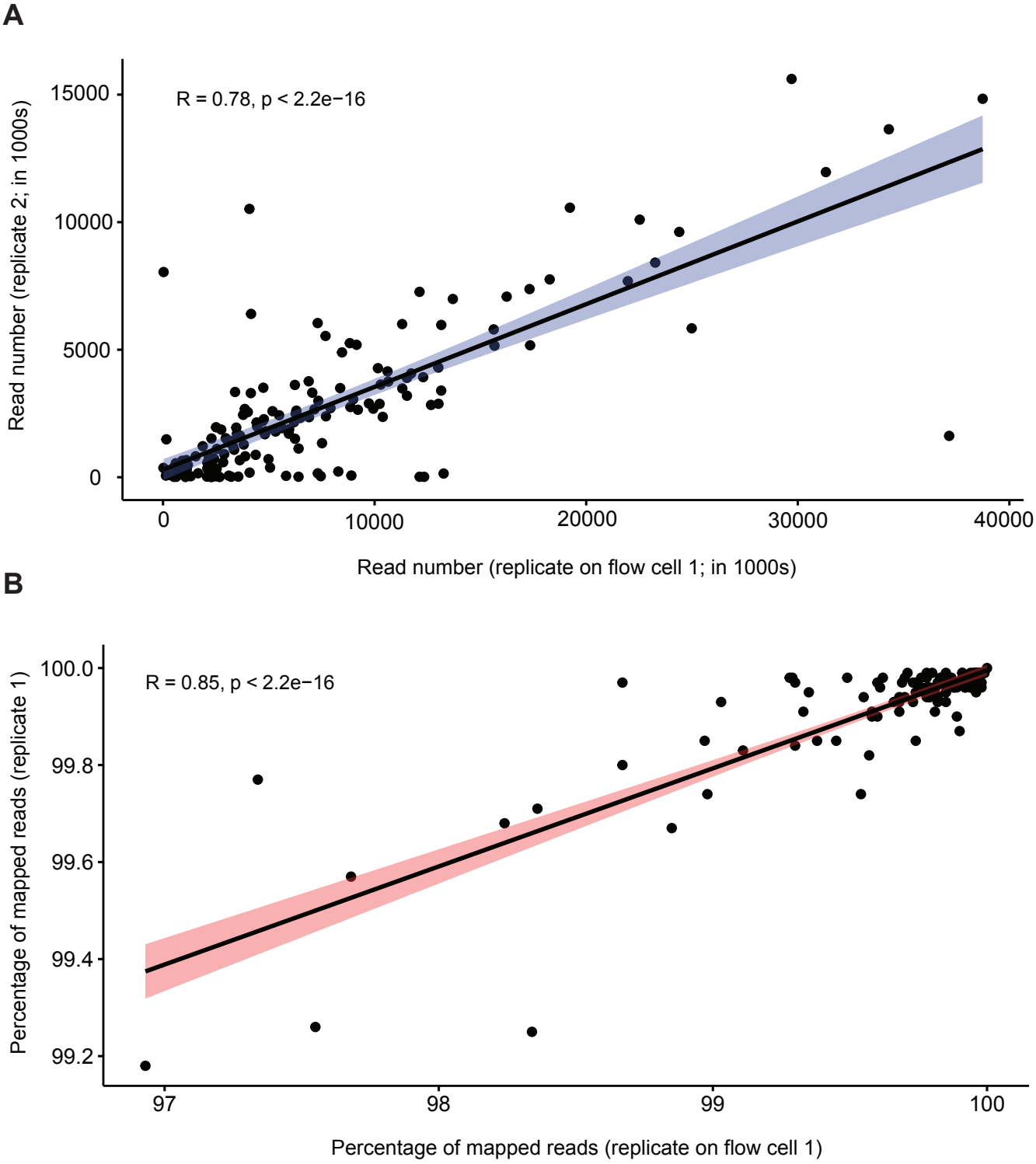


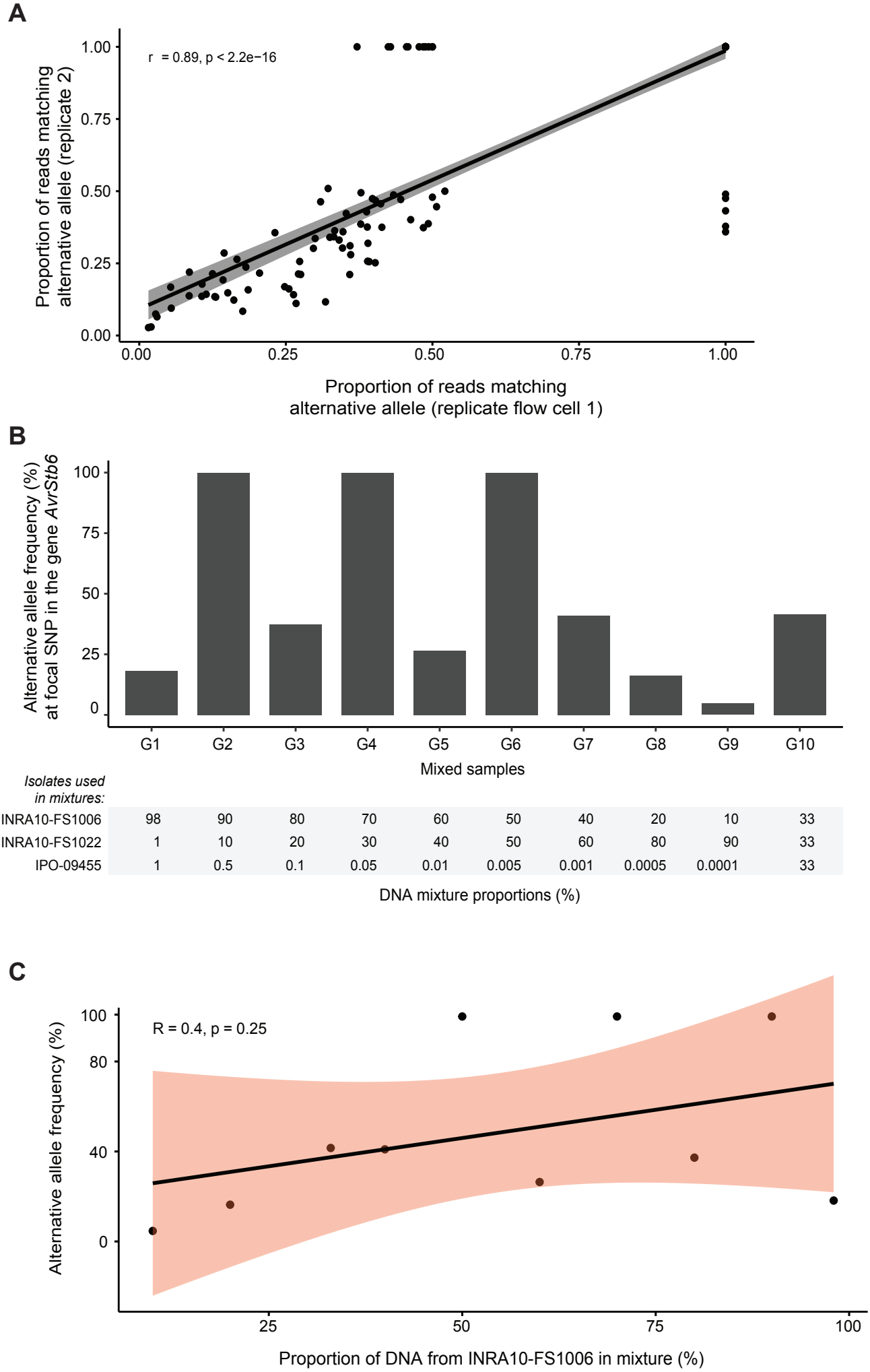
Figure 6: Population structure analyses based on genome-wide markers genotyped on leaf-extracted assemblies of *Zymoseptoria tritici* strains. A) Principal component analysis of wheat leaf

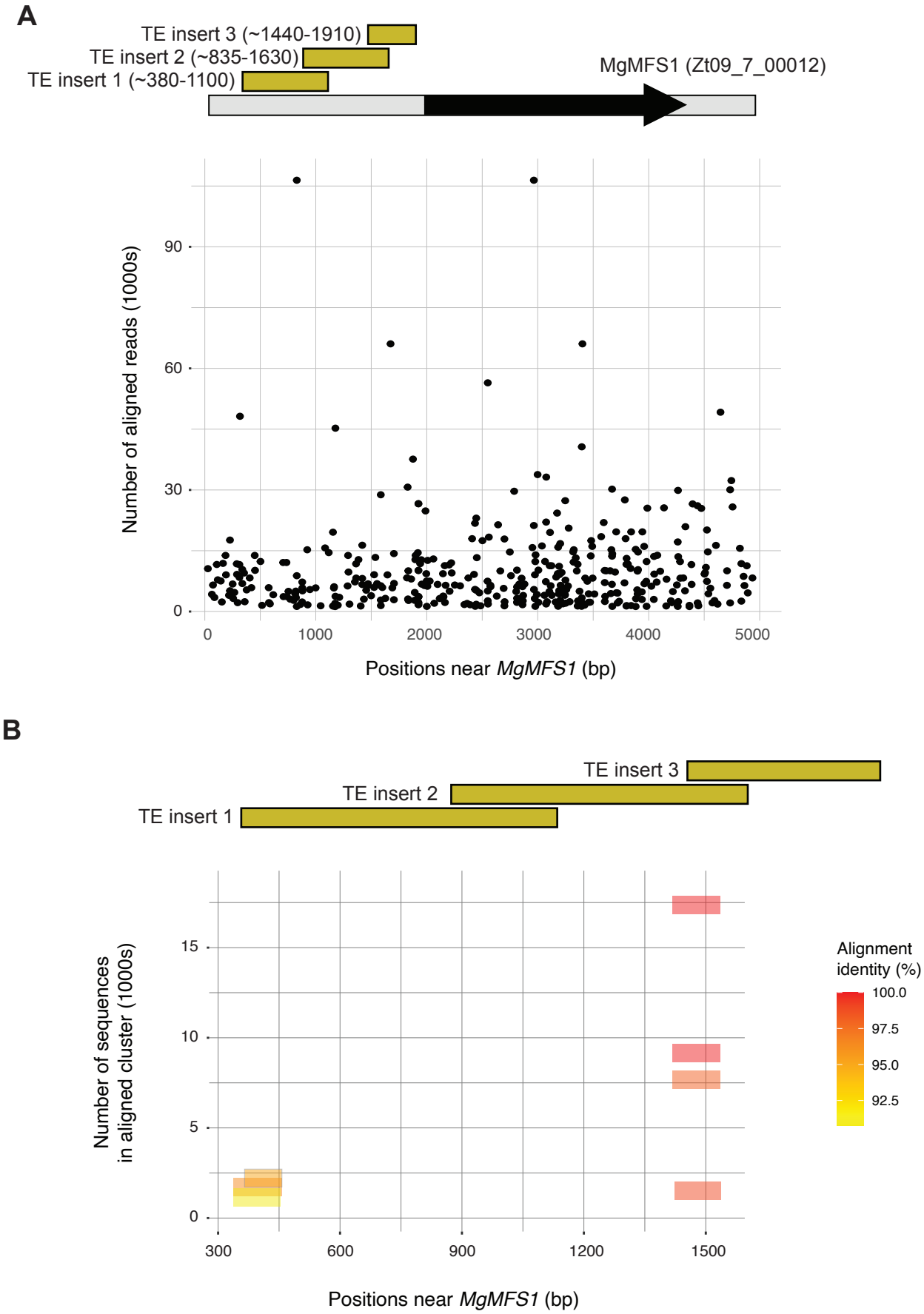
samples collected in France, Belgium, Ireland and the United Kingdom and B) the subset of wheat leaf samples collected in France colored by region.


A
Masked polymorphic sites
Forward and reverse primers
Focal SNP
Amplicon
Chromosomal sequence
B
Up to 192 samples
Amplification in microfluidic chambers on a chip
Assayed: 24 pools with 798 primer pairs in total
24 pools of up to 100 primer pairs per pool possible
Collected Adapter Ligation & Quality Control
Illumina sequencing
C
Individual samples
(mixed or pure DNA)
Mapped Illumina reads
Genotyped alleles
Focal SNP
Effector genes
Fungicide resistance genes
Genome-wide markers

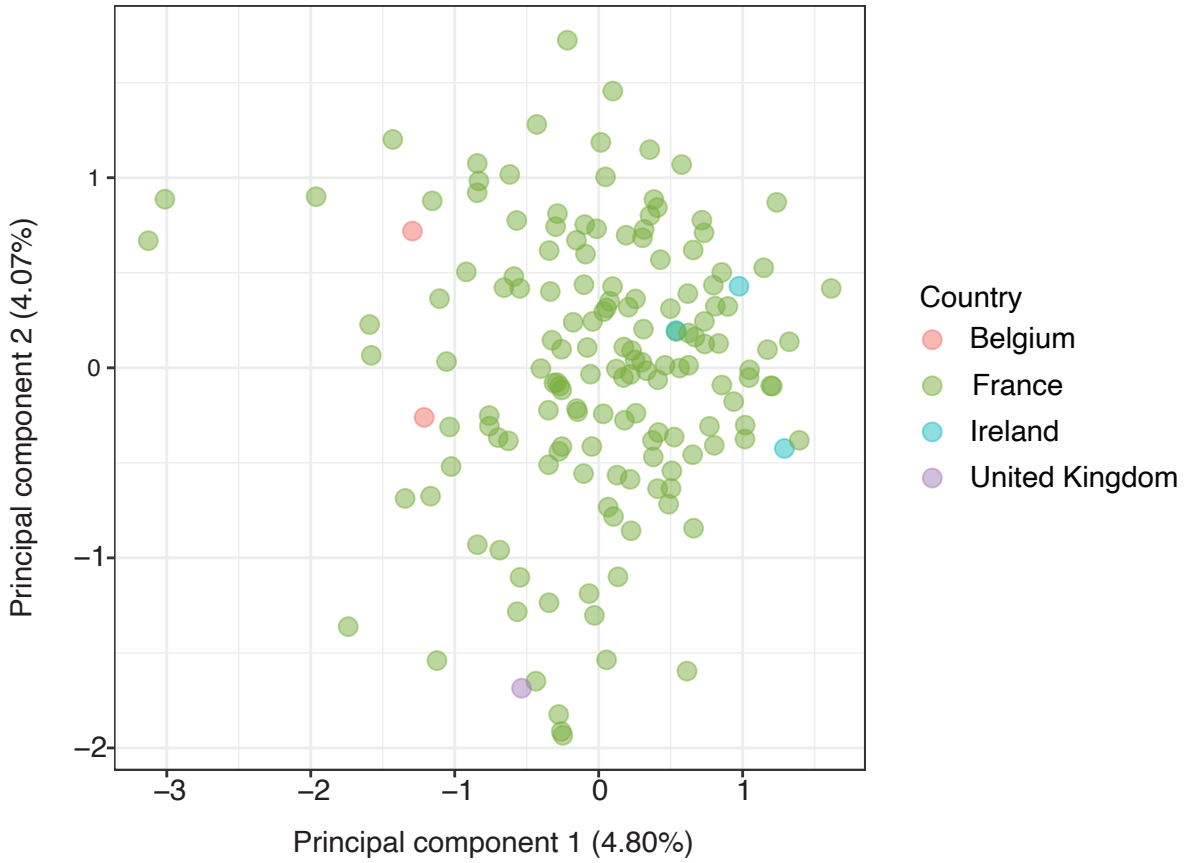

A
Read pairs (millions)
200.000
100.000
50.000
20.000
10.000
5.000
2.000
1.000
0.500
0.200
0.100
0.050
0.020
0.010
Samples (ranked by mean number ofreads)
B
% mapped reads
100
80
60
40
20
0
Samples (ranked by % mapped reads)
C
Number of loci
Effector genes
Fungicide resist. genes
Genome-wide
Marker type
Read coverage
10
100
1000
10000
1e+05
1e+06
D
Number of loci
Retained
no
yes
Effector genes
Fungicide resist. genes
Genome-wide
Marker type
E
Chromosomes
mt
Position on chromosome (kb)
Marker type
Effector genes
Fungicide resist. genes
Genome-wide


A
R = 0.78, p < 2.2e−16
Read number (replicate 2; in 1000s)
15000
10000
5000
0
0
10000
20000
30000
40000
Read number (replicate on flow cell 1; in 1000s)
B
R = 0.85, p < 2.2e−16
Percentage of mapped reads (replicate 1)
100.0
99.8
99.6
99.4
99.2
97
98
99
100
Percentage of mapped reads (replicate on flow cell 1)

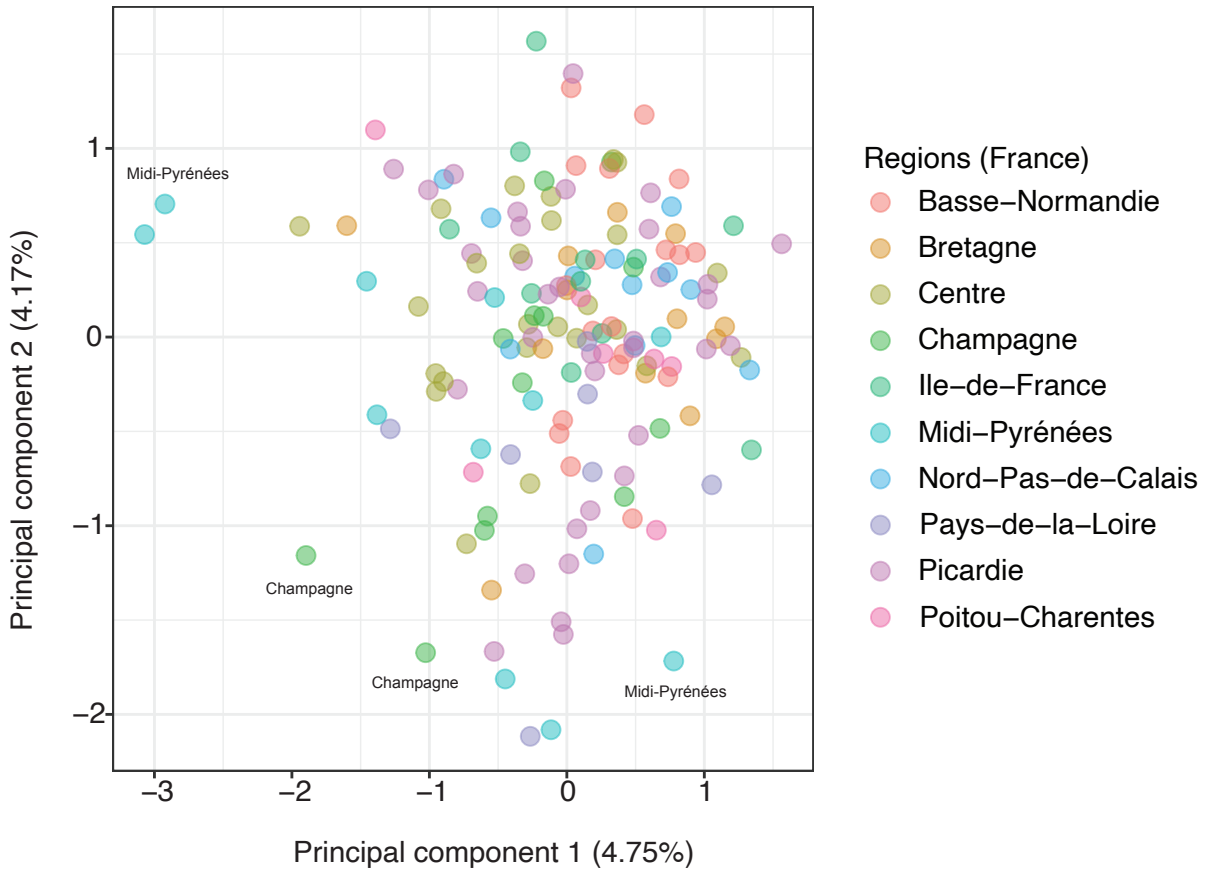

A
r = 0.89, p < 2.2e−16
Proportion of reads matching alternative allele (replicate 2)
Proportion of reads matching alternative allele (replicate flow cell 1)
B
Alternative allele frequency (%) at focal SNP in the gene AvrStb6
G1
G2
G3
G4
G5
G6
G7
G8
G9
G10
Mixed samples
Isolates used in mixtures:
INRA10-FS1006 98 90 80 70 60 50 40 20 10 33
INRA10-FS1022 1 10 20 30 40 50 60 80 90 33
IPO-09455 1 0.5 0.1 0.05 0.01 0.005 0.001 0.0005 0.0001 33
DNA mixture proportions (%)
C
R = 0.4, p = 0.25
Alternative allele frequency (%)
Proportion of DNA from INRA10-FS1006 in mixture (%)


A
TE insert 3 (~1440-1910)
TE insert 2 (~835-1630)
TE insert 1 (~380-1100)
MgMFS1 (Zt09_7_00012)
Number of aligned reads (1000s)
90
60
30
0
0
1000
2000
3000
4000
5000
Positions near MgMFS1 (bp)
B
TE insert 3
TE insert 2
TE insert 1
Number of sequences in aligned cluster (1000s)
15
10
5
0
300
600
900
1200
1500
Positions near MgMFS1 (bp)
Alignment identity (%)
100.0
97.5
95.0
92.5

A


Principal component 2 (4.07%)
Principal component 1 (4.80%)
1
0
−1
−2
−3
−2
−1
0
1
Country
Belgium
France
Ireland
United Kingdom

B


Principal component 2 (4.17%)
Principal component 1 (4.75%)
Midi-Pyrénées
Champagne
Champagne
Midi-Pyrénées
1
0
−1
−2
−3
−2
−1
0
1
Regions (France)
Basse−Normandie
Bretagne
Centre
Champagne
Ile−de−France
Midi−Pyrénées
Nord−Pas−de−Calais
Pays−de−la−Loire
Picardie
Poitou−Charentes