Population biology and invasion history of puccinia striformis F.SP. tritici at worldwide and local scale

Ali Ali Sajid

UNIVERSITE PARIS-SUD – UFR des Sciences

ÉCOLE DOCTORALE SCIENCES DU VEGETAL

Thèse

Pour obtenir le grade de

DOCTEUR EN SCIENCES DE L'UNIVERSITÉ PARIS SUD

Discipline : **Genetics, Evolution and Population Biology**

Par

Sajid ALI

Population biology and invasion history of *Puccinia striiformis* f.sp. *tritici* at worldwide and local scale

Soutenue publiquement le **10 septembre 2012**, devant le jury d'examen :

Jacqui SHYKOFF	DR, CNRS-Université Paris-Sud 11, Orsay, France – Président
James K.M. BROWN	Professor, John Innes Centre, UK – Rapporteurs
Didier THARREAU	DR, UMR BGPI, CIRAD, France – Rapporteurs
Tatiana GIRAUD	DR, CNRS-Université Paris-Sud 11, Orsay, France – Examinateur
Kumarse NAZARI	Senior Cereal Pathologist, ICARDA, Syria – Examinateur
Mogens S. HOVMØLLER	Professor, Aarhus University, Denmark – Examinateur
Claude DE VALLAVIEILLE-POPE	DR, UR BIOGER-CPP, INRA Grignon, France – Directrice de thèse

Co-directeurs

Jérôme ENJALBERT	UMR Génétique Végétale, INRA du Moulon, France
Pierre Gladieux	Department of Plant and Microbial Biology, University of California, CA 94720-3102

Thesis Title: **Population biology and invasion history of *Puccinia striiformis* f.sp. *tritici* at worldwide and local scale**

Candidate: **Sajid ALI**

PhD session: **October 2009 to September 2012**

Specialization: **Genetics, Evolution, Population Biology, Epidemiology, Resistance Breeding**

University: **Doctoral School of Plant Sciences, UNIVERSITE PARIS-SUD – UFR des Sciences – France**

Laboratory of research: **UR BIOGER-CPP, INRA Versailles-Grignon, France**

HEC, Pakistan program: **Overseas scholarship for MS/MPhil leading to PhD in selected field phase-II batch II - 2008**

PhD Supervision:

Supervisor: **Claude de Vallavieille-Pope,** INRA UR 1290 BIOGER-CPP, BP01, F-78850 Thiverval-Grignon, France

Co-supervisor: **Jerome Enjalbert,** INRA UMR 320 Génétique Végétale, Ferme du Moulon, F-91190 Gif sur Yvette, France

Co-supervisor: **Mogens S. Hovmøller,** Department of Agroecology, Aarhus University, Flakkebjerg, DK-4200 Slagelse, Denmark

Co-supervisor (invited): **Pierre Gladieux,** Department of Plant and Microbial Biology, University of California, Berkeley, CA 94720-3102

Résumé

BIOLOGIE DES POPULATIONS ET HISTOIRE DES INVASIONS DE *PUCCINIA STRIIFORMIS* F.SP. *TRITICI* A L'ECHELLE MONDIALE ET LOCALE

L'étude de la structure génétique des populations d'agents pathogènes à grandes échelles permet d'identifier des réservoirs de diversité et des routes de migration, de retracer les trajectoires évolutives et de repérer notamment des sources potentielles de nouvelles invasions. *Puccinia striiformis* f.sp. *tritici* (PST), responsable de la rouille jaune du blé, réparti mondialement, constitue un modèle fongique d'intérêt pour les études d'invasion étant donné sa capacité de migration sur de longue distance et l'apparition récurrente de nouvelles souches localement. Nous avons analysé la structure des populations de PST à l'échelle mondiale, à l'aide de marqueurs microsatellites sur un échantillon de 409 isolats issus des six continents (représentatif des collections INRA, France et AU, Danemark). Les génotypes ont été répartis en six groupes génétiques correspondant à leur origine géographique grâce à des méthodes de classification bayésienne et multivariée. Les analyses indiquent une forte hétérogénéité géographique de déséquilibre de liaison et diversité génotypique, avec des signatures de recombinaison dans les régions de l'Himalaya (Népal et Pakistan) et à proximité en Chine. La structure reste clonale pour les populations des autres régions. L'assignation des isolats aux différents groupes génétiques a permis de déterminer l'origine des invasions (récentes ou anciennes). Ainsi, les souches agressives adaptées à de hautes températures, répandues de par le monde depuis 2000, sont originaires de Mer rouge-Moyen Orient ; les isolats d'Amérique du Nord et du Sud et d'Australie proviennent d'Europe du Nord-Ouest. Par ailleurs, les isolats d'Afrique du Sud appartiennent au groupe génétique de la zone méditerranéenne. La subdivision marquée entre les différentes zones géographiques indique qu'elles ne sont pas fortement marquées par les migrations récentes. De plus, les voies de migration identifiées attestent de l'importance des activités humaines dans la dispersion de PST à longue distance. La biologie des populations des zones les plus diverses (Chine et Pakistan) a été finement étudiée à l'aide d'échantillonnages réalisés deux années consécutives. Une population échantillonnée en 2004 et 2005 dans la vallée de Tianshui, (province de Gansu, Chine), s'est révélée très diverse, fortement recombinante et non structurée spatialement et temporellement. L'observation de clones identiques entre les deux échantillons temporels a permis de développer un estimateur du taux de sexualité, i.e. du rôle relatif de la reproduction sexuée par rapport à celui de la reproduction asexuée dans le maintien de la population. Ce taux de reproduction sexuée est estimé à 74 %, alors que la taille efficace de la population est de 1735, ce qui donne les premières indications du rôle du cycle sexué pour cette espèce. L'échantillonnage réalisé au Nord du Pakistan a permis de décrire quatre groupes génétiques ayant tous une grande diversité génotypique et une structure recombinante. Le très faible taux de ré-échantillonnage de génotypes identiques au cours de deux années suggère le rôle prédominant de la reproduction sexuée dans le maintien temporel des populations locales. La forte diversité génétique et génotypique, la signature de recombinaison et la capacité à la reproduction sexuée de PST dans la région himalayenne suggèrent que cette zone est le centre d'origine potentielle de PST. Les analyses d'approximations bayésiennes confirment la thèse d'une dispersion à partir de l'Himalaya vers les autres régions du monde. Le Pakistan et Chine seront les populations les plus ancestrales des populations testées. La variabilité pour la capacité à produire des téleutosores, spores indispensables à l'initiation de la phase sexuée, a été analysée (56 isolats mondiaux), et s'avère liée à la variabilité génotypique et au taux de recombinaison. Ce résultat conforte la thèse de l'apparition de la sexualité dans la zone himalayenne et à proximité de cette zone et de la perte de sexualité lors de migrations dans les zones où l'hôte alternant est absent et où le cycle épidémique est essentiellement asexué. La description de l'origine, des voies de migration de populations mondiales de PST ainsi que de son centre de diversité contribue à la compréhension du potentiel évolutif de PST et à la construction de stratégies de gestion de lutte contre l'agent pathogène.

Abstract

POPULATION BIOLOGY AND INVASION HISTORY OF *PUCCINIA STRIIFORMIS* F.SP. *TRITICI* AT WORLDWIDE AND LOCAL SCALE

Analyses of the large-scale population structure of pathogens enable the identification of migration patterns, diversity reservoirs or longevity of populations, the understanding of current evolutionary trajectories and the anticipation of future ones. Once identified, a detailed analysis of populations in centre of diversity, with emphasis on demographic and reproductive parameters should enable to infer the adaptive capacity of the pathogen and identify potential sources for new invasions. *Puccinia striiformis* f.sp. *tritici* (PST) is the causal agent of wheat yellow/stripe rust, and despite a worldwide distribution, this fungus remains a model species for invasion studies, due to its long-distance migration capacity and recurrent local emergence of new strains. Little is known about the ancestral relationship of the worldwide PST population, while its center of origin is still unknown. We used multilocus microsatellite genotyping to infer the worldwide population structure of PST and the origin of new invasions, analysing a set of isolates representative of sampling performed over six continents. Bayesian and multivariate clustering methods partitioned the isolates into six distinct genetic groups, corresponding to distinct geographic areas. The assignment analysis confirmed the Middle East-Red Sea Area as the most likely source of newly spreading, high-temperature-adapted strains; Europe as the source of South American, North American and Australian populations; and Mediterranean-Central Asian populations as the origin of South African populations. The existence of strong population subdivision at worldwide level shows that major genetic groups are not markedly affected by recent dispersal events. However, the sources for recent invasions and the migration routes identified emphasize the importance of human activities on the recent long-distance spread of the disease. The analyses of linkage disequilibrium and genotypic diversity indicated a strong regional heterogeneity in levels of recombination, with clear signatures of recombination in the Himalayan (Nepal and Pakistan) and near-Himalayan (China) regions and a predominant clonal population structure in other regions. To explain the variability in diversity and recombination of worldwide PST populations, we assessed their sex ability in terms of telial production, the sex-specific structures that are obligatory for PST sexual cycle, in a set of 56 isolates representative of these worldwide geographical origins. We confirmed that the variability in genotypic diversity/ recombination was linked with the sex ability, pinpointing the Himalayan region as the possible center of origin of PST, from where it then spread worldwide. The reduced sex ability in clonal populations certainly reflects a loss of sexual function, associated to migration in areas where sexual alternate host is lacking, or not necessary for the completion of epidemic cycle. Approximate Bayesian computation analyses confirmed an out of Himalaya spread of PST, with Pakistan and China being the most ancestral population. An in-depth study was made in the Himalayan region (Pakistan) and Near Himalayan region (Gansu, China) to better describe the population maintenance in two areas of high diversity. Analyses of Pakistani population at regional level revealed the existence of a strong population subdivision, a high genotypic diversity and the existence of recombination signature at each location reflecting the role of sexual recombination in the temporal maintenance at local level. A time spaced sampling of PST in the valley of Tianshui (China) inspired the development of a new estimator, allowing to quantify the relative contribution of sexual reproduction and effective population size on the basis of clonal resampling within and between years. A sexual reproduction rate of 74% (95% confidence interval [CI]: 38-95%) and effective population size of 1735 (95% CI: 675-2800) was quantified in Chinese PST population. The description of the origin and migration routes of PST populations worldwide and at its centre of diversity contributes to our understanding of PST evolutionary potential, and is helpful to build disease management strategies.

SYNTHESE EN FRANÇAIS

Biologie des populations et histoire des invasions de *Puccinia striiformis* f.sp. *tritici* à l'échelle mondiale et locale

L'étude de la structure génétique des populations d'agents pathogènes permet l'identification des réservoirs de diversité et des routes de migration, et ainsi de décrire les trajectoires évolutives passées et prévoir celles à venir. Une fois l'origine des sources d'invasions passées identifiées, une analyse approfondie des paramètres démographiques et de la reproduction permettra d'inférer les capacités d'adaptation de l'agent pathogène. *Puccinia striiformis* f.sp. *tritici* (PST), responsable de la rouille jaune du blé, reste un modèle fongique important pour étudier les invasions, du fait de sa capacité à migrer sur de longues distances. On connaît mal l'histoire évolutive des populations de PST à l'échelle mondiale en particulière centre d'origine et de diversité, ainsi que l'origine des émergences récentes.

Dans une première analyse, nous avons donc mené une analyse basée sur des microsatellites pour déduire la structure de la population de PST à l'échelle mondiale et l'origine de nouvelles invasions à l'aide d'un échantillon représentatif de la distribution du l'agent pathogène sur les six continents. Les méthodes de classification bayésienne et multivariée partitionnent l'ensemble des génotypes observés sur un échantillon de 409 isolats en six groupes génétiques distincts associés à leur origine géographique probable. Les analyses de déséquilibre de liaison et de la diversité génotypique indiquent une forte hétérogénéité régionale dans les niveaux de recombinaison, avec des signatures claires de recombinaison dans les régions de l'Himalaya (Népal et Pakistan) et à proximité de l'Himalaya (Chine) et une structure clonale de la population dans les autres régions.

L'analyse a confirmé le Moyen-Orient et la région de la Mer Rouge comme étant l'origine géographique des souches agressives adaptées à haute température, qui ont envahi le monde depuis 2000. Nous confirmons que l'Europe du Nord-Ouest est à l'origine des souches isolées en Amérique du Sud, Amérique du Nord et Australie. La zone méditerranéenne est à l'origine des populations d'Afrique du Sud. La subdivision forte des populations observée au niveau mondial montre que les populations du centre de diversité ne sont pas à l'origine les événements de dispersion récents. Les sources et les voies de migration identifiées pour les invasions récentes montrent l'importance des activités humaines dans la dispersion de PST à longue distance.

Pour expliquer la variabilité dans la diversité et la recombinaison des populations de PST à travers le monde, nous avons dans une seconde étude évalué les variations d'aptitude au cycle sexué chez PST, au travers d'une capacité à produire des téliosores sur les plantules de blé. Cette structure est nécessaire à la production de basidiospores qui ont la capacité d'infecter l'hôte alternant, *Berberis spp*. En analysant un échantillon de 56 isolats représentatifs de la population mondiale, nous avons confirmé que la variabilité observée de la diversité génotypique et du taux de recombinaison est liée à la capacité à la production de téliosores. Ce résultat est en faveur de l'hypothèse d'un rôle important de la sexualité dans la zone himalayenne et à proximité de cette zone et d'une perte de sexualité lors de migration dans les zones où l'hôte alternant est absent, et où la phase sexuée du cycle n'est pas nécessaire pour l'achèvement du cycle épidémique de PST.

Deux études complémentaires ont approfondi l'analyse de la région himalayenne du Pakistan et de la région de Gansu en Chine pour décrire la structure des populations à l'aide de paramètres démographiques et reproductifs. La population pakistanaise est très diverse, elle est subdivisée en quatre

groupes génétiques ayant tous une grande diversité génotypique et une signature de recombinaison génétique. La comparaison des populations au cours de deux années montre l'absence de différentiation au niveau F_{ST}. Le très faible taux de ré-échantillonnage des clones sur les deux années consécutives suggère le rôle prédominant de la reproduction sexuée dans le maintien temporel des populations locales.

La population chinoise de la province de Gansu, zone de survie estivale et hivernale de l'agent pathogène, montre une grande diversité et une signature de recombinaison mais ne montre pas de structuration en plusieurs groupes géographique. Le ré-échantillonnage d'un faible nombre de clones entre deux années (2004 et 2005), suggère le rôle prédominant de la reproduction sexuée par rapport à la reproduction asexuée dans le maintien de la population dans cette zone. Un estimateur a été développé pour quantifier la contribution relative de la reproduction sexuée et la taille efficace de la population sur la base du ré-échantillonnage clonal au sein et entre les populations temporelles. Un taux de reproduction sexuée de 74 % (intervalle de confiance à 95 % [IC]: 38-95 %) et la taille efficace de la population de 1735 (95% CI: 675-2800) ont été estimés dans cette population chinoise de PST. La méthode est disponible en ligne (http://ciam.inra.fr/biosp/nease) et peut être utilisée pour quantifier les paramètres de reproduction et démographiques chez tout type d'organismes. L'information sur les paramètres démographiques et reproductifs contribuera à une meilleure compréhension de la biologie des populations de ravageurs et de microorganismes.

A l'échelle mondiale, les résultats de notre étude sur la diversité génétique et génotypique, la signature de recombinaison et la capacité à la reproduction sexuée de PST dans la région himalayenne suggèrent cette zone comme le centre d'origine possible de PST, d'où il se serait ensuite propagé dans le monde entier. Les analyses bayésiennes approximatives (ABC) ont confirmé une dispersion à partir de l'Himalaya vers

les autres régions du monde ainsi que le Pakistan et la Chine étant les populations les plus ancestrales. Cette étude a permis d'identifier ou de confirmer l'origine des invasions récentes de PST. La description des centres de diversité et des voies de migration des populations dans le monde entier de PST contribuera à la compréhension du potentiel évolutif de PST, et sera utile pour construire des stratégies de gestion de variétés résistantes pour lutter contre cette maladie. Les résultats obtenus dans cette thèse contribueront à comprendre la structure des populations d'agents pathogènes au niveau mondial et local dont l'intérêt dépasse le modèle biologique étudié.

ACKNOWLEDGEMENTS

Thanks to Allah for His beneficence that enabled me to accomplish this task and His prophet Muhammad (SAW), being the main source of inspiration in my life, not well understood in the Muslim society of today.

I am thankful to France to allow me utilize her resources to complete my PhD in France. I am also thankful to all French people with whom I interacted during my stay and by-and-large I found them very friendly, kind and co-operative.

Thanks to you Claude for supervising my PhD work and benefiting me from the expertise and materials available in her lab. It was a kindness to provide me possibilities to attend various scientific conferences, meetings and workshops. I am indebted to you for your professional, social and moral support throughout my PhD. I am grateful to mention that you showed confidence in me and let me modify my PhD work time to time to bring it to the present shape. This enabled me to work independently and, I think, a part of my efficiency, if any. Thanks also for your social support regarding administrative tasks.

Jerome, special thanks will definitely go to you, being the co-supervisor and always willing to provide me directions regarding my PhD. Your professional, moral and social support always helped me to work efficiently. I really benefited from the discussions we had at different occasions. Your constructive comments and critics on experiments, data analyses, interpretation and write-up were always helpful.

I am thankful to Mogens also, whose guidance and help assisted to better orient various aspects of my work. His kindness was also to allow me exploit the materials available at his lab, without which I would not have the worldwide representative isolates. Thanks to all the Aarhus University group members, especially Annemarie. Thank you Sarah for inviting us to your chateau. Thank you Stephanie for your constructive comments on my work and for your social support.

Special thanks go to Pierre, who substantially contributed to my PhD as co-supervisor, although he was not declared officially. I learnt a lot from you, regarding population genetics analyses and interpretations as well as about technical writing. It was through you that I got acquaintance to latest population genetics tools and was able to interpret the output.

Thanks to Jacqui, for her willingness to help in any aspect regarding technical or administrative tasks. Her moral, social and scientific support also contributed to my work.

I am also thankful to Samuel Soubeyrand for his contribution to the work on estimation of effective population size and sex rate.

Thanks to Marc for his technical help during epidemiological work and for providing spores for different part of my work. Without his nice management of the INRA PST spore collection the present work would not have been possible.

A very special thanks to Leurent who was always with me for spore multiplication, pathotyping and teliospore production tests. A part from his professional support, he were kind enough with me to give me moral and social support, especially regarding my residence card. Thanks to Nathalie for her help in spore multiplication and teliospore production test.

Angelique, thanks to you for giving me the training for molecular genotyping and for the follow up. You were always kind to help me at any point during my practical work and even sometime postponing your own tasks to help me.

I am grateful to all the staff of BIOGER, especially Marc-Henri, Anne-Sophie, Dorothée, Gustavo, Aurélie, Thierry, Sian, Virginie, Catherine, Henriette, Ivan, Fréderic, Nicolas, Sandrine, Lilian, Jean-Marc, Anaïs, Martin, Anne-Catherine, Philippe and Christian for their facilitation during my stay at the group.

Thanks to Dr. H. Rahman, Dr. H. Ahmed, Shahab and Fiaz for their assistance which enabled to achieve the work on Pakistani population. Special thanks to Dr. Rahman and Shahab, for their sampling in July, when they were trapped in the flood of 2010, but still they continued to accomplish the sampling from Skardu.

Thanks to all of the members of Reid-Champignon, specially Babeth, Didier, Tatiana, Pascal, Cyril, Fabien and Benoit, from whom I learnt a lot during different conferences and meetings. Thanks to them for giving me a friendly environment, and due to which I got adapted to French scientific community with ease.

Apart from this, I am thankful to all of my Pakistani and French friends and teachers, especially Shahab, Dr. Sayyar Sb, Jabbar, Awais, Ali, Dr. Manzoor Sb., Suhaib, Sana, Zaigham, Christophe and Pierre.

Thanks to all French scientists from whom I got guidance either professionally or administratively. Particularly, Serge Kreteir, Manuel Plantegenest, Michel Tharaud and Claire Neema coopération.

Special thanks to Dr. M. Arif, with whom I started my first field research. Thanks to Dr. Farhatulla and Dr. Zahoor A. Swati for their kind facilitation of my post-PhD dossier in Pakistan. Thanks to Dr. Jawad A. Shah, with whom I started my career on wheat diseases and it was in his lab that I got the information about the Claude's and Jerome's work.

I am thankful to my nation for providing funds for my stay in France through the Government of Pakistan and HEC, Pakistan. I would like to thank all members of HEC, particularly Mr. Rana Shafeeq and Rizwan Ahmed. Special thanks to SFERE for facilitating our stay in France, particularly France Lamiscarre and Jean-François Grongnet.

Thank you Marwen and Youcef, you really were very helpful to me. Your advices regarding social and sometimes professional helped me a lot.

To keep me sane in different tough moments, it was the support of my family members, who enabled me to stay here in France and accomplish this task through their moral support, love and prayers. Thank you abujee, moorjani, bhaijan, Khurshid, Asif, Saima and Fatima jano for your encouragements and prayers as well as taking the pain that I was away from you in special moment of life. Thank you Saiqa, for giving me support through your prayers and love and then taking the adventure to stay with me at the most difficult part of my PhD, the writing stage. After getting married with you, it increased my sense of responsibility and the courage to face difficult moments. Special love to my sweet Humna, and thanks to her that she increased the joy of my life. Thanks to my family-in-law for their prayers and moral support, especially Khalid Uncle.

The last thanks should definitely go to my computer DELL. I have cheated with you a bit, when I said that I am buying you to enjoy songs and movies with you. You were pretty happy about that, but then all a sudden you had to face complex, boring and tasteless (at least for common man) softwares of population genetics. And the result is that today, you are even not able to run a movie longer than one hour, but still thanks that I used you till the end of my PhD.

Sajid

To

Moorjani & Abujee

My brothers & sisters

My wife & Humna

CONTENTS

GENERAL INTRODUCTION

General introduction

In natural ecosystems, plant pathogens have co-evolved with their host plants in a general context of high genetic diversity (for both pathogens and hosts) and under diverse environments. A complex pattern of ecological interactions exists in these natural ecosystems, with varying impact on population structure of both pathogen and host (Gilbert, 2002; Parker and Gilbert, 2004; von Broembsen, 1989). The complexities of these interactions were reduced at the advent of plant domestication and agricultural development, giving rise to new ecosystems: the agro-ecosystems (Salamini *et al.*, 2002; Stukenbrock and McDonald, 2008). These agro-ecosystems are characterized by a reduced inter- and intra- specific diversity of plant communities along with a common cropping system over a large area and substantial human intervention (Altieri, 1999; Stukenbrock and McDonald, 2008). This reduced diversity in agro-ecosystem compared to the natural ecosystems turned in the favor of pathogen present in the surrounding natural ecosystems and capable of attacking the domesticated crops, resulting in outbreaks of crop disease epidemics (Wolfe, 1985). Furthermore, the migration of man from one region to the other resulted not only in the exchange of seeds, plant materials and crop produce at continental and worldwide scale but also provided new encounters of pathogens and plants (Dean *et al.*, 2012; Parker and Gilbert, 2004; Wellings and McIntosh, 1990). When one of these newly introduced pathogens becomes successfully established in a given area, a new disease epidemic is caused. A number of such disease emergences have been reported causing enormous impact on human society, cropping system and the overall agro- and natural ecosystem (Anagnostakis, 1987; Anagnostakis, 2001; Smart and Fry, 2001). Today, the increased homogeneity in terms of host and cropping techniques in modern agriculture, enhanced human mobility with expanding global trade and the climatic changes render the modern agro-ecosystems prone to biological invasions and disease emergences (Anderson *et al.*, 2004; Harvell *et al.*, 2002).

Disease emergence could be defined as the increased incidence of a disease in recent past due to the incursion of a new pathogen or its new strain/pathotype, the change in host range of an existing pathogen and/or the acquisition of increased fitness by the pathogen. Immense losses to economy, environment and overall human welfare have been reported due to these diseases, often called emerging infectious diseases (EID). Many crop pathogens originate from natural pathogens of their wild ancestors, via a host tracking of pathogens, so that co-evolution is maintained during the domestication of crop plants (Stukenbrock and McDonald, 2008). Some others, which evolved on non-crop plants, became important crop pathogens through a host shift when crop plants were introduced into the region out of their centre of origin (Giraud *et al.*, 2010). In a third case the incursion of a pathogen (or its new strain/pathotype) has occurred from outside and become established in the region, termed as biological invasion (Dilmaghani *et al.*, 2012; Parker and Gilbert, 2004; Pimentel *et al.*, 2005). A number of recently reported EIDs were considered as the result of pathogen invasions through the incursion of either a new disease or a new pathogen strain with increased fitness from outside (Desprez-Loustau *et al.*, 2007; Dutech *et al.*, 2012; Hovmøller *et al.*, 2008; Singh *et al.*, 2006). Thus worldwide distribution of crops is increasing the probabilities for new encounters and new invasions; even the pathogens endemic in certain parts of the world could invade to other regions. In the case of worldwide distributed pathogens and not considered the context of invasions (Palm, 2001), some regions in the world could serve as a source for individuals with higher fitness (e.g., with a new virulence (Singh *et al.*, 2006), increased aggressiveness (Hovmøller *et al.*, 2008) or fungicide resistance) to the rest of the world, rendering even these widely spread pathogens important in the context of disease re-emergence.

To bring such invasions and emergence under better human control, previously reported invasions should be carefully analyzed and the potential sources and targets of new invasions should be identified (Campbell, 2001; Desprez-Loustau *et al.*, 2007; Perrings *et al.*, 2002). Identification of the sources and migration patterns of the previous invasions would help to trace past invasions routes, and

to forecast future ones. Although stochasticity is a main component of invasions in several aspects, ranging from dispersal to host encounter and favorable environment (Wingen *et al.*, 2012), identification of the potential sources of new strains and the corresponding host targets would contribute to the overall forecasting of new invasions. Both of these will require the knowledge of worldwide pathogen population structure and diversity centre(s).

The identification of centre(s) of diversity and grographical zones with lesser diversity should be accompanied with the knowledge of the regional adaptive potential of the pathogen and host population structure, as centre(s) of diversity would not serve as the sole source of new invasions in all cases. The selection pressure could select new advantageous mutations with infinite individuals during epidemic cycle even in less diverse populations and thus could serve as the source of invasions. This necessitates the comprehension of various selection pressures and the adaptive potential of the pathogen at regional level, which can be achieved through an in depth description of the genetic structure of neutral and adaptive traits. The knowledge on this adaptive potential coupled with various selection pressures (which is not addressed in the main manuscript but is addressed in the ANNEXE III) would reflect on the capacity of the pathogen to establish in the newly invaded regions.

The adaptive potential of a pathogen is strongly linked to its mode of reproduction. Both sexual and asexual reproduction have their own advantages and disadvantages in the context of evolutionary potential and adaptation capacity (Fisher, 2007; Milgroom, 1996). Asexual/clonal reproduction would allow a rapid amplification of the fittest individuals (de Meeûs *et al.*, 2007; Fisher, 2007), while recombination would allow the production of new gene combinations that unite beneficial mutations that appeared in distinct lineages (Goddard *et al.*, 2005; McDonald and Linde, 2002; Taylor *et al.*, 1999). In addition, sexual recombination could involve the differentiation of certain robust structures, enabling survival of the population during adverse environmental/host conditions (e.g. teliospores;

Anikster, 1986). Although variability in modes of reproduction does exist among pathogens from complete clonality to obligate sexual cycle (Milgroom, 1996; Taylor *et al.*, 1999), pathogens provided with both sexual and asexual reproduction will benefit from the advantages of both of these modes of reproduction (Gladieux *et al.*, 2011; McDonald and Linde, 2002). Hence, the same pathogen could have different modes of reproduction in different geographically spaced population (Ali *et al.*, 2010; Barrett *et al.*, 2008; Saleh *et al.*, 2012b).

This variability in the mode of reproduction can be caused by the loss of sexuality after migration to a newly invaded region. Thus loss of sexuality could be caused by the lack of one of the two mating types (e.g. Magnaporthe grisea; Saleh *et al.*, 2012b) or alternate host (e.g. Puccinia graminis f. sp. tritici; Groth and Roelfs, 1982) in the newly introduced region; or by the degradation of sexual reproduction ability through deleterious mutations accumulation during the process of invasion (e.g. Puccinia striiformis f. sp. tritici; Ali *et al.*, 2010; e.g. Magnaporthe grisea; Saleh *et al.*, 2012a). Whatever the reason, the information on modes of reproduction over a pathogen geographic range would not only enable us to infer its adaptive potential in different regions (Taylor *et al.*, 1999), but would also reflect on the possible centre of origin of the pathogen, from where the pathogen would have migrated to the rest of the world (sexuality considered as ancestral). This, however, will not be the case for pathogens with an obligate sexual cycle as the sexual reproduction is maintained throughout the invasion process (Gladieux *et al.*, 2008). In the case of obligate clonal pathogens, this loss of sexual reproduction could not be useful to track the centre of origin, for example if the populations in the centre of origin had disappeared or if the event of loss of sexual reproduction had already occurred in the ancestral species. However, this approach would be useful for the pathogens with both clonal and sexual reproduction separated in space and/or time. Thus, it will be important to study the variability in sexual reproduction for geographically spaced populations.

Thus once the regions presenting a high diversity and sexual reproduction are identified, it would be important to carry out detailed population analyses at the local level. One main focus of these studies should address the description of modes of reproduction, as they are essential for understanding both demography, epidemiology and adaptive potential of the pathogen (Barrett *et al.*, 2008; Taylor *et al.*, 1999). Through the production of high survival spores, one essential effect of sexuality is to influence temporal maintenance of populations, and therefore impact their effective population sizes. Although pathogens produce millions and trillions of spores during the epidemics, the across year survival is highly limited due to lack of susceptible host and/or favorable environment. The effective number of individuals (effective population size) contributing to the next year's population would be orders of magnitude lower than the number of spores produced at epidemic climax. Although effective population sizes have been estimated for plants and animals (Caballero, 1994; Frankham, 1995), little efforts have been made to estimate the effective population sizes of pathogens, especially in natural populations. Thus combining information on effective size and mode of reproduction are central for understanding the adaptive potential of a pathogen at a local level, and could enable us to pinpoint the potential zones of future invasions.

Despite the increased cases of invasions and emergence, little is known about the sources and migration patterns of previously reported invasions of these pathogens. One explanation is that little is known about the global population structure, centre(s) of diversity and ancestral relationship of worldwide distributed pathogens. *Puccinia striiformis* Westend. f.sp. *tritici* (PST), the causal agent of wheat yellow/stripe rust, is one of these pathogens that remained important in the context of invasion (Hovmøller *et al.*, 2008; Singh *et al.*, 2004; Wellings and McIntosh, 1990) owing to its long distance migration capacity (Brown and Hovmøller, 2002).

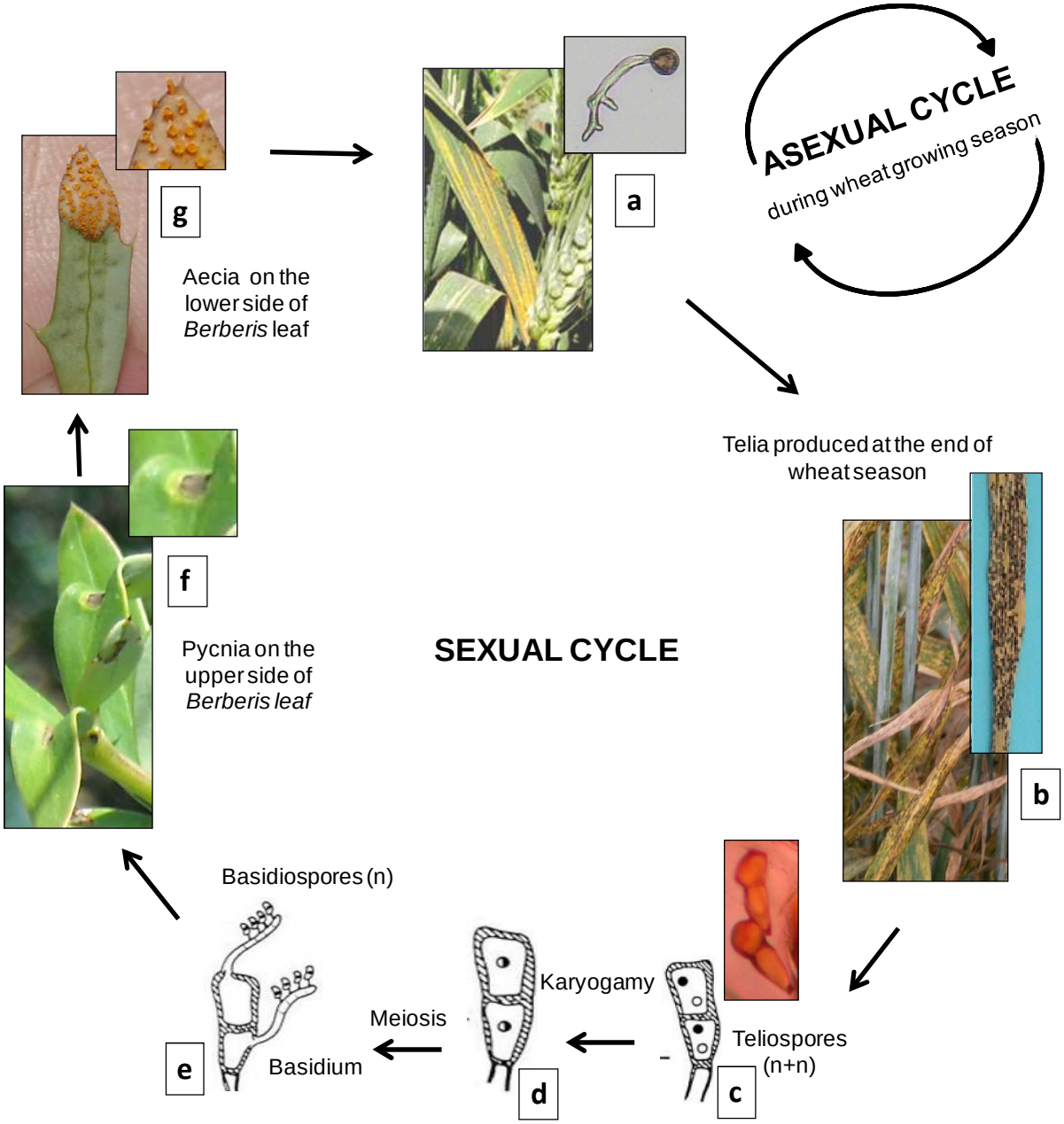


Fig. 1. Life cycle of *Puccinia striiformis*. It passes its epidemic stage on wheat through several asexual cycles in the form of uredinial stage (a), producing telia (b) at the end of wheat season. The telia contains teliospores (c), which undergo kayogamy (d) and meiosis to produce a basidium developing into four basidiospores (e). The basidiospores infect the alternate host (*Berberis* spp.) to produce the disease symptoms at the pycnial (f) and aecial (g) stages on the upper and lower leaf surface, respectively, as observed on *Berberis* spp. at Pakistan.

Wheat yellow rust is an economically important disease in all wheat growing regions with temperate/cold climate or with warm climate but cold and humid nights (Ali *et al.*, 2009; de Vallavieille-Pope *et al.*, 2012; Hau and de Vallavieille-Pope, 2006; Hovmøller *et al.*, 2011; Hovmøller *et al.*, 2010; Singh *et al.*, 2004; Stubbs, 1985; Wellings, 2007; Wellings, 2011; Zadoks, 1961). This pathogen had a particular importance for the development of the discipline of phytopathology, as it was used to demonstrate the Mendelian genetic control of disease resistance in a host for the first time (Biffen, 1905). Since then, it remained an important model pathogen in different studies aimed at understanding the genetics of host-pathogen interaction, such as virulence evolution in response to host resistance deployment (de Vallavieille-Pope *et al.*, 2012; Dedryver *et al.*, 2009; Johnson, 1992).

PST has been considered until very recently to reproduce only asexually through dicaryotic uredospores on wheat with the teliospores produced at the end of cropping season unable to encounter an alternate/aecial host to finalize the sexual cycle (Cummins, 1971; Stubbs, 1985). However, very recently PST has been shown to complete its sexual cycle on different species of *Berberis* (Jin *et al.*, 2010). Although the detailed life cycle for PST is not yet fully described, its analogy with wheat stem rust (*Puccinia graminis* f.sp. *tritici*) could provide an appropriate idea, as all the types of spores have been recovered in the above described study of Jin et al. (2010). The teliospores produced at the end of the epidemic cycle undergo karyogamy and meiosis to give the haploid basidiospores. These basidiospores infect the upper side of *Berberis* spp. leaves to produce pycnia, which produce two types of gametes i.e., pycniospores and receptive hyphae. The pycniospores can fertilise a receptive hypha of opposite mating type producing a dicaryotic mycelium, which produce aecia containing aeciospores on the lower side of the leaf (Fig. 1). The aecia contain aeciospores capable of infecting the wheat crop to start uredial cycles, at the origin of disease epidemics. Although a recombinant population structure was reported in Gansu, China (Mboup *et al.*, 2009), the role of sexual cycle has never been shown under natural field conditions. The sexual cycle has been suggested to play important role in offseason survival of many uredinales, while for PST clonal survival has been assumed on volunteer plants and

wheat seedlings, at least in clonal populations like in NW Europe (Ali *et al.*, 2010; de Vallavieille-Pope *et al.*, 2012). Similarly, long distance migration has also been suggested to play an important role in re-establishment of PST each year (Brown and Hovmøller, 2002; de Vallavieille-Pope *et al.*, 2012; Enjalbert *et al.*, 2005; Hovmøller *et al.*, 2002).

The long distance migration capacity of the pathogen made it important in the context of invasion. The isolates with virulence against a host resistance gene (*Yr9*), first reported in Ethiopia in 1986, invaded the whole Middle East and South Asia within 12 years (Singh *et al.*, 2004). Similarly, PST was absent in South Africa until 1996, date of its first report, probably due to a migration from North and East Africa (Boshoff *et al.*, 2002). While these migrations are mostly caused by wind dispersal of uredospores, the incursion of PST into Australia in 1978/79 was suggested to have been caused by some human intervention (Wellings, 2007). Similar impact of human activity has also been suggested for the more recent increase in geographic range of PST, reaching relatively warm climate zones in south-east USA after 2000 (Chen, 2005). This emergence was linked to the incursion of two closely related atypical aggressive strains (Hovmøller and Justesen, 2007) with high temperature adaptation (Markell and Milus, 2008; Milus *et al.*, 2009). Hovmøller et al. (2008) reported that these aggressive isolates consisted of two closely related PST clonal lineages, one present in USA and Australia (PstS1) and another in Eurasia and Europe (PstS2). However, nothing is known about the source population and migration pattern of these aggressive strains.

Despite the documentation of these continental dispersal events (Brown and Hovmøller, 2002; Singh *et al.*, 2004; Wellings and McIntosh, 1990), little is known about the worldwide population structure and the centre of diversity and origin of the pathogen. The centre of origin (and diversity) for cultivated *triticae* is the Fertile Crescent in the Middle East (Badr *et al.*, 2000; Heun, 1997; Salamini *et al.*, 2002). However, the centre(s) of origin/diversity of pathogens may differ from the centre of

origin/diversity of the corresponding host (Brunner *et al.*, 2007; Zaffarano *et al.*, 2006). Individual studies carried out at the regional and continental level have suggested very low genetic diversity of PST populations in most parts of the world (Bahri *et al.*, 2009; Chen *et al.*, 1993; Enjalbert *et al.*, 2005; Wellings, 2007), while a high diversity was reported in Pakistan (Bahri *et al.*, 2011) and recombinant population structure in China (Duan *et al.*, 2010; Mboup *et al.*, 2009). A detailed study is thus required to describe the population structure of PST at worldwide and local level as well as infer the sources of recent invasions and migration patterns as well as identify its centre of diversity, while inferring on its centre of origin.

The present Ph.D. work was designed in the above described context to address the PST population structure at worldwide and local level and the invasion history of the recently reported invasions. The study was designed utilising the materials and expertise of the two laboratories, INRA France and Aarhus University, Denmark. The overall work was divided into four chapters / sub-parts along with a preliminary task to develop an efficient molecular genotyping technique (presented in ANNEXE II). Each of these parts is discussed in four chapters, while the details of some additional tasks achieved during the Ph.D. are given at the end in annexes.

1. In the first part, a representative set of isolates was assembled from the worldwide collection of isolates to address the worldwide population structure of PST and infer the source populations and migration patterns of recent invasions.

2. The second part was designed to assess whether the difference in diversity of worldwide populations is linked with the difference in sexual reproduction ability. As a proxy for this sex-ability, we measured the telial production, a sex-specific structures that are obligatory for sexual cycle in a subset of isolates representative of worldwide geographical origins.

3. The third part was designed to describe the detailed population structure in Pakistan, while addressing the temporal maintenance of population and spatial population structure at local level.

4. The fourth part was designed to address the temporal maintenance of the recombinant Chinese population through the estimation of the effective population size and the relative contribution of sexual and asexual reproduction.

Literature cited:

Ali S, Leconte M, Walker A-S, Enjalbert J, de Vallavieille-Pope C (2010) Reduction in the sex ability of worldwide clonal populations of *Puccinia striiformis* f.sp. *tritici*. *Fungal Genetics and Biology* **47**, 828-838.

Ali S, Shah SJA, Rahman H, Saqib MS, Ibrahim M, Sajjad M (2009) Variability in wheat yield under yellow rust pressure in Pakistan. *Turkish Journal of Agriculture and Forestry* **33**, 537-546.

Altieri MA (1999) The ecological role of biodiversity in agroecosystems. *Agriculture, Ecosystems & Environment* **74**, 19-31.

Anagnostakis SL (1987) Chestnut blight: The classical problem of an introduced pathogen. *Mycologia* **29**, 23-37.

Anagnostakis SL (2001) American chesnut sprout survival with biological control of the chesnut-blight fungus population. *Forest Ecology and Management* **152**, 225-233.

Anderson P, Cunningham A, Patel N, Morales F, Epstein P, Daszak P (2004) Emerging infectious diseases of plants: pathogen pollution, climate change and agrotechnology drivers. *Trends in Ecology & Evolution* **19**, 535-544.

Anikster Y (1986) Teliospore germination in some rust fungi. *Phytopathology* **76**, 1026-1030.

Badr A, Muller K, Schafer-Pregl R, El Rabey H, Effgen S, Ibrahim HH, Pozzi C, Rohde W, Salamini F (2000) Badr, A. et al. On the origin and domestication history of barley. *Molecular Biology and Evolution* **17**, 499-510.

Bahri B, Leconte M, Ouffroukh A, de Vallavieille-Pope C, Enjalbert J (2009) Geographic limits of a clonal population of wheat yellow rust in the Mediterranean region. *Molecular Ecology* **18**, 4165-4179.

Bahri B, Shah SJA, Hussain S, Leconte M, Enjalbert J, de Vallavieille-Pope C (2011) Genetic diversity of wheat yellow rust population in Pakistan and its relationship with host resistance. *Plant Pathology* **60**, 649-660.

Barrett LG, Thrall PH, Burdon JJ, Nicotra AB, Linde CC (2008) Population structure and diversity in sexual and asexual populations of the pathogenic fungus *Melampsora lini*. *Molecular Ecology* **17**, 3401-3415.

Biffen RH (1905) Mendel's laws of inheritance and wheat breeding. *Journal of Agricultural Science* **1**, 4-48.

Boshoff WHP, Pretorius ZA, van Niekerk BD (2002) Establishment, distribution, and pathogenicity of *Puccinia striiformis* f.sp. *tritici* in South Africa. *Plant Disease* **86**, 485-492.

Brown JKM, Hovmøller MS (2002) Aerial dispersal of fungi on the global and continental scales and its consequences for plant disease. *Science* **297**, 537-541.

Brunner PC, Schurch S, McDonald BA (2007) The origin and colonization history of the barley scald pathogen Rhynchosporium secalis. *Journal of Evolutionary Biology* **20**, 1311-1321.

Caballero A (1994) Developments in the prediction of effective population size. *Heredity* **73**, 657—679.

Campbell F (2001) The science of risk assessment for phytosanitary regulation and the impact of changing trade regulations. *BioScience* **51**, 148-153.

Chen X, Line RF, Leung H (1993) Relationship between virulence variation and DNA polymorphism in *Puccinia striiformis*. *Phytopathology* **83**, 1489-1497.

Chen XM (2005) Epidemiology and control of stripe rust [*Puccinia striiformis f. sp. tritici*] on wheat. *Canadian Journal of Plant Pathology* **27**, 314–337.

Cummins GB (1971) *The rusts fungi of cereals, grasses, and bamboos* Springer-Verlag, New York.

de Meeûs T, Prugnolle F, Agnew P (2007) Asexual reproduction: genetics and evolutionary aspects. *Cellular and Molecular Life Science* **64**, 1355–1372.

de Vallavieille-Pope C, Ali S, Leconte M, Enjalbert J, Delos M, Rouzet J (2012) Virulence dynamics and regional structuring of *Puccinia striiformis* f. sp. *tritici* in France between 1984 and 2009. *Plant Disease* **96**, 131-140.

Dean R, Van Kan J, Pretorius ZA, Hammond-Kosack KE, Pietro A, Spanu PD, Rudd JJ, Dickman M, Kahmann R, Ellis J, Foster GD (2012) The top 10 fungal pathogens in molecular plant pathology. *Molecular Plant Pathology* **13**, 414-430.

Dedryver F, Paillard S, Mallard S, Robert O, Trottet M, Nègre S, Verplancke G, Jahier J (2009) Characterization of genetic components involved in durable resistance to stripe rust in the bread wheat 'Renan'. *Phytopathology* **99**, 968-973.

Desprez-Loustau M-L, Robin C, Buée M, Courtecuisse R, Garbaye J, Suffert F, Sache I, Rizzo DM (2007) The fungal dimension of biological invasions. *Trends in Ecology and Evolution* **22**, 472–480.

Dilmaghani A, Gladieux P, Gout L, Giraud T, Brunner PC, Grandaubert J, Stachowiak A, Balesdent MH, Rouxel T (2012) Migration patterns and changes in population biology associated with the worldwide spread of the oilseed rape pathogen *Leptosphaeria maculans*. *Molecular Ecology* **doi: 10.1111/j.1365-294X.2012.05535.x**.

Duan X, Tellier A, Wan A, Leconte M, de Vallavieille-Pope C, Enjalbert J (2010) *Puccinia striiformis* f.sp. *tritici* presents high diversity and recombination in the over-summering zone of Gansu-China. *Mycologia* **102**, 44-53.

Dutech C, Barres B, Bridier J, Robin C, Milgroom MG, Ravigné V (2012) The chestnut blight fungus world tour: successive introduction events from diverse origins in an invasive plant fungal pathogen. *Molecular Ecology* **doi: 10.1111/j.1365-294X.2012.05575.x**.

Enjalbert J, Duan X, Leconte M, Hovmøller MS, de Vallavieille-Pope C (2005) Genetic evidence of local adaptation of wheat yellow rust (*Puccinia striiformis* f. sp. *tritici*) within France. *Molecular Ecology* **14**, 2065-2073.

Fisher MC (2007) The evolutionary implications of an asexual lifestyle manifested by *Penicillium marneffei*. In: *Sex in fungi: molecular determination and evolutionary implications.* (ed. Heitman JK, J.W.; Taylor, J.W.; Casselton, L.A.), p. 542. ASM Press, Washington.

Frankham R (1995) Effective population-size adult-population size ratios in wildlife — a review. *Genetics Research* **66**, 95-107.

Gilbert GS (2002) Evolutionary ecology of plant diseases in natural ecosystems. *Annual Review of Phytopathology* **40**, 13-43.

Giraud T, Gladieux P, Gavrilets S (2010) Linking the emergence of fungal plant diseases with ecological speciation. *Trends in Ecology & Evolution* **25**, 387-395.

Gladieux P, Byrnese E, Fisher MC, Aguleta G, Heitman J, T. G (2011) Epidemiology and evolution of fungal pathogens, in plants and animals. In: *Genetics and Evolution of Infectious Diseases* (ed. M T), p. 772. Elsevier.

Gladieux P, Zhang X-G, Afoufa-Bastien D, Sanhueza R-MV, Sbaghi M, Le Cam B (2008) On the origin and spread of the scab disease of apple: out of Central Asia. *PLoS ONE* **3**, e1455.

Goddard MR, Godfray HCJ, Burt A (2005) Sex increases the efficacy of natural selection in experimental yeast populations. *Nature* **434**, 636-640.

Groth JV, Roelfs AP (1982) Effect of sexual and asexual reproduction on race abundance in cereal rust fungus populations. *Phytopathology* **72**, 1503-1507.

Harvell CD, Mitchell CE, Ward JR, Altizer S, Dobson AP, Ostfeld RS, Samuel MD (2002) Climate Warming and Disease Risks for Terrestrial and Marine Biota. *Science* **296**, 2158.

Hau B, de Vallavieille-Pope C (2006) Wind-dispersed diseases. In: *The Epidemiology of Plant Diseases*, pp. 387-416. Springer Netherlands, Netherlands.

Heun M (1997) Site of einkorn wheat domestication identified by DNA fingerprinting. *Science and Justice* **278**, 1312-1314.

Hovmøller MS, Justesen AF (2007) Appearance of atypical *Puccinia striiformis* f. sp. *tritici* phenotypes in north-western Europe. *Australian Journal of Agricultural Research* **58**, 518-524.

Hovmøller MS, Justesen AF, Brown JKM (2002) Clonality and long-distance migration of *Puccinia striiformis* f.sp. *tritici* in north-west Europe. *Plant Pathology* **51**, 24-32.

Hovmøller MS, Sørensen CK, Walter S, Justesen AF (2011) Diversity of *Puccinia striiformis* on cereals and grasses. *Annual Review of Phytopathology* **49**, 197-217.

Hovmøller MS, Walter S, Justesen AF (2010) Escalating threat of wheat rusts. *Science* **329**, 369.

Hovmøller MS, Yahyaoui AH, Milus EA, Justesen AF (2008) Rapid global spread of two aggressive strains of a wheat rust fungus. *Molecular Ecology* **17**, 3818-3826.

Jin Y, Szabo LJ, Carson M (2010) Century-old mystery of Puccinia striiformis life history solved with the identification of Berberis as an Alternate Host. *Phytopathology* **100**, 432-435.

Johnson R (1992) Reflection of a plant pathologist on breeding for disease resistance, with emphasis on yellow rust and eyespot of wheat. *Plant Pathology* **41**, 239-254.

Markell SG, Milus EA (2008) Emergence of a Novel Population of *Puccinia striiformis* f. sp. *tritici* in Eastern United States. *Phytopathology* **98**, 632-639.

Mboup M, Leconte M, Gautier A, Wan AM, Chen W, de Vallavieille-Pope C, Enjalbert J (2009) Evidence of genetic recombination in wheat yellow rust populations of a Chinese oversummering area. *Fungal Genetics and Biology* **46**, 299-307.

McDonald BA, Linde C (2002) Pathogen population genetics, evolutionary potential, and durable resistance. *Annual Review of Phytopathology* **40**, 349-379.

Milgroom MG (1996) Recombination and the multilocus structure of fungal populations. *Annual Review of Phytopathology* **34**, 457-477.

Milus EA, Kristensen K, Hovmøller MS (2009) Evidence for increased aggressiveness in a recent widespread strain of *Puccinia striiformis* f.sp. *tritici* causing stripe rust of wheat. *Phytopathology* **99**, 89-94.

Palm M (2001) Systematics and the impact of invasive fungi on agriculture in theUnited States. *BioScience* **51**, 141-147.

Parker IM, Gilbert GS (2004) The evolutionary ecology of novel plant-pathogen interactions. *Annual Review of Ecology, Evolution, and Systematics* **35**, 675-700.

Perrings C, Williamson M, Barbier E, Delfino D, Dalmazzone S, J. S, Simmons P, Watkinson A (2002) Biological invasion risks and the public good: an economic perspective. *Conservation Ecology* **6**, Article 1 (online).

Pimentel D, Zuniga R, Morrison D (2005) Update on the environmental and economic costs associated with alieninvasive species in the United States. *Ecological Economics* **52**, 273-288.

Salamini F, Ozkan H, Brandolini A, Schafer-Pregl R, Martin W (2002) Genetics and geography of wild cereal domestication in the Near East. *Nature Reviews Genetics* **3**, 429-441.

Saleh D, Milazzo J, Adreit H, Tharreau D, Fournier E (2012a) Asexual reproduction induces rapid and permanent loss of sexual reproduction ability of the fungal rice pathogen, *Magnaporthe oryzae*: results from in vitro experimental evolution assays. *BMC Evolutionary Biology* **12**, 42.

Saleh D, Xu P, Shen Y, Li C, Adreit H, Milazzo J, Ravigné V, Bazin E, Nottéghem J-L, Fournier E, Tharreau D (2012b) Sex at the origin: an Asian population of the rice blast fungus Magnaporthe oryzae reproduces sexually. *Molecular Ecology* **DOI: 10.1111/j.1365-294X.2012.05469.x**.

Singh RP, Hodson DP, Jin Y, Huerta-Espino J, Kinyua MG, Wanyera R, Njau P, Ward WR (2006) Current status, likely migration and strategies to mitigate the threat to wheat production from race Ug99 (TTKS) of stem rust pathogen. *CAB Reviews* **1**, 054.

Singh RP, William HM, Huerta-Espino J, Rosewarne G (2004) Wheat rust in Asia: meeting the challenges with old and new technologies.

Smart C, Fry W (2001) Invasions by the late blight pathogen: renewed sex and enhanced fitness. *Biological Invasions* **3**, 235-243.

Stubbs RW (1985) Stripe rust. In: *The Cereal Rusts, Diseases, Distribution, Epidemiology and Control* (eds. Roelfs AP, Bushnell WR), pp. 61–101. Academic Press, London, UK.

Stukenbrock EH, McDonald BA (2008) The origins of plant pathogens in agro-ecosystems. *Annual Review of Phytopathology* **46**, 75-100.

Taylor JW, Jacobson DJ, Fisher MC (1999) The evolution of asexual fungi: Reproduction, speciation and classification. *Annual Review of Phytopathology* **37**, 197-246.

von Broembsen SL (1989) Invasions of natural ecosystems by plant pathogens. In: *Biological invasion: a global perspective* (ed. Drake JA). SCOPE. John Wiley & Sons Ltd.

Wellings CR (2007) *Puccinia striiformis* in Australia: A review of the incursion, evolution and adaptation of stripe rust in the period 1979-2006. *Australian Journal of Agricultural Research* **58**, 567-575.

Wellings CR (2011) Global status of stripe rust: a review of historical and current threats. *Euphytica* **179**, 129-141.

Wellings CR, McIntosh RA (1990) *Puccinia striiformis* f.sp. *tritici* in Australia: pathogenic changes during the first 10 years. *Plant Pathology* **39**, 316-325.

Wingen LU, Shaw MW, Brown JKM (2012) Long-distance dispersal and its influence on adaptation to host resistance in a heterogeneous landscape. *Plant Pathology* **Doi: 10.1111/j.1365-3059.2012.02621.x**.

Wolfe MS (1985) The current status and prospects of multiline cultivars and variety mixtures for disease control. *Annual Review of Phytopathology* **23**, 251-273.

Zadoks JC (1961) Yellow rust on wheat, studies in epidemiology and physiologic specialization. *Tijdschrift over Plantenziekten* **67**.

Zaffarano PL, McDonald BA, Zala M, Linde CC (2006) Global hierarchical gene diversity analysis suggests the Fertile Crescent is not the center of origin of the barley scald pathogen *Rhynchosporium secalis*. *Phytopathology* **96**, 941-950.

CHAPTER-I

Origin, migration routes and genetic structure of worldwide populations of the wheat yellow rust pathogen, *Puccinia striiformis* f.sp. *tritici*

Sajid Ali, Pierre Gladieux, Jérôme Enjalbert, Marc Leconte, Angélique Gautier, Annemarie F. Justesen, Mogens S. Hovmøller, Claude de Vallavieille-Pope

Origin, migration routes and genetic structure of worldwide populations of the wheat yellow rust pathogen, *Puccinia striiformis* f.sp. *tritici*

The article will be submitted to an international journal

Origin, migration routes and genetic structure of worldwide populations of the wheat yellow rust pathogen, *Puccinia striiformis* f.sp. *tritici*

Sajid Ali[1]*, Pierre Gladieux[2,3], Jérôme Enjalbert[4], Marc Leconte[1], Angélique Gautier[1], Annemarie F. Justesen[5], Mogens S. Hovmøller[5], Claude de Vallavieille-Pope[1]

[1] INRA UR 1290 BIOGER-CPP, BP01, F-78850 Thiverval-Grignon, France

[2] UMR 8079 Ecologie Systematique Evolution, Univ. Paris-Sud 11, F-91405 Orsay, France; CNRS, 91405 Orsay, France

[3] Department of Plant and Microbial Biology, University of California, Berkeley, CA 94720-3102
[4] INRA UMR 320 Génétique Végétale, Ferme du Moulon, F-91190 Gif sur Yvette, France

[5] Department of Agroecology, Aarhus University, Flakkebjerg, DK-4200 Slagelse, Denmark

Short title: Worldwide wheat yellow rust population structure

* Corresponding author (sajid.ali@versailles.inra.fr)

ABSTRACT

Analyses of the large-scale population structure of pathogens enable the identification of migration patterns, diversity reservoirs or longevity of populations, the understanding of current evolutionary trajectories and the anticipation of future ones. This is particularly important for long-distance migrating fungal pathogens such as *Puccinia striiformis* f.sp. *tritici* (PST), which is capable of rapid spread to new regions and crop varieties. Although a range of recent PST invasions at continental scales are well documented, the worldwide population subdivisions and the center of origin are still unknown. In this study, we used multilocus microsatellite genotyping to infer the worldwide

population structure of PST and the origin of new invasions based on 409 isolates representative of the distribution of the fungus on six continents. Bayesian and multivariate clustering methods partitioned the set of multilocus genotypes into six distinct genetic groups associated with their likely geographical origin. Analyses of the linkage disequilibrium and genotypic diversity indicated a strong regional heterogeneity in levels of recombination, with clear signatures of recombination in the Himalayan (Nepal and Pakistan) and near-Himalayan regions (China) and a predominant clonal population structure in other regions. Our analysis confirmed the Middle East-Red Sea Area as the most likely source of newly spreading, high-temperature-adapted strains; Europe as the source of South American, North American and Australian populations; and Mediterranean-Central Asian populations as the origin of South African populations. The higher genotypic diversity, recombinant population structure and high sex ability of the pathogen in the Himalayan region pinpoints this region as the possible center of origin of PST. Although most geographic populations are not markedly affected by recent dispersal events, this study emphasizes the importance of human activities on the recent long-distance spread of the disease.

Keywords: fungal pathogen, stripe rust, biological invasions, disease emergence, basidiomycete

AUTHOR SUMMARY

Domestication of ecosystems, climate change and expanding global trade have accelerated the pace of disease emergence, caused by their introduction into new areas with susceptible host or the spread of new damaging genotypes. Wheat yellow rust pathogen is one of these pathogens with recent reports of invasions, significantly affecting worldwide wheat production. However, its origin and ancient migration routes remain unclear and the source of newly spreading strains is largely unknown. This information is important for understanding the trajectories of current invasions and forecast the future spread of PST, and more generally develop risk-assessment models of pathogen emergence. We analyzed a set of worldwide representative isolates of PST, which enabled us to identify six different area-specific populations. Using population genetics tools, we identified its centre of diversity in the

Himalayan and near Himalayan region. We also identified the origin of recently emerged populations; Middle East-Red Sea area as the source of high-temperature-adapted strains; Europe as the source of American and Australian populations; and Mediterranean-Central Asian populations as the origin of South African populations. We demonstrate the importance of human activities on the recent long-distance spread of the disease, though most geographic populations are not markedly affected by recent dispersal events.

INTRODUCTION

Domestication of ecosystems, climate change and expanding global trade have accelerated the pace of pathogen emergence and spread (Parker and Gilbert, 2004; Stukenbrock and McDonald, 2008). Widely distributed and genetically homogenous crop genotypes are conducive for rapid pathogen emergence and subsequent propagation across large areas. Crop plants with a worldwide distribution are potentially vulnerable to new invasions by their pathogens. Even when these pathogens are initially endemic to restricted geographical areas, they can rapidly emerge in new regions, provided they encounter a farmland with susceptible hosts and favorable environmental conditions. For widely grown food crops, many pathogens were introduced long ago and are now so geographically widespread that they do not come to mind as invasive pathogens (Palm, 2001). Although introductions could have occurred centuries ago, the evolution of such ubiquitous pathogens remains a significant cause of concern due to the risk of re-emergence caused by accidental dissemination of new, multi-virulent races (Singh *et al.*, 2006; Singh *et al.*, 2004) or new, highly aggressive strains (Hovmøller *et al.*, 2008). An accurate understanding of the origin, distribution of diversity reservoirs and past and recent migration routes of these pathogens is crucial for understanding current epidemics, developing risk-assessment models and alleviating the potential adverse effects of disease emergence (Campbell, 2001; Perrings *et al.*, 2002). This is particularly true for pathogens capable of long-distance migration, for which any newly advantageous mutant (increased virulence, aggressiveness or resistance to fungicides) has the potential to invade a large geographical area (Brown and Hovmøller, 2002).

Yellow (stripe) rust on wheat, which is caused by *Puccinia striiformis* f.sp. *tritici* (PST), is present in most wheat-growing regions of the world (Ali *et al.*, 2009a; Ali *et al.*, 2009b; de Vallavieille-Pope *et al.*, 2012; Hovmøller *et al.*, 2011; Hovmøller *et al.*, 2010; Singh *et al.*, 2004) and has major negative impacts on wheat production due to re-emergences and invasions (Hovmøller *et al.*, 2008; Singh *et al.*, 2004; Wellings, 2007). As for most ubiquitous pathogens of major crops, the origin, introduction pathways and current population structure of wheat yellow rust remains largely unknown. Although an origin in Transcaucasia has been hypothesized based on disease prevalence and geographical barriers (Stubbs, 1985; Stubbs, 1988), it has never been assessed in light of new knowledge on the population structure of PST. Long-distance dispersal by wind is thought to play a key role in the dissemination of the disease. The fungus is capable of long-distance migration, with well-documented cases of recurrent re-establishment of pathogen populations in areas where there are no host plants during summer/winter to allow the pathogen survival, such as the main wheat-growing provinces of north-eastern China (Brown and Hovmøller, 2002). Such spread can be due to successive jumps from field to field by this polycyclic disease throughout the season, as in the USA (Kolmer, 2005), as well as direct long-distance migration caused by winds, as documented between England and Denmark (Justesen *et al.*, 2002). Accidental spore transport via human travel may also contribute to the intercontinental dispersal of the pathogen, as exemplified by the introduction of PST into Australia in 1979 from Europe, certainly through contaminated clothing (Wellings, 2007).

Despite the capacity for long-distance migration, the worldwide spread of PST is relatively recent, with most emergences reported only within the last several decades. The pathogen reached South Africa in 1996 from an unknown source, but the first pathotypes detected were similar to those present in the Middle East and Mediterranean regions (Boshoff *et al.*, 2002; Hovmøller *et al.*, 2008). PST was also first reported in South America in the early 20th century, with an unknown origin (Rudorf and Job, 1934; Stubbs, 1985). More recently, an expansion of the geographic range of PST into the warm

climate of south-eastern USA (Chen, 2005) was shown to be due to the emergence of an aggressive strain adapted to higher temperatures than usually reported to be optimal for PST (Markell and Milus, 2008; Milus et al., 2009). The same strain was found in Australia, while another closely related one was observed in Europe, Central and West Asia and the Red Sea Area (Hovmøller *et al.*, 2008). In addition to recently colonized areas, the disease is known to periodically re-emerge through the acquisition of new virulences. These events are well documented for pathotypes carrying virulence against resistance gene *Yr9*, with a first report in 1986 in Eastern Africa (Ethiopia) and subsequent invasions of the Middle East, Pakistan and India, reaching Bangladesh in only 12 years (Singh *et al.*, 2004). The geographical origins of most of the emerging strains are unknown. However, the population structure of PST is, therefore, likely to display the hallmarks of a complex mixture of re-emergences over continuous wheat-growing areas and rare founder events due to long-distance migration. Recent spreads of the disease are likely to induce marked changes in patterns of population differentiation among regions, potentially erasing the ancient evolutionary history.

Our understanding of the population structure of PST has been changed very recently. Independent studies using different genetic markers revealed a clonal population structure with low genetic diversity in Europe (Enjalbert *et al.*, 2005; Hovmøller *et al.*, 2002), USA (Chen *et al.*, 1993) and Australia (Wellings, 2007). This was in accordance with the accepted hypothesis of strictly asexual reproduction of the fungus. However, a study performed in the Gansu province of China revealed a PST population with an almost fully recombinant genetic structure (Mboup *et al.*, 2009). Similarly, isolates from Pakistan presented a higher diversity than normally found in clonal populations (Bahri *et al.*, 2011). Together, these studies suggested differences in the level of diversity among PST populations, with Asian populations being more diverse than other populations and probably experiencing regular recombination.

Table 1. Sampling regions and number of isolates selected for a world-wide phylogeographical analysis of *Puccinia striiformis* from wheat.

Continent	Country	Number of isolates	Sampling year
Asia	China	71	2004-05
	Nepal	55	2008
	Pakistan	68	2004-2006, 2008
	Afghanistan	7	2009
	Cyprus	9	2005-2006
	Iran	17	2005
	Israel	7	2005-2006
	Lebanon	5	2006
	Turkey	10	2005
	Yemen	12	2003, 2005
	Azerbaijan	11	2005
	Kazakhstan	6	2005
	Kyrgyzstan	7	2005
	Uzbekistan	1	2003
Africa	South Africa	6	1996-2004
	Eritrea	23	2002-2005
	Tunisia	4	2005, 2007
	Morocco	2	2006
	Algeria	7	2005-2006
Europe	Italy	2	1998-2006
	Portugal	4	2006
	Spain	2	2006
	Denmark	14	1995- 2006
	France	6	1997-2008
	United Kingdom	5	1975, 1978, 1991-1998
North America	Mexico	4	1989, 2002-2003
	USA	16	1981-1983, 1991-1997
South America	Argentina	1	2010
	Brazil	6	2010
	Chile	7	2010
	Uruguay	10	2010
Australia	Australia	4	2002-2004
Total		409	

In the present study, we assembled a representative set of isolates from a larger collection of isolates from the worldwide geographical range of PST and analyzed their genetic variability using a single set of highly variable genetic markers. Our objectives were the following: i) to identify the main genetic groups in modern PST populations; ii) to identify a possible centre of diversity; and iii) to identify the geographical origin of recently emerged populations and assess the ancestral relationships among geographically spaced populations.

MATERIALS AND METHODS

Selection of isolates. A set of 409 isolates was selected to represent 11 geographical regions on six continents (Africa, Asia, Australia, Europe, North America and South America) from a collection of more than 4,000 isolates available at Institut National de la Recherche Agronomique (INRA), France and Aarhus University, Denmark. The selection was made to maximize the representation of each population (partially assessed previously by AFLP, microsatellites and virulence profiles (Ali *et al.*, 2010; Bahri *et al.*, 2009; Bahri *et al.*, 2011; de Vallavieille-Pope *et al.*, 2012; Enjalbert *et al.*, 2005; Hovmøller *et al.*, 2008; Mboup *et al.*, 2009)) such that isolates from different genotypic groups were present in any given geographical region. Isolates representative of aggressive strains were selected from the two recently emerged aggressive strains, PstS1 (associated with the post-2000 epidemics in the USA and Australia), and the European strain, PstS2, as well as a set of aggressive isolates frequently reported in Europe, PstS3, which were lesser aggressive than PstS1 and PstS2 (Milus *et al.*, 2009). Details regarding the number of isolates are shown in Table. 1.

Molecular genotyping. For most isolates, DNA was already available, having been previously extracted through modified CTAB protocols (Enjalbert *et al.*, 2002; Justesen *et al.*, 2002). For isolates received from Pakistan and Nepal in 2008 and China in 2005, DNA was extracted from 5 mg of spores following Ali et al.(2011; ANNEX II of thesis). All of the isolates were multiplied from single pustule lesions to avoid a mixture of genotypes. Molecular genotyping was carried out using a set of 20

microsatellite loci in three multiplex reactions, with subsequent separation of the PCR products using a Beckman Coulter CEQ-8000 DNA Analyzer. Electrophorograms were processed using the CEQ-8000 Genetic Analysis System Software (Beckman Coulter) (ANNEX II of thesis; Ali *et al.*, 2011).

Analyses of population subdivision. The level of population subdivision among different geographical regions was assessed using both model-based Bayesian and non-parametric, multivariate clustering approaches. We used the model-based Bayesian method implemented in STRUCTURE 2.2 (Pritchard *et al.*, 2000). The rationale of this method is to assign multilocus genotypes to different clusters while minimizing the Hardy-Weinberg disequilibrium and the gametic phase disequilibrium between loci within clusters (where the number of clusters may be unknown). The Monte Carlo Markov Chain (MCMC) sampling scheme was run for 200,000 iterations with a 100,000 burn-in period, with K ranging from 1 to 10 and 20 independent replications for each K. The STRUCTURE outputs were processed with CLUMPP (Jakobsson and Rosenberg, 2007); a G'-statistic greater than 80% was used to assign groups of runs to a common clustering pattern.

Because STRUCTURE can overestimate the number of clusters when there is relatedness among some genotypes (e.g., due to asexual reproduction; Gao *et al.*, 2007), we also analyzed the level of population subdivision using a non-parametric approach that does not rely on a particular population model. We used discriminate analyses of principal components (DAPC), implemented in the ADEGENET package in the R environment (Jombart *et al.*, 2010). The number of clusters was identified based on the Bayesian Information Criterion (BIC), as suggested by Jombart *et al.* (2010). The relatedness among populations was plotted using a neighbor-joining population tree based on the genetic distance D_A (Nei *et al.*, 1983), as implemented in the POPULATION program (Langella, 2008). Significance was assessed using 1000 bootstraps. The level of population differentiation was assessed using pairwise F_{ST} statistics among pairs of populations (GENETIX 4.05.2 (Belkhir *et al.*, 2004)).

Analyses for genetic variability and recombination. The quality of the set of markers for inferring population structure was tested by assessing the ability of the set of microsatellite loci to detect multilocus genotypes (MLGs) under panmixia, using GENCLONE (Arnaud-Haond and Belkhir, 2007). The redundancy of the set of loci was tested by estimating the linkage disequilibrium among different loci and generating 1000 random permutations with GENETIX 4.05.2 (Belkhir *et al.*, 2004). Within-population variability was assessed using allele richness and gene diversity, calculated with FSTAT 2.9.3 (Goudet, 2001). Private allelic richness was estimated using a rarefaction approach, implemented in ADZE (Szpiech *et al.*, 2008). Observed (Ho) and unbiased expected heterozygosity (He) were computed using GENETIX 4.05.2 (Belkhir *et al.*, 2004). The null hypothesis of Hardy-Weinberg equilibrium within each population was tested using the exact test implemented in GENEPOP 4.0 (Raymond and Rousset, 1995). Calculations were performed both on the whole dataset and on the clone-corrected data (i.e., a dataset in which only one representative of each repeated MLG was kept). Only the clone-corrected data are reported in cases where the two datasets yielded different results because the sampling during epidemics would result in over-representation of certain clones due to the recent/epidemic clonality resulting from epidemic clonal structure (Maynard-Smith *et al.*, 1993).

Ancestral relationship and migration patterns among populations: Different competing scenarios were tested to infer the ancestral relationship among populations through Approximate Bayesian Computations (ABC) analyses implemented in DIYABC (Cornuet *et al.*, 2010; Cornuet *et al.*, 2008). The method has been reported to be appropriate for complex population genetic models (Cornille *et al.*, 2012; Dilmaghani *et al.*, 2012), as instead of exact likelihood estimation, the method estimates the posterior probabilities of given scenarios based on the posterior distributions of demographic parameters from observed and simulated datasets.

We used a "hierarchal" strategy for comparing different scenarios, based on our understanding of the population structure in different regions. In a first step we made the comparison between three populations at a time, termed as "triplets". We started with the three recombinant populations in the centre of diversity (Pakistan, Nepal and China), and then compared the rest of the populations among each other using the same triplet strategy and with these recombinant populations (Supporting information_on_ABC_Analyses). In the second step we explored the relationship of Middle-Eastern, Central Asian and Mediterranean populations with those of the centre of diversity. In the third step we included the NW European population to explore the ancestral relationship among the overall world populations. For each dataset, parameters were estimated for the most appropriate scenarios. The results shown in the thesis chapter will only be based on "triplet" results.

A total of 10^6 simulated data was generated for each scenario under the generalized stepwise mutation model, with two parameters i.e., the mean mutation rate (I) and the mean parameter (P) of the geometric distribution used to model the length of mutation events (in number of repeats). Due to the lack of empirical estimates of mutation rate for microsatellites in PST, the mean mutation rate was drawn from a uniform distribution of 10^{-4} to 10^{-3}, while the mutation rate at each locus was drawn from a gamma distribution (mean = μ, shape =2). The parameter P was kept in the range of 0.1 to 0.3. A range of 40 contiguous allelic states was kept for each locus, characterized by the individual value of mutation rate (l_L) and the parameter of the geometric distribution (P_L), which were obtained from a Gamma distribution (with mean =1, range 5 x 10^{-5} to 5 x 10^{-2} for I_L; and mean = P, shape = 2, shape 0.01-0.09 for P_L). Mean number of alleles per locus, mean genetic diversity (Nei, 1978), mean variance in allele size, genetic differentiation between pairwise groups, F_{ST} (Weir and Cockerham, 1984), and genetic distance $\delta\mu$ ($\delta\mu$; Goldstein *et al.*, 1995) were used as summary statistics.

A polychotomous logistic regression procedure (Fagundes *et al.*, 2007) was used to estimate the relative posterior probabilities of different scenarios using the 1% of simulated datasets closest to the

observed data. The limiting distribution of the maximum likelihood estimators was used to compute the confidence interval of the posterior probabilities. The posterior distribution of parameters were estimated for the most likely scenario using the local linear regression (Beaumont *et al.*, 2002; Cornuet *et al.*, 2008) on 1% simulated datasets closest to the observed data. Confidence in model choice was assessed using a leave-one-out method (Csilléry *et al.*, 2011). For each model we drew 500 of the 1,000,000 simulated datasets used for model selection and treated them as observed datasets (*i.e.*, pseudo-observed datasets). Posterior probabilities of competing models were evaluated for each pseudo-observed dataset, using all remaining simulated datasets and the same methodology as described for the observed dataset. Confidence in model choice was then estimated using the number of pseudo-observed dataset that gave higher posterior probability to the model they had been simulated with. In tests of goodness-of-fit (i.e., model checking), we simulated datasets of similar numbers of markers as observed datasets and calculated for each dataset the average across loci of several test quantities. The set of test quantities included the summary statistics used in analyses of the observed dataset. Because using the same statistics in parameter inference and model checking can overestimate the quality of the fit (Cornuet *et al.*, 2010), we selected additional summary statistics that had not been used in parameter inferences: mean allele size variance across loci, mean index of classification and mean gene diversity across loci. Test statistics computed from observed data were then ranked against the distributions obtained from simulated datasets (Cornuet *et al.*, 2010). (The results of model checking and confidence in scenario choice are not shown in the thesis chapter and were only discussed as future perspectives during the oral PhD defense. This will be the addressed in the version of this chapter submitted to an international journal.)

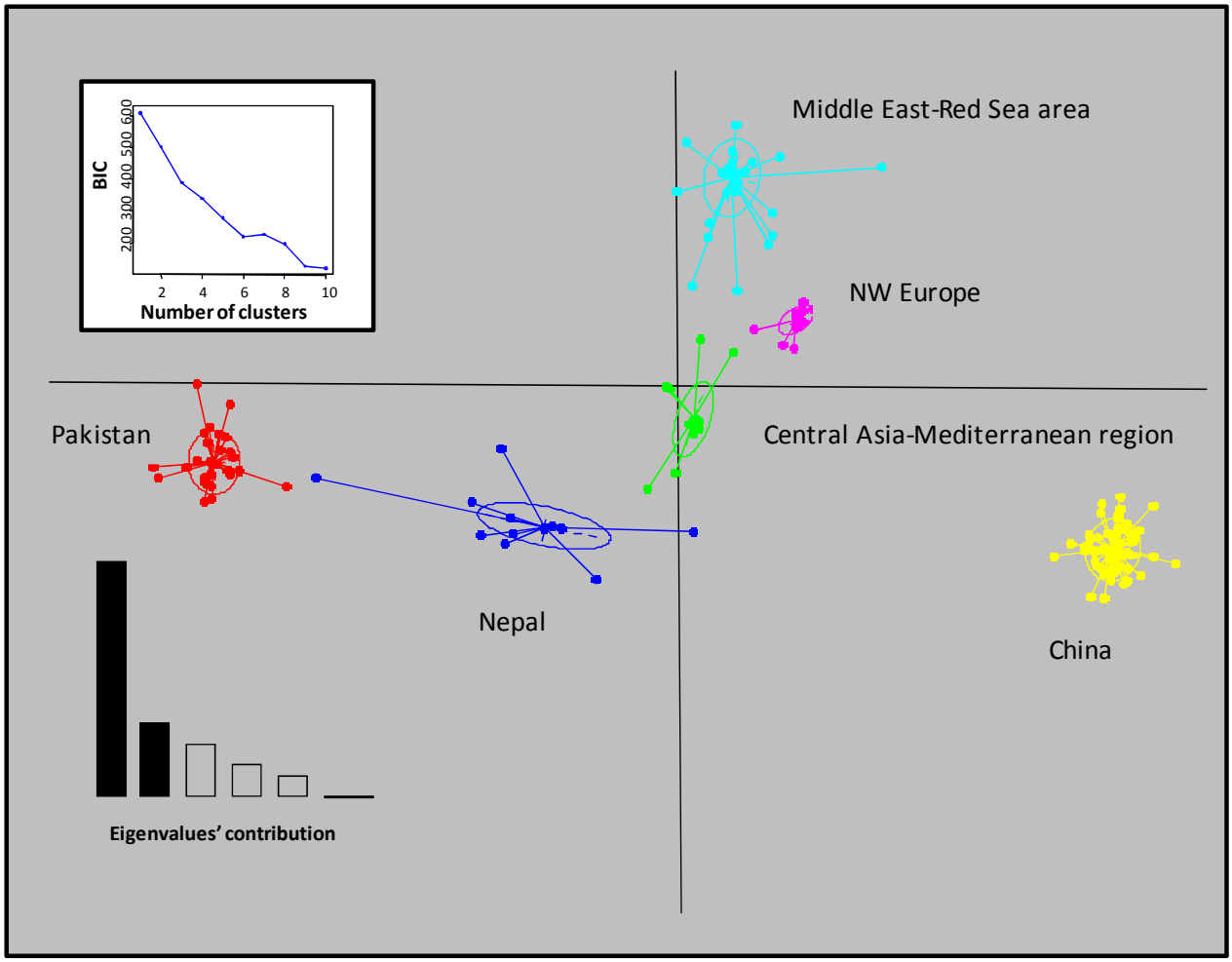


Fig. 1. Discriminate analysis of principal components (DAPC) analysis of worldwide PST populations sampled from different geographical regions, identified six distinct genetic groups of PST.

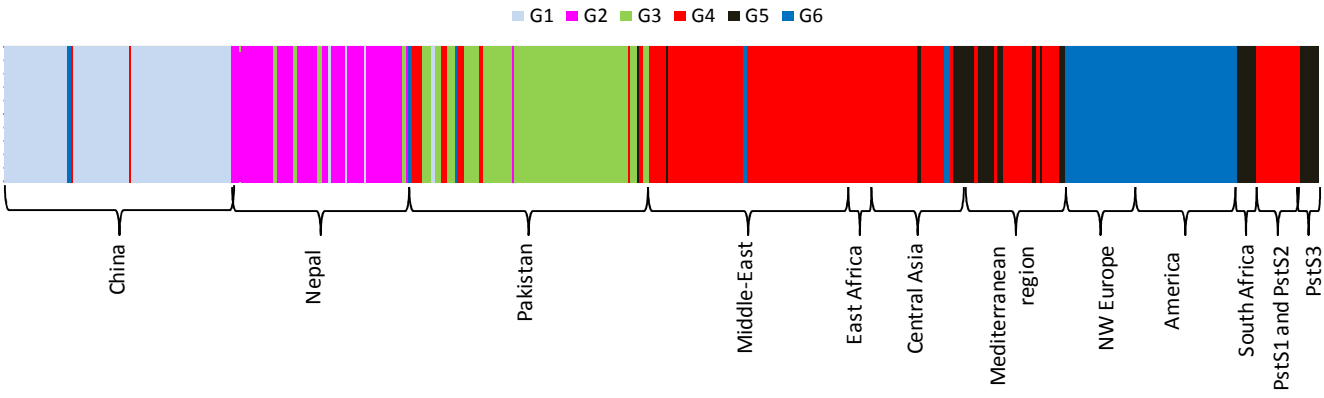


Fig. 2. Assignment of 409 PST isolates representing worldwide geographical regions to genetic groups for the optimal K-value (K =6) in the DAPC analysis. PstS1 and PstS2 refers to the two closely related aggressive strains, while PstS3 refers to the older aggressive isolates regularly reported in Europe.

RESULTS

Summary of genetic variation

We performed multilocus genotyping of 409 PST isolates, representatives of a worldwide collection, using a set of 20 microsatellite markers. Plotting the multilocus genotypes detected against the number of loci re-sampled showed that the full set of SSRs was sufficient for discriminating clonal lineages (supplementary files; Fig. S1). No significant linkage disequilibrium was found among SSR loci (data not shown), suggesting a lack of redundancy among markers. Some of the loci were monomorphic in certain geographical areas, except that China had no fixed loci and Pakistan had only one monomorphic locus (RJN-12; supplementary files; Table S1).

Population subdivision

Genotypes clearly clustered according to their geographical origin in the analyses with the model-based clustering method implemented in STRUCTURE, with an optimal number of clusters (*K*) equal to 6, based on the rate of change in the log probability of data across successive *K* values (Evanno *et al.*, 2005). At $K = 2$, Middle Eastern, Mediterranean and Central Asian populations were assigned to one group; the Chinese population was assigned to the other group; and Nepalese, Pakistani and NW European populations had a mixed assignment of the two groups (Fig. S2). Increasing *K* to 3 individualized a Pakistan-specific group, while increasing *K* to 4 split the cluster of Middle East, Central Asia and Mediterranean region into two groups, one specific to the Middle East and East Africa and the other specific to the Central-Asia and Mediterranean region, with substantial admixture from the Middle East. The Middle Eastern and East African populations had no differentiation from each other and are termed as Middle East-Red Sea Area, onward. At $K = 5$, the NW European populations were separated from the Chinese population, and at $K = 6$, the Nepalese group individualized (Fig. S2). Increasing *K* above 6 did not reveal any further subdivisions. We confirmed that the presence of some of the clonal populations would not result in strong deviation from the STRUCTURE results, as the existence of six genetic groups was further supported by the

non-parametric DAPC analysis (Fig. 1 and Fig.2). The BIC curve in the DAPC analyses also supported K=6 with a clear discrimination of genotypes from China, Pakistan, Nepal, Middle East-Red Sea Area, NW Europe and Central Asia-Mediterranean region (Fig. 2).

Population differentiation among the different groups was estimated by means of pairwise F_{ST}. Populations showed a high differentiation, with strong and significant F_{ST} values for all pairs except for PST from the Middle Eastern, Central Asian and Mediterranean regions (Table 2), confirming a relatively recent shared ancestry or significant gene flow among these populations. Chinese, Pakistani and Nepalese populations were differentiated from one another and from the Middle Eastern and Mediterranean populations. These two latter populations were not highly differentiated from one another (Fig. S2; Table 2). The NW European population showed a strong differentiation from Mediterranean and Middle Eastern populations but was closer to the Chinese population (Fig. S2 and Fig. S3).

Table 2. Estimates of F_{ST} (upper diagonal) and its significance (lower diagonal) based on 20 microsetillite loci for 386 PST isolates representing worldwide geographically spaced populations. The lower two lines shows F_{ST} and its p-value for isolates representing the post 2000 emerged strains.

	NW Europe	**North America**	**South America**	**Mediterranean Region**	**Central Asia**	**South Africa**	**East Africa**	**Middle East**	**Nepal**	**Pakistan**	**China**
NW Europe	-	**0.039**	**0.001**	0.420	0.380	0.498	0.500	0.380	0.370	0.410	0.390
North America	**0.100**	-	**0.046**	0.409	0.368	0.485	0.490	0.378	0.364	0.400	0.398
South America	**0.410**	**0.100**	-	0.435	0.396	0.514	0.511	0.393	0.379	0.416	0.405
Mediterranean Region	0.000	0.000	0.000	-	0.020	**0.109**	0.150	**0.009**	0.280	0.280	0.390
Central Asia	0.000	0.000	0.000	0.000	-	**0.044**	0.160	0.040	0.230	0.260	0.340
South Africa	0.000	0.000	0.000	**0.010**	**0.190**	-	**0.229**	0.160	0.298	0.314	0.419
East Africa	0.000	0.000	0.000	0.000	0.000	**0.600**	-	0.140	0.380	0.280	0.540
Middle-East	0.000	0.000	0.000	**0.020**	0.000	0.000	0.000	-	0.260	0.250	0.360
Nepal	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-	0.220	0.210
Pakistan	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-	0.450
China	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-
F_{ST} for aggressive strains	0.420	0.408	0.434	**0.000**	**0.010**	**0.099**	0.150	**0.010**	0.270	0.270	0.390
P-value for aggressive strains	0.000	0.000	0.000	**0.860**	**0.070**	**0.040**	0.000	**0.020**	0.000	0.000	0.000

Non-significant F_{ST} values (> 0.01) are shown in bold.

Geographical patterns of genotypic variability

Populations from NW Europe, USA, South America, Australia, South Africa, Eritrea, the Middle East and the Mediterranean region displayed low genotypic diversity as well as an excess of heterozygosity compared to expectations under HWE, confirming their long-term clonality. Samples from Pakistan, Nepal and China did not depart from HWE, suggesting the occurrence of recombination within the populations (Fig. 3). Himalayan (Nepalese and Pakistani) and near Himalayan (Chinese) populations had a higher genotypic diversity, higher number of alleles and higher allele richness (Fig. 4) than the populations from Middle East-Red Sea area. The latter two were themselves more diverse than the European and Mediterranean populations, where the maximum clonal resampling was observed. Thus, Asian populations appeared as the zone of the highest diversity of the pathogen. A similar pattern was observed for private allele richness, with Pakistan possessing the highest number of private alleles (Fig. 4). Isolates representing NW Europe also had high private allele richness, probably due to their strict clonality (Enjalbert *et al.*, 2005; Justesen *et al.*, 2002) and isolation from other populations.

Source of recently emerged populations

We detected only a few recent migrants, admixed and unassigned isolates in each geographical region, in the clustering analyses (Fig. 2 and Fig. S2). Clear migration footprints were only found when focusing on recently colonized areas. Analyses confirmed NW Europe as the source of the North American and Australian populations, and the Mediterranean region and Central Asia appeared to be the source of the South African population (Fig. 2 and Fig. S2; non-significant F_{ST}, Table 2). Additionally, the South American isolates were assigned to NW European isolates and displayed very low diversity, revealing another incursion from NW Europe.

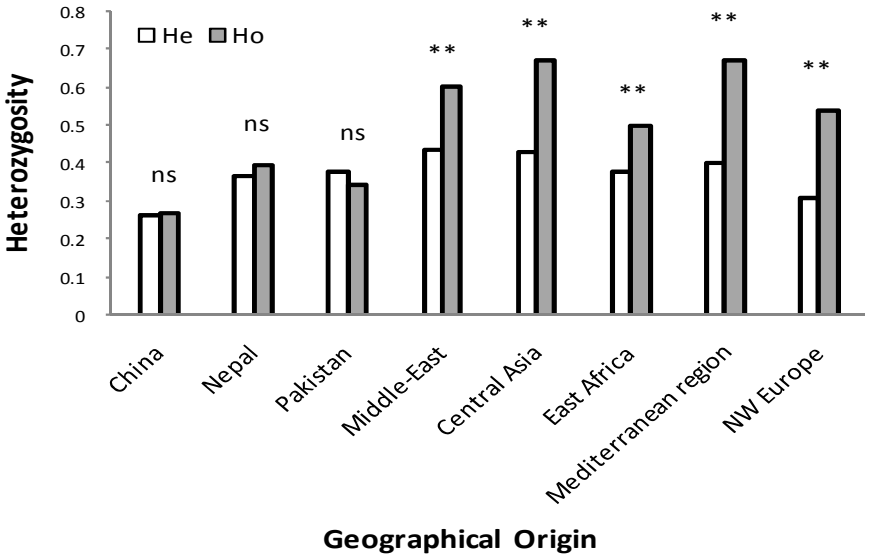


Fig. 3. Expected (He) and observed (Ho) heterozygosity for clone-corrected data based on 20 polymorphic microsatellite loci for PST isolates sampled from diverse geographical regions.

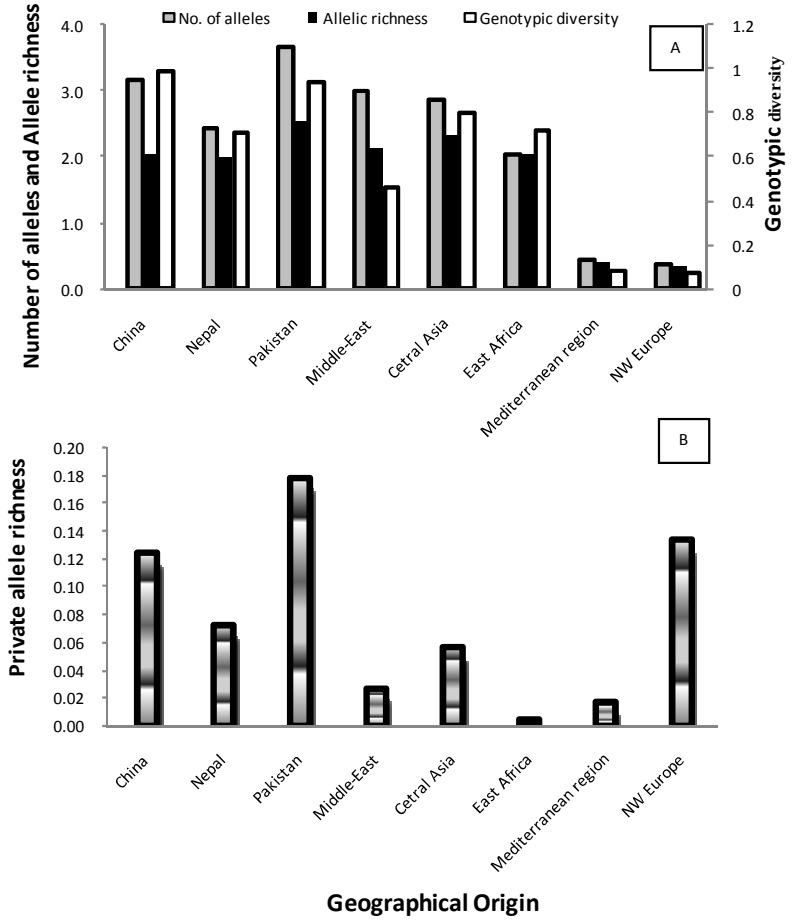


Fig. 4. Diversity index, allele richness (A) and private allele richness (B) for PST populations from worldwide geographical regions.

For both of the two recently emerged, aggressive strains (PstS1 and PstS2), PstS1 isolates associated with the post-2000 epidemics in the USA and Australia, consisted of only a single multilocus genotype (MLG-99) resampled in other geographical regions as well. PstS1 was closely related to the European strain PstS2, which consisted of different multilocus genotypes, including this MLG-99. Both PstS1 and PstS2 were assigned to the Middle Eastern-Red Sea area group, suggesting a source in Middle East-Red Sea area for these widely spread, aggressive strains (Fig. 2 and Table 2). An older set of aggressive isolates frequently reported in Europe (Enjalbert *et al.*, 2005), although less aggressive than PstS1 and PstS2 (Milus *et al.*, 2009), were assigned to the Central Asian-Mediterranean genetic group (represented as PstS3 in Fig. 2).

Ancestral relationship and migration patterns of populations

The results from the ABC analyses carried on "triplets" of different populations were combined to give a preliminary vision about the ancestral relationship among populations, summarised in Fig. 5 (detailed in Supporting information_on_ABC_analyses). The "triplet" results revealed that the subdivision between Pakistani and Chinese populations is the most ancestral split among the three Himalayan recombinant populations. Nepal was confirmed to be the result of an admixture between Pakistan and China. A similar pattern was observed for Middle-East, East Africa and Central Asia, when compared with Pakistan and China individually. Analyses also revealed that the Mediterranean population was an admixture between Middle East and Central Asia, whereas the Central Asian population itself was a divergence from Middle Eastern population. The comparison of NW European population in "triplet" with other populations revealed that this population resulted from an admixture from Pakistan and China. These results, however, need to be confirmed through the analyses of more complex models including more than three populations i) to confirm the above mentioned historical relationships and ii) to identify the relative positioning of different divergence and admixture events in time.

DISCUSSION

We investigated the origin, migration routes and population structure of wheat-infecting *Puccinia striiformis* f.sp. *tritici* (PST) using a comprehensive set of isolates from six continents and 20 highly informative SSR markers. We showed that despite the long-distance dispersal and the recent global spread of aggressive strains, the worldwide PST population has a clear genetic structure and is separated into six groups corresponding to main epidemic areas. The maintenance of a clear genetic structure despite substantial gene flow results from both strong clonality in many PST populations, and strong founder effects in recently emerged populations.

Strong population subdivision despite long-distance dispersal

We report the existence of a strong population subdivision within PST, with a clustering of isolates according to their geographical origin despite the capacity for long-distance migration of the fungus (Brown and Hovmøller, 2002; Hovmøller et al., 2011). This pattern stands in stark contrast with the previous understanding of the worldwide population structure of PST, which considers the potential replacement of the local population by new invasions (Brown and Hovmøller, 2002; Singh et al., 2004). On the basis of pathotype surveys that monitor the occurrence of strains with newly acquired virulences that defeat recently deployed *Yr* (resistance) genes, the population structure of PST was considered to be shaped by a continual replacement of pre-existing populations by the newly emerged and spread pathotypes, or aggressive strains. This process is well known as the boom and bust cycle (de Vallavieille-Pope *et al.*, 2012; Singh *et al.*, 2004). However, such surveys were designed to track the spread of a new, virulent race and, therefore, were potentially biased due to sampling only from varieties with the defeated resistance gene(s) in question but not from local landraces or other varieties. These observations lead to Asia (except China) being considered as a single epidemiological zone, with rapid and recurrent spread of new virulences over the whole zone, as in the case of virulence matching the *Yr9* resistance gene (Fig. 6 (Singh *et al.*, 2004)) and the recent virulence matching the stem rust resistance gene *Sr31* (Singh *et al.*, 2006). Indeed, such geographic migrations

are also documented in our study, but recently spread genotypes appear to coexist with and are dominated by older populations specific to the main geographic areas, suggesting that migrants do not replace the local populations in recombinant population in Asia despite the capacity for recurrent and long-distance dispersal. In contrast, the invasion of new genotypes in clonal populations would result in a population sweep and would replace the original population. This was observed in the USA, where the post-2000 PST population is dominated by the pathotypes characteristic of the aggressive strain, PstS1, or its derivatives, shown above to have originated from the Red Sea-Middle Eastern region. These two patterns might emerge depending on the importance of *i)* the better adaptation of local populations, more competitive in their region of origin; and *ii)* the selective advantage of migrants, e.g. virulence to new widely deployed resistance genes.

Regional differences in levels of recombination

PST has long been considered a strictly asexual pathogen on wheat due to the lack of knowledge of the alternate host for sexual reproduction at the end of the crop season (Cummins, 1971; Stubbs, 1985). Population genetic surveys that revealed clonal populations with very low diversity in USA (Chen *et al.*, 1993), Europe (Enjalbert *et al.*, 2005) and Australia (Wellings, 2007) were consistent with this hypothesis. Recently, populations with higher diversity were reported in the Middle East (Bahri *et al.*, 2009) and Pakistan (Bahri *et al.*, 2011), and a recombinant population structure was found in China (Mboup *et al.*, 2009). Herein, we identified a recombinant population structure and high diversity in Nepal and Pakistan and confirmed previous findings in China, suggesting the existence of possible sexual reproduction in PST populations from a broad area ranging from the Himalayan region to the Mongolian plateau. This possibility also recently gained indirect experimental support, with *Berberis spp.* being shown to serve as an alternate host for PST in laboratory conditions by Jin et al. (2010) and a high sex ability (in the form of telial production) reported in the Asian populations of PST (Ali *et al.*, 2010). Although the role of *Berberis spp.* for the life cycle of PST under natural conditions remains to be further investigated, the presence of *Berberis spp.* in Pakistan, Nepal and China (Perveen and

Qaiser, 2010; Ray et al., 2011) are consistent with the possible existence of a sexual cycle of the pathogen in Asia.

On the source of new incursions and emergences

Assignment analyses allowed us to identify the source of new incursions and emergence (Fig. 6). The source of the Australian and North American populations was confirmed to be NW Europe, in accordance with previous findings, suggesting the migration of PST from NW Europe to Australia in 1979 (Wellings, 2007) and probably earlier to North America (Carleton, 1915; Hovmøller *et al.*, 2011). We also identified the NW European source of the South American PST population, which was reported earlier in the 20th century with no inference on its source (Rudorf and Job, 1934; Stubbs, 1985). This suggests that the PST incursion into both North and South America was made from NW Europe, probably through human intervention. We also identified the Mediterranean-Central Asian population as the source of South African populations, first reported in 1996 (Boshoff *et al.*, 2002), which might have resulted from wind dispersal or human intervention (Fig. 6).

Two closely related strains of PST, distinct from local populations, were recently reported in North America, Australia and Europe (Hovmøller *et al.*, 2008). These strains were shown to be highly aggressive and adapted to high temperature (Milus *et al.*, 2009). One of the two strains (PstS1) was responsible for PST epidemics in south-central USA, a region previously considered too warm for yellow rust epidemics (in 2000; Chen, 2005; Milus *et al.*, 2006), and in Western Australia (in 2002; Wellings *et al.*, 2003). Another strain (PstS2), closely related to the first one, was reported in NW Europe with similar aggressiveness and strong differentiation from local PST populations. Hovmøller et al. (2008) reported the existence of the two strains in the Mediterranean region and the Middle East-Red Sea area. Our analyses revealed that PstS1 representative isolates had a single multilocus genotype (MLG-99), while PstS2 consisted of different, but closely related, MLGs. Assignment analyses revealed that both strains originated from the Middle East-Red Sea Area, and such global

patterns of dispersal certainly involve accidental spore transport by human activities. The incursion of PST into the Americas and the spread of aggressive strains are most probably the direct consequence of human-associated dispersal, as suggested already for the initial introduction of PST in Australia (Wellings, 2007). These results reflect the importance of human intervention in the emergence of PST epidemics.

Origin of the worldwide PST populations

Transcaucasia had previously been suggested as the centre of origin for PST, mainly based on its diversity of virulence and distribution of pathotypes (Stubbs, 1985; Stubbs, 1988). However, the diversity of virulence and the distribution of pathotypes are strongly influenced by the resistance in host populations as well, and this might lead to biased inferences of the location of the centre of diversity of the pathogen. In our analyses, the representative isolates from Transcaucasia are less diverse, clonal and do not exhibit strong divergence from the rest of the Oriental populations. In contrast, the existence of high levels of diversity, private alleles, a recombinant population structure, the ability to produce sex-related structures (Ali *et al.*, 2010) and the independent maintenance of PST populations in the Himalayan region identify the Himalayan region as a more plausible centre of origin for PST. If one considers that the centre of origin of PST is in the Himalayan region, then PST would have adapted to wheat through a host shift and not following host-tracking co-evolution with early wheat domesticates in the Fertile Crescent. This would add to the increasingly adopted view that host-shift speciation is a major route of disease emergence in fungal pathogens (Giraud *et al.*, 2010).

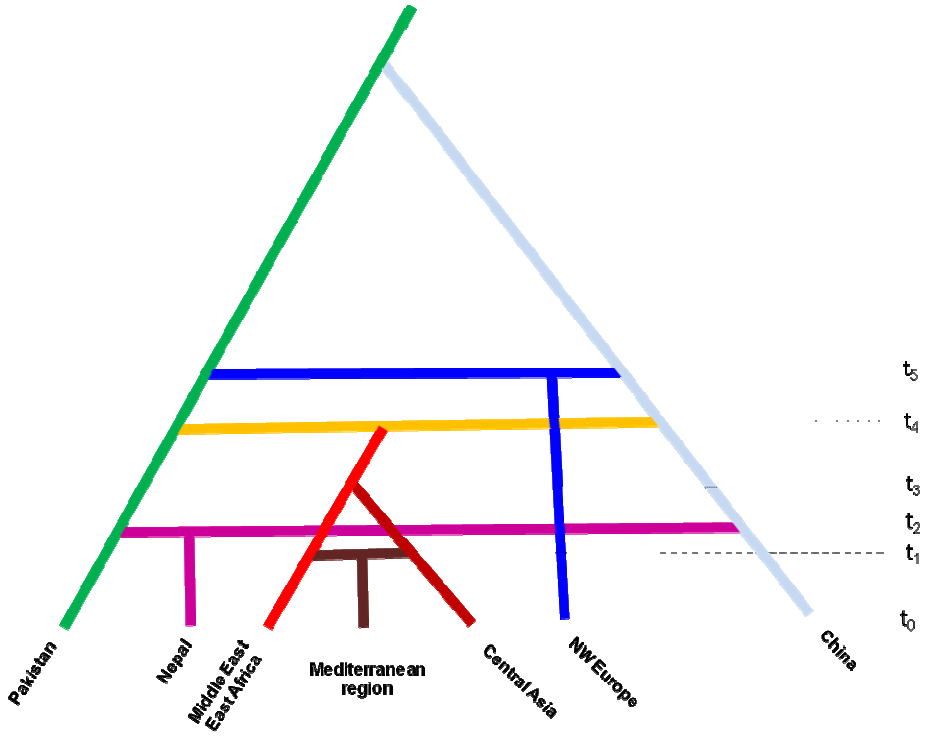


Fig. 5. Ancestral relationship among worldwide PST populations as inferred from the preliminary analyses of approximate Bayesian computations. The relative contribution of Pakistan and China (and Middle East and Central Asia in case of Mediterranean population) and the relative positioning in time for the admixture and divergence events need to be confirmed.

Proposed scenario for historical migration routes of PST

Once adapted to wheat in its centre of origin in the Himalayan region, PST would have spread to the rest of the world while evolving independently in different parts of the Himalayan region, resulting in population subdivision within the native area. The ABC analyses confirmed the Himalayan populations to be the ancestral populations and the differentiation between Pakistan and China to be the most ancestral split in this region (Fig. 5). PST would have spread northward from the Himalayas to the Mongolian plateau in China, where it maintained sexual reproduction and a high diversity in some parts, with an acquisition of virulences to the Chinese wheat population. The Eastward admixture of Pakistan and China could have resulted in the Nepali population. On the westward side of the Himalayan region, the populations of Middle East, Central Asia East Africa and Mediterranean region seemed to be the result of an admixture between Pakistan and China (Fig. 5). These populations, termed together as CWANA (Central and West Asia and North Africa), could have an

independent evolution and a high gene flow within Eastern Africa, the Middle East, Central Asia and the Mediterranean regions. The Mediterranean, Middle Eastern, East African and Central Asian populations diverged less, despite the fact that they cover a large geographical area. The off-season maintenance in some regions, its subsequent spread to CWANA regions and the lack of local off-season survival through volunteers or sexual reproduction in all regions could result in a source and sink relationship at the scale of the whole CWANA region, as previously suggested (Bahri *et al.*, 2009). The most recent incursion from this CWANA population was into South Africa, where PST was absent before 1996 (Boshoff *et al.*, 2002). The lesser divergence between the NW European population and the Chinese population, compared to the CWANA population, suggests that PST would have spread from China to NW Europe, probably through some human intervention rather than an airborne incursion from the Middle Eastern, Central Asian and/or Mediterranean regions. The preliminary ABC analyses revealed NW Europe to be a result of admixture between Pakistan and China, however, the more detailed analyses needs to be carried out for its confirmation. This NW European population succeeded in terms of off-season survival on volunteers in coastal areas and resulted in a reduced sex ability (Ali *et al.*, 2010). The clonal evolution within the NW European population resulted in a strong negative F_{IS} value, owing to a clonal divergence of the dikaryotic genomes (Enjalbert *et al.*, 2005), in line with what is expected according to the Meselson Effect (Bultin, 2002). From NW Europe, PST would have then been recently introduced to North and South America and even more recently to Australia (Fig. 6).

Although this study points out the Himalayan region as the likely centre of origin of PST and confirms the previous migration hypotheses (invasion of USA and Australia from NW Europe and the source of aggressive strains), we also proposed a scenario for worldwide PST migrations routes raising new hypotheses like the migration to South America or the source of the European population (Fig. 6). More extended surveys are necessary to provide an accurate view of the diversity and structure of PST in Asia and confirm some of the hypotheses raised above, taking advantage of the historical/epidemiological records of PST emergences.

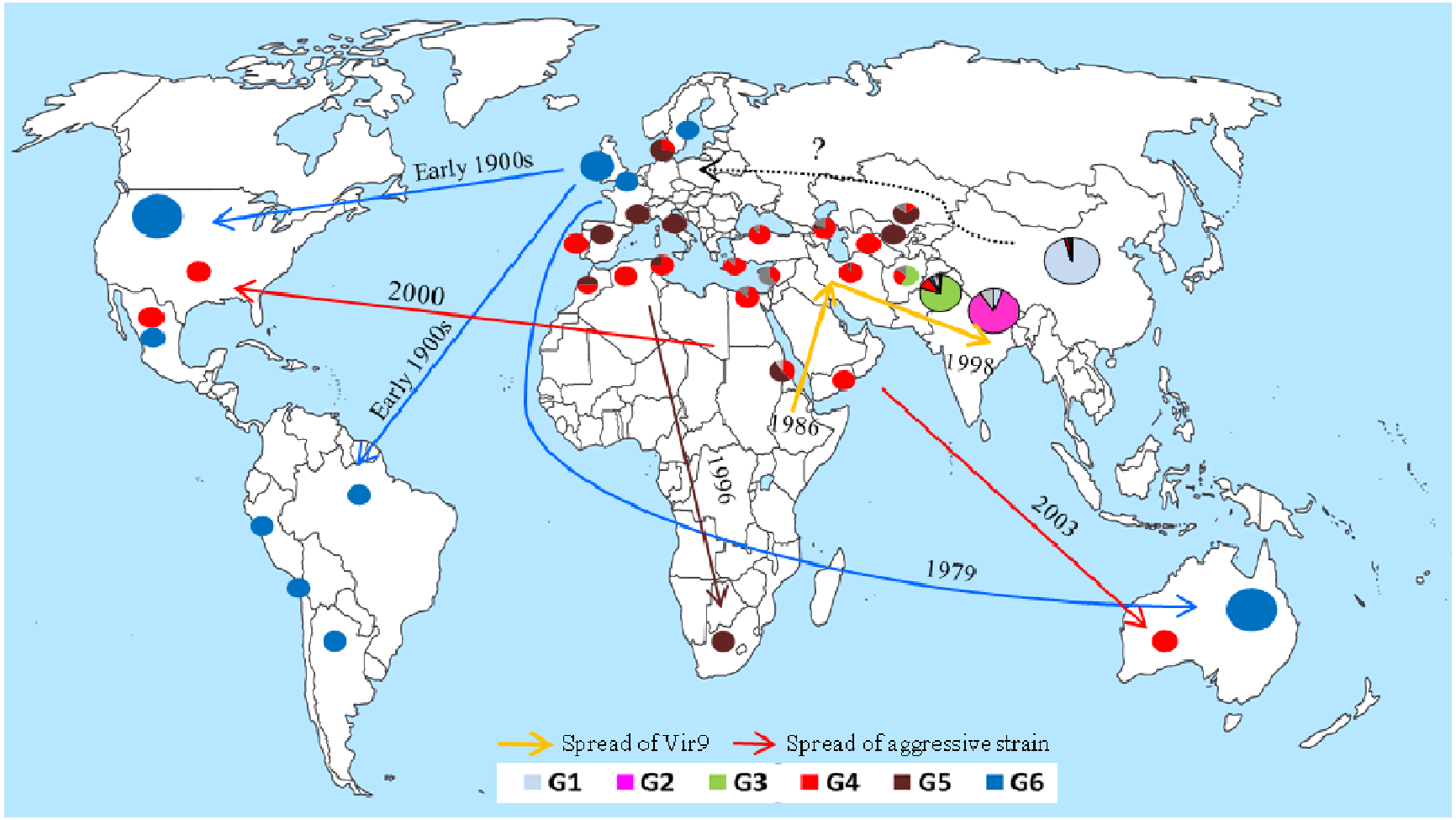


Fig. 6. Origin and migration routes of recently emerged populations of wheat yellow rust pathogen identified or confirmed through the population genetic analyses of a worldwide representative set of isolates. The years shows the first report of incidence based on present work and previous work (Aggressive strain = Hovmøller et al., 2011; South African population = Boshoff et al., 2002; Spread of Vir9 = Singh et al., 2004; incursion into Australia = Wellings and McIntosh, 1990;).

Conclusions

The existence of a high genotypic diversity, a high sex ability as well as the independent maintenance of strongly differentiated populations in the Himalayan region pinpoint this region as the possible centre of origin of PST. Differences in the levels of diversity and mode of reproduction among geographically distant populations are particularly relevant in the context of risk-assessment for disease emergence: Asian populations (China, Nepal and Pakistan) with a high level of recombination, diversity and sex ability could serve as possible sources for the emergence of new, virulent and aggressive strains. The maintenance of populations specific to geographical regions in Asia suggests a survival of local populations in these regions, potentially through sexual reproduction. For integrated disease management, it would be important to quantify the relative contribution of sexual vs. asexual reproduction to the diversity in different populations around the world and identify sexual host(s) or clonal over-summering/-wintering pathways. Finally, this study emphasizes how human inter-continental travel has been a major driver in the emergence of PST. According to the evolution of the political context, the intensification of business and tourism activities in regions known as major sources of pathogen diversity should be considered an increased risk for re-emergence of PST worldwide.

AUTHOR CONTRIBUTIONS

SA carried out the molecular genotyping and population genetic analyses. AG and CP supervised the molecular genotyping. JE and PG supervised population genetic analyses. MSH, AJ, ML, JE and CP provided isolates (spores or DNA) for genotyping. SA, PG, JE and MSH prepared the manuscript. AJ and CP revised the manuscript. JE, MSH and CP conceived and designed the study. All of the authors have read and approved the manuscript.

ACKNOWLEDGEMENTS

The authors are grateful to all their colleagues who participated in establishing the yellow rust collection used for this study: W. Chen and X. Duan (China); B. Bahri and A. Ouffroukh (Tunisia); M. Mboup (Senegal); S. Lohani (Nepal); and S. Hussain and S.J.A Shah (Pakistan), to name a few. We would like to acknowledge Stephanie Walter for her constructive comments, which improved the manuscript. We also thank Laurent Gérard and Nathalie Galet for their technical assistance during the multiplication of yellow rust isolates. This work received financial support from the European Integrated Project BIOEXPLOIT, FOOD-CT-2005-513959 and a French project EMERFUNDIS, ANR 07-BDIV-003. Sajid Ali was financially supported by a grant from the Higher Education Commission, Pakistan.

REFERENCES

Ali S, Gautier A, Leconte M, Enjalbert J, De Vallavieille-Pope C (2011) A rapid genotyping method for an obligate fungal pathogen, *Puccinia striiformis* f.sp. *tritici*, based on DNA extraction from infected leaf and Multiplex PCR genotyping. *BMC Research Notes* **4**, 240.

Ali S, Leconte M, Walker A-S, Enjalbert J, de Vallavieille-Pope C (2010) Reduction in the sex ability of worldwide clonal populations of *Puccinia striiformis* f.sp. *tritici*. *Fungal Genetics and Biology* **47**, 828-838.

Ali S, Shah SJA, Khalil IH, Rahman H, Maqbool K, Ullah W (2009a) Partial resistance to yellow rust in introduced winter wheat germplasm at the north of Pakistan. *Australian Journal of Crop Science* **3**, 37-43.

Ali S, Shah SJA, Rahman H, Saqib MS, Ibrahim M, Sajjad M (2009b) Variability in wheat yield under yellow rust pressure in Pakistan. *Turkish Journal of Agriculture and Forestry* **33**, 537-546.

Arnaud-Haond S, Belkhir K (2007) GenClone 1.0: A new program to analyse genetics data on clonal organisms. *Molecular Ecology Notes* **7**, 15-17.

Bahri B, Leconte M, Ouffroukh A, de Vallavieille-Pope C, Enjalbert J (2009) Geographic limits of a clonal population of wheat yellow rust in the Mediterranean region. *Molecular Ecology* **18**, 4165-4179.

Bahri B, Shah SJA, Hussain S, Leconte M, Enjalbert J, de Vallavieille-Pope C (2011) Genetic diversity of wheat yellow rust population in Pakistan and its relationship with host resistance. *Plant Pathology* **60**, 649-660.

Beaumont MA, Zhang W, Balding DJ (2002) Approximate Bayesian computation in population genetics. *Genetics* **162**, 2025-2035.

Belkhir KBP, Chikhi L, Raufaste N, Bonhomme F (2004) GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5000: Université de Montpellier II, Montpellier (France) http://www.univ-montp2.fr/~genetix/genetix/genetix.htm.

Boshoff WHP, Pretorius ZA, van Niekerk BD (2002) Establishment, distribution, and pathogenicity of *Puccinia striiformis* f.sp. *tritici* in South Africa. *Plant Disease* **86**, 485-492.

Brown JKM, Hovmøller MS (2002) Aerial dispersal of fungi on the global and continental scales and its consequences for plant disease. *Science* **297**, 537-541.

Bultin R (2002) The costs and benefits of sex: new insights from old asexual lineages. *Nature Reviews Genetics* **3**, 313-317.

Campbell F (2001) The science of risk assessment for phytosanitary regulation and the impact of changing trade regulations. *BioScience* **51**, 148-153.

Carleton MA (1915) A serious new wheat rust in this country. *Science* **42**, 58-59.

Chen X, Line RF, Leung H (1993) Relationship between virulence variation and DNA polymorphism in *Puccinia striiformis*. *Phytopathology* **83**, 1489-1497.

Chen XM (2005) Epidemiology and control of stripe rust [*Puccinia striiformis f. sp. tritici*] on wheat. *Canadian Journal of Plant Pathology* **27**, 314–337.

Cornille A, Gladieux P, Smulders MJM, Roldan-Ruiz I, Laurens F, Le Cam B, Nersesyan A, Clavel J, Olonova M, Feugey L, Gabrielyan I, Zhang X-G, Tenaillon MI, Giraud T (2012) New Insight into the History of Domesticated Apple: Secondary Contribution of the European Wild Apple to the Genome of Cultivated Varieties. *PLoS Genetics* **8**, e1002703. doi:1002710.1001371/journal.pgen.1002703.

Cornuet JM, Ravigne V, Estoup A (2010) Inference on population history and model checking using DNA sequence and microsatellite data with the software DIYABC (v1.0). *BMC Bioinformatics* **11**, 401.

Cornuet JM, Santos F, Beaumont MA, Robert CP, Marin JM, al. e (2008) Inferring population history with DIY ABC: a user-friendly approach to approximate Bayesian computation. *Bioinformatics* **24**, 2713-2719.

Csilléry K, François O, Blum M (2011) abc: an R package for approximate Bayesian computation (ABC). *Methods in Ecology and Evolution* **3**, 475–479.

Cummins GB (1971) *The rusts fungi of cereals, grasses, and bamboos* Springer-Verlag, New York.

de Vallavieille-Pope C, Ali S, Leconte M, Enjalbert J, Delos M, Rouzet J (2012) Virulence dynamics and regional structuring of *Puccinia striiformis* f. sp. *tritici* in France between 1984 and 2009. *Plant Disease* **96**, 131-140.

Dilmaghani A, Gladieux P, Gout L, Giraud T, Brunner PC, Grandaubert J, Stachowiak A, Balesdent MH, Rouxel T (2012) Migration patterns and changes in population biology associated with the worldwide spread of the oilseed rape pathogen *Leptosphaeria maculans*. *Molecular Ecology* **doi: 10.1111/j.1365-294X.2012.05535.x**.

Enjalbert J, Duan X, Giraud T, Vautrin C, de Vallavieille-Pope C, Solignac M (2002) Isolation of twelve microsatellite loci, using an enrichment protocol, in the phytopathogenic fungus *Puccinia striiformis* f. sp. *tritici*. *Molecular Ecology Notes* **2**, 563-565.

Enjalbert J, Duan X, Leconte M, Hovmøller MS, de Vallavieille-Pope C (2005) Genetic evidence of local adaptation of wheat yellow rust (*Puccinia striiformis* f. sp. *tritici*) within France. *Molecular Ecology* **14**, 2065-2073.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE : a simulation study. *Molecular Ecology* **14**, 2611-2620.

Fagundes NJR, Ray N, Beaumont MA, Neuenschwander S, Salzano FM, Bonatto SL, Excoffier L (2007) Statistical evaluation of alternative models of human evolution. *Proceedings of the National Academy of Science of the United States of America* **104**, 17614-17619.

Gao H, Williamson S, Bustamante CD (2007) A markov chain monte carlo approach for joint inference of population structure and inbreeding rates from multilocus genotype data. *Genetics* **176**, 1635-1651.

Giraud T, Gladieux P, Gavrilets S (2010) Linking the emergence of fungal plant diseases with ecological speciation. *Trends in Ecology & Evolution* **25**, 387-395.

Goldstein DB, Ruiz LA, Cavalli-Sforza LL, Feldman MW (1995) An evaluation of genetic distances for use with microsatellite loci. *Genetics* **139**, 463-471.

Goudet J (2001) FSTAT, a program to estimate and test gene diversities and fixation indices. Available from http://www2.unil.ch/popgen/ softwares/fstat.htm.

Hovmøller MS, Justesen AF, Brown JKM (2002) Clonality and long-distance migration of *Puccinia striiformis* f.sp. *tritici* in north-west Europe. *Plant Pathology* **51**, 24-32.

Hovmøller MS, Sørensen CK, Walter S, Justesen AF (2011) Diversity of *Puccinia striiformis* on cereals and grasses. *Annual Review of Phytopathology* **49**, 197-217.

Hovmøller MS, Walter S, Justesen AF (2010) Escalating threat of wheat rusts. *Science* **329**, 369.

Hovmøller MS, Yahyaoui AH, Milus EA, Justesen AF (2008) Rapid global spread of two aggressive strains of a wheat rust fungus. *Molecular Ecology* **17**, 3818-3826.

Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* **23**, 1801-1806.

Jin Y, Szabo LJ, Carson M (2010) Century-old mystery of Puccinia striiformis life history solved with the identification of Berberis as an Alternate Host. *Phytopathology* **100**, 432-435.

Jombart T, Devillard S, Balloux F (2010) Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics* **11**, 95 http://www.biomedcentral.com/1471-2156/1411/1494.

Justesen AF, Ridoutb CJ, Hovmøller MS (2002) The recent history of *Puccinia striiformis* f.sp. *tritici* in Denmark as revealed by disease incidence and AFLP markers. *Plant Pathology* **51**, 13-23.

Kolmer JA (2005) Tracking wheat rust on a continental scale. *Current Opinion in Plant Biology* **8**, 441-449.

Langella O (2008) POPULATIONS 1.2.3, Logiciel de génétique des populations (ed. Langella O). Laboratoire Populations, Génétique et évolution, CNRS, Gif-sur-Yvette, France.

Markell SG, Milus EA (2008) Emergence of a Novel Population of *Puccinia striiformis* f. sp. *tritici* in Eastern United States. *Phytopathology* **98**, 632-639.

Maynard-Smith J, Smith NH, O'rourke M, Spratt BG (1993) How clonal are bacteria? **90**, 4384-4388.

Mboup M, Leconte M, Gautier A, Wan AM, Chen W, de Vallavieille-Pope C, Enjalbert J (2009) Evidence of genetic recombination in wheat yellow rust populations of a Chinese oversummering area. *Fungal Genetics and Biology* **46**, 299-307.

Milus EA, Kristensen K, Hovmøller MS (2009) Evidence for increased aggressiveness in a recent widespread strain of *Puccinia striiformis* f.sp. *tritici* causing stripe rust of wheat. *Phytopathology* **99**, 89-94.

Milus EA, Seyran E, McNew R (2006) Aggressiveness of *Puccinia striiformis* f. sp. *tritici* isolates in the South-Central United States. *Plant Disease* **90**, 847-852.

Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* **89**, 583-590.

Nei M, Tajima F, Tateno Y (1983) Accuracy of estimated phylogenetic trees from molecular data. *Journal of Molecular Evolution* **19**, 153-170.

Palm M (2001) Systematics and the impact of invasive fungi on agriculture in theUnited States. *BioScience* **51**, 141-147.

Parker IM, Gilbert GS (2004) The evolutionary ecology of novel plant-pathogen interactions. *Annual Review of Ecology, Evolution, and Systematics* **35**, 675-700.

Perrings C, Williamson M, Barbier E, Delfino D, Dalmazzone S, J. S, Simmons P, Watkinson A (2002) Biological invasion risks and the public good: an economic perspective. *Conservation Ecology* **6**, Article 1 (online).

Perveen A, Qaiser M (2010) Pollen flora of Pakistan–LXV. berberidaceae. *Pakistan Journal of Botany* **42**, 1-6.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* **155**, 945-959.

Ray R, Gururaja KV, Ramchandra TV (2011) Predictive distribution modeling for rare Himalayan medicinal plant *Berberis aristata* DC. *Journal of Environmental Biology* **32**, 725-730.

Raymond M, Rousset F (1995) GENEPOP Version 1.2: population genetics software for exact tests and ecumenicism. *Journal of Heredity* **86**, 248-249.

Rudorf W, Job M (1934) Untersuchungen bezüglich des Spezialisierung von *Puccinia graminis tritici. Puccinia trilicina* und *Puccinia glumarum tritici*, sowie über Resistenz und ihre Vererbung in verschiedenen Kreuzungen. *Pflanzenzüchtung* **19**, 333-365.

Singh RP, Hodson DP, Jin Y, Huerta-Espino J, Kinyua MG, Wanyera R, Njau P, Ward WR (2006) Current status, likely migration and strategies to mitigate the threat to wheat production from race Ug99 (TTKS) of stem rust pathogen. *CAB Reviews* **1**, 054.

Singh RP, William HM, Huerta-Espino J, Rosewarne G (2004) Wheat rust in Asia: meeting the challenges with old and new technologies.

Stubbs RW (1985) Stripe rust. In: *The Cereal Rusts, Diseases, Distribution, Epidemiology and Control* (eds. Roelfs AP, Bushnell WR), pp. 61–101. Academic Press, London, UK.

Stubbs RW (1988) Pathogenicity analysis of yellow (stripe) rust of wheat and its significance in a global context. In: *Breeding strategies for resistance to the rusts of wheat* (eds. Simmonds NW, Rajaram S). CIMMYT D.F. Mexico.

Stukenbrock EH, McDonald BA (2008) The origins of plant pathogens in agro-ecosystems. *Annual Review of Phytopathology* **46**, 75-100.

Szpiech ZA, Jakobsson M, Rosenberg NA (2008) ADZE: a rarefaction approach for counting alleles private to combinations of populations. *Bioinformatics* **24**, 2498–2504.

Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution Psychiatrique* **38**, 1358-1370.

Wellings CR (2007) *Puccinia striiformis* in Australia: A review of the incursion, evolution and adaptation of stripe rust in the period 1979-2006. *Australian Journal of Agricultural Research* **58**, 567-575.

Wellings CR, Wright DG, Keiper F, Loughman R (2003) First detection of wheat stripe rust in Western Australia: evidence for a foreign incursion. *Australasian Plant Pathology* **32**, 321-322.

Table S1. Gene diversity for 20 SSR loci in PST populations of diverse geographical origin.

SSR Locus	NW Europe N America **	Mediterranean region	Central Asia	East Africa	Middle-East	Nepal	Pakistan	China
No. of isolates	**29**	**21**	**25**	**24**	**67**	**54**	**68**	**70**
No. of MLGs	**3**	**2**	**8**	**4**	**12**	**13**	**25**	**51**
RJN5	0.50	0.50	0.54	0.68	0.52	0.51	0.22	0.15
RJN6	0.31	0.40	0.41	**0.00***	0.05	0.46	0.19	0.34
RJN13	**0.00***	0.50	0.50	0.47	0.49	**0.00***	0.09	0.17
RJN3	**0.00***	**0.00***	0.35	0.33	0.27	0.11	0.75	0.30
RJO21	**0.00***	0.65	0.62	0.63	0.62	0.44	0.64	0.03
RJN4	0.09	0.50	0.50	0.47	0.50	0.47	0.40	0.27
RJN8	**0.00***	0.50	0.50	0.47	0.49	**0.00***	0.09	0.14
RJO18	0.50	0.50	0.54	0.47	0.51	0.55	0.22	0.47
RJN12	**0.00***	**0.00***	**0.00***	**0.00***	**0.00***	**0.00***	**0.00***	0.02
RJN10	0.50	0.40	0.43	0.47	0.55	0.50	0.26	0.22
RJO24	0.50	0.78	0.77	0.67	0.73	0.68	0.69	0.70
RJO4	0.50	0.65	0.64	0.72	0.65	0.67	0.59	0.06
RJO20	**0.00***	0.17	0.08	**0.00***	0.22	0.56	0.58	0.41
RJN11	0.45	0.50	0.61	0.47	0.77	0.71	0.58	0.63
RJN9	0.50	**0.00***	**0.00***	**0.00***	0.05	**0.00***	0.06	0.02
RJN-2	0.80	0.83	0.79	0.68	0.85	0.59	0.83	0.64
WU-6	**0.00***	**0.00***	**0.00***	**0.00***	**0.00***	**0.00***	0.03	0.17
RJO-3	0.50	0.50	0.50	0.47	0.50	0.35	0.14	0.45
WU-12	0.50	0.47	0.48	0.47	0.44	0.32	0.63	0.04
RJO-27	0.63	0.50	0.58	0.72	0.60	0.59	0.74	0.06
Mean	0.31	0.42	0.44	0.41	0.44	0.38	0.39	0.26

* One allele fixed, i.e., no diversity (0.00)

** North American isolates prior to 2000

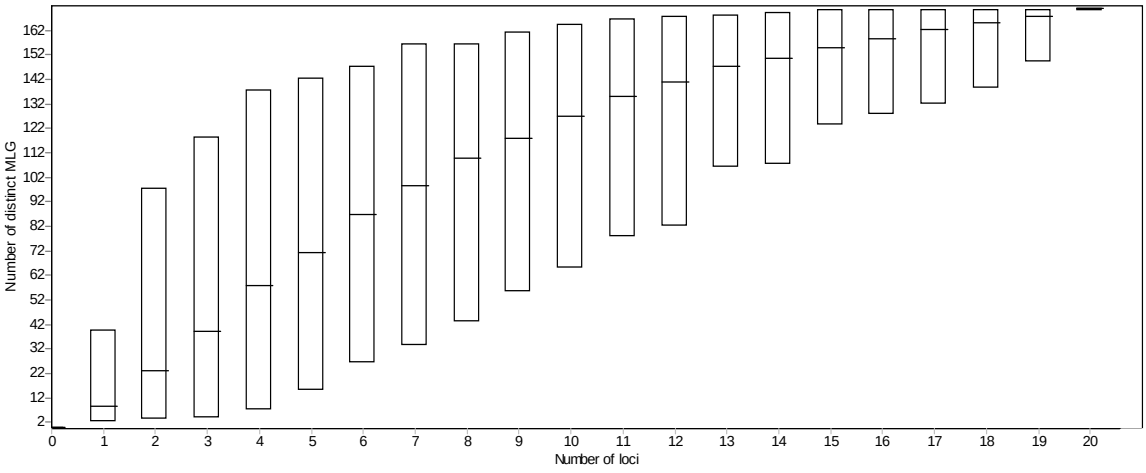


Fig S1. Box-plot of number of PST genotypes detected as a function of the number of loci re-sampled 1000 times within the 20 microsatellite markers (from 0 -20) using GENECLONE software. The box represents the average, minimum and maximum numbers of MLGs detected when re-sampling on loci.

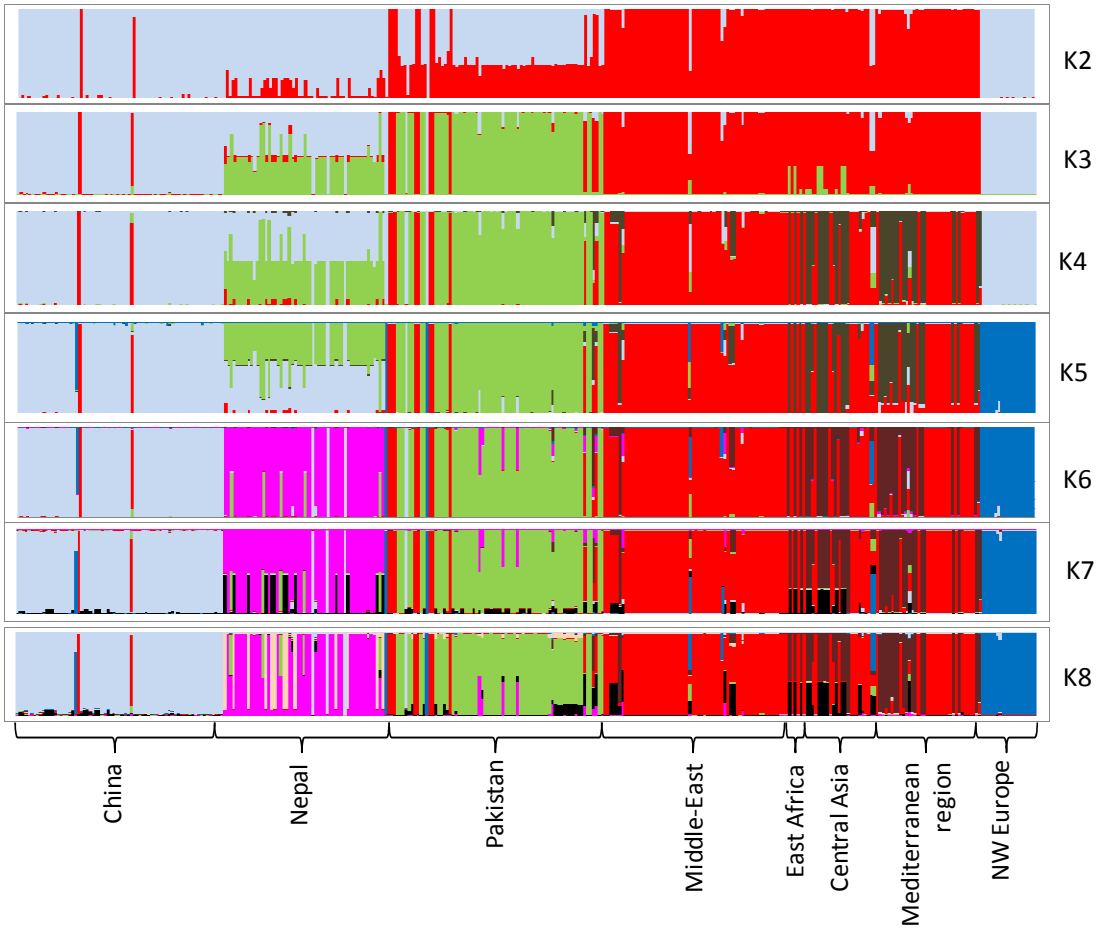


Fig. S2. Assignment of PST isolates from worldwide geographical regions to genetic groups, using the STRUCTURE software and different K-values (genetic groups). The chart represents the consensus assignment obtained by analysis of result from 20 runs of the STRUCTURE analysis with the CLUMPP software. Each colour represents a different genetic group.

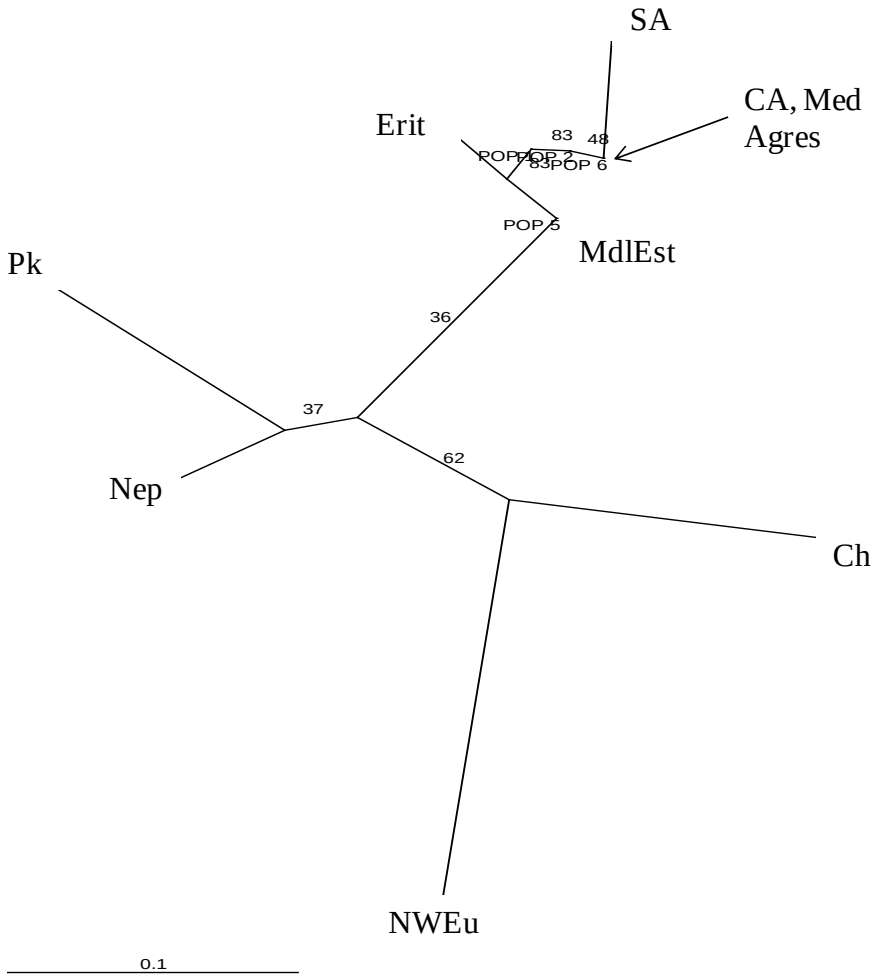
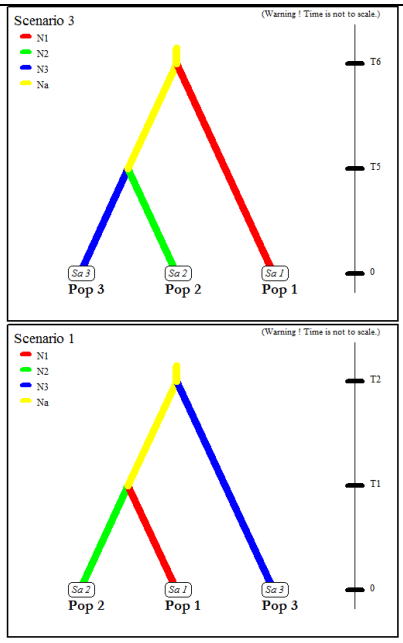


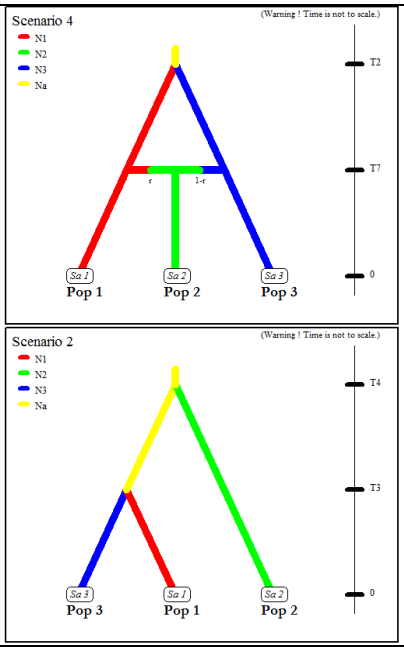
Fig. S3. Microsatellite distance-based Neighbor-Joining tree for PST isolates from worldwide geographically spaced populations. Agr = Aggressive strain, CA = Central Asia, Ch = China, Erit = East Africa, MdlEst = Middle East, Med = Mediterranean region, Nep = Nepal, NWEu = NW European group, Pk = Pakistan and SA = South Africa.

SUPPORTING INFORMATION ON ABC

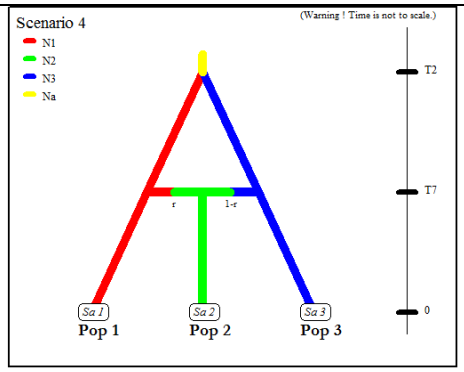
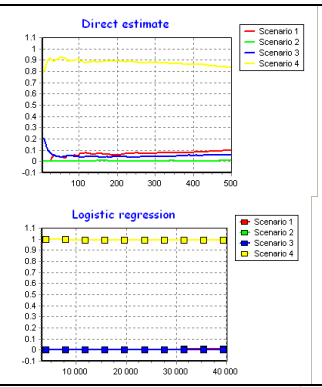
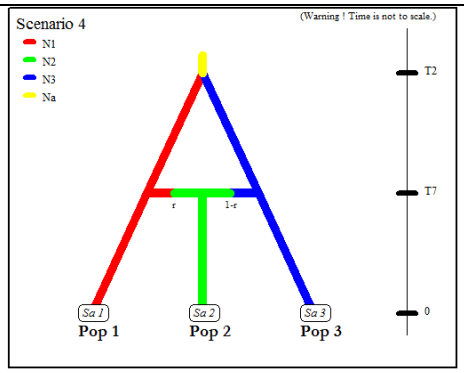
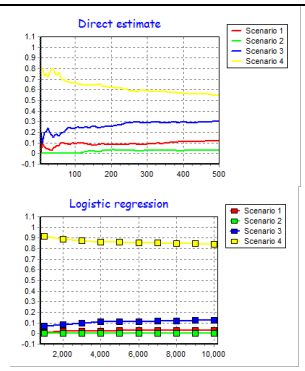
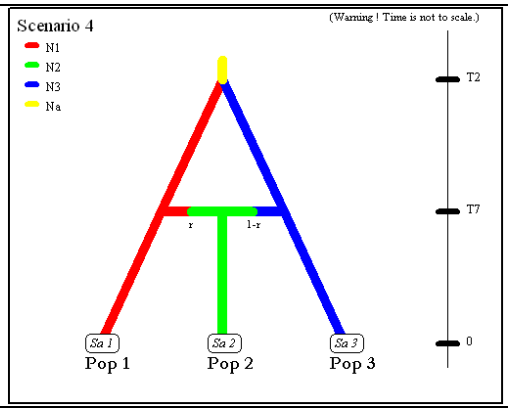
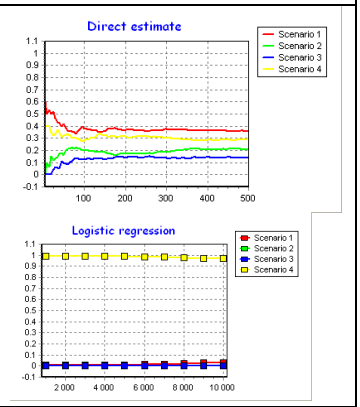
The overall results can be summarized as:
- Pakistan and China are ancestral to all populations
- Nepal is a recent admixture between Pakistan and China, in which Pakistan is ancestral to both China and Nepal
- Then the Middle-East population was an admixture between China and Pakistan
- The Mediterranean population resulted from an admixture from Middle East and Central Asia, which itself diverged from Middle East
- NW Europe would be an admixture between China and Pakistan

The four scenarios tested for all combinations are

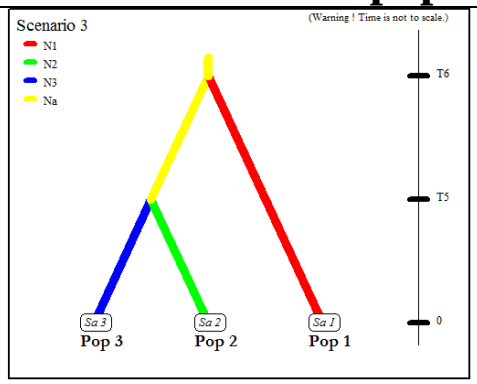
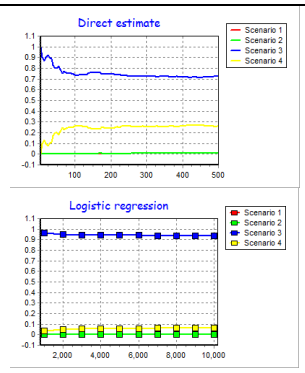
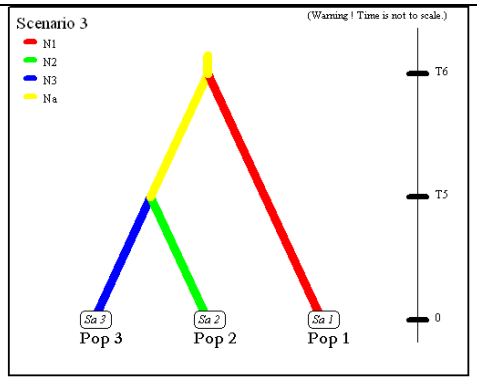
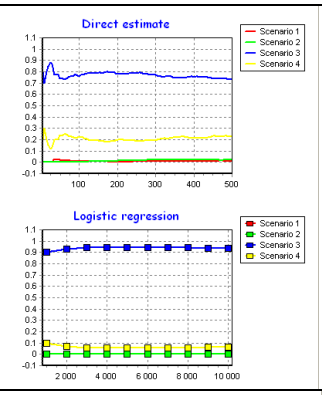
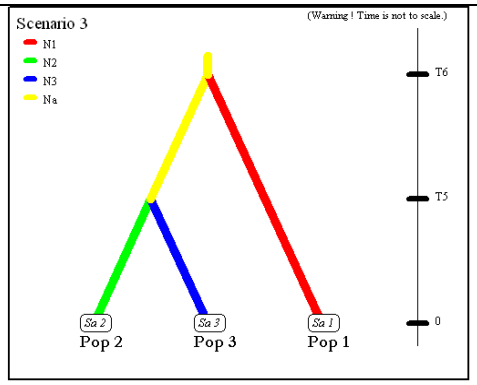
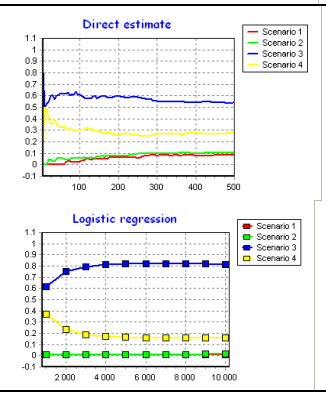
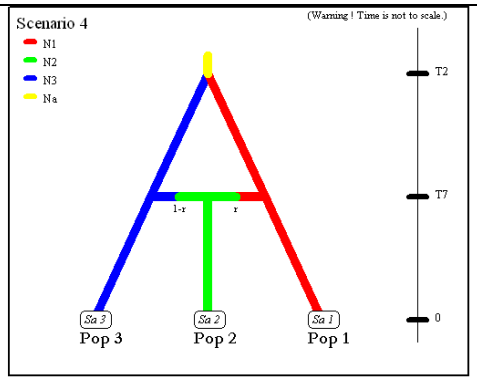
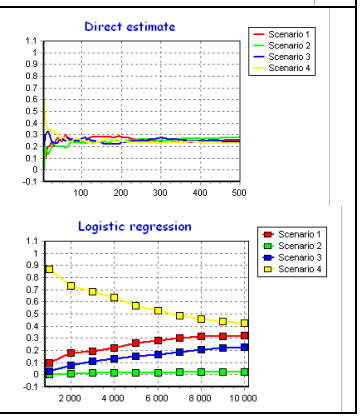


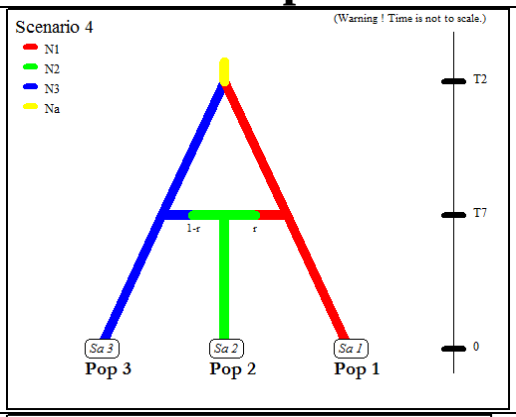
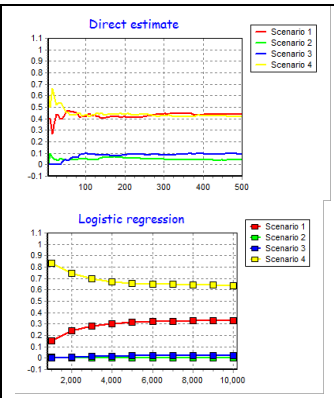
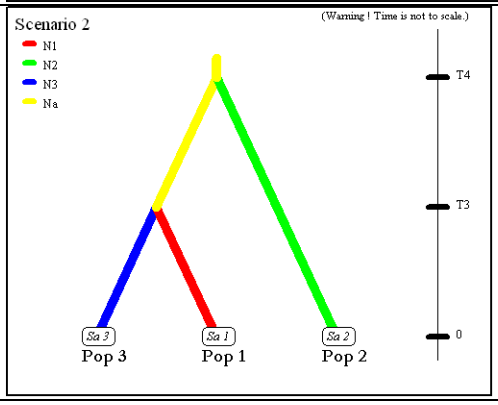
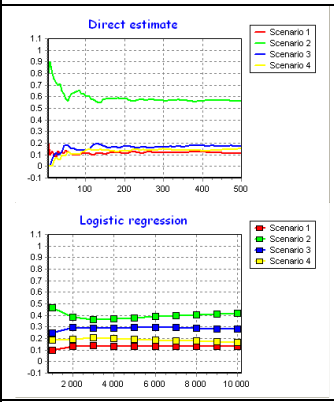
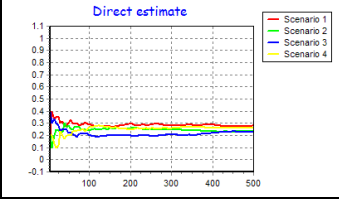


China and Pakistan with other populations

Details	Scenario selected	Selection criteria
ChNpPk Pop. 1 is Ch Pop. 2 is Np Pop. 3 is Pk Ne-estimate: Ch = 313, Np=<402, Pk=>409 Na=499		
ChMdlEstPk Pop. 1 is Ch Pop. 2 is MdlEst Pop. 3 is Pk Ne-estimate: Ch=307, MdlEst= 193, Pk= 393 Na= 421		
PkCACH Pop. 1 is Pk Pop. 2 is CA Pop. 3 is Ch Ne-estimate: Pk= 310, CA= <146, Ch= 633 Na= 950		

Pakistan and Oriental populations

PkCAMdlEst Pop. 1 is Pak Pop. 2 is CA Pop. 3 is MdlEst Ne-estimate: Pk=434, CA=203, MdlEst=182 Na= 557		
PkMdtMdlEst Pop. 1 is Pk Pop. 2 is Mdt Pop. 3 is MdlEst Ne-estimate:		
PkEAfMdlEst Pop. 1 is Pk Pop. 2 is EAf Pop. 3 is MdlEst Ne-Estimate: Pk = >427, EAf = <510, MdlEst = 190 Na = 689		
MdlEstMdtCa Pop. 1 is MdlEst Pop. 2 is Mdt Pop. 3 is CA Ne-estimate: MdlEst =316 , Mdt=<531, Ca=307 Na=899		

Europe		
ChEuPk Pop. 1 is Ch Pop. 2 is Eu Pop. 3 is Pk Ne-estimate: Ch = 291, Eu=<490, Pk=>402 Na=433		
PkEuMdlEst Pop. 1 is Pk Pop. 2 is Eu Pop. 3 is MdlEst Ne-estimate: Pk =>403 , Eu= 116, MdlEst=> 271 Na=683		
EuSAmNAm Pop. 1 is Eu Pop. 2 is SAm Pop. 3 is NAm	No scenario can be selected	

CHAPTER-II

Reduction in the sex ability of worldwide clonal populations of *Puccinia striiformis* f.sp. *tritici*

Sajid Ali, Marc Leconte, Anne-Sophie Walker, Jérôme Enjalbert and Claude de Vallavieille-Pope.

2010: Fungal Genetics and Biology: 47, 828-838

Reduction in the sex ability of worldwide clonal populations of *Puccinia striiformis* f.sp. *tritici*

Published in **Fungal Genetics and Biology: 47, 828-838**

Reduction in the sex ability of worldwide clonal populations of *Puccinia striiformis* f.sp. *tritici*

Sajid Ali[1]*, Marc Leconte[1], Anne-Sophie Walker[1], Jérôme Enjalbert[2] and Claude de Vallavieille-Pope[1]

[1] INRA UR 1290 BIOGER-CPP, BP01, F-78850 Thiverval-Grignon, France

[2] INRA UMR 320 Génétique Végétale, Ferme du Moulon, F-91190 Gif sur Yvette, France

Fax: +33130815306

* Corresponding author: sajid.ali@versailles.inra.fr

ABSTRACT

Puccinia striiformis f.sp. *tritici* (PST), has so far been considered to reproduce asexually with until very recently no known alternate host, has a clonal population structure in the USA, Australia and Europe. However, recently, high genotypic diversity in Eastern Asia and recombinant populations in China has been reported. We assessed whether variations in the ability for sexual reproduction could provide an explanation for such a geographical variability in genotypic diversity and recombination. In order to address this hypothesis, we tested for the existence of a relationship between the ability to produce telia, sex-specific structures that are obligatory for sexual cycle, and the genetic diversity of populations measured using neutral markers, in a set of 56 isolates representative of six worldwide geographical origins. Molecular genotyping and subsequent clustering methods assigned these isolates

to five genetic groups corresponding to their geographical origin, with eight inter-group hybrid individuals. Isolates representing China, Nepal and Pakistan displayed the highest telial production, while clonal populations from France and the Mediterranean region displayed very low telial production. The geographic variability in telial production corresponded to the variability of genotypic diversity described during previous studies, showing a clear difference in telial production between clonal vs. diverse/recombinant populations. The higher mean Qst value (0.822) for telial production than the Fst value (0.317) suggested that telial production has more probably evolved through direct or indirect selection rather than genetic drift alone. The existence of high telial production in genetically diverse populations and its reduction in clonal populations is discussed with regard to evolution of sex in the context of PST centre of origin and distribution of its alternative host.

Keywords: *wheat yellow (stripe) rust, genetic structure, teliospore production, microsatellites*

INTRODUCTION

Pathogens generally display alternating sexual and asexual stages, thus harbouring complex life cycles that strongly impact their adaptative potential to both host and environmental selective pressures (McDonald, Linde, 2002). Sexual reproduction allows the rapid production of advantageous gene combinations, while clonality enables the rapid amplification of strongly fitted individuals (de Meeûs *et al.*, 2007; Samils *et al.*, 2001). A selective advantage of sexual over asexual strains has been effectively found in case of yeast submitted to a changing environment (Goddard *et al.*, 2005). In addition, the sexual phase often involves the differentiation of specific robust structures that enable the pathogen to survive adverse environmental/host conditions (Anikster, 1986; Barrett *et al.*, 2008). Because of these two genetic and demographic effects, the sexual phase can therefore be expected to enhance genotypic diversity (more allelic combinations). As a consequence, populations differing in terms of their ratio of sexual vs. asexual phases can be expected to diverge with respect to allelic, genotypic and phenotypic diversity. Accordingly, Balloux et al. (2003) showed by theoretical approaches that genotypic diversity decreases with increasing rates of asexual reproduction, even if

populations reproducing almost exclusively asexually can maintain high genotypic diversity (Bengtsson, 2003). Indeed, a greater diversity of sexual vs. asexual populations has been reported in different plant pathogen species such as *Puccinia graminis* f.sp. *tritici* (Burdon, Roelfs, 1985), *Puccinia recondita* (Groth, Roelfs, 1982), *Melampsora lini* (Barrett *et al.*, 2008) and *Melampsora larici-epitea* (Samils *et al.*, 2001) regarding both molecular and/or virulence traits. In a context of selection, the reduced diversity of weakly recombining populations is reinforced by hitchhiking effects affecting neutral genetic diversity linked to loci subjected to strong positive selection (Brown, 1995; Kim, Stephan, 2000). In extreme cases, in a strictly asexual population, a mutation conferring a strong fitness advantage on an individual will lead to a complete genomic sweep, where the mutant genotype completely replaces the initial population. This has been observed in the context of boom and bust cycles of wheat yellow rust in response to resistance genes in Europe and Australia (Enjalbert *et al.*, 2005; Hovmøller *et al.*, 2008; Wellings, 2007). Sexual and asexual phases thus present distinct adaptive advantages, and interestingly most plant pathogenic fungi combine the two strategies: they reproduce predominantly through asexual reproduction but this is accompanied by more or less regular sexual phases (Milgroom, 1996). Moreover, in a given species, populations from different geographical areas may diverge with respect to their reproductive cycle (Barrett *et al.*, 2008; Burdon, Roelfs, 1985; Groth, Roelfs, 1982). In order to explain this variability in the mode of reproduction, an issue related to the evolution of sex in a species (McDonald, Linde, 2002; Michod *et al.*, 2008), different hypothesis have been put forward and have mainly focused on the long term vs. short term adaptive advantages of sex, the resistance of sexual spores to adverse environments (Anikster, 1986; Barrett *et al.*, 2008) and the use of alternate hosts that may be available at appropriate times (Ironside, 2007). Knowledge of the identity of species and populations and their reproductive modes, while adopting a broad view of pathogen behaviour in space and time, should improve the abilities of pathologists for better disease management and even predict pathogen behaviour (Taylor *et al.*, 1999). In this regard, *Puccinia striiformis* f.sp. *tritici* (PST, the causal agent of wheat yellow/stripe rust) is an interesting model to address the evolution of sex, because both fixed and genetically diverse populations co-exist in different parts of the world (Bahri *et al.*, 2009b; Enjalbert *et al.*, 2005; Hovmøller, 2002; Mboup *et al.*, 2009).

PST epidemics result from multiple infectious cycles initiated by wind dispersed uredospores produced asexually on wheat (*Triticum aestivum*), which is almost exclusively the only host (Stubbs, 1985). At the end of the cropping season, telia are produced and contain teliospores. These teliospores germinate to form haploid basidiospores (gametes) through meiosis, which may serve to attack an alternative host for sexual reproduction, as in the case of wheat black/stem rust and brown/leaf rust (Fig. 1: *supplementary files*) (Stubbs, 1985). Until very recently, no alternate host was known for PST, and PST was considered to reproduce asexually. Accordingly, efforts made to describe the genetic structure of PST population in North-Western Europe (Enjalbert *et al.*, 2005; Hovmøller, Justesen, 2007; Hovmøller, 2002; Newton *et al.*, 1985), the USA (Chen *et al.*, 1993), Australia and New Zealand (Steele *et al.*, 2001) have confirmed their strong clonality. But as reported in the case of some other species initially considered to be strictly asexual species and subsequently shown to be recombinant (Burt *et al.*, 1996; Gavino *et al.*, 2000; O'Gorman *et al.*, 2009; Sujkowski *et al.*, 1994), recent studies by our team revealed the presence of the Hardy-Weinberg equilibrium and a lack of linkage disequilibrium in a Chinese PST population, suggesting the presence of genetic recombination (Mboup *et al.*, 2009). Furthermore, the presence of relatively higher genotypic diversity has been reported in PST populations in the Middle East (Bahri *et al.*, 2009b) and Pakistan (Bahri *et al.*, submitted), together with that of inter genetic group hybrid genotypes, also in accordance with the presence of genetic recombination. The recent description of *Berberis chinensis* as an alternative host for PST in laboratory conditions (Jin *et al.*, 2010) might explain the origin of these recombination signatures, even if the true role of this species on evolution and genetic structure of world PST populations remains to be studied. The different genetic studies realised so far are only partially covering the wheat-growing areas affected by PST epidemics, but they suggest variability in genetic diversity from North-Western Europe and North Africa to Eastern Asia, with a parallel increase in levels of recombination. Is such a pattern linked to a cline in the rate of sexual reproduction, resulting itself from the ability to reproduce sexually and the availability of the alternate host, or is it due to other causes? We assumed that the presence of a sexual cycle would be supported in the event of a

correlation between an ability to differentiate sexual structure and the presence of genetic diversity and recombination in populations. Indeed, in clonal populations that derive from sexual populations, sex-related structures might be expected to degenerate, because traits that do not contribute to fitness are lost during the course of evolution (Eckert *et al.*, 1999). Indeed the fact that the East-West variability of genotypic diversity do not match the traditional scheme of a PST origin in Transcausia (Stubbs, 1985) raises the alternative hypothesis of a host jump in Asia. In order to assess the impact of sexual recombination on the geographical variability in a setting of neutral genetic diversity, we therefore tried to clarify the existence of a correlation between genetic diversity and an ability to produce sex-specific structures, i.e. telia containing teliospores, which would give rise to haploid gametes (basidiospores). We studied this sex-related trait in 56 isolates representative of the six worldwide geographical origins that display differential genotypic diversity. Our objectives were to i) to see whether any variability does exist for sexual reproduction ability and ii) whether this variability in sexual reproduction could explain the variability in genetic diversity and recombination.

Table 1. List of PST isolates from different geographical origins, selected for the assessment of teliosore production relative to their geographical origin and phylogenetic position.

Geographical origin	Isolate Code	Graph code[a]	Country	Year[b]	Pathotype[c]	Reference
Northern	J89108 MC	F1	France	1989	2, 3, 4, 9, Sd, Su	-
France	J9850	F2	France	1998	1, 2, 3, 4, 9, 17, Sd, Su	Mboup, 2008
	J99198	F3	France	1999	1, 2, 3, 4, 6, 9, 17, Sd, Su	Mboup, 2008
	J08124	F4	France	2008	3, 4, 6, 32, A, Sd, Su	-
Mediterranean	J8617	M1	France	1986	2, 6, 7, 8	Mboup, 2008
Region	J05253	M2	Tunisia	2005	2, 6, 7, 8, 9, A	Bahri et al., 2009
	Beja 2007	M3	Tunisia	2007	2, 6, 7, 8, A	Bahri et al., 2009
	J06283m1	M4	Algeria	2006	1, 2, 3, 4, 6, 9, 17, Sd, Su	Bahri et al., 2009
	J05386	M5	Algeria	2005	2, 6, 7, 8, 9, A, (Sd)	Bahri et al., 2009
	J0610am1	M6	Spain	2006	2, 6, 7, 8	Bahri et al., 2009
	J9706 M1	M7	France	1997	2, 6, 7, 8	Bahri et al., 2009
	J04003	M8	France	2004	2, 6, 7, 8, 9, A	Bahri et al., 2009
Middle-East	WRY 85-22	ME1	Lebanon	1985	2, 6, 7, 9	Calonnec et al., 2002
	J05183 M2	ME2	Iran	2005	2, 6, 7, 9, A	Bahri et al., 2009
	J06448 M1	ME3	Lebanon	2006	2, 6, 7, 8, A	Bahri et al., 2009
	J05682 M1	ME4	Turkey	2005	2, 6, 7, 8, 9, A	Bahri et al., 2009
	J05001	ME5	Cyprus	2005	2, 6, 7, 8, A	Bahri et al., 2009
	J06007 M1	ME6	Cyprus	2006	2, 6, 7, 9, A	Bahri et al., 2009
	J05674 M2	ME7	Turkey	2005	2, 6, 7, 9, A	Bahri et al., 2009
	J05684 M2	ME8	Turkey	2005	2, 6, 7, 8, A	Bahri et al., 2009
Pakistan	J0624 M2	P1	Pakistan	2006	1, 2, 6, 7, 9, Su	Bahri et al., in prep
	J06208 M1	P2	Pakistan	2006	2, 6, 7, 8, Su	Bahri et al., in prep
	J06147 M1	P3	Pakistan	2006	1, 2, 3, 6, 7, 8, 9, A, Su	Bahri et al., in prep
	J0695 M1	P4	Pakistan	2006	2, 6, 7, 9, A, Su	Bahri et al., in prep
	J06075 M3	P5	Pakistan	2006	2, 6, 7, 8, A	Bahri et al., in prep
	J06187 M1	P6	Pakistan	2006	2, 6, 7, 8, A	Bahri et al., in prep
	J06050 M1	P7	Pakistan	2006	2, 6, 7, 8, Su	Bahri et al., in prep
	J06105 M1	P8	Pakistan	2006	1, 2, 6, 7, 9, Su	Bahri et al., in prep
	J06089 M1	P9	Pakistan	2006	2, 6, 7, 8, A, Su	Bahri et al., in prep
	J08X-P14 M1	P10	Pakistan	2008	1, 2, 6, 7, 9, A, Su	-
	J08X-P46M1	P11	Pakistan	2008	1, 2, 6, 7, 9, Su	-
Nepal	J08X-N011M1	N1	Nepal	2008	1, 2, 6, 7, 8, A, Su	-
	J08X-N071M2	N2	Nepal	2008	1, 2, 6, 7, 8, A, Su	-
	J08X-N095 M1	N3	Nepal	2008	1, 2, 6, 7, 8, A, Su	-
	J08X-N097 M1	N4	Nepal	2008	1, 2, 6, 7, 8, A, Su	-
	J08X-N103 M1	N5	Nepal	2008	1, 2, 6, 7, 8, A, Su	-
	J08X-N112 M3	N6	Nepal	2008	1, 2, 3, 4, 7, 9, A, Sd, Sp, Su	-
	J08X-N126 M1	N7	Nepal	2008	2, 6, 7, 8, A, Su	-
	J08X-N144 M2	N8	Nepal	2008	1, 2, 6, 7, 8, A, Su	-
China	CH4/11-07	C1	China	2004	-	Mboup et al., 2009
	CH4/01-04	C2	China	2004	-	Mboup et al., 2009
	CH4/15-09	C3	China	2004	-	Mboup et al., 2009
	CH4/23-08	C4	China	2004	-	Mboup et al., 2009
	CH4/30-08	C5	China	2004	1, 2, 3, 6, 9, 17, A, Sp, Su	Mboup et al., 2009
	CH4/37-11	C6	China	2004	1, 2, 3, 4, 6, 7, 9, A, Sd, Sp, Su	Mboup et al., 2009
	CH4/44-01	C7	China	2004	1, 2, 3, 6, 8, 9, 17, A, Sp, Su	Mboup et al., 2009
	CH5/08-03 M1	C8	China	2005	-	Mboup et al., 2009
	CH5/24-04 M1	C9	China	2005	-	Duan et al., 2010
	CH5/29-01 M1	C10	China	2005	1, 2, 3, 9, 17, A, Su	Duan et al., 2010
	CH2/018	C11	China	2001	-	Duan et al., 2010
	CH2/088	C12	China	2001	-	Duan et al., 2010
	CH2/141	C13	China	2001	-	Duan et al., 2010
	CH2/164	C14	China	2001	-	Duan et al., 2010
	CH2/197	C15	China	2001	-	Duan et al., 2010
Recently spread	DK66/02	R1	Denmark	2002	(2), 6, 7, 8, 9, (Sd)	Milus et al., 2009
US-Euopean clone	DK80/01	R2	Denmark	2001	(2), 6, 7, 8, 9, (Sd)	Dr. M. Hovmøller

[a] Isolate code used in phylogentic tree [b] Year of collection [c] The virulence values given in brackets are of an intermediate infection type (5-6); other virulence values are of a high infection type (7-9) on a 0-9 scale after infection on different cultivars(de Vallavieille-Pope *et al.*, 1995). Similarly, isolates not tested for their pathotype or not reported are marked with ' -`.

MATERIALS AND METHODS

Selection and multiplication of isolates

In order to represent the genetic diversity of PST, a total of 56 isolates were selected from six areas in three continents: Europe, Africa and Asia. A selection was made among the two thousand isolates available in our collection in order to maximise the diversity of each population (assessed in terms of AFLP, microsatellites and virulence) by choosing isolates from the different genotypic groups present in a given geographical region. Details regarding the number of isolates and their virulence profiles are given in Table 1.

All isolates were purified from a single lesion and were then multiplied simultaneously on the susceptible cultivar Cartago. Spores were produced in a fully confined S3-type greenhouse, with an 8 h 15°C dark period and a 16 h 19°C light period (with a light intensity of 300 µmol quanta m^{-2} s^{-1}). Fifteen days after inoculation, uredospores were harvested, placed in a silica gel-filled desiccator for 4 days at 4°C and then stored in liquid nitrogen. The spores were heat-shocked at 40°C for 10 minutes before inoculation to achieve their maximum germination capacity (de Vallavieille-Pope *et al.*, 2002).

Analysis of molecular markers

Genomic DNA was extracted from 10 mg of spores using a modified CTAB protocol (Enjalbert *et al.*, 2002). The extracted DNA was quantified on agarose gel and was stored at -20°C until further use. A total of 16 simple sequence repeat (SSR) markers were used for molecular genotyping, i.e. RJO-4, RJO15, RJO-20, RJO-21 and RJO-24 (Enjalbert *et al.*, 2002) and RJN-2, RJN-3, RJN-4, RJN-5, RJN-6, RJN-8, RJN9, RJN-10, RJN-11, RJN-12 and RJN-13 (Bahri *et al.*, 2009a). The isolates were amplified with SSR markers for 25-50 ng of DNA, using an iCycler (Biorad) thermocycler with 35 cycles of 94°C for 30 s, 52°C for 45 s and 72°C for 30 s. Each reaction mix (10 µL) contained 4.75 µL of deionized water, 2 µL of 1X reaction buffer, 0.60 µL of 2.5 mM $MgCl_2$, 0.80 µL of 2.5 mM

solution of dNTPs, 0.05 µL of 5 U of Taq-polymerase and 0.4 µL of 0.2 µM of the marker together with 1 µL of the DNA. The PCR products were analyzed in 6% polyacrylamide gel and visualized by silver staining (Chalhoub *et al.*, 1997).

The GENETIX 4.05.2 software (Belkhir *et al.*, 2004) was used to compute the number of alleles per loci, linkage disequilibrium among loci and expected heterozygosity (genetic diversity). The GENECLONE software (Arnaud-Haond, Belkhir, 2007) was then applied to test the ability of the set of microsatellite loci to identify clones, by computing the probability of re-sampling any recurrent MLG (multilocus genotype) under panmixia. Population sub-division was analyzed using the STRUCTURE 2.2 software (Pritchard *et al.*, 2000), which implements a Monte Carlo Markov Chain (MCMC) approach in a Bayesian framework to cluster individuals in K clusters so as to minimize the Hardy-Weinberg disequilibrium and gametic phase disequilibrium between loci within clusters. The MCMC scheme was run for 100,000 iterations with a 10,000 burn-in period, as recommended by default, with K ranging from 1 to 10 with at least 20 repetitions to check the convergence of likelihood value for each value of K. The number of populations best representing the observed data was determined by plotting the graph of estimated values of Ln likelihood for each K, and selecting the value which maximized the Ln likelihood of the data, as proposed by Evanno et al., (2005).

Because assumptions regarding the population genetic model implemented under STRUCTURE could be violated in our system, population subdivision was also analyzed using non-parametric multivariate analyses; i.e. factorial correspondence analysis (FCA) with GENETIX software (Belkhir *et al.*, 2004) and construction of a phylogenetic tree with the POPULATION program (Langella, 2008) using the neighbour-joining (NJ) method with an Nei genetic distance and the most parsimonious phylogenetic tree constructed with PHYLIP (Fig. S4). It should be noted that the lower bootstraps values could result due to the inter-group hybrid isolates as the parsimony approach is not appropriate when a hybrid occurs, along with the existence of certain group specific rare alleles.

Assessment of telial production ability

The Cartago cultivar was chosen to assess the telial production ability of the selected isolates, because cv. Victo (previously considered to be highly susceptible) had displayed resistance to some Pakistani isolates. Seeds were sown in 125 cm^3 pots. Of the six pots per isolate, three pots were treated with 20 mL maleic acid (0.25 g L^{-1}, leaf senescence inhibitor) when the seedlings were 1 cm tall, in order to observe the effect of senescence on telial production. Applications of maleic acid (hereinafter referred to as the "senescence inhibitor") delay first leaf senescence and increase uredospore production by blocking the growth of upper leaves. The plants were thinned to 15 plants per pot when the seedlings were 7 day-old, and then inoculated by spraying them with a spore suspension of one mg spores/300 µL soltrol (mineral oil). All inoculated plants were left to evaporate the soltrol and then placed in humid plastic bags that were placed in a climatic chamber at 9°C for 24 hours, to serve as a dew chamber. The plants were then transferred to a greenhouse with optimum conditions (15°C for 8 h dark period and 19°C for 16 h-light period). A week after inoculation, cellophane bags were placed over each pot to prevent cross-contamination between the isolates. During the first set of experiments, the temperature was kept lower than 20°C in March and April. During the second set of experiments, the average temperature was higher (25°C), which thus affected plant growth.

A weekly scoring system was implemented, starting after four weeks of inoculation and continuing until seven weeks post-inoculation. Scores were determined on a 0-10 scale based on the percentage of leaf covered by telia, taking account of its incidence and relative intensity over the 15 plants per pot (Fig. 1). The whole experiment was repeated twice. Telial production data were analyzed using the general linear model to test the telial production area (0-10) as a function of isolates nested in six geographical regions. Analyses were performed separately at all four dates. Data for the four scoring dates were converted into the area under the telial production curve (AUTPC) in order to reflect overall telial production as affected by the presence (SI-P) or absence (SI-A) of the senescence inhibitor. Analyses of variance were performed using the R statistical software (version 2.8.0).

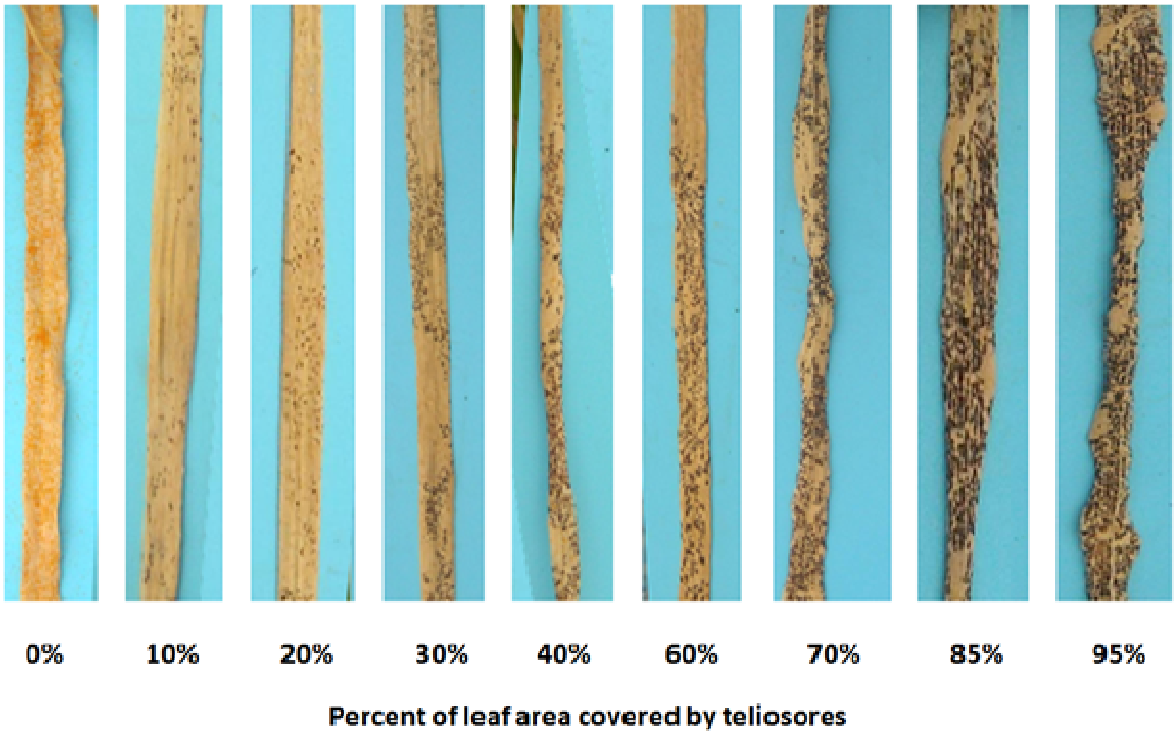


Fig. 1. The overall scoring of telial production was performed on a 0–10 scale (with 10 as the maximum score) based on the percentage of leaf area covered by telia while considering its incidence and relative intensity in 15 plants per pot.

Relationship between genotypic and phenotypic parameters

The relationship between genetic structure and telial production was assessed in the five genetic groups by comparing their Qst calculated for AUTPC and their Fst estimated from SSR data. Qst represents the degree of differentiation between populations for a quantitative trait, and was calculated according to the method described by Spitze (1993). Fst, the degree of differentiation between populations for neutral genetic markers (here SSR), was estimated by the GENETIX software using a pair-wise Fst estimation with 2000 permutations. The significance of the correlation between the two differentiation matrices was assessed by a Mantel test (Mantel, 1967), using a bilateral probability of 0.05 and 1000 random permutations, performed with the XLSTAT 2009 software.

RESULTS

Genetic grouping vs. geographical distribution of PST isolates

Isolates were genotyped with 16 SSR markers, with one primer pair amplifying two loci (RJO-20low and RJO-20up) so that a total of 17 loci were scored. All markers were polymorphic when all 56 isolates were considered, with the number of alleles ranging from two (RJN-8 and RJN-12) to 11 (RJN-2) (Table 2). Some loci were fixed in some of the populations; for example, six loci were monomorphic in Northern French isolates, while RJN-12 was fixed in all populations except those from China (data not shown). Genetic diversity expressed in terms of expected heterozygosity (He) was found to reach its maximum for RJN-2 (0.83) and RJO-24 (0.80), whereas the lowest value was recorded for RJN-12 (He = 0.03), as only two individuals presented a different allele. The linkage disequilibrium (LD) between pairs of SSR loci was estimated with GENETIX software and its significance was tested using 2000 permutations (Table 3). The overall low LD across all isolates suggested a lack of any redundancy for utilisation of the current set of loci. However, this LD could be higher if it was estimated separately in each geographical group. The asymptote reached when plotting the number of multilocus genotypes detected against the number of loci re-sampled showed that almost a full set of markers was necessary to differentiate all the selected genotypes (Fig. S2).

Table 2. Number of alleles, expected (He) and observed heterozygosity (Ho) estimated for 17 polymorphic SSR loci in 56 PST isolates from six geographical origins.

SSR locus	No. of alleles	Expected Heterozygosity	Observed Heterozygosity
RJN-2	11	0.83	0.37
RJN-3	3	0.66	0.21
RJN-4	4	0.51	0.53
RJN-5	4	0.59	0.72
RJN-6	4	0.25	0.25
RJN-9	5	0.75	0.59
RJN-10	3	0.59	0.64
RJN-11	5	0.74	0.87
RJN-13	3	0.45	0.33
RJO-15	3	0.46	0.6
RJN-8	2	0.28	0.33
RJN-12	2	0.03	0.03
RJO-4	3	0.59	0.64
RJO-20up	3	0.55	0.28
RJO-20low	7	0.76	0.79
RJO-21	5	0.77	0.17
RJO-24	10	0.80	0.86
Mean	4.53	0.57	0.48

Table 3. Permutation percentage (%) of real values (upper values) and the presence of a linkage disequilibrium (bold type) between SSR loci in the 56 PST isolates tested.

	RJN-2	**RJN-3**	**RJN-4**	**RJN-5**	**RJN-6**	**RJN-9**	**RJN-10**	**RJN-11**	**RJN-13**	**RJO-15**	**RJN-8**	**RJN-12**	**RJO-4**	**RJO-20up**	**RJO-20low**	**RJO-21**	**RJO-24**
RJN-2	-	0.10	27.85	25.15	0.25	77.45	15.35	4.05	3.30	83.75	0.10	90.65	1.60	7.55	0.30	0.20	12.70
RJN-3	-	-	4.15	2.95	93.90	29.65	**0.00****	1.30	0.40	92.15	0.05	74.35	19.95	0.25	**0.00****	0.05	53.50
RJN-4	-	-	-	0.15	59.85	73.20	0.20	69.80	65.20	80.85	6.00	87.30	0.05	**0.00****	1.40	**0.00****	76.90
RJN-5	-	-	-	-	32.15	95.45	**0.00****	69.60	5.65	5.85	1.65	15.50	0.40	0.70	1.05	**0.00****	31.85
RJN-6	-	-	-	-	-	49.25	73.25	88.00	16.20	61.25	35.90	90.00	33.75	30.40	66.45	25.30	1.55
RJN-9	-	-	-	-	-	-	40.00	56.30	39.65	8.95	62.30	42.65	4.55	62.05	87.50	15.30	75.50
RJN-10	-	-	-	-	-	-	-	29.10	9.10	3.25	7.05	75.70	0.60	0.05	**0.00****	0.75	14.30
RJN-11	-	-	-	-	-	-	-	-	17.45	32.10	0.25	98.20	27.80	**0.00****	**0.00****	1.75	20.80
RJN-13	-	-	-	-	-	-	-	-	-	50.90	**0.00****	100.00	50.55	1.60	0.95	0.15	89.45
RJO-15	-	-	-	-	-	-	-	-	-	-	52.75	76.50	26.35	96.25	8.55	0.55	63.35
RJN-8	-	-	-	-	-	-	-	-	-	-		86.70	17.25	0.15	**0.00****	**0.00****	70.70
RJN-12	-	-	-	-	-	-	-	-	-	-	-	-	62.40	100.00	78.25	76.20	7.20
RJO-4	-	-	-	-	-	-	-	-	-	-	-	-	-	0.40	**0.00****	**0.00****	64.30
RJO-20up	-	-	-	-	-	-	-	-	-	-	-	-	-	-	**0.00****	**0.00****	74.10
RJO-20low	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	**0.00****	24.50
RJO-21	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	4.45
RJO-24	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

** P < 0.05%; a value of <10 over 2000 permutations (calculated with GENETIX software) was considered to be significantly associated

Using STRUCTURE analysis to assess the presence of the genetic structure in the dataset (Pritchard *et al.*, 2000), the Log likelihood of the data was drawn against K (the number of clusters) in order to identify the K value that best fitted the dataset proposed by Evanno et al. (2005). The Log likelihood increased strongly up to five clusters (K = 5), which appeared to be the optimum subdivision of our dataset (Fig. S3: *supplementary file*). These five genetic groups corresponded well to their predefined geographical origins, with an addition of nine inter-group hybrid individuals (Fig. 2). However, some of the isolates were assigned to a genetic group that was not related to their geographical origin and could be considered as migrants. French isolates were grouped in one genetic group (G1), Mediterranean and Middle-East isolates in a second cluster (G2), Pakistani isolates in a third group (G3), Chinese isolates in a fourth group (G4), and finally Nepali isolates in a fifth group (G5). When inter-group hybrid individuals were considered, they were mostly assigned in accordance with previous results, except for J06448 (Middle-East), which was now fully assigned to the Middle-East/Mediterranean genotype (G2) using the larger set of SSR markers used during the present study. Among the Pakistani isolates, eight were grouped into a separate, Pakistan-specific cluster (G3), while only two; i.e. J05060M1 and J06147M1, were strongly assigned to the Mediterranean-Middle-East population (G2). Isolate J06208 was assigned partially to G3 and partially to G2, while J0695M1 was not assigned to any group. Among the Nepali isolates, all were assigned partially to G5 (Nepal-specific profile) with some partially assigned to either G3 or G4. Most of the Chinese isolates were assigned to a distinct group (G4), specific to China, with a few being recombinant with any of the other three genotypic groups (G2, G3 and G5). Interestingly, two Chinese isolates were assigned to G1; i.e. Northern French isolates (CH2/141 and CH2/088). Similarly, of the two isolates representative of the recently worldwide spread aggressive clone, DK80/01 was assigned to the Middle-East/Mediterranean (G2) population while DK66/02 was partially assigned to G2 (Mediterranean/Middle-East-specific genotype) and G4 (Nepal-specific genotype), with a genotype that was similar to that of two Middle-Eastern recombinant isolates, i.e. J05001 and J06007M1. However, this G5 genotypic group did not remain stable when the K value was reduced to 4 (data not shown), and DK66/02 was fully assigned to G2. This assignation was further affirmed by the tree generated using PHYLIP (Fig. S4).

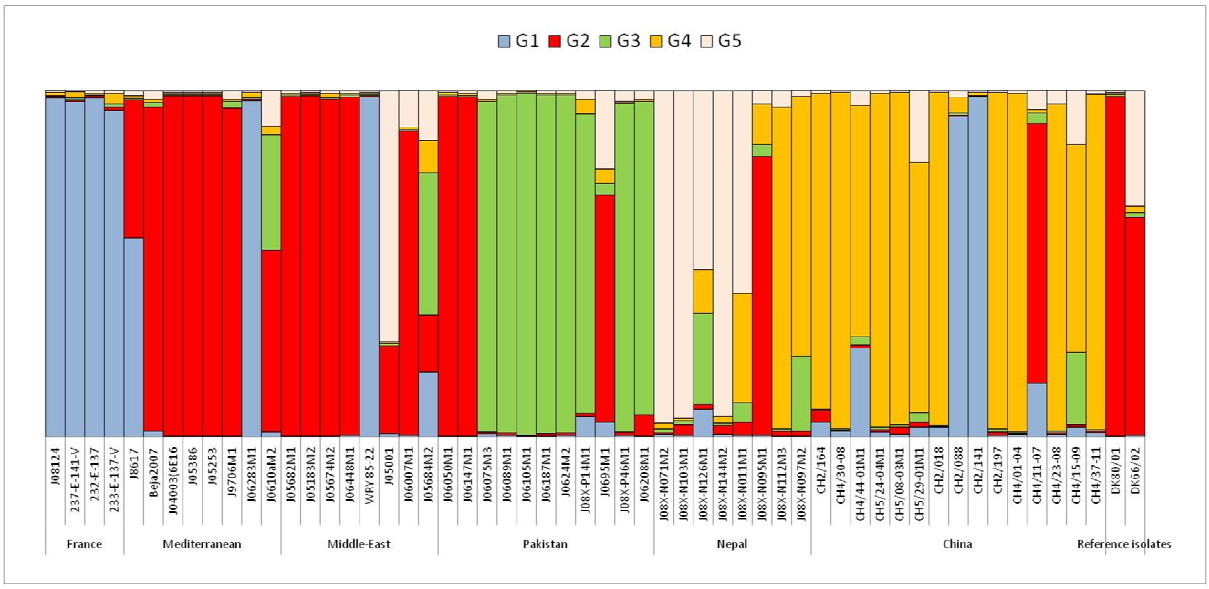


Fig. 2. STRUCTURE software results showing the assignment of PST isolates to the five inferred genotypic groups (G1–G5) on the basis of SSR polymorphism. Each colour represents a different genotypic group.

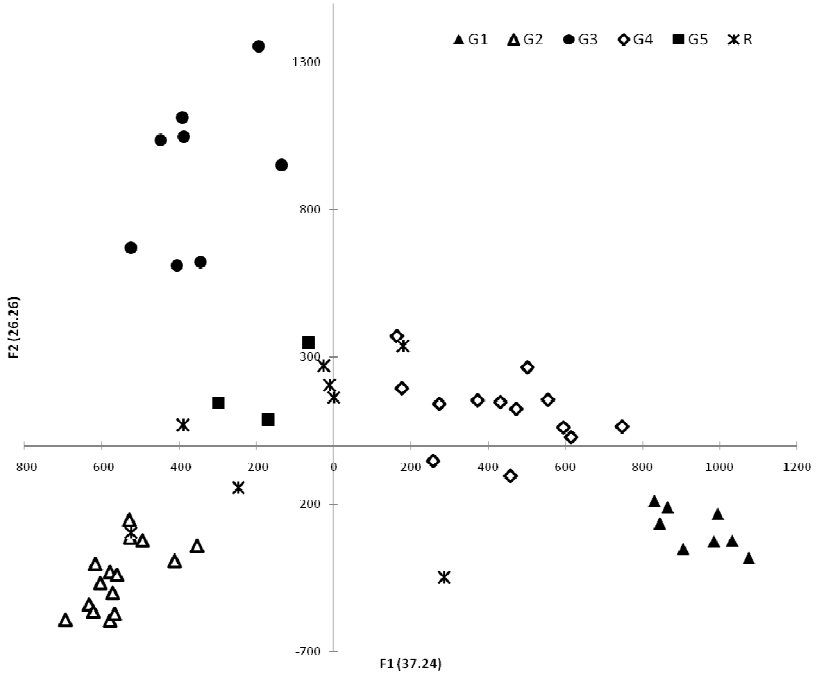


Fig. 3**.** Factorial Correspondence Analysis of SSR data of PST isolates, showing five genetic groups representative of six geographical origins, i.e. G1 = France; G2 = Mediterranean/Middle-East; G3 = Pakistan; G4 = China; G5 = Nepal; R = recombinant isolates.

The relatedness of isolates was also confirmed by Factorial Correspondence Analysis (FCA) on the SSR data (Fig. 3). On the first two FCA axes (63.30% of total genetic variance), the five STRUCTURE groups were clearly separated, while inter-group hybrid isolates were found to be intermediate to them. The reliability of this structure was also indicated by the neighbour-joining tree constructed using POPULATION software (Fig. S4), and the most parsimonious phylogenetic tree constructed with PHYLIP (not shown). The low bootstrap values are explained by both the presence of inter-group hybrids with unsteady position in the tree, as well as the presence of few rare alleles at the origin of the discrimination of some of the groups.

Ability for telial production

Highly significant differences between both geographical origins and isolates were observed at all the scoring dates ($p < 0.01$, Table 4). The repeatability of the results was demonstrated (Fig. 4) when analysing the correlations between the final scorings of the two experiments. However, a higher pace of telial production occurred during the second replication, when the third scoring was equivalent to the fourth scoring of the first replication, while the fourth scoring was abandoned because of drying of the plants. For this reason, the kinetics and scorings of the first replicate are presented hereinafter because they are more detailed (one additional scoring date), while statistical analysis was performed on the entire dataset. We therefore present the first replication data, instead of a mean over two replications, knowing that the overall pattern remained the same across the two replications.

All the isolates studied produced telia with a clear increase in density over time (Fig. 5) during the two repetitions. Telial production was higher for China and Nepal isolates, followed by Pakistan isolates, at all scoring dates. The mean value for Northern France and Mediterranean PST isolates was the lowest at all scoring dates. The Middle-East-Mediterranean isolates and the recently disseminated high temperature adapted clone representing isolates were between the two categories. Mean values for all the geographical origins tested reached their maximum at the third scoring date and remained at the same level on the fourth date, with a slight increase for Middle-East isolates. It should be noted

that there was a decrease after the third scoring of some isolates, which may have been due to scoring error, as some leaves had started to dry so that the telia could not be scored with total accuracy.

Table 4. Analysis of variance of teliospore production at four scoring dates for 56 PST isolates, during the first experiment.

	Weeks after inoculation			
	4th week	**5th week**	**6th week**	**7th week**
	F-calculated	F-calculated	F-calculated	F-calculated
Geographical origin	55.25 ***	143.44 ***	161.11 ***	147.47 ***
Isolate	6.90 ***	17.36 ***	11.20 ***	10.54 ***
Multiple R	0.71	0.86	0.85	0.83

*** Displaying highly significant differences.

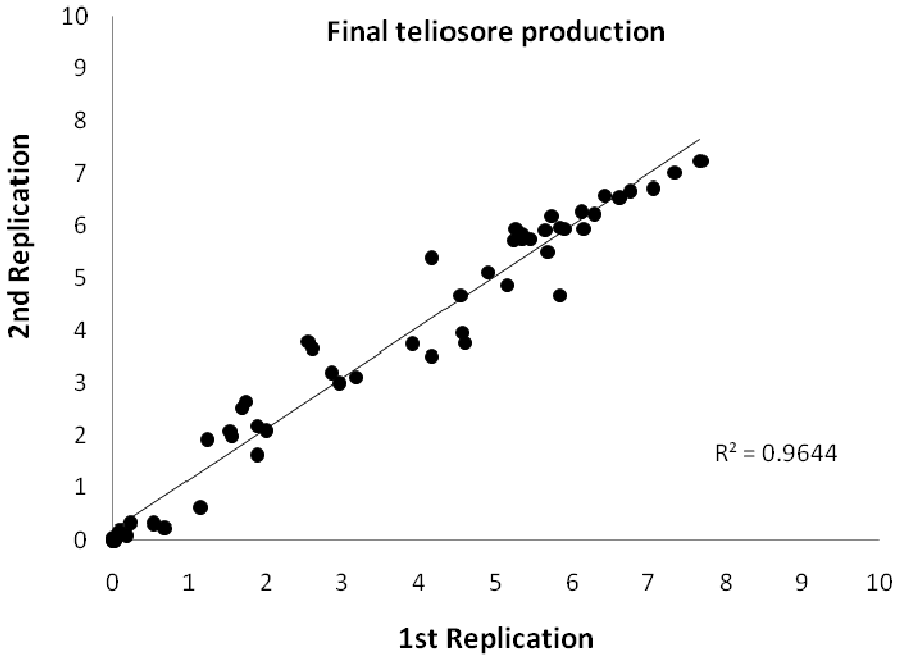


Fig. 4. Final scoring of telial production (based on a 0–10 scale) for two replications performed independently.

When the isolates were regrouped with respect to the five genetic groups identified by the STRUCTURE analysis, a very similar picture was obtained (Fig. 6, shown in terms of AUTPC). The group containing all the hybrid isolates also displayed a higher level of telial production. Application of the senescence inhibitor resulted in higher telial production than otherwise in all the genetic groups (Fig. 6). The four French PST isolates produced the smallest quantities of telia, while the highest production was recorded for Chinese isolates, reported to be from a panmictic population (Mboup *et al.*, 2009), with highly variable telial production (mean values ranging from 0.1 to 7.2). Inter-group hybrid isolates displayed high variability (0.32-6.58), however, hybrid isolates from China, Nepal and Pakistan had a higher telial production than hybrid isolates from the other genetic groups.

Relationship between phenotypic trait and genetic structure

The relationship between genetic structure and telial production was assessed by testing the correlation between Qst and Fst in each pair of the five inferred genetic groups (Table 5). The pairwise Qst estimated for AUTPC revealed that all five genetic groups differed significantly in terms of their telial production. High telial producing genetic groups, i.e. China (G4), Pakistan (G3) and Nepal (G5), displayed the maximum phenotypic distance from the low telial producing genetic groups of France (G1) and Mediterranean/Middle-East (G2). The minimum distance was found between G5 and G3 (Qst = 0.397), followed by G5 and G4 (0.535).

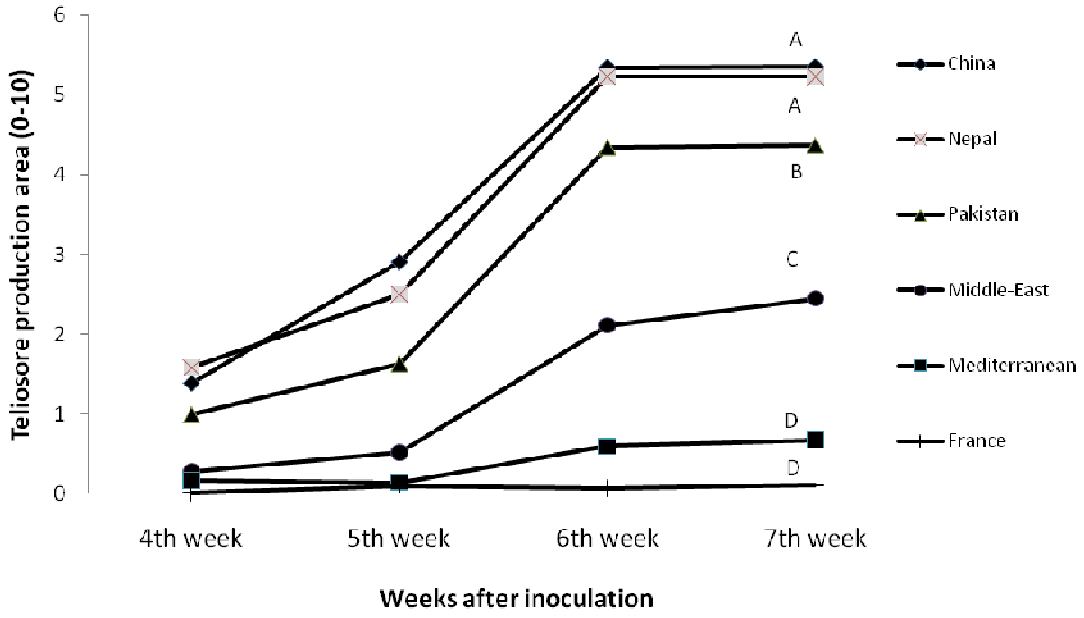


Fig. 5. Telial production pattern on a 0–10 scale across four scoring dates (above) recorded 4–7 weeks after inoculation for the six geographical origins. Values with different letters differ significantly at P 0.05, based on the LSD test.

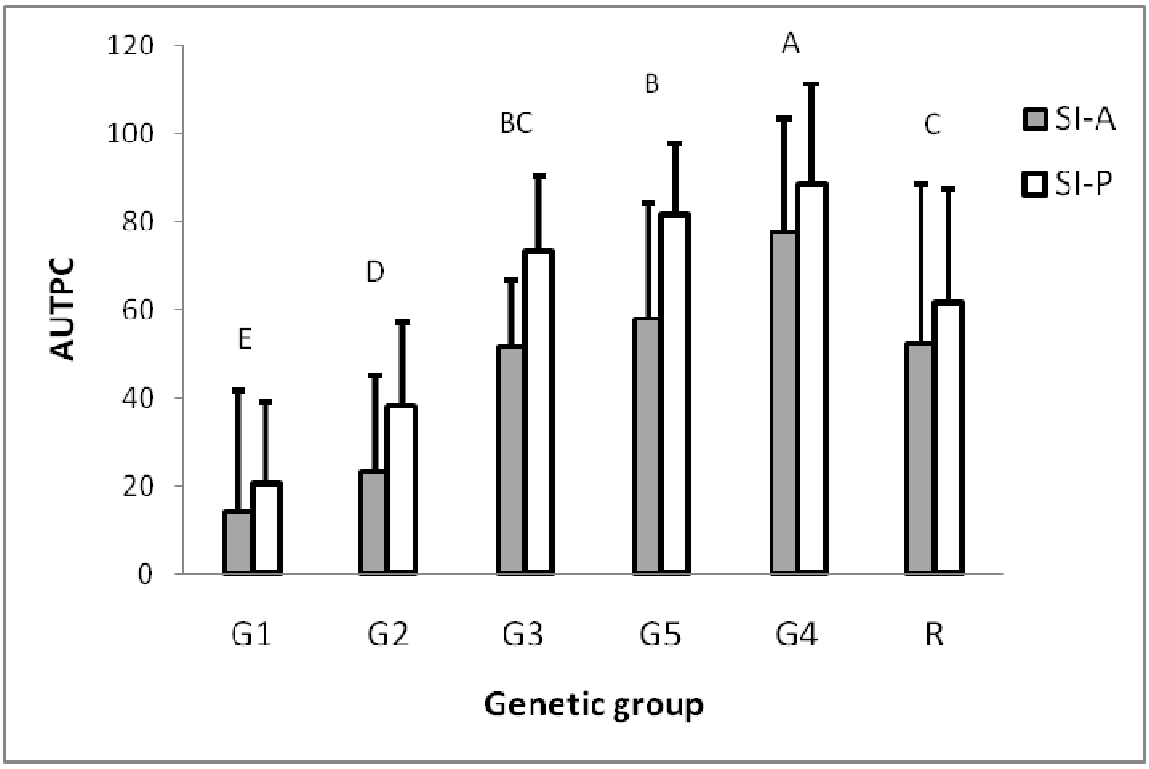


Fig. 6. Area under the telial production curve (AUTPC) of five genetic groups and recombinant isolates (R) of PST isolates representative of six geographical origins. SI-P and SI-A represent the presence or absence of the senescence inhibitor applied to the pot, respectively. (G1 = France; G2 = Mediterranean/Middle-East; G3 = Pakistan; G4 = China; G5 = Nepal; R = recombinant isolates).

Pairwise Fst values between the six geographical origins showed that five out of the six geographical origins were significantly distant from one another (Table S1), while a similar pattern was displayed for genetic groups made following STRUCTURE analysis (Table 5). The maximum Fst was between Northern France and Pakistan (0.411), while the minimum was recorded for Pakistan and Nepal (0.197). The Mantel test performed on the Qst and Fst matrices revealed a weak but non-significant (Mantel r value of 0.432, p = 0.23) relationship between telial production behaviour and genetic differentiation (Fig. 7). Indeed, populations that were related on the basis of their SSR diversity were similar from a phenotypic point of view, while phenotypically distinct populations might or might not be distinct genetically. When considering the overall mean Fst and Qst values, Qst (0.822) was higher than Fst (0.317), thus revealing a marked differentiation in terms of telial production when compared to the neutral genetic markers. Thus a strong genetic variability was recorded for telial production across the populations, with a higher pace of evolution for telial production than for neutral markers.

DISCUSSION

We report here on the presence of highly significant differences regarding telial production between PST isolates representative of the five major genetic groups identified during previous worldwide studies. The higher telial production found for Asiatic isolates agreed with the recent report of diverse/recombinant populations in Pakistan and China (Bahri *et al.*, 2009b; Duan *et al.*, 2010; Mboup *et al.*, 2009). By contrast, the low teliospore producing European/Mediterranean isolates matched with their clonal genetic structure (Hovmøller, Justesen, 2007; Newton *et al.*, 1985). These findings therefore support the existence of sexual recombination in Asian populations and highlight the importance of the evolution of sex related traits to understand PST population genetics.

Table 5. Estimates of genetic (Fst, upper diagonal) and phenotypic (Qst, lower diagonal)differentiation measurements for the teliospore production of PST isolates from five genetic groups representative of six geographical origins, i.e. G1 = France; G2 = Mediterranean/ Middle-East; G3 = Pakistan; G4 = China; G5 = Nepal.

Genetic group	**G1**	**G2**	**G3**	**G4**	**G5**
G1	-	0.6997	0.9585	0.9767	0.9391
G2	0.3698	-	0.9492	0.9801	0.9347
G3	0.4109	0.3024	-	0.8485	0.3972
G4	0.2693	0.3357	0.3775	-	0.5346
G5	0.3523	0.2506	0.1967	0.3052	-
Mean Fst	**0.3170**			**Mean Qst**	**0.8218**

All Fst values were significantly different at 5% (100 out of 2000 permutations).

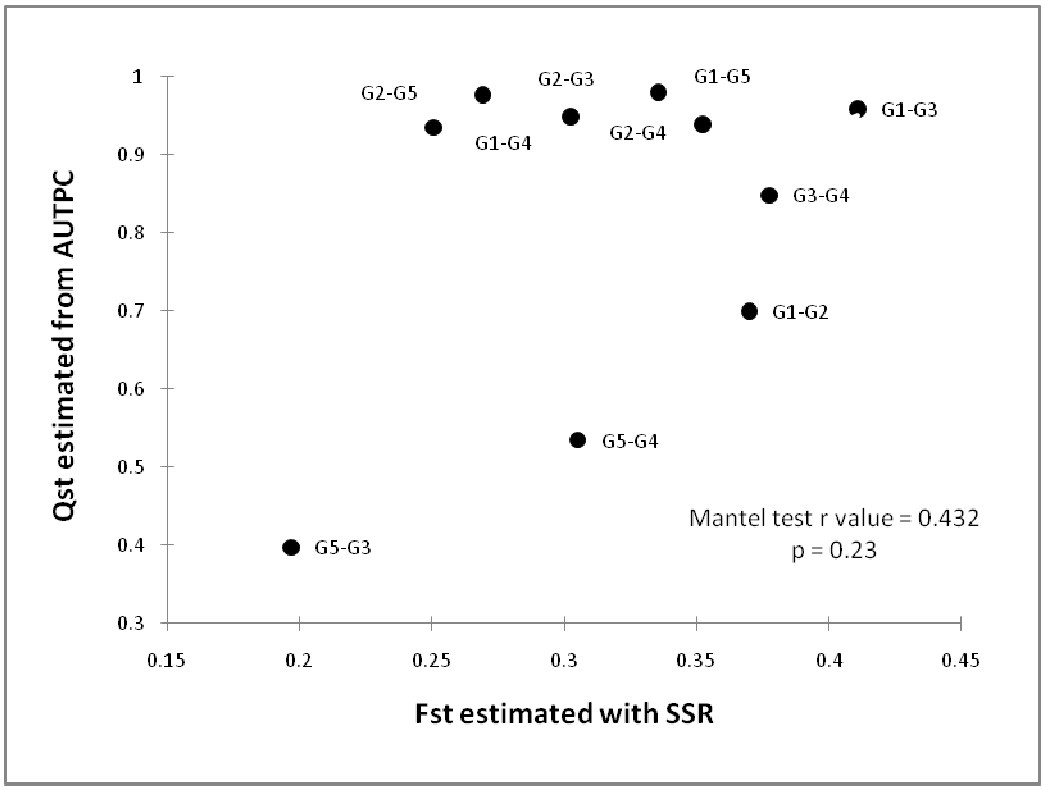


Fig. 7. Mantel test results for associations between phenotypic differentiation measurements (Qst) and genetic differentiation measurements (Fst) regarding the distance between five genetic groups of PST isolates representative of six geographical origins (G1 = France; G2 = Mediterranean/Middle-East; G3 = Pakistan; G4 = China; G5 = Nepal).

Phylogenetic position of isolates and genetic grouping of populations

Molecular marker analysis showed that selected isolates constituted a good reflection of geographical origins, separating them into five genetic groups, i.e. Northern France (Europe), Mediterranean and Middle-East, Pakistan, Nepal and China (Fig. 4). The overall assignation of isolates to the five genetic groups was in accordance with previous population studies, despite the use of a new set of SSR markers that showed an increase in genetic diversity from western to eastern origins, and a parallel increase in the frequency of hybrid or recombinant isolates (Bahri *et al.*, 2009b; Bahri *et al.*, submitted; Duan *et al.*, 2010; Enjalbert *et al.*, 2005; Mboup *et al.*, 2009). Interestingly, this study also enabled the assignment of two isolates representative of recently emerging PST clones (Hovmøller, Justesen, 2007) to a Mediterranean/Middle-East origin. Previously, Bahri et al. (2009b) had also suggested that the recently emerging PST clones of Southern France originated from the Mediterranean/Middle-East. Only a few genotypes were identified as migrants, because of their assignment to a genetic group endemic in a distant geographic region. Globally, the selected isolates were well assigned to their genetic groups, and more or less representative of that geographical origin.

Sex-related traits vs. genetic diversity

We also report the presence of highly significant differences in telial production between isolates and between genetic groups or geographical origins. As the geographical and genetic groups were strongly concordant, and because of the presence of only a few migrants, we discuss below the phenotypic variability found within genetic groups (i.e. geographical origins corrected inter-group hybrids and migrants). Isolates from the three genetic groups for China, Nepal and Pakistan had a high telial production, while Northern French and Mediterranean/Middle-East isolates had very low or no telial production. Isolates identified as recombinant between two groups with a high telial production also displayed high telial production. As expected, the high telial production found for Chinese, Pakistani and Nepali isolates was associated with populations displaying higher genetic diversity, as previously found with a large set of isolates (Bahri *et al.*, submitted; Mboup *et al.*, 2009), and *vice versa* for

clonal and low telial producing European isolates. This concordance between telial production and genetic structure was therefore in line with a direct role for sexual recombination in the structure of PST genetic diversity, without there being definite proof of its actual occurrence. Indeed, the sexual cycle is not the only means of recombination in fungi, and asexual recombination or parasexuality has been described to various degrees in different fungal species (Clutterbuck, 1996; Noguchi, Fujita, 2006; Spring, Zipper, 2006). In PST, parasexuality has been suggested as a means of generating new virulence phenotypes by the rearrangement of whole nuclei (Little, Manners, 1969; Wright, Lennard, 1980), even if a complete parasexual cycle involving chromosomal recombination has never been demonstrated to date (Mboup *et al.*, 2009). However, if asexual recombination does occur significantly in some areas, a lack of correlation between telia and genetic diversity would be observed, with some populations exhibiting a strongly recombining genetic structure with poor telial production, which was not the case here. So although there is little proof of parasexuality, it nonetheless merits further investigation.

In the absence of any important role for sexual recombination in population evolution, sex-related traits should evolve as neutral traits, and therefore display a genetic structure that is coherent with the marker structure. This should provide another opportunity to test for the evolutionary importance of telial production in PST populations. Thus the existence of a quantitative variation in a sex-related trait (a concept that has been the subject of little studied and described), raises questions concerning an evolutionary basis for variations in telial production.

Evolution of sex-related trait

The divergence between populations regarding phenotypic traits and neutral markers provides a means of comparing the relative importance of selection forces and genetic drift to the evolution of a trait (Kaeuffer *et al.*, 2006; McKay, Latta, 2002; Reed, 2007). On the one hand, fitness-related traits are subject to selection, while molecular markers are assumed to be neutral, evolving through genetic

drift (Kimura, 1983; McKay, Latta, 2002) or via a hitchhiking effect when in linkage disequilibrium with QTLs. If divergent selection does occur in recently separated populations, differences between populations with respect to the phenotypic trait will be greater than for neutral markers (Merilä, Crnokrak, 2001; Miller *et al*., 2008). On the other hand, if a similar set of selection forces is driving a phenotypic trait across different populations, greater differentiation of neutral markers could be expected (Lewontin, Krakauer, 1973). The overall comparison between telial production and genetic structure revealed a weak and non-significant relationship: populations with a similar phenotype were equally similar genetically, but phenotypically distant populations could or could not be close genetically. In addition, because the mean Qst value (0.822) was higher than that of Fst (0.317), this suggests the presence of divergent selection or a higher pace of evolution of telial production across populations than for neutral markers (Clifford, Clothier, 1974; Lynch *et al*., 1999; McKay, Latta, 2002; Merilä, Crnokrak, 2001). Therefore, genetic drift alone (Brace, 1963; Wilkens, 1988) could not explain this variation in telial production across populations, so that some direct/indirect selection may be involved, as suggested previously (Poulson, White, 1969; Regal, 1977). However, the direction of the selection towards loss of telia having a cost or relaxed selection on their maintenance remains an open question because different hypotheses could lead to higher Qst. Similarly, the relationship between Qst and Fst within populations displaying partial or total clonality has not yet been studied, and this point deserves further theoretical and experimental exploration.

Hypothetical scenario for the evolution of sex in PST

The hypothesis concerning evolution of this trait was an underlying feature of the present study: the differences observed in telial production might have been related to the importance of sexual reproduction to PST cycles in different regions. Sexual reproduction is a plesiomorphic (ancestral) trait in the *Puccinia* genus, as in all fungi, so it could have been lost in the more or less recent past of PST populations. Note that this loss of sexual cycle is only viable for this biotrophic fungus in areas where there is an almost continuous presence of wheat, as urediniospores have a very low viability

(max. 15 days). In Europe, the clonal survival in absence of wheat in summer can be explained by the presence of volunteers in coastal regions (more rainy), responsible for a "green bridge" and source of primary inoculum for next epidemics. Reciprocally, a functional sexual cycle in some recombinant PST populations also requires the presence of the aecial host, an assumption supported by the distribution of the very recently reported *Berberis* spp. as aecial host for PST in China, Central Asia and Caucasus (Jin *et al.*, 2010). Thus, migrations of PST in regions where this host is absent should result instantly in redundancy of the sexual cycle. As selection is released from this trait, deleterious mutations could be expected to accumulate in genes dedicated to the differentiation of sexual structures, resulting in a deterioration of sexual functions. Therefore, the degradation of sex-related functions such as telial production could be expected to be proportional to the time elapsing since a population turned to a strictly clonal behaviour. Few examples of sex trait deterioration with clonality have been described so far. The absence/loss of teliospore production, or their germination capacity in clonally reproducing populations, has been reported in *Melampsora lini* (Barrett *et al.*, 2008) and observed to some extent in *Puccinia graminis* f.sp. *tritici* in the USA, where the eradication of alternative hosts has affected both the genetic structure of the population (Groth, Roelfs, 1982) and its telial production capacity (Roelfs, personal communication). In the case of *Magnaporthe grisea*, fertile female isolates lose their fertility when they are reproduced several times asexually (Didier Tharreau, personal communication). Similarly, a loss of sex-related traits in clonal populations has also been reported in some plant species (Eckert *et al.*, 1999; Kennington, James, 1997; McKee, Richards, 1996). Thus the evolution of these sex-related traits could be explained by an important feature of evolution, i.e. the loss of traits that do not contribute to fitness, a point that has been acknowledged since Darwin (1859) and further addressed by others (Brace, 1963; Maughan *et al.*, 2007; Poulson, White, 1969; Prout, 1964; Ralph, 1966; Regal, 1977; Wilkens, 1988).

The loss of traits that do not contribute to fitness could be explained by two major hypotheses; one based on neutral genetic drifts (Brace, 1963; Wilkens, 1988), and the other on the direct or indirect selective advantage of the mutations involved in this loss (Poulson, White, 1969; Regal, 1977). As mentioned above, Qst and Fst comparisons suggest that differences in telial production could not be

explained solely by genetic drift, because it exerts an equal impact on both telia and neutral markers. The other hypothesis could be answered by addressing the advantages linked to this loss. A fitness advantage conferred by loss of sex implies the existence of a trade-off between teliospore and uredospore production. However, initial attempts to compare the differences in both telial and uredospore production in clonal (French) and recombinant (Chinese) isolates showed that Chinese isolates produced more uredospores and telia than French isolates (Fig. 8). This could, however, be explained by the accumulation of deleterious mutations in clonal populations, causing fitness penalties due to Muller's ratchet effect, and resulting in less aggressiveness than recombinant and diverse populations. Detailed studies will be necessary to assess this trade-off concept in Chinese populations using a large set of isolates.

Conclusion

The higher levels of telial production by recombinant isolates than by clonal populations, and the concordance of the telial production variation with the variation in genetic diversity, have led us to support the hypothesis of the presence of sexual recombination in this species, which might facilitate identification of the cryptic alternate host for PST. Because sex is an ancestral trait in *Puccinia* spp., the evolution of sex is certainly a good angle from which to throw light on the centre of origin of this pathogen, which also remains unknown. The existence of a greater diversity and higher telial production in China and Asia along with the distribution of the recently reported aecial host i.e., *Berberis* spp. (Jin *et al.*, 2010), could suggest an Asiatic origin of PST, and therefore support the hypothesis of a host jump, rather than the generally assumed theory of co-evolution at the centre of origin of host i.e. the fertile crescent (see however McDonald, Linde, 2002). Gaining a clearer understanding of the importance of sex in worldwide PST populations is crucial to the development of integrated disease management. It will not only help to clarify the role of recombination for adaptation, as well as widely documented mutation events (Steele *et al.*, 2001), but will also determine alternative pathways for the over-wintering/over-summering of this pathogen.

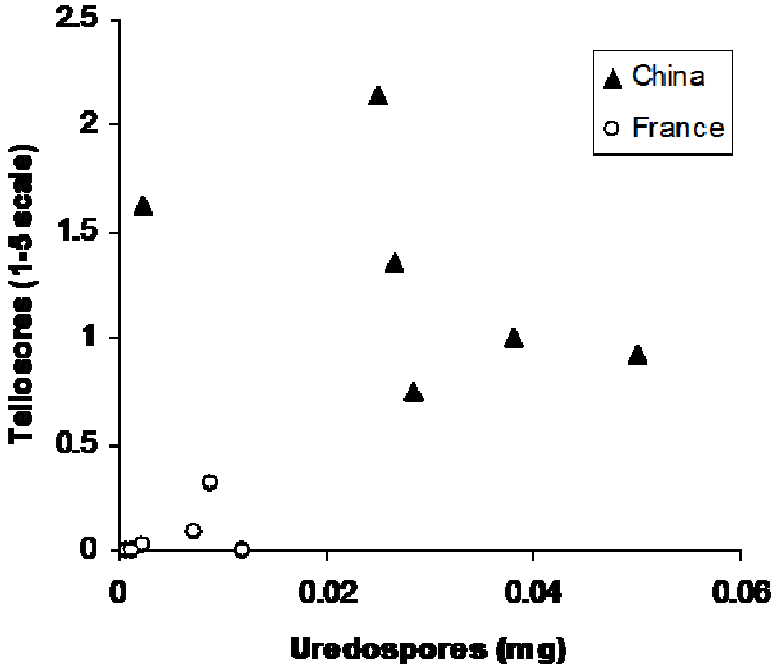


Fig. 8. Comparison of telial and urediospore production in Chinese (recombinant) and French (clonal) isolates.

ACKNOWLEDGEMENTS

The authors are grateful to all their colleagues who participated in establishing the yellow rust collection used for this study. Our thanks go to Pierre Gladeux and Elisabeth Fournier for their constructive comments which enabled improvements to the manuscript. We also thank Laurent Gérard and Nathalie Galet for their technical assistance during the multiplication of yellow rust isolates. This work received support from the European Integrated Project BIOEXPLOIT, FOOD-CT-2005-513959 and EMERFUNDIS, ANR 07-BDIV-003. Sajid Ali was financially supported by a grant from the Higher Education Commission for the Government of Pakistan.

REFERENCES

Anikster Y (1986) Teliospore germination in some rust fungi. *Phytopathology* **76**, 1026-1030.

Arnaud-Haond S, Belkhir K (2007) GenClone 1.0: a new program to analyse genetics data on clonal organisms. *Molecular Ecology Notes* **7**, 15-17.

Bahri B, Leconte M, de Vallavieille-Pope C, Enjalbert J (2009a) Isolation of ten microsatellite loci in an EST library of the phytopathogenic fungus *Puccinia striiformis* f.sp. *tritici*. *Technical note, Conservation Genetics* **10**, 1425-1428.

Bahri B, Leconte M, Ouffroukh A, de Vallavieille-Pope C, Enjalbert J (2009b) Geographic limits of a clonal population of wheat yellow rust in the Mediterranean region. *Molecular Ecology* **18**, 4165-4179.

Bahri B, Shah SJA, Hussain S, Leconte M, Enjalbert J, De Vallavieille-Pope C (submitted) Genetic diversity of wheat yellow rust population in Pakistan and its relationship with host resistance. *Plant Pathology* **Submitted**.

Balloux F, Lehmann L, de Meeu T (2003) The population genetics of clonal and partially clonal diploids. *Genetics* **164**, 1635-1644.

Barrett LG, Thrall PH, Burdon JJ, Nicotra AB, Linde CC (2008) Population structure and diversity in sexual and asexual populations of the pathogenic fungus *Melampsora lini*. *Molecular Ecology* **17**, 3401-3415.

Belkhir KBP, Chikhi L, Raufaste N, Bonhomme F (2004) GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5000: Université de Montpellier II, Montpellier (France) http://www.univ-montp2.fr/~genetix/genetix/genetix.htm.

Bengtsson BO (2003) Genetic variation in organisms with sexual and asexual reproduction. *Journal of Evolutionary Biology* **16**, 189-199.

Brace CL (1963) Structural reduction in evolution. *The American Naturalist* **97**.

Brown JKM (1995) Recombination and selection in populations of plant pathogens. *Plant Pathology* **44**, 279-293.

Burdon JJ, Roelfs AP (1985) Isozyme and virulence variation in asexually reproducing populations of *Puccinia graminis* and *P. recondita* on wheat. *Phytopathology* **75**, 907-913.

Burt A, Carter DA, Koenig GL, White TJ, Taylor JW (1996) Molecular markers reveal cryptic sex in the human pathogen *Coccidioides immitis*. *Proceedings of the National Academy of Science of the United States of America* **93**, 770-773.

Calonnec A, Johnson R, de Vallavieille-Pope C (2002) Genetic analyses of resistance of the wheat differential cultivars Carstens V and Spaldings Prolific to two races of *Puccinia striiformis*. *Plant Pathology* **51**, 777-786.

Chalhoub BA, Thibault S, Laucou V, Rameau C, Hofte H, Cousin R (1997) Silver staining and recovery of AFLP amplification products on large denaturing polyacrylamide gels. *Biotechniques* **22**, 216-220.

Chen X, Line RF, Leung H (1993) Relationship between virulence variation and DNA polymorphism in *Puccinia striiformis*. *Phytopathology* **83**, 1489-1497.

Clifford BC, Clothier RB (1974) Physiologic specialization of Puccinia hordei on barley hosts with non-hypersensitive resistance. *Transactions of the British Mycological Society* **63**, 421-430.

Clutterbuck AJ (1996) Parasexual recombination in fungi. *Journal of Genetics* **75**, 281-286.

Darwin C (1859) *On the origin of species.* John Murray, London.

de Meeûs T, Prugnolle F, Agnew P (2007) Asexual reproduction: genetics and evolutionary aspects. *Cellular and Molecular Life Science* **64**, 1355–1372.

de Vallavieille-Pope C, Huber L, Leconte M, Bethenod O (2002) Preinoculation effects of light quantity on infection efficiency of *Puccinia striiformis* and *P. triticina* on wheat seedlings. *Phytopathology* **92**, 1308-1314.

de Vallavieille-Pope C, Huber L, Leconte M, Goyeau H (1995) Comparative effects of temperature and interrupted wet periods on germination, penetration, and infection of *Puccinia recondita* f.sp. *tritici* and *P. striiformis* on wheat seedlings. *Phytopathology* **85**, 409-415.

Duan X, Tellier A, WAN A, Leconte M, de Vallavieille-Pope C, Enjalbert J (2010) *Puccinia striiformis* f.sp. *tritici* presents high diversity and recombination in the over-summering zone of Gansu-China. *Mycologia* **102**, 44-53.

Eckert CG, Dorken ME, Mitchell SA (1999) Loss of sex in clonal populations of a flowering plant, *Decondon verticillatus* (lythraceae). *Evolution* **53**, 1079-1092.

Enjalbert J, Duan X, Giraud T, Vautrin C, de Vallavieille-Pope C, Solignac M (2002) Isolation of twelve microsatellite loci, using an enrichment protocol, in the phytopathogenic fungus *Puccinia striiformis* f. sp. *tritici*. *Molecular Ecology Notes* **2**, 563-565.

Enjalbert J, Duan X, Leconte M, Hovmøller MS, de Vallavieille-Pope C (2005) Genetic evidence of local adaptation of wheat yellow rust (*Puccinia striiformis* f. sp. *tritici*) within France. *Molecular Ecology* **14**, 2065-2073.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE : a simulation study. *Molecular Ecology* **14**, 2611-2620.

Gavino PD, Smart CD, Sandrock RW, Miller JS, Hamm BP, Lee TY, Davis RM, Fry WE (2000) Implications of sexual reproduction for *Phytophthora infestans* in the United States: Generation of an agressive lineage. *Plant Disease* **84**, 731-735.

Goddard MR, Godfray HCJ, Burt A (2005) Sex increases the efficacy of natural selection in experimental yeast populations. *Nature* **434**, 636-640.

Groth JV, Roelfs AP (1982) Effect of sexual and asexual reproduction on race abundance in cereal rust fungus populations. *Phytopathology* **72**, 1503-1507.

Hovmøller MS, Justesen AF (2007) Appearance of atypical *Puccinia striiformis* f. sp. *tritici* phenotypes in north-western Europe. *Australian Journal of Agricultural Research* **58**, 518-524.

Hovmøller MS, Justesen, A. F., Brown, J. K. M. (2002) Clonality and long-distance migration of *Puccinia striiformis* f.sp. *tritici* in north-west Europe. *Plant Pathology* **51**, 24-32.

Hovmøller MS, Yahyaoui AH, Milus EA, Justesen AF (2008) Rapid global spread of two aggressive strains of a wheat rust fungus. *Molecular Ecology* **17**, 3818-3826.

Ironside JE (2007) Multiple losses of sex within a single genus of *Microsporidia*. *BMC Evolutionary Biology* **7:48 doi:10.1186/1471-2148-7-48**.

Jin Y, Szabo LJ, Carson M (2010) Century-old mystery of *Puccinia striiformis* life history solved with the identification of *Berberis* as an Alternate Host. *Phytopathology* **100**, 432-435.

Kaeuffer R, Coltman DW, Chapuis J-L, Pontier D, Re´ale D (2006) Unexpected heterozygosity in an island mouflon population founded by a single pair of individuals. *Proceeding of Royal Society London Series B-Biological Sciences*.

Kennington WJ, James SH (1997) The effect of small population size on the mating system of a rare clonal mallee, *Eucalyptus argutifolia* (Myrtaceae) *Heredity* **78**, 252-260.

Kim Y, Stephan W (2000) Joint effects of genetic hitchhiking and background selection on neutral variation. *Genetics* **155**, 1415-1427.

Kimura M (1983) *The Neutral Theory of Molecular Evolution* Cambridge University Press.

Langella O (2008) POPULATIONS 1.2.3, Logiciel de génétique des populations (ed. Langella O). Laboratoire Populations, Génétique et évolution, CNRS, Gif-sur-Yvette, France.

Lewontin RC, Krakauer J (1973) Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms. *Genetics* **74**, 175-195.

Little R, Manners J (1969) Somatic recombination in yellow rust of wheat (*Puccinia striiformis*): I. The production and possible origin of two new physiologic races. *Transactions of the British Mycological Society* **53**, 251-258.

Lynch M, Pfrender M, Spitze K, Lehman N, Hicks H, Allen D, Latta L, Ottene M, Bogue F, Colbourne J (1999) The quantitative and molecular genetic architecture of a subdivided species. *Evolution* **53**, 100-110.

Mantel N (1967) The detection of disease clustering and a generalized regression approach. *Cancer Research* **27**, 209-220.

Maughan H, Masel J, Birky CW, Jr., Nicholson WL (2007) The roles of mutation accumulation and selection in loss of sporulation in experimental populations of *Bacillus subtilis*. *Genetics* **177**, 937-948.

Mboup M, Leconte M, Gautier A, Wan AM, Chen W, de Vallavieille-Pope C, Enjalbert J (2009) Evidence of genetic recombination in wheat yellow rust populations of a Chinese oversummering area. *Fungal Genetics and Biology* **46**, 299-307.

McDonald BA, Linde C (2002) Pathogen population genetics, evolutionary potential, and durable resistance. *Annual Review of Phytopathology* **40**, 349-379.

McKay JK, Latta RG (2002) Adaptive population divergence: markers, QTL and traits. *Trends in Ecology & Evolution* **17**, 285-291.

McKee J, Richards AJ (1996) Variation in seed production and germinability in common reed (*Phragmites australis*) in Britain and France with respect to climate. *New Phytologist* **133**, 233-243.

Merilä J, Crnokrak P (2001) Comparison of genetic differentiation at marker loci and quantitative traits. *Journal of Evolutionary Biology* **14**, 892-903.

Michod RE, Bernstein H, Nedelcu AM (2008) Adaptive value of sex in microbial pathogens. *Infection, Genetics and Evolution* **8**, 267-285.

Milgroom MG (1996) Recombination and the multilocus structure of fungal populations. *Annual Review of Phytopathology* **34**, 457-477.

Miller JR, Wood BP, Hamilton MB (2008) Fst and Qst under neutrality. *Genetics* **180**, 1023–1037.

Milus EA, Kristensen K, Hovmøller MS (2009) Evidence for increased aggressiveness in a recent widespread strain of *Puccinia striiformis* f.sp. *tritici* causing stripe rust of wheat. *Phytopathology* **99**, 89-94.

Newton AC, Caten CE, Johnson R (1985) Variation for isozymes and double-stranded RNA among isolates of *Puccinia striiformis* and two other cereal rusts. *Plant Pathology* **34**, 235–247.

Noguchi MTY, N., Fujita Y (2006) Evidence of genetic exchange by parasexual recombination and genetic analysis of pathogenicity and mating type of parasexual recombinants in rice blast fungus, *Magnaporthe oryzae*. *Phytopathology* **96**, 746-750.

O'Gorman CM, Fuller HT, Dyer PS (2009) Discovery of a sexual cycle in the opportunistic fungal pathogen *Aspergillus fumigatus*. *Nature* **457**, 471-474.

Poulson LT, White WB (1969) The Cave Environment. *Science* **165**, 971 - 981.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* **155**, 945-959.

Prout T (1964) Observations on structural reduction in evolution. *The American Naturalist* **98**, 239-249.

Ralph LH, Jr. (1966) Structural reduction through the ldquoprobable mutation effectrdquo. A critique with questions regarding human evolution. *American Journal of Physical Anthropology* **25**, 7-11.

Reed DH (2007) Selection and the rate of loss of genetic variation: Natural selection and genetic diversity. *Heredity* **99**, 1-2.

Regal PJ (1977) Evolutionary loss of useless features: Is it molecular noise suppression? *The American Naturalist* **111**, 123-133.

Samils B, V. S, Lagercrantz U, Lascoux M, Gullberg U (2001) Genetic diversity in relation to sexual and asexual reproduction in populations of *Melampsora larici-epitea*. *European Journal of Plant Pathology* **107**, 871-881.

Spitze K (1993) Population structure in *Daphnia obtusa*: Quantitative genetic and allozymic variation. *Genetics* **135**, 367-374.

Spring O, Zipper R (2006) Evidence for asexual genetic recombination in sunflower downy mildew, *Plasmopara halstedii*. *Mycological Research* **110**, 657-663.

Steele KA, Humphreys E, Wellings CR, Dickinson MJ (2001) Support for a stepwise mutation model for pathogen evolution in Australasian *Puccinia striiformis* f. sp. *tritici*. *Plant Pathology* **50**, 174-180.

Stubbs RW (1985) Stripe rust. In: *The Cereal Rusts, Diseases, Distribution, Epidemiology and Control* (eds. Roelfs AP, Bushnell WR), pp. 61–101. Academic Press, London, UK.

Sujkowski LS, Goodwin SB, Dyer AT, Fry WE (1994) Increased genotypic diversity via migration and possible occurrence of sexual reproduction of Phytophthora infestans in Poland. *Phytopathology* **84**, 201-207.

Taylor JW, Jacobson DJ, Fisher M (1999) The evolution of asexual fungi: reproduction, speciation and classification. *Annual Review of Phytopathology* **37**, 197-246.

Wellings CR (2007) *Puccinia striiformis* in Australia: A review of the incursion, evolution and adaptation of stripe rust in the period 1979-2006. *Australian Journal of Agricultural Research* **58**, 567-575.

Wilkens H (1988) Evolution and genetics of epigean and cave *Astyanax fasciatus* (Characidae, Pisces): support for the neutral mutation theory. *Evolutionary Biology* **23**, 271-367.

Wright RG, Lennard JH (1980) Origin of a new race of *Puccinia striiformis*. *Transactions of the British Mycological Society* **74**, 283-287.

SUPPORTING INFORMATION

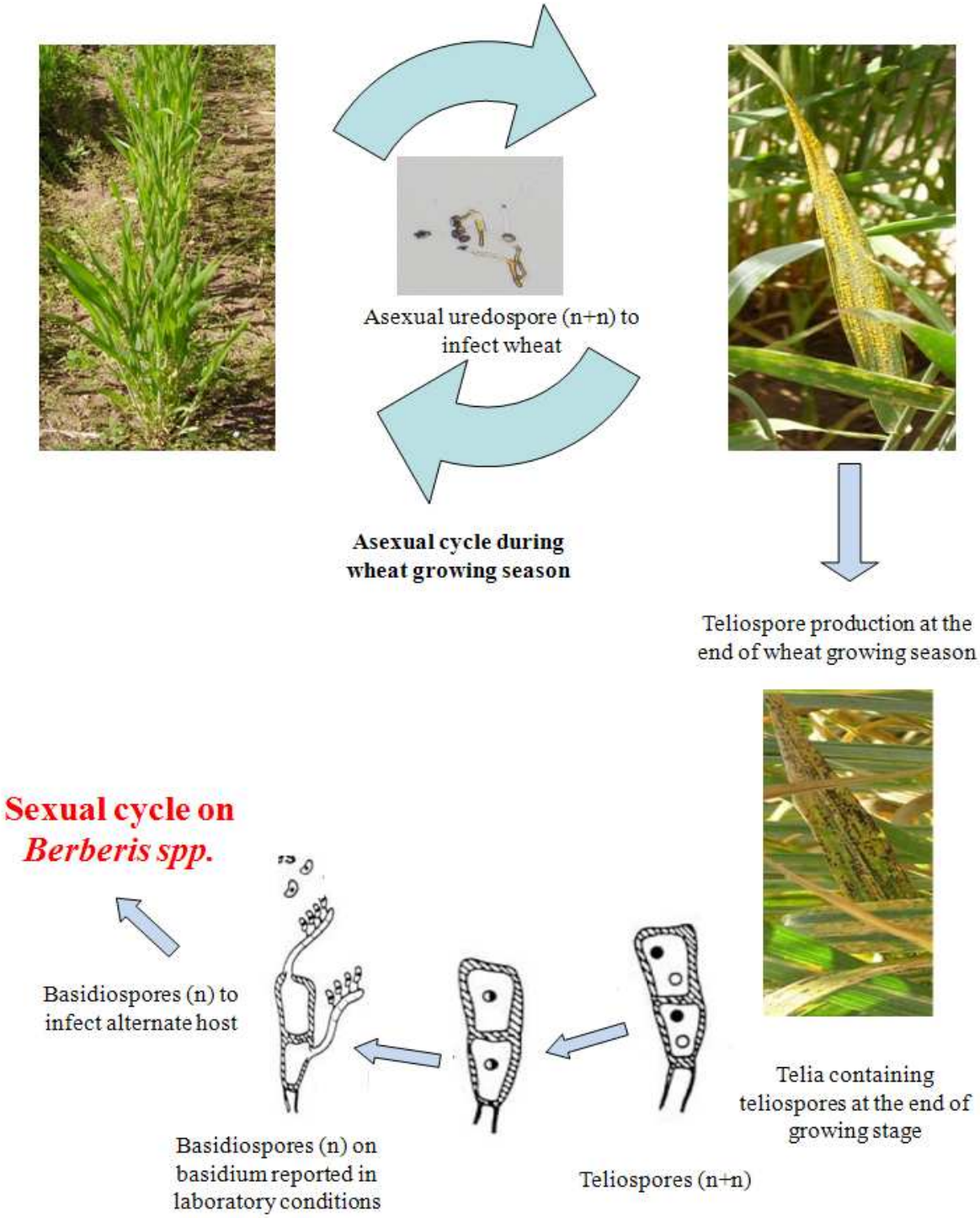


Fig. S1. Life cycle of *Puccinia striiformis* f.sp. *tritici*, showing the asexual cycle (Stubbs, 1985; Jin et al., 2010).

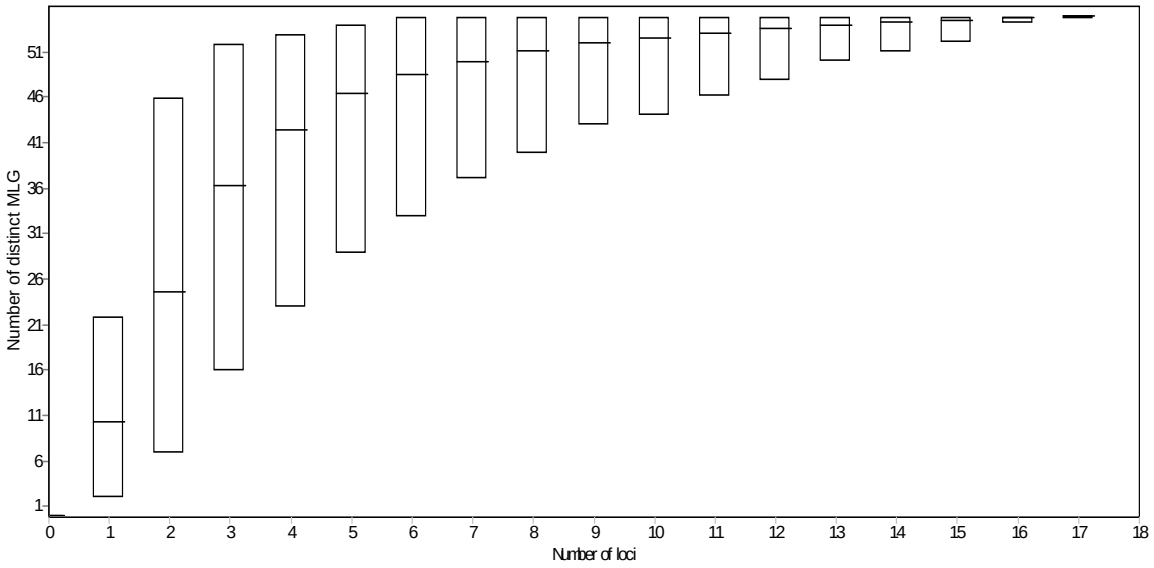


Fig. S2. Number of PST genotypes detected as a function of the number of loci resampled within the 17 microsatellite markers using GENECLONE software. The box represents the average, minimum and maximum numbers of MLSTs detected when resampling on loci.

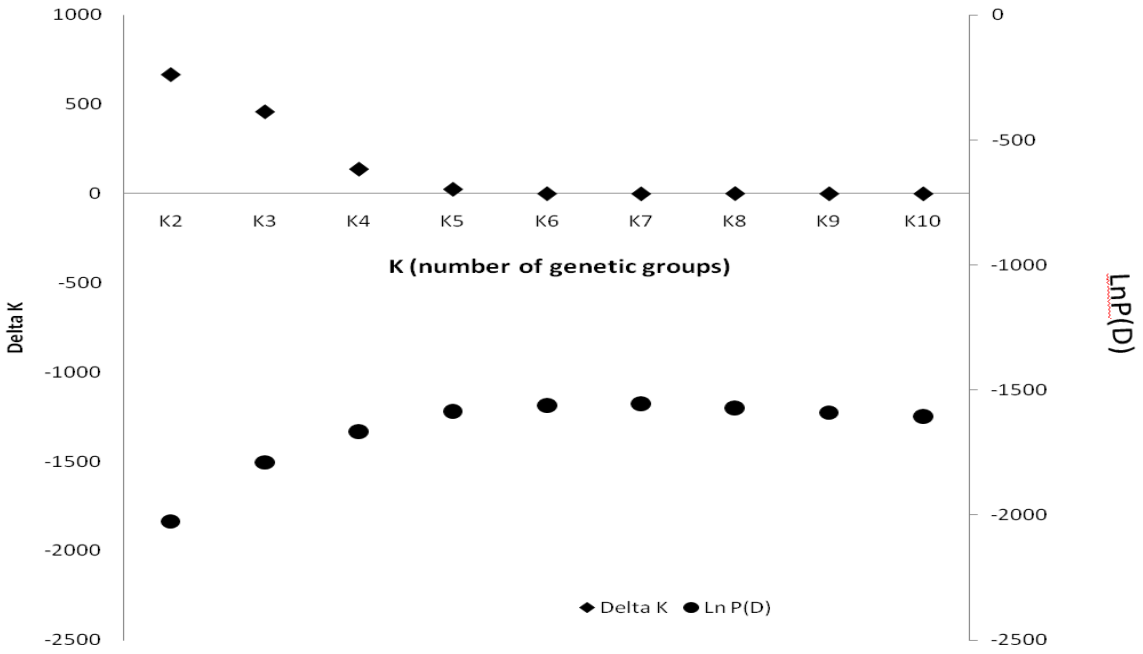


Fig. S3. Distribution of mean Ln P(D), averaged across 20 runs for 10 K (number of populations) values estimated using STRUCTURE software, in order to select the appropriate K value.

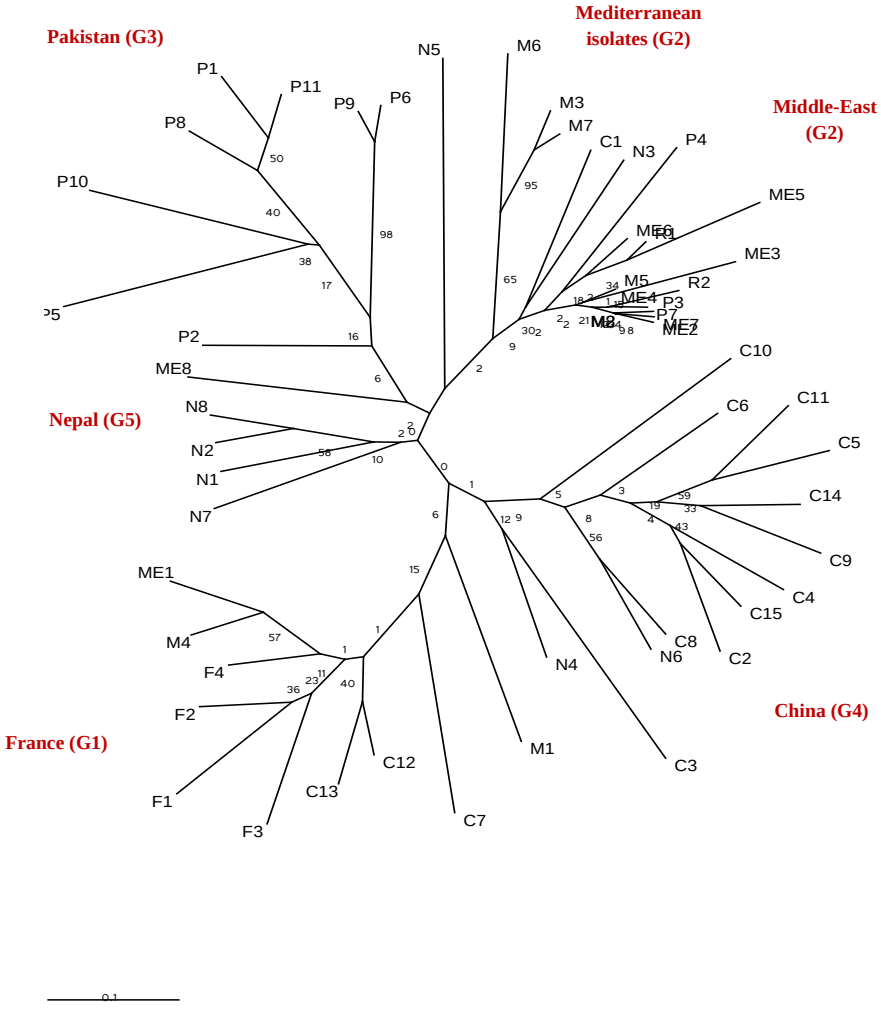


Fig. S4. The phylogenetic tree generated by the POPULATION software (1000 bootstraps) for 56 PST isolates from five genetic groups of six geographical origins. The low bootstraps values could result due to the inter-group hybrid isolates and the existence of group specific rare alleles.

Table S1. Estimates of Fst (upper diagonal) and its significance (lower diagonal) based on 17 SSR loci for PST isolates from six geographical origins.

	Northern France	**Mediterranean**	**Middle-East**	**Pakistan**	**China**	**Nepal**
Northern France	-	0.2340	0.2413	0.3313	0.1786	0.2757
Mediterranean	0.01	-	0.0000	0.1357	0.1873	0.1332
Middle-East	0.02	**0.42**	-	0.1186	0.1988	0.1054
Pakistan	0.00	0.01	0.01	-	0.2521	0.0920
China	0.00	0.00	0.00	0.00	-	0.1097
Nepal	0.00	0.00	0.00	0.00	0.00	-

CHAPTER-III

The role of sexual recombination and over-summering in temporal maintenance of *Puccinia striiformis* f.sp. *tritici* at its centre of diversity, the Himalayan region of Pakistan

Sajid Ali, Pierre Gladieux, Hidayatur Rahman, Muhammad Shahab Saqib, Muhammad Fiaz, Habib Ahmed, Marc Leconte, Angélique Gautier, Annemarie Fejer Justesen, Mogens Støvring Hovmøller, Jérôme Enjalbert and Claude de Vallavieille-Pope

The role of sexual recombination and over-summering in temporal maintenance of *Puccinia striiformis* f.sp. *tritici* at its centre of diversity, the Himalayan region of Pakistan

The article is under review by *Molecular Ecology*

The role of sexual recombination and over-summering in temporal maintenance of *Puccinia striiformis* f.sp. *tritici* at its centre of diversity, the Himalayan region of Pakistan

Sajid Ali*[1], Pierre Gladieux[2], Hidayatur Rahman[3], Muhammad Shahab Saqib[3], Muhammad Fiaz[4], Habib Ahmed[4], Marc Leconte[1], Angélique Gautier[1], Annemarie Fejer Justesen[5], Mogens Støvring Hovmøller[5], Jérôme Enjalbert[6] and Claude de Vallavieille-Pope[1]

[1] INRA UR 1290 BIOGER-CPP, BP01, 78850 Thiverval-Grignon, France

[2] Department of Plant and Microbial Biology, University of California, Berkeley, CA 94720-3102, USA

[3] Department of Plant Breeding and Genetics, KPK Agricultural University Peshawar, Pakistan

[4] Hazara University Mansehra, Pakistan

[5] Department of Agroecology, Aarhus University, Flakkebjerg, DK-4200 Slagelse, Denmark

[6]INRA UMR 320 Génétique Végétale, Ferme du Moulon, 91190 Gif sur Yvette, France

Keywords: invasive fungi, wheat yellow rust, population dynamics

*Corresponding author (sajid.ali@versailles.inra.fr)

Running title: Fungal temporal maintenance at centre of diversity

ABSTRACT

Understanding the temporal maintenance of fungal pathogens in their centre of diversity is crucial for comprehending the fungal ecology in the context of invasion. The centre of diversity of *Puccinia striiformis* f.sp. *tritici* (PST) has recently been reported in the Himalayan region, with Pakistan and China harbouring ancestral to all other worldwide populations. We investigated the between season "temporal" maintenance of PST in the Himalayan region of Pakistan through microsatellite genotyping of 684 isolates from 14 locations in 2010 and 2011. High genotypic and pathotypic

diversity, recombinant population structure across all locations and shallow population subdivision was observed. The genotypes were assigned to four genetic groups in a complex mosaic, with a clear subdivision only between a zone with no alternate host (*Berberis*), and the *Berberis* zone. The non-*Berberis* zone population was closely related to the diverse over-summering population of *Berberis* zone. This over-summering population seems to serve as a source for the non-*Berberis* zone. Despite the lack of any differentiation in populations across two sampling years (non-significant F_{ST}), only two multilocus genotypes (MLGs) at the maximum were resampled over both years at a given location, suggesting a limited across-year survival through clonality. A genotyping of infected *Berberis* leaves confirmed the presence of *Puccinia striiformis*; however, the *formae specialis* needs to be confirmed. Our results thus suggest the contribution of both sexual recombination (in the *Berberis* zone) and over-summering survival (in the non-*Berberis* zone) to the temporal maintenance of PST in the Himalayan region of Pakistan.

INTRODUCTION

Knowledge of the population structure and temporal maintenance of fungal pathogens in ecological context is critical for the development of efficient disease management strategies. Due to their role in recent disease emergences, fungi are being increasingly recognized as a major threat to food security and ecosystem health (Dean *et al.*, 2012; Fisher *et al.*, 2012; Gladieux *et al.*, 2011; Simwami *et al.*, 2011). Modern molecular epidemiology tools can help better comprehend fungal ecology as they can (i) provide insights into pathogen transmissibility, the longevity of populations and their maintenance over seasons, (ii) allow the identification of infection reservoirs, and (iii) enable prediction of the evolutionary potential of extant populations (McDonald, Linde, 2002; Milgroom, Peever, 2003; Taylor, Fisher, 2003). An accurate understanding of population structure is particularly important in the centre of diversity of pathogens, as the zones with highest diversity represent a significant threat, acting as reservoirs for rare alleles involved in pathogenicity, drug-resistance or adaptive response to abiotic factors. Individuals with such advantageous alleles have an increased potential for invading

new host populations or new territories. The importance of such invasions becomes even more crucial for evolutionary trajectories of pathogens lacking genetic diversity outside of these diversity zones (e.g. pathogen with clonal population structure), especially if the pathogen has long distance migration capacity (Brown, Hovmøller, 2002).

Wheat yellow rust is an economically important wheat disease of all wheat-growing regions with temperate to mild-cold climate (Ali *et al.*, 2009b; de Vallavieille-Pope *et al.*, 2012; Hovmøller *et al.*, 2011; Hovmøller *et al.*, 2010; Singh *et al.*, 2004). The disease is caused by a basidiomycete fungus, *Puccinia striiformis* f.sp. *tritici* (PST), which has a long distance migration capacity (Brown, Hovmøller, 2002). The pathogen remains important in the context of invasions and re-emergence (Hovmøller *et al.*, 2008; Singh *et al.*, 2004; Wellings, 2007), however, our understanding of its biogeography and colonization history was limited until very recently. Population genetics surveys revealed a clonal population structure in North-western Europe, USA, Australia and Mediterranean region (Bahri *et al.*, 2009; Chen *et al.*, 1993; Enjalbert *et al.*, 2005; Hovmøller *et al.*, 2002; Wellings, 2007), and a high diversity in China (Duan *et al.*, 2010; Mboup *et al.*, 2009) and Pakistan (Bahri *et al.*, 2011). In many countries, PST reproduces exclusively asexually on its main (uredial) host, wheat, yielding a highly clonal population structure. In these regions, summer and autumn survival of PST is assumed to be dependent on volunteers or self-sown wheat and, to a lesser extent, on other grass species (Ali *et al.*, 2010; de Vallavieille-Pope *et al.*, 2012; Enjalbert *et al.*, 2005). New epidemics can also be initiated by migration from regions where PST is capable of over-summering (Brown, Hovmøller, 2002). Differences in wheat maturation between regions of contrasted altitudes can allow an overlapping period between the lowland sowing and highland harvest periods, creating a "green bridge" (Brown, Hovmøller, 2002; Wan *et al.*, 2004). The sexual phase of the life cycle of PST has only recently been documented on *Berberis spp*. under laboratory condition (Jin *et al.*, 2010), ending decades of unfruitful search (Stubbs, 1985). In Pucciniales like stem and leaf rust (*Puccinia graminis* or *Puccinia triticina*), telia are produced at the end of season, and remain dormant during the inter-

cropping season. In early spring, the teliospores germinate to form haploid basidiospores, which then infect an alternate ("aecial") host to complete the sexual cycle. After a complex process, recombinant dikaryotic aeciospores are produced in late spring on the alternate host, responsible for the production of an initial inoculum re-infecting the uredial host (wheat) and initiating new epidemics (Fig.1).

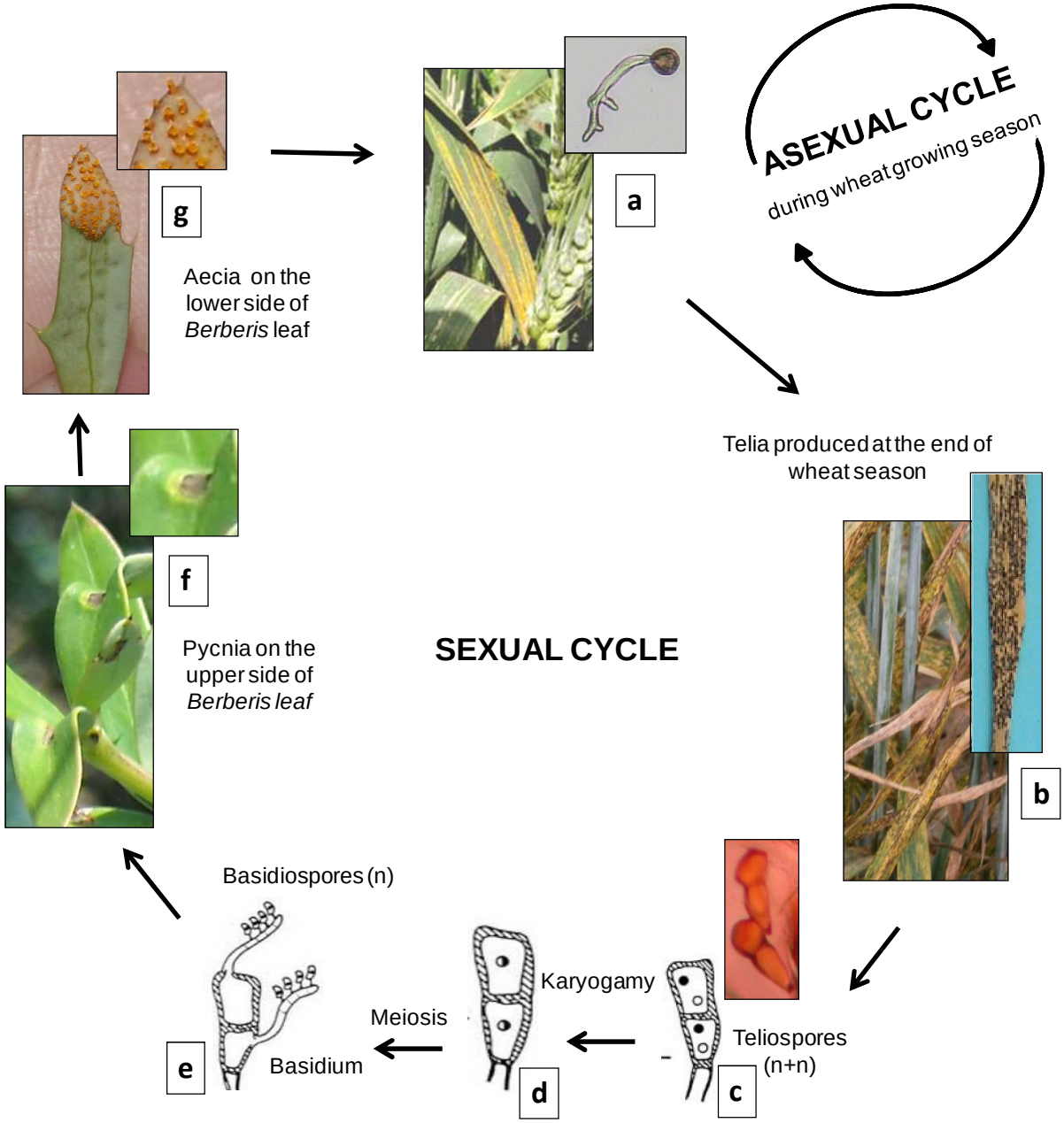


Fig. 1. Life cycle of *Puccinia striiformis*. It passes its epidemic stage on wheat through several asexual cycles in the form of uredinial stage (a), producing telia (b) at the end of wheat season. The telia contains teliospores (c), which passes through karyogamy (d) and meiosis to produce a basidium developing into four basidiospores (e). The basidiospores infect the alternate host (*Berberis* spp.) to produce the disease symptoms at the pycnial (f) and aecial (g) stages on the upper and lower leaf surface, respectively, as observed on *Berberis* spp. at Pakistan.

Sexual reproduction has been suspected in populations of PST that exhibit footprints of recombination in their genetic structure, such as in China, Nepal or Pakistan (Ali *et al.*, in preparation-a; Bahri *et al.*, 2011; Duan *et al.*, 2010; Mboup *et al.*, 2009), but the effective role of an alternate host species in the biology of a PST population has not yet been identified. A detailed study using a comprehensive set of samples collected worldwide identified the centre of diversity of the pathogen in the Himalayan region ranging from Pakistan to Nepal, and confirmed the recombining population structure of PST in this region and its clonal population structure in other parts of the world (Ali *et al.*, in preparation-a). Coalescent-based analyses of microsatellite variation also pinpointed populations from Pakistan and China as the most likely ancestral source of worldwide PST populations (Ali *et al.*, in preparation-a). Pakistan is thus a good candidate area for describing PST populations involving, at least partly, the sexual cycle on an alternate host. The identification of PST through molecular genotyping of rust infected *Berberis* leaves would be a first step to demonstrate the role of sexual reproduction on *Berberis*. However, the biological test for completing the life cycle would be necessary to be more conclusive.

PST has remained an economically important wheat disease throughout the Pakistan, particularly in the northern Himalayan region (Ali *et al.*, 2009b; Rizwan *et al.*, 2010; Singh *et al.*, 2004) with genetic resistance being used as the most economical and environment friendly control measure. As successive epidemics resulted from the breakdown of the major yellow rust resistance genes that were deployed, mainly *Yr2*, *Yr9* and *Yr27* (Duveiller *et al.*, 2007; Singh *et al.*, 2004), efforts have also been made for the deployment of minor genes based partial resistance (Ali *et al.*, 2007; Ali *et al.*, 2009a; Shah *et al.*, 2009). The pathogen, however, has proven capability of overcoming such partial resistance as well (Huerta-Espino, Singh, 1996; Park, McIntosh, 1994). This capacity of the pathogen to adapt to host resistance is likely a result of high within population diversity and sexual reproduction by the pathogen (Giraud *et al.*, 2010; Giraud *et al.*, 2006; Gladieux *et al.*, 2011; McDonald, Linde,

2002; Sobsey, Pillai, 2009; Stukenbrock, McDonald, 2008) as well as if the populations are maintained stably over time.

Several hypotheses have been proposed regarding the temporal maintenance of PST in Pakistan. Over-summering of PST in the Himalayan region on summer wheat crop has been suggested as a source for re-establishment in the winter wheat cropping regions (Bahri *et al.*, 2011; Chaube, Pundhir, 2005; Hassan, 1968). Several lines of evidence suggest a possible role of the sexual reproduction in the temporal maintenance of PST in the Himalayan region of Pakistan: i) the existence of a high genotypic diversity and recombinant population structure (Ali *et al.*, in preparation-a; Bahri *et al.*, 2011), ii) a high sexual reproduction ability, estimated from telial production capacity (Ali *et al.*, 2010), and iii) the presence of the supposed alternate host, *Berberis* spp., in the region (Perveen, Qaiser, 2010; Ray *et al.*, 2011). Others suggested repeated migrations from the Middle-East and central Asia to establish the PST populations in Pakistan, exemplified by the incursion of pathotypes with virulence "*V9*" (Singh *et al.*, 2004).

The temporal maintenance of PST can be investigated by comparing a series of temporally spaced population samples to assess the role of over-summering, migration and/or sexual reproduction. The temporal maintenance of isolates through volunteer plants would result in resampling of clonemates, in direct relation with the effective size of the population (Ali *et al.*, In preparation-b; Burt *et al.*, 1996). Effective population size itself will mostly depend on the number of individuals establishing the new epidemics, and therefore on the efficiency of over-summering clonal and/or sexual survival. Presence of a "green bridge" would prevent a strong bottleneck, and allow the maintenance of numerous clonal lineages, while survival through sexual reproduction would re-shuffle alleles into new genotypes, and break up clonal lineages. Thus temporal shifts in population along with the comparison with over-summering populations would provide a description of the temporal

maintenance of PST in Pakistan. Confirmation of the presence of yellow rust on *Berberis* leaves will support the role of sexual reproduction in PST temporal stability.

The present study was based on samples collected on wheat cultivars in the *Berberis* and non-*Berberis* zones of the Himalayan Mountains in Northern Pakistan, where wheat is cultivated from November to May. A single sample was also collected from the summer wheat growing region with wheat season from June to September. Aecia infected *Berberis* leaves were also collected from different locations of *Berberis* zone. Multilocus microsatellite typing of a comprehensive set of isolates was used i) to investigate the existence of recombination and the geographic distribution of genetic and pathotypic variability; ii) to assess the existence of spatial population subdivision; iii) to compare populations sampled at the same place in two consecutive years to infer temporal maintenance of PST in these areas; iv) to confirm the presence of yellow rust on *Berberis* plants near wheat fields in the Himalayan region of Pakistan; v) and to test whether summer wheat can serve as a "green-bridge" that maintains inoculums during the off-season.

MATERIALS AND METHODS

Sampling, multiplication and molecular genotyping

PST isolates were collected in the wheat-growing region in the Himalayan Mountains of Pakistan, ranging from Khyber in the Hindu-Kush at the frontier of Afghanistan in the west to Skardu at the frontier of Kashmir in the East (Fig. 2 and Table 1). First sampling was done from 10 different locations in April 2010 for all regions (wheat growing season: November-May) except for Skardu, in the summer wheat growing region, which was sampled in August 2010. This site was central for testing the existence of a "green-bridge", where summer wheat fields would ensure the over-summering of the PST populations. In 2011, a second sampling was done at 11 locations, eight of which were the same as in 2010. Thus a total of 14 locations were included in the present study.

Berberis spp. were absent in only two lowland locations (Peshawar and Nowshera, termed as "non-*Berberis* zone"), while *Berberis* spp. were abundant in all other locations (termed as "*Berberis* zone"), including the over-summering area. *Berberis* spp. leaves infected with rust were collected from nine locations in 2011 to confirm whether *Puccinia striiformis* was present.

DNA was extracted from either 5 mg of spores or from a single lesion from infected wheat leaves following Ali et al. (2011). A set of 20 microsatellites amplified in three multiplex PCR reactions was used for molecular genotyping. PCR products were analysed with a Beckman Coulter CEQ-8000 DNA Analyzer, allowing to record fragment lengths for each allele (Ali *et al.*, 2011). From a subset of eight *Berberis* infected leaves, isolated clusters of aecial cups were cut out of the leaf for DNA extraction and microsatellite genotyping using the same procedure.

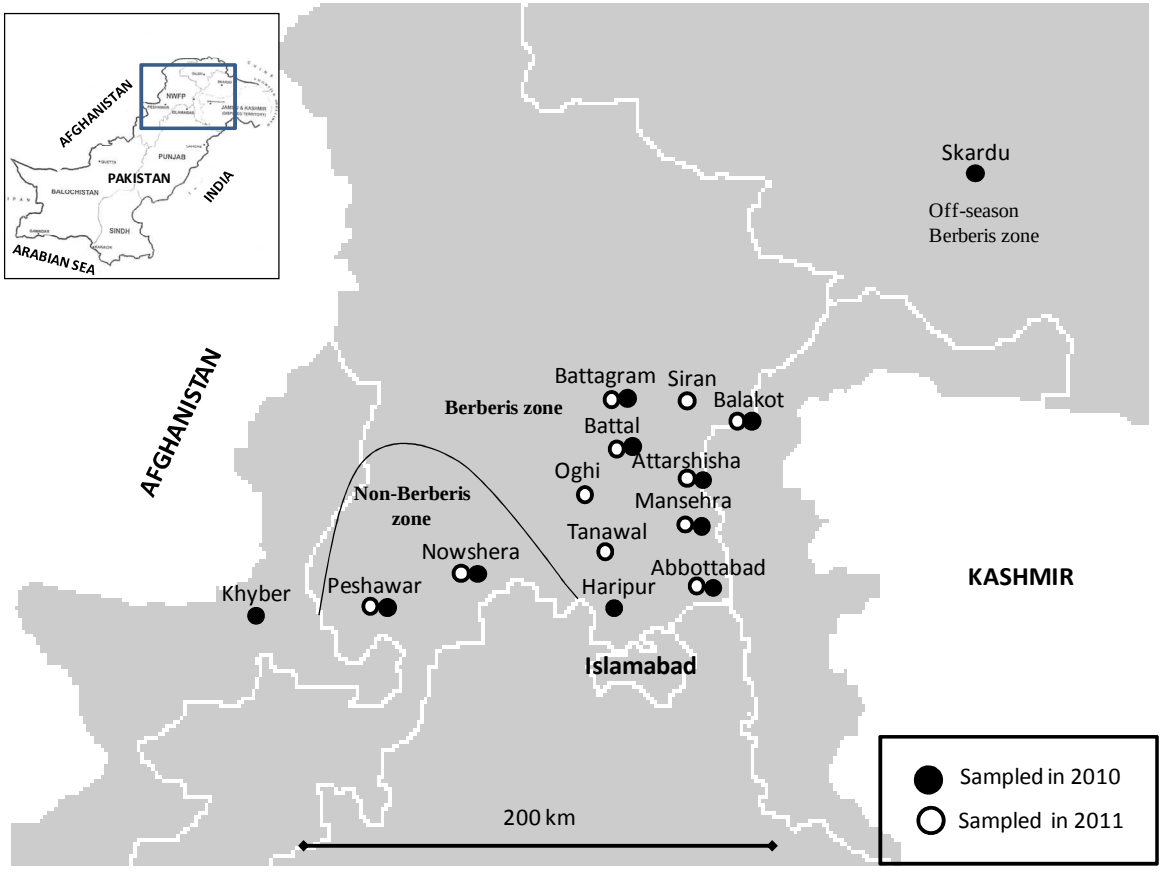


Fig. 2. Sampling locations of *Puccinia striiformis* (PST) populations from the Himalayan region of Pakistan surveyed during 2010 and 2011 to understand PST population structure and temporal maintenance at its centre of diversity.

Analyses of population structure

We first checked the suitability of the markers for detecting multilocus genotypes (MLGs) under panmixia by plotting the number of MLGs detected against the number of loci using GENECLONE (Arnaud-Haond, Belkhir, 2007). The non-redundancy of the set of loci was also investigated by estimating linkage disequilibrium among marker pairs using GENETIX 4.05.2 (Belkhir *et al.*, 2004).

Observed (Ho) and unbiased expected heterozygosity (He) were computed using GENETIX 4.05.2 software (Belkhir *et al.*, 2004). The null hypothesis of Hardy-Weinberg equilibrium within each population was tested using the exact test implemented in GENEPOP 4.0 (Raymond, Rousset, 1995). Within population diversity was studied using allelic richness and gene diversity calculated using FSTAT 2.9.3 (Goudet, 2001) and genotypic diversity was calculated using MULTILOCUS (Agapow, Burt, 2001). Calculations were performed both on the full dataset and on a clone-corrected dataset (i.e., a dataset in which only one representative of each repeated MLG is kept) to avoid the bias resulting from epidemic clonal amplification (Maynard-Smith *et al.*, 1993).

Spatial population structure was assessed using both model-based Bayesian and non-parametric multivariate clustering methods. We used the model-based Bayesian method implemented in the STRUCTURE 2.2 software (Pritchard *et al.*, 2000), which assigns multilocus genotypes into clusters while minimizing the Hardy-Weinberg disequilibrium and gametic phase disequilibrium between loci within clusters. A total of 20 independent runs was made with a burn-in period of 100,000 and a total of 200,000 iterations, selecting *K* ranging from 1 to 20. STRUCTURE outputs were processed with CLUMPP (Jakobsson, Rosenberg, 2007) to assign groups of runs to a common clustering pattern (G'-statistic greater than 80%). The optimal *K* value was determined using the method of Evanno et al. (2005) based on the rate of change in the log probability of data between successive *K* values.

As STRUCTURE can overestimate the number of clusters when there is relatedness among some genotypes (e.g. due to asexual reproduction; Gao *et al.*, 2007) we also analysed population subdivision using a non-parametric approach, which does not rely on a particular genetic model. We used discriminant analysis of principal components (DAPC) implemented in ADEGENET package (Jombart *et al.*, 2010), in *R* environment. For DAPC, the number of clusters was identified based on the Bayesian Information Criterion (BIC) as suggested by Jombart *et al.*(2010). The relationships among populations were represented using a neighbour-joining (NJ) tree based on Nei's genetic distance, as implemented in the POPULATION software (Langella, 2008). Significance was assessed using 1000 bootstraps. The extent of differentiation between population pairs was assessed using pairwise F_{ST}, with 1000 random permutations to test for significance (GENETIX 4.05.2; Belkhir *et al.*, 2004). Additionally, analyses of molecular variance (AMOVA) were carried out with ARLEQUIN 3.1 (Excoffier *et al.*, 2005) to estimate the distribution of molecular variation within location, across locations and over years.

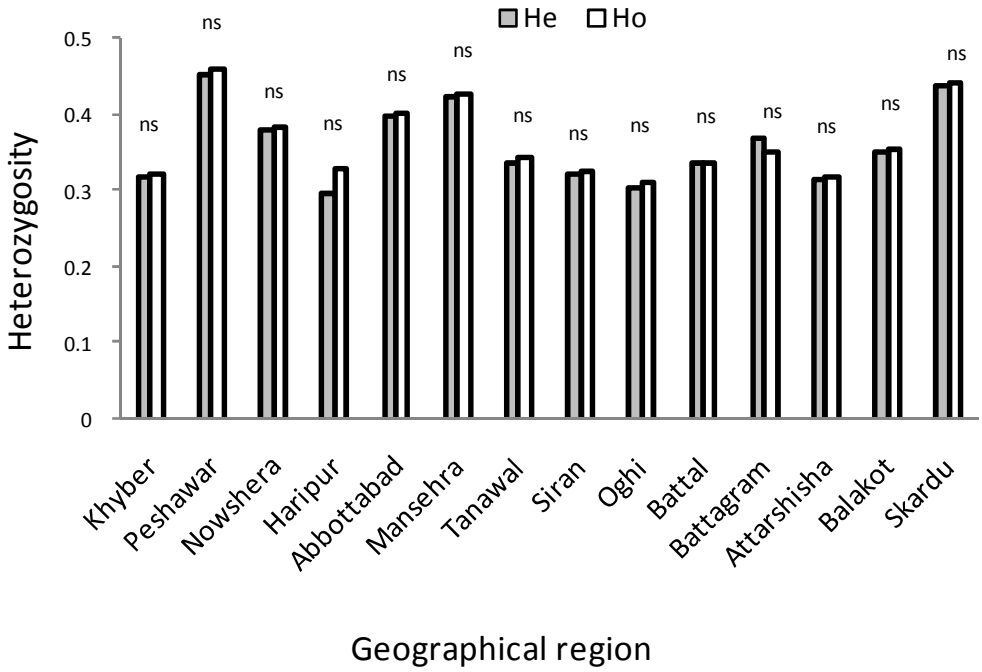


Fig. 3. Mean expected (He) and observed (Ho) heterozygosity over 20 microsatellite markers for *Puccinia striiformis* f.sp. *tritici* populations from the Himalayan region of Pakistan.

Characterization of virulence profiles

Virulence profiles for a subset of 127 isolates were determined to assess the pathotypic diversity of the Himalayan PST population and to compare it with the previously characterized world and Pakistani isolates. Each of the 127 isolates was multiplied from a single lesion on the susceptible cultivars Cartago or Michigan Amber in a fully confined S3-type greenhouse as described previously (Bahri *et al.*, 2011; de Vallavieille-Pope *et al.*, 2002). Virulence profiles were determined using a set of 36 differential lines from the world, European and Chinese sets along with 17 Avocet wheat lines using a qualitative infection-type scale of 0-9, as described previously (de Vallavieille-Pope *et al.*, 2012; de Vallavieille-Pope *et al.*, 1990; Johnson *et al.*, 1972).

Temporal maintenance of PST populations

The temporal maintenance of populations was investigated by examining temporal shifts in genetic composition at the eight locations sampled both in 2010 and 2011. Temporal shifts were assessed by measuring F_{ST} between the two years, inspecting the resampling of multilocus genotypes and estimating the effective population size (based on temporal variance of allele frequencies: Waples, 1989). Additionally "historic" estimates, measuring the long-term effective population size from one single sample, were estimated for all populations. Historic estimates were made based on two linkage disequilibrium methods i.e., the method of Pudovkin et al. (1996) implemented in Ne-ESTIMATOR software and the method of Waples and Do (2010) implemented in LDNe software. The over-summering wheat population was compared with the other 13 locations, to assess whether some genotypes were found in common, suggesting the presence of a "green-bridge". In that case, the 2010 samples would be considered as potential sources of inoculum for the over-summering area sampled in late 2010, while 2011 locations would reversely be considered as recipient populations.

Table 1. Gene and genotypic diversity indices for *Puccinia striiformis* populations for geographically spaced populations from the Himalayan region of Pakistan, sampled in 2010 and 2011.

				Genotypic diversity				Gene diversity		
Year	**Location**	**Presence of *Berberis* spp.**	**Number of individuals**	**MLGs**	**Genotypic diversity**	**Most abundant MLG's frequency**	**Linkage disequilibrium (Rd)**	**Gene diversity**	**No. of alleles**	**Allelic Richness**
2010	Khyber	Yes *	35	25	0.971	5	0.31	0.323	3.100	2.435
	Peshawar	No	12	9	0.939	3	0.303	0.422	2.650	2.575
	Nowshera	No	19	8	0.673	11	0.701	0.356	2.400	2.200
	Haripur	Yes *	42	20	0.864	15	0.33	0.301	3.050	2.360
	Abbottabad	Yes	34	31	0.993	3	0.112	0.333	3.200	2.401
	Mansehra	Yes	28	27	0.997	2	0.174	0.349	3.200	2.531
	Battal	Yes	46	33	0.967	8	0.305	0.328	3.300	2.489
	Battagram	Yes	9	9	1	1	0.092	0.341	2.500	2.500
	Attarshisha	Yes	30	22	0.966	5	0.207	0.284	2.600	2.148
	Balakot	Yes	24	18	0.975	3	0.152	0.368	3.100	2.600
	Skardu	Yes (over-summering wheat)	63	54	0.991	5	0.115	0.441	3.600	2.798
2011	Peshawar	No	33	23	0.973	4	0.294	0.447	3.350	2.818
	Nowshera	No	16	10	0.825	7	0.609	0.337	2.150	2.079
	Abbottabad	Yes	18	9	0.706	10	0.593	0.386	2.450	2.322
	Mansehra	Yes	61	54	0.991	6	0.273	0.449	3.350	2.892
	Tanawal	Yes	22	19	0.983	3	0.249	0.345	3.050	2.629
	Siran	Yes	51	33	0.939	11	0.199	0.336	3.000	2.382
	Oghi	Yes	19	16	0.977	3	0.16	0.312	2.250	2.152
	Battal	Yes	35	30	0.99	3	0.126	0.319	2.450	2.198
	Battagram	Yes	20	17	0.984	2	0.257	0.373	3.050	2.700
	Attarshisha	Yes	40	31	0.953	9	0.174	0.302	3.100	2.290
	Balakot	Yes	27	21	0.977	3	0.178	0.328	2.750	2.306

* *Berberis* spp. present only on the surrounding mountains in close vicinity to the fields.

RESULTS

Summary of genetic variation

Multilocus microsatellite genotyping of 684 isolates of PST sampled from 14 locations in the Himalayan region of Pakistan resulted in the detection of 470 multilocus genotypes (232 MLGs in 2010 and 238 in 2011). The full set of markers was sufficient to discriminate different clonemates, as shown by the saturating curve of MLGs plotted against marker number (Fig. S1, supporting information). There was no redundancy among markers, as revealed by the overall non-significant linkage disequilibrium among microsatellites (data not shown).

Microsatellite genotyping of the eight infected leaves of *Berberis* spp. from the *Berberis* zone of Pakistan amplified the same alleles as amplified in *Puccinia striiformis* in our collections. The alleles found were mostly the same as observed in the wheat infecting PST collection of Pakistan, but with some additional alleles not observed before. These samples presented three or four alleles at some SSR loci, which was explained by the presence different genotypes within the same aecial lesion.

Geographical patterns of genetic variability

All populations revealed a clear signature of recombination and high genotypic diversity. The difference between observed (Ho) and expected (He) heterozygosity was non-significant across all geographically spaced populations, consistent with expectations under Hardy-Weinberg equilibrium (Fig. 3). Genotypic diversity was high at all locations for both sampling years, with the minimum value of a 0.673 for the non-*Berberis* zone population of Nowshera in 2010. All locations from the *Berberis* zone had a genotypic diversity above 0.90, except for Abbottabad in 2011 (0.71). Populations from *Berberis* zone had higher gene diversity, number of alleles and allelic richness than the non-*Berberis* zone populations (Table 1).

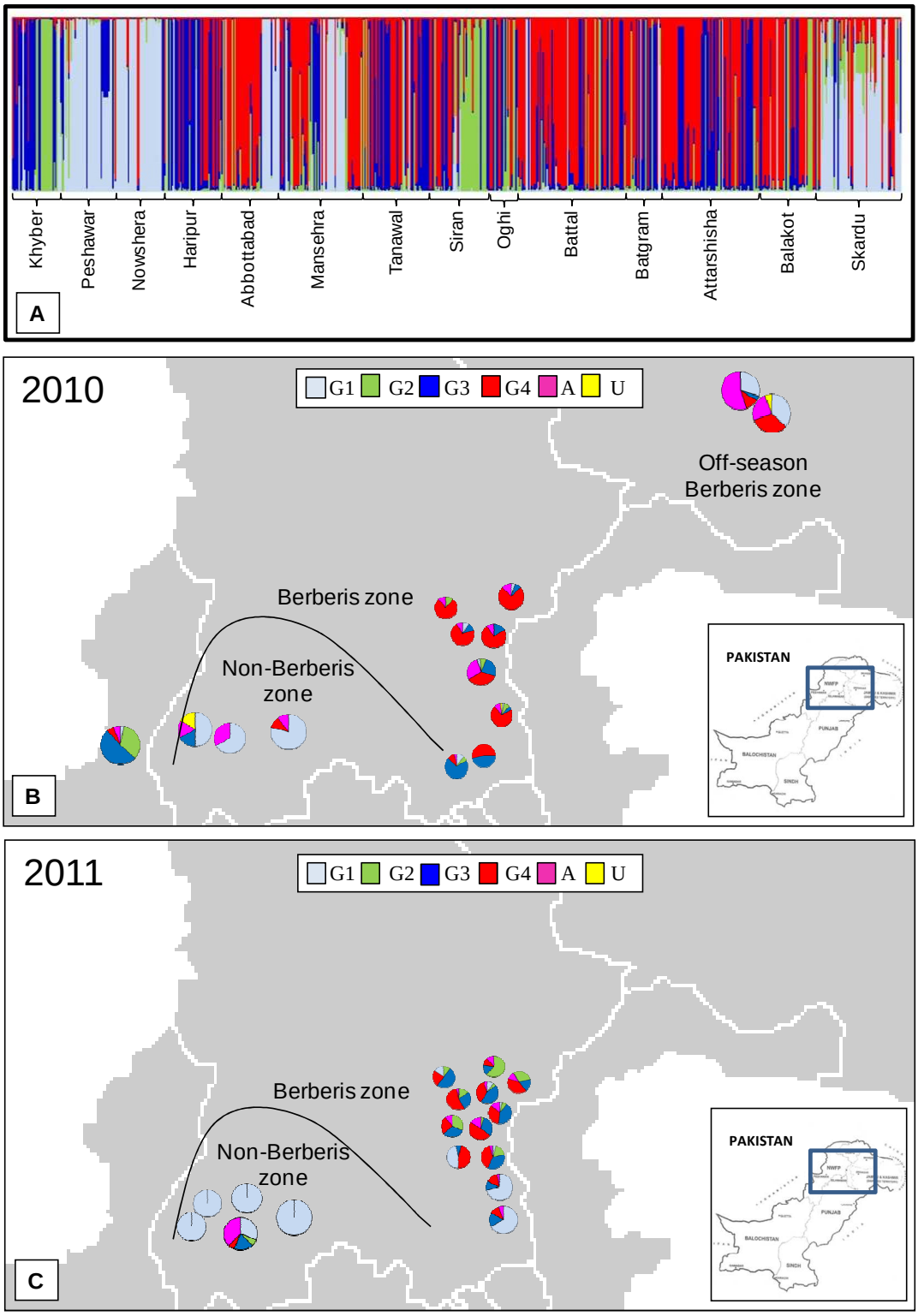


Fig. 4. Assignment analyses of 684 *Puccinia striiformis* f.sp. *tritici* isolates from 14 different locations from the Himalayan region of Pakistan. A = Assignment of isolates using STRUCTURE software. B and C = Map of spatial distribution of genotypes based on their membership coefficients in 4 genetic groups for sampling years of 2010 and 2011. Isolates with at least 40% assignment to two different groups were termed as "admixed (A)", while those not assigned to any specific group were considered as "unassigned (U)".

Population subdivision

Clustering analyses with the STRUCTURE software clearly identified population subdivision in the Pakistani samples. The clustering of isolates was assessed for *K* ranging from 2 to 10 (Fig. S2), with *K*=4 as the optimal number of clusters (Evanno *et al.*, 2005: Fig. S3). *K*=2 separated the non-*Berberis* zone (Peshawar and Nowshera) from the rest of the isolates with a mixed assignment for the over-summering population from Skardu (Fig. S2). At *K*=4, two groups were discriminated in the *Berberis* zone, respectively in the Siran Valley and in Haripur, with substantial resampling in other areas (Fig. 4).

Non-parametric analyses mostly confirmed STRUCTURE analyses (Fig. S4). The overall data analyses could separate only three populations i.e., one group comprised of the two non-*Berberis* zone and over-summering *Berberis* zone populations together, another group of Siran valley population and a third group with the rest of the *Berberis* zone (Fig. S4). Similarly, the described population subdivision was further confirmed by pairwise F_{ST} estimates, with a strong differentiation of non-*Berberis* zone populations from all other locations except the over-summering population from the *Berberis* zone (Skardu: Table 2). Although pairwise F_{ST} values were significant between most of the populations, the level of differentiation was lower among populations belonging to the *Berberis* zone. The AMOVA revealed that 86% of variation was present within populations and 14% of the variation was due to the between groups, while there was no variation due to sampling years (Table S1).

Comparison of temporally spaced populations and effective population size estimation

A non-significance differentiation was found between years for the eight locations sampled in both 2010 and 2011 (non-significant F_{ST}, mean 0.0981 ± 0.0618 (SD), Table 3) except for Abbottabad (0.2375) and Attarshisha (0.1359). The number of multilocus genotypes (MLGs) resampled through both years ranged from 0 to 2, suggesting a limited across year survival of populations through

clonality (Table 3). The temporal effective population size was the lowest at Abbottabad (9.1), while for the rest of locations it ranged from 20 for Nowshera, to 35.2 for Battagram (Table 3). Estimates based on linkage disequilibrium resulted in low values of "historic" effective population size ranging from 0.3 (Nowshera-2011) to 8.2 (Battagram-2010).

Table 2. Divergence of georgraphically spaced *Puccinia striiformis* populations from the Himalayan region of Pakistan in terms of Fst (upper diagonal) and its significance (lower diagonal) based on 20 microsatillite loci.

	Khyber	**Peshawar**	**Nowshera**	**Haripur**	**Abbottabad**	**Mansehra**	**Tanawal**	**Siran**	**Oghi**	**Battal**	**Battagram**	**Attarshisha**	**Balakot**	**Skardu**
Khyber	-	0.233	0.345	**0.037**	0.078	0.039	0.062	0.028	**0.018**	0.072	0.020	0.059	0.040	0.204
Peshawar	0.000	-	0.030	0.225	0.120	0.145	0.178	0.263	0.221	0.243	0.210	0.240	0.223	**0.063**
Nowshera	0.000	0.000	-	0.328	0.184	0.225	0.280	0.366	0.342	0.323	0.311	0.336	0.311	0.112
Haripur	**0.020**	0.000	0.000	-	0.061	0.047	0.009	0.092	0.041	0.061	0.043	**0.039**	0.064	0.195
Abbottabad	0.000	0.000	0.000	0.000	-	0.017	0.031	0.084	0.060	0.043	0.038	0.054	0.047	0.092
Mansehra	0.000	0.000	0.000	0.000	0.010	-	0.031	0.040	0.046	0.028	**0.004**	0.033	0.022	0.131
Tanawal	0.000	0.000	0.000	0.000	0.000	0.000	-	0.109	**0.026**	0.056	0.044	0.021	0.069	0.152
Siran	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-	0.055	0.072	**0.021**	0.079	**0.017**	0.224
Oghi	**0.070**	0.000	0.000	0.010	0.000	0.000	**0.060**	0.000	-	0.071	0.043	0.045	0.054	0.178
Battal	0.000	0.000	0.000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	-	0.018	0.042	0.051	0.193
Battagram	0.010	0.000	0.000	0.000	0.000	**0.220**	0.000	**0.050**	0.000	0.000	-	0.029	**0.009**	0.185
Attarshisha	0.000	0.000	0.000	**0.100**	0.000	0.000	0.010	0.000	0.000	0.000	0.000	-	0.050	0.210
Balakot	0.000	0.000	0.000	0.000	0.000	0.000	0.000	**0.020**	0.000	0.000	**0.120**	0.000	-	0.188
Skardu	0.000	**0.030**	0.000	0.000	0.090	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-

Non-significant F_{ST} values (> 0.01) are shown as bold.

F_{ST} values were significant for all comparisons of the over-summering population with the other populations, except for Peshawar in 2010 (F_{ST} = 0.028). The Lowest F_{ST} values were found for the non-*Berberis* location (Peshawar) in both years, as well as one non-*Berberis* location (Nowshera) in 2010 and two *Berberis* locations in 2011 (Table S3). However, none of the locations shared the multilocus genotypes resampled in this over-summering location, except for one single MLG that was resampled at most of the locations.

Virulences and pathotypic diversity at the Himalayan region of Pakistan

Virulences were detected to 20 out of the 26 different yellow rust resistance (*Yr*) genes, with regional differences in virulence diversity (Table 5). Thus virulence was absent only to six resistances genes i.e., *Yr3*, *Yr10*, *Yr17*, *Yr26*, *Yr32* and *YrSp*. Virulence to *Yr2* and *Yr9* were fixed in all populations, while *V6* and *V8* were fixed in four out of 11 tested populations. Virulences to *Yr5*, *Yr15*, *VrSd* and *Yr24* were rare and found only at some locations (Table 5). These virulences were present in 53 different virulence combinations (pathotypes) for the 127 tested isolates (Table 2), suggesting a high pathotypic diversity when compared to other worldwide clonal populations of PST (de Vallavieille-Pope *et al.*, 2012). None of the pathotypes was prevalent in the overall region, as no pathotype was found at all locations and no one was predominant at more than two regions. Similarly, at a given location no single pathotype had a frequency higher than 30% in the *Berberis* zone. The highest number of pathotypes was observed at Battal (18 pathotypes) while the lowest was detected at Peshawar (2 pathotypes).

Table 3. Differentiation between temporally spaced *Puccinia striiformis* f.sp. *tritici* populations (F_{ST}), temporal effective population size (Ne) and the number of resampled multilocus genotypes (MLGs) over 2010 and 2011 from 8 locations of Pakistan.

Locations	Sample size	F_{ST} between 2010 and 2011	Ne ± 95% CI	Number of resampled MLGs and their frequency over two years*
Peshawar	45	0.0719	32.6 [17.7-61.6]	1 (5:3)
Nowshera	35	0.0863	20.0 [10.0-38.0]	1 (13:1)
Abbottabad	52	**0.2375**	9.1 [5.5-14.0]	1 (1:11)
Mansehra	89	0.0522	32.6 [19.7-51.8]	2 (2:2) (1:6)
Battal	81	0.0740	31.1 [18.4-50.0]	0
Battagram	29	0.0661	35.2 [16.6-88.7]	0
Attarshisha	70	**0.1359**	23.1 [13.8-36.7]	0
Balakot	51	0.0609	33.7 [18.5-60.5]	1 (1:5)

***** The first digit shows the number of multilocus genotypes (MLGs) resampled, while the frequency in 2010 and 2011 are shown in brackets.

Table 4. Estimates of effective population size (± 95% CI) for *Puccinia striiformis* f.sp. *tritici* populations from different locations in the Himalayan region of Pakistan using two methods based on linkage disequilibrium.

Year	Location	Sample size	LDNe software [a]	Ne-estimator [b]
2010	**Khyber**	35	1.2 (1.1-1.4)	2.4 (2.2-2.6)
	Peshawar	12	0.7 (0.6-0.9)	- -
	Nowshera	19	0.3 (0.3-0.4)	0.7 (0.7-0.8)
	Haripur	42	1.5 (1.3-1.7)	2.9 (2.7-3.2)
	Abbottabad	34	7.2 (4.1-10.1)	5 (4.5-5.5)
	Mansehra	28	2.7 (2.3-3.1)	5.3 (4.8-5.9)
	Battal	46	1.6 (1.4-1.8)	2.9 (2.7-3.1)
	Battagram	9	8.2 2.6-44.0)	8.5 (6.2-12.5)
	Attarshisha	30	1.6 (1.3-1.9)	2.8 (2.5-3.1)
	Balakot	24	2 (1.6-2.4)	4 (3.6-4.4)
	Skardu	63	3.3 (2.9-3.6)	7.3 (6.8-7.9)
2011	**Peshawar**	33	1.2 (1.1-1.9)	- -
	Nowshera	16	0.3 (0.3-0.4)	- -
	Abbottabad	18	0.4 (0.3-0.5)	0.8 (0.7-0.9)
	Mansehra	61	1.5 (1.3-1.6)	3.1 (2.9-3.3)
	Tanawal	22	1.5 1.3-1.8)	2.7 (2.5-3.0)
	Siran	51	1.5 (1.3-1.7)	3.6 (3.3-3.9)
	Oghi	19	1.4 (1.0-1.8)	3 (2.6-3.5)
	Battal	35	1.8 (1.5-2.2)	3.6 (3.1-4.0)
	Battagram	20	1.5 (1.3-1.7)	2.6 (2.3-2.8)
	Attarshisha	40	4.4 (3.5-7.1)	4.7 (4.3-5.2)
	Balakot	27	1.5 (1.3-1.8)	3.5 (3.1-3.8)

[a] Following the algorithm of Waples and Do (2010); [b] Following the algorithm of Pudovkin et al. (1996)

DISCUSSION

We used multilocus microsatellite genotyping and pathogenicity tests to infer the population structure of PST in the Himalayan region of Pakistan, a recently described centre of diversity (Ali *et al.*, in preparation-a). Our analyses of natural PST populations sampled from wheat fields revealed the existence of a recombining population structure, high genotypic and high pathotypic diversity at each location. This level of diversity stands in stark contrasts with the very low genetic variability of most PST populations around the world (Ali *et al.*, in preparation-a). The high genotypic diversity of PST in Pakistan reported by Bahri et al. (2011) on the basis of a limited set of isolates, raised the question whether the observed high diversity in Pakistan could be the result of migration or local recombination. Here our results suggest the temporal stability of populations, despite the presence of a spatial pattern of population subdivision at regional level, thus supporting the importance of local survival through sexual cycle. The genotyping of *Berberis* leaves infected with rust revealed the presence of *Puccinia striiformis sensu lato* (Liu, Hambleton, 2010), supporting the role of sexual recombination. The contribution of over-summering population, which seems to be maintained locally, was very low to the populations of *Berberis* zone in comparison to non-*Berberis* zone.

High levels of genotypic diversity and recombining population structure

Analyses of microsatellite variation in Himalayan PST populations of Pakistan revealed the existence of high levels of genotypic diversity, no significant difference between observed and expected heterozygosity and low levels of linkage disequilibrium between markers consistent with the existence of recombination in the area. Pakistani populations were clearly differentiated from other world populations (Ali *et al.*, in preparation-a) despite the capacity of long distance migration in PST up to continental scales (Brown, Hovmøller, 2002; Hovmøller *et al.*, 2008), suggesting the local stability of Pakistani genetic groups. This stability is also supported by the fact that genotypes closely related to the recently spread aggressive strain PstS2 (Hovmøller *et al.*, 2010) were present in Pakistan but did not dominate the local Pakistani PST population in contrast to the populations in the Middle-East (R.

Table 5. Frequency (%) of different virulence factors in *Puccinia striiformis* populations from different locations of the Himalayan region of Pakistan sampled in 2010.

Virulence*	Khyber	Peshawar	Nowshera	Haripur	Abbottabad	Mansehra	Battal	Battagram	Attarshisha	Balakot	Skardu	Overall population
Vr1	72	0	0	30	7	60	8	33	0	27	25	24
Vr2	100	100	100	100	100	100	100	100	100	100	100	100
Vr3	0	0	0	0	0	0	0	0	0	0	0	0
Vr4	6	0	0	0	0	0	0	0	38	9	50	8
Vr5	0	0	0	0	0	0	8	0	0	0	0	2
Vr6	78	100	100	100	93	60	88	83	25	55	100	78
Vr7	89	100	100	100	100	100	100	100	100	100	100	98
Vr8	61	100	100	45	64	40	75	100	81	82	100	71
Vr9	100	100	100	100	100	100	100	100	100	100	100	100
Vr10	0	0	0	0	0	0	0	0	0	0	0	0
Vr15	0	0	20	0	0	0	0	0	0	0	0	1
Vr17	0	0	0	0	0	0	0	0	0	0	0	0
Vr24	0	0	0	0	0	0	4	0	0	0	0	1
Vr25	11	100	80	5	14	20	17	33	0	9	75	19
Vr26	0	0	0	0	0	0	0	0	0	0	0	0
Vr27	83	0	20	85	79	60	79	33	25	45	50	62
Vr32	0	0	0	0	0	0	0	0	0	0	0	0
VrA	22	100	80	25	36	40	21	17	6	18	100	29
VrSd	0	50	0	0	0	0	0	0	0	0	50	3
VrSu	94	0	20	95	86	80	92	100	100	91	25	85
VrSp	0	0	0	0	0	0	0	0	0	0	0	0
VrEp	11	100	80	5	14	20	38	17	0	9	100	23
VrVic	11	100	80	5	14	20	25	17	0	27	100	22
VrAv	50	100	80	50	79	60	46	50	56	55	100	58
Isolates tested	14	11	24	6	20	18	5	5	4	4	16	127

* Virulence factors for A, Sd, Su, Sp, Ep, Vic, Av refer to resistance genes of cvs Anza, Strubes Dickkopf , Suwon92xOmar, Spalding Prolific, Early Premium,Victo, Avocet, respectively.

Amil, unpublished data) where PstS2 dominated the vast majority of the area. The temporal stability, high genotypic diversity and recombining population structure of local populations is therefore most likely the result of sexual reproduction locally rather than recurrent immigration of genotypes from other regions. The occurrence of local sexual recombination is further supported by a high telial production ability of Pakistani populations (Ali *et al.*, 2010), and by a high abundance of *Berberis* spp. (the alternate host for the sexual cycle of PST: Jin *et al.*, 2010) in the Himalayan uplands (Perveen, Qaiser, 2010; Ray *et al.*, 2011). Thus we propose that local sexual recombination and not migration from other world populations causes the above described high genotypic diversity of PST in Pakistan. We identified plants of *Berberis* spp. severely infected with rust in all the prospected locations of *Berberis* zone (Fig. 1). The microsatellite genotyping of eight samples from these infected leaves confirmed that they likely belonged to *Puccinia striiformis sensu lato* (Liu, Hambleton, 2010), as our SSRs should be specific to yellow rust i.e., we verified that they do not amplify *P. triticina* (leaf rust) or *P. graminis* f.sp. *tritici* (stem rust). However, the amplification of closely related species and *formae specialis* needs to be confirmed through additional sequencing based analyses.

Pathotypic diversity in the Himalayan region of Pakistan

Analyses of virulence profiles of Pakistani PST populations revealed the existence of high diversity in the Himalayan region, confirming previous results (Bahri *et al.*, 2011; Rizwan *et al.*, 2010). The high frequency or fixation of different virulences (*Vr2*, *Vr9*, *Vr6* and *Vr8*) is in accordance with the cultivation of resistance varieties with the matching *Yr* genes (Bahri *et al.*, 2011; Rizwan *et al.*, 2010), except for *Vr8* that is not deployed in Pakistan. Similar cases of virulences not associated to a corresponding host induced selection were found (*Yr5*, *Yr24*, *Yr15*, *YrSd*), and are in accordance with the high diversity of the pathogen in this area. Additionally avirulence was observed for Victo, which has been reported generally susceptible to PST isolates in other parts of the world (de Vallavieille-Pope *et al.*, 2012). Our results showed no dominance of any pathotype in any of the locations. Additionally none of the pathotypes was specific to a given genetic group, except P-26 (with *Vr2, 6, 7,*

8, 9, 25, A, Ep, Vic, Av), with nine isolates belonging to the genetic group dominant in Peshawar valley.

Comparison of the pathotypes observed in the present study with a previous study conducted on PST populations from Pakistan revealed the existence of almost the same virulences but in more diverse combinations. In Peshawar and Nowshera, sampled in the two studies, pathotype data were available over three years. The dominant pathotype of 2006 (Bahri *et al.*, 2011) and 2008 (Shah, 2010) were not the same as in 2010, but they were still resampled in 2010. This could not arise from any host shift, as varieties based on the same *Yr* genes were cultivated during the period, but possibly a result of founder events, with recurrent re-establishment of the population from *Berberis* zone (likely from the over-summering area). Comparison with the worldwide distributed pathotypes suggested very high diversity in Pakistan, with pathotypes not related to other geographic area (Bahri *et al.*, 2009; de Vallavieille-Pope *et al.*, 2012; Hovmøller *et al.*, 2002; Mboup *et al.*, 2009).

Spatial patterns of population subdivision

Analyses of population subdivision revealed the existence of a complex, albeit weak, pattern population structure in PST populations from the Himalayan region of Pakistan. Both Bayesian and non-parametric analyses revealed the existence of at least four genetic groups (Fig. 4 and Fig. S2-S4). The most contrasted populations were found in Peshawar and Nowshera (non-*Berberis* zone), as they mostly belong to a unique genetic group. In contrast, the other locations are a genetic mosaic of genotypes from the different groups, together with "admixed" and unassigned individuals. The populations of Peshawar and Nowshera located in the non-*Berberis* zone clustered together with the diverse over-summering population in the *Berberis* zone (Skardu), suggesting its role of source population to establish PST in this non-*Berberis* zone.

Our analyses reveal the existence of population subdivision at the regional level for a pathogen characterized for long distance migration (Brown, Hovmøller, 2002; Hovmøller *et al.*, 2011) and lacking population subdivision elsewhere at continental and regional levels (Hovmøller *et al.*, 2002; Mboup *et al.*, 2009). Previously, a stable population subdivision over years for PST was only reported between the north and south of France (Enjalbert *et al.*, 2005), which was explained by the difference in host resistance gene deployment and differential adaptation temperature (Mboup *et al.*, 2012). However, in the present study there was no significant difference in resistance gene deployment across locations and at least the same two varieties (Inqilab-91 and PS-2004) were present and sampled at all locations, suggesting host pressure to be less probable at the origin of this spatial structure. Similarly, this pattern could not arise from differential adaptation to temperature, as differentiation existed between different valleys in the upland Himalayan region despite the prevalence of the same climatic conditions, while no differentiation existed between non-*Berberis* zone and the over-summering *Berberis* zone despite their contrasting climatic conditions (Table 2). Alternatively, the observed population structure, with local coexistence of different genetic groups, could be explained by a specific demography and mode of reproduction of PST population in Pakistan.

Population dynamics across years

When testing for a difference between years effect for populations sampled in two consecutive years, no significant effect was detected by AMOVA, demonstrating the absence of any global genetic shift in the surveyed area between years (Table S1). This temporal stability was further supported by the non-significant F_{ST} between the eight temporally spaced populations. However, we observed a limited re-sampling of the same multilocus genotypes over both years at these locations, which could result from sexual recombination and/or a high clonal diversity of the population. The limited clonal resampling would suggest both a high global effective size, and a limited survival through clonality, contrasting with the clonal populations of NW Europe and Mediterranean region, where the same multilocus genotypes are resampled over years (Bahri *et al.*, 2009; Enjalbert *et al.*, 2002; Hovmøller *et*

al., 2002). When analyzing more carefully the temporal patterns, the low values of the estimated effective size (ranging from 9 to 35) in the different areas are giving a slightly different picture: significant fluctuations in group frequencies/allele frequencies are occurring between years. The genetic composition of populations in the Himalayan area of Pakistan would be slightly fluctuating through the effect of drift, or migration. The signatures of recombination, high genotypic and pathotypic diversity are consistent with a significant role of sexual reproduction on *Berberis* in the local survival of the pathogen. However, in the case of a predominant local maintenance of inoculum through sexual cycle, all genetic groups would infect the same local *Berberis* plants, leading to the recombination of all groups. Hence, a single generation of survival over *Berberis* would generate numerous hybrids, erasing after few generations any pattern of population subdivision. Different hypotheses can be raised to explain the situation observed. First, a significant part of the population can survive through clonal survival on volunteers, which is less probable due to extreme summer temperatures (while severe cold temperature for over-summering zone). Then, specialization on the aecial host, i.e. *Berberis* species, could be at the origin of a lack of hybridization between genetic groups. Finally, some non-sampled areas might be mainly composed of a single genetic group, locally maintaining the integrity of the group and being a source of migrants for the other populations. PST populations in the *Berberis* zone of the Himalayan region would thus be maintained at least partly through sexual reproduction on the alternate host *Berberis spp.* abundantly present in the region. The over-summering *Berberis* zone population would have little impact on the local population in these *Berberis* zone locations, as revealed by strong differentiation of these populations from the over-summering population.

Significant differentiation of the over-summering population with these populations over both years did not support the idea of a “green-bridge”, except for non-*Berberis* zone. As the alternate host is absent in the non-*Berberis* zones like the Peshawar valley, PST populations have to be re-established each year either through local survival on volunteers or through re-establishment resulting from

migrations from other regions. The resampling of genotypes dominant in the non-*Berberis* zone in populations of the *Berberis* zone (including the over-summering zone) suggest a source and sink relationship between the non-*Berberis* zone and *Berberis* zone. Our analyses revealed that this population is closer to the over-summering *Berberis* zone population and highly diverged from the rest of the *Berberis* zone. This would thus suggest that this population could have a sink relationship with the over-summering *Berberis* zone, where the over-summering population would serve as the source population.

Conclusion

The present work confirmed the sexual recombination and to some extent over-summering survival at the origin of the high genotypic and pathotypic diversity of Pakistan, the western part of the recently shown centre of diversity of PST (Ali *et al.*, in preparation-a). The populations were shown to have a shallow spatial population structure with each population displaying a signature of recombination and high diversity. The *Berberis* zone populations are maintained, at least partly, through sexual recombination with a potential role of *Berberis* spp., while the non-*Berberis* zone population of Peshawar and Nowshera must be re-established by the migrants coming from the other populations, apparently from the over-summering *Berberis* zone population. The high diversity coupled with sexual recombination would grant high adaptation capacity and evolutionary potential to the pathogen to generate new pathotypes. Indeed virulences were present to genes not deployed in Pakistan or even worldwide (*Yr5*). The information generated in the study has several practical applications and would enable a better disease management at the local level in Pakistan as well as at a worldwide level in the context of invasions. The role of the alternate host in temporal maintenance of PST in the region needs to be fully assessed. Once the practical implication of the alternate host is confirmed, its eradication from the zones of diversity would enable a better disease management, as previously done for stem rust to bring the disease to the level with little economic impact. Finally, our methodology based on

the comparison of temporally and geographically spaced populations with genotyping based tracking could be used for other fungi to better comprehend fungal ecology.

ACKNOWLEDGEMENTS

The authors are grateful to all their colleagues in Pakistan who facilitated the sampling and provided valuable comments regarding the wheat production status in the region, mainly Muhammad Sartaj (NTRI, Mansehra, Pakistan), Muhammad Latif (ARS, Mansehra, Pakistan) and Abdul Hakim Khan (ARS, Mansehra, Pakistan). We pay a special thanks to Laurent Gérard for his remarkable assistance during spore multiplication in confined high security greenhouse. We also thank Nathalie Galet and Nicolas Lecutier for their technical assistance. This work received support from the European Integrated Project BIOEXPLOIT, FOOD-CT-2005-513959 and EMERFUNDIS, ANR 07-BDIV-003. Pierre Gladieux was supported by Marie Curie grant PIOF-GA-2010-273086 – GEDiNdis. Sajid Ali was financed by a grant from the Higher Education Commission, Pakistan and partially by the above mentioned projects.

AUTHOR CONTRIBUTIONS

SA carried out the molecular genotyping and population genetic analyses. AG supervised the molecular genotyping. JE and PG supervised population genetic analyses. SA, HR, MSS, HA and MF carried out the sampling. SA, ML and CP did the spore multiplication and pathotyping. SA, PG, CP and JE prepared the manuscript. MSH and AFJ supervised the *Berberis* part and revised the manuscript. SA, JE, HR, MSH and CP conceived and designed the study. All authors read and approved the manuscript.

REFERENCES

Agapow PM, Burt A (2001) Indices of multilocus linkage disequilibrium. *Molecular Ecology Notes* **1**, 101-102.

Ali S, Enjalbert J, Gladieux P, Leconte M, Gautier A, Justesen AF, Hovmoller MS, De Vallavieille-Pope C (in preparation-a) Origin, migration routes and genetic structure of worldwide populations of the wheat yellow rust pathogen.

Ali S, Gautier A, Leconte M, Enjalbert J, de Vallavieille-Pope C (2011) A rapid genotyping method for an obligate fungal pathogen, *Puccinia striiformis* f.sp. *tritici*, based on DNA extraction from infected leaf and Multiplex PCR genotyping. *BMC Research Notes* **4**, 240 doi:210.1186/1756-0500-1184-1240.

Ali S, Leconte M, Walker A-S, Enjalbert J, de Vallavieille-Pope C (2010) Reduction in the sex ability of worldwide clonal populations of *Puccinia striiformis* f.sp. *tritici*. *Fungal Genetics and Biology* **47**, 828-838.

Ali S, Rahman H, Shah SS, Khalil IA, Khalil IH (2007) Evaluation of outcrossed maize progeny for grain yield traits. *Sarhad Journal of Agriculture* **23**, 629-632.

Ali S, Shah SJA, Khalil IH, Rahman H, Maqbool K, Ullah W (2009a) Partial resistance to yellow rust in introduced winter wheat germplasm at the north of Pakistan. *Australian Journal of Crop Science* **3**, 37-43.

Ali S, Shah SJA, Rahman H, Saqib MS, Ibrahim M, Sajjad M (2009b) Variability in wheat yield under yellow rust pressure in Pakistan. *Turkish Journal of Agriculture and Forestry* **33**, 537-546.

Ali S, Soubeyrand S, Gladieux P, Leconte M, Gautier A, Mboup M, de Vallavieille-Pope C, Enjalbert J (In preparation-b) Recapturing clones to estimate sexuality and population size of pathogens. **In preparation**.

Arnaud-Haond S, Belkhir K (2007) GenClone 1.0: A new program to analyse genetics data on clonal organisms. *Molecular Ecology Notes* **7**, 15-17.

Bahri B, Leconte M, Ouffroukh A, de Vallavieille-Pope C, Enjalbert J (2009) Geographic limits of a clonal population of wheat yellow rust in the Mediterranean region. *Molecular Ecology* **18**, 4165-4179.

Bahri B, Shah SJA, Hussain S, Leconte M, Enjalbert J, de Vallavieille-Pope C (2011) Genetic diversity of wheat yellow rust population in Pakistan and its relationship with host resistance. *Plant Pathology* **60**, 649-660.

Belkhir KBP, Chikhi L, Raufaste N, Bonhomme F (2004) GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5000: Université de Montpellier II, Montpellier (France) http://www.univ-montp2.fr/~genetix/genetix/genetix.htm.

Brown JKM, Hovmøller MS (2002) Aerial dispersal of fungi on the global and continental scales and its consequences for plant disease. *Science* **297**, 537-541.

Burt A, Carter DA, Koenig GL, White TJ, Taylor JW (1996) Molecular markers reveal cryptic sex in the human pathogen *Coccidioides immitis*. *Proceedings of the National Academy of Science of the United States of America* **93**, 770-773.

Chaube HS, Pundhir VS (2005) *Crop diseases and their management* Prentice-Hall, New Delhi, India.

Chen X, Line RF, Leung H (1993) Relationship between virulence variation and DNA polymorphism in *Puccinia striiformis*. *Phytopathology* **83**, 1489-1497.

de Vallavieille-Pope C, Ali S, Leconte M, Enjalbert J, Delos M, Rouzet J (2012) Virulence dynamics and regional structuring of *Puccinia striiformis* f. sp. *tritici* in France between 1984 and 2009. *Plant Disease* **96**, 131-140.

de Vallavieille-Pope C, Huber L, Leconte M, Bethenod O (2002) Preinoculation effects of light quantity on infection efficiency of *Puccinia striiformis* and *P. triticina* on wheat seedlings. *Phytopathology* **92**, 1308-1314.

de Vallavieille-Pope C, Picard-Formery H, Radulovic S, Johnson R (1990) Specific resistance factors to yellow rust in seedlings of some French wheat varieties and races of *Puccinia striiformis* Westend. in France. *Agronomie* **10**, 103-113.

Dean R, Van Kan J, Pretorius ZA, Hammond-Kosack KE, Pietro A, Spanu PD, Rudd JJ, Dickman M, Kahmann R, Ellis J, Foster GD (2012) The top 10 fungal pathogens in molecular plant pathology. *Molecular Plant Pathology* **13**, 414-430.

Duan X, Tellier A, Wan A, Leconte M, de Vallavieille-Pope C, Enjalbert J (2010) *Puccinia striiformis* f.sp. *tritici* presents high diversity and recombination in the over-summering zone of Gansu-China. *Mycologia* **102**, 44-53.

Duveiller E, Singh RP, Nicol JM (2007) The challenges of maintaining wheat productivity: pests, diseases, and potential epidemics. *Euphytica* **157**, 417-430.

Enjalbert J, Duan X, Giraud T, Vautrin C, de Vallavieille-Pope C, Solignac M (2002) Isolation of twelve microsatellite loci, using an enrichment protocol, in the phytopathogenic fungus *Puccinia striiformis* f. sp. *tritici*. *Molecular Ecology Notes* **2**, 563-565.

Enjalbert J, Duan X, Leconte M, Hovmøller MS, de Vallavieille-Pope C (2005) Genetic evidence of local adaptation of wheat yellow rust (*Puccinia striiformis* f. sp. *tritici*) within France. *Molecular Ecology* **14**, 2065-2073.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE : a simulation study. *Molecular Ecology* **14**, 2611-2620.

Excoffier L, Laval G, Schneider S (2005) Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* **1**, 47-50.

Fisher M, Henk D, Briggs C, Brownstein JS, Madoff LC, McCraw SL, Gurr SJ (2012) Emerging fungal threats to animal, plant and ecosystem health. *Nature* **484**, 186-194.

Gao H, Williamson S, Bustamante CD (2007) A markov chain monte carlo approach for joint inference of population structure and inbreeding rates from multilocus genotype data. *Genetics* **176**, 1635-1651.

Giraud T, Gladieux P, Gavrilets S (2010) Linking the emergence of fungal plant diseases with ecological speciation. *Trends in Ecology & Evolution* **25**, 387-395.

Giraud T, Villareal L, Austerlitz F, Le Cam B, Lavigne C (2006) Importance of the life cycle in host race formation and sympatric speciation in parasites. *Phytopathology* **96**, 280-287.

Gladieux P, Byrnese E, Fisher MC, Aguleta G, Heitman J, T. G (2011) Epidemiology and evolution of fungal pathogens, in plants and animals. In: *Genetics and Evolution of Infectious Diseases* (ed. M T), p. 772. Elsevier.

Goudet J (2001) FSTAT, a program to estimate and test gene diversities and fixation indices. Available from http://www2.unil.ch/popgen/ softwares/fstat.htm.

Hassan SF (1968) Cereal rusts situation in Pakistan. In: *Proceedings of the European and Mediterranean Cereal Rusts Conference*, pp. 124-125, Oeiras, Portugal.

Hovmøller MS, Justesen AF, Brown JKM (2002) Clonality and long-distance migration of *Puccinia striiformis* f.sp. *tritici* in north-west Europe. *Plant Pathology* **51**, 24-32.

Hovmøller MS, Sørensen CK, Walter S, Justesen AF (2011) Diversity of *Puccinia striiformis* on cereals and grasses. *Annual Review of Phytopathology* **49**, 197-217.

Hovmøller MS, Walter S, Justesen AF (2010) Escalating threat of wheat rusts. *Science* **329**, 369.

Hovmøller MS, Yahyaoui AH, Milus EA, Justesen AF (2008) Rapid global spread of two aggressive strains of a wheat rust fungus. *Molecular Ecology* **17**, 3818-3826.

Huerta-Espino J, Singh RP (1996) Misconceptions on the durability of some adult plant resistance genes in wheat. In: *Proceedings of the 9th European and Mediterranean Cereal Rusts and Powdery Mildew Conference*, pp. 109-111, Lunteren The Netherlands.

Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* **23**, 1801-1806.

Jin Y, Szabo LJ, Carson M (2010) Century-old mystery of Puccinia striiformis life history solved with the identification of Berberis as an Alternate Host. *Phytopathology* **100**, 432-435.

Johnson R, Stubbs RW, Fuchs E, Chamberlain NH (1972) Nomenclature for physiologic races of *Puccinia striiformis* infecting wheat. *Transactions of the British Mycological Society* **58**, 475-480.

Jombart T, Devillard S, Balloux F (2010) Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics* **11**, 95 http://www.biomedcentral.com/1471-2156/1411/1494.

Langella O (2008) POPULATIONS 1.2.3, Logiciel de génétique des populations (ed. Langella O). Laboratoire Populations, Génétique et évolution, CNRS, Gif-sur-Yvette, France.

Liu M, Hambleton S (2010) Taxonomic study of stripe rust, *Puccinia striiformis* sensu lato, based on molecular and morphological evidence. *Fungal Biology* **114**, 881–899.

Maynard-Smith J, Smith NH, O'rourke M, Spratt BG (1993) How clonal are bacteria? *Proceedings of National Academy of Science* **90**, 4384-4388.

Mboup M, Bahri B, Leconte M, de Vallavieille-Pope C, Kaltz O, Enjalbert J (2012) Genetic structure and local adaptation of European wheat yellow rust populations: the role of temperature-specific adaptation. *Evolutionary Applications* **DOI: 10.1111/j.1752-4571.2011.00228.x**.

Mboup M, Leconte M, Gautier A, Wan AM, Chen W, de Vallavieille-Pope C, Enjalbert J (2009) Evidence of genetic recombination in wheat yellow rust populations of a Chinese oversummering area. *Fungal Genetics and Biology* **46**, 299-307.

McDonald BA, Linde C (2002) Pathogen population genetics, evolutionary potential, and durable resistance. *Annual Review of Phytopathology* **40**, 349-379.

Milgroom MG, Peever TL (2003) Population biology of plant pathogens: the synthesis of plant disease epidemiology and population genetics. *Plant Disease* **87**, 608-617.

Park RF, McIntosh RA (1994) Adult plant resistances to Puccinia recondita f. sp. tritici in wheat. *New Zealand Journal of Crop and Horticultural Science* **22**, 151-158.

Perveen A, Qaiser M (2010) Pollen flora of Pakistan–LXV. berberidaceae. *Pakistan Journal of Botany* **42**, 1-6.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* **155**, 945-959.

Pudovkin AI, Zaykin DV, Hedgecock D (1996) On the potential for estimating the effective number of breeders from heterozygote-excess in progeny. *Genetics* **144**, 383-387.

Ray R, Gururaja KV, Ramchandra TV (2011) Predictive distribution modeling for rare Himalayan medicinal plant *Berberis aristata* DC. *Journal of Environmental Biology* **32**, 725-730.

Raymond M, Rousset F (1995) GENEPOP Version 1.2: population genetics software for exact tests and ecumenicism. *Journal of Heredity* **86**, 248-249.

Rizwan S, Iftikhar A, Kazi AM, Sahi GM, Mirza JI, Rattu AR, Ashraf M (2010) Virulence variation of *Puccinia striiformis* Westend. f. sp. *tritici* in Pakistan. *Archives of Phytopathology and Plant Protection* **43**, 875-882.

Shah SJA (2010) Aerobiology, slow rusting resistance, gene postulation and virulence diversity of *Puccinia striformis* Westend. f. sp.*tritici* in NWFP, Pakistan. *Department of Plant Pathology, Khyber pakhtunkhwa Agricultural University Peshawar, Pakistan.* **PhD thesis**.

Shah SJA, Imtiaz M, Hussain S (2009) Phenotypic and molecular characterization of wheat for slow rusting resistance against *Puccinia striiformis* Westend. f.sp. *tritici*. *Journal of Phytopathology* **doi: 10.1111/j.1439-0434.2009.01631.x**.

Simwami SP, Khayhan K, Henk DA, Aanensen DM, Boekhout T, Hagen F, Brouwer AE, Harrison TS, Donnelly CA, Fisher MC (2011) Low diversity Cryptococcus neoformans variety grubii multilocus sequence types from Thailand are consistent with an ancestral African origin. *Plos Pathogen* **7**, e1001343.

Singh RP, William HM, Huerta-Espino J, Rosewarne G (2004) Wheat rust in Asia: meeting the challenges with old and new technologies.

Sobsey MD, Pillai SD (2009) Where future emerging pathogens will come from and what approaches can be used to find them, besides VFARs. *Journal of Water and Health* **7**, S75–S93.

Stubbs RW (1985) Stripe rust. In: *The Cereal Rusts, Diseases, Distribution, Epidemiology and Control* (eds. Roelfs AP, Bushnell WR), pp. 61–101. Academic Press, London, UK.

Stukenbrock EH, McDonald BA (2008) The origins of plant pathogens in agro-ecosystems. *Annual Review of Phytopathology* **46**, 75-100.

Taylor JW, Fisher MC (2003) Fungal multilocus sequence typing - it's not just for bacteria. *Current Opinion in Microbiology* **6**, 351-356.

Wan A, Zhao Z, Chen X, He Z, Jin S, Jia Q, Yao G, Yang J, Wang B, Li G, BI Y, Yuan Z (2004) Wheat stripe rust epidemic and virulence of *Puccinia striiformis* f.sp. *tritici* in China in 2002. *Plant Disease* **88**, 896-904.

Waples RS (1989) A generalized approach for estimating effective population size from temporal changes in allele frequency. *Genetics* **121**, 379-391.

Waples RS, Do C (2010) Linkage disequilibrium estimates of contemporary Ne using highly variable genetic markers: a largely untapped resource for applied conservation and evolution. *Evolutionary Applications* **3**, 244–262.

Wellings CR (2007) *Puccinia striiformis* in Australia: A review of the incursion, evolution and adaptation of stripe rust in the period 1979-2006. *Australian Journal of Agricultural Research* **58**, 567-575.

SUPPORTING INFORMATION

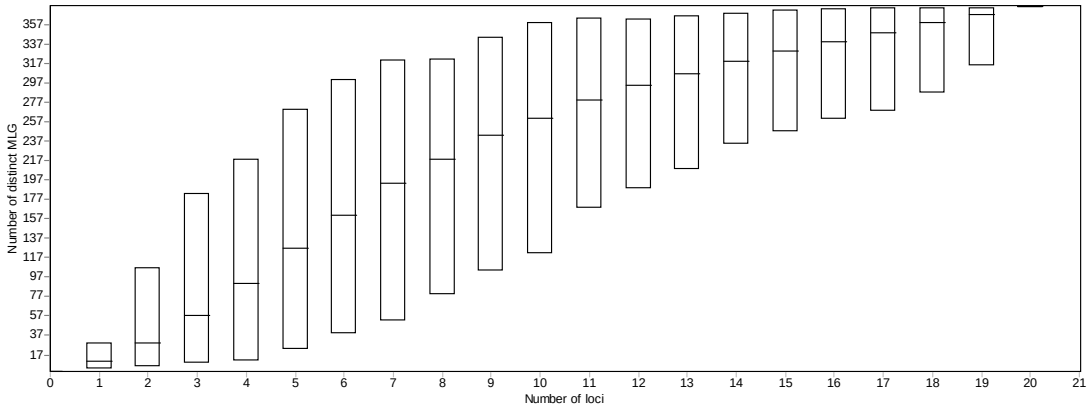


Fig. S1. Number of *Puccinia striiformis* f.sp. *tritici* multilocus genotypes (MLG) detected as a function of the number of loci re-sampled within the 20 microsatellite markers using GENECLONE software. The box represents the average, minimum and maximum numbers of MLGs detected when re-sampling on loci.

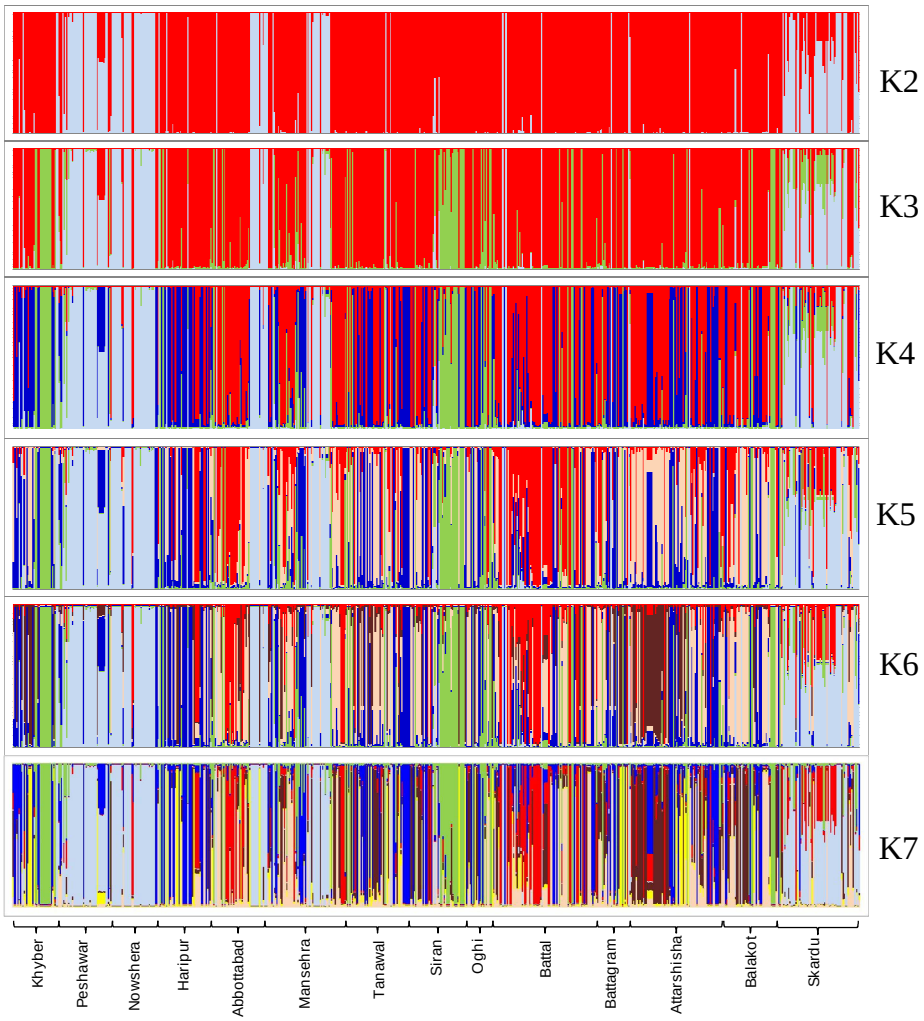


Fig. S2. Assignment of *Puccinia striiformis* isolates from geographically spaced populations from the Himalayan region of Pakistan, using the STRUCTURE software for different K-values (genetic groups).

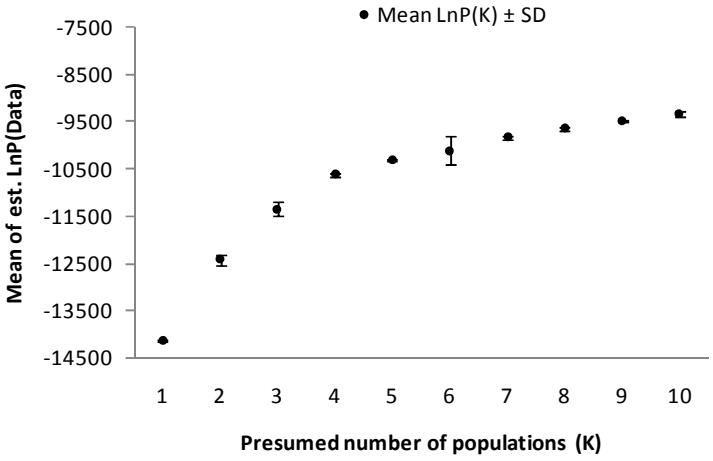


Fig. S3. Distribution of mean Ln P(D) values estimated with STRUCTURE software for *Puccinia striiformis* f.sp. *tritici* populations from the Himalayan region of Pakistan, suggesting K =4 as the optimal number of populations (K) following Evano et al., 2005).

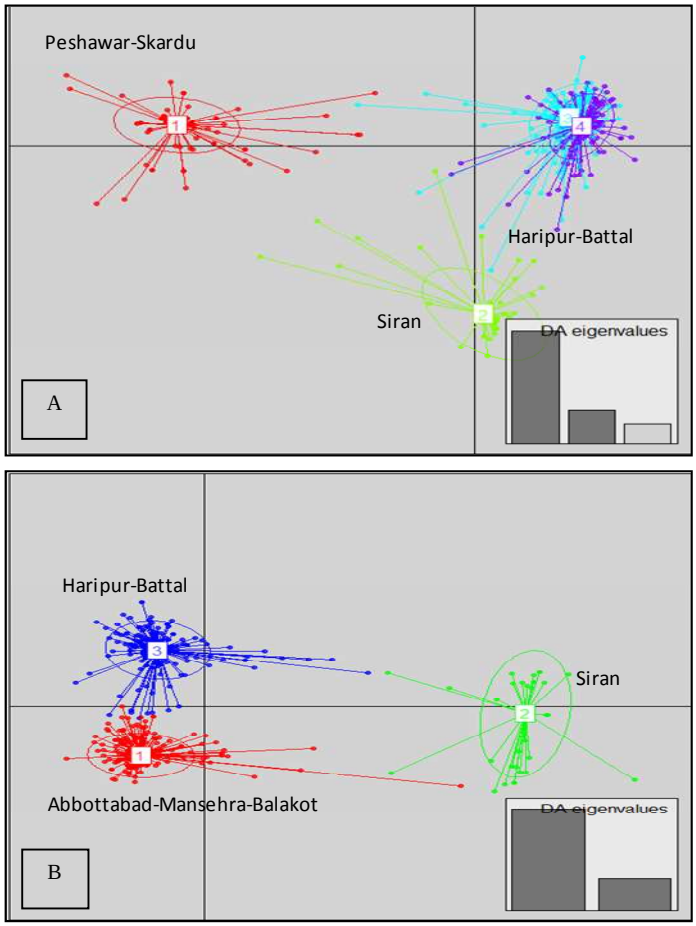


Fig. S4. Discriminate analyses of principal components (DAPC) for *Puccinia striiformis* isolates form the Himalayan region of Pakistan, suggesting at least four sub-populations. A = DAPC analyses was made for all isolates, including the Peshawar and Skardu populations. B = A sub analyses was made for all isolates except G1 of STRUCTURE analyses i.e., Peshawar-Skardu populations to explore the divergence of the other three populations.

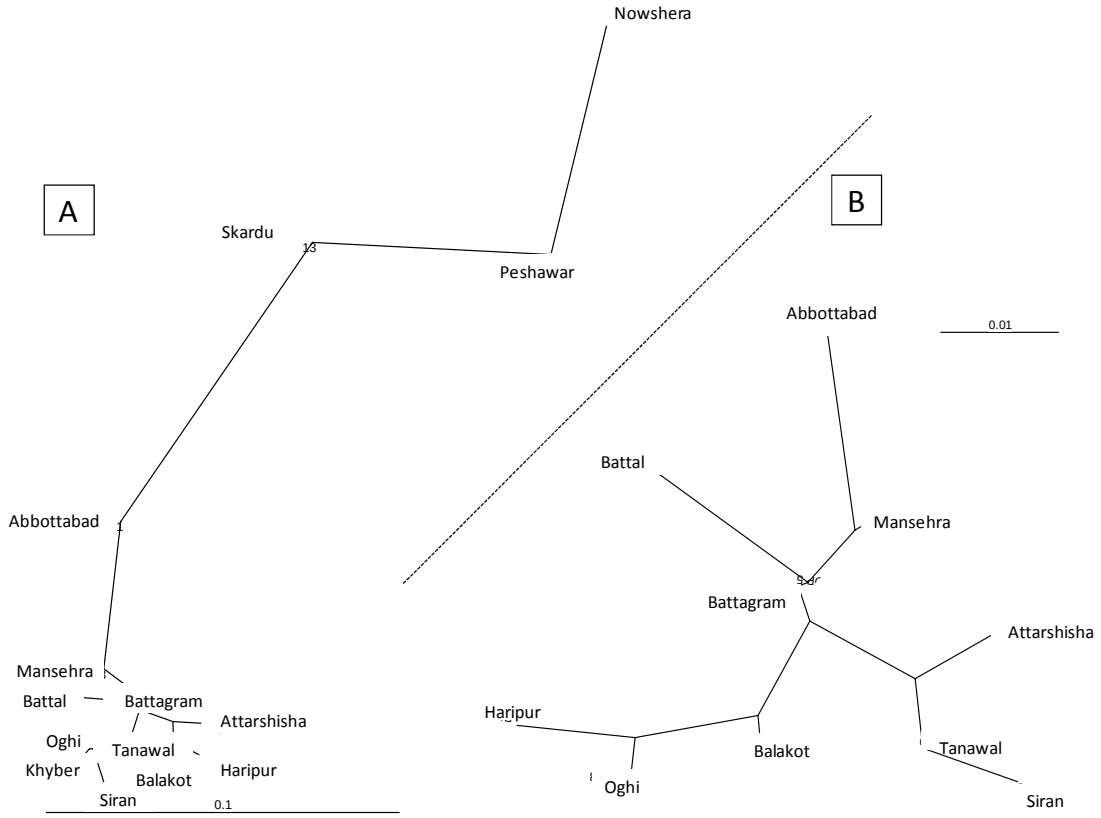


Fig. S5. Neighbour-joining tree based on microsatellite distance for Puccinia striiformis population in the Himalayan region of Pakistan. A. All 14 populations were used to construct the tree; B. 10 populations were used in analyses, excluding Peshawar valley, Khyber and Skardu populations.

Table S1 : Analyses of molecular variance for *Puccinia striiformis* f.sp. *tritici* isolates from 14 different locations sampled over 2010 and 2011 in the Himalayan region of Pakistan.

Source of Variation	d.f.	Sum of squares	Variance components	Percentage of variation
Among populations	1	25.029	-0.03302 Va	-0.81
Among populations within groups	20	779.828	0.58260 Vb	14.22
Within populations	1346	4775.041	3.54758 Vc	86.59
Total	1367	5579.897	4.09717	

Table S2. Distribution of 53 pathotypes (P-1-P-53) of *Puccinia striiformis* across different locations of the Himalayan region of Pakistan sampled in 2010, using differential lines able to identify Vr1, Vr2, Vr3, Vr4, Vr5, Vr6, Vr7, Vr8, Vr9, Vr10, Vr15, Vr17, Vr24, Vr25, Vr, Vr26, Vr27, Vr32, VrA, VrSd, VrSu, VrSp, VrEp, VrVic, VrAv.

Pathotypes	Virulence profile	Khyber	Peshawar	Nowshera	Haripur	Abbottabad	Mansehra	Battal	Battagram	Attarshisha	Balakot	Skardu	Overall Population
P-1	,2,,,,,,9,,,,,27,,,,Su,,,,	6	-	-	-	-	-	-	-	-	-	-	1
P-2	,2,,,,,7,,9,,,,,,,,,,Su,,,,	-	-	-	-	-	-	4	-	13	-	-	2
P-3	,2,,,,,7,,9,,,,,,27,,,,Su,,,,	-	-	-	-	-	-	4	-	-	-	-	1
P-4	,2,,,,,7,8,9,,,,,,,,,,Su,,,,	-	-	-	-	-	20	-	-	-	-	-	1
P-5	,2,,,,,7,8,9,,,,,,,,,,Su,,,,Av	-	-	-	-	-	-	-	-	19	9	-	3
P-6	,2,,,,,7,8,9,,,,,,27,,,,,Su,,,,	-	-	-	-	7	-	4	-	6	-	-	2
P-7	,2,,,,6,7,,9,,,,,,,,,,,,,Av	-	-	-	-	7	-	-	-	-	-	-	1
P-8	,2,,,,6,7,,9,,,,,,,,,,Su,,,,	-	-	-	-	-	-	-	-	-	9	-	1
P-9	,2,,,,6,7,,9,,,,,,,,,,Su,,Vic,,	-	-	-	-	-	-	4	-	-	-	-	1
P-10	,2,,,,6,7,,9,,,,,,27,,,,,Su,,,,	-	-	-	15	-	-	-	-	-	9	-	3
P-11	,2,,,,6,7,,9,,,,,,27,,,,,Su,,,,Av	-	-	-	10	14	-	-	-	-	-	-	3
P-12	,2,,,,6,7,,9,,,,,,27,,A,,,Su,,,,Av	-	-	-	5	7	-	4	-	-	-	-	2
P-13	,2,,,,6,7,,9,,,,,25,,,A,,,,EP,Vic,,Av	-	-	-	5	-	-	4	-	-	-	-	2
P-14	,2,,,,6,7,,9,,,,,25,27,,,,,Su,,,,	-	-	-	-	7	-	-	-	-	-	-	1
P-15	,2,,,,6,7,8,9,,,,,,,,,,Su,,,,	6	-	20	5	-	-	4	17	6	-	-	5
P-16	,2,,,,6,7,8,9,,,,,,,,,,Su,,,,Av	-	-	-	5	7	-	-	17	6	-	-	3
P-17	,2,,,,6,7,8,9,,,,,,27,,,,Su,,,,	-	-	-	10	-	-	13	-	6	-	-	5
P-18	,2,,,,6,7,8,9,,,,,27,,,,Su,,,,Av	6	-	-	5	14	-	8	-	-	9	-	6
P-19	,2,,,,6,7,8,9,,,,,,27,,,,Su,,Vic,,	-	-	-	-	-	-	-	-	-	9	-	1
P-20	,2,,,,6,7,8,9,,,,,,27,,,,Su,,Vic,,Av	-	-	-	-	-	-	-	-	-	9	-	1
P-21	,2,,,,6,7,8,9,,,,,,27,,,,Su,Ep,,,	-	-	-	-	-	-	8	-	-	-	-	2
P-22	,2,,,,6,7,8,9,,,,,,27,,,,Su,Ep,,,Av	-	-	-	-	-	-	13	-	-	-	-	2
P-23	,2,,,,6,7,8,9,,,,,,27,,,,Su,Ep,Vic,,	-	-	-	-	-	-	4	-	-	-	-	1
P-24	,2,,,,6,7,8,9,,,,,,27,,A,,,Su,,,,Av	6	-	-	10	21	-	-	-	6	-	-	6
P-25	,2,,,,6,7,8,9,,,,,25,,,,,Su,,,,	-	-	-	-	-	-	-	17	-	-	-	1
P-26	,2,,,,6,7,8,9,,,,,25,,,A,,,,Ep,Vic,,Av	6	50	60	-	7	20	-	-	-	9	-	7
P-27	,2,,,,6,7,8,9,,,,,25,,,A,SD,,,Ep,Vic,,Av	-	50	-	-	-	-	-	-	-	-	-	2
P-28	,2,,,,6,7,8,9,,,,,25,27,,A,,,,Ep,Vic,,Av	-	-	-	-	-	-	4	-	-	-	25	2
P-29	,2,,,,6,7,8,9,,,,,25,27,,A,,,Su,,,,Av	-	-	-	-	-	-	4	-	-	-	-	1
P-30	,2,,,,6,7,8,9,,,,,25,27,,A,,,Su,Ep,Vic,,Av	-	-	-	-	-	-	-	17	-	-	-	1
P-31	,2,,,,6,7,8,9,,15,,,25,27,,A,,,,Ep,Vic,,Av	-	-	20	-	-	-	-	-	-	-	-	1
P-32	,2,,,5,6,7,8,9,,,,,25,27,,A,,,Su,,Vic,,Av	-	-	-	-	-	-	4	-	-	-	-	1
P-33	,2,,,5,6,7,8,9,,,,24,,,,,,Su,Ep,Vic,,	-	-	-	-	-	-	4	-	-	-	-	1
P-34	,2,,4,,,7,,9,,,,,,,,,,Su,,,,Av	-	-	-	-	-	-	-	-	6	-	-	1
P-35	,2,,4,,,7,8,9,,,,,,,,,,Su,,,,	-	-	-	-	-	-	-	-	13	-	-	2
P-36	,2,,4,,,7,8,9,,,,,,,,,,Su,,,,Av	-	-	-	-	-	-	-	-	13	9	-	2
P-37	,2,,4,,,7,8,9,,,,,27,,,,Su,,,,Av	-	-	-	-	-	-	-	-	6	-	-	1
P-38	,2,,4,,6,7,8,9,,,,,25,,,A,SD,,,Ep,Vic,,Av	-	-	-	-	-	-	-	-	-	-	25	1
P-39	,2,,4,,6,7,8,9,,,,,25,,,A,SD,,Su,Ep,Vic,Ju,Av	-	-	-	-	-	-	-	-	-	-	25	1
P-40	1,2,,,,,,8,9,,,,,,27,,,,,Su,,,,	6	-	-	-	-	-	-	-	-	-	-	1
P-41	1,2,,,,,,7,,9,,,,,,27,,,,,Su,,,,Av	-	-	-	-	-	20	-	-	-	-	-	1
P-42	1,2,,,,,,7,8,9,,,,,,,,,,Su,,,,	-	-	-	-	-	-	-	17	-	9	-	2
P-43	1,2,,,,,,7,8,9,,,,,,,,,,Su,,,,Av	6	-	-	-	-	-	-	-	-	9	-	2
P-44	1,2,,,,,,7,8,9,,,,,,27,,,,,Su,,,,Av	6	-	-	-	-	-	-	-	-	-	-	1
P-45	1,2,,,,,,7,8,9,,,,,,27,,A,,Su,,,,	-	-	-	-	-	-	-	-	-	9	-	1
P-46	1,2,,,,6,7,,9,,,,,,27,,,,,Su,,,,	28	-	-	15	-	20	4	-	-	-	-	8
P-47	1,2,,,,6,7,,9,,,,,,27,,A,,,Su,,,,Av	6	-	-	5	-	20	-	-	-	-	-	2
P-48	1,2,,,,6,7,8,9,,,,,,27,,,,,Su,,,,	6	-	-	5	-	-	-	-	-	-	-	2
P-49	1,2,,,,6,7,8,9,,,,,,27,,,,,Su,,,,Av	6	-	-	5	-	-	4	17	-	-	-	3
P-50	1,2,,,,6,7,8,9,,,,,,27,,,,,Su,Ep,Vic,,	-	-	-	-	7	-	-	-	-	-	-	1
P-51	1,2,,,,6,7,8,9,,,,,,27,,A,,,,Ep,Vic,,Av	-	-	-	-	-	-	-	-	-	-	25	1
P-52	1,2,,,,6,7,8,9,,,,,25,27,,A,,,Su,Ep,Vic,,Av	6	-	-	-	-	-	-	-	-	-	-	1
P-53	1,2,,4,,6,7,8,9,,,,,,27,,,,,Su,,,,Av	6	-	-	-	-	-	-	-	-	-	-	1
Number of isolates tested		18	4	5	20	14	5	24	6	16	11	4	127

Table S3. F_{ST} Values estimated for over-summering populations (2010) with all other 13 populations of PST resampled in the Himalayan region of Pakistan.

Between offseason population and :	year	F_{ST}	Significance
Khyber	2010	0.168	0.000
Peshawar	2010	0.028	0.027
	2011	0.076	0.000
Nowshera	2010	0.085	0.000
	2011	0.135	0.000
Haripur	2010	0.164	0.000
Abbottabad	2010	0.143	0.000
	2010	0.063	0.000
Mansehra	2010	0.170	0.000
	2011	0.083	0.000
Siran	2011	0.176	0.000
Tanawal	2011	0.124	0.000
Oagi	2011	0.153	0.000
Attarshisha	2010	0.234	0.000
	2011	0.147	0.000
Balakot	2010	0.138	0.000
	2011	0.166	0.000
Battal	2010	0.146	0.000
	2011	0.192	0.000
Battagram	2010	0.142	0.000
	2011	0.119	0.000

CHAPTER-IV

Recapturing clones to estimate the sexuality and population size of pathogens

Sajid Ali, Samuel Soubeyrand *, Pierre Gladieux, Marc Leconte, Angélique Gautier, Mamadou Mboup, Wanquan Chen and Claude de Vallavieille-Pope, Jérôme Enjalbert

* The mathematical work, the development of R-Package and simulations were carried out by Samuel Soubeyrand.

Recapturing clones to estimate the sexuality and population size of pathogens

The article will be submitted to an international journal

Recapturing clones to estimate the sexuality and population size of pathogens

Sajid Ali[1], Samuel Soubeyrand[2*], Pierre Gladieux[3,4], Marc Leconte[1], Angélique Gautier[1], Mamadou Mboup[5], Wanquan Chen[6] and Claude de Vallavieille-Pope[1], Jérôme Enjalbert[7]

[1] UR1290, BIOGER-CPP, INRA-AgroParisTech, BP01, 78850 Thiverval-Grignon, France

[2] UR0546, Unité Biostatistique et Processus Spatiaux (BioSP), INRA, Avignon, France

[3] Laboratoire Ecologie, Systématique et Evolution, Univ. Paris Sud, UMR8079, Orsay, France; CNRS, Orsay, France; AgroParisTech, Orsay, France

[4] Department of Plant and Microbial Biology, University of California, Berkeley, CA 94720-3102

[5] LMU Biocenter - Evolutionary Biology, Munich, Germany

[6] Laboratory of Biology of Plant Diseases and Insect Pests, Institute of Plant Protection, Chinese Academy of Agricultural Sciences, Beijing 100094, China

[7] INRA UMR320/ UMR 8120 Génétique Végétale, F-91190 Gif sur Yvette, France

sajid.ali@grignon.inra.fr; enjalbert@moulon.inra.fr; Samuel.Soubeyrand@avignon.inra.fr

* The mathematical work, the development of R-Package and simulations was carried out by Samuel Soubeyrand.

ABSTRACT

Quantifying reproductive and demographic parameters is crucial for understanding the evolutionary potential of pathogens. Molecular typing techniques can provide useful data for understanding the factors that affect the spatial and temporal dynamics of diseases, but there are multiple methodological challenges that are involved in inferring the key population genetic parameters. Here, we describe a new, fast and accurate method for jointly quantifying the relative contribution of sexual

and asexual reproduction and the effective population size based on clonal identity within and across generations. The performance of the method was assessed using simulations, which demonstrated that reliable estimates could be obtained for a wide range of parameters. The method was then applied to the fungal pathogen *Puccinia striiformis* f.sp. *tritici* (PST), which causes stripe/yellow rust, a significant wheat disease. We estimated a sexual reproduction rate of 74% (95% confidence interval [CI]: 38-95%) using two temporally spaced samples of a Chinese PST population, which suggests a high contribution from sexual reproduction in this region where recombination was previously described. We estimated an effective population size of 1735 (95% CI: 675-2800), which fell within the estimates obtained using other methods for estimating the instantaneous population size based on the variance in allele frequencies (779 to 4748, depending upon the method used). The new method presented here can be used to quantify the reproductive and demographic parameters of populations of any organism given clonemate sampling at two different times and will contribute to better understanding of the population biology of pests and microbes.

Keywords: Sex rate, effective population size, microorganisms, clones

INTRODUCTION

Estimating the reproductive mode and effective population size of pathogens is a major concern for scientists studying microbial population biology (Awadalla, 2003). These basic population parameters play a key role in the emergence of new virulent or drug-resistant strains and knowledge of them can assist in attempts to control disease-causing agents (Williams, 2010). Asexual reproduction, which is widespread in microorganisms, can boost the adaptive response through multiplication of the fittest individuals (de Meeûs *et al.*, 2007), while recombination facilitates the fixation of beneficial mutations (Goddard *et al.*, 2005) and the maintenance of favorable allelic combinations within individual genomes. The effective population size drives the capture of novel adaptive mutations or the purge of deleterious ones (Glémin, 2003; Paland, Schmid, 2003).

Because microbial pathogens are by nature inconspicuous, the use of molecular markers has become a favored method for investigating their population biology (Maiden *et al.*, 1998; Taylor, Fisher, 2003). The use of these markers coupled with population genetics tools have revolutionized our understanding of the biology of microbes, proving the usefulness of the markers in determining if species are purely clonal or if they also undergo recombination (Burt *et al.*, 1996; Taylor *et al.*, 1999). However, recombination rates can only be quantified by using direct experimental estimates of the mutation and recombination rates. Measures of genetic diversity from population genomics data can then be used to estimate the frequency of different types of reproduction in the pathogen's life cycle or to infer the effective population size (Tsai *et al.*, 2008). Because of these limitations, the sex rate and the effective population size (*Ne*) remain unknown for the vast majority of microorganisms, including several model species.

Although microbial pathogens have a large census population size, many pathogens experience recurrent bottlenecks due to the low probability of transmission to a new host, periodic fluctuations in host availability, host resistance and extreme environmental conditions. In such cases, although pathogen populations can reach millions or billions of individuals within a single host, the effective population size is often many orders of magnitude lower. Very few studies have attempted to estimate *Ne* in microbial populations (Fisher, 2007). The classical approach for estimating the *Ne* uses measures of genetic diversity, mutation rate and assumes mutation-drift equilibrium (Fisher, 2007; Zhan *et al.*, 2001). This approach provides an estimate of the effective population size over a long period of time (Charlesworth, 2009), although it tells little or nothing about the effective population size in the short term, though this is more relevant for the emergence of new, damaging pathogen genotypes (Karasov *et al.*, 2010). Alternate methods for inferring the instantaneous effective size have been proposed (Caballero, 1994; Wang, 2005). One of the most reliable methods infers *Ne* from temporal variation in allele frequency, which is studied by sequentially sampling populations across

generations (Wang, 2005; Waples, 1989). To our knowledge, this method has not yet been applied to microorganisms.

The rate of sexual versus asexual reproduction (sex rate) is also very difficult to estimate, primarily because key population genetic parameters, such as fixation indices F_{IS} or F_{ST}, have similar values in both obligate sexual and nearly completely clonal populations (de Meeûs, Balloux, 2005). The first evidence of clonality in a population is usually the detection of clonemates, i.e., sets of genotypes that are identical by descent and that have identical genomic sequences. Burt et al. (1996) first highlighted that genetic drift, which reduces clonal diversity in a predominantly asexual population, can be countered by sexual reproduction, which increases clonal diversity. On this basis, he proposed the use of the frequency of clonemates to quantify the sex rate. However, as the probability of clonemate resampling depends on the product of the sex rate and the effective population size (Burt *et al.*, 1996), this approach requires an independent estimate of the effective population size.

To untangle the interplay between the effective size and the sex rate, we extended the model of Burt et al. (1996) by including the transmission of clonal lineages and resampling within and across generations to develop a new method for estimating the sex rate and *Ne* based on the sequential sampling of a population over two generations. The clonemate resampling probability within a sample and across two generations allows for the joint estimation of the sex rate and the effective population size. We applied our method to a Chinese population of *Puccinia striiformis* f.sp. *tritici* (PST), a fungal pathogen that causes wheat yellow/stripe rust, an important disease in wheat (Ali *et al.*, 2009; Hovmøller *et al.*, 2010; Singh *et al.*, 2004). The low confidence intervals obtained for the sex rate and *Ne* estimates demonstrates the utility of the method for quantifying demographic and life-cycle parameters in microorganisms. This approach is applicable to any organism provided that the same population is sampled on two successive dates, the samples contain at least some clonemates and information on the number of generations between the two sampling dates is available. The method is

freely available online in the form of R-script "NeASE" (*Ne* And *S* Estimator) at http://ciam.inra.fr/biosp/nease.

THEORETICAL MODEL

Within generation clonality: Burt et al. (1996) postulated that the probability of resampling a clone in a population depends on the product of the sex rate and the effective population size. They derived the probability of resampling two clonemates in a fixed generation (*f*) for a population of constant size *Ne* that has a constant sex rate *s* at each cycle:

$$f' = \frac{1}{Ne}(1-s)^2 + \left(1-\frac{1}{Ne}\right) f\,(1-s)^2 \qquad [1]$$

where *f* and *f'* are the probability of resampling two clonemates in the first and second generation, respectively. The two terms on the right-hand side of Equation [1] represents the probability of two individuals being derived from clonal reproduction of i) the same parent in the last generation and ii) the same parent in any other preceding generation. At equilibrium, *f'=f*, this equation allows for the estimation of *sNe* using the formula

$$sNe \approx \frac{1}{2}\left(\frac{1}{f}-1\right). \qquad [2]$$

This equation can be used to estimate the product of the sex rate and the effective population size from the clonemate resampling probability based on sampling a single generation.

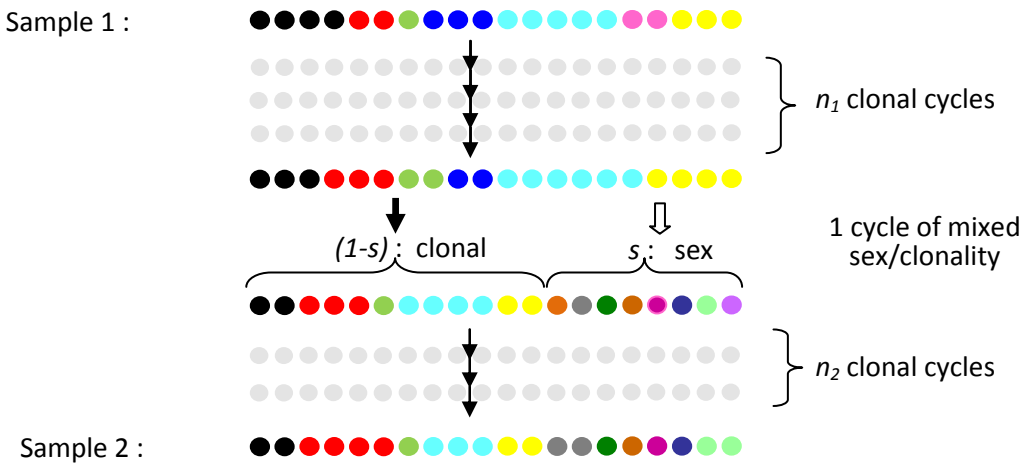


Fig. 1. The succession of n1 asexual cycles, followed by one sexual cycle and then followed by n2 asexual cycles, between the two temporally spaced populations.

Across-generation clonality: With two samples separated in time, the probability of clonemate resampling across two generations provides a way of independently estimating the sex rate and the effective size of the population. We extended the model of Burt et al. (1996) to jointly estimate the sex rate and the effective population size by including information on clonal lineage transmission and resampling within and across generations.

Consider a population reproducing with a sex rate *s* and sampled in two successive generations. Let *g* denote the probability that two random isolates, say *I* (sampled from a given generation) and *I'* (sampled from the following generation), are clonemates. This probability can be written as *g* = (probability that *I'* is not sexually derived) × [(probability that *I* is the parent of *I'*) + (probability that *I* is not the parent of *I'*) × (probability that *I* and the parent of *I'* are clonemates)]:

$$g = (1-s)\left\{\frac{1}{Ne} + \left(1-\frac{1}{Ne}\right) f\right\} = \frac{f'}{1-s}. \qquad [3]$$

A multi-cycle model: In real case studies, there may be several biological cycles during the gap between collecting the two samples. We propose incorporating this feature into the model. Suppose that after collecting the first sample, n_1 cycles of strict clonality occur (epidemic cycles) followed by one cycle of partial sexual recombination at a rate of s (1-s = the proportion of the population that continues to reproduce clonally), which is followed by n_2 additional cycles of strict clonal reproduction before the second sample is collected. In this case, n_1+1+n_2 cycles separate the two samples (Fig. 1). This sequence corresponds to a clonal pathogen passing through one overwintering sexual stage during which sexual reproduction occurs among several individuals in the population.

Let F' denote the probability of sampling two clonemates when the second sample is collected. Iterating Equation [1] for n_1+1+n_2 cycles with a sex rate equal to zero for n_1 and n_2 cycles leads to

$$F' = \begin{cases} (1-s)^2\left[\frac{1}{Ne}\left\{\sum_{i=0}^{n_1}\left(1-\frac{1}{Ne}\right)^i\right\}+\left(1-\frac{1}{Ne}\right)^{n_1+1} f\right] & \text{if } n_2 = 0 \\ \frac{1}{Ne}\left\{\sum_{i=0}^{n_2-1}\left(1-\frac{1}{Ne}\right)^i\right\}+(1-s)^2\left(1-\frac{1}{Ne}\right)^{n_2}\left[\frac{1}{Ne}\left\{\sum_{i=0}^{n_1}\left(1-\frac{1}{Ne}\right)^i\right\}+\left(1-\frac{1}{Ne}\right)^{n_1+1} f\right] & \text{if } n_2 > 0 \end{cases} \quad [4]$$

where f is still the probability of sampling two clonemates in the first sample (see supplementary material A). When n_1=0 and n_2=0, Equation [4] is identical to Equation [1].

Let G denote the probability that two random isolates, say I (from the first sample) and I' (from the second sample), are clonemates. In that case,

$$G = (1-s)\left\{\frac{1}{Ne}+\left(1-\frac{1}{Ne}\right) f\right\}. \quad [5]$$

G is equal to g in Equation [3] when there is only one cycle with sexual reproduction in the n_1+1+n_2 cycles.

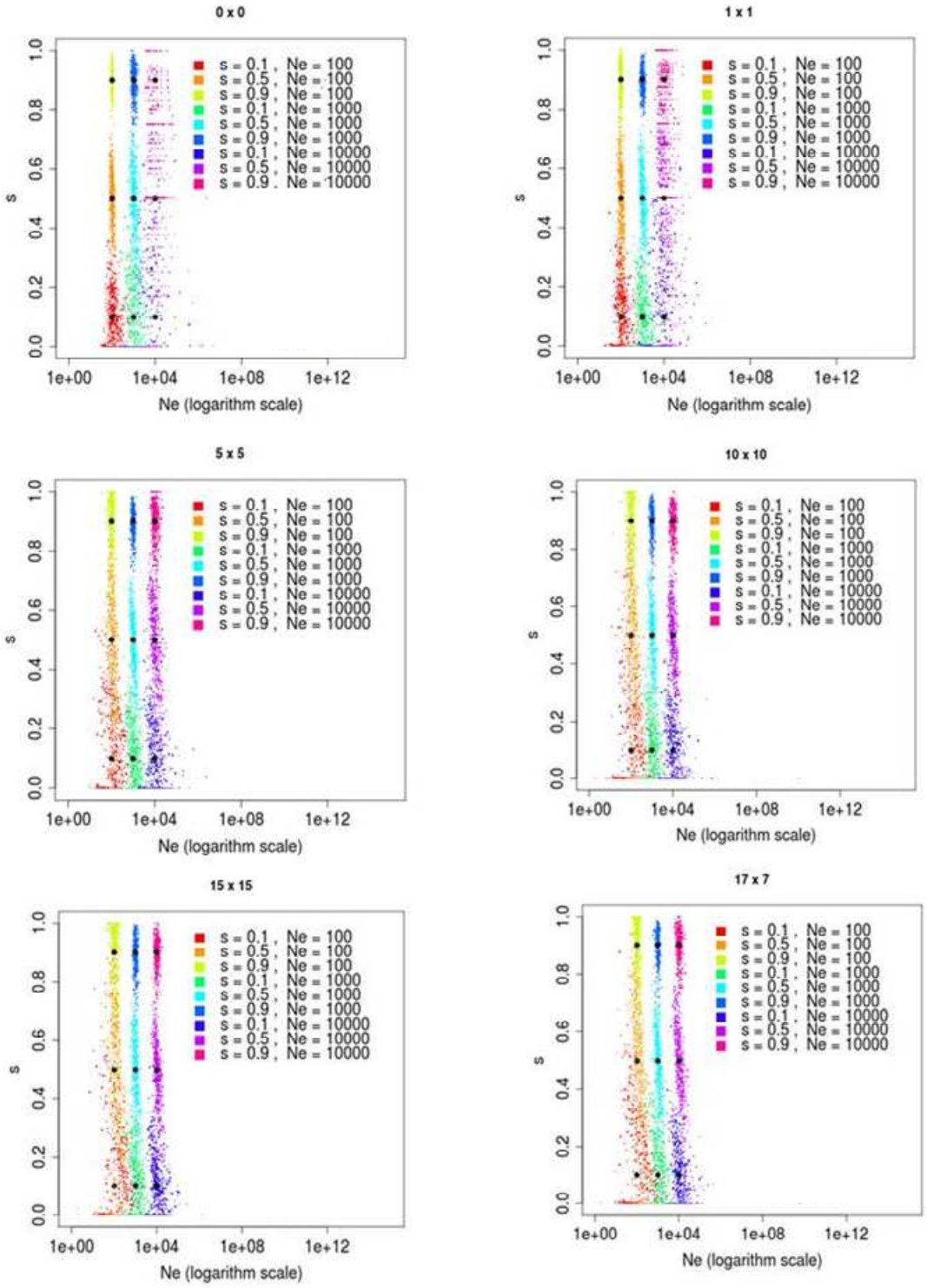


Fig. 2. Raw estimates (colored points) of effective size *Ne* and sex rate *s* obtained for simulations performed under various Ne and s values (black points) and various numbers of cycles between the two samples (n_1=n_2=0, n_1=n_2=1, n_1=n_2=5, n_1=n_2=10, n_1=n_2=15 and n_1=17 and n_2=7).

Estimation based on a contrast function: Estimates $\hat{s}$ and $\hat{Ne}$ can be obtained by minimizing the following contrast function:

$$\left(\hat{s},\hat{Ne}\right)=\text{argmin}_{(s,N)\in[0,1]\times\{1,2,3,\ldots\}}\left\{\left(F'-\tilde{F}'\right)^2+\left(G-\tilde{G}\right)^2\right\} \qquad [6]$$

where $\tilde{F}'$ and $\tilde{G}$ are the measurements of *F'* and *G* based on the two samples. *F'* and *G* are obtained through Equations [4] and [5] when *f* is replaced by its measurement, $\tilde{f}$. The expressions of asymptotically unbiased measurements of *f, F'* and *G* are provided in supplementary material B. The minimization of the contrast was performed using the algorithm from Nelder and Mead (1965). Confidence intervals and, when necessary, bias correction for $\hat{s}$ and $\hat{Ne}$ can be calculated using a parametric bootstrap approach (Efron, Tibshirani, 1993). The NeASE code provides both raw and bias-corrected point estimates of *s* and *Ne*, confidence intervals (CIs) and their actual coverage.

PERFORMANCE IN SIMULATIONS

Simulations were used to assess the performance of the method when *Ne*, *s*, n_1 and n_2 varied (see supplementary material C). The accuracy of our method was assessed on data sets generated using forward and individual-centered simulations under various parameter values. Five hundred data sets were generated for each value of *s* in [0.1, 0.5, 0.9] and each value of *Ne* in [10^2, 10^3, 10^4]. The contrast function (Equation [6]) was minimized for each simulated data set. Fig. 2 illustrates the distribution of estimates for various values of n_1 and n_2, including n_1=17 and n_2=7 (which correspond to the real case studied hereafter). The method had overall satisfactory results with respect to bias and variance and was generally able to distinguish the values of *Ne* and *s* that were used to generate the simulations. Inaccuracies were observed only under very specific parameter conditions; the relative bias and variance were large with small values for *Ne* and *s* and large values for *n1* and *n2* as well as when *Ne* was large and n_1 and n_2 were small.

Table 1. Description of temporally sampled *Puccinia striiformis* f.sp. *tritici* (PST) populations (2004 and 2005) from China, used for effective population size and clonality rate estimation.

		2004	2005
Sampling scheme	**Number of individuals**	268	268
	Number of sites sampled[1]	3	3
Clonality	**Distinct MLGs**	232	220
	Number of clones re-sampled	36	48
	Frequency of most abundant MLG	18	12
	Raw estimate of clonemates re-sampling probability	$f = 0.057$	$f' = 0.050$
Genetic diversity	**Expected heterozygosity ***	0.274	0.276
	Observed heterozygosity	0.287	0.288
	Mean number of alleles per locus	3.389	3.556
	Genotypic diversity	0.994	0.995

* There is a non-significant difference between expected and observed heterozygosity, demonstrating its recombinant structure. [1] The sites sampled shows the three geographical regions, which remained the same for the two years, see supplementary file Fig S3.

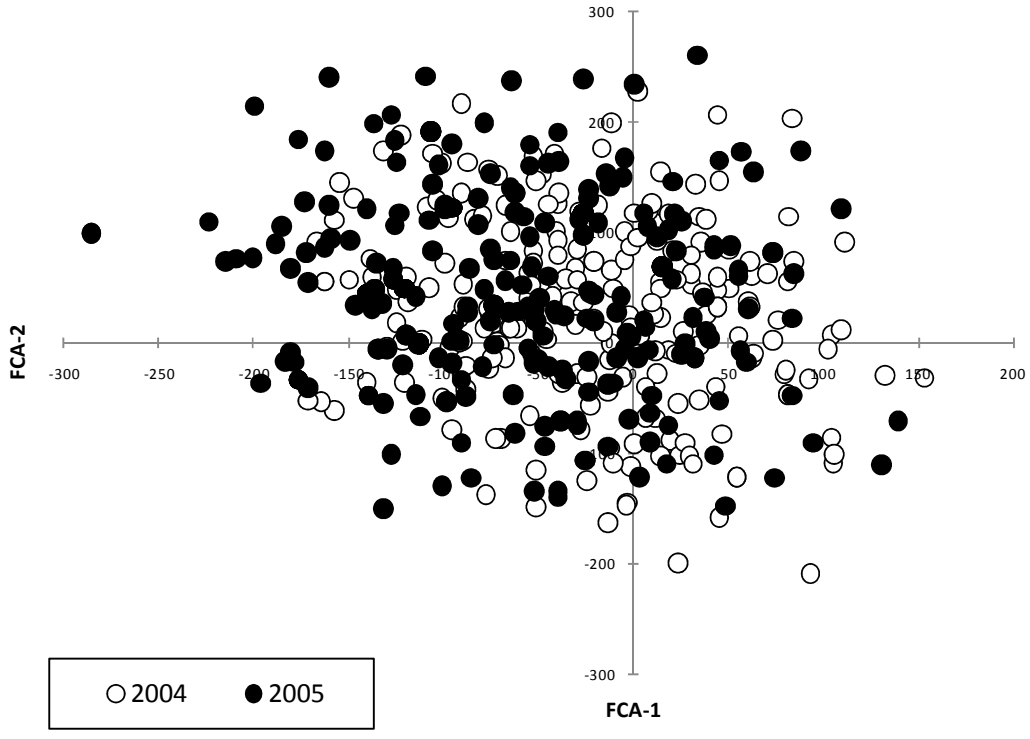


Fig. 3. Factorial correspondence analysis for Chinese PST populations sampled in two years (2004 and 2005).

Small values for the *Ne* and *s* effectively led to the fixation of single clone in the population, while large values for the *Ne* maintained a high degree of diversity and resulted in very little clone resampling. Increasing n_1 and n_2 resulted in a higher relative contribution from sexual reproduction and therefore resulted in a lower proportion of clonemates. Overall, the estimation method seems to be quite accurate. There are, however, for particular values of *s*, *Ne*, n_1 and n_2, some inaccuracies. Table S4 provides medians and interquartile ranges of parameter estimates (those shown in Fig. 2). The medians of estimates are relatively close to true values of *Ne* and *s* whatever these true values are. However, the variation in the estimations may be large for a few combinations of parameters as shown by the interquartile ranges. Indeed, for small n_1+n_2 and $Ne=10^4$, the method leads to strongly varying estimates (large interquartile ranges). Besides, when *s* is small, the estimates of *Ne* are strongly varying and can even be excessively large (see the cloud of blue points around $Ne=10^{10}$). Finally, it has to be noted that the measurements proposed in Appendix B led to better results than the raw empirical measurements (results not shown).

APPLICATION TO A FUNGAL PATHOGEN

We used our method to estimate the sex rate and the *Ne* of a natural population of wheat yellow rust (*Puccinia striiformis* fsp *tritici*: PST) from the Gansu province of China. In many countries, PST reproduces exclusively asexually on its main (uredial) host, wheat; seasonal epidemics encompass multiple cycles of asexual propagation. At the end of the crop season, teliospores can be produced and germinate to form haploid basidiospores, but no alternate (called “aecial”) host has so far been described as being infected by these spores. The sexual phase of the life cycle of PST has only recently been documented on *Berberis spp*. under laboratory conditions (Jin *et al.*, 2010). Under natural conditions, sexual reproduction is suspected to occur in populations that exhibit recombination footprints in their genetic structure (Ali *et al.*, in preparation; Ali *et al.*, 2010), such as in Gansu (Mboup *et al.*, 2009), but an alternate host species has not yet been identified. The sexual phase would occur once a year through the production of sexual spores that infect the alternate host. Therefore, a PST population includes a proportion *p* of clonal individuals that result from survival on volunteers

(Ali *et al.*, 2010) and a proportion 1-*p* that result from sexual reproduction. To estimate the contribution of the sexual cycle to the population biology of PST in the Tianshui area of Gansu, we analyzed the genetic and clonal diversity in samples separated by two years, which were taken before and after the putative sexual cycle.

Table 2. Raw and bias corrected estimates of *s* and *Ne* as well as the 95% confidence intervals of the parameters and the coverage rate of these intervals (250 iterations of parametric bootstrap). The number of cycles separating the two samples is assumed to be 28 (n_1=17 and n_2=7).

	Raw estimates		Bias-corrected estimates	
	s	*Ne*	*s*	*Ne*
Estimate	0.745	1950	0.756	1735
95% Confidence interval	0.383-0.950	1450-3510	0.469-1.000	675-2800
Coverage	0.93	0.90	0.90	0.94

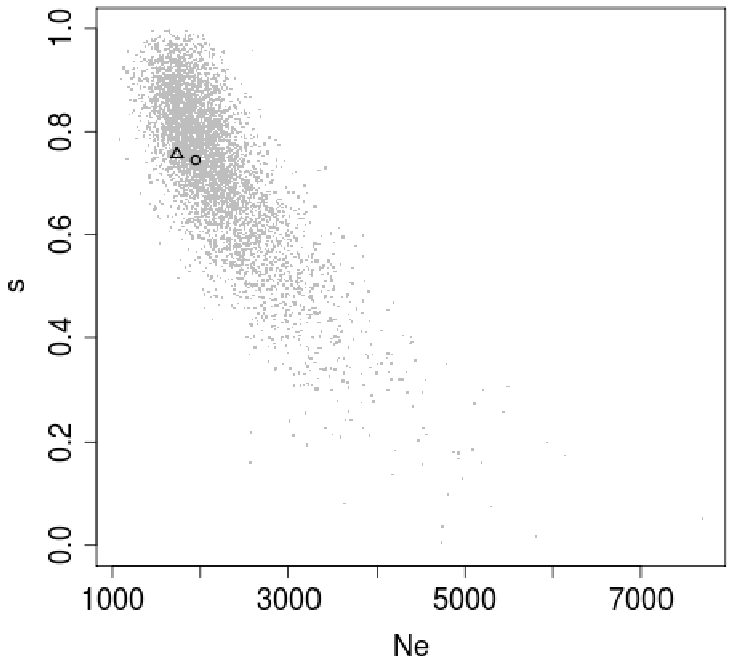


Fig. 4. Sex rate (s) plotted against effective population size (*Ne*). Raw estimation (open circle), bias corrected estimation (triangle) and estimations obtained for the 500 bootstrap samples (dots).

Two temporally separated samples of a Gansu PST population were collected in 2004 and 2005. The sample collections were separated by 18 months and covered one cycle of sexual recombination (early summer) and 24 clonal cycles (estimated as explained below). The genotyping of 536 isolates with 20 SSR markers revealed 17 polymorphic SSRs with a variable number of alleles (ranging from 2 to 10) and genetic diversity (ranging from 0.0037 to 0.688; Table SI-1, supporting information). Neither intra- nor inter-population subdivisions were observed via STRUCTURE and factorial correspondence analysis (Fig. 3). The resulting F_{ST} values were extremely low and non-significant (Table SI-2). This confirms that the two samples belonged to a single population. A high genotypic diversity (Table 1), a lack of deviation from Hardy-Weinberg equilibrium (Fig SI-1) and a very low overall linkage disequilibrium (*Ld*) between markers (Table SI-3) confirmed that recombination occurred in both populations. Clonal populations are expected to exhibit exactly the opposite indices (Gladieux *et al.*, 2010; Milgroom, 1996). Despite the presence of recombination, the resampling of clonemates within and across the years revealed the persistence of clones in the area throughout the time covered by the survey (Fig. SI-2). We confirmed the absence of any correlation in the distribution of the identified clones, as the preferential resampling of clones within a same field, or neighbor fields, in a given year, or between the two sampled years (data no shown).

Sex rate and Ne estimates: The sex rate and the *Ne* of the PST population were inferred using empirical estimates of the clonemate resampling probability at the given generation ($\tilde{f}$ = 0.057 for 2004 and $\tilde{F}'$ = 0.050 for 2005) and the inter-generation clonemate resampling probability ($\tilde{G}$ = 0.0014). To infer the number of generations that occurred between sample collections we considered that asexual production of urediopores can take from 10 days to 2 months depending on the temperature. The expected number of asexual cycles occurring between the two samplings was estimated using the Zadoks method (Rapilly, 1976; Rapilly, 1979; Zadoks, 1961) and the recorded temperatures for the period. Because sexual spores are released in the Spring, the 18 months separating the two sampling dates corresponds to n_1=17 clonal generations before the sexual cycle and n_2 =7

clonal generations after the sexual cycle. When applying NeASE to the PST population data (Fig. 4 and Table 2), we estimated *s* = 74% (95% CI: 38-95%) and *Ne* = 1735 (95% CI: 675-2800). Very satisfactory coverage of the confidence interval was obtained (over 90% of the simulations had the true values of *Ne* and *s* fall within their 95% confidence intervals).

Table 3. Various estimates of effective population size (*Ne*) for two temporally sampled PST populations (2004 and 2005) from China.

	Method used	Software used	Year	Sample size	Ne	CI	Method Reference
Temporal estimate*	**Clonal resampling based**		-	536	1735	675-2800	This article
	Moments based	**MLNe**	-	536	779	-	(Waples, 1989)
		NeEstimator	-	536	4748	1958-23923	
	Likelihood based	**MLNe**	-	536	999	833-1000	(Wang and Whitlock, 2003)
	MCMC based	**MCLEEPS**	-	536	1300	1000-1600	(Anderson et al., 2000)
Historic/one sample estimation	**Heterozygosity excess**	**NeEstimator**	2004	268	159	-	(Pudovkin et al. 1996)
			2005	268	96	-	
	Linkage disequilibrium	**NeEstimator**	2004	268	558	350-1193	(Hill, 1981)
			2005	268	128	108-155	
		LDNe	2004	268	82	67-99	(Waples and Do, 2010)
			2005	268	44	38-50	
	Theta based		2004	268	487	-	Ohta and Kimura (1973)
			2005	268	525	-	

* The number of generations between two samples was 28

We compared the *Ne* value to the estimates produced by other available methods (Table 3). The *Ne* value fell in the range of the estimates obtained using temporally spaced samples (range from 779 to 4748, Table 3). As with our method, these methods, which are based on temporal change in allele frequencies, give "instantaneous/contemporary" estimates of the effective population size. Additional methods based on the genetic parameters of a single population (e.g., excess of heterozygosity, linkage disequilibrium and genetic diversity) reflect ancient demographic changes (Wang, 2005). These "historic" estimates were lower than the temporal methods and ranged from 44 to 552 (Table 3). This discrepancy may result from the occurrence of rare bottlenecks or selective sweeps in the ancient past of the population that affect the "historic" estimates but not the instantaneous ones (Wang, 2005). Clonality, which is not accounted for in the historic approaches, may also contribute to the discrepancy. Because of the absence of alternative methods for estimating *s*, the reliability of the inferred sex rate was confirmed through simulation-based criteria (as discussed above).

DISCUSSION

To our knowledge, the method we developed is the first joint estimate of sex rate and effective population size performed on a natural population, despite the importance of the two parameters in shaping the genetic diversity of populations. Sex rate has been only estimated once in a plant pathogen population, albeit using an artificial inoculation with traceable genotypes (Zhan *et al.*, 2001) and the method employed was the subject of controversy (Brown, 2000; Zhan *et al.*, 2000; Zhan *et al.*, 2001). Recent estimates have been performed on yeasts using genomic sequences (Tsai *et al.*, 2008; Zeyl, Otto, 2007), but these methods integrate the long-time, species-wide, evolutionary history and do not provide estimates at the population level. Although methods have been proposed for testing the presence or absence of recombination (de Meeûs, Balloux, 2005; de Meeûs *et al.*, 2007), no quantitative methods have been available to date. The development of population-level estimates was limited by the fact that most of the relevant genetic parameters are only affected by a very strong level of clonality, while the structure of a population with low rates of recombination cannot be

distinguished from that of a population with 100% sexual reproduction (Balloux *et al.*, 2003). This effect is illustrated clearly in our rust study, where clone resampling was the only effect of clonality detected (no deviation from Hardy-Weinberg equilibrium). As the presence of clonemates in a dataset affects diversity level, linkage disequilibrium and population genetic structure, a usual practice in population genetics studies is to perform a double analysis, on a full dataset first and then to apply a "clone-correction" by keeping a single copy of each multi-locus genotype (Halkett *et al.*, 2005). NeASE method has the distinct advantage that it does not require any double analysis of the data, as it uses the clonal resampling as a direct input for estimating the sex rate and effective population size. Like any population genetic inference, the NeASE estimator relies upon various assumptions such as the absence of selection, spatial structure or migration. However, instantaneous estimates based on population resampling are mainly affected by strong selective pressures, as few generations are involved. In such cases genetic analyses allow testing for the presence of spatial structure or gene flow in the sample, as shown in our case-study. Specific attention should thus be paid to the sampling scale, in order to avoid oversampling of clones due to a patchy distribution (leaf or plant level in our Pst case). One central and main difficulty involves the estimation of the number of clonal generations between sampling events and recombination (*n1* and *n2*). Note that this caveat also concerns any effective size estimate using the "Wapples-like" sampling scheme (samplings spaced in time), when performed on inconspicuous microbial organism. Thus, the method is best applied to populations where the number of generations and the timing of sexuality can be fairly accurately estimated. Note that when testing for the effect of inaccuracies in values of *n1* and *n2*, only small effects were observed for mean and variance of estimates. Another important model assumption is constant population size assumed by the model itself. For fluctuating populations, as often occur during epidemic growth of microorganisms, the estimated size will be the geometric mean size of successive sizes between the two sampled generations. When there is a severe bottleneck, as observed for PST populations in summer, when the host disappears (wheat harvest), estimated size will be close to this bottleneck size.

In our case study, the NeASE method presents the same range of variation, but with more reliable estimates than the multilocus based estimators. To date, the most reliable estimates on mating system are based on progeny analyses. However, in most microbial populations progeny analysis is impossible, both because progeny cannot be isolated or harvested, and because sexual/clonal reproduction cycles are not occurring at the same place or time. The NeASE method therefore provides a simple and rapid means of estimating the rate of sex and effective population size for any organism providing sampling at two different dates and the existence of at least some direct evidence of clonality through clonemate resampling. The resampling probabilities developed here will mostly be relevant for inferential methods (e.g., Approximate Bayesian Computation approaches) used to investigate the contemporary and historical structure of pathogen populations. The impact of alternating clonal and sexual reproduction on population structure is crucial to the fields of agricultural and medical molecular epidemiology, as many pathogens display such lifecycles (Gladieux *et al.*, 2010; Milgroom, 1996). Obtaining information on pathogen population size and levels of clonality is also important for mathematical epidemiology and modeling (McDonald, Linde, 2002; Regula *et al.*, 2009) and critical for understanding the causes for pathogen emergence and spread.

MATERIALS AND METHODS

Genetic description of two temporally spaced populations: A total of 536 isolates from two temporally spaced PST populations (268 isolates from each, taken from 44 fields in 2004 and 30 fields in 2005, and different plants within field) sampled during May 2004 and December 2005 in Gansu Province, China, at three geographical regions (North-East, North-West and South; supplementary file Fig. SI-3) were genotyped using 20 SSR markers in three multiplex reactions following Ali et al. (2011; ANNEX II of thesis). The suitability of the markers was assessed by plotting the number of recurrent multi-locus genotypes (MLGs) detected by resampling the 20 SSR loci under panmixia and reached an asymptote with the full set of SSRs (Fig. S4), suggesting that the set of markers used was

sufficient for accurately genotyping the populations (GENECLONE software; Arnaud-Haond, Belkhir, 2007). The recombination signature and level of population polymorphism were assessed using the number of alleles per loci, the linkage disequilibrium among loci, the observed and expected heterozygosity levels (GENETIX 4.05.2 software; Belkhir *et al.*, 2004) and the estimation of the genotypic diversity and the index of association for disequilibrium among the loci (MULTILOCUS software; Agapow, Burt, 2001). Departures from Hardy-Weinberg equilibrium were investigated using the exact test (GENEPOP 4.0 software; Raymond, Rousset, 1995). The spatial population structure was analyzed through F_{ST} (GENETIX 4.05.2 software; Belkhir *et al.*, 2004) and non-parametric multivariate analyses (factorial correspondence analysis, FCA; (Belkhir *et al.*, 2004). These estimations were also made using clone-corrected data to avoid an impact from recent clonality on our estimates as a result of epidemic clonal structure (Maynard-Smith *et al.*, 1993).

Estimation of the number of generations: The number of generations between two samples was estimated using the method proposed by Zadoks (Rapilly, 1976; Rapilly, 1979; Zadoks, 1961). This method estimates the latent period (the time from spore penetration to first sporulation) as a function of the temperature. The temperature determines the length of an asexual PST generation. We estimated the total number of generations from May 1, 2004, to December 31, 2005, assuming no generational overlap and the presence of wheat throughout the year.

Estimation of the sex rate: We estimated the sex rate using our method (see *Theoretical model development* above) with the information on the probability of clonemate resampling within (f) and across generations (g). The clonal origin of these clonemates was confirmed using the *psex* value, which is the probability of a given MLG being observed in two independent sexual progenies (Arnaud-Haond, Belkhir, 2007). An empirical estimate of f and g was obtained from the data as the probability of re-sampling each clonemate while summing the probability of all MLGs sampled at least twice.

Effective population size: We also estimated the effective population size using previously published methods. Both temporally separated populations were used to estimate the "instantaneous" or short-term effective population size and the "historic" or long-term effective population size for each population individually. Temporal estimates were based on the change in allele frequencies between the two temporally repeated samples (2004 and 2005) using the moment-based point estimate proposed by Waples (1989; Ne-ESTIMATOR software), the likelihood approach used by Wang and Whitlock (MlNe software; 2003) and an MCMC-based approach (MCLEEPS software; Anderson *et al.*, 2000). Furthermore, because recent clonal reproduction could affect the effective population size, both the full dataset and the clone-corrected data sets were used to estimate the *Ne* with these previously developed methods.

"Historic" or single-sample estimates are calculated using three parameters: the amount of heterozygosity in the population, the linkage disequilibrium among the different loci and the amount of diversity in the population. The heterozygosity excess-based method (following Pudovkin *et al.*, 1996) implemented in the *Ne*-ESTIMATOR software and the linkage disequilibrium-based method implemented in both the *Ne*-ESTIMATOR software (following Hill, 1981) and the LDNe software (following Waples, 2006) were used. Estimation of the *Ne* based on the diversity observed at the microsatellite markers was obtained according to Ohta and Kimura's estimate of the population mutation parameter θ (Ohta, Kimura, 1973).

ACKNOWLEDGEMENTS

We are thankful to Laurent Gérard and Nathalie Galet for their technical assistance in the maintenance of the yellow rust collection. We would like to acknowledge Bruce McDonald, Tatiana Giraud, Tony Brown, Etienne Klein and Isabelle Goldringer for their constructive comments on the manuscript. This work received financial support from the European Integrated Project BIOEXPLOIT, FOOD-CT-

2005-513959, ANR-07-BDIV-003 EMERFUNDIS and ANR-09-BLAN-01 EMILE. Sajid Ali was funded by a grant from the Higher Education Commission, The Government of Pakistan.

AUTHOR CONTRIBUTIONS

SA performed the molecular genotyping. JE and AG supervised the molecular genotyping. SA, PG and JE conducted the population genetic analyses. SS and JE developed the theoretical model. ML, MM, JE, CP and WC contributed to the sampling and multiplication of the isolates. SA, ML and CP estimated the number of PST asexual generations based on temperature data. SA, JE, SS and PG prepared the manuscript. JE and CP designed the study. All of the authors have read and approved the manuscript.

REFERENCES

Agapow PM, Burt A (2001) Indices of multilocus linkage disequilibrium. *Molecular Ecology Notes* **1**, 101-102.

Ali S, Enjalbert J, Gladieux P, Leconte M, Gautier A, Justesen AF, Hovmoller MS, De Vallavieille-Pope C (in preparation) Origin, migration routes and genetic structure of worldwide populations of the wheat yellow rust pathogen.

Ali S, Gautier A, Leconte M, Enjalbert J, De Vallavieille-Pope C (2011) A rapid genotyping method for an obligate fungal pathogen, *Puccinia striiformis* f.sp. *tritici*, based on DNA extraction from infected leaf and Multiplex PCR genotyping. *BMC Research Notes* **4**, 240.

Ali S, Leconte M, Walker A-S, Enjalbert J, de Vallavieille-Pope C (2010) Reduction in the sex ability of worldwide clonal populations of *Puccinia striiformis* f.sp. *tritici*. *Fungal Genetics and Biology* **47**, 828-838.

Ali S, Shah SJA, Rahman H (2009) Multi-locations variability in Pakistan for partial resistance in wheat to *Puccinia striiformis* West. *tritici*. *Phytopathologia Mediterranea* **48**, 269–279.

Anderson EC, Williamson EG, Thompson EA (2000) Monte Carlo evaluation of the likelihood for Ne from temporally-spaced samples. *Genetics* **156**, 2109-2118.

Arnaud-Haond S, Belkhir K (2007) GenClone 1.0: A new program to analyse genetics data on clonal organisms. *Molecular Ecology Notes* **7**, 15-17.

Awadalla P (2003) The evolutionary genomics of pathogen recombination. *Nature Reviews Genetics* **4**, 50-60.

Balloux F, Lehmann L, De Meeus T (2003) The population genetics of clonal and partially clonal diploids. *Genetics* **164**, 1635-1644.

Belkhir KBP, Chikhi L, Raufaste N, Bonhomme F (2004) GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5000: Université de Montpellier II, Montpellier (France) http://www.univ-montp2.fr/~genetix/genetix/genetix.htm.

Brown JKM (2000) Estimation of rates of recombination and migration in populations of plant pathogens. *Phytopathology* **90**, 320-323.

Burt A, Carter DA, Koenig GL, White TJ, Taylor JW (1996) Molecular markers reveal cryptic sex in the human pathogen *Coccidioides immitis*. *Proceedings of the National Academy of Science of the United States of America* **93**, 770-773.

Caballero A (1994) Developments in the prediction of effective population size. *Heredity* **73**, 57—679.

Charlesworth B (2009) Effective population size and patterns of molecular evolution and variation. *Nature Reviews Genetics* **10**, 195-205.

de Meeûs T, Balloux F (2005) F-statistics of clonal diploids structured in numerous demes. *Molecular Ecology* **14**, 2695-2702.

de Meeûs T, Prugnolle F, Agnew P (2007) Asexual reproduction: genetics and evolutionary aspects. *Cellular and Molecular Life Science* **64**, 1355–1372.

Efron B, Tibshirani RJ (1993) *An Introduction to the Bootstrap* Chapman and Hall, New York.

Fisher MC (2007) The Evolutionary Implications of an Asexual Lifestyle Manifested by *Penicillium marneffei*. In: *Sex in fungi: molecular determination and evolutionary implications.* (ed. Heitman JK, J.W.; Taylor, J.W.; Casselton, L.A.), p. 542. ASM Press, Washington.

Gladieux P, Byrnese E, Fisher MC, Aguleta G, Heitman J, T. G (2010) Epidemiology and evolution of fungal pathogens, in plants and animals. In: *Genetics and Evolution of Infectious Diseases* (ed. M T), p. 772. Elsevier.

Glémin S (2003) How are deleterious mutations purged? Drift versus nonrandom mating. *Evolution* **57**, 2678-2687.

Goddard MR, Godfray HCJ, Burt A (2005) Sex increases the efficacy of natural selection in experimental yeast populations. *Nature* **434**, 636-640.

Halkett F, Simon J, Balloux F (2005) Tackling the population genetics of clonal and partially clonal organisms. *Trends in Ecology & Evolution* **20**, 194-201.

Hill WG (1981) Estimation of effective population size from data on linkage disequilibrium. *Genetic Research* **38**, 209-216.

Hovmøller MS, Walter S, Justesen AF (2010) Escalating threat of wheat rusts. *Science* **329**, 369.

Jin Y, Szabo LJ, Carson M (2010) Century-old mystery of Puccinia striiformis life history solved with the identification of Berberis as an Alternate Host. *Phytopathology* **100**, 432-435.

Karasov T, Messer SW, Petrov DA (2010) Evidence that adaptation in *Drosophila* is not limited by mutation at single sites. *PLoS Genet* **6**, e1000924. doi:1000910.1001371/journal.pgen.1000924.

Maiden MCJ, Bygraves JA, Feil E, Morelli G, Russell JE, Urwin R, Zhang Q, Zhou JJ, Zurth K, Caugant DA, Feavers IM, Achtman M, Spratt BG (1998) Multilocus sequence typing: A portable approach to the identification of clones within populations of pathogenic microorganisms. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 3140-3145.

Maynard-Smith J, Smith NH, O'rourke M, Spratt BG (1993) How clonal are bacteria? **90**, 4384-4388.

Mboup M, Leconte M, Gautier A, Wan AM, Chen W, de Vallavieille-Pope C, Enjalbert J (2009) Evidence of genetic recombination in wheat yellow rust populations of a Chinese oversummering area. *Fungal Genetics and Biology* **46**, 299-307.

McDonald BA, Linde C (2002) Pathogen population genetics, evolutionary potential, and durable resistance. *Annual Review of Phytopathology* **40**, 349-379.

Milgroom MG (1996) Recombination and the multilocus structure of fungal populations. *Annual Review of Phytopathology* **34**, 457-477.

Nelder JA, Mead R (1965) A simplex method for function minimization. *Computation Journal* **7**, 308–313.

Ohta T, Kimura M (1973) A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a finite population. *Genetic Resources Cambridge* **22**, 201-204.

Paland S, Schmid B (2003) Population size and the nature of genetic load in *Gentianella germanica*. *Evolution* **57**, 2242-2251.

Pudovkin AI, Zaykin DV, Hedgecock D (1996) On the potential for estimating the effective number of breeders from heterozygote-excess in progeny. *Genetics* **144**, 383-387.

Rapilly F (1976) Essai d'explication de l'epidemie de rouille jaune sur ble en 1975. *Sélectionneur Français* **22**, 47-52.

Rapilly F (1979) Yellow rust epidemiology. *Annual Review of Phytopathology* **17**, 59-73.

Raymond M, Rousset F (1995) GENEPOP Version 1.2: population genetics software for exact tests and ecumenicism. *Journal of Heredity* **86**, 248-249.

Regula T, Medlock J, Galvani A (2009) The discounted reproductive number for epidemiology. *Methematical Biosciences and Engineering* **6**, 379-395.

Singh RP, William HM, Huerta-Espino J, Rosewarne G (2004) Wheat rust in Asia: meeting the challenges with old and new technologies

Taylor JW, Fisher MC (2003) Fungal multilocus sequence typing - it's not just for bacteria. *Current Opinion in Microbiology* **6**, 351-356.

Taylor JW, Jacobson DJ, Fisher MC (1999) The evolution of asexual fungi: Reproduction, speciation and classification. *Annual Review of Phytopathology* **37**, 197-246.

Tsai IJ, Bensasson D, Burt A, Koufopanou V (2008) Population genomics of the wild yeast Saccharomyces paradoxus: Quantifying the life cycle. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 4957-4962.

Wang J (2005) Estimation of effective population sizes from data on genetic markers. *Philosophical Transaction of the Royal Society B* **360**, 1395-1409.

Wang JL, Whitlock MC (2003) Estimating effective population size and migration rates from genetic samples over space and time. *Genetics* **163**, 429-446.

Waples RS (1989) A generalized approach for estimating effective population size from temporal changes in allele frequency. *Genetics* **121**, 379-391.

Waples RS (2006) A bias correction for estimates of effective population size based on linkage disequilibrium at unlinked gene loci. *Conservation Genetics* **7**, 167-184.

Williams PD (2010) Darwinian interventions: taming pathogens through evolutionary ecology. *Trends in Parasitology* **26**, 83-92.

Zadoks JC (1961) Yellow rust on wheat, studies in epidemiology and physiologic specialization. *Tijdschrift over Plantenziekten* **67**.

Zeyl CW, Otto SP (2007) A short history of recombination in yeast. *Trends in Ecology & Evolution* **22**, 223-225.

Zhan J, Mundt CC, McDonald BA (2000) Estimation of rates of recombination and migration in populations of plant pathogens-A reply. *Phytopathology* **90**, 324-326.

Zhan J, Mundt CC, McDonald BA (2001) Using restriction fragment length polymorphisms to assess temporal variation and estimate the number of ascospores that initiate epidemics in field populations of *Mycosphaerella graminicola*. *Phytopathology* **91**, 1011-1017.

Appendix A: Expression of F' (Eq. [4])

Let f_n denote the probability of sampling two clonemates after cycle *n* (if the population after cycle *n* would be sampled); $f_0 = f$ and $f_{n_1+1+n_2} = F'$ correspond to the two actually sampled populations. For the first n_1 cycles, the Burt et al (1996) formula without sexuality is $f_n = (1/N) + (1-1/N) f_{n-1}$ with *n* in {1,…,n_1} if n_1>0. Iterating this formula leads to [A1] (for n_1>0):

$$f_{n_1} = (1/N)\left(\sum_{n=0}^{n_1-1}(1-1/N)^n\right) + (1-1/N)^{n_1} f_0 \qquad \text{[A1]}$$

Then, at cycle n_1+1, the Burt et al (1996) formula including sexuality gives [A2]:

$$f_{n_1+1} = (1-s)^2\left\{(1/N) + (1-1/N) f_{n_1}\right\} \qquad \text{[A2]}$$

Finally, like above, for the last n_2 cycles, there is no sexuality and we get [A3] (for n_2>0):

$$f_{n_1+1+n_2} = (1/N)\left(\sum_{n=0}^{n_2-1}(1-1/N)^n\right) + (1-1/N)^{n_2} f_{n_1+1} \qquad \text{[A3]}$$

Combining [A1], [A2] and [A3] leads to the equation [4]:

$$F' = \begin{cases} (1-s)^2\left[\dfrac{1}{Ne}\left\{\sum_{i=0}^{n_1}\left(1-\dfrac{1}{Ne}\right)^i\right\} + \left(1-\dfrac{1}{Ne}\right)^{n_1+1} f\right] & \text{if } n_2 = 0 \\ \dfrac{1}{Ne}\left\{\sum_{i=0}^{n_2-1}\left(1-\dfrac{1}{Ne}\right)^i\right\} + (1-s)^2\left(1-\dfrac{1}{Ne}\right)^{n_2}\left[\dfrac{1}{Ne}\left\{\sum_{i=0}^{n_1}\left(1-\dfrac{1}{Ne}\right)^i\right\} + \left(1-\dfrac{1}{Ne}\right)^{n_1+1} f\right] & \text{if } n_2 > 0 \end{cases}$$

Appendix B: Measurements of within- and across-generation clonality *f*, *F'* and *G*

The sex rate *s* and the effective size *Ne* are estimated by minimizing a contrast function (see Eq. [6]) which depends on measurements $\tilde{f}$, $\tilde{F}'$ and $\tilde{G}$ of *f*, *F'* and *G*. Let Q_i (resp. Q_i') be the effective number of individuals of clone *i* in the population from which the first (resp. second) sample was drawn. The index *i* varies from 1 to *I* (*I* is the number of clonal lineages in both populations) and $\sum_{i=1}^{I} Q_i = \sum_{i=1}^{I} Q_i' = Ne$. The probability *f*, *F'* and *G* satisfy:

$$f = \sum_{i=1}^{I} \frac{Q_i(Q_i-1)}{Ne(Ne-1)}, \quad F' = \sum_{i=1}^{I} \frac{Q_i'(Q_i'-1)}{Ne(Ne-1)} \quad \text{and} \quad G = \sum_{i=1}^{I} \frac{Q_i Q_i'}{Ne^2}.$$

The raw (empirical) estimates of *f*, *F'* and *G* are

$$\breve{f} = \sum_{i=1}^{\breve{I}} \frac{q_i(q_i-1)}{m(m-1)}, \quad \breve{F}' = \sum_{i=1}^{\breve{I}} \frac{q_i'(q_i'-1)}{m'(m'-1)} \quad \text{and} \quad \breve{G} = \sum_{i=1}^{\breve{I}} \frac{q_i q_i'}{mm'},$$

where *m* (resp. *m'*) is the size of the first (resp. second) sample, q_i (resp. q_i') is the number of individuals of clone *i* in the first (resp. second) sample, and $\breve{I}$ is the total number of clonal lineages in both samples. $\breve{I}$ tends to *I* when the sample sizes *m* and *m'* increase.

The limit expectation (when $\breve{I}$ goes to *I*) of $\breve{G}$ is *G*. However, the estimates $\breve{f}$ and $\breve{F}'$ are asymptotically biased: $E\left(\breve{f}\right) = \sum_{i=1}^{\breve{I}} (Q_i/Ne)^2 \xrightarrow[\breve{I}\to I]{} \sum_{i=1}^{I} (Q_i/Ne)^2$ and $E\left(\breve{F}'\right) = \sum_{i=1}^{\breve{I}} (Q_i'/Ne)^2 \xrightarrow[\breve{I}\to I]{} \sum_{i=1}^{I} (Q_i'/Ne)^2$. Thus, we used $\tilde{G} = \breve{G}$ as an estimate of *G* but we searched for other estimates of *f* and *F'*.

The probability *f* can be written: $f = \frac{Ne}{Ne-1}\sum_{i=1}^{I}\left(\frac{Q_i}{Ne}\right)^2 - \frac{1}{Ne-1}$. Consequently, we propose to estimate *f* by the asymptotically unbiased estimate:

$$\tilde{f} = \frac{Ne}{Ne-1}\breve{f} - \frac{1}{Ne-1},$$

and we considered a similar estimate of *F'*:

$$\tilde{F}' = \frac{Ne}{Ne-1}\breve{F}' - \frac{1}{Ne-1},$$

These estimates cannot be used in general because *Ne* is unknown, however, they can be used in the contrast minimization. Indeed, by explicitly writing the dependences in *Ne*, *s* and the data, the contrast satisfies:

$$\left(F'-\tilde{F}'\right)^2 + \left(G-\tilde{G}\right)^2 = \left(F'\left(Ne, s, \tilde{f}\,(Ne, \text{data})\right) - \tilde{F}'(Ne, \text{data})\right)^2 + \left(G\left(Ne, s, \tilde{f}\,(Ne, \text{data})\right) - \tilde{G}\,(\text{data})\right)^2$$

and can be minimized with respect to *Ne* and *s*.

Appendix C: Simulation-based study of the performance of the estimation method

The method of estimation of the sex rate *s* and the effective population size *Ne* was applied to data sets generated under the simulation model constructed for the parametric bootstrap approach. The simulation model was run under various values of the sex rate, the effective size and the number of cycles. For each value of *s* in [0.1, 0.5, 0.9] and each value of *Ne* in [10^2, 10^3, 10^4], 500 data sets were generated with $n_1=n_2=0$, $n_1=n_2=1$, $n_1=n_2=5$, $n_1=n_2=10$, $n_1=n_2=15$ and $n_1=15$ and $n_2=12$ (like for the real data set). Then, the contrast function, equation [6], was minimized for each simulated data set. In all the cases, the sample sizes were equal to the sample sizes in the real data, that is to say 268 isolates per sample. Fig. S5 (below) shows the estimations which were obtained (these are the raw estimations, not the bias corrected estimations).

Overall, the estimation method seems to be quite accurate. There are, however, for particular values of *s*, *Ne*, n_1 and n_2, some inaccuracies. Table S4 provides medians and interquartile ranges of parameter estimates (those shown in Figure 2). The medians of estimates are relatively close to true values of *Ne* and *s* whatever these true values are. However, the variation in the estimations may be large for a few combinations of parameters as shown by the interquartile ranges. Indeed, for small n_1+n_2 and $Ne=10^4$, the method leads to strongly varying estimates (large interquartile ranges). Besides, when *s* is small, the estimates of *Ne* are strongly varying and can even be excessively large (see the cloud of blue points around $Ne=10^{10}$). Finally, it has to be noted that the measurements proposed in Appendix B led to better results than the raw empirical measurements (results not shown).

A large part of the estimation biases and variances is due to the use of measurements of the within- and across-generation clonality *f*, *F'* and *G* in the estimation method (even if the measurements that we used were better than the raw measurements; see Appendix B). Indeed, Fig. S5 (below) shows (for $n_1=n_2=0$ and $n_1=n_2=10$) the estimates which are obtained with the same method as above except that

the true values of *f*, *F'* and *G* are used instead of the measurements $\tilde{f}$, $\tilde{F}'$ and $\tilde{G}$. The accuracy lost due to the use of $\tilde{f}$, $\tilde{F}'$ and $\tilde{G}$ is large when *Ne* is large because the samples of size 268 are much smaller than the effective population size. In such a case, to improve the estimation method, more accurate measurements of ***f*, *F'* and *G*** should be used. Moreover, even with the true values of ***f*, *F'* and *G***, the estimation of small Ne and s is poor for large numbers of cycles. In such a case, to improve the estimation method additional information could be integrated in the contrast function (i.e. *F'* and *G* should be used together with other statistics, as for example variations in allelic frequencies). It must be noted that the bias correction which is proposed in the estimation procedure aims to take into account the possible bias in the estimates.

Fig. S5. Estimates (colored points) of effective size Ne and sex rate s obtained for simulations carried out under various Ne and s values (black points) and various numbers of cycles between the two samples (n_1=n_2=0 and n_1=n_2=10). The estimates are not the raw ones but those obtained by replacing in the estimation method the measurements of *f*, *F'* and *G* by their true values.

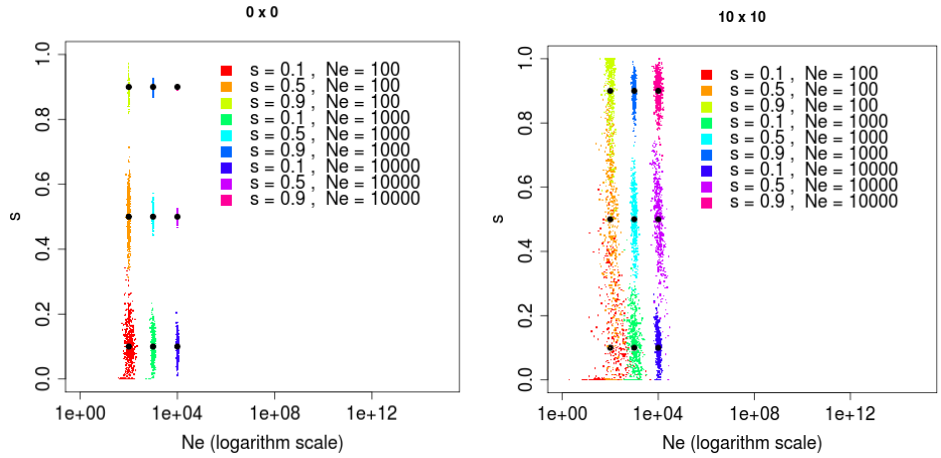


Table S4. Medians and interquartile ranges of raw estimates of effective size *Ne* and sex rate *s* obtained for simulations carried out under various *Ne* and *s* values and various numbers of cycles between the two samples (n_1=n_2=0, n_1=n_2=1, n_1=n_2=5, n_1=n_2=10, n_1=n_2=15 and n_1=17 and n_2=7).

Median of estimates of Ne									
Cycle number	Ne=100			Ne=1,000			Ne=10,000		
	s=0.1	s=0.5	s=0.9	s=0.1	s=0.5	s=0.9	s=0.1	s=0.5	s=0.9
0x0	107	101	101	963	1042	1002	6174	10717	11926
1x1	105	105	101	1005	1031	1016	8415	11115	10363
5x5	114	110	105	1027	1042	1012	9310	10620	10274
10x10	121	117	108	1031	1057	1014	9681	10094	10105
15x15	135	118	110	1027	1053	1030	10039	10474	10117
17x7	116	118	107	1010	1041	1014	10008	10644	10134
Interquartile range of estimates of Ne									
Cycle number	Ne=100			Ne=1,000			Ne=10,000		
	s=0.1	s=0.5	s=0.9	s=0.1	s=0.5	s=0.9	s=0.1	s=0.5	s=0.9
0x0	69	18	9	1019	376	225	4×10^{10}	14417	10644
1x1	68	27	14	752	282	185	3×10^{5}	11493	6426
5x5	110	46	24	666	296	156	9172	5765	3372
10x10	193	65	34	636	305	155	9227	3708	2350
15x15	267	78	38	755	309	184	7421	3578	2374
17x7	202	74	27	801	371	162	7734	4860	2630
Median of estimates of s									
Cycle number	Ne=0.1			Ne=1,000			Ne=10,000		
	s=0.1	s=0.5	s=0.9	s=0.1	s=0.5	s=0.9	s=0.1	s=0.5	s=0.9
0x0	0.091	0.505	0.907	0.087	0.492	0.903	0.086	0.502	0.974
1x1	0.113	0.503	0.909	0.091	0.506	0.901	0.089	0.502	0.917
5x5	0.092	0.513	0.921	0.100	0.490	0.902	0.084	0.502	0.907
10x10	0.091	0.535	0.936	0.098	0.500	0.910	0.099	0.494	0.903
15x15	0.077	0.524	0.953	0.114	0.507	0.906	0.094	0.502	0.904
17x7	0.082	0.522	0.933	0.107	0.507	0.906	0.090	0.493	0.904
Interquartile range of estimates of s									
Cycle number	Ne=100			Ne=1,000			Ne=10,000		
	s=0.1	s=0.5	s=0.9	s=0.1	s=0.5	s=0.9	s=0.1	s=0.5	s=0.9
0x0	0.135	0.111	0.053	0.189	0.159	0.063	0.252	0.448	0.199
1x1	0.167	0.171	0.075	0.143	0.137	0.056	0.234	0.308	0.166
5x5	0.229	0.258	0.109	0.148	0.134	0.057	0.180	0.182	0.076
10x10	0.236	0.297	0.140	0.164	0.151	0.062	0.162	0.150	0.064
15x15	0.224	0.355	0.150	0.175	0.175	0.071	0.147	0.138	0.061
17x7	0.242	0.291	0.118	0.183	0.144	0.062	0.160	0.163	0.076

Table S5. Effect of inaccuracy on n_1 and n_2 values on estimates of effective size *Ne* and sex rate. $n_1$and n_2 where generated through random sampling in two Gaussian distributions centered on 17 and 7, respectively (25% quantiles corresponding to intervals 15-19 and 5-9 respectively).

	True values n_1=17; n_2=7		*Simulated inaccuracies* n_1=17+σ; n_2=7+σ	
	Mean	c.i.	*Mean*	c.i.
Ne (raw)	1950	1450,3510	1950	1390,3540
Ne (bias corr.)	1740	670,2800	1730	610,2840
s (raw)	0.75	0.38, 0.95	0.75	0.37,0.95
s (bias corr.)	0.76	0.46,1.00	0.76	0.47,1.0

SUPPORTING INFORMATION

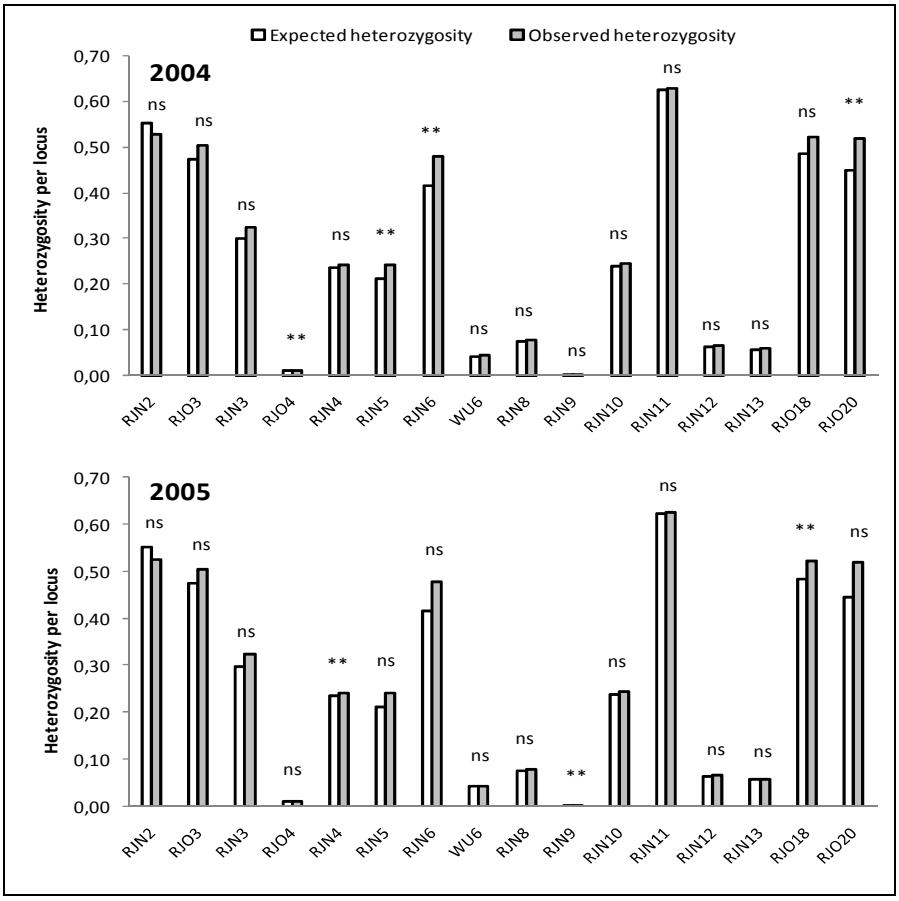


Fig. S1. Expected (He) and observed (Ho) heterozygosity for 17 polymorphic microsatellite loci for PST population of 2004 (above) and 2005 (below).

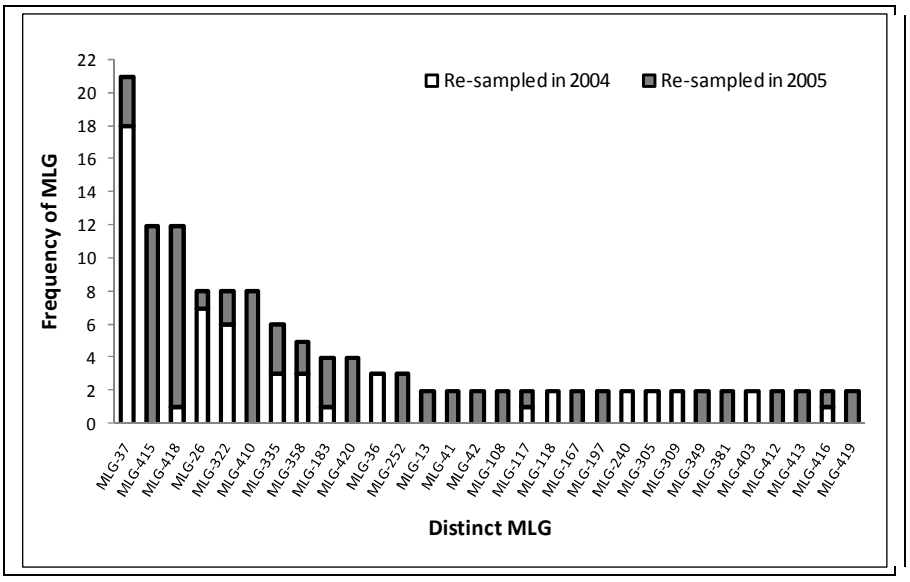


Fig S2. Frequency of multi-locus genotypes (MLG) re-sampled across two Chinese PST populations of 2004 and 2005.

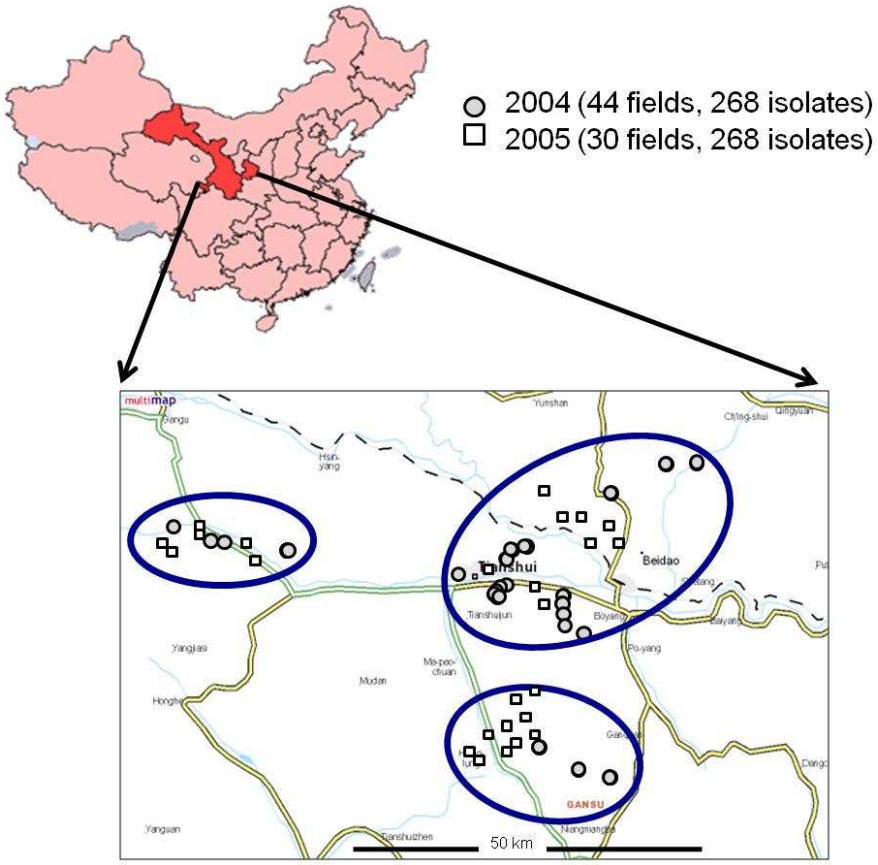


Fig. S3. The sampling plane for two years (2004 and 2005), showing three sub-populations in Tianshui county, Gansu province-China.

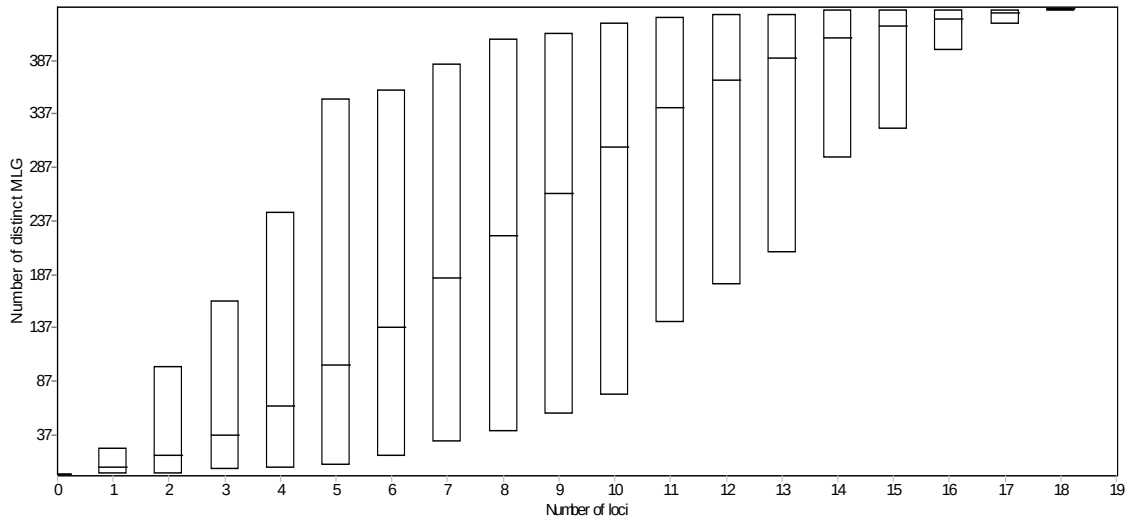


Fig. S4. Number of PST genotypes detected as a function of the number of loci re-sampled within the 17 polymorphic microsatellite markers using GENECLONE software. The box represents the average, minimum and maximum numbers of MLSTs detected when re-sampling on loci.

Table S1. Number of alleles and expected (He) estimated for 17 polymorphic SSR loci in the Chinese PST population with 536 isolates.

	2004		2005		
	Number of alleles	**Expected heterozygosity**	**Number of alleles**	**Expected heterozygosity**	**p-value[1]**
RJN2	9	0,5515	10	0,6003	0,01
RJO3	2	0,4741	2	0,4127	**0,00**
RJN3	5	0,2988	5	0,2371	0,16
RJO4	2	0,0111	3	0,0074	0,25
RJN4	2	0,2353	3	0,3642	**0,00**
RJN5	3	0,2136	3	0,1386	**0,00**
RJN6	3	0,4157	4	0,3718	0,23
WU6	2	0,0438	2	0,2046	**0,00**
RJN8	5	0,0757	3	0,1004	**0,00**
RJN9	2	0,0037	3	0,0111	0,63
RJN10	2	0,238	2	0,1221	**0,00**
RJN11	6	0,6239	6	0,603	0,01
RJN12	2	0,0649	2	0,0855	0,43
RJN13	2	0,0579	2	0,1901	**0,00**
RJO18	3	0,4855	2	0,4669	0,20
RJO20	2	0,4473	3	0,3833	**0,00**
RJO24	8	0,6878	8	0,6683	**0,00**

Table S2. Estimates of Fst (upper diagonal) and its significance (lower diagonal) based on 15 SSR loci for two Chinese PST populations during 2004 and 2005.

	G-1 (2004)	**G-2 (2004)**	**G-3 (2004)**	**G-1 (2005)**	**G-2 (2005)**	**G-3 (2005)**
G-1 (2004)	-	0.010	0.009	0.011	0.003	0.035
G-2 (2004)	0.993	-	0.002	0.022	0.009	0.041
G-3 (2004)	0.998	0.743	-	0.013	0.007	0.038
G-1 (2005)	1.000	1.000	1.000	-	0.006	0.010
G-2 (2005)	0.906	0.993	1.000	0.989	-	0.028
G-3 (2005)	1.000	1.000	1.000	0.990	1.000	-

Table S3. Permutation percentage (%) of real values and the presence of a linkage disequilibrium (bold type) between 17 polymorphic SSR loci in the Chinese PST population with 536 isolates.

	RJN2	**RJO3**	**RJN3**	**RJO4**	**RJN4**	**RJN5**	**RJN6**	**WU6**	**RJN8**	**RJN9**	**RJN10**	**RJN11**	**RJN12**	**RJN13**	**RJO18**	**RJO20**	**RJO24**
RJN2	-	0.0459	0.0346	0.0168	0.0404	0.0304	0.0305	0.1170	0.0311	0.0121	0.0612	0.0486	0.0555	0.1208	0.0402	0.1992	0.0432
RJO3	0.1258	-	0.0561	0.0374	0.1237	0.0350	0.0442	0.2290	0.0547	0.0332	0.0186	0.0576	0.0372	0.2412	0.0896	0.0826	0.0426
RJN3	0.1227	0.0201	-	0.0168	0.0345	0.0451	0.0417	0.0804	0.0247	0.0178	0.0588	0.0341	0.0291	0.1023	0.0378	0.0434	0.0486
RJO4	0.9999	0.1755	1.0000	-	0.0324	0.0197	0.0266	0.0259	0.0107	0.0044	0.0310	0.0295	0.0337	0.0257	0.0277	0.0139	0.0198
RJN4	0.2283	0.0010	0.8869	0.8888	-	0.0605	0.0598	0.2189	0.0538	0.0142	0.1259	0.0471	0.0769	0.2376	0.0207	0.0615	0.0677
RJN5	0.2297	0.3027	0.3133	0.8600	0.1271	-	0.0557	0.0929	0.0416	0.0181	0.0939	0.0568	0.08955	0.0783	0.5382	0.0849	0.0536
RJN6	0.9579	0.4585	0.4557	0.9386	0.1948	0.0608	-	0.1158	0.0360	0.0129	0.0398	0.0389	0.0238	0.1325	0.0635	0.0464	0.0486
WU6	0.0001	0.0001	0.0039	0.8619	0.0001	0.0049	0.0003	-	0.1206	0.0274	0.1536	0.0778	0.0968	0.5991	0.0298	0.1486	0.1169
RJN8	0.0574	0.0191	0.9988	1.0000	0.0213	0.5435	0.3560	0.0001	-	0.0113	0.0452	0.0392	0.0255	0.1424	0.0249	0.0504	0.0305
RJN9	1.0000	0.6258	0.9990	1.0000	0.9744	0.9899	0.9867	0.7417	1.0000	-	0.0545	0.0244	0.0200	0.0272	0.0332	0.0365	0.0196
RJN10	0.0001	0.6698	0.0653	0.7671	0.0005	0.0024	0.4516	0.0007	0.3397	0.1516	-	0.0767	0.0907	0.1391	0.0944	0.0633	0.0812
RJN11	0.0001	0.0207	0.7526	0.6237	0.1065	0.0002	0.3095	0.0001	0.3199	0.9159	0.0001	-	0.0265	0.1087	0.0475	0.0414	0.0499
RJN12	0.0001	0.4020	0.6799	0.0029	0.6794	0.0001	0.7817	0.0348	0.6384	0.8485	0.0409	0.6522	-	0.1160	0.0466	0.0118	0.0721
RJN13	0.0001	0.0001	0.0001	0.8645	0.0001	0.0307	0.0001	0.0001	0.0001	0.7449	0.0021	0.0001	0.0113	-	0.0344	0.1635	0.0998
RJO18	0.0148	0.0286	0.6754	0.8227	0.7828	0.0001	0.1360	0.7050	0.8400	0.7264	0.0189	0.2967	0.4001	0.6287	-	0.0611	0.0561
RJO20	0.0001	0.0276	0.2979	0.9720	0.1935	0.0023	0.4623	0.0001	0.1082	0.3074	0.1403	0.0521	0.9283	0.0001	0.2126	-	0.0500
RJO24	0.0001	0.1853	0.0001	0.9985	0.0001	0.0013	0.0523	0.0001	0.4805	0.9668	0.0001	0.0001	0.0001	0.0001	0.0075	0.0162	-

** P < 0.01%; a value of < 2 over 2000 permutations (calculated with GENETIX software) was considered to be significantly associate.

GENERAL CONCLUSION

GENERAL CONCLUSION

The present work is the first comprehensive study to describe the worldwide population structure and invasion history of *Puccinia striiformis* f.sp. *tritici* (PST) and to explore the role of sexual reproduction in PST temporal maintenance. In the first chapter, the results on worldwide populations revealed strong population subdivision and the existence of clonal population structure worldwide except in the Himalayan region (Pakistan and Nepal) and in its proximity in China. The second chapter's results suggested that this variability in recombination and diversity was in accordance with the worldwide variability in sexual reproduction capacity; the recombinant population had high telial production. The detailed study of populations from Pakistan (chapter 3) and China (chapter 4) further endorsed the role of sexual reproduction in the temporal maintenance of these recombinant PST populations. It is also worth mentioning that our vision of worldwide PST diversity is also changed, as we considered a gradient of diversity from East to West (Ali et al. 2010), but this study clearly shows the existence of recombination only in Himalayan and Near Himalayan region, with clonal/less diverse structure in the rest of the world populations. As these results are thoroughly discussed in each chapter, the conclusion addresses the worldwide population subdivision, identification of centre of diversity and the role of sexual reproduction in the temporal maintenance of recombinant populations in the context of invasions and pathogen evolutionary potential, while pointing out future perspectives.

Impact of invasion on the worldwide population subdivision

Wheat yellow rust has been characterised for its long distance migration capacity evidenced through its regional (Zeng and Luo 2006), continental (Brown and Hovmøller 2002; Hovmøller et al. 2002; Singh et al. 2004) and intercontinental dispersal (Singh et al. 2004). Our results clearly demonstrated the existence of worldwide population subdivision, which permitted us to infer the sources of recent invasions and their migration routes. Some of these recent migrations, affecting different continents in

few years, testify the role of human activities on long distance dispersal of PST, previously reported for other pathogens as well (Dilmaghani et al. 2012; Gladieux et al. 2008; Stukenbrock et al. 2006; Zaffarano et al. 2006). However, the existence of genetic groups specific to different geographical areas demonstrated that dispersal has not erased the major genetic groups, especially in centre of diversity. The replacement of local populations by these invasions would depend on the balance between local adaptation of native populations vs. the selective advantage of the migrants. In clonal populations invasive strains/pathotypes may have a high probability of replacing local populations if they are more fit than the native population. This could be the case for Mediterranean populations, which were dominated by the *Vr9* carrying pathotypes from the Middle East (Bahri et al. 2009). Similarly, the PST population in the USA after 2000 is apparently dominated by a PstS1 type pathotype or their derivatives after their invasion in 2000 (Chen 2005; Hovmøller et al. 2008) from a source population in the Red Sea-Middle East. However, molecular genotyping of the recent population in the USA would be necessary to confirm the relative proportion of the pre-2000 and post-2000 populations. In case of recombinant populations, the introgression of advantageous traits into the native population (Desprez-Loustau et al. 2007) would not result in the replacement of the local population. This could be the case of Himalayan populations. Pathotypes carrying virulence against *Yr*9, with a well described migration route starting in 1986 in Ethiopia, then spreading through the Middle East (Singh et al. 2004), did not replace the Himalayan populations. An alternative is that the virulence could have been already present in the native Pakistani population, though the available studies suggest the first detection in 1990s (Bahri et al. 2011; Rizwan et al. 2010; Singh et al. 2004).

Centre of diversity vs. sources of invasion

Identification of the diversity reservoirs in the centre of diversity as the potential source of invasion while the clonal populations outside this zone as the potential targets could be helpful to predict future invasion trajectories and to adopt better preventive measures (Campbell 2001; Desprez-Loustau et al.

2007; Perrings et al. 2002). However, in the context of worldwide invasions, the regional adaptive potential must also be known as the centre(s) of diversity would not always serve as the unique source of new invasions. The success of invasions would depend on the pre-existence of new virulences/advantageous alleles in the centre of diversity, the probability of their migration to other regions (possibly through human interventions) and the selection pressure on mutants outside of the centre of diversity. Although the diverse populations of the pathogen could serve as a reservoir for new virulences or other advantageous alleles, the selection pressure during the epidemic cycle on the pathogen population outside the centre of diversity, with virtually infinite population size at the climax of epidemics, could also allow rare mutants in clonal populations with low/no genetic diversity to appear and be selected. This is shown by the adaptation of clonal populations to resistance genes (Andrivon and de Vallavieille-Pope 1995; de Vallavieille-Pope et al. 2012; Goyeau et al. 2006; Wellings 2007) or to local climatic conditions (Enjalbert et al. 2005; Mboup et al. 2012). The probability that migration brings new invaders to a target host population would be linked with the migration capacity of the pathogen. In the case of PST, as invoked above, human activities are very important in PST invasions. Thus despite the existence of a centre of diversity in Himalayan and near Himalayan region, most recent invasions came from outside of this region, most probably linked with some human intervention. Thus along with the consideration of diversity zone, the deployment of resistance genes and the patterns of human activities (mobility, exchange of material etc.) should also be considered while predicting future invasion trajectories.

Worldwide variability in diversity, recombination and sexual reproduction

The analyses on diversity and recombination revealed the predominance of clonal population structure in worldwide PST populations, except in its centre of diversity in the Himalayan region and nearby in China. The observed difference in diversity and recombination signature was correlated with a reduced telial production capacity outside the Himalayan and near Himalayan region. This could be

the result of a partial loss of sex ability in clonal populations during its worldwide spread. The reduction/loss of sexual reproduction reported in other pathogens may be linked to the lack of one of the two mating types due to genetic drift or difference in relative fitness (Andrivon 1996; Saleh et al. 2012b), lack of alternate host (Barrett et al. 2008; Groth and Roelfs 1982) or the degradation of sexual reproduction capacity when the pathogen passes through several clonal cycles (Saleh et al. 2012a). Reduction in PST sexual reproduction ability during its worldwide spread could result from its migration into areas where the alternate host is lacking or the sexual cycle is not necessary for a year-around epidemic cycle. Future studies should test for the possibility of a sexual cycle of the recombinant PST population on different *Berberis* spp. present in the clonal region and the role of *Berberis* spp. on population structure of PST population in the centre of diversity.

Centre of origin and ancestral relationship among populations

The results on the worldwide population subdivision, diversity, recombination signature and sexual reproduction capacity suggest the Himalayan region to be the most likely centre of origin of PST. This was further endorsed by the approximate Bayesian computation analyses, which identified the western part of Himalayan chain (north of Pakistan) to be the most ancestral population, while Nepal at the eastern part of Himalayan chain to be a result of admixture between Pakistan and China. The existence of the centre of origin in the Himalayan region suggests host shift instead of co-evolution with wheat in the Fertile Crescent, supporting the increasing evidence of host-shift in invasive fungi (Giraud et al. 2010; Zaffarano et al. 2006). Further studies would be necessary to elaborate the population structure in the whole Himalayan region, including western Afghanistan, southern China and northern India. Furthermore, although most of the worldwide populations were represented in the present work, populations from Mongolia, the northern part of Central Asia and Russia must also be studied to confirm the origin of NW European populations.

The role of* Berberis *in the temporal maintenance of PST

We have shown the first evidence of the role of sexual reproduction on *Berberis* spp. in the temporal maintenance of natural PST populations in the Himalayan region of Pakistan and suggest a similar role in the recombinant population of Gansu, China. The temporal maintenance through a sexual cycle on the alternate host resulted in the high genotypic and pathotypic diversity observed in the region. As Pakistani populations were shown to be the most ancestral in the present collection of isolates, detailed studies on the yellow rust on local grasses, the host barriers of the local yellow rust population, the diversity in resistance genes in the local wheat landraces and the comparison of wheat vs. *Berberis* infecting PST population would be needed.

Empirical estimates of parameters on temporal maintenance of PST

The estimation of demographic and reproductive parameters through the information on the resampling of clones within and across generations enabled us to address the role of sexual reproduction in the temporal maintenance of the Chinese recombinant population. Our method could be applied to other PST populations, for which the same population is resampled across generations with no spatial structuring and at least some clonal resampling over generations. It can also be used on other organisms with mixed reproductive modes to infer effective population size and the relative contribution of sexual and asexual reproduction, as many invasive pathogens have a mixed mode of reproduction (Fisher 2007; Gladieux et al. 2011; Milgroom 1996). A comparison of the estimates from our methods with other estimates in certain macro-organism, where clonal survival could be traced, would help us to assess the accuracy of the method. Information on demographic and reproductive parameters are important for mathematical epidemiology and modeling and is critical for understanding the causes for pathogen emergence and spread (Fisher 2007; McDonald and Linde 2002; Regula et al. 2009).

Concluding remarks

The information generated in the present study would improve our understanding of PST population biology and invasion history. The existence of worldwide population subdivision, identification of diversity reservoirs and target populations and the sources and migration patterns of recent invasions would help predict future invasion risks. The role of human activities in recent invasions must also be considered for preventing future invasions through better quarantine measure and increasing awareness among people. Identification of a centre of genetic diversity with frequent sexual reproduction would encourage orienting future efforts on determining other centre(s) of diversity. The evidence of the role of *Berberis* in temporal maintenance would reflect on the adaptive potential of the pathogen in its centre of diversity. The information on the temporal maintenance in the centre of diversity would be useful for better management of wheat yellow rust and resistance gene deployment in the region, while considering the adaptive potential of the pathogen inside the centre of diversity linked with the diversity and the role of sexual reproduction.

Literature cited:

Ali, S., Leconte, M., Walker, A.-S., Enjalbert, J., De Vallavieille-Pope, C., 2010. Reduction in the sex ability of worldwide clonal populations of *Puccinia striiformis* f.sp. *tritici*. Fungal Genetics and Biology 47, 828-838.

Andrivon, D., 1996. The origin of *Phytophthora infestans* populations present in Europe in the 1840s: a critical review of historical and scientific evidence. Plant Pathology 45, 1027-1035.

Andrivon, D., de Vallavieille-Pope, C., 1995. Race diversity and complexity in selected populations of fungal biotrophic pathogens of cereals. Phytopathology 85 (8), 897-905.

Bahri, B., Leconte, M., Ouffroukh, A., de Vallavieille-Pope, C., Enjalbert, J., 2009. Geographic limits of a clonal population of wheat yellow rust in the Mediterranean region. Molecular Ecology 18, 4165-4179.

Bahri, B., Shah, S.J.A., Hussain, S., Leconte, M., Enjalbert, J., de Vallavieille-Pope, C., 2011. Genetic diversity of wheat yellow rust population in Pakistan and its relationship with host resistance. Plant Pathology 60, 649-660.

Barrett, L.G., Thrall, P.H., Burdon, J.J., Nicotra, A.B., Linde, C.C., 2008. Population structure and diversity in sexual and asexual populations of the pathogenic fungus *Melampsora lini*. Molecular Ecology 17, 3401-3415.

Brown, J.K.M., Hovmøller, M.S., 2002. Aerial dispersal of fungi on the global and continental scales and its consequences for plant disease. Science 297, 537-541.

Campbell, F., 2001. The science of risk assessment for phytosanitary regulation and the impact of changing trade regulations. Bioscience 51, 148-153.

Chen, X.M., 2005. Epidemiology and control of stripe rust [*Puccinia striiformis f. sp. tritici*] on wheat. Canadian Journal of Plant Pathology 27, 314–337.

de Vallavieille-Pope, C., Ali, S., Leconte, M., Enjalbert, J., Delos, M., Rouzet, J., 2012. Virulence dynamics and regional structuring of *Puccinia striiformis* f. sp. *tritici* in France between 1984 and 2009. Plant Disease 96, 131-140.

Desprez-Loustau, M.-L., Robin, C., Buée, M., Courtecuisse, R., Garbaye, J., Suffert, F., Sache, I., Rizzo, D.M., 2007. The fungal dimension of biological invasions. Trends in Ecology and Evolution 22, 472–480.

Dilmaghani, A., Gladieux, P., Gout, L., Giraud, T., Brunner, P.C., Grandaubert, J., Stachowiak, A., Balesdent, M.H., Rouxel, T., 2012. Migration patterns and changes in population biology associated with the worldwide spread of the oilseed rape pathogen *Leptosphaeria maculans*. Molecular Ecology doi: 10.1111/j.1365-294X.2012.05535.x.

Enjalbert, J., Duan, X., Leconte, M., Hovmøller, M.S., de Vallavieille-Pope, C., 2005. Genetic evidence of local adaptation of wheat yellow rust (*Puccinia striiformis* f. sp. *tritici*) within France. Molecular Ecology 14 (7), 2065-2073.

Fisher, M.C., 2007. The evolutionary implications of an asexual lifestyle manifested by *Penicillium marneffei*. In: Sex in fungi: molecular determination and evolutionary implications., Heitman, J.K., J.W.; Taylor, J.W.; Casselton, L.A. (Ed.), ASM Press, Washington, p. 542.

Giraud, T., Gladieux, P., Gavrilets, S., 2010. Linking the emergence of fungal plant diseases with ecological speciation. Trends in Ecology & Evolution 25, 387-395.

Gladieux, P., Byrnese, E., Fisher, M.C., Aguleta, G., Heitman, J., T., G. (Eds.), 2011. Epidemiology and evolution of fungal pathogens, in plants and animals. Elsevier.

Gladieux, P., Zhang, X.-G., Afoufa-Bastien, D., Sanhueza, R.-M.V., Sbaghi, M., Le Cam, B., 2008. On the origin and spread of the scab disease of apple: out of Central Asia. PLoS ONE 3, e1455.

Goyeau, H., Park, R., Schaeffer, B., Lannou, C., 2006. Distribution of Pathotypes with Regard to Host Cultivars in French Wheat Leaf Rust Populations. Plant Pathology 96, 264-273.

Groth, J.V., Roelfs, A.P., 1982. Effect of sexual and asexual reproduction on race abundance in cereal rust fungus populations. Phytopathology 72, 1503-1507.

Hovmøller, M.S., Justesen, A.F., Brown, J.K.M., 2002. Clonality and long-distance migration of *Puccinia striiformis* f.sp. *tritici* in north-west Europe. Plant Pathology 51 (1), 24-32.

Hovmøller, M.S., Yahyaoui, A.H., Milus, E.A., Justesen, A.F., 2008. Rapid global spread of two aggressive strains of a wheat rust fungus. Molecular Ecology 17 (17), 3818-3826.

Mboup, M., Bahri, B., Leconte, M., De Vallavieille-Pope, C., Kaltz, O., Enjalbert, J., 2012. Genetic structure and local adaptation of European wheat yellow rust populations: the role of temperature-specific adaptation. Evolutionary Applications DOI: 10.1111/j.1752-4571.2011.00228.x.

McDonald, B.A., Linde, C., 2002. Pathogen population genetics, evolutionary potential, and durable resistance. Annual Review of Phytopathology 40 (1), 349-379.

Milgroom, M.G., 1996. Recombination and the multilocus structure of fungal populations. Annual Review of Phytopathology 34 (1), 457-477.

Perrings, C., Williamson, M., Barbier, E., Delfino, D., Dalmazzone, S., J., S., Simmons, P., Watkinson, A., 2002. Biological invasion risks and the public good: an economic perspective. Conservation Ecology 6, Article 1 (online).

Regula, T., Medlock, J., Galvani, A., 2009. The discounted reproductive number for epidemiology. Methematical Biosciences and Engineering 6, 379-395.

Rizwan, S., Iftikhar, A., Kazi, A.M., Sahi, G.M., Mirza, J.I., Rattu, A.R., Ashraf, M., 2010. Virulence variation of *Puccinia striiformis* Westend. f. sp. *tritici* in Pakistan. Archives of Phytopathology and Plant Protection 43, 875-882.

Saleh, D., Milazzo, J., Adreit, H., Tharreau, D., Fournier, E., 2012a. Asexual reproduction induces rapid and permanent loss of sexual reproduction ability of the fungal rice pathogen, *Magnaporthe oryzae*: results from in vitro experimental evolution assays. BMC Evolutionary Biology 12, 42.

Saleh, D., Xu, P., Shen, Y., Li, C., Adreit, H., Milazzo, J., Ravigné, V., Bazin, E., Nottéghem, J.-L., Fournier, E., Tharreau, D., 2012b. Sex at the origin: an Asian population of the rice blast fungus Magnaporthe oryzae reproduces sexually. Molecular Ecology DOI: 10.1111/j.1365-294X.2012.05469.x.

Singh, R.P., William, H.M., Huerta-Espino, J., Rosewarne, G., 2004. Wheat rust in Asia: meeting the challenges with old and new technologies. In: "New directions for a diverse planet": 4th International Crop Science Congress, Brisbane, Australia.

Stukenbrock, E.H., Banke, S., McDonald, B.A., 2006. Global migration patterns in the fungal wheat pathogen *Phaeosphaeria nodorum*. Molecular Ecology 15, 2895-2904.

Wellings, C.R., 2007. *Puccinia striiformis* in Australia: A review of the incursion, evolution and adaptation of stripe rust in the period 1979-2006. Australian Journal of Agricultural Research 58, 567-575.

Zaffarano, P.L., McDonald, B.A., Zala, M., Linde, C.C., 2006. Global hierarchical gene diversity analysis suggests the Fertile Crescent is not the center of origin of the barley scald pathogen *Rhynchosporium secalis*. Phytopathology 96, 941-950.

Zeng, S.-M., Luo, Y., 2006. Long-distance spread and interregional epidemics of wheat stripe rust in China. Plant Disease 90, 980-988.

ANNEXES

ANNEXES

ADDITIONAL TASKS

During my PhD, apart from my PhD work I collaborated to the following projects.

1. I actively participated in data presentation and manuscript writing of the article describing PST virulence dynamics over last 25 years in France. The article was published in Plant disease.

   VIRULENCE DYNAMICS AND REGIONAL STRUCTURING OF *PUCCINIA STRIIFORMIS* F. SP. *TRITICI* IN FRANCE BETWEEN 1984 AND 2009.

   de Vallavieille-Pope C, **Ali S.**, Leconte M, Enjalbert, J., Delos, M., Rouzet, J. 2012. **Plant Disease.** 96: 131-140.

2. I participated to the PhD of Rola El-Amil in defining her PhD work and providing her the basic training in molecular genotyping and population genetic analyses. An abstract has been submitted to BGRI meeting, from that work :

   STATUS OF STRIPE (YELLOW) RUST IN SYRIA IN 2010-2011

   R. El Amil, K. Nazari, C. de Vallavieille-Pope, **S. Ali**, A. Yahyaoui and M. Hovmøller

3. I Participated to the post-doc work of Stephanie Walter on the origin, distribution and evolution of the aggressive strains of PST, adapted to high temperature. I did the SSR genotyping as well as contributed to the data interpretation and writing. An abstract has been submitted to 13 ICRPMC, from that work:

   RECENT INVASION OF WORLD-WIDE WHEAT GROWING AREAS BY TWO AGGRESSIVE STRAINS OF *PUCCINIA STRIIFORMIS*

   Stephanie WALTER[1], **Sajid ALI**, Annemarie Fejer JUSTESEN, Jérôme ENJALBERT, Claude de VALLAVIEILLE-POPE, Mogens Stovring HOVMØLLER

4. I collaborated to the work on newly spreading multivirulent pathotypes in collaboration with Dr. Mogens S. Hovmøller group. I carried out the SSR genotyping of the isolates from these pathotypes. An abstract from the work on the recent changes in European PST population, has been submitted to 13 ICRPMC:

   THE INFLUENCE OF MUTATION, RECOMBINATION AND EXOTIC INCURSIONS ON THE RECENT DYNAMICS OF *PUCCINIA STRIIFORMIS* IN EUROPE

   Mogens Stovring HOVMØLLER, Stephanie WALTER, **Sajid ALI**, Annemarie Fejer JUSTESEN, Jens Grønbech HANSEN, Poul LASSEN, Rosemary BAYLES, Kerstin FLATH, Claude de VALLAVIEILLE-POPE

5. In collaboration with the researchers from CIRAD, Montpellier working on *Megnaporthe oryzae*, we developed a first version of research project. In this project, an effort will be made to develop a joint collaborative effort to understand the population structure of *Megnaporthe oryzae* in the Himalayan region. This will be done after the completion of my PhD.

CONFERENCES ATTENDED

I attended following conferences, workshops, colloquia and meetings during my PhD and benefited from the work presented by experts from diverse fields including population genetics, fungal biology, pathology, epidemiology and plant breeding and genetics.

1. 05-08 June 2012- **8ème colloque de la Société Française de Phytopathologie**. Paris, France.
2. 07-11 May, 2012 – **Meeting on ongoing yellow rust research work at global rust reference centre (GRRC)**, Denmark.
3. 19-23 September 2011- **XVI Congress of European Mycologists**, Halkidiki, Greece.
4. 25-29 June, 2011 - **Jacques Monod Conference**, Roscoff, France.
5. 18th may 2011 **– Invited talk** at Institute of Biotechnology and Genetic Engineering, KP Agricultural University Peshawar, Pakistan.
6. 18th may 2011 **– Invited talk** at Department of Genetics, Hazara University Mansehra, Pakistan.
7. 15-17 February, 2011 – **Conference of BIOEXPLOIT final meeting**, Wageningen, The Netherlands.
8. 25-26 November, 2010 - **Annual meeting of EMERFUNDIS** at Lauret-France.
9. 2-4 June, 2010 **–** Colloquium of **"2eme Journee doctorant SPE-2010",** Sophia-Antipolis, France
10. 4th may 2010 **– Invited talk** at Midipile at University Paris Sud-11.
11. 9-10 February, 2010 **-** Technology Workshop on **"Marker Assisted Selection: from discovery to application"**. Wageningen, The Netherlands.
12. 25-29 January, 2010 **- 8èmes Rencontres de Phytopathologie – Mycologie, JJC 2010,** Aussois, France.
13. 8-9 December, 2009 **- 9th International Conference on Plant Diseases (CIMA-2009),** Tours, France.
14. 27 November, 2009 - **Journee REID-Champignon** at AgroParisTech, Paris-France.
15. 25-26 November, 2009 - **Annual meeting of EMERFUNDIS** at Paris-France.
16. 12 November, 2009 **- Colloquium on "Evolution de l'evolution"** at AgroParisTech, Paris-France.
17. 3-4 November, 2009 **–** Workshop on **Durable disease resistance management.** INRA-Paris-France.

ABSTRACTS PUBLISHED

8-9 December, 2009 - 9th International Conference on Plant Diseases (CIMA-2009), Tours, France.

TEMPERATURE ADAPTATION AND TELIOSORE PRODUCTION APTITUDE OF *PUCCINIA STRIIFORMIS* F.SP. *TRITICI* POPULATIONS IN RELATION TO THEIR GEOGRAPHICAL AND PHYLOGENETIC POSITION

Sajid ALI, Marc LECONTE, Jérôme ENJALBERT and Claude de VALLAVIEILLE-POPE

UMR 1290 BIOGER-CPP, INRA-AgroParisTech, Equipe Epidémiologie BP 01, 78 850 Thiverval-Grignon, France
sajid.ali@grignon.inra.fr

The recent expansion of yellow rust epidemics (*Puccinia striiformis* f.sp. *tritici* : PST) in new areas with a climate considered so far as too hot (South USA, West Australia, South Africa) has been related to the emergence and global spread of new high temperature adapted races from a narrow genetic basis (Hovmøller et al. 2008). Furthermore, Mboup (2008) and Bahri (2008) reported the presence of a gradient of diversity in PST populations from North-western Europe to Pakistan and China, along with a higher teliosore production for recombinant (Chinese) than clonal (French) isolates. We report an overall variability for temperature adaptation in terms of germination rate (at 20°C and 25°C) and relative infection efficiency (RIE) for the isolates in accordance with their phylogenetic position and geographical origin. A non-significant difference among geographical origins was observed for germination rate at 8°C. The isolates from Nepal, Middle-East and Pakistan, the regions having hot climate, were high temperature adapted with even higher germination rate and/or RIE than the two high temperature adapted reference isolates. This suggests that the high temperature adapted isolates could be a result of migration from the already adapted populations present in these three regions. In the case of teliosores, isolates from Pakistan, Nepal and China had high teliosore production, while the clonal populations of Northern France and Mediterranean region had very low teliosore production. This represents a correspondence between the gradient in teliosore production and the gradient of genetic diversity already described (Bahri 2008; Mboup 2008), suggesting a strong relation between genetic diversity and teliosore production aptitude. The phylogenetic position of most of the isolates was in accordance with their geographical origin with some off-type isolates. A general concordance found between phenotypic profile and genetic structure along with the climatic conditions of sampled locations reveals the presence of local adaptation in PST populations for high temperature adaptation and teliosore production. The probable expansion risk of aggressive isolates from the zones with higher diversity (Asia) to the areas with clonal populations should thus be considered in the context of climatic change and durable crop protection.

Keywords: *wheat yellow rust, geographical distribution, spore germination rate, microsatellites*

25-29 January, 2010 - 8èmes Rencontres de Phytopathologie – Mycologie, JJC 2010, Aussois, France.

GEOGRAPHICAL GRADIENT FOR TELIOSORE PRODUCTION OF *PUCCINIA STRIIFORMIS* F.SP. *TRITICI* ISOLATES IN RELATION TO THEIR PHYLOGENETIC POSITION

Sajid ALI, Marc LECONTE, Laurent GERARD, Jérôme ENJALBERT and Claude de VALLAVIEILLE-POPE
UMR 1290 BIOGER-CPP, INRA-AgroParisTech, BP 01, 78 850 Thiverval-Grignon, France
sajid.ali@grignon.inra.fr

Puccinia striiformis f.sp. *tritici* (PST) is a basidiomycete, considered to reproduce asexually through dicaryotic uredospores on wheat. At the end of cropping season, teliosores are differentiated to produce the teliospores, which in turns can germinate to form haploid basidiospores after meiosis. In Uredinales, basidiospores infect an alternate host, not described so far for PST. Accordingly, genetic studies had shown a clonal PST population structure in Europe, Australia and USA. However, our team has recently shown the genetic signature of recombination in Chinese populations along with the presence of genetically diverse populations in Middle-East and Pakistan, suggesting an overall gradient of diversity from Western Europe to Eastern Asia (Bahri *et al.*, 2009; Mboup *et al.*, 2009). A higher teliosore production for recombinant (Chinese) isolates than clonal (French) isolates was equally observed. Here we attempted to assess the difference for teliosore production in populations belonging to different geographical origins, and exhibiting different levels of genetic diversity. A set of 56 isolates from six geographical origins were genotyped with 16 SSR markers, which amplified 17 loci. Genetic analysis of SSR data assigned them to 5 genetic groups corresponding to their geographical origin with some off-type and hybrid isolates. An overall variability for teliosore production was observed in accordance with their phylogenetic position and geographical origin. Isolates from China, Nepal and Pakistan had high teliosore production, while the clonal populations of France and Mediterranean region had very low teliosore production. This geographic gradient in teliosore production corresponded to the gradient of genetic diversity mentioned above, suggesting a clear relationship of teliosore production aptitude with genetic diversity. This could help in orienting the efforts to search for its alternate host and centre of diversity. The overall relation between teliosore production, phylogenetic position and geographical origin, as detected by Fst and Qst analysis, suggests that the selection pressure on teliosore production could be stronger than those influencing neutral markers.

Keywords: *wheat yellow rust, genetic structure, microsatellites, sexual reproduction trait*

Bahri *et al.* (2009) *Molecular Ecology*, 18, 4165-4179.
Mboup *et al.* (2009) *Fungal Genetics and Biology*, 46, 299-307.

2-4 June, 2010 – Colloquium of "2eme Journee doctorant SPE-2010", Sophia-Antipolis, France

HIGH TEMPERATURE ADAPTATION IN WORLDWIDE POPULATIONS OF *PUCCINIA STRIIFORMIS* F. SP. *TRITICI* IN RELATION WITH THEIR PHYLOGENETIC POSITION

Sajid ALI1, Marc LECONTE1, Laurent GERARD1, Anne-Sophie WALKER1, Bochra BAHRI1, Mamadou MBOUP1, Jérôme ENJALBERT2 and Claude de VALLAVIEILLE-POPE1

1UMR 1290 BIOGER-CPP, INRA-AgroParisTech, BP 01, 78850 Thiverval-Grignon, France

2UMR 320 Génétique Végétale, INRA, Ferme du Moulon, 91190 Gif sur Yvette, France

The genetic diversity of wheat yellow rust (PST: *Puccinia striiformis* f.sp. *tritici*) is strongly structured both by a clonal life cycle and by host selective pressures. Along with the description of "boom and bust cycle" through arm-race between host and pathogen, we have suggested that other selective pressures, like climatic conditions, can be at the origin of the genetic structure of yellow rust populations at a regional scale (Enjalbert et al. 2005). The role of this climate adaptation is further questioned by the recent expansion of PST epidemics to the areas previously considered hot for PST (South USA, West Australia, South Africa) due to recent spread of high temperature adapted isolates (Hovmøller et al. 2008). Furthermore, the strict clonality of PST can be questioned as Mboup *et al.* (2009) and Bahri (2009) have reported the presence of recombinant populations or genotypes in Chinese and Pakistani populations. On the basis of a worldwide sampling of PST, we studied microsatellite polymorphism, high temperature adaptation and teliosore production to test whether two climatic and sex-related traits are linked to the phylogeography of populations. The existence of high temperature adapted isolates in Asiatic populations with higher diversity suggests that they are the possible source for new emergence. Similarly, we report a higher teliosore production, a sex-specific trait, for isolates from genetically diverse populations than clonal populations.

25-29 June, 2011 - Jacques Monod Conference, Roscoff, France.

WORLDWIDE GEOGRAPHICAL CLINE FOR GENOTYPIC DIVERSITY CORRELATES WITH SEX ABILITY IN *PUCCINIA STRIIFORMIS* F.SP. *TRITICI*

Sajid Ali, Jérôme Enjalbert, Marc Leconte, Angélique Gautier, Mogens S. Hovmøler, Claude de Vallavieille-Pope

Puccinia striiformis f.sp. *tritici* (PST) populations have been reported clonal in the USA, Australia and Europe with a recent report of a high genotypic diversity in Pakistan (Bahri et al. 2011a) and recombinant populations in China (Mboup et al. 2009), suggesting an apparent gradient of diversity. To confirm the existence of this phylogeographic diversity gradient we genotyped a set of 385 isolates representative of worldwide geographical origins, with 20 SSR markers. Phylogenetic analyses (STRUCTURE, DAPC and AFC) clustered these isolates into 6 distinct genetic groups according to their geographical origins. Asiatic populations (China, Nepal and Pakistan) were found recombinant, while Australian-European, African, and North American populations were clonal. The East-West genotypic diversity gradient was also confirmed, with PST populations from Middle-East and Central Asia of intermediate diversity but lacking the recombination signature. We assessed whether this gradient in diversity/recombination could be linked to difference in the ability for sexual recombination. As a proxy for this sex-ability, we measured the telial production, a sex-specific structures that are obligatory for sexual cycle. Telial production was measured in a set of 56 isolates representative of these worldwide geographical origins. We observed a geographic cline in telial production in accordance with the observed gradient of genotypic diversity/recombination, a result defending the occurrence of sexuality in PST populations. The higher mean Qst value (0.822) for telial production than the Fst value (0.317) might indicate a direct or indirect selection affecting telial production, rather than evolution under genetic drift alone. We suggest that these gradients might reflect on the evolution of sex in PST population during its worldwide migration, with a loss of sexual function in areas where sexual alternate host is lacking, or not necessary for the completion of epidemic cycle.

19-23 September 2011- XVI Congress of European Mycologists, Halkidiki, Greece.

ASIAN POPULATIONS OF THE WHEAT STRIPE RUST PATHOGEN AS A POTENTIAL SOURCE OF NEW EMERGENCES, DUE TO THEIR HIGH GENOTYPIC AND PHENOTYPIC DIVERSITY

Sajid Ali[1]*, Jérôme Enjalbert[2], Pierre Gladieux[3], Marc Leconte[1], Angélique Gautier[1], Mogens S. Hovmøler[4], Claude de Vallavieille-Pope[1]

[1] UMR1290 BIOGER-CPP, INRA-AgroParisTech, 78850 Thiverval-Grignon, France
[2] UMR 320 Génétique Végétale, INRA, Ferme du Moulon, 91190 Gif sur Yvette, France
[3] Laboratoire Ecologie Systématique Evolution, UMR8079, Univ. Paris-Sud, CNRS, AgroParisTech, Orsay, France
[4] University of Aarhus, Forsøgsvej 1, DK-4200 Slagelse, Denmark
* sajid.ali@grignon.inra.fr, bioscientist122@yahoo.com

Keywords: *Puccinia striiformis* f.sp. *tritici*, invasive fungi, microsatellites, telial production

Understanding the origin, migration routes, and possible changes in population biology of fungal pathogens is essential for durable disease management. Here, we used microsatellite genotyping to infer the invasion history of the wheat pathogen *Puccinia striiformis* f.sp. *tritici* (PST), based on a set of 385 worldwide isolates. The pathogen has been reported clonal with cryptic molecular variation in Europe, USA and Australia, but recent works highlighted the existence of a high diversity in Asian PST populations (Bahri et al. 2011a; Mboup et al. 2009). In the last decade, a high temperature adapted strain, first reported in USA, has spread globally, with no known origin of emergence. Our analyses, based on Bayesian and multivariate methods, clustered worldwide PST populations into six distinct groups, corresponding to their geographical origin. Asiatic populations (China, Nepal and Pakistan) were found recombinant and diverse, while Australian-European, African, and North American populations were clonal. Molecular data support a Mediterranean-Middle-Eastern origin of the recently spread aggressive strain. The distribution of genetic variability was consistent with an East-West genotypic diversity gradient, with PST populations from Middle-East and Central Asia of intermediate diversity but lacking the recombination signature. To assess the role of sex in this diversity cline, we used telial production as a proxy for sex-ability; telia being the sex-specific obligatory structures for sexual cycle. The geographic cline in telial production was found significantly correlated with observed gradient of genotypic diversity/recombination in a subset of 56 representative isolates, defending the scenario of a loss of sex ability in clonal PST populations. Temperature adaptation tests for the same 56 isolates demonstrated the existence of high variability in temperature adaptation in Middle-East and Asian populations. We conclude that Asian populations having recombination/high diversity, high sex ability and high variability in temperature adaptation could serve as potential source of new emergences.

25-29 January, 2012 - 9èmes Rencontres de Phytopathologie – Mycologie, JJC 2010, Aussois, France.

ASIAN POPULATIONS OF *PUCCINIA STRIIFORMIS* F.SP. *TRITICI* AS A CENTRE OF DIVERSITY WITH HIGH SEX ABILITY

Sajid Ali, Jérôme Enjalbert, Marc Leconte, Angélique Gautier, Mogens S. Hovmøller, Claude de Vallavieille-Pope

The worldwide population analysis is crucial for understanding the evolutionary potential and identifying centre of diversity for long distance migrating pathogens. *Puccinia striiformis* f.sp. *tritici* (PST) is one of these pathogens with worldwide distribution but unknown centre of diversity and origin. Early population analysis of PST has shown clonal population structure in the USA, Australia and Europe. Our recent work showed a high genotypic diversity in Pakistan (Bahri et al. 2011a) and recombinant populations in China (Mboup et al. 2009), suggesting the existence of diverse population at least in some Asian populations. To explore the worldwide population structure of PST a set of 385 isolates representative of worldwide geographical origins, were genotyped with 20 SSR markers. These isolates were clustered into 6 distinct clusters reflecting their geographical origins. Recombination signature was present in Asiatic populations (China, Nepal and Pakistan) while absent in Australian-European, African, and American populations, confirming the East-West diversity gradient. To infer on the possible origin of sexual recombination for this gradient of diversity, we measured the telial production, a sex-specific structures that are obligatory for sexual cycle, as a proxy for this sex-ability. Telial phenotyping was made for 56 isolates selected as representative of the above mentioned geographical regions. Indeed, the telial production was high in populations with recombination signature (Pakistan, Nepal and China) while low in clonal populations. The cline observed in telial production was in accordance with the observed gradient of genotypic diversity/recombination, supporting the hypothesis of occurrence of sexuality in PST populations.

05-08 June 2012- 8ème colloque de la Société Française de Phytopathologie. Paris, France.

WORLDWIDE POPULATION GENETIC ANALYSES SUGGESTS HIMALAYA AS A CENTRE OF DIVERSITY FOR *PUCCINIA STRIIFORMIS* F.SP. *TRITICI* ALONG WITH HIGH SEX ABILITY AND TEMPERATURE ADAPTATION

Sajid Ali[1]*, Jérôme Enjalbert[2], Pierre Gladieux[3,4], Marc Leconte[1], Angélique Gautier[1], Annemarie F. Justesen[5], Mogens S. Hovmøller[5], Claude de Vallavieille-Pope[1]

Identification of centre of genotypic and phenotypic diversity and zone of sex occurrence is crucial for understanding the evolutionary potential of crop pathogens and subsequent disease management. *Puccinia striiformis* f.sp. *tritici* (PST) is an economically important pathogen with worldwide distribution but unknown centre(s) of diversity and previously considered clonal. However, high diversity has been shown very recently in Pakistan (Bahri et al. 2011b) and recombination in China (Mboup et al. 2009). To explore the worldwide population structure of PST and identify the centre of diversity and potential zones of sex occurrence, a set of 409 isolates were genotyped with 20 SSR markers representative of worldwide geographical origins. We confirmed the existence of population subdivision at worldwide level with 6 genetic clusters according to their geographical origin, suggesting the independent maintenance of each population. A very low diversity and clonal population structure was observed in NW European, Middle-Eastern, Mediterranean, African, and American populations, while a high genotypic diversity and recombination signature was present in Asiatic populations near to Himalayan Chain (China, Nepal and Pakistan). These populations had also a high sex ability in terms of telial production and more phenotypic variance for temperature adaptation. Our results thus suggest Himalayan populations as centre of diversity for PST and could be the potential zone of sex occurrence and thus probable centre of origin. Additionally we identified the Middle-East as the origin of the high temperature adapted aggressive strain (Milus et al. 2009), and confirmed the NW European origin of South American population while the Mediterranean origin of South African isolates. This information could be useful for understanding the evolutionary potential of the pathogen and its subsequent management at worldwide and local level.

[1] INRA UR 1290 BIOGER-CPP, BP01, 78850 Thiverval-Grignon, France
[2] INRA UMR 320 Génétique Végétale, Ferme du Moulon, 91190 Gif sur Yvette, France
[3] UMR 8079 Ecologie Systematique Evolution, Univ. Paris-Sud 11, 91405 Orsay, France; CNRS, 91405 Orsay, France
[4] Department of Plant and Microbial Biology, University of California, Berkeley, CA 94720-3102
[5] Department of Agroecology, Aarhus University, Flakkebjerg, DK-4200 Slagelse, Denmark
* sajid.ali@grignon.inra.fr

Ali et al. BMC Research Notes 2011, 4:240
http://www.biomedcentral.com/1756-0500/4/240

BMC Research Notes

TECHNICAL NOTE Open Access

A rapid genotyping method for an obligate fungal pathogen, *Puccinia striiformis* f.sp. *tritici*, based on DNA extraction from infected leaf and Multiplex PCR genotyping

Sajid Ali[1*], Angélique Gautier[1], Marc Leconte[1], Jérôme Enjalbert[2] and Claude de Vallavieille-Pope[1]

Abstract

Background: *Puccinia striiformis* f.sp. *tritici* (PST), an obligate fungal pathogen causing wheat yellow/stripe rust, a serious disease, has been used to understand the evolution of crop pathogen using molecular markers. However, numerous questions regarding its evolutionary history and recent migration routes still remains to be addressed, which need the genotyping of a large number of isolates, a process that is limited by both DNA extraction and genotyping methods. To address the two issues, we developed here a method for direct DNA extraction from infected leaves combined with optimized SSR multiplexing.

Findings: We report here an efficient protocol for direct fungal DNA extraction from infected leaves, avoiding the costly and time consuming step of spore multiplication. The genotyping strategy we propose, amplified a total of 20 SSRs in three Multiplex PCR reactions, which were highly polymorphic and were able to differentiate different PST populations with high efficiency and accuracy.

Conclusion: These two developments enabled a genotyping strategy that could contribute to the development of molecular epidemiology of yellow rust disease, both at a regional or worldwide scale.

Background

Puccinia striiformis f.sp. *tritici* (PST), an obligate basidiomycete that causes wheat yellow/stripe rust, a serious disease in all major wheat growing regions [1-5]. The development of different molecular markers has aided the description of possible PST migration patterns [6], the emergence of high temperature-adapted strains [7,8] and the existence of recombination [9,10]. Despite these recent developments, numerous questions still need to be addressed, e.g. the evolutionary history of PST, its centre of origin, its historic migration pathways or more recent migrations causing new epidemics. These studies necessitate the genotyping of a large number of isolates, a process that is limited by both DNA extraction and genotyping methods.

Two or three cycles of PST spore multiplication on plants are usually necessary after sampling before DNA extraction. Because of its obligate nature, PST cannot be cultured on routine media to obtain sufficient biomass for DNA extraction [11]. Spore production may be further complicated when dealing with exotic isolates, which involve the mandatory use of expensive, time-consuming and wholly-contained facilities. In addition, using a given set of susceptible varieties to increase the spores of exotic isolates may give rise to bias. We had previously observed very low levels of infection or even resistance reactions in previously considered fully susceptible varieties such as cv. Victo [12], Michigan Amber and Cartago when inoculated with Pakistani isolates. This can result in the loss of isolates having avirulence factors recognized by unknown resistance genes in varieties used to increase spores. One alternative is to extract DNA from one or few spores, and then increase it through a whole genome multiple displacement amplification [11] before performing genotyping. However, the sophistication required for this method, as

* Correspondence: sajid.ali@grignon.inra.fr
[1]UMR1290 BIOGER-CPP, INRA-AgroParisTech, BP01, 78850 Thiverval-Grignon, France
Full list of author information is available at the end of the article

well as the need to prevent any contamination from other organisms, limits its use. We therefore tested here a third procedure, i.e. the direct extraction of fungal and plant DNA from single spore-infected leaf.

Another issue was the availability of a set of molecular markers sufficient to describe the population structure of a pathogen. The use of microsatellites/simple sequence repeat (SSR) markers with co-dominance and high polymorphism is of considerable value to the study of dikaryotic fungi such as PST [13]. When SSR detection and allele sizing are performed using an automated DNA fragment analyzer based on the separation of fluorescently-labeled amplicons, accurate and efficient genotyping can be achieved [14]. A good way to further enhance the efficiency of SSR genotyping is to multiplex SSR amplifications. Multiplex PCR refers to the simultaneous amplification of several markers in a single reaction, thereby saving the time and money required to manage each PCR reaction separately [15]. This method has been reported as achieving the same specificity as single conventional PCR reactions [14]. We report here a protocol that enabled the amplification of a set of highly informative SSR markers for PST studies, by means of three multiplex PCR reactions for PST.

Findings

Improvements to PST DNA extraction

To extract DNA from infected leaves, we selected both sporulating leaves from the first cycle of spore multiplication and leaves sampled in the field with single sporulating lesion, i.e. infections presumably resulting from a single spore infection. A Qiagen DNeasy® plant mini DNA extraction kit protocol was used to extract DNA from the sample. The infected leaves were put in 1.5 μL Eppendorf tubes containing 70 μL lysis (AP-1) solution (65°C) and one tungsten bead, and then ground for 3 min at 30 rps with a Retsch-MM300 grinder. An additional 70 μL of lysis (AP-1) solution was added and the liquid material was transferred into special tubes for extraction using the Qiagen DNA 'Biorobot 3000' extraction robot. Further details regarding the protocol are available in the manual (available at http://www.qiagen.com). The DNA thus extracted was then quantified on Nanodrop spectrophotometer (Thermo Scientific) and stored at -20°C until further use for molecular genotyping. The quantity of DNA was within the range 2 ng μL^{-1} to 6 ng μL^{-1}, depending on the size of the leaf lesion. It is essential to remember that these extracts were containing the DNA of both the plant and the fungus. Three PCR multiplex reactions (targeting 20 SSRs, see below) were used to determine whether the PST DNA was of sufficiently good quality for genotyping. The gel pictures for the three multiplexes run with DNA extracted from both the spores and the infected leaves are shown in Figure 1. Although the DNA extractions from infected leaves produced smaller quantities of PCR-amplicons, the amplifications were always sufficient to produce unambiguous signals for all loci when run on a sequencer. To have an idea about the relative amount of PST DNA in wheat DNA, we made a dilution of DNA extracted from PST infected lesion and DNA extracted from pure PST spores. In case of infected leaf extracted DNA, a dilution of two times resulted in no or very weak amplifications, while for DNA extracted from PST spores, a dilution of 20 times was necessary to reach to the lowest concentrations for PCR amplification. This indicates that the infected leaf DNA contains 1/5 of fungal DNA and 4/5 of leaf DNA. Thus at least 3 μL of the infected leaf extracted DNA must be taken to have enough PCR amplification. Further dilution will reduce the amplification to be able to read the alleles unambiguously. Extraction made from infected leaves of different ages, different varieties or with different infection times gave equal amplification. However, the quantity of the infection lesion is important. An infection lesion of at least 2 cm long must be used to have enough quantity of pathogen spores and hyphae for DNA extraction and subsequent amplification. Indeed, the amplification of DNA extracted from non-infected leaves from two wheat varieties (Michigan Amber and Sogood) did not achieve any amplification (Figure 1). The extractions from infected leaves were first validated in four different isolates (two from France and two from Pakistan), for which SSR typing was strictly identical to the genotypes obtained from the DNA extracted from spores (data not shown). This technique was then applied to a set of more than 100 isolates sampled in Pakistan in 2010 which displayed little or no spore production during multiplication. The SSR amplification was highly efficient, while allelic patterns were matching those of isolates sampled at the same location but genotyped after DNA extraction from spores (data not shown). This technique could also be used for single sporulating lesion infected leaves sampled in the field as such an infection normally results from a single spore and has the same genotype. However, in the case of leaves infected with several sporulating lesions, they need to be cloned to obtain single spore lesions that can then be used in the same manner as discussed above. In both cases however, at least half of the sporulating lesions must be preserved in order to retrieve the isolate if it is required in the future. One option would be to isolate samples through mono-chlorosis isolation or single spore inoculation, and then use the infected leaf for DNA extraction while at the same time retrieving the spores for future pathotyping and other biological/epidemiological studies. This direct use of infected leaves for DNA extraction is applicable for any wheat pathogen while considering the purity of races per lesion, eliminating the costly and critical step of exotic isolate multiplication in full confinement.

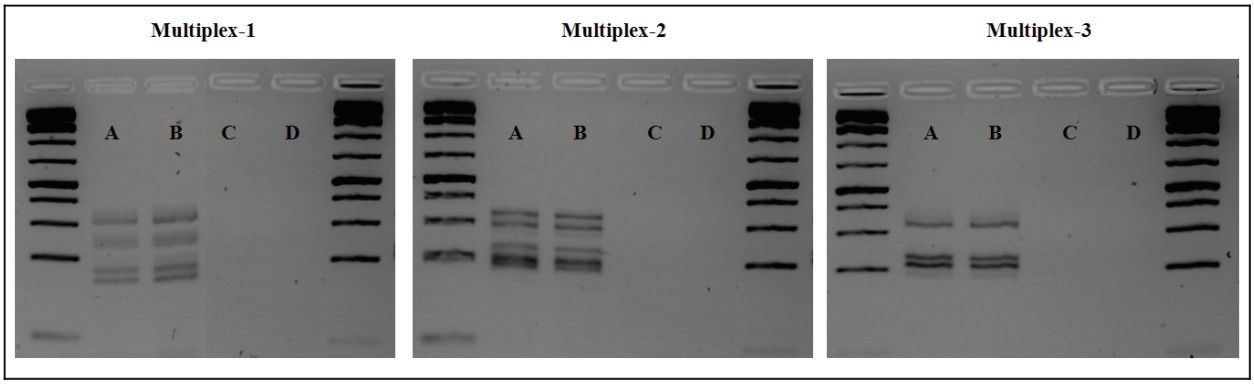


Figure 1 Gel run for the amplification of three multiplexes for DNA extracted directly from spores and from leaves infected with PST. A: J1023M2 spore extracted, B: DNA extracted from plant infected with J1023M2, C: Non-infected cv. Michigan Amber seedling leaf DNA, D: Non-infected cv. Sogood adult plant DNA.

Development of an efficient and multiplexed set of SSR for the PST population study

Another aim of this study was to obtain a reliable combination of different SSR markers that could be amplified during a small number of PCR reactions, rather than a separate PCR amplification for each SSR. We benefited from a total set of 22 previously developed SSRs [13,16] and a set of 13 SSRs sequences kindly provided by Dr. X. Chen (Washington State University, Pullman, USA), out of which only nine were amplified and two were polymorphic. Amplifications were performed first of all on each individual SSR and they were then combined in three multiplexes as a function of their allele sizes. Based on their polymorphism (assessed using a set of eight isolates representing different genetic groups of global PST populations) and after selection for a successful amplification in multiplex PCR reactions, a total of 20 SSRs were chosen. For each SSR locus, the forward primer was labelled with black, green or blue florescent dyes, the red fluorescent dye being reserved for the size marker (Figure 2). SSR loci

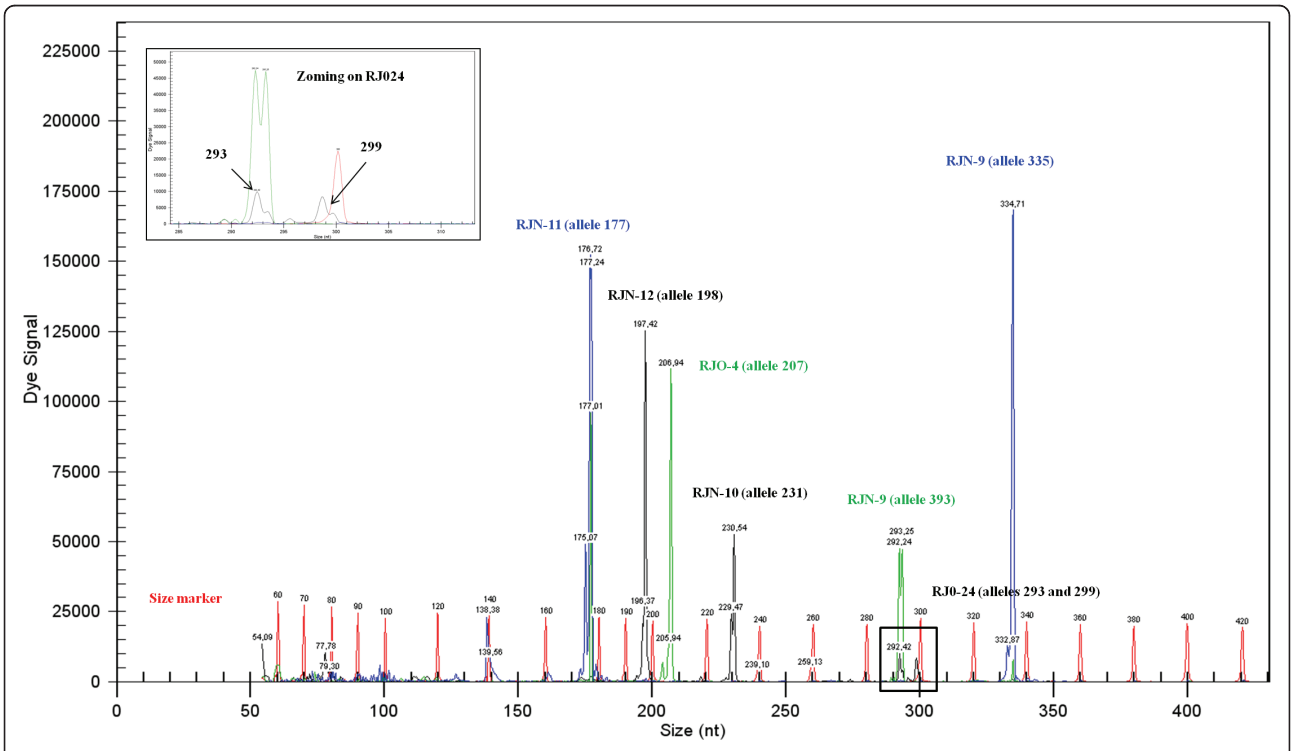


Figure 2 Chromatogram of Multiplex-2 with seven SSR loci labelled with blue (2 SSRs), green (2SSRs) and black (3 SSRs) fluorescent dyes, while the red dye represents the length markers. The RJO-24 locus with a low strength peak could be read after zooming (top left).

in the same or closed allele size range were then labelled with different florescent dyes to achieve the maximum possible number of loci per run of the sequencer. Table 1 provides details on the SSRs combined in each multiplex PCR reaction.

The PCR reactions were performed using a QIAGEN kit containing a single mix of Taq-polymerase, MgCl, dNTPs and buffer, referred to as the Type-it microsatellite kit specially designed for Multiplex PCR reactions. Each reaction contained 2 μL water, 1 μL Q-solution, 1 μL of the primer mix (containing 2 μM of each SSR primer), 5 μL of the Type-it mix and 1 μL (15 ng) of DNA. The amount of DNA was increased to 3 μL and no water was added for infected plant leaves, as the PST DNA was diluted with plant DNA. An optimization step was performed to identify the optimum melting temperature for all the SSRs in a given Multiplex. The optimum PCR conditions thus determined were the same for all three multiplexes, with preheating at 95°C for 5 min 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 at 60°C for 30 minutes, using an iCycler (Biorad) thermocycler. The PCR products were run on a 2% agarose gel to reveal the amplification products. The amounts of PCR product added to 35 μL of the Sample Loading Solution, containing 0.4 μL Beckman Coulter 400-bp size standard, varied as follows:1.5 μL for Multiplex-1 and Multiplex-3 and 2.2 for Multiplex-2 in the case of spore DNA; 3 μL for Multiplex-1 and Multiplex-2 and 4 μL for Multiplex-2 in the case of infected leaf DNA. Amplicon fragments were separated using a Beckman Coulter CEQ-8000 DNA Analyzer with the default FRAG-3 run method. The fragments were read using CEQ-8000 Genetic Analysis System Software (Beckman Coulter) to record alleles manually for each locus according to amplicon fragment lengths. All the alleles were readable and there was no difference between allele lengths whether the PCR was performed for each SSR separately or in a multiplex reaction. Allele sizes were within the range of previously reported alleles for previously developed SSR markers [13,16], while for two new SSRs, the allele sizes were within the range of 211-213 bp for WU-6 and 325-334 bp for WU-12. The technique was then used to genotype more than a thousand isolates representing the worldwide PST population and we found no ambiguity in terms of allele reading, together with an efficient discriminating power for this set of SSRs. The data on these worldwide set of isolates would be used to infer about the PST phylogeny and evolutionary biology. The development of this multiplex-based amplification technique, together with the reading of allele length through a sequencer, achieved gains in time and accuracy, as well as regarding the reproducibility of the results.

Conclusions

The extraction of DNA from infected leaves, together with a Multiplex-based PCR reaction and the reading of fluorescently labeled alleles through a sequencer thus provides a ready-to-use method for the efficient

Table 1 Description of three PCR multiplexes enabling the genotyping of 20 *Puccinia striiformis* f.sp. *tritici* SSR

Multiplex	SSR Locus	Florescence	Allele size range	Reference
Multiplex 1	RJN5	Black	225-231	[13]
	RJN6	Black	310-322	[13]
	RJN13	Green	150-153	[13]
	RJN3	Green	339-347	[13]
	RJO21	Blue	167-182	[16]
	RJN4	Blue	258-264	[13]
	RJN8	Blue	306-318	[13]
	RJO18	Blue	334-360	[16]
Multiplex 2	RJN12	Black	192-200	[13]
	RJN10	Black	225-231	[13]
	RJO24	Black	273-308	[16]
	RJO4	Green	201-207	[16]
	RJO20	Green	282-293	[16]
	RJN11	Blue	173-185	[13]
	RJN9	Blue	335-337	[13]
Multiplex 3	RJN2	Black	172-196	[13]
	WU6	Black	211-213	Provided by X. Chen
	RJO3	Green	202-204	[16]
	WU12	Green	325-334	Provided by X. Chen
	RJO27	Blue	217-243	[16]

genotyping of PST, and enables clear gains in terms of time, money, reproducibility and accuracy. Because of the high level of polymorphism of the SSR markers selected, the proposed SSR set could also constitute a genotyping reference at worldwide level, enabling the rapid comparison of genetic analyses. Such easily comparable sets of markers constitute an essential tool for molecular epidemiology and to trace emerging races in a fungus that is known for its highly efficient long distance migration [17]. Furthermore, the genotyping of large number of isolates from different geographical regions coupled with recent population genetics analyses would assist to address ancestral relationship between different geographically spaced populations, describe the ancient migration routes and the role of host and geography on pathogen population structuring. This will help us to understand overall evolution of pathogens and to consequently orientate disease management strategies.

Acknowledgements
We would like to thank Laurent Gérard and Nathalie Galet for their technical assistance during the multiplication of yellow rust isolates. We thank X. Chen for kindly sharing the list of 13 sequences for candidate SSRs. This work was supported by the European Integrated Project Bioexploit, FOOD-CT-2005-513959 and the ANR (Agence Nationale de la Recherche) through the EMERFUNDIS project (ANR 07-BDIV-003). Sajid Ali was supported by the Higher Education Commission, the Government of Pakistan.

Author details
[1]UMR1290 BIOGER-CPP, INRA-AgroParisTech, BP01, 78850 Thiverval-Grignon, France. [2]UMR 320 Génétique Végétale, INRA, Ferme du Moulon, 91190 Gif sur Yvette, France.

Authors' contributions
SA, AG, CP and JE designed the study; SA, ML and AG carried out the experimental work; SA, JE and CP prepared the manuscript. All authors have read and approved the manuscript.

Competing interests
The authors declare that they have no competing interests.

Received: 21 February 2011 Accepted: 20 July 2011
Published: 20 July 2011

References

1. Ali S, Shah SJA, Rahman H: **Multi-locations variability in Pakistan for partial resistance in wheat to *Puccinia striiformis* West. *tritici*.** *Phytopathologia Mediterranea* 2009, **48**:269-279.
2. Singh RP, William HM, Huerta-Espino J, Rosewarne G: **Wheat rust in Asia: meeting the challenges with old and new technologies.** *New directions for a diverse planet: 4th International Crop Science Congress: 26 Sep - 1 Oct 2004; Brisbane, Australia* 2004.
3. Hovmøller MS: **Disease severity and pathotype dynamics of *Puccinia striiformis* f.sp. *tritici* in Denmark.** *Plant Pathology* 2001, **50**:181-189.
4. Hovmøller MS, Walter S, Justesen AF: **Escalating threat of wheat rusts.** 2010, **329**:369.
5. De Vallavieille-Pope C, Ali S, Leconte M, Enjalbert J, Delos M, Rouzet J: **Virulence dynamics and regional structuring of *Puccinia striiformis* f. sp. *tritici* in France between 1984 and 2009.** *Plant Disease* 2011.
6. Bahri B, Leconte M, Ouffroukh A, de Vallavieille-Pope C, Enjalbert J: **Geographic limits of a clonal population of wheat yellow rust in the Mediterranean region.** *Molecular Ecology* 2009, **18**:4165-4179.
7. Ali S, Leconte M, Walker A-S, Enjalbert J, de Vallavieille-Pope C: **Reduction in the sex ability of worldwide clonal populations of *Puccinia striiformis* f.sp. *tritici*.** *Fungal Genetics and Biology* 2010, **47**:828-838.
8. Hovmøller MS, Justesen AF: **Appearance of atypical *Puccinia striiformis* f. sp. *tritici* phenotypes in north-western Europe.** *Australian Journal of Agricultural Research* 2007, **58**:(6):518-524.
9. Mboup M, Leconte M, Gautier A, Wan AM, Chen W, de Vallavieille-Pope C, Enjalbert J: **Evidence of genetic recombination in wheat yellow rust populations of a Chinese oversummering area.** *Fungal Genetics and Biology* 2009, **46**:(4):299-307.
10. Duan X, Tellier A, Wan A, Leconte M, de Vallavieille-Pope C, Enjalbert J: ***Puccinia striiformis* f.sp. *tritici* presents high diversity and recombination in the over-summering zone of Gansu-China.** *Mycologia* 2010, **102**:44-53.
11. Wang Y, Zhu M, Zhang R, Yang H, Wang Y, Sun G, Jin S, Hsiang T: **Whole genome amplification of the rust *Puccinia striiformis* f. sp. *tritici* from single spores.** *Journal of Microbiological Methods* 2009, **77**:229-234.
12. Bahri B, Shah SJA, Hussain S, Leconte M, Enjalbert J, De Vallavieille-Pope C: **Genetic diversity of wheat yellow rust population in Pakistan and its relationship with host resistance.** *Plant Pathology* 2011.
13. Bahri B, Leconte M, de Vallavieille-Pope C, Enjalbert J: **Isolation of ten microsatellite loci in an EST library of the phytopathogenic fungus *Puccinia striiformis* f.sp. *tritici*.** *Conservation Genetics* 2009, **10**:1425-1428.
14. 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: doi:10.1186/1471-2164-1189-1180.
15. Henegariu O, Heerema NA, Dlouhy SR, Vance GH, Vogt PH: **Multiplex PCR: Critical parameters and step-by-step protocol.** *Biotechniques* 1997, **23**:504-511.
16. Enjalbert J, Duan X, Giraud T, Vautrin C, de Vallavieille-Pope C, Solignac M: **Isolation of twelve microsatellite loci, using an enrichment protocol, in the phytopathogenic fungus *Puccinia striiformis* f. sp. *tritici*.** *Molecular Ecology Notes* 2002, **2**:563-565.
17. Brown JKM, Hovmøller MS: **Aerial dispersal of fungi on the global and continental scales and its consequences for plant disease.** *Science* 2002, **297**:537-541.

doi:10.1186/1756-0500-4-240
Cite this article as: Ali *et al.*: A rapid genotyping method for an obligate fungal pathogen, *Puccinia striiformis* f.sp. *tritici*, based on DNA extraction from infected leaf and Multiplex PCR genotyping. *BMC Research Notes* 2011 **4**:240.

Virulence Dynamics and Regional Structuring of *Puccinia striiformis* f. sp. *tritici* in France Between 1984 and 2009

Claude de Vallavieille-Pope, Sajid Ali, and **Marc Leconte,** UMR 1290 BIOGER-CPP, INRA-AgroParisTech, BP01, 78850 Thiverval-Grignon, France; **Jérôme Enjalbert,** UMR 320 Génétique Végétale, INRA, Ferme du Moulon, 91190 Gif sur Yvette, France; **Marc Delos,** SRAl, Cité administrative, Bat E, bd Armand Duportal, 31074 Toulouse cedex, France; and **Jacques Rouzet,** SRAl, rue Alco, BP 3056, 34034 Montpellier cedex France

Abstract

de Vallavieille-Pope, C., Ali, S., Leconte, M., Enjalbert, J., Delos, M., and Rouzet, J. 2012. Virulence dynamics and regional structuring of *Puccinia striiformis* f. sp. *tritici* in France between 1984 and 2009. Plant Dis. 96:131-140.

Understanding of long-term virulence dynamics of pathogen populations in response to host resistance gene deployment is of major importance for disease management and evolutionary biology. We monitored the virulence dynamics of *Puccinia striiformis* f. sp. *tritici*, the causal agent of wheat stripe rust, over 25 years in France. Virulence dynamics was explained by estimates of area associated with resistance genes carried by farmers' cultivars. The epidemics assessed through disease severity significantly correlated with the number of *P. striiformis* f. sp. *tritici* isolates collected each year, used to describe virulence dynamics. In the south, the dominance of the Mediterranean pathotype 6E16 and the cultivation of a susceptible cultivar were associated with an epidemic from 1997 to 1999. In the north, five epidemics occurred due to successive acquisition of virulence to the resistance genes *Yr7*, *Yr6*, *Yr9*, *Yr17*, and *Yr32*, either by acquisition of the virulence in the previous dominant pathotype or by incursion or selection of one or two new pathotypes. Frequency of pathotypes with *Vr7* and *Vr6* declined with the reduction in the cultivation of corresponding *Yr* gene cultivars, whereas the virulence *Vr9* persisted longer than the cultivation of *Yr9* cultivars. Although the first pathotypes carrying *Vr9* decreased, this virulence persisted in other pathotypes even in the absence of *Yr9* cultivars. At the regional level, *Yr9* cultivars in the north caused a shift from high *Vr6* frequency to high *Vr9* frequency whereas, in the central region, where *Yr9* cultivars were rare, *Vr6* remained prevalent.

Stripe rust (yellow rust) of wheat (*Triticum aestivum* L.), caused by *Puccinia striiformis* f. sp. *tritici* Eriks., is a disease of worldwide economic importance (2,23,30,50,52). The deployment of resistance genes has been suggested to be the most economical and environmentally friendly measure to control the disease (2,22,48,50). At least 53 yellow rust resistance genes (*Yr*) have been identified (42,53). These resistance genes are designated with *Yr* and a suffix (i.e., *Yr1, Yr2, Yr3*, and so on) according to the sequence of their genetic characterization (42), and the matching virulences to these resistance genes are designated as *Vr1, Vr2, Vr3*, and so on, respectively. These resistance genes are mostly race specific, fitting the well-described gene-for-gene theory (14,15,26). Most of these genes are expressed at all growth stages of the plant but some are effective only at the adult-plant stage. Most of these resistance genes have been deployed in different cultivars, at different geographical scales (region, country, and continent), and for different periods of time, and, in general, have been associated with short-term stripe rust control (8,33,52).

When resistance genes are deployed in popular cultivars on a large area, strong selective pressure occurs in the pathogen population (33,35,41,52,56). Selection pressure caused by the deployment of resistance genes within host populations is particularly effective for biotrophic pathogens such as *P. striiformis* f. sp. *tritici*, because their entire life cycle depends on the availability of susceptible living host tissue. Thus, virulent pathotypes are rapidly selected, leading to a loss of resistance-gene efficacy and, subsequently, to a reduction in the cultivation of cultivars carrying this sole source of resistance (8,41). The overall pathogen population may then either become more complex through virulence accumulation in the same pathotype (22,36,52) via strong direct or indirect (hitchhiking) selection (32), or remain quite stable, having the same mean number of virulences per pathotype (4), depending on the disease and the region studied. Despite considerable interest in understanding how host selection pressure and pathogen adaptive potential interact in the temporal dynamics of epidemics, few studies have described the changes in virulence frequency over a relatively long period of time and, hence, taken into account the temporal evolution of area allocated to the corresponding resistance genes.

In France, stripe rust most frequently occurred in the north, where devastating epidemics in the 1960s ('Joss Cambier', *Yr11*) led to greater emphasis on breeding for resistance to this disease (49). Since that time, a number of resistance genes have been deployed in wheat cultivars. The *P. striiformis* f. sp. *tritici* population pathotype survey began in 1984 in order to monitor the evolution of virulence frequencies in the French *P. striiformis* f. sp. *tritici* population using a set of differential cultivars (21,22). The northwestern European *P. striiformis* f. sp. *tritici* population, which included northern French *P. striiformis* f. sp. *tritici* populations, was described as being predominantly clonal (25,29) and, therefore, most similar to that reported for the Australian (55) and northern American (12) *P. striiformis* f. sp. *tritici* populations, while Oriental and Asiatic populations appeared to have varying level of diversity (1,24,40). Despite this clonal behavior, a succession of pathotypes has been detected in France across six major epidemics, due to stepwise mutations and acquisition of new virulences (25). The clonality of *P. striiformis* f. sp. *tritici* populations in France and the succession of a low number of pathotypes dominating the country provided a unique model to

Corresponding author: C. de Vallavieille-Pope,
E-mail: pope@grignon.inra.fr

Accepted for publication 13 July 2011.

http://dx.doi.org/10.1094/PDIS-02-11-0078

study the temporal dynamics of adaptation of *P. striiformis* f. sp. *tritici* to resistance gene selective pressures.

Despite the long-distance migration capacity of *P. striiformis* f. sp. *tritici* (11), a clear divergence has been reported in the northern and southern French *P. striiformis* f. sp. *tritici* populations (25), partly due to temperature adaptation (39) and, apparently, to resistance gene deployment. This highlights the role of host genetic composition at the regional level. To better understand the role of host gene deployment on regional population structure, we used the case of *Yr6* and *Yr9* cultivation in two regions. We assessed whether the cultivation of the *Yr9*-based 'Slejpner' in northern France and of the *Yr6*-based 'Récital' in both the northern and central regions (GNIS, http://www.gnis.fr/) (22) might influence the *P. striiformis* f. sp. *tritici* population structure at the regional level.

The first objective of this study was to determine whether the pathotype dynamics of *P. striiformis* f. sp. *tritici* in France may be explained by changes in the deployment of resistance genes at the national level, with particular emphasis on the subsequent evolution of the corresponding virulences over a long period of time. The second objective was to see whether different host population structures at the regional level made it possible to study their effect on the pathogen population structure. We used the annual pathotype and epidemic severity surveys conducted at the national level to find answers to these questions.

Materials and Methods

Temporal fluctuations in the severity of stripe rust epidemics. Annual surveys of the severity of stripe rust epidemics were carried out by the Plant Protection Service, Ministry of Agriculture, Paris. These estimates were used as indicators of epidemic intensity. Detailed observations on disease severity were available from 1975 in all production regions within France and had been undertaken by a specialized network concerning cereal diseases. Disease severity was assessed by visual scoring in farmers' fields throughout France. For the present study, the scores averaged across all sites from northern and southern France each year were represented on a scale from 0 to 3, where 0 = no disease; 0.5 = rare epidemic cases observed; 1 = limited disease severity, either low severity across a large area or high severity limited to a very small area; 2 = severe disease severity over a large area with maximum severity (i.e., 70%); and 3 = high disease severity with a larger area severely infected (disease severity >70%). The disease severity score presented herein was the median based on the distribution of severity scores per year for representative regions in the north and south of France. Each region comprised about 300 farmers' fields and 10 different cultivars. The disease evaluation in each reference field in an area not treated with fungicides was based on weekly visual observations of diseased leaf area on susceptible or highly susceptible cultivars.

Survey of *P. striiformis* f. sp. *tritici* pathotypes. The frequency of *P. striiformis* f. sp. *tritici* pathotypes was monitored throughout France in farmers' fields and noninoculated breeding nurseries beginning in 1984. For the present study, information was obtained from a range of institutes, including Service de la Protection des Végétaux, Ministry of Agriculture, Paris; Arvalis—Institut du Végétal, Paris; Chambers of Agriculture; agrochemical companies; and wheat breeders. Randomly selected samples of diseased wheat leaves were collected from fields that were not inoculated. The leaves were wrapped in glycine bags or paper envelopes and sent to the INRA laboratory in Versailles-Grignon. An isolate of a single sporulating lesion was obtained from each sample and then used for subsequent pathotyping.

The *P. striiformis* f. sp. *tritici* survey performed in northern France during the period from 1987 to 1992 was used to assess the regional structure of the *P. striiformis* f. sp. *tritici* population in relation to resistance gene distribution of cultivars in farmers' fields. In one region of northern France that included the cities of Lille (50°38′14″ N, 3°03′48″ E) and Amiens (49°54′0″ N, 2°18′0″ E), Slejpner, which carried the resistance gene *Yr9* (7), was widely cultivated at that time; whereas, in one region of central France, including Paris (48°51′24″ N, 2°21′07″ E) and Orléans (47°54′09″ N, 1°54′32″ E), Slejpner was not grown and Récital (carrying *Yr6*; 22) was predominant. 'Thésée' (*Yr2*) was grown in both regions and was susceptible to all prevalent pathotypes (either avirulent or virulent to *Yr6* and *Yr9*), because all pathotypes found during this period carried *Vr2* (22). The frequencies of *Vr9* and *Vr6* were compared between the two regions to assess the regional structure of the *P. striiformis* f. sp. *tritici* population relative to the distribution of resistance genes.

Spore purification and multiplication. The *P. striiformis* f. sp. *tritici* samples collected from the field were purified and multiplied in a greenhouse for further pathotyping. Five wheat seeds per pot were planted in square pots (7 by 7 by 8 cm) filled with standard peat soil. In order to enhance spore production, the plants were treated with 15 ml of maleic hydrazide solution (0.25 g/liter) per pot when seedlings were 1 cm high, 4 to 5 days after sowing. All isolates were single lesion and urediniospores were produced through inoculation on seedlings at the two-leaf stage of the susceptible wheat 'Michigan Amber' or 'Victo'. The latter cultivar is highly susceptible to the northwestern European stripe rust population and resistant to most of the leaf (brown) rust pathotypes (23). Starting in the early 2000s, the plants were kept in a greenhouse under continuous light for 6 h before inoculation to increase infection efficiency, with a light intensity of 300 µmol quanta m^{-2} s^{-1} (19). Inoculated plants were kept in a dark dew chamber at 8°C for 24 h and then placed in a greenhouse where the temperature was maintained between 14°C at night and a maximum of 20°C during day time. The daylight period was extended to 16 h through the use of sodium vapor lamps (with a light intensity of 300 µmol quanta m^{-2} s^{-1}). Plastic cylinders or cellophane bags were placed over each pot to prevent any cross-contamination. Plants bearing a single lesion on a leaf were isolated in a cellophane bag to produce single-lesion isolates, and the spores collected from this single lesion were multiplied for further use. Seedlings were inoculated with urediniospores mixed with talc. Urediniospores from plants inoculated 15 to 21 days earlier were harvested using a cyclone spore collector. These spores were placed in a desiccator (40% relative humidity) for 4 days at 5°C and then stored in liquid nitrogen (20). After removal from storage, the urediniospores were heat shocked (40°C for 10 min) before inoculation and used for either spore multiplication or virulence profile tests.

Characterization of virulence profiles of *P. striiformis* f. sp. *tritici*. The virulences of *P. striiformis* f. sp. *tritici* isolates were determined using world and European sets of 15 differential lines (22,34) plus additional varieties: 'Chinese 166' (*Yr1*), 'Kalyonsona' (*Yr2*), 'Heines VII' (*Yr2*+), 'Vilmorin 23' (*Yr3*), 'Nord Desprez' (*Yr3*+), 'Hybrid 46' (*Yr4*), *T. spelta album* (*Yr5*), 'Austerlitz' (*Yr6*), 'Heines Kolben' (*Yr6*+), 'Heines Peko' (*Yr6* and *Yr2*), 'Lee' (*Yr7*), 'Reichersberg 42' (*Yr7*+), 'Compair' (*Yr8*+), 'Federation 4×/Kavkaz' (*Yr9*), 'Clement' (*Yr9*+), 'Moro' (*Yr10*), 'VPM1' (*Yr17*), line TP981 (*Yr25*), Avocet isoline (*Yr27*), 'Carstens V' (*Yr32*), 'Strubes Dickkopf' (*YrSd*), 'Spaldings Prolific' (*YrSp*), and 'Suwon92/Omar' (*YrSu*) (http://www.ars.usda.gov/SP2UserFiles/ad_hoc/36400500Resistancegenes/Yrgene.xls). Each differential line carried at least one race-specific resistance gene (*Yr*) expressed at the seedling stage. The "+" sign after the name of the *Yr* gene indicated additional unknown resistance. During each test, Michigan Amber or Victo was included as a susceptible control (22,23). Each test included six plants for each isolate–wheat genotype combination and was conducted twice. All the conditions of plant growth and inoculum production were as described above. The seedlings were inoculated 8 to 10 days after sowing, when the first leaf was fully expanded and the second leaf emerging. Seedlings were inoculated using 4 mg of spores suspended in 650 µl of Soltrol 170 mineral oil (Chevron-Phillips Chemical Co., Houston). After 15 min for oil evaporation, inoculated seedlings were kept in a dark dew chamber for 24 h at 8°C before being placed in a greenhouse (as described above) to promote disease development. Seedling responses were assessed 17 to 19 days after

inoculation, using a qualitative infection-type scale of 0 to 9 defined by the presence of necroses and chloroses and the intensity of sporulation (43) (http://public.wsu.edu/~wheaties/rustpic.html). Compatible reactions (host susceptibility and pathogen virulence) were defined as infection types 7 to 9. Infection types 0 to 4 were defined as incompatible reactions (host resistance and pathogen avirulence) and infection types 5 to 6 were considered to be intermediate reactions.

The virulence phenotypes of the isolates were used according to the pathotyping nomenclature (34) and their complete virulence profile. The pathotype code is binary; the first number corresponded to the avirulent or virulent infection types on the world differential set and the second number described avirulent or virulent infection types on the European differential set. The final number indicated avirulent or virulent infection types on the additional differential lines for *Yr17* and *Yr27*.

Postulation of stripe rust resistance genes. The geographical distribution of cultivars was provided by the Office National Interprofessionnel des Grandes Cultures, Paris (47) and Arvalis—Institut du Végétal, Paris. Resistance genes were determined at the seedling stage using a set of 12 French pathotypes with complementary virulence patterns: 2 southern *P. striiformis* f. sp. *tritici* pathotypes, 6E16 (*Vr2, 6, 7*, and *8*) and 6E16V27 (*Vr2, 6, 7, 8, 9, 25*, *27,* and *Sd*); and 10 northern *P. striiformis* f. sp. *tritici* pathotypes: 237E141V17 (*Vr1, 2, 3, 4, 6, 9, 17, Sd,* and *Su*), 169E136V17 (*Vr1, 2, 3, 9, 17,* and *Sd*), 40E8 (*Vr3* and *Sd*), 43E138 (*Vr1, 2, 3, 7,* and *Sd*), 106E139 (*Vr2, 3, 4, 7, Sd*, and *Su*), 109E141 (*Vr1, 2, 3, 4, 6, Sd*, and *Su*), 232E137 (*Vr2, 3, 4, 9, Sd*, and *Su*), 233E169V17 (*Vr1, 2, 3, 4, 9, 17, 32, Sd*, and *Su*), 237E141 (*Vr1, 2, 3, 4, 6, 9, Sd*, and *Su*), and 239E141V17 (*Vr1, 2, 3, 4, 6, 7, 9, 17, Sd*, and *Su*) (22). The cultivars grown during the study period were tested at the seedling stage, with the same conditions as described above for pathotype determination (22).

Results

Prevalence and severity of stripe rust epidemics. In total, 1,574 isolates were collected over the 25-year period: 1,362 isolates from the north and 212 from the south (Fig. 1). The geographical distribution of the number of isolates sampled per department was six times higher in the north than in the south. North-central France (departments of Calvados, Essonne, Eure, Eure et Loir, Ille et Vilaine, Loir et Cher, Nord, Oise, Pas de Calais, Seine et Marne, Seine Maritime, Somme, and Yvelines) had the highest number of isolates collected in the north, ranging from 25 to 206 isolates per department. In southwestern and eastern France, where the area of wheat cultivation was lower (www.gnis.fr), fewer isolates were received. The analyses of pathotypes and disease severity considered the north and south of France separately. The maximum number of isolates received (n = 200) was in 1999, with 180 of them originating from northern regions with a disease severity score of 2, corresponding to the 1999 to 2002 *P. striiformis* f. sp. *tritici* epidemics (Table 1; Fig. 2A). The number of isolates received in 1989 (121 isolates) corresponded to the 1988 to 1991 *P. striiformis* f. sp. *tritici* epidemics. The number of isolates in 2008 (124 isolates) that corresponded to the 2007 to 2009 *P. striiformis* f. sp. *tritici* epidemics was also high. A moderate epidemic observed in 1994 (57 isolates) corresponded to the 1993 to 1995 *P. striiformis* f. sp. *tritici* epidemics. For southern France, the maximum number of isolates (n = 72) was collected in 1997 and corresponded to the 1996 to 2001 *P. striiformis* f. sp. *tritici* epidemics that were due to the cultivation of Victo. Those epidemic periods corresponded to disease severity scores higher than 2 on the 0-to-3 scale (Fig. 2A and B). There were four periods with no epidemics: in 1986, 1992, 1996, and 2004 to 2006. For these years, the number of isolates received was below 10 isolates per year and the disease severity score was lower than 0.5.

The sources of isolates were the surveys conducted by the Plant Protection Service, and it was hypothesized that the number of isolates collected each year was indicative of epidemic severity of that year. The severity of stripe rust epidemics and the number of isolates received each year were found to be strongly correlated (for the north: R^2 = 0.661 and P < 0.001, Fig. 2A; for the south: R^2 = 0.608 and P < 0.001, Fig. 2B), showing that the number of *P. striiformis* f. sp. *tritici* isolates received annually varied in line with epidemic severity (Fig. 2A and B). Because we were interested in the relative frequency of pathotypes and virulences, henceforth, epidemic status was described using the number of isolates received per year rather than disease severity.

The virulence frequency over the 25-year period in the north revealed a succession of five epidemics, due to the sequential selection of the virulences *Vr7*, *Vr6*, *Vr9*, *Vr17*, and *Vr32* (Fig. 3A). The *Vr7* frequency was high in 1984 and then declined. The *Vr6* frequency fluctuated three times from high to low and then to high again. The *Vr9* and *Vr17* virulences remained fixed in the population after their selection. Finally, the *Vr32* frequency was only high at the end of the period. In the south, there was a single large epidemic, which was caused by a pathotype combining *Vr7* and *Vr6* (Fig. 3B). *Vr9* and *Vr17* were detected in 2000 in the south, corresponding to some pathotypes characteristic of the north.

In total, 23 pathotypes were detected from the 1,574 *P. striiformis* f. sp. *tritici* isolates collected during the study period. In all, 20 pathotypes were identified in northern France and 14 in southern France (Table 1). In the north, a succession of pathotypes was detected during the 25-year period. During the epidemic in 1984, pathotype 43E138 containing *Vr7* virulence was dominant but decreased thereafter, coinciding with the diminishing production area of Talent (*Yr7*) (Fig. 4A). The epidemics of 1987 to 1989 were due to two new pathotypes (45E140 and 109E141) with *Vr6* virulence, collected mainly from *Yr6*-based Récital and Austerlitz (Fig. 4B). The next most frequent pathotype was 232E137, containing *Vr9* virulence, which was detected in 1986 and became prevalent between 1990 and 1994 before declining. The prevalence of 232E137 was mainly due to the cultivation of Slejpner (*Yr9*). In 1989, another pathotype (237E141) appeared carrying *Vr6* in addition to *Vr9*. This was concomitant with the cultivation of 'Equinox' and 'Madrigal', containing both *Yr6* and *Yr9*. In 1998, the two major newly detected pathotypes (233E137V17 and 169E136V17) contained both the *Vr9* and *Vr17*

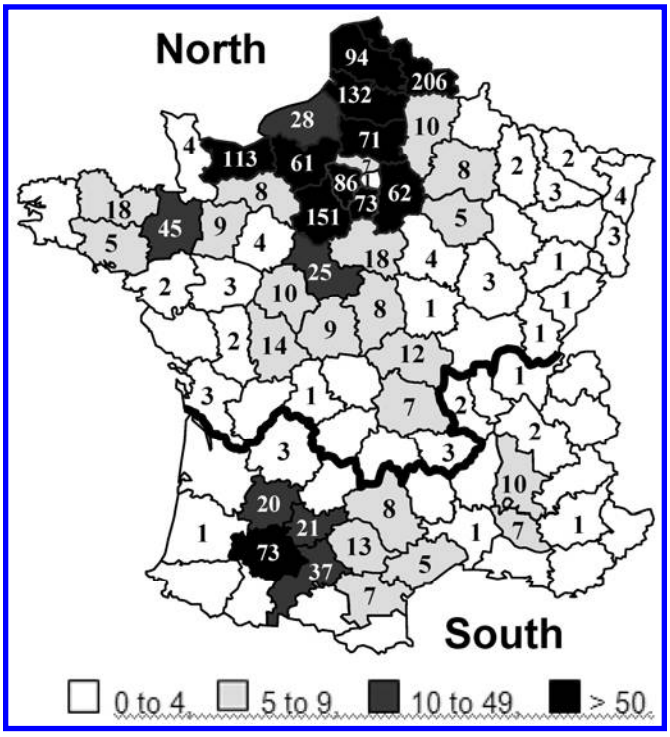


Fig. 1. Distribution of the 1,574 *Puccinia striiformis* f. sp. *tritici* isolates collected between 1984 and 2009 in 93 French departments. Range of numbers of isolates per department: white areas, 0 to 4; light gray areas, 5 to 9; dark gray areas, 10 to 49, and black areas, > 50.

virulences. The *Vr9* virulence was already widespread and the *Vr17* virulence was newly detected. The pathotypes 233E137V17 and 169E136V17 were found from 1998 to 2004. In 1999, a new pathotype (237E141V17) possessing *Vr9, Vr6,* and *Vr17* was identified, associated with the cultivation of 'Ornicar' and 'Cardos', containing both *Yr6* and *Yr17*. This pathotype frequency declined after 2003 and increased again in 2009.

Another pathotype (233E169V17), containing *Vr9, Vr17,* and *Vr32,* was detected in 2002 and became frequent during the epidemics of 2007 to 2009, while the pathotype 104E41, containing *Vr32* without *Vr9* and *Vr17,* was detected in 2008 and 2009. The *Yr32* gene was present mainly in 'Toisondor', 'Aldric', and 'Alixan'. Some other minor pathotypes were also detected during this 25-year period but at a very low frequency (Table 1).

By contrast, the southern French *P. striiformis* f. sp. *tritici* population was clearly dominated by the pathotype 6E16 (150 isolates), with the remaining pathotypes, including four specific to the south, being encountered much less frequently (Table 1). The highest number of 6E16 isolates sampled (53 isolates out of the total of 150 such isolates collected during the 25-year period) was observed in 1997. The earlier 1986 epidemic, due to infection of spring wheat 'Prinqual' in the Rhône-Alpes region (south of France), was also caused by the pathotype 6E16.

Moreover, despite the marked divergence between northern and southern *P. striiformis* f. sp. *tritici* populations, some pathotypes from the north were found at very low frequencies in the south during the years of severe epidemics (Table 1). Similarly, the dominant pathotype in southern France (6E16) was found at a very

Table 1. Annual frequency and virulence profile of *Puccinia striiformis* f. sp. *tritici* pathotypes detected in northern and southern France between 1984 and 2009

	Year[a]																									
Pathotype[b]	**84**	**85**	**86**	**87**	**88**	**89**	**90**	**91**	**92**	**93**	**94**	**95**	**96**	**97**	**98**	**99**	**00**	**01**	**02**	**03**	**04**	**05**	**06**	**07**	**08**	**09**	**Virulence[c]**
North																											
43E138	83	90	+	+	7	2	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	1,2,3,7,25,Sd
171E138	17	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	1,2,3,7,9,25,Sd
41E136	–	10	–	17	6	1	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	1,2,3,25,Sd
232E137	–	–	+	–	3	21	44	52	+	44	46	38	+	–	–	–	–	–	–	–	–	–	–	–	–	–	2,3,4,9,25,Sd,Su
109E141	–	–	–	6	15	14	13	3	–	–	2	3	–	+	3	–	–	2	–	–	–	–	–	–	–	–	1,2,3,4,6,25,Sd,Su
169E136	–	–	–	11	–	1	3	6	+	38	26	13	+	–	–	–	–	–	–	–	–	–	–	–	–	–	1,2,3,9,25,Sd
45E140	–	–	–	56	67	58	35	6	–	3	–	5	+	–	–	–	1	1	–	–	–	–	–	–	–	–	1,2,3,6,25,Sd
233E137	–	–	–	6	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	1,2,3,4,9,25,Sd,Su
235E139	–	–	–	–	3	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	1,2,3,4,7,9,25,Sd,Su
237E141	–	–	–	–	–	3	3	30	+	16	26	41	–	+	–	–	–	–	1	–	+	–	–	-	2	–	1,2,3,4,6,9,25,Sd,Su
106E139	–	–	–	–	–	–	2	–	–	–	–	–	–	+	–	–	–	–	–	–	+	–	–	-	–	–	2,3,4,7,25,Sd,Su
169E136V17	–	–	–	–	–	–	–	–	–	–	–	–	–	+	26	5	2	6	2	21	–	–	–	3	–	–	1,2,3,9,17,Sd
233E137V17	–	–	–	–	–	–	–	–	–	–	–	–	–	+	68	89	91	75	41	43	+	–	–	1	–	–	1,2,3,4,9,17,25,Sd,Su
237E141V17	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	1	1	11	48	29	–	–	–	1	–	42	1,2,3,4,6,9,17,25,Sd,Su
173E140V17	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	1	2	–	+	–	–	–	–	–	1,2,3,6,9,17,25,Sd
239E143V17	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	1	–	–	–	–	–	–	–	–	1,2,3,4,6,7,9,17,25,Sd,Su
233E169V17	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	5	7	–	–	–	96	77	31	1,2,3,4,9,17,32,25,Sd,Su
104E41	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	21	27	3,4,6,32,25,Sd,Su
6E16	–	–	+	–	–	–	–	3	–	–	–	–	–	–	3	4	4	4	2	–	–	–	–	–	–	–	2,6,7,8
40E8	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	2,3,25,Sd
0E0	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	_
32E0	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	1	–	–	–	–	–	–	–	–	–	–	25,Sd
6E16V27	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	2,6,7,8,9,25,27,(Sd)
Isolate no.	12	10	4	18	72	121	63	33	4	32	57	39	8	8	34	180	90	108	105	14	7	0	0	115	124	104	
South																											
43E138	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	1,2,3,7,25,Sd
171E138	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	1,2,3,7,9,25,Sd
41E136	–	–	–	+	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	1,2,3,25,Sd
232E137	–	+	–	–	–	–	–	–	–	–	+	–	–	1	3	–	7	–	–	+	–	–	–	–	–	–	2,3,4,9,25,Sd,Su
109E141	–	–	–	–	–	–	–	–	–	–	–	–	–	6	–	–	–	–	–	–	–	–	–	–	–	–	1,2,3,4,6,25,Sd,Su
169E136	–	–	–	–	–	–	–	–	–	–	–	–	–	–	3	–	–	–	–	+	–	–	–	–	–	–	1,2,3,9,25,Sd
45E140	–	–	–	–	–	–	+	–	–	–	–	–	–	4	3	–	–	–	–	–	–	–	–	–	–	–	1,2,3,6,25,Sd
233E137	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	1,2,3,4,9,25,Sd,Su
235E139	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	1,2,3,4,7,9,25,Sd,Su
237E141	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	1,2,3,4,6,9,25,Sd,Su
106E139	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	2,3,4,7,25,Sd,Su
169E136V17	–	–	–	–	–	–	–	–	–	–	–	–	–	–	5	–	–	3	–	–	–	–	–	–	–	–	1,2,3,9,17,Sd
233E137V17	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	50	–	–	+	–	–	–	–	–	–	1,2,3,4,9,17,25,Sd,Su
237E141V17	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	1,2,3,4,6,9,17,25,Sd,Su
173E140V17	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	1,2,3,6,9,17,25,Sd
239E143V17	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	1,2,3,4,6,7,9,17,25,Sd,Su
233E169V17	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	+	–	1,2,3,4,9,17,32,25,Sd,Su
104E41	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	+	+	3,4,6,32,25,Sd,Su
6E16	–	–	+	–	–	–	–	–	–	–	–	–	+	74	67	100	43	97	+	–	–	–	–	–	–	–	2,6,7,8
40E8	–	–	–	–	–	–	–	–	–	–	–	–	+	8	3	–	–	–	–	–	–	–	–	–	–	–	2,3,25,Sd
0E0	–	–	–	–	–	–	–	–	–	–	–	–	–	4	–	–	–	–	–	–	–	–	–	–	–	–	_
32E0	–	–	–	–	–	–	–	–	–	–	–	–	–	3	18	–	–	–	–	–	–	–	–	–	–	–	25,Sd
6E16V27	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	+	–	–	–	–	–	2,6,7,8,9,25,27,(Sd)
Isolate no.	0	1	6	2	0	0	1	0	0	0	1	0	9	72	39	20	14	33	1	3	4	0	0	0	5	1	

[a] Symbols: + indicates that the pathotype was present but the frequency was not calculated because the number of isolates sampled that year was <10 and – indicates that the pathotype was not detected.

[b] Pathotypes were coded according to Johnson et al. (34). The virulences and avirulences tested were *Vr1,2,3,4,6,7,8,9,10,17,25, 27,32,Sd,Sp,Su*.

[c] Virulence profile. Factors *Sd, Sp*, and *Su* referred to the resistance genes of the 'Strubes Dickkopf', 'Spalding Prolific'; and 'Suwon92' × 'Omar' cultivars, respectively. Sd infection types 5 to 6 were considered as intermediate reactions and were shown in parentheses.

low frequency in the north (Alsace and Pas de Calais regions) on Victo, which was rarely grown in the north. Interestingly, each new pathotype was detected following epidemic-free years, as in the case of *Vr6* (after the epidemic-free year of 1986), *Vr17* (after the epidemic-free years 1996 to 1997), and *Vr32* (after the epidemic-free years 2005 to 2006).

Virulence dynamics versus resistance gene distribution. The dynamics of *P. striiformis* f. sp. *tritici* virulence factors during the past 25 years have been strongly influenced by the relative distribution of resistance genes (Fig. 4). At the beginning of each epidemic, two to three new pathotypes have been detected carrying virulences that correlated with recent and widespread resistance genes, one of them subsequently becoming dominant. The frequency of this prevailing pathotype remained high for a few years and then declined, mainly in response to a reduction in the matching *Yr* gene (as in the cases of *Vr7* and *Vr6*; Fig. 4A and B). However, the virulence itself persisted for even longer in other pathotypes, even after the disappearance of the *Yr* genes (as in the case of *Vr9;* Fig. 4C) or reappeared in some new pathotypes with additional virulences to the newly deployed resistance genes associated with the target resistance gene (as in the case of *Vr17;* Fig. 4D). In the case of *Vr32,* the corresponding resistance gene *Yr32* was present up to the end of the study in the grown cultivars, and the *Vr32* was selected (Fig. 4E).

The presence of *Vr7* in *P. striiformis* f. sp. *tritici* isolates collected at the beginning of the study period was correlated with the relative distribution of a *Yr7*-based cultivar ('Talent') and, indeed, its frequency declined during subsequent years as the cultivation of Talent decreased (Fig. 4A). The 1988 epidemic was more severe. Samples from seven regions (Nord, Picardie, Pays de la Loire, Ile de France, Bourgogne, Rhône-Alpes, and Bretagne) showed that the epidemic was caused by the 45E140 pathotype first identified in 1987 (Fig. 4B). This pathotype possessed the virulence that matched *Yr6*, a resistance gene carried by the widely cultivated Récital and Austerlitz. Interestingly, *Vr6* first appeared when Récital was at the seed multiplication stage in nurseries, before its commercial cultivation. The frequency of the two pathotypes (45E140 and 109E141) declined gradually as the cultivation of *Yr6*-based cultivars decreased through 1992.

A pathotype analysis of the epidemics between 1990 and 1995 showed that the appearance of new pathotypes carrying *Vr9* was strongly correlated with the increase of cultivation of *Yr9*-based cultivars (mainly Slejpner) during this period. After 1996, the *Vr9*-containing pathotypes initially detected declined significantly, although the presence of *Vr9* remained constant in the population and was still being observed in other pathotypes during the post-1996 epidemics (Fig. 4C). The epidemics between 1998 and 2002 were caused by pathotypes containing *Vr17* and *Vr9* and some *Vr6*, not found in French pathotypes before 1997. The distribution of *Yr17* cultivars had been limited in France before reaching 10% of grown wheat area in 1998 and 12% in 1999 (8). Thus, the appearance of this virulence also corresponded with an increase in the production area of corresponding *Yr17*-carrying cultivars, mainly 'Audace', at the initiation of the epidemics (Fig. 4D). Although the initially detected *Vr17*-containing pathotypes decreased in frequency, some new pathotypes containing *Vr17* were identified during 2008 and 2009. *Vr17* was associated with virulence *Vr32*, which overcomes the resistance gene *Yr32*, present in Toisondor, cultivated in France since 2004 (Fig. 4E).

Regional structure of *P. striiformis* f. sp. *tritici* pathotypes. The relative production area planted with Récital (containing *Yr6*), Slejpner (containing *Yr9*), and Thésée (susceptible, containing neither *Yr6* nor *Yr9*) varied in two regions of northern and central France during the period 1987 to 1992 (Fig. 5). The relative area planted with Slejpner (*Yr9*) was high in the north whereas it was rarely cultivated in the central region. Récital, on the other hand, was grown in both regions and was more widely cultivated in the central region than in the northern region. The frequency of iso-

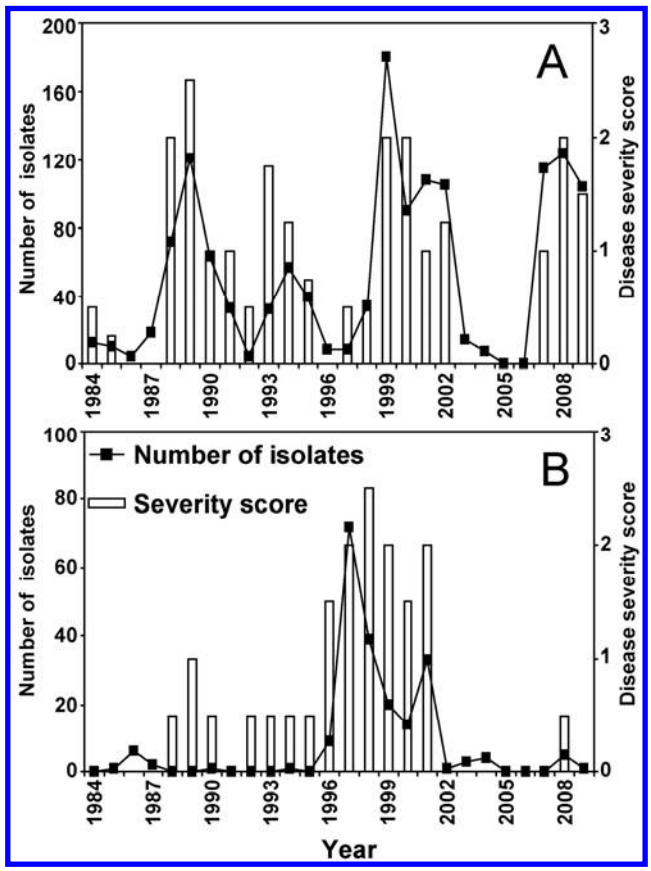


Fig. 2. Number of *Puccinia striiformis* f. sp. *tritici* isolates collected (black squares) over 25 years in **A,** northern and **B,** southern France. Disease severity score (0 to 3) is shown on a secondary axis.

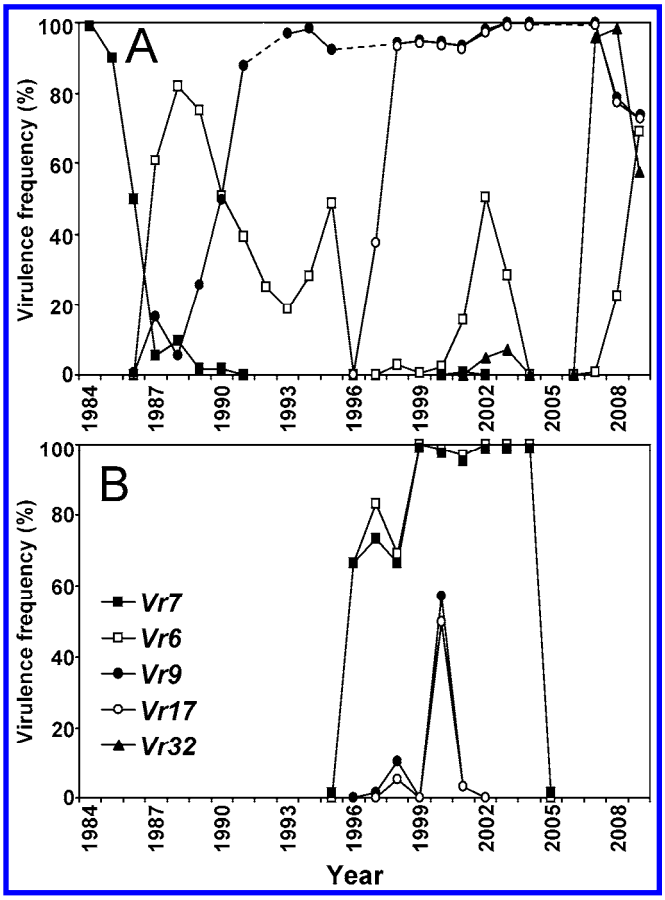


Fig. 3. Virulence (*Vr*) frequencies of *Puccinia striiformis* f. sp. *tritici* isolates collected each year between 1984 and 2009 in **A,** northern and **B,** southern France.

lates with *Vr6* or *Vr9* detected on these three cultivars varied in the two regions during the period. *Vr6* was detected in both regions from 1988 to 1991, and no epidemic was detected in 1992 due to unfavorable climatic conditions (*data not shown*). The *Vr9* was more frequent in the northern region than in the central region. A shift of virulence frequency was detected in the northern region, where Slejpner (*Yr9*) was cultivated, but in the central region, where Récital (*Yr6*) cultivation was higher than in the northern region, *Vr6* remained dominant. The regional distribution of the cultivars determined the regional structure of the *P. striiformis* f. sp. *tritici* population.

Detection of sources of durable resistance. This survey demonstrated the presence of some sources of durable resistance in 'Camp Rémy', 'Renan', 'Soissons' and 'Apache', which were found to be resistant for at least 10 years despite their cultivation on a large acreage (Table 2). Soissons and Apache were successively the most popular cultivars in France for 10 years. The resistance of Camp Rémy and Apache was expressed from the seedling stage while that of Soissons and Renan was expressed at the adult-plant stage. *Yr* postulation at the seedling stage of Apache revealed the presence of *Yr7* and *Yr17*, whereas no pathotype was widespread in France with combined virulences to both *Yr7* and *Yr17*.

Discussion

A clearer understanding of the appearance of pathotypes with new virulences, and of their subsequent disappearance in line with the cultivation of corresponding resistance genes, and of the existence of host selection pressure at a regional level may be helpful in crop disease management. This survey has described the temporal dynamics of different virulence genes, explained by the distribution of resistance genes. The severity of stripe rust epidemics varied annually as a function of the deployment of resistant cultivars. Epidemic severity was consistently different between the north and the south of France over the study period, with epidemics being more frequent and severe in the former. This may have been due to more favorable environmental conditions for sporadic

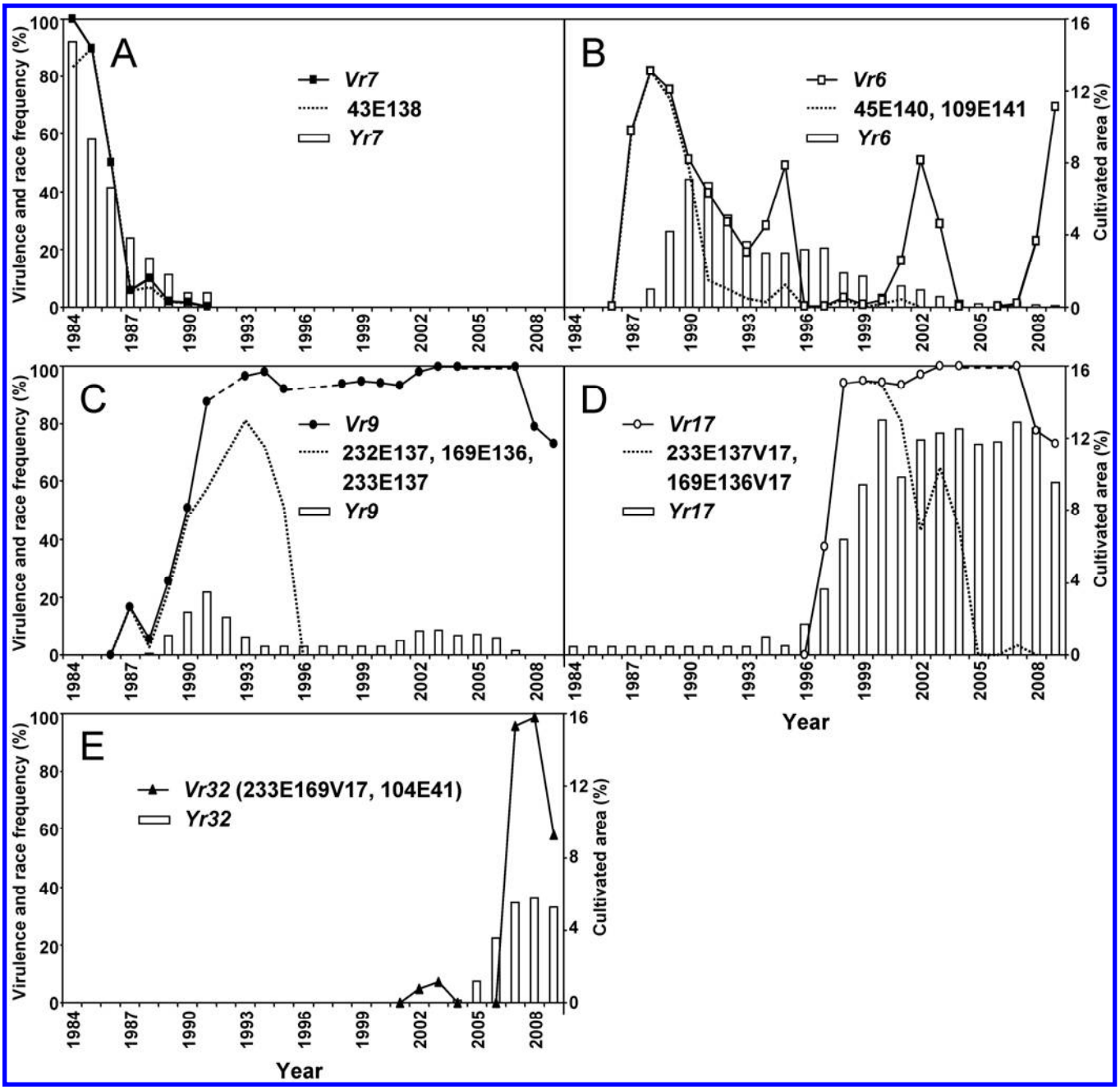


Fig. 4. Virulence frequencies of *Puccinia striiformis* f. sp. *tritici* **A,** *Vr7*; **B,** *Vr6*; **C,** *Vr9*; **D,** *Vr17*; and **E,** *Vr32* (continuous lines) and the relative cultivated area of cultivars with the corresponding *Yr* gene (as bars) in northern France over the 25-year period. Dotted line represents the frequency of the races first detected with this virulence.

disease development; more often in the north than in the south, where temperatures are generally higher and rainfall lower as required for *P. striiformis* f. sp. *tritici* (27). Additionally, the dominance of pathotype 6E16 in southern France coincided with the cultivation of the highly susceptible Victo, which bears no known *Yr* genes and was widely cultivated (5 to 6% of total wheat production area in southern France during 1997 and 1998) because of its resistance to leaf rust, the prevailing type of wheat rust in southern France (27). The pathotype 6E16 was characterized as being exotic because of a virulence spectrum that markedly diverged from that of northern pathotypes (avirulent on *YrSd* and virulent on *Yr7* and *Yr8*). This pathotype was also found in a sample from the durum wheat 'Regal', and its appearance seemed to be constant and characteristic of southern *P. striiformis* f. sp. *tritici* epidemics. Using molecular markers, the two populations were shown to be highly divergent at a genotypic level (25), with southern isolates belonging to a Mediterranean *P. striiformis* f. sp. *tritici* genetic group (6) and northern isolates belonging to the northwestern European genetic group (25). This divergence may have resulted from both the cultivation of distinct cultivars and the adaptation of the pathogen to different climatic conditions (39).

The years that were epidemic-free were 1986, 1992, 1996, 2004, and 2006. These could be explained by environmental variables, and especially winter temperatures. One important measure was the cardinal temperatures (i.e., 0°C for infection) (49), which reduced the viability of the inoculum. In 1986, 11% more days with minimum temperature below 0°C than the average over the 25-year period in the north were recorded. Similarly, the absence of epidemics in 1992, in spite of the presence of both a susceptible host (*Yr9*) and of the corresponding virulence (*Vr9*), could be due to the extremely low winter temperatures in December 1991 and February 1992 (with 20% more days than average with minimum temperature below 0°C). In 1996 and 2006, the percentages were 54 and 63%, respectively, while, in 2004, a dry spring (12% relative humidity below the average) may account for the absence of an epidemic. On the other hand, in years with severe epidemics, such as 1999, where there were 11% fewer days than average with minimum temperature below 0°C, temperature was mild and relative humidity was high, conditions conducive to the disease (*data not shown*). In Europe, clonal survival in the absence of wheat during the summer can be explained by the presence of volunteers that were responsible for a "green bridge" in humid coastal regions and that constituted a source of primary inoculum for future epidemics (1). This may explain the absence of epidemics between 2003 and 2005, when severe drought conditions in the summer of 2003 could have caused the destruction of over-summering *P. striiformis* f. sp. *tritici* spores. The mean temperature in the north from June to September 2003 was the highest of the 25 years under study, 11% higher than the average temperature over the period. Overall, climatic factors played an important role in the survival and size of

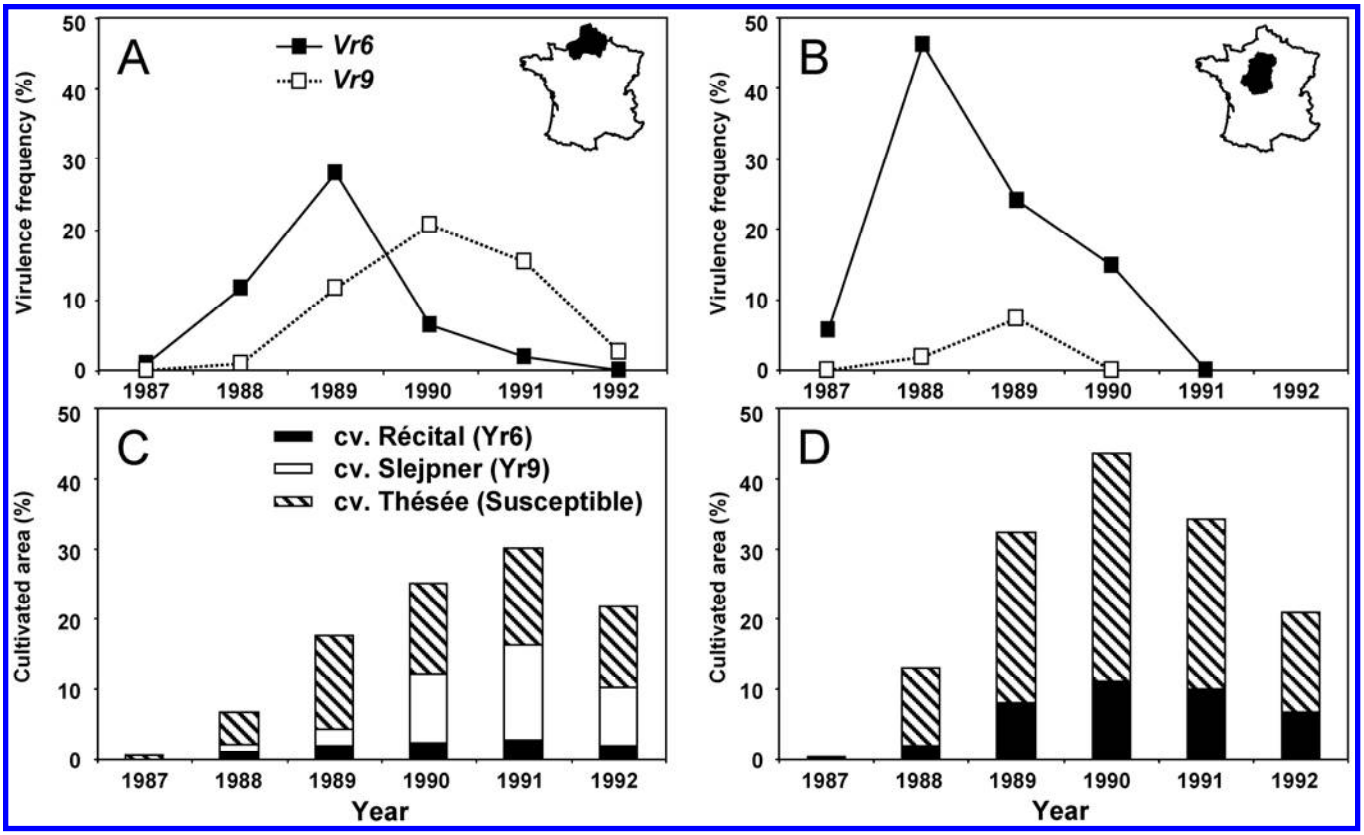


Fig. 5. Regional distribution of *Puccinia striiformis* f. sp. *tritici Vr6* and *Vr9* virulence frequency in two regions of France (**A,** north and **B,** center) and the proportionate distribution of 'Récital' (*Yr6*), 'Slejpner' (*Yr9*), and 'Thésée' (susceptible) in **C,** north and **D,** center of France. The two regions are indicated on the two maps of France (A and B).

Table 2. Cultivars with durable resistance according to the *Puccinia striiformis* f. sp. *tritici* survey performed during the period 1984 to 2009 in France

Cultivar	Registration year	Duration of cultivation on major area	Maximum annual cultivar area (%)	Seedling resistance genes	Adult-plant resistance genes	Reference
Camp Rémy	1980	1981–present	≈11	*QYr.inra-2BL*	*QYr.inra-2DS, QYr.inra-5BL.2*	37
Renan	1989	1990–present	≈5	*Yr17*	*QYr.inra-2BS, QYr.inra-3BS, QYr.inra-6B*	17,22
Soissons	1988	1989–present	≈40	*Yr3*	Undetermined	22
Apache	1998	1999–present	≈24	*Yr7, Yr17*	*QYr.inra-2AS2*	*Unpublished data*

the pathogen population, an issue outside of the scope of the current study.

The overall virulence dynamics and succession of different pathotypes in northern France were in accordance with the deployment of host resistance genes. Each epidemic in this region was characterized by the breakdown of a newly deployed or redeployed resistance gene. New pathotypes were detected in the years following the absence of epidemics, as in the case of *Vr6* in 1987 (no epidemics in 1986), *Vr17* in 1998 (no epidemics in 1996 to 1997), and *Vr32* in 2007 (no epidemics in 2005 to 2006). A similar case was reported in Denmark (28,29) where, following the absence of epidemics in 1996, pathotypes with a new virulence (*Vr17*) were detected and shown to be migrants from the United Kingdom. In the case of *Vr9*, however, the epidemics in northern France that occurred in 1989 followed the *Vr6* epidemics without a break. Furthermore, pathotypes containing *Vr9* remained at the same frequencies despite an epidemic-free year (1992). We also observed the acquisition of new virulences in the French *P. striiformis* f. sp. *tritici* population, well described as a clonal population (25). This showed the adaptation of a clonal pathogen population to a host through mutation (as has been mentioned in the Australian *P. striiformis* f. sp. *tritici* population; 51) or migration (e.g., northwestern-European populations; 29). This followed the arms race between the host and the pathogen in nature (16,38) but with human intervention where the cultivation on large areas of a resistant cultivar led to a "boom and bust cycle" (54). This resulted from the selection of a new mutant or migrant individual carrying the virulence to the newly deployed cultivar with resistance to the prevalent pathogen population. The frequency of this pathotype would increase in the overall population, discouraging cultivation of the cultivar in question. This was the case of the breakdown of *Yr9* and *Yr17* in northwestern Europe (8) and of *Yr10* in the United States (13). The frequency of the pathotype would decrease with the deployment of a new cultivar resistant to this pathotype if there is a fitness cost. This is repeated for new pathotypes and resistance genes.

In the case of northern France, the evolution of pathotypes was based on their chronological appearance and virulence profiles, as well as on the deployment of *Yr* genes (Fig. 6A) and the phylogenetic position of the pathotypes previously studied (25). This evolution followed a stepwise mutation model in a clonal population, as initially suggested for the evolution of pathotypes in Australia (55). Given that the most probable origin of the pathotypes in France was migration (11), the proposed scheme needs to be considered at the level of northwestern Europe. Virulence to a newly deployed *Yr* gene can be acquired by any existing pathotype clade but only one or two pathotypes were frequently found. As mentioned earlier, the epidemics in France were due to the successive breakdown of *Yr7*, *Yr6*, *Yr9*, *Yr17*, and *Yr32* resistance genes. When new *Yr* genes were deployed, pathotypes generally tended to evolve over time toward complexity and the acquisition of new virulences. However, in the present case, this adaptive process was not achieved through successive accumulation of new virulences in previously dominant pathotypes but through a more stochastic emergence of virulence in existing clonal lineages. The introduction of cultivars with two resistance genes (i.e., *Yr6+Yr9* and *Yr6+Yr17*) led to the rapid selection of pathotypes with the two corresponding virulences (Fig. 6A). Thus, these new French pathotypes probably originated through migration rather than mutation in the existing population, because most of them displayed a virulence profile that differed for more than two virulences from previously prevailing pathotypes. In most cases, these pathotypes had already been described in other northwestern European countries such as the United Kingdom (9) and Denmark (28) (http://www.eurowheat.org). Hovmøller et al. (29) described the case of *Vr17* epidemics at the European level, where pathotypes migrated in line with the cultivation of *Yr17*-based cultivars from the United Kingdom to Denmark and then on to Germany and France (8). Migration played a major role in the *P. striiformis* f. sp. *tritici* structure, with the northern French populations belonging to a clonal northwestern European population (25,29) while

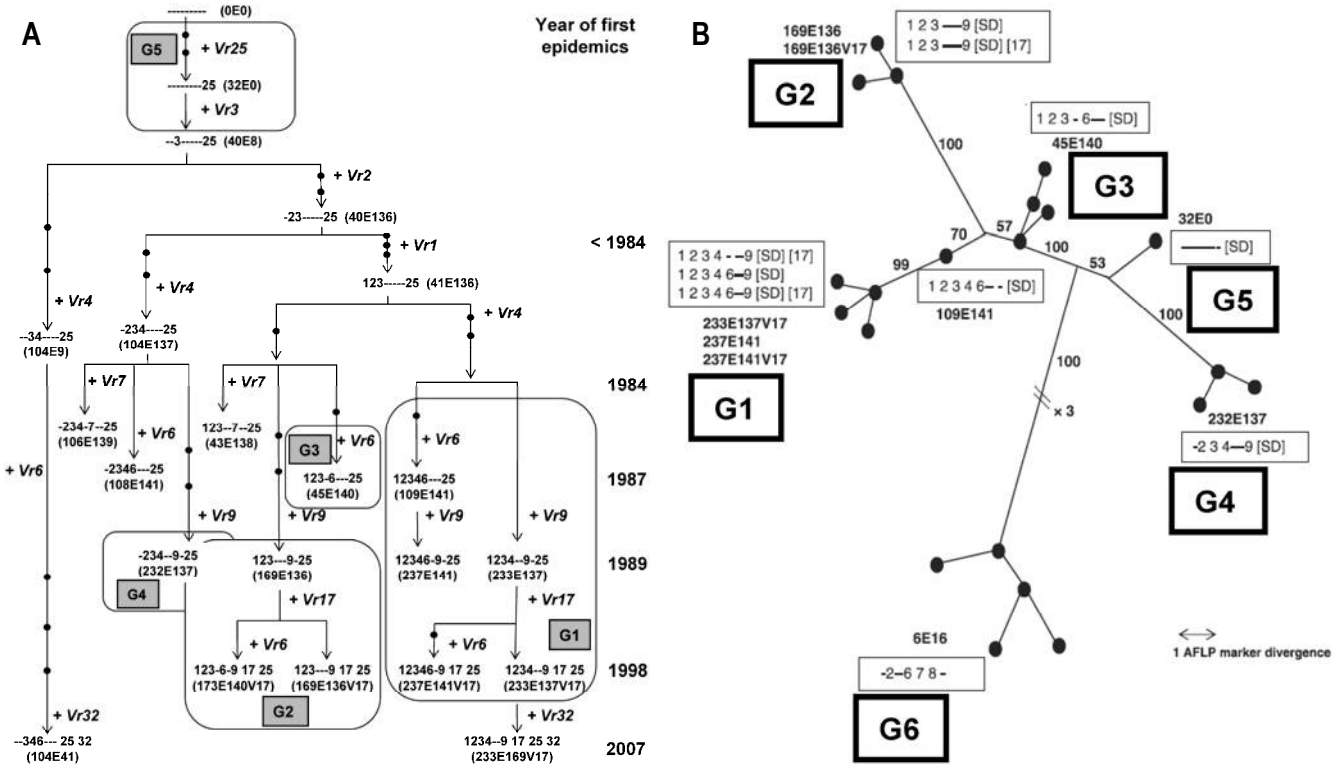


Fig. 6. A, Hypothetical scheme for the pathotype evolution of *Puccinia striiformis* f. sp. *tritici* populations in northern France, based on virulence profile, chronological appearance, and pathotype prevalence. Virulence profiles included virulences *Vr1, Vr2, Vr3, Vr4, Vr6, Vr7, Vr8, Vr9, Vr17, Vr25,* and *Vr32.* Pathotypes are indicated in brackets. Genetic groups (G1 to G6) corresponded to B. Points indicate the number of amplified fragment length polymorphism (AFLP) markers that differed between the strains (25). **B,** Phylogenetic tree constructed with AFLP genetic data from Enjalbert et al. (25) used to compare genetic subgroups (G1 to G6). SD (Strubes Dickkof) virulence was shown in brackets.

southern French populations belonged to a clonal Mediterranean population (6).

The phylogenetic tree (25) supported our hypothesis by grouping ancestral and descendant pathotypes within closed genetic groups (Fig. 6A). For example, we proposed that 109E141 isolates gave rise to 237E141 while placing 233E137V17 on another closed phylogenetic branch that gave rise to 237E141V17 and 233E137V17. This scheme took into account the acquisition of the same virulence independently in different genetic groups (i.e., *Vr17* in G2 for 169E136V17 and in G1 for 233E137V17, the latter being the most frequent pathotype with *Vr17*).

Although the frequency of a pathotype decreased in line with reduction in the cultivation of corresponding resistance gene-containing cultivars, virulence itself either disappeared or remained as an unnecessary virulence in other pathotypes carrying additional virulences. The loss of utility of host-specific resistance due to selection of the corresponding virulence was a major reason to question the use of such resistance genes. In our case, either the frequency of a critical virulence decreased due to a reduction of that specific *Yr* gene (e.g., *Vr7* and *Vr6*; Fig. 4A and B) or the critical virulence remained present in new pathotypes with additional virulence to the newly deployed resistance gene (as in the case of *Vr9* and *Vr17*; Fig. 4C and D). This could be explained by the concept of the cost of virulence, the trade-off between the ability to establish an infection (i.e., virulence) and other characteristics of the parasite (e.g., within-host growth or the production of transmission stages; 10). Such virulence trade-offs penalized the pathogen in the absence of corresponding resistance genes (5). Because of this cost, all unnecessary virulence factors tend to be eliminated from a population, and pathotypes with virulences that matched the particular resistance of a host can be expected to be favored (54). Such a virulence cost was exemplified in the evolution of both *Vr7* and *Vr6*. The latter has already been shown to be coupled with a substantial fitness cost (5). Virulences that did not generate costs could be expected to evolve as neutral traits in the absence of corresponding resistance genes, or may even tend to become fixed if the corresponding resistance gene is still cultivated to a lesser extent. This may be true for *Vr9* and *Vr17* in our case. *Vr9* has been suggested to have compensated for its initial cost before becoming neutral (5). This seemed logical when considering the fact that all *Vr17*-harboring pathotypes carried *Vr9*. However, such frequency evolution must be analyzed with caution in a clonal population, because it can also be due to hitchhiking effects; a set of virulences or avirulences can increase in frequency because they are associated (in one or more pathotypes) to another resistance gene submitted to positive selection (29).

We reported here on the existence of a regional structure of *P. striiformis* f. sp. *tritici* populations in terms of *Vr6* and *Vr9* virulence frequency in line with the distribution of cultivars bearing *Yr6* and *Yr9* resistance genes. Although only a limited number of isolates were available, preventing an exhaustive survey, we were still able to determine a regional structure for *P. striiformis* f. sp. *tritici*, despite their ability for long-distance migration (11). The regional management of cultivars with different resistance genes could be a key component in landscape disease management because it can prevent the global spread of a major pathotype by reducing the genetic uniformity of the host. Strategies to increase the diversity of disease resistance genes could prevent the widespread prevalence of one or more major pathotypes (46,57). Previously, the utilization of cultivar mixtures has been suggested as a means to increase intrafield diversity (18,46). However, it has rarely been adopted on a large scale in intensive crop production because it leads to some complications arising from mixing cultivars with different cropping requirements. The concept of landscape management may be useful in this respect, because individual fields can be managed in a conventional manner while cultivars can be chosen to maximize the interfield diversity.

Both temporal and regional approaches should be adopted regarding host management in order to achieve durable rust resistance. Temporal alteration of cultivars containing different resistance genes may be useful to prevent the development of complex pathotypes. At the regional level, the cultivation of cultivars with different resistance genes would prevent the selection and uniform spread of a specific pathotype. On a larger scale, this could help to prevent the rapid spread of new pathotypes, as was observed for the highly aggressive, temperature-adapted, and closely related strains (44) that were first detected in the United States (45) and that spread rapidly to Australia (55) and Europe (31). However, these two strains have not been detected yet in northern France because of a lack of certain virulences (*Vr3, Vr4, Vr17*, and *Vr25*), against which the corresponding *Yr* genes are frequently present in grown cultivars (22). Similarly, at a continental level, the differential deployment of *Yr* genes has resulted in pathotypes with different virulences; for example, the presence of *Vr9* in Europe (21) and Asia (50) and its absence from the United States (21), and vice versa for *Vr10* (21). Thus, variability in terms of resistance genes at the regional, national, and continental levels, coupled with an alternation of cultivars over time and the use of partial resistance (issues addressed elsewhere; 2,3,48,50), could help to avoid successions of virulences and their fixation and, thus, achieve sustainable management of resistance. Similarly, the 25-year survey demonstrated the presence of sources of durable resistance in four cultivars. Those cultivars had quantitative trait loci (QTLs) of resistance expressed at the adult-plant stage, and QTLs of resistance have now been selected from Camp Rémy (37) and Renan (17) and could be used in breeding programs to increase the durability of resistance to stripe rust. We suggested using adult-plant resistance in addition to the major resistance genes in a same-host genotype, as observed for the cultivars with durable resistance, and diversifying the sources of resistance both temporally and regionally. The scale of landscape diversification could be the region, because we showed that the regional distribution of resistant cultivars affected the pathotype frequency.

Acknowledgments

Part of this research was carried out under the European COST 8.17 project (Agriculture and Biotechnology) between 1995 and 1999. This work received support from the European Integrated Project BIOEXPLOIT, FOOD-CT-2005-513959 (2006 to 2011), and from the ENDURE project, Eurowheat (2008 to 2010), under the 6th Framework Programme. S. Ali was supported by a grant from the Higher Education Commission, the Government of Pakistan. We thank D. Caron, C. Maumené, and their colleagues from ARVALIS Institut du Végétal; our colleagues from Service Régional de la Protection des Végétaux; all the breeders who participated in the sampling of *P. striiformis* f. sp. *tritici* isolates; members of the COST 8.17 stripe rust subgroup (R. Bayles, R. Johnson [deceased], and M. Hovmøller) for their contribution; E. A. Milus for his constructive comments for the improvement of the manuscript; and our technical staff, notably L. Gérard and N. Galet.

Literature Cited

1. Ali, S., Leconte, M., Walker, A.-S., Enjalbert, J., and de Vallavieille-Pope, C. 2010. Reduction in the sex ability of worldwide clonal populations of *Puccinia striiformis* f. sp. *tritici*. Fungal Genet. Biol. 47:828-838.
2. Ali, S., Shah, S. J. A., Khalil, I. H., Rahman, H., Maqbool, K., and Ullah, W. 2009. Partial resistance to yellow rust in introduced winter wheat germplasm at the north of Pakistan. Aust. J. Crop Sci. 3:37-43.
3. Ali, S., Shah, S. J. A., and Rahman, H. 2009. Multi-locations variability in Pakistan for partial resistance in wheat to *Puccinia striiformis* West. *tritici*. Phytopathol. Mediterr. 48:269-279.
4. Andrivon, D., and de Vallavieille-Pope, C. 1995. Race diversity and complexity in selected populations of fungal biotrophic pathogens of cereals. Phytopathology 85:897-905.
5. Bahri, B., Kaltz, O., Leconte, M., de Vallavieille-Pope, C., and Enjalbert, J. 2009. Tracking costs of virulence in natural populations of the wheat pathogen, *Puccinia striiformis* f. sp. *tritici*. Online. BMC Evol. Biol. Online publication. doi:10.1186/1471-2148-1189-1126.
6. Bahri, B., Leconte, M., Ouffroukh, A., de Vallavieille-Pope, C., and Enjalbert, J. 2009. Geographic limits of a clonal population of wheat yellow rust in the Mediterranean region. Mol. Ecol. 18:4165-4179.
7. Bayles, R. A., Channell, M. H., and Stigwood, P. L. 1989. Yellow rust of wheat. Pages 11-17 in: U.K. Cereal Pathogen Virulence Survey, 1987. Annual Report, U.K. Cereal Pathogen Virulence Survey Committee, Cambridge.
8. Bayles, R. A., Flath, K., Hovmøller, M. S., and de Vallavieille-Pope, C. 2000. Breakdown of the *Yr17* resistance to yellow rust of wheat in northern

Europe—a case study by the yellow rust sub-group of COST 817. Agronomie (Paris) 20:805-811.
9. Bayles, R. A., and Stigwood, P. L. 1999. Yellow rust of wheat. Pages 27-42 in: U.K. Cereal Pathogen Virulence Survey, 1998. Annual Report, U.K. Cereal Pathogen Virulence Survey Committee, Cambridge.
10. Brown, J. K. M. 2003. Little else but parasites. Science 299:1680-1681.
11. Brown, J. K. M., and Hovmøller, M. S. 2002. Aerial dispersal of fungi on the global and continental scales and its consequences for plant disease. Science 297:537-541.
12. Chen, X., Line, R. F., and Leung, H. 1993. Relationship between virulence variation and DNA polymorphism in *Puccinia striiformis*. Phytopathology 83:1489-1497.
13. Chen, X. M., Moore, M., Milus, E. A., Long, D. L., Line, R. F., Marshall, D., and Jackson, L. 2002. Wheat stripe rust epidemics and races of *Puccinia striiformis* f. sp. *tritici* in the United States in 2000. Plant Dis. 86:39-46.
14. Chisholm, S. T., Coaker, G., Day, B., and Staskawicz, B. J. 2006. Host-microbe interactions: shaping the evolution of the plant immune response. Cell 124:803-814.
15. Dangl, J. L., and Jones, J. D. 2001. Plant pathogens and integrated defence responses to infection. Nature 411:826-833.
16. Dawkins, R., and Krebs, J. R. 1979. Arm races between and within species. Proc. R. Soc. Lond. B Biol. Sci. 205:489-511.
17. Dedryver, F., Paillard, S., Mallard, S., Robert, O., Trottet, M., Nègre, S., Verplancke, G., and Jahier, J. 2009. Characterization of genetic components involved in durable resistance to stripe rust in the bread wheat 'Renan'. Phytopathology 99:968-973.
18. de Vallavieille-Pope, C. 2004. Management of disease resistance diversity of cultivars of a species in single fields: controlling epidemics. C. R. Biol. 327:611-620.
19. de Vallavieille-Pope, C., Huber, L., Leconte, M., and Bethenod, O. 2002. Preinoculation effects of light quantity on infection efficiency of *Puccinia striiformis* and *P. triticina* on wheat seedlings. Phytopathology 92:1308-1314.
20. de Vallavieille-Pope, C., Huber, L., Leconte, M., and Goyeau, H. 1995. Comparative effects of temperature and interrupted wet periods on germination, penetration, and infection of *Puccinia recondita* f. sp. *tritici* and *P. striiformis* on wheat seedlings. Phytopathology 85:409-415.
21. de Vallavieille-Pope, C., and Line, R. F. 1990. Virulence of North American and European races of *Puccinia striiformis* on North American, world and European differentials. Plant Dis. 74:739-743.
22. de Vallavieille-Pope, C., Picard-Formery, H., Radulovic, S., and Johnson, R. 1990. Specific resistance factors to yellow rust in seedlings of some French wheat varieties and races of *Puccinia striiformis* Westend. in France. Agronomie (Paris) 10:103-113.
23. de Vallavieille-Pope, C., Rouzet, J., Leconte, M., Delos, M., and Mistou, M. N. 2000. La rouille jaune du blé en France: des épidémies déclenchées par une nouvelle race, un hiver doux et un printemps humide. Phytoma 527:22-29.
24. Duan, X., Tellier, A., Wan, A., Leconte, M., de Vallavieille-Pope, C., and Enjalbert, J. 2010. *Puccinia striiformis* f. sp. *tritici* presents high diversity and recombination in the over-summering zone of Gansu-China. Mycologia 102:44-53.
25. Enjalbert, J., Duan, X., Leconte, M., Hovmøller, M. S., and de Vallavieille-Pope, C. 2005. Genetic evidence of local adaptation of wheat yellow rust (*Puccinia striiformis* f. sp. *tritici*) within France. Mol. Ecol. 14:2065-2073.
26. Flor, H. H. 1971. The current status of the gene-for-gene concept. Annu. Rev. Phytopathol. 9:275-296.
27. Hau, B., and de Vallavieille-Pope, C. 2006. Wind-dispersed diseases. Pages 387-416 in: The Epidemiology of Plant Diseases, 2nd ed. B. M. Cooke, D. Gareth Jones, and B. Kaye, eds. Springer, Dordrecht, The Netherlands.
28. Hovmøller, M. S. 2001. Disease severity and pathotype dynamics of *Puccinia striiformis* f. sp. *tritici* in Denmark. Plant Pathol. 50:181-189.
29. Hovmøller, M. S., Justesen, A. F., and Brown, J. K. M. 2002. Clonality and long-distance migration of *Puccinia striiformis* f. sp. *tritici* in north-west Europe. Plant Pathol. 51:24-32.
30. Hovmøller, M. S., Walter, S., Justesen, A. F., 2010. Escalating threat of wheat rusts. Science 329:369.
31. Hovmøller, M. S., Yahyaoui, A. H., Milus, E. A., and Justesen, A. F. 2008. Rapid global spread of two aggressive strains of a wheat rust fungus. Mol. Ecol. 17:3818-3826.
32. Huang, R., Kranz, J., and Welz, H. G. 1994. Selection of pathotypes of *Erysiphe graminis* f. sp. *hordei* in pure and mixed stands of spring barley. Plant Pathol. 43:458-470.
33. Johnson, R. 1992. Past, present and future opportunities in breeding for disease resistance, with examples from wheat. Euphytica 63:3-22.
34. Johnson, R., Stubbs, R. W., Fuchs, E., and Chamberlain, N. H. 1972. Nomenclature for physiologic races of *Puccinia striiformis* infecting wheat. Trans. Br. Mycol. Soc. 58:475-480.
35. Line, R. F. 2002. Stripe rust of wheat and barley in North America: a retrospective historical review. Annu. Rev. Phytopathol. 40:75-118.
36. Line, R. F., and Qayoum, A. 1991. Virulence, aggressiveness, evolution, and distribution of races of *Puccinia striiformis* (the cause of stripe rust of wheat) in North America, 1968-87. U. S. Dep. Agric. Tech. Bull. 1788.
37. Mallard, S., Gaudet, D., Aldeia, A., Abelard, C., Besnard, A. L., Sourdille, P., and Dedryver, F. 2005. Genetic analysis of durable resistance to yellow rust in bread wheat. Theor. Appl. Genet. 110:1401-1409.
38. Maor, R., and Shirasu, K. 2005. The arms race continues: battle strategies between plants and fungal pathogens. Curr. Opin. Microbiol. 8:399-404.
39. Mboup, M., Bahri, B., Leconte, M., de Vallavieille-Pope, C., Katz, O., and Enjalbert, J. Genetic structure and local adaptation of European wheat yellow rust populations: The role of temperature-specific adaptation. Evolutionary Applications. In press.
40. Mboup, M., Leconte, M., Gautier, A., Wan, A. M., Chen, W., de Vallavieille-Pope, C., and Enjalbert, J. 2009. Evidence of genetic recombination in wheat yellow rust populations of a Chinese oversummering area. Fungal Genet. Biol. 46:299-307.
41. McDonald, B. A., and Linde, C. 2002. Pathogen population genetics, evolutionary potential, and durable resistance. Annu. Rev. Phytopathol. 40:349-379.
42. McIntosh, R. A., Dubcovsky, J., Rogers, W. J., Morris, C., Appels, R., and Xia, X. C. 2010. Catalogue of gene symbols. KOMUGI Integrated Wheat Science Database. http://www.shigen.nig.ac.jp/wheat/komugi/genes/symbolClassList.jsp;jsessionid=689B192F53CED7CB561A9DC624FB9518.lb1,%202010
43. McNeal, F. H., Konzak, C. F., Smith, E. P., Tate, W. S., and Russel, T. S. 1971. A uniform system for recording and processing cereal research data. Pages 34-121 in: U. S. Agric. Res. Serv. Bull. Washington, DC.
44. Milus, E. A., Kristensen, K., and Hovmøller, M. S. 2009. Evidence for increased aggressiveness in a recent widespread strain of *Puccinia striiformis* f. sp. *tritici* causing stripe rust of wheat. Phytopathology 99:89-94.
45. Milus, E. A., Seyran, E., and McNew, R. 2006. Aggressiveness of *Puccinia striiformis* f. sp. *tritici* isolates in the south-Central United States. Plant Dis. 90:847-852.
46. Mundt, C. C., and Browning, J. A. 1985. Genetic diversity and cereal rust management. Pages 527-560 in: The Cereal Rusts. A. P. Roelfs and W. R. Bushnell, eds. Academic Press, Orlando, FL.
47. Office National Interprofessionnel des Grandes Cultures. 2008. Répartition des variétés—récoltes 1999-2008 (blé tendre, orge, blé dur, triticale). Technical Report: Les Cahiers de l'ONIGC. France Agrimer. Paris.
48. Pathan, A. K., and Park, R. F. 2007. Evaluation of seedling and adult plant resistance to stem rust in European wheat cultivars. Euphytica 155:87-105.
49. Rapilly, F. 1979. Yellow rust epidemiology. Annu. Rev. Phytopathol. 17:59-73.
50. Singh, R. P., William, H. M., Huerta-Espino, J., and Rosewarne, G. 2004. Wheat rust in Asia: meeting the challenges with old and new technologies. In: New Directions for a Diverse Planet: 4th Int. Crop Sci. Congr. Brisbane, Australia.
51. Steele, K. A., Humphreys, E., Wellings, C. R., and Dickinson, M. J. 2001. Support for a stepwise mutation model for pathogen evolution in Australasian *Puccinia striiformis* f. sp. *tritici* by use of molecular markers. Plant Pathol. 50:174-180.
52. Stubbs, R. W. 1985. Stripe rust. Pages 61-101 in: The Cereal Rusts, Diseases, Distribution, Epidemiology and Control. A. P. Roelfs and W. R. Bushnell, eds. Academic Press, London.
53. USDA. 2010. Catalog of Rust Resistance Genes in Small Grains http://www.ars.usda.gov/Main/docs.htm?docid=10342.
54. Van der Plank, J. E. 1968. Disease Resistance in Plants. Academic Press, New York and London.
55. Wellings, C. R. 2007. *Puccinia striiformis* in Australia: a review of the incursion, evolution and adaptation of stripe rust in the period 1979-2006. Aust. J. Agric. Res. 58:567-575.
56. Wellings, C. R., and McIntosh, R. A. 1990. *Puccinia striiformis* f. sp. *tritici* in Australasia: pathogenic changes during the first 10 years. Plant Pathol. 39:316-325.
57. Wolfe, M. S., and Barrett, J. A. 1980. Can we lead the pathogen astray. Plant Dis. 64:148-155.

BIOLOGIE DES POPULATIONS ET HISTOIRE DES INVASIONS DE *PUCCINIA STRIIFORMIS* F.SP. *TRITICI* A L'ECHELLE MONDIALE ET LOCALE

L'étude de la structure génétique des populations d'agents pathogènes à grandes échelles permet d'identifier des réservoirs de diversité et des routes de migration, de retracer les trajectoires évolutives et de repérer notamment des sources potentielles de nouvelles invasions. *Puccinia striiformis* f.sp. *tritici* (PST), responsable de la rouille jaune du blé, réparti mondialement, constitue un modèle fongique d'intérêt pour les études d'invasion étant donné sa capacité de migration sur de longue distance et l'apparition récurrente de nouvelles souches localement. Nous avons analysé la structure des populations de PST à l'échelle mondiale, à l'aide de marqueurs microsatellites sur un échantillon de 409 isolats issus des six continents (représentatif des collections INRA, France et AU, Danemark). Les génotypes ont été répartis en six groupes génétiques correspondant à leur origine géographique grâce à des méthodes de classification bayésienne et multivariée. Les analyses indiquent une forte hétérogénéité géographique de déséquilibre de liaison et diversité génotypique, avec des signatures de recombinaison dans les régions de l'Himalaya (Népal et Pakistan) et à proximité en Chine. La structure reste clonale pour les populations des autres régions. L'assignation des isolats aux différents groupes génétiques a permis de déterminer l'origine des invasions (récentes ou anciennes). Ainsi, les souches agressives adaptées à de hautes températures, répandues de par le monde depuis 2000, sont originaires de Mer rouge-Moyen Orient ; les isolats d'Amérique du Nord et du Sud et d'Australie proviennent d'Europe du Nord-Ouest. Par ailleurs, les isolats d'Afrique du Sud appartiennent au groupe génétique de la zone méditerranéenne. La subdivision marquée entre les différentes zones géographiques indique qu'elles ne sont pas fortement marquées par les migrations récentes. De plus, les voies de migration identifiées attestent de l'importance des activités humaines dans la dispersion de PST à longue distance. La biologie des populations des zones les plus diverses (Chine et Pakistan) a été finement étudiée à l'aide d'échantillonnages réalisés deux années consécutives. Une population échantillonnée en 2004 et 2005 dans la vallée de Tianshui, (province de Gansu, Chine), s'est révélée très diverse, fortement recombinante et non structurée spatialement et temporellement. L'observation de clones identiques entre les deux échantillons temporels a permis de développer un estimateur du taux de sexualité, i.e. du rôle relatif de la reproduction sexuée par rapport à celui de la reproduction asexuée dans le maintien de la population. Ce taux de reproduction sexuée est estimé à 74 %, alors que la taille efficace de la population est de 1735, ce qui donne les premières indications du rôle du cycle sexué pour cette espèce. L'échantillonnage réalisé au Nord du Pakistan a permis de décrire quatre groupes génétiques ayant tous une grande diversité génotypique et une structure recombinante. Le très faible taux de ré-échantillonnage de génotypes identiques au cours de deux années suggère le rôle prédominant de la reproduction sexuée dans le maintien temporel des populations locales. La forte diversité génétique et génotypique, la signature de recombinaison et la capacité à la reproduction sexuée de PST dans la région himalayenne suggèrent que cette zone est le centre d'origine potentielle de PST. Les analyses d'approximations bayésiennes confirment la thèse d'une dispersion à partir de l'Himalaya vers les autres régions du monde. Le Pakistan et la Chine seront les populations les plus ancestrales des populations testées. La variabilité pour la capacité à produire des téleutosores, spores indispensables à l'initiation de la phase sexuée, a été analysée (56 isolats mondiaux), et s'avère liée à la variabilité génotypique et au taux de recombinaison. Ce résultat conforte la thèse de l'apparition de la sexualité dans la zone himalayenne et à proximité de cette zone et de la perte de sexualité lors de migrations dans les zones où l'hôte alternant est absent et où le cycle épidémique est essentiellement asexué. La description de l'origine, des voies de migration de populations mondiales de PST ainsi que de son centre de diversité contribue à la compréhension du potentiel évolutif de PST et à la construction de stratégies de gestion de lutte contre l'agent pathogène.

POPULATION BIOLOGY AND INVASION HISTORY OF *PUCCINIA STRIIFORMIS* AT WORLDWIDE AND LOCAL SCALE

Analyses of the large-scale population structure of pathogens enable the identification of migration patterns, diversity reservoirs or longevity of populations, the understanding of current evolutionary trajectories and the anticipation of future ones. Once identified, a detailed analysis of populations in centre of diversity, with emphasis on demographic and reproductive parameters should enable to infer the adaptive capacity of the pathogen and identify potential sources for new invasions. *Puccinia striiformis* f.sp. *tritici* (PST) is the causal agent of wheat yellow/stripe rust, and despite a worldwide distribution, this fungus remains a model species for invasion studies, due to its long-distance migration capacity and recurrent local emergence of new strains. Little is known about the ancestral relationship of the worldwide PST population, while its center of origin is still unknown. We used multilocus microsatellite genotyping to infer the worldwide population structure of PST and the origin of new invasions, analyzing a set of isolates representative of sampling performed over six continents. Bayesian and multivariate clustering methods partitioned the isolates into six distinct genetic groups, corresponding to distinct geographic areas. The assignment analysis confirmed the Middle East-Red Sea Area as the most likely source of newly spreading, high-temperature-adapted strains; Europe as the source of South American, North American and Australian populations; and Mediterranean-Central Asian populations as the origin of South African populations. The existence of strong population subdivision at worldwide level shows that major genetic groups are not markedly affected by recent dispersal events. However, the sources for recent invasions and the migration routes identified emphasize the importance of human activities on the recent long-distance spread of the disease. The analyses of linkage disequilibrium and genotypic diversity indicated a strong regional heterogeneity in levels of recombination, with clear signatures of recombination in the Himalayan (Nepal and Pakistan) and near-Himalayan (China) regions and a predominant clonal population structure in other regions. To explain the variability in diversity and recombination of worldwide PST populations, we assessed their sex ability in terms of telial production, the sex-specific structures that are obligatory for PST sexual cycle, in a set of 56 isolates representative of these worldwide geographical origins. We confirmed that the variability in genotypic diversity/ recombination was linked with the sex ability, pinpointing the Himalayan region as the possible center of origin of PST, from where it then spread worldwide. The reduced sex ability in clonal populations certainly reflects a loss of sexual function, associated to migration in areas where sexual alternate host is lacking, or not necessary for the completion of epidemic cycle. Approximate Bayesian computation analyses confirmed an out of Himalaya spread of PST, with Pakistan and China being the most ancestral populations. An in-depth study was made in the Himalayan region (Pakistan) and Near Himalayan region (Gansu, China) to better describe the population maintenance in two areas of high diversity. Analyses of Pakistani population at regional level revealed the existence of a strong population subdivision, a high genotypic diversity and the existence of recombination signature at each location reflecting the role of sexual recombination in the temporal maintenance at local level. A time spaced sampling of PST in the valley of Tianshui (China) inspired the development of a new estimator, allowing to quantify the relative contribution of sexual reproduction and effective population size on the basis of clonal resampling within and between years. A sexual reproduction rate of 74% (95% confidence interval [CI]: 38-95%) and effective population size of 1735 (95% CI: 675-2800) was quantified in Chinese PST population. The description of the origin and migration routes of PST populations worldwide and at its centre of diversity contributes to our understanding of PST evolutionary potential, and is helpful to build disease management strategies.