

RESEARCH

Open Access



Molecular evidences for population differentiation and the migration from south to north of *Puccinia triticina* in eastern China

Hongfu Li^{1,2}, Qinqin Zhang^{1,2}, Gui Wang¹, Jifeng Wang¹, Zhiyong Chen¹, Wuchao Zhao^{1,2,3}, Xinyue Zheng^{1,2}, Li Gao^{1,2}, Bo Liu¹, Lijian Xu³, Wanquan Chen^{1,2*} and Taiguo Liu^{1,2,3,4*} 

Abstract

Wheat leaf rust is caused by *Puccinia triticina* (*Pt*), leading to serious wheat yield loss in the world. To study the population structure and reveal the transmission routes of *Pt* in eastern China, leaf samples were collected from the main wheat-producing areas from April to June 2020. Total of 372 *Pt* strains were amplified by 13 SSR makers and a high level of genetic diversity was revealed with 289 multi-locus genotypes (MLG) identified. STRUCTURE analysis suggests that all *Pt* strains were assigned to 3 clusters, and 11 populations were further defined by considering geographic locations. All 55 pairwise populations had number of migration (Nm) values > 1 , indicating moderate genetic differentiation and frequent exchanges among populations. The genetic structure was significant different among populations in the northern and southern regions bounded by the Qinling Mountains-Huaihe River line. *Pt* strains in the southern regions, such as Jiangsu, Anhui and Zhejiang provinces, had higher level of genetic diversity and genetic variation, and Jiangsu might play an important role in the epidemic and population structure of *Pt*. Both genetic communication and horizontal wind field analyses showed that *Pt* had higher level of gene flow from the southern to northern regions than that of the reverse direction. The demonstrated genetic structure and dispersal route of Chinese eastern *Pt* populations would provide valuable information for epidemiological studies and disease control.

Keywords *Puccinia triticina*, Population genetic structure, Genetic diversity, Gene flow

Background

Wheat leaf rust caused by *Puccinia triticina* (*Pt*) is one of the most common diseases on wheat worldwide and causing the largest loss of global wheat yield (Savary et al. 2019). China is the world's largest wheat producer, with production of about 134 million tons in 2020, accounting for 17.6% of the world's wheat production (Food and Agriculture Organization of the United Nations 2022). However, wheat leaf rust occurs annually on about 15 mha, with approximately 3 million tons of wheat production loss (Huerta-Espino et al. 2011). Wheat leaf rust affects almost all wheat production areas in China, and usually occurs in spring and summer. It occurs in late April in the central and southwest of China, and then eastern regions, southeast regions from May to June, and

*Correspondence:

Wanquan Chen
wqchen@ippcaas.cn

Taiguo Liu
liutaiguo@caas.cn

¹ State Key Laboratory for the Biology of Plant Diseases and Insect Pests, Institute of Plant Protection, Chinese Academy of Agricultural Science, Beijing 100193, China

² National Agricultural Experimental Station for Plant Protection, Gangu, Ministry of Agriculture and Rural Affairs, Gansu 741200, China

³ College of Advanced Agriculture and Ecological Environment, Heilongjiang University, Harbin 150080, China

⁴ Control of Biological Hazard Factors (Plant Origin) for Agri-Product Quality and Safety, Ministry of Agriculture and Rural Affairs, Beijing 100193, China



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

occurs in northern regions from June to July (Liu et al. 1989; Huang et al. 2005; Liang 2019).

Pt is a macrocyclic and heteroecious rust fungus with five spore stages, and the urediniospores play a major role in the epidemics in China. *Pt* does not kill its host immediately after infection, and urediniospores can survive on the leaf surface for a long period in the wheat growing season. After an outbreak in one area, urediniospores spread to the surrounding areas and cause reinfections, and various strains form heterogeneous populations that differ in their virulence and parasitic fitness (Prasad et al. 2020). Geographical barriers, such as oceans and mountains, hinder the dispersal of pathogens. For example, the Tianshan Mountains are regarded as the barrier that impedes the movement of *Pt* strains from Kazakhstan to Uzbekistan, Kyrgyzstan, and Tajikistan in Central Asia (Kolmer and Ordoñez 2007). Kolmer (1998) first identified new pathotypes of *Pt* in North America and Canada in 1996. And isolates with the same virulence and SSR genotypes were then detected in South America (Ordoñez et al. 2010) and Europe (Kolmer et al. 2013). However, these strains were not found in China (Kolmer 2015). China is an independent epidemic zone of *Pt*, and the virulence structure of *Pt* in China is different from other countries or regions of the world (Liu and Chen 2012). The variable factors, such as mountains, plateaus, basins, crisscrossing rivers, climate differences caused by the wide range of landscape, cultivars differences and growth periods in different wheat productions affect the population formation of *Pt*, which is shown in different virulence frequency among different provinces (Ge et al. 2015).

Wheat leaf rust is particularly frequent in eastern China (Peng and Yu 2015; Huang et al. 2022), where the top five wheat production areas, consisting of Henan, Shandong, Anhui, Hebei, and Jiangsu and accounting for 80.2% of the national output were located (National Bureau of Statistics, <https://data.stats.gov.cn/adv.htm?m=advquery&cn=E0103>). The continuous planting of winter wheat provided adequate hosts for *Pt* infection. Airflow is regarded as the main driving force for the long-distance transmission of urediniospores of *Pt* (Li and Zeng 2002) and is one of the reasons for the prevalence of wheat leaf rust (Zhao et al. 2016). Continuous plains in eastern China facilitate the rapid spread of *Pt* strains by the wind. The airflow study may help to understand the driving force of *Pt* migration. For instance, Li et al. (2021) combined a genetic structure analysis with airflow analysis and illustrated the important role of Gansu and Yunnan regions in *Puccinia striiformis* f. sp. *tritici* (*Pst*) overwintering. Wind field analysis has been used to predict the long-distance migration of rust spores (Pan

et al. 2006), and the relationship among geographical populations of fungi (Cooke et al. 2006).

Although some studies have analyzed the genetic structure of *Pt* in different regions of China, few have included the Jiangsu and Zhejiang regions, and studies on the driving force of airflow for *Pt* isolates have not been conducted (Ma et al. 2020). Nevertheless, our previous studies have detected a high level of virulence polymorphism of Jiangsu *Pt* population, that may play an important role in the epidemic of wheat leaf rust in eastern China (unpublished data). Studies of the genetic structure of *Pt* from the population perspective may improve our understanding of the main characteristics of *Pt* populations via comparisons of the composition and distribution of populations from different regions, and facilitate analyses of the nature and the spread of this disease.

In this study, leaf samples of *Pt* were collected from the Hebei, Henan, Shandong, Anhui, Jiangsu, Zhejiang, and Beijing regions from April to June 2020, and the genotypes were identified using 16 SSR primers to clarify the genetic structure of *Pt* populations in eastern China. Moreover, a horizontal wind field analysis was performed to elucidate the driving force for the migration of *Pt* urediniospores and to predict the possible migration directions of *Pt* strains in eastern China. Thus, providing reliable reference for wheat leaf rust prediction and control.

Results

Population structure analysis

The 16 SSR markers showed the number of alleles (N_A) from 3 to 11 with the mean value of 5.69, and the mean value of unbiased expected heterozygosity (uHe) was 0.448 ± 0.014 , suggesting that the 16 markers were appropriate for genetic analyses in this study (Table 1). Genotype accumulation curve indicated that it is sufficient to discriminate unique individuals for the given loci (Additional file 1: Figure S1).

The samples were assigned into different clusters by STRUCTURE software, and $K=3$ was the best K value for this set of data (Additional file 1: Figure S2). Samples from different clusters within the same province were divided into 11 populations for further analysis (Table 2). Cluster 1 contained H1_HB-S, H3_BJ, H4_HN, H5_SD, and H8_JS-S; Cluster 2 contained H2_HB-N, H6_JS-N, and H11_ZJ-N2; Cluster 3 contained H7_JS, H9_AH, and H10_ZJ-N1. DAPC analysis of 11 populations showed that the populations were divided into 3 clusters same as the results of STRUCTURE software (Additional file 1: Figure S3). Separate DAPC and clustering analyses were performed for each cluster to examine the rationalization of the division of populations (Additional file 1:

Table 1 Sequence information, number of alleles (N_A), observed heterozygosity (H_O), expected heterozygosity (H_E), and unbiased expected heterozygosity (uHe) of 16 codominant simple sequence repeat (SSR) markers employed in this study

Locus	EMBL	Repeat motif	Primer sequence (5'-3')	Ta (°C)	N_A (size: bp)	H_O	H_E	uHe	Reference
RB17	AJ508886	(TGC) ₅ +(TGG) ₆	F: CTTCGGTAG GATTTTCGAGCG R: CAGCTCCAA ATCCTTTGCC	58	6 (192–200)	0.067 ± 0.040	0.127 ± 0.061	0.129 ± 0.062	Duan et al. (2003)
RB28	AJ508890	(TGG) ₅	F: CATCTGGCT GGTGAGGTC GC R: GAAGCCCCGC CGAGCAGC	58	8 (319–335)	0 ± 0.008	0.278 ± 0.045	0.282 ± 0.046	Duan et al. (2003)
RB4	AJ508880	(GT) ₈	F: CAGTATTGT GGTGGTTGG ATG R: ACTCAAGAA TAATGGGGA ACAC	60	6 (247–256)	1.000 ± 0.016	0.5 ± 0.017	0.508 ± 0.017	Duan et al. (2003)
RB12	AJ508884	(AG) ₅₊₃	F: CCACAAGCA ACACATAC CACC R: TGGTCCATG AAGAAGTCT CTGAAC	58	11 (266–300)	0.800 ± 0.049	0.491 ± 0.027	0.499 ± 0.028	Duan et al. (2003)
RB8	AJ508881	(TGG) ₇	F: CGCCGTTCC CATCGTTC R: TAAAACACT CCACCCACG CC	60	5 (143–155)	0.267 ± 0.087	0.32 ± 0.062	0.325 ± 0.062	Duan et al. (2003)
RB1	AJ508879	(GT) ₅	F: TTGTCGTTT TGGAATGATGC R: TGCCACAAA CCCTCCTC	58	6 (127–139)	0.833 ± 0.046	0.486 ± 0.008	0.494 ± 0.008	Duan et al. (2003)
RB35	AJ508893	(AC) ₉ +(TA) ₅ +(AG) ₅	F: ACTGCGATA TCCAGTACA CACAC R: TGATGGGCT CGCAGTGG	60	6 (245–250)	0.700 ± 0.062	0.654 ± 0.027	0.665 ± 0.027	Duan et al. (2003)
PtSSR151A	DQ789156	(AAC) ₁₂	F: TCATCGCAC TCCACTCAGAC R: ATGCTGCCC AACCTGCTC	61	4 (469–475)	0.667 ± 0.080	0.464 ± 0.031	0.472 ± 0.032	Szabo and Kolmer (2007)
PtSSR161	DQ789160	(TC) ₁₃	F: ACTGCCTCC TGTGCCTTCT R: TAGTCCGAG GGTGACGAA GT	60	6 (217–222)	0.900 ± 0.044	0.509 ± 0.021	0.518 ± 0.021	Szabo and Kolmer (2007)
RB11	AJ508883	(CA) ₁₇	F: AGCAGTGAG CAGCAGCGTC R: ACTACTGTG AGTGTGCGC TTGG	60	6 (176–202)	0.833 ± 0.035	0.486 ± 0.024	0.494 ± 0.024	Duan et al. (2003)
PtSSR152	DQ789157	(TG) ₇ +(TG) ₃ +(TG) ₃	F: CTCCGTTCC TCTTTCTGTCG R: CCATCGCAA CCAACAAACA	61	4 (380–393)	0.233 ± 0.071	0.255 ± 0.044	0.259 ± 0.044	Szabo and Kolmer (2007)

Table 1 (continued)

Locus	EMBL	Repeat motif	Primer sequence (5'-3')	Ta (°C)	N _A (size: bp)	H _O	H _E	uHe	Reference
PtSSR173	DQ789162	(TC) ₃ +(TC) ₁₂	F: CTCAGCGAC CTCAAAGAA CC R: GAGACG ACGGATGTC AACAA	60	6 (212–222)	0.533 ± 0.077	0.475 ± 0.045	0.483 ± 0.046	Szabo and Kolmer (2007)
PtSSR164	DQ789161	(TC) ₁₃	F: GTGGAAGTG AGCGGAAGA AG R: GGAGAT GGCAGATG AGGTA	61	5 (218–228)	0.567 ± 0.068	0.459 ± 0.025	0.467 ± 0.026	Szabo and Kolmer (2007)
PtSSR55	DQ789150	(TC) ₁₀	F: AGCTTACGG TCCTCAATCG R: AGTGAAAGG GGCTGGGAGT	61	3 (304–308)	0.567 ± 0.064	0.406 ± 0.024	0.413 ± 0.024	Szabo and Kolmer (2007)
PtSSR13	DQ789148	(TC) ₉	F: CGAATTCGC GTTTTATGTCC R: TGATCCAAT CGAACCTAG CC	61	5 (125–130)	0.567 ± 0.092	0.663 ± 0.042	0.674 ± 0.042	Szabo and Kolmer (2007)
RB16	AJ508885	(TGG) ₇ +(GT) ₄	F: CATTGAGC CACGGTTGA CTG R: AGACAT GGTCGCAGC CACTG	61	3 (281–284)	0 ± 0.051	0.278 ± 0.052	0.282 ± 0.053	Duan et al. (2003)
Mean					5.69	0.588 ± 0.026	0.440 ± 0.013	0.448 ± 0.014	

Figures S4, S5), and both of the results validated the population division.

For K=2, populations H1–H6 (located in the north) and H8 were clustered in the first cluster, while populations H7, and H9–H11 (located in the south) were clustered in the second (Fig. 1), which revealed a south-north differentiation of *Pt* population structure. Sampling sites for the 11 populations were shown in the geographic topographic map (Fig. 2a). For K=3 (Fig. 2b and Table 2), Cluster 1 was composed of samples from Hebei, Shandong, Henan, Jiangsu, and Beijing in the northern regions, accounting for 13.70–25.6%. Cluster 2 was mainly composed of samples from northern Jiangsu, Hebei and Zhejiang, accounting for 57.10, 23.80, and 10.5%, respectively. Cluster 3 was mainly composed of samples from Jiangsu, Zhejiang, and Anhui, accounting for 88.20%. Cluster 1 was composed of samples mainly located in the northern regions of eastern China, while Cluster 3 was located in the southern regions.

The population structure of *Pt* in the same province was also different between the northern and southern

regions. H1_HB-S of Cluster 1 consisted of samples mainly from Handan (11/30) and Xingtai (8/30), which was located in southern Hebei. H2_HB-N of Cluster 2 consisted of samples mainly from Tangshan (20/25), which was located in northeastern Hebei. H6_JS-N from northern Jiangsu was mainly assigned to samples from Cluster 2, while H8_JS-S in southern Jiangsu was mainly assigned to samples from Cluster 1 (Table 2).

Genetic diversity analysis

A total of 289 multi-locus genotypes (MLG) were identified from 372 samples, indicating a high level of genetic diversity (Table 3). For 11 populations of the smallest sample size ≥ 10 based on rarefaction, H4_HN showed the least genetic diversity, as it had the smallest values of the number of expected MLG (eMLG: 7.81), *Ne* (1.629 ± 0.129), *I* (0.521 ± 0.082), and *uHe* (0.330 ± 0.056) indexes. H7_JS showed the highest genetic diversity, for the largest values of indexes mentioned above. *Pt* in Jiangsu was a mixed population, and samples from Jiangsu were divided into 3 different populations due

Table 2 Composition of 11 populations combined with the structure and geographical location data

Cluster	Population	Province	County	No. from the county	No. from the province	Sampling time	
Cluster 1	H1_HB-S	Hebei	Handan	11	30	2020-05-28-2020-06-10	
			Xingtai	8			
			Tangshan	7			
			Hengshui	2			
			Baoding	2			
Cluster 2	H2_HB-N	Hebei	Tangshan	20	25	2020-06-01-2020-06-11	
			Qinhuangdao	2			
			Xingtai	2			
			Cangzhou	1			
Cluster 1	H3_BJ	Beijing	Beijing	19	19	2020-06-06-2020-06-08	
Cluster 1	H4_HN	Henan	Xinxiang	11	22	2020-05-01-2020-05-22	
			Zhoukou	6			
			Puyang	5			
Cluster 1	H5_SD	Shandong	Tai'an	26	43	2020-05-13-2020-05-23	
			Linyi	17			
Cluster 2	H6_JS-N	North Jiangsu	Lianyungang	51	59	2020-05-12-2020-05-20	
			Xuzhou	8			
Cluster 3	H7_JS	Jiangsu	Xuzhou	13	63	2020-05-12-2020-05-20	
			Lianyungang	4			
			Zhenjiang	16			2020-04-21-2020-04-24
			Jurong	11			
			Yangzhou	9			
			Yancheng	8			
Cluster 1	H8_JS-S	South Jiangsu	Yangzhou	15	20	2020-04-10	
			Zhenjiang	5			
Cluster 3	H9_AH	Anhui	Hefei	11	31	2020-04-02-2020-05-08	
			Wuhu	10			
			Huai'an	5			
			Bengbu	4			
			Tongling	1			
Cluster 3	H10_ZJ-N1	North Zhejiang	Jiaxing	32	48	2020-04-21	
			Hangzhou	16			
Cluster 2	H11_ZJ-N2	North Zhejiang	Hangzhou	12	12	2020-04-21	
	Total				372		

The 11 populations are intending to distribute from north to south by province basically

to their genetic differentiation. We combined the geographically close populations within the same province (combined H6, H7, and H8 from Jiangsu into one, and combined H10 and H11 from Zhejiang into one), and the Jiangsu population had the highest level of genetic diversity, with $uHe = 0.516 \pm 0.038$. Jiangsu, Anhui, and Zhejiang in the south all had $uHe > 0.50$, which was larger than that of Shandong (0.481 ± 0.032), Hebei (0.434 ± 0.05 and 0.435 ± 0.037), Beijing (0.426 ± 0.042), and Henan (0.33 ± 0.056) in the north.

The private allele is an allele that is present in only one of many populations sampled and is presumed to be responsible for adaptation to a stressful environment (Konecka et al. 2019). Private alleles appeared more frequently in the Jiangsu and Anhui regions. The combination population H6-H8 in Jiangsu contained the largest number of private alleles (0.813 ± 0.306), followed by H9 in Anhui (0.5 ± 0.258), suggesting that *Pt* strains with more genetic variation accumulated in Jiangsu and Anhui.

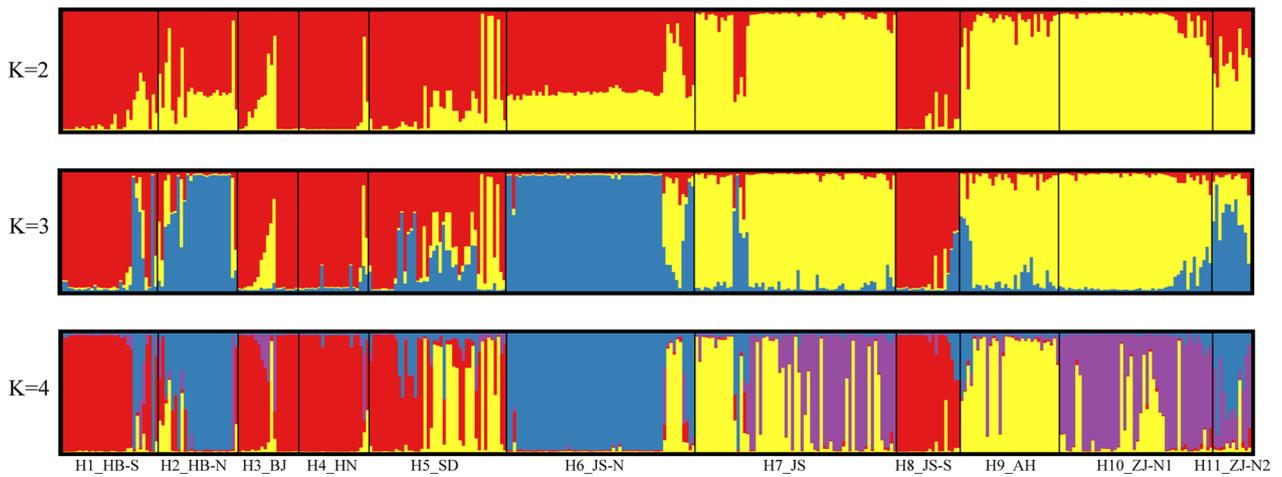


Fig. 1 The division of 11 populations based on different K values (from K = 2 to 4) in the STRUCTURE analysis, and the best value is K = 3, with red, blue, and yellow, representing Cluster 1, Cluster 2, and Cluster 3, respectively. 11 populations are distributed in order from south to north. Single samples are represented by vertical bars, and the colors and proportions of color indicate that samples originated from different clusters. Populations are separated by black vertical lines

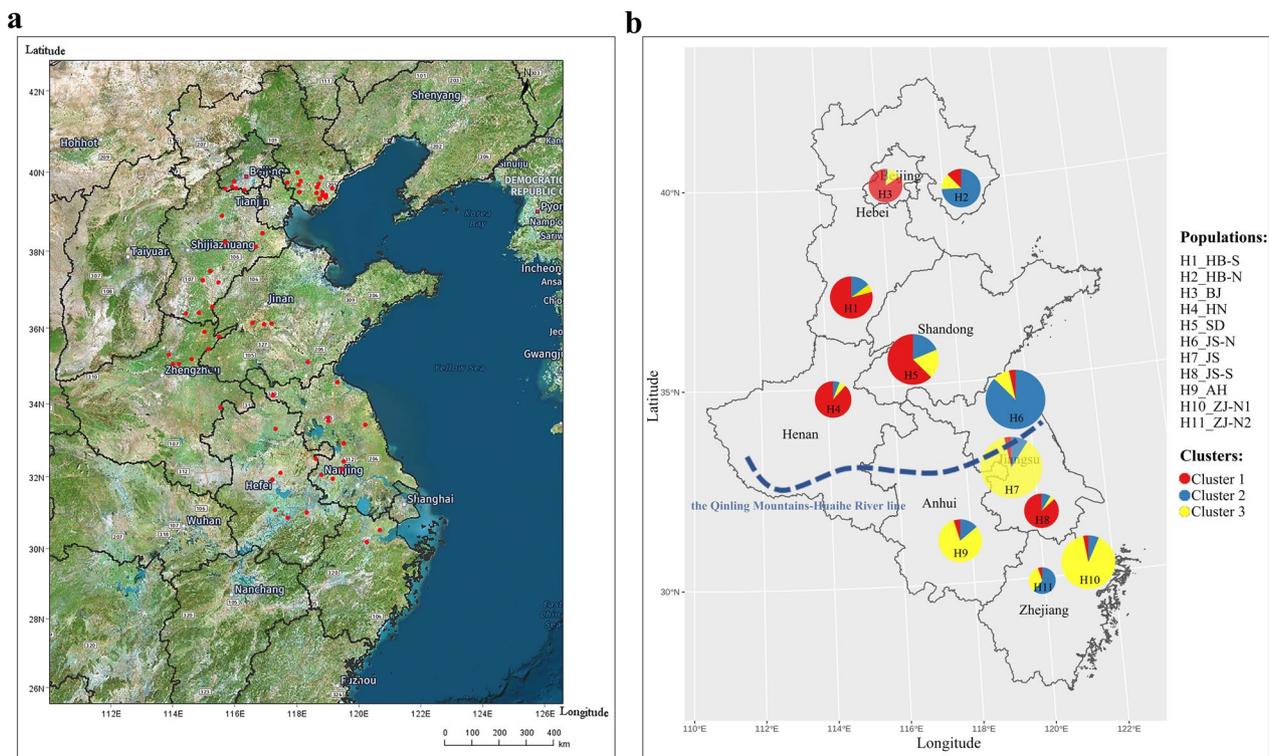


Fig. 2 Populations in eastern China from the Anhui, Beijing, Hebei, Henan, Jiangsu, Shandong, and Zhejiang provinces in 2020. **a** Geographic topographic map of sampling sites. **b** Pie chart of 11 populations labeled on the map. Population size is shown in the area of the circle, and population composition is shown with different sections in red (Cluster 1), blue (Cluster 2), and yellow (Cluster 3). The Qinling Mountains-Huaihe River line is shown in the blue line, which represents the geographical dividing line between north and south China

Population divergence and genetic communication

The majority of the variation in *Pt* populations was detected within populations (85.89%, $P < 0.001$) (Table 4).

F_{ST} and N_m were calculated between 11 pairwise populations. F_{ST} values between 0.15 and 0.25 indicated moderate differentiation, and 21/55 of the pairwise populations

Table 3 Analysis of the genetic diversity of 11 *Puccinia triticina* populations from 7 regions showed an overall high level of genetic diversity

Region	Cluster	Population	N	MLG	eMLG	Ne	I	He	uHe	No. of private alleles
Hebei	Cluster 1	H1_HB-S	30	25	11.05	1.861 ± 0.125	0.688 ± 0.065	0.428 ± 0.036	0.435 ± 0.037	0 ± 0
	Cluster 2	H2_HB-N	25	20	10.65	1.893 ± 0.124	0.680 ± 0.080	0.425 ± 0.049	0.434 ± 0.050	0 ± 0
Beijing	Cluster 1	H3_BJ	19	19	12	1.833 ± 0.131	0.694 ± 0.074	0.415 ± 0.041	0.426 ± 0.042	0.063 ± 0.063
Henan	Cluster 1	H4_HN	22	13	7.81	1.629 ± 0.129	0.521 ± 0.082	0.323 ± 0.054	0.330 ± 0.056	0 ± 0
Shandong	Cluster 1	H5_SD	43	33	11.1	1.995 ± 0.101	0.761 ± 0.057	0.476 ± 0.032	0.481 ± 0.032	0.188 ± 0.101
Jiangsu	Cluster 2	H6_JS-N	59	29	8.66	1.827 ± 0.118	0.615 ± 0.080	0.402 ± 0.052	0.406 ± 0.053	0.063 ± 0.063
	Cluster 3	H7_JS	63	63	12	2.442 ± 0.236	0.965 ± 0.090	0.532 ± 0.044	0.537 ± 0.044	0.625 ± 0.239
	Cluster 1	H8_JS-S	20	16	10.46	1.691 ± 0.087	0.611 ± 0.050	0.384 ± 0.033	0.394 ± 0.034	0 ± 0
		H6-H8	142	/	/	2.232 ± 0.156	0.917 ± 0.073	0.514 ± 0.038	0.516 ± 0.038	0.813 ± 0.306
	Anhui	Cluster 3	H9_AH	31	31	12	2.362 ± 0.247	0.899 ± 0.117	0.500 ± 0.056	0.508 ± 0.057
Zhejiang	Cluster 3	H10_ZJ-N1	48	47	11.94	2.190 ± 0.179	0.825 ± 0.088	0.491 ± 0.047	0.496 ± 0.047	0.125 ± 0.085
	Cluster 2	H11_ZJ-N2	12	12	12	1.889 ± 0.156	0.645 ± 0.081	0.410 ± 0.051	0.428 ± 0.054	0 ± 0
		H10-H11	60	/	/	2.201 ± 0.172	0.835 ± 0.084	0.498 ± 0.044	0.502 ± 0.045	0.125 ± 0.085
Total/Mean			372	289	11.68	1.980 ± 0.047	0.726 ± 0.024	0.440 ± 0.013	0.448 ± 0.014	/

Standard error (SE) values are listed after the ± symbol; the combination from Jiangsu (H6-H8) and Zhejiang (H10-H11) are shown in bold and not included again in the total

N, mean number of individuals observed; MLG, multi-locus genotypes observed; eMLG, number of expected MLG at the smallest sample size ≥ 10 based on rarefaction; I, mean Shannon information index; He, mean expected heterozygosity; uHe, mean unbiased expected heterozygosity; No. of private alleles, mean number of alleles unique to a single population

Table 4 Analysis of molecular variance (AMOVA) within and among populations of *Puccinia triticina*

Source of variation	Degrees of freedom	Sum of squares	Variance components	Percentage of variation (%)	P value
Among populations	10	431.597	0.59805	14.11	0.001
Within populations	733	2669.366	3.6417	85.89	0.001
Total	743	3100.964	4.23975	100	/

Table 5 F_{ST} values for pairwise populations revealed a moderate level of genetic differentiation

Populations	H1_HB-S	H2_HB-N	H3_BJ	H4_HN	H5_SD	H6_JS-N	H7_JS	H8_JS-S	H9_AH	H10_ZJ-N1	H11_ZJ-N2
H1_HB-S		0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
H2_HB-N	0.11691		0.001	0.001	0.001	0.018	0.001	0.001	0.001	0.001	0.001
H3_BJ	0.01617	0.15632		0.001	0.001	0.001	0.001	0.05	0.001	0.001	0.001
H4_HN	0.02407	0.17633	0.04083		0.001	0.001	0.001	0.018	0.001	0.001	0.001
H5_SD	0.07444	0.108	0.04983	0.12094		0.001	0.001	0.001	0.001	0.001	0.001
H6_JS-N	0.1837	0.00723	0.21614	0.23371	0.13982		0.001	0.001	0.001	0.001	0.001
H7_JS	0.17368	0.10311	0.1746	0.24723	0.12371	0.11274		0.001	0.001	0.001	0.001
H8_JS-S	0.01566	0.17762	0.01329	0.01128	0.09122	0.2309	0.21459		0.001	0.001	0.001
H9_AH	0.20931	0.10289	0.2085	0.27231	0.13874	0.09394	0.04751	0.23355		0.001	0.001
H10_ZJ-N1	0.19948	0.12787	0.21587	0.28631	0.1575	0.15698	0.03964	0.2633	0.08008		0.001
H11_ZJ-N2	0.13429	0.02165	0.17972	0.21498	0.13409	0.05675	0.09126	0.21023	0.11942	0.1012	

F_{ST} indicates the inbreeding number of populations relative to the total population, F_{ST} values > 0.25 indicate strong differentiation and are shown in bold. The P value for F_{ST} is shown above the diagonal

F_{ST} mean pairs of alleles between individuals within populations

had F_{ST} values ranging from 0.15 to 0.25 (Table 5). Moreover, the pairwise populations of H4 with H9, and H10, as well as H8 with H10, had F_{ST} values greater than 0.25, indicating strong population divergences.

Estimates of Nm reflect gene flow in pairwise populations caused by all mechanisms result in the movement of genes (Slatkin 1985). The Nm values of 55 pairwise populations were all greater than 1, indicating that genetic communication was common between *Pt* populations in eastern China (Fig. 3). H1, H2, H3, H4, and H5 had low levels of genetic exchange with H9, H10, and H11, and the former populations were located in the north, while the latter were located in the south. This might reflect that geographical distance affected the density of genetic communication. Nm values >4 indicate frequent migration between populations, and 15/55 of the pairwise populations had Nm values >4, indicating strong gene flows between these populations. Among these, genetic exchanges were even more frequent between H1 and H3, H1 and H8, H2 and H6, and H4 and H8, with Nm values >10. The 4 pairwise populations had F_{ST} values

between 0.007 and 0.016, indicating a low level of population divergence. Thus, frequent genetic communication of populations helped weaken population divergences, although they were far apart.

The relative migration network reflected the direction of asymmetric gene flows between pairwise populations (Fig. 4). H6, H7, and H8 from Jiangsu played important roles in gene communication, since H1 and H9 accepted gene flows from H6, and H5, H4, and H2 accepted gene flows from H8. Under the filter threshold of 0.1, H3, H10, and H11 did not have such obvious asymmetric gene flows with other populations. Remote gene flows were detected from H8 to H2, H4 to H2, and H9 to H1, which were from southern regions to northern regions.

Horizontal wind field analysis

Airflow potentially carries urediniospores for long distances, not only to different counties in the same region (Alam et al. 2021), but also to different provinces (Craigie 1945) or even farther (Eversmeyer and Kramer 2000) after some time. The horizontal wind field analysis

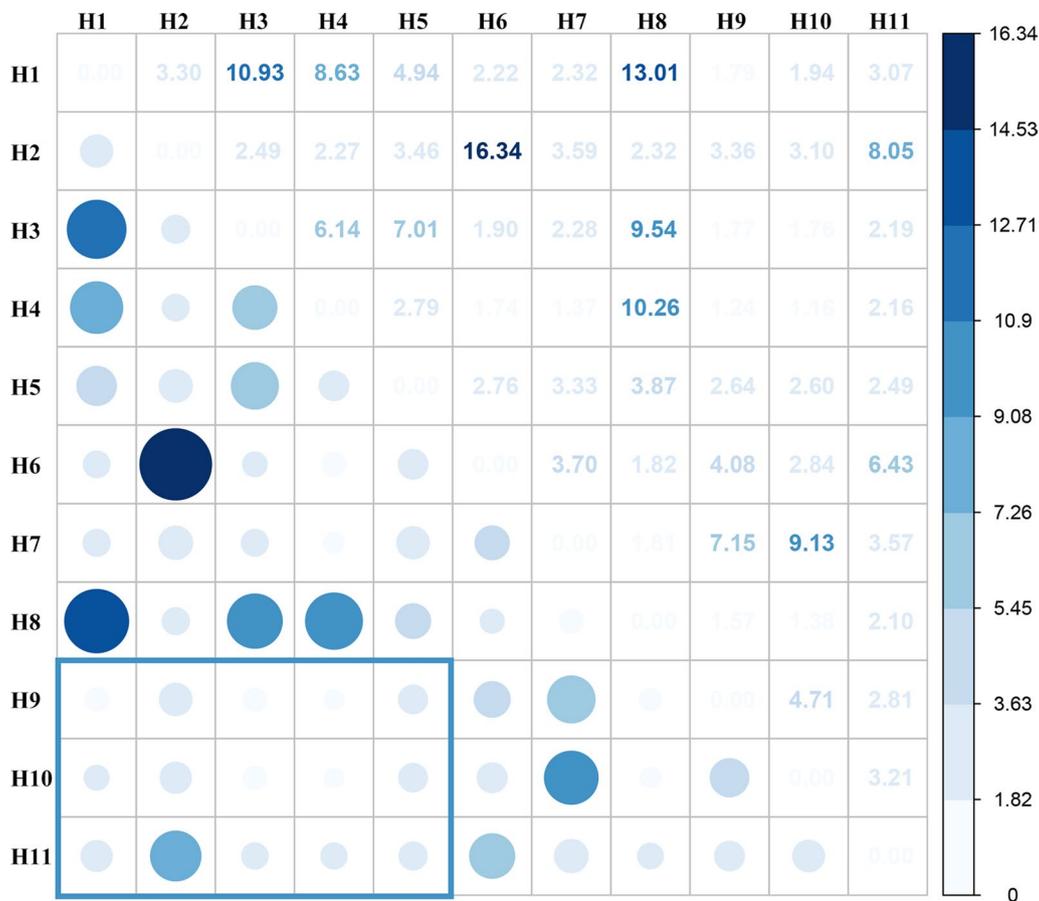


Fig. 3 Pairwise population Nm values indicate frequent migrations of *Pt* strains between geographically close populations. The shade of the color and size of the circle represents the Nm values. Populations in long distances have low Nm values in the bottom left box (H1–H5 with H9–H11)

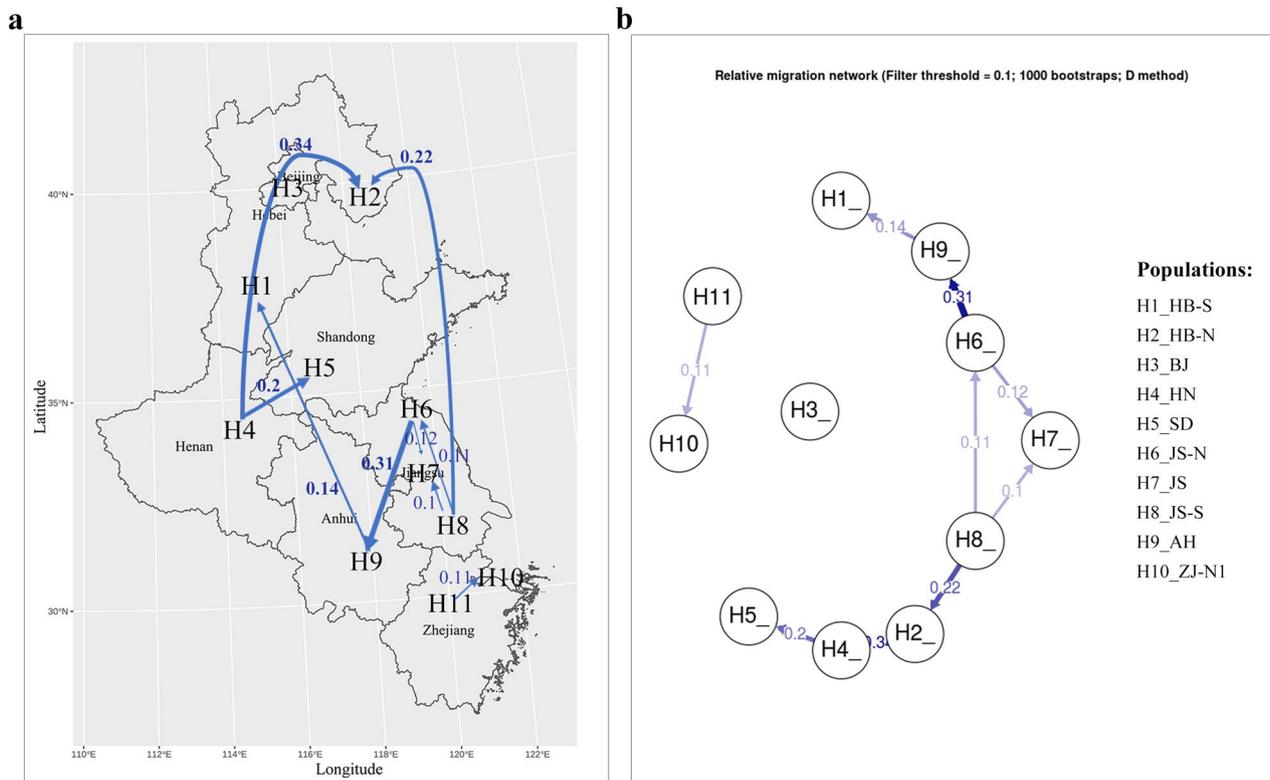


Fig. 4 Relative migration network of 11 populations reflecting the direction of gene flow in eastern China. **a** Abridged general view on the map. **b** Relative migration network with filter threshold = 0.1. The arrows show the direction of gene flow between pairwise populations, and the shades of color correspond to the values of relative gene flow

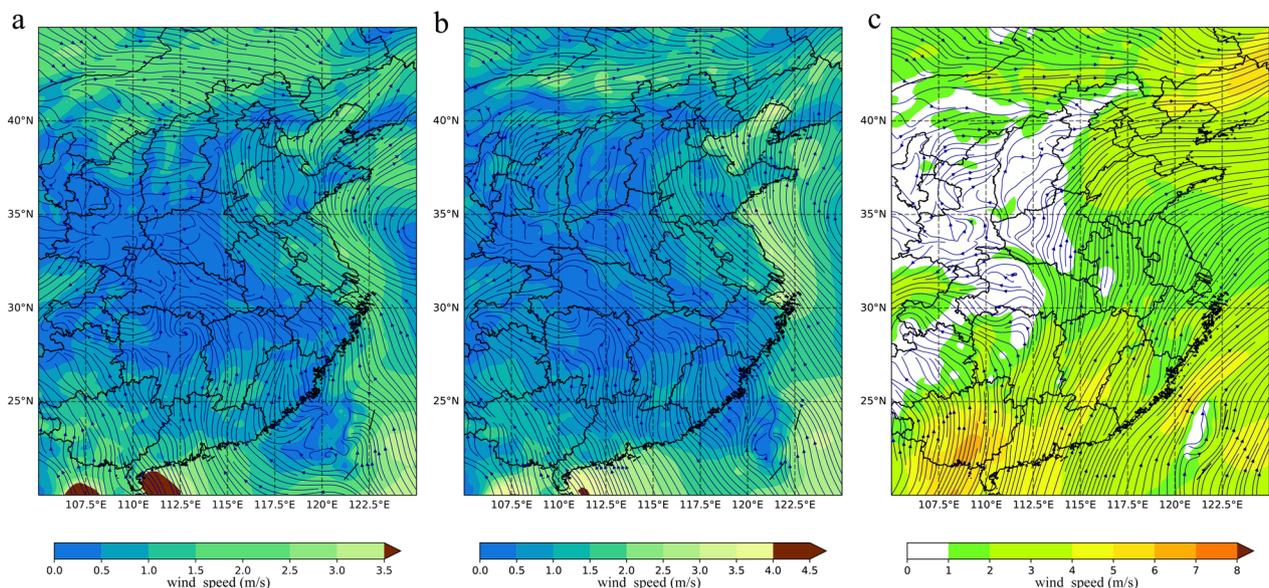


Fig. 5 Average horizontal wind field from April to June supports airflow driving force for *Pt* urediniospores from the south to north. Colored areas show different wind speeds, and the wind field streamlines and directions are shown with dark blue arrows. **a** Average horizontal wind field at 10 m altitude from April to June 2020. **b** Average horizontal wind field at 10 m altitude from April to June in 2016–2020. **c** Average horizontal wind field at 900 hPa from April to June in 2016–2020

(Fig. 5) revealed that horizontal wind streamlines generally followed the direction from south to north over eastern China in 2020 and 2016–2020 from April to June, not only at 10 m altitude but also at high altitudes of 900 hPa, which provided the driving force to carry urediniospores of *Pt* from south to north.

Discussion

Revealed by genetic structure analyses, there were significant differences in genetic structure between northern and southern populations of *Pt* bounded by the north–south division of the Qinling Mountains-Huaihe River line, and populations in Jiangsu, Anhui, and Zhejiang had high levels of genetic diversity and genetic variation. Genetic communication analyses showed that *Pt* had more gene flows from south to north, which was supported by horizontal wind field analysis.

The central–periphery hypothesis (CPH), also known as the central–marginal hypothesis, is a long-standing postulate (Brussard 1984). It states that genetic variation and demographic performance of a species decrease from the center to the edge of its geographic range (Eckert et al. 2008; Pironon et al. 2017). In the present study, the evidence suggests that *Pt* in southern regions migrated northward and affected the structure of northern populations. First, populations in Hebei, Beijing, and northern Henan had a lower level of genetic diversity, and according to the CPH, population with a low level of genetic diversity was regarded as the peripheral population; while Jiangsu, Anhui, and Zhejiang were regarded as the central populations due to their high level of genetic diversity. Shandong was seemingly in the middle of the central and peripheral regions because of its medium level of genetic diversity. Second, the results of the relative gene flow analysis also supported migration by revealing the direction of genetic migration from south to north. Third, the horizontal wind fields at 10 m height and 900 hPa altitude both showed that the conditions were favorable to transport urediniospores of *Pt* from southern to northern regions. In conclusion, *Pt* strains migrated from south to north in eastern China.

Mutation and natural selection are intrinsic and extrinsic driving forces affecting the virulence and genetic structure of *Pt* populations, respectively. Through complex driving forces, *Pt* strains acquire different evolutionary rates on genomes under different conditions. A study based on the same reference genome revealed that the average total SNPs in *Pt* strains growing on durum wheat (416,611–450,544 SNPs) vary from those growing on common wheat (310,033–422,333 SNPs), and the SNPs in different *Pt* strains growing on common wheat also vary substantially (Fellers et al. 2021). In eastern China, the main varieties, disease-resistance gene layout, and

growth period of wheat in different provinces differed. *Pt* populations in the southern and northern regions are exposed to different driving forces such as different climate conditions, wheat varieties, and intensities of disease control. *Pt* strains that migrate from south to north experienced different selective pressures, and *Pt* populations in different regions are evolving in different ways, resulting in genetic differentiation.

Our data suggest that the *Pt* isolates from Jiangsu were assigned into 3 clusters (Table 2), and the 3 populations play important roles in gene exchange (Fig. 5). Wheat production in Jiangsu is divided into the Huang-Huai winter wheat area (northern areas of the Huaihe River) and the middle and lower reaches of the Yangtze River area (southern areas of the Huaihe River). In the northern areas of Jiangsu, wheat is planted in early October, and cultivars are mostly winter wheat or weak winter wheat, with a growth period of about 230 days. And in the southern areas of Jiangsu, wheat is planted from late October to early November, and cultivars are usually weak winter wheat or spring wheat, with less growth period of about 200 days. Different conditions of wheat production affect the population of *Pt* in Jiangsu, and the complexity of population structure might imply the importance of Jiangsu in the virulence accumulation of epidemic in eastern China. The warm climate conditions in Jiangsu provide possibilities for urediniospores to overwinter in plants in local areas; thus, strains with variation might accumulate continuously from the last prevailing period to the next. Meanwhile, the largest number of private alleles might reflect that the strains in Jiangsu are in the process of being adapted to a stressful environment.

However, some questions remain to address. First, the Jiangsu, Anhui, and Zhejiang regions might not be the only source of *Pt* in eastern China. Epidemiological studies on *Pst* have revealed that the Guanzhong and Huabei regions are overwintering areas for *Pst* (Chen et al. 2014). Few *Pst* strains overwinter in spring endemic areas, and the major spring inocula were from the Chengdu Plain and Jiangnan River Basin regions. Unlike *Pst*, *Pt* had a wider temperature range from 10 to 25°C (Bolton et al. 2008) and might be able to overwinter on volunteer seedlings or winter wheat in the middle and lower reaches of the Yangtze River and the Huang-Huai regions. Ma et al. (2020) studied the population structure of *Pt* in central and western China, separating the Hubei population from other regional populations, and revealed a high level of genetic diversity in the Hubei population. Hubei is adjacent to Anhui and Henan provinces, and Anhui and Henan populations might take in *Pt* strains from the western regions. However, researchers have not determined whether the western strains affect the genetic

structure of eastern populations and how these afferent strains exert their effects. Second, relevant records about the onset time and places of wheat leaf rust in eastern China are lacking, and migration inference of *Pt* stains based on population genetic structures is indirect. Therefore, it is difficult to precisely construct the transmission route of *Pt* strains. Regarding questions of where the strains overwintered, and the mechanisms by which the strains form the preliminary inocula, complete reproduction, and expand from one place to another, further studies and verification are required.

Although serving as an important wheat disease that causes substantial production loss, the division and transmission routes of wheat leaf rust pathogens in ecological zones in China have not been defined clearly. These uncertainties impede further research and effective control. Wheat stripe rust, the same genus as wheat leaf rust, is one of the main diseases in China. Previously wheat stripe rust caused catastrophic damage to wheat production in China. After identifying its ecological zones and aerial dispersal routes, advanced prediction and various effective measures, such as a logical layout of resistant varieties and pesticide application, were applied in regions with major outbreaks of *Pst* and reduced the loss successfully. Regarding the inoculation regions, wheat leaf rust harms broader areas of wheat than wheat stripe rust (Bolton et al. 2008). Although wheat leaf rust is not taken as seriously as wheat stripe rust in China, it is already the disease causing the highest global wheat yield losses (Savary et al. 2019). Further research on wheat leaf rust in advance must be conducted to facilitate the control of disease epidemics on a large scale and ensure the stable production of wheat (Chen et al. 2014).

Conclusions

In this study, we report the genetic differentiation of southern and northern *Pt* populations in eastern China and conclude that *Pt* strains in southern regions of eastern China of Jiangsu, Anhui, and Zhejiang may migrate northward and affect the structure of northern populations. This research provides a basis for understanding the spread and prevalence of wheat leaf rust in eastern regions of China, which is also essential for future work of epidemiological research and disease control.

Methods

Sample collection

During the period of wheat leaf rust occurrence in eastern China, from April to June 2020, researchers from the Institute of Plant Protection, Chinese Academy of Agricultural Sciences took wheat leaf rust samples along a field survey route, started from Anhui, Jiangsu, and Zhejiang provinces in April, and then northward to Henan

and Shandong provinces in May. Samples in Hebei province and Beijing were collected from late April to June. A total of 372 wheat leaf samples infected with *Pt* were collected from 76 sampling sites. Single-leaf tissue samples were sandwiched between drying papers. After drying, the samples were stored in a kraft paper bag at 4°C. Each leaf segment (approximately 1 cm²) containing inoculated uredinium was cut and served as the material for DNA extraction of a sample (Ali et al. 2011). The sampling map was drawn using MeteoInfoMap software v 2.3.2 (<http://www.meteothink.org/>) (Wang et al. 2009) and the *ggplot2* v 3.3.5 package (Wickham 2016) in R software v 4.1.1 (Team 2021).

SSR amplification

Each leaf tissue was placed into a 2 mL grinding tube with 1 grain of 5.0 mm grinding stainless steel beads (YA3032-500 g, Beijing Solarbio Technology Co., Ltd., China) and 0.3 g of 1.0 mm glass grinding beads (BE6061-500, Beijing Easybio Technology Co., Ltd., China). The tubes were frozen with liquid nitrogen and then ground using Fastprep tissue homogenizers (MpBio China) at 1800 strokes/min for 30 s. A plant genomic DNA kit (TIANGEN Biotech Co. Ltd., Beijing, China) was used for genomic DNA extraction. The DNA quality and concentration were determined using a DS-11 spectrophotometer/fluorometer series (DeNovix) to ensure the authenticity of missing loci, requiring 260 nm/280 nm values ranging from 1.8–2.0 and concentrations greater than 20 ng/μL. DNA solutions were diluted to between 20 ng/μL and 50 ng/μL for SSR genotyping.

A preliminary experiment have done for detecting the amplification efficiency and polymorphism of markers, and 16 primers for the SSR loci (RB17, RB28, RB4, RB12, RB8, RB1, RB35, RB11, and RB16 (Duan et al. 2003), as well as PtSSR151A, PtSSR161, PtSSR152, PtSSR173, PtSSR164, PtSSR55, and PtSSR13 (Szabo and Kolmer 2007) were suitable and used for SSR genotyping in this study. The primers were modified by fluorescent dyes FAM, HEX, CY3, or ROX (Beijing Qingke Biotechnology Co., Ltd., China) at the 5' end. Each PCR reaction contained 10 μL of 2× M5 HiPer plus Taq HiFi PCR mix (Beijing Mei5 Biotechnology Co., Ltd China), 7 μL of ddH₂O, 1 μL (10 ng/μL) of each primer, and 1 μL of diluted DNA solution, for a total volume of 20 μL. Thermal cycling conditions included an initial denaturation step at 95°C for 3 min; followed by 35 cycles of denaturation at 94°C for 25 s, annealing at 58°C to 61°C for 15 s and elongation at 72°C for 5 min; and final elongation step at 72°C for 5 min. PCRs were conducted using a Veriti 96-well machine (Applied Biosystems, China Branch).

For genetic analysis, PCR products were diluted at 1:50 with sterile water, and then 1 μL of diluted PCR products

was added into a 9 μ L mixture of HiDi Formamide (Applied Biosystems): GeneScan 500 LIZ Size Standard (Applied Biosystems)=1000:15. After centrifugation for 2 min (OSE-MP26, TIANGEN Biotech Co. Ltd, Beijing, China), the mixtures were heated at 95°C for 5 min and then placed in an ice bath for 3 min. Fluorescence signals were detected using a 3500 Genetic Analyzer (Applied Biosystems). GeneMarker software v 2.7.0 (SoftGenetics) was used to read the peak signals and determine the genotypes.

Population genetic structure

The preliminary analysis of the data was performed using the *poppr* v 2.9.3 package (Kamvar et al. 2014, 2015) in R (Team 2021), and primers with locus loss rates greater than 5%, as well as samples with locus loss rates greater than 10%, were removed to ensure the accuracy of analysis. The diversity analysis of SSR primers was performed using GenAEx v 6.502 (Peakall and Smouse 2006, 2012).

After data cleaning, the population structure analysis was conducted using STRUCTURE v 2.3 (Hubisz et al. 2009), which implements a model-based clustering method for inferring population structure using genotype data consisting of unlinked markers (Pritchard et al. 2000). After the linkage disequilibrium analysis, the genetic data were analyzed using STRUCTURE with no preclustering. The parameters were set as follows (Evanno et al. 2005): the length of the burn-in period was 10,000, the number of MCMC repetitions after burn-in was 100,000, and K was set from 1–10 to perform 10 independent runs. The results obtained from the STRUCTURE model were packaged and submitted to Structure Harvester Web v 0.6.94 (Earl and vonHoldt 2012) to estimate the best K value based on ΔK . CLUMPP v 1.1.2 (Jakobsson and Rosenberg 2007) was used to determine optimal alignments of replicate cluster analyses of the same data. The output from CLUMPP was submitted to Distruct v 1.1 (Rosenberg 2004) to visualize the population structure results; this software is capable of showing individuals as line fragments of different colors from K estimated clusters.

Discriminant analysis of principal components (DAPC) is a method using sequential K means and model selection to infer genetic clusters; principal component analysis (PCA) is first performed to transform the data, and then a discriminant analysis (DA) is performed to identify clusters (Jombart et al. 2010). Clone correction of populations was performed internally before the DAPC analysis. DAPC was performed using the *adegenet* v 2.0.0 package (Jombart 2008; Jombart and Ahmed 2011) in R (Team 2021) to visualize genetic clusters.

Clustering analyses were performed based on the unweighted pair group method with arithmetic means

(UPGMA) and the neighbor-joining (NJ) methods using the *ape* v 5.6–2 package (Paradis and Schliep 2019) in R. And phylogenetic trees were visualized and annotated using *ggtree* v. 3.2.1 package (Yu et al. 2018) in R.

Analysis of molecular variance (AMOVA) was performed using Arlequin v. 3.5 with 9999 permutations to detect differences between populations based on evolutionary distance (Excoffier et al. 1992, 2005). Arlequin v. 3.5 was used to calculate F_{ST} (Wright 1951), and GenAEx v 6.502 (Peakall and Smouse 2006, 2012) was used to calculate genetic diversity indexes and the number of migrations (Nm) (Slatkin 1985) to clarify the genetic diversity and genetic divergence between populations. The analysis of genetic diversity was based on indexes of the number of effective alleles (N_e), Shannon's information index (I) (Shannon 1948), the number of alleles unique to a single population, the expected heterozygosity (H_e) (Nei 1973), and the unbiased expected heterozygosity (uH_e) (Nei 1978). F_{ST} (pairs of alleles between individuals within populations) indicates the inbreeding number of populations relative to the total population (Weir and Cockerham 1984). Nm was calculated based on the migration rate, $Nm = [(1/F_{ST}) - 1]/4$.

Relative migration rates were used to estimate the directional components of genetic divergence and asymmetric gene flow between population pairs (Sundqvist et al. 2016). SSR genotype information was submitted to divMigrate-online (<https://popgen.shinyapps.io/divMigrate-online/>), with the number of bootstraps set to 1000. Alpha was set to 0.05, and D (JOST 2008) was used to calculate the relative migration statistic.

Horizontal wind field analysis

The horizontal wind field analysis revealed the direction of wind streamlines at a specific height or barometric surface, helping to clarify the wind-driven pathogen migration. EAR5 (ECMWF Reanalysis v5) was the fifth generation ECMWF (European Centre for Medium-Range Weather Forecasts) reanalysis of the global climate and weather for the past 4 to 7 decades. EAR5 reanalysis data were provided by Copernicus Climate Change Service (C3S) of ECMWF. 'ERA5 hourly data on single levels from 1959 to present' had an hourly temporal resolution and $0.25^\circ \times 0.25^\circ$ horizontal resolution (atmosphere) and was used to extract the 10 m u-component of wind and the 10 m v-component of wind to analyze the average horizontal wind field at 10 m altitude from April to June 2020. 'ERA5 monthly mean data on single levels from 1959 to present' had a monthly temporal resolution and $0.25^\circ \times 0.25^\circ$ horizontal resolution (atmosphere) and was used to extract the 10 m u-component of wind and the 10 m v-component of wind to analyze the average horizontal wind field at

10 m altitude from April to June in 2016–2020. ‘ERA5 monthly mean data on pressure levels from 1959 to present’ had a monthly temporal resolution, $0.25^\circ \times 0.25^\circ$ horizontal resolution, 1000 hPa to 1 hPa vertical coverage, and 37 pressure levels of vertical resolution. It was used to extract the u-component and v-component of wind at 900 hPa and to analyze the average horizontal wind field at 900 hPa from April to June 2016–2020. The horizontal wind field analysis was performed using Python v 3.9 and the third-party libraries xarray, pandas, numpy, cartopy, and matplotlib to analyze ERA5 reanalysis data during wheat leaf rust occurrence seasons in eastern China (from April to June) from 2016 to 2020.

Abbreviations

AMOVA	Analysis of molecular variance
CPH	The central–periphery hypothesis
DA	Discriminant analysis
DAPC	Discriminant analysis of principal components
eMLG	Number of expected MLG at the smallest sample size ≥ 10 based on rarefaction
F_{ST}	Pairs of alleles between individuals within populations
H_e	The expected heterozygosity
I	Shannon’s information index
MLG	Multi-locus genotypes
N_e	Indexes of the number of effective alleles
NJ	The neighbor-joining methods
N_m	Number of migrations
PCA	Principal component analysis
<i>Pt</i>	<i>Puccinia triticina</i>
SSR genotype	A series of SSR loci with different numbers of base pairs
SSR	Simple sequence repeat
uH_e	The unbiased expected heterozygosity
UPGMA	The unweighted pair group method with arithmetic means

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s42483-023-00163-3>.

Additional file 1: Figure S1. Genotype accumulation curve is used to assess if it is sufficient to discriminate between unique individuals for the given loci. It had reached the plateau of 289 MLGs, indicated that the set of SSR loci we used is enough for distinguishing all the observed MLGs. **Figure S2.** deltaK showed the best K value was 3 in this study. **Figure S3.** DAPC analysis of 11 *Pt* populations in eastern China shows 3 clusters on the coordinate axis. H1, H3, H4, H5, and H8 from Cluster 1 are located on the left side of the Y axis, whereas H2, H6, and H11 from Cluster 2 are located below the X axis, and H7, H9, H10 from Cluster 3 are located on the right side of the Y axis. **Figure S4.** Separate DAPC analyses of the 3 clusters show that after division according to geographical areas, Clusters 2 (b) and 3 (c) still exhibit clear differentiation between subpopulations. A large mixture of samples exists in Cluster 1 (a). **Figure S5.** Phylogenetic trees of the 3 clusters constructed based on the unweighted pair group method with arithmetic means (UPGMA) and the neighbor-joining (NJ) method. The red dots represent nodes, and the blue triangles represent samples. Clades are shown on the right side with different colors of labels.

Acknowledgements

We thank Prof. Hao Zhang (Institute of Plant Protection, Chinese Academy of Agricultural Science) for his help in sample collections.

Authors’ contributions

TL and WC designed the research, BL, LG, WZ, and LX provided sampling help, HL performed the research, QZ, ZC, and XZ analyzed the data, JW, GW, and ZC provided assistance in meteorological data analysis, HL and TL wrote the manuscript. All authors read and approved the final manuscript.

Funding

This study was supported by grants from the National Key Research and Development Program (2021YFD1401000), National Natural Science Foundation of China (31671967), China Agriculture Research System (CARS-3), Agricultural Science and Technology Innovation Program (CAAS-ASTIP, CAAS-ZDRW202002), and Epidemic Detection and Control of Crop Diseases and Insect Pests (2130108).

Availability of data and materials

Not applicable.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Received: 26 October 2022 Accepted: 19 January 2023

Published online: 28 February 2023

References

- Alam MA, Li H, Hossain A, Li M. Genetic diversity of wheat stripe rust fungus *Puccinia striiformis* f. sp. tritici in Yunnan, China. *Plants*. 2021;10:1735. <https://doi.org/10.3390/plants10081735>.
- Ali S, Gautier A, Leconte M, Enjalbert J, Vallavieille-Pope Cd. 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 Res Notes*. 2011;4:240. <https://doi.org/10.1186/1756-0500-4-240>.
- Bolton MD, Kolmer JA, Garvin DF. Wheat leaf rust caused by *Puccinia triticina*. *Mol Plant Pathol*. 2008;9:563–75. <https://doi.org/10.1111/j.1364-3703.2008.00487.x>.
- Brussard PF. Geographic patterns and environmental gradients: the central-marginal model in drosophila revisited. *Annu Rev Ecol Syst*. 1984;15:25–64. <https://doi.org/10.1146/annurev.es.15.110184.000325>.
- Chen W, Wellings C, Chen X, Kang Z, Liu T. Wheat stripe (yellow) rust caused by *Puccinia striiformis* f. sp. tritici. *Mol Plant Pathol*. 2014;15:433–46. <https://doi.org/10.1111/mpp.12116>.
- Cooke BM, Gareth JD, Kaye B. *The epidemiology of plant diseases*. Netherlands: Springer; 2006.
- Craigie JH. Epidemiology of stem rust in western Canada. *Sci Agric*. 1945;25:285–401.
- Duan X, Enjalbert J, Vautrin D, Solignac M, Giraud T. Isolation of 12 microsatellite loci, using an enrichment protocol, in the phytopathogenic fungus *Puccinia triticina*. *Mol Ecol Notes*. 2003;3:65–7. <https://doi.org/10.1046/j.1471-8286.2003.00350.x>.
- Earl DA, vonHoldt BM. Structure Harvester: a website and program for visualizing structure output and implementing the Evanno method. *Conserv Genet Resour*. 2012;4:359–61. <https://doi.org/10.1007/s12686-011-9548-7>.
- Eckert CG, Samis KE, Loughheed SC. Genetic variation across species’ geographical ranges: the central-marginal hypothesis and beyond. *Mol Ecol*. 2008;17:1170–88. <https://doi.org/10.1111/j.1365-294X.2007.03659.x>.
- Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol*. 2005;14:2611–20. <https://doi.org/10.1111/j.1365-294X.2005.02553.x>.

- Eversmeyer MG, Kramer CL. Epidemiology of wheat leaf and stem rust in the central Great Plains of the USA. *Annu Rev Phytopathol.* 2000;38:491–513. <https://doi.org/10.1146/annurev.phyto.38.1.491>.
- Excoffier L, Smouse PE, Quattro JM. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics.* 1992;131:479–91. <https://doi.org/10.1093/genetics/131.2.479>.
- Excoffier L, Laval G, Schneider S. Arlequin (version 3.0): An integrated software package for population genetics data analysis. *Evol Bioinf.* 2005;1:47–50. <https://doi.org/10.1177/117693430500100003>.
- Fellers JP, Sakthikumar S, He F, McReil K, Bakkeren G, Cuomo CA, et al. Whole-genome sequencing of multiple isolates of *Puccinia triticina* reveals asexual lineages evolving by recurrent mutations. *G3: Genes Genom Genet.* 2021;11:jkab219. <https://doi.org/10.1093/g3journal/jkab219>.
- Food and Agriculture Organization of the United Nations. Production quantities of Wheat by country. 2022.
- Ge R, Liu T, Gao L, Liu B, Chen W. Virulence of *Puccinia triticina* from 6 Provinces in China in 2011–2012. *Acta Phytopathol Sin.* 2015;45:175–80. [https://doi.org/10.13926/j.cnki.apps.2015.02.008\(inChinese\)](https://doi.org/10.13926/j.cnki.apps.2015.02.008(inChinese)).
- Huang J, Zhang B, Sun Z, Jia Q, Cao S, Luo H, et al. Population structure and diversity analysis of *Puccinia triticina* in Gansu province from 2016 to 2019. *J Triticeae Crops.* 2022;42:764–72. **(in Chinese)**.
- Huang G, Yao G, Xia X, Liu Z. Overwintering and oversummer of wheat leaf and stem rust in Sichuan. *Plant Prot* 2005, pp 67–8. **(in Chinese)**.
- Hubisz MJ, Falush D, Stephens M, Pritchard JK. Inferring weak population structure with the assistance of sample group information. *Mol Ecol Resour.* 2009;9:1322–32. <https://doi.org/10.1111/j.1755-0998.2009.02591.x>.
- Huerta-Espino J, Singh RP, Germán S, McCallum BD, Park RF, Chen WQ, et al. Global status of wheat leaf rust caused by *Puccinia triticina*. *Euphytica.* 2011;179:143–60. <https://doi.org/10.1007/s10681-011-0361-x>.
- Jakobsson M, Rosenberg NA. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics.* 2007;23:1801–6. <https://doi.org/10.1093/bioinformatics/btm233>.
- Jombart T. *adeigenet*: a R package for the multivariate analysis of genetic markers. *Bioinformatics.* 2008;24:1403–5. <https://doi.org/10.1093/bioinformatics/btn129>.
- Jombart T, Ahmed I. *adeigenet* 1.3–1: new tools for the analysis of genome-wide SNP data. *Bioinformatics.* 2011;27:3070–1. <https://doi.org/10.1093/bioinformatics/btr521>.
- Jombart T, Devillard S, Baloux F. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genet.* 2010;11:94. <https://doi.org/10.1186/1471-2156-11-94>.
- Jost L. GST and its relatives do not measure differentiation. *Mol Ecol.* 2008;17:4015–26. <https://doi.org/10.1111/j.1365-294X.2008.03887.x>.
- Kamvar ZN, Tabima JF, Grunwald NJ. *Poppr*: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ.* 2014;2:e281. <https://doi.org/10.7717/peerj.281>.
- Kamvar ZN, Brooks JC, Grunwald NJ. Novel R tools for analysis of genome-wide population genetic data with emphasis on clonality. *Front Genet.* 2015;6:208. <https://doi.org/10.3389/fgene.2015.00208>.
- Kolmer JA. Physiologic specialization of *Puccinia recondita* f. sp. *tritici* in Canada in 1996. *Can J Plant Pathol.* 1998;20:176–81. <https://doi.org/10.1080/07060669809500424>.
- Kolmer JA. Collections of *Puccinia triticina* in different provinces of China are highly related for virulence and molecular genotype. *Phytopathology.* 2015;105:700–6. <https://doi.org/10.1094/phyto-11-14-0293-r>.
- Kolmer JA, Ordoñez ME. Genetic differentiation of *Puccinia triticina* populations in Central Asia and the Caucasus. *Phytopathology.* 2007;97:1141–9. <https://doi.org/10.1094/PHYTO-97-9-1141>.
- Kolmer JA, Hanzalova A, Goyeau H, Bayles R, Morgounov A. Genetic differentiation of the wheat leaf rust fungus *Puccinia triticina* in Europe. *Plant Pathol.* 2013;62:21–31. <https://doi.org/10.1111/j.1365-3059.2012.02626.x>.
- Konecka A, Tereba A, Studnicki M, Nowakowska JA. Rare and private alleles as a measure of gene pool richness in Scots pine planting material. *Sylwan.* 2019;163:948–56. <https://doi.org/10.26202/sylwan.2019068>.
- Li ZQ, Zeng SM. Wheat rust in China. Beijing: China Agriculture Press; 2002. **(in Chinese)**.
- Li MJ, Zhang YH, Chen WQ, Duan XY, Liu TG, Jia QZ, et al. Evidence for Yunnan as the major origin center of the dominant wheat fungal pathogen *Puccinia striiformis* f. sp. *tritici*. *Australas Plant Pathol.* 2021;50:241–52. <https://doi.org/10.1007/s13313-020-00770-0>.
- Liang B. Characteristics and control techniques of wheat leaf rust in Pinglu County. *Agri Technol Equip.* 2019, pp 108–109. <https://kns.cnki.net/kcms/detail/detail.aspx?FileName=NJTU201912054&DbName=CJFQ2019> **(in Chinese)**.
- Liu TG, Chen WQ. Race and virulence dynamics of *Puccinia triticina* in China during 2000–2006. *Plant Dis.* 2012;96:1601–7. <https://doi.org/10.1094/pdis-06-10-0460-re>.
- Liu Q, Tang Y, Zhang B. Study on occurrence and control of wheat leaf rust in Shandong Province. *Shandong Agri Sci.* 1989; pp 10–3. <https://kns.cnki.net/kcms/detail/detail.aspx?FileName=AGRI198901002&DbName=CJFQ1989> **(in Chinese)**.
- Ma YT, Liu TG, Liu B, Gao L, Chen WQ. Population genetic structures of *Puccinia triticina* in five provinces of China. *Eur J Plant Pathol.* 2020;156:1135–45. <https://doi.org/10.1007/s10658-020-01956-4>.
- Nei M. Analysis of gene diversity in subdivided populations. *Proc Natl Acad Sci USA.* 1973;70:3321–3. <https://doi.org/10.1073/pnas.70.12.3321>.
- Nei M. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics.* 1978;89:583–90. <https://doi.org/10.1093/genetics/89.3.583>.
- Ordoñez ME, Germán SE, Kolmer JA. Genetic differentiation within the *Puccinia triticina* population in South America and comparison with the North American population suggests common ancestry and intercontinental migration. *Phytopathology.* 2010;100:376–83. <https://doi.org/10.1094/phyto-100-4-0376>.
- Pan Z, Yang XB, Pivonia S, Xue L, Pasken R, Roads J. Long-term prediction of soybean rust entry into the Continental United States. *Plant Dis.* 2006;90:840–6. <https://doi.org/10.1094/pd-90-0840>.
- Paradis E, Schliep K. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics.* 2019;35:526–8. <https://doi.org/10.1093/bioinformatics/bty633>.
- Peakall R, Smouse PE. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Mol Ecol Notes.* 2006;6:288–95. <https://doi.org/10.1111/j.1471-8286.2005.01155.x>.
- Peakall R, Smouse PE. GenALEX 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics.* 2012;28:2537–9. <https://doi.org/10.1093/bioinformatics/bts460>.
- Peng H, Yu SQ. Epidemiological analysis and control measures of wheat leaf rust in Henan Province in 2015. *Rural Sci Technol.* 2016;3:91–3. <https://doi.org/10.19345/j.cnki.1674-7909.2016.03.071>. **(in Chinese)**.
- Pironon S, Papuga G, Vilellas J, Angert AL, Garcia MB, Thompson JD. Geographic variation in genetic and demographic performance: new insights from an old biogeographical paradigm. *Biol Rev.* 2017;92:1877–909. <https://doi.org/10.1111/brv.12313>.
- Prasad P, Savadi S, Bhardwaj SC, Gupta PK. The progress of leaf rust research in wheat. *Fungal Biol.* 2020;124:537–50. <https://doi.org/10.1016/j.funbio.2020.02.013>.
- Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. *Genetics.* 2000;155:945–59. <https://doi.org/10.1093/genetics/155.2.945>.
- Rosenberg NA. DISTRUCT: a program for the graphical display of population structure. *Mol Ecol Notes.* 2004;4:137–8. <https://doi.org/10.1046/j.1471-8286.2003.00566.x>.
- Savary S, Willcoquet L, Pethybridge SJ, Esker P, McRoberts N, Nelson A. The global burden of pathogens and pests on major food crops. *Nat Ecol Evol.* 2019;3:430–9. <https://doi.org/10.1038/s41559-018-0793-y>.
- Shannon CE. A mathematical theory of communication. *Bell Syst Tech J.* 1948;27(379–423):623–56. <https://doi.org/10.1002/j.1538-7305.1948.tb01338.x>.
- Slatkin M. Gene flow in natural populations. *Annu Rev Ecol Syst.* 1985;16:393–430. <https://doi.org/10.1146/annurev.es.16.1.393>.
- Sundqvist L, Keenan K, Zackrisson M, Prodohl P, Kleinhans D. Directional genetic differentiation and relative migration. *Ecol Evol.* 2016;6:3461–75. <https://doi.org/10.1002/ece3.2096>.
- Szabo LJ, Kolmer JA. Development of simple sequence repeat markers for the plant pathogenic rust fungus *Puccinia triticina*. *Mol Ecol Notes.* 2007;7:708–10. <https://doi.org/10.1111/j.1471-8286.2007.01686.x>.
- Team RC. R: a Language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. 2021. <https://www.R-project.org/>.

- Wang YQ, Zhang XY, Draxler RR. TrajStat: GIS-based software that uses various trajectory statistical analysis methods to identify potential sources from long-term air pollution measurement data. *Environ Model Software*. 2009;24:938–9. <https://doi.org/10.1016/j.envsoft.2009.01.004>.
- Weir BS, Cockerham CC. Estimating F -statistics for the analysis of population structure. *Evolut Int J Organ Evolut*. 1984;38:1358–70. <https://doi.org/10.1111/j.1558-5646.1984.tb05657.x>.
- Wickham H. *ggplot2: elegant graphics for data analysis*. New York: Springer; 2016.
- Wright S. The genetical structure of populations. *Ann Eugen*. 1951;15:323–54. <https://doi.org/10.1111/j.1469-1809.1949.tb02451.x>.
- Yu G, Lam TT-Y, Zhu H, Guan Y. Two methods for mapping and visualizing associated data on phylogeny using ggtree. *Mol Biol Evol*. 2018;35:3041–3. <https://doi.org/10.1093/molbev/msy194>.
- Zhao J, Wang MN, Chen XM, Kang ZS. Role of alternate hosts in epidemiology and pathogen variation of cereal rusts. *Annu Rev Phytopathol*. 2016;54:207–28. <https://doi.org/10.1146/annurev-phyto-080615-095851>.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

