



# Genetic Diversity and Population Genetic Structure of *Setosphaeria turcica* From Sorghum in Three Provinces of China Using Single Nucleotide Polymorphism Markers

Linkai Cui<sup>††</sup>, Junli Deng<sup>††</sup>, Linxi Zhao<sup>1</sup>, Yanhong Hu<sup>1\*</sup> and Tingli Liu<sup>2\*</sup>

## OPEN ACCESS

### Edited by:

Mohammad Arif,  
University of Hawaii at Manoa,  
United States

### Reviewed by:

Francisco Javier Flores,  
University of the Armed Forces  
(ESPE), Ecuador  
Sowmya Ramachandran,  
Foreign Disease-Weed  
Science Research (USDA-ARS),  
United States

### \*Correspondence:

Yanhong Hu  
9902618@haust.edu.cn  
Tingli Liu  
tliu2100@163.com

<sup>††</sup>These authors have contributed  
equally to this work

### Specialty section:

This article was submitted to  
Microbe and Virus Interactions with  
Plants,  
a section of the journal  
Frontiers in Microbiology

**Received:** 13 January 2022

**Accepted:** 10 February 2022

**Published:** 02 March 2022

### Citation:

Cui L, Deng J, Zhao L, Hu Y and  
Liu T (2022) Genetic Diversity and  
Population Genetic Structure of  
*Setosphaeria turcica* From Sorghum  
in Three Provinces of China Using  
Single Nucleotide Polymorphism  
Markers.  
*Front. Microbiol.* 13:853202.  
doi: 10.3389/fmicb.2022.853202

<sup>1</sup>College of Horticulture and Plant Protection, Henan University of Science and Technology, Luoyang, China, <sup>2</sup>Jiangsu Key Laboratory for Horticultural Crop Genetic Improvement, Institute of Leisure Agriculture, Jiangsu Academy of Agricultural Sciences, Nanjing, China

*Setosphaeria turcica* is a heterothallic fungus that is the causal agent of northern leaf blight (NLB), which is a devastating foliar disease of sorghum and maize. Despite of its adversary to crop production, little is known about the genetic diversity and population genetic structure of this pathogen from sorghum. In this study, we explored the utilization of single nucleotide polymorphism (SNP) molecular markers and three mating type-specific primers to analyze the genetic diversity, population genetic structure, and mating type distribution of 87 *S. turcica* isolates that had been collected in sorghum production areas from three provinces, including Henan, Shaanxi, and Shanxi in China. The populations are featured with moderate genetic diversity and relatively equal mating type distribution of MAT1-1 and MAT1-2. The genetic differentiation was significant ( $p < 0.05$ ) among different populations except those from Henan and Shanxi provinces that showed particularly frequent gene flow between them. Neither the maximum likelihood phylogenetic tree, nor principal coordinate analysis, nor genetic structure analysis was able to completely separate the three populations. The relatively low genetic distance and high genetic identification were also observed among the three populations. Nevertheless, the genetic variation within populations was the major source of variation as revealed by AMOVA analysis. The findings of this study have improved our current understanding about the genetic diversity, population genetic structure, and the distribution of mating type of *S. turcica*, which are useful for unraveling the epidemiology of NLB and developing effective disease management strategies.

**Keywords:** *Setosphaeria turcica*, genetic diversity, population genetic structure, mating type, SNP molecular marker

## INTRODUCTION

Sorghum (*Sorghum bicolor*) is an important crop with multiple uses for food, brewing, fodder, and biofuel applications in many countries in the world (Lu and Dahlberg, 2001); and it is reemerging to prominence as an alternative crop for a number of high yielding crops, such as corn, due to its stronger drought tolerance, especially in the current circumstances with

increasing occurrence of drought and change of weather patterns (Nieuwoudt et al., 2018). However, the production of sorghum is severely constrained by the infestation of northern leaf blight (NLB) which is caused by a pathogen known as *Setosphaeria turcica* (Luttr.) Leonard and Suggs (1974); anamorph: *Exserohilum turcicum* (Pass.) Leonard and Suggs (1974). Apart from sorghum, NLB is also a devastating foliar disease of maize and some other related grass species wherever they are cultivated (Leonard and Suggs, 1974; Tang et al., 2015). NLB is very prevalent, spreading more widely, especially in the environment, such as warm and humid conditions, that is conducive to the development of the disease (Ferguson and Carson, 2004), which causes considerable economic destruction. The disease mainly infects sorghum leaves, whose typical symptoms are long fusiform, light brown to brown in the center, purple at the edge, and irregular ring patterns can be seen in the early stage (Dai et al., 2021). Upon NLB infection, the yield of sorghum could be severely compromised as the result of reduction in photosynthetic potential of the plant.

As a heterothallic ascomycete, the individual isolate of *S. turcica* cannot mate with itself, and its sexual reproduction requires two isolates with complementary mating type genes. The perfect stage (teleomorph or sexual stage) of *S. turcica* was first described in 1958 (Luttrell, 1958), which is controlled by a specific mating type (MAT) gene locus that may appear in two alternative allele forms, namely, MAT1-1 and MAT1-2 (Luttrell, 1958). Although rarely occurs in the field, the sexual stage of *S. turcica* is can be observed under laboratory conditions (Luttrell, 1958; Fan et al., 2007). Previous research results have revealed that the pathogenicity and parasitic adaptability of some sexual progenies are significantly enhanced after sexual reproduction (Guo et al., 2013). In addition, various races of *S. turcica* with different mating type could mate randomly to produce new races of different pathogenicity (Hou et al., 2006; Fan et al., 2007). It is generally recognized that the genetic variability and pathogenicity of *S. turcica* are the key factors that affect host-plant resistance and develop feasible strategies for disease management (Sree et al., 2012). This necessitates studies in these areas to better understand the relationship between genetic diversity and mating type of *S. turcica*.

Molecular markers have been successfully used to analyze genetic diversity, population genetic structure, and mating type distribution of *S. turcica* isolates, such as universally primed polymerase chain reaction (UP-PCR), random amplified polymorphic DNA (RAPD), simple sequence repeat (SSR), and sequence-related amplified polymorphism (SRAP). RAPD molecular markers were developed to analyze genetic diversity and population genetic structure of 264 *S. turcica* isolates from four different continents, which displayed the variability in the genetic diversity and mating type distribution of different regions and states (Borchardt et al., 1998). This was corroborated by further studies by using UP-PCR molecular markers that the pathogenic specialization of *S. turcica* isolates was more correlated with genetic diversity than geographic distribution (Tang et al., 2015). Abundant genetic diversity among various *S. turcica* isolates from different geographic locations was also demonstrated by SSR molecular marker (Li et al., 2019) and SRAP molecular markers (Ma et al., 2020a). A significant correlation between

mating types and cluster groups of *S. turcica* isolates was also established by using SSR markers (Li et al., 2019).

Despite the genetic diversity of *S. turcica* is well documented, a systematic and comprehensive analysis of the genetic structure and genetic diversity of the pathogen by using the single nucleotide polymorphism (SNP) molecular marker is lacking. SNP is the third-generation molecular marker which is characteristic of high stability, large number, wide distribution, and the ability to detect the most frequent type of DNA variation (Tang et al., 2012). What's more, there are relatively fewer studies on *S. turcica* from sorghum compared to *S. turcica* from corn in China, because sorghum is not normally considered as a food crop there. In the present study, we analyzed the genetic diversity and population genetic structure of 87 *S. turcica* isolates from Henan, Shaanxi, and Shanxi provinces in China, by using SNP molecular markers for the first time. We aim to improve our current understanding the genetic variation and evolution of *S. turcica*, which may provide useful information for the disease management of NLB in sorghum.

## MATERIALS AND METHODS

### Collection and Isolation of Samples

In 2019, the diseased leave with typical symptom were collected randomly at the middle and late stages of sorghum growth from Henan, Shanxi, and Shaanxi provinces of China (Figure 1). A total of 87 isolates were obtained. Thirty isolates were from Henan province, thirty isolates were from Shanxi province, and 27 isolates were from Shaanxi province (Table 1).

*S. turcica* were isolated by using the tissue separation methods. In brief, the infected sorghum leaves were cut into 2–3 mm pieces, which were disinfected with 75% alcohol for 20 s (sec), followed by 3% sodium hypochlorite for 2–3 min (min), and washed with sterile water for three times, prior to being transferred onto the potato dextrose agar (PDA) medium containing rifampicin (RFP) at 50 µg/ml and incubated at 25°C for 3–5 days (d). All the cultures were checked for conidiophore and conidium under microscopy to verify that isolates were *S. turcica* before being stored at 4°C until further use.

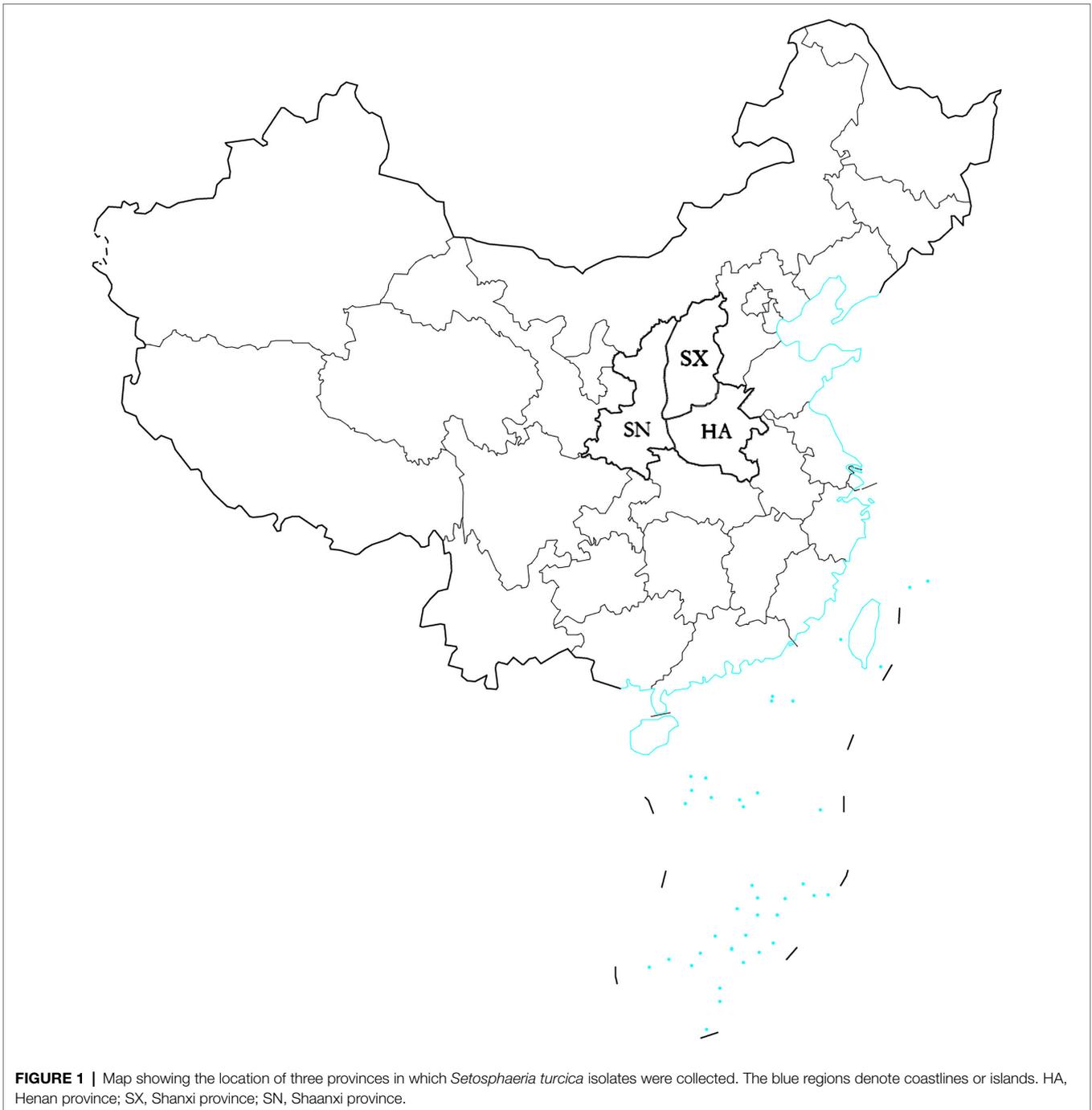
### DNA Extraction

*Setosphaeria turcica* isolates were cultured in liquid potato dextrose broth for 3 d at 25°C to collect mycelium, from which total DNA was extracted using a rapid blood genomic DNA isolation kit (Sangon, Shanghai, China), following the manufacturer's instruction. The quality and concentration of the obtained DNA was determined by using Qubit 2.0 fluorometer (Thermo Fisher Scientific, Waltham, MA, United States).

### Primer Design

A total of 172,067 potential SNP loci were detected through the whole-genome resequencing of 15 *S. turcica* isolates<sup>1</sup> using the GATK software, from which 80 polymorphic SNP loci were selected based on a minimum distance of 300 kb between

<sup>1</sup><https://www.ncbi.nlm.nih.gov/sra/PRJNA803488>



SNP loci. The primers were designed by the online tool Batchprimer <sup>32</sup> to amplify the target sequence in the range of 140bp to 270bp, the primer length was between 18bp and 35bp, and G/C content in any primer was between 35 to 65%. Finally, 77 SNP primer pairs were successfully designed (**Supplementary Table S1**), which were commercially synthesized by Sangon Biotech (Shanghai, China).

<sup>32</sup><http://wheat.pw.usda.gov/demos/BatchPrimer3/>

### Multiplex PCR and Sequencing

A panel consisting of 77 target SNP sites was designed, and libraries were prepared by using two-step PCR. The PCR reaction consisted of 1  $\mu$ l each of the forward and reverse amplicon PCR primers (10  $\mu$ M), 15  $\mu$ l kapa hifi 2  $\times$  PCR ready mix (Kapa Biosystems, Cape Town, South Africa), 2  $\mu$ l genomic DNA (10 ng/ $\mu$ l), and sterile water to 25  $\mu$ l. PCR reaction was executed in a T100<sup>TM</sup> Thermal Cycler (BIO-RAD, Hercules, CA, United States) with the following conditions: an initial

**TABLE 1** | Information of 87 *Setosphaeria turcica* isolates from 3 provinces in China, obtained from sorghum samples collected in 2019.

Isolates	Locations	Isolates	Locations
HAS19-02	Luoyang, Henan	SNS19-16	Weinan, Shaanxi
HAS19-03	Luoyang, Henan	SNS19-18	Weinan, Shaanxi
HAS19-04	Luoyang, Henan	SNS19-19	Weinan, Shaanxi
HAS19-05	Luoyang, Henan	SNS19-20	Weinan, Shaanxi
HAS19-06	Luoyang, Henan	SNS19-21	Weinan, Shaanxi
HAS19-07	Luoyang, Henan	SNS19-22	Weinan, Shaanxi
HAS19-09	Sanmenxia, Henan	SNS19-23	Weinan, Shaanxi
HAS19-10	Sanmenxia, Henan	SNS19-25	Weinan, Shaanxi
HAS19-11	Sanmenxia, Henan	SNS19-26	Weinan, Shaanxi
HAS19-13	Sanmenxia, Henan	SNS19-27	Weinan, Shaanxi
HAS19-14	Sanmenxia, Henan	SNS19-28	Weinan, Shaanxi
HAS19-15	Sanmenxia, Henan	SNS19-30	Weinan, Shaanxi
HAS19-16	Sanmenxia, Henan	SNS19-31	Weinan, Shaanxi
HAS19-18	Sanmenxia, Henan	SXS19-01	Yuncheng, Shanxi
HAS19-19	Sanmenxia, Henan	SXS19-02	Yuncheng, Shanxi
HAS19-21	Sanmenxia, Henan	SXS19-03	Yuncheng, Shanxi
HAS19-22	Sanmenxia, Henan	SXS19-04	Yuncheng, Shanxi
HAS19-24	Sanmenxia, Henan	SXS19-05	Yuncheng, Shanxi
HAS19-25	Sanmenxia, Henan	SXS19-06	Yuncheng, Shanxi
HAS19-27	Sanmenxia, Henan	SXS19-08	Yuncheng, Shanxi
HAS19-29	Luoyang, Henan	SXS19-09	Yuncheng, Shanxi
HAS19-31	Luoyang, Henan	SXS19-10	Yuncheng, Shanxi
HAS19-32	Luoyang, Henan	SXS19-11	Yuncheng, Shanxi
HAS19-33	Luoyang, Henan	SXS19-13	Yuncheng, Shanxi
HAS19-35	Luoyang, Henan	SXS19-14	Yuncheng, Shanxi
HAS19-36	Luoyang, Henan	SXS19-15	Yuncheng, Shanxi
HAS19-37	Luoyang, Henan	SXS19-18	Yuncheng, Shanxi
HAS19-38	Luoyang, Henan	SXS19-19	Yuncheng, Shanxi
HAS19-39	Luoyang, Henan	SXS19-20	Yuncheng, Shanxi
HAS19-40	Luoyang, Henan	SXS19-22	Yuncheng, Shanxi
SNS19-01	Xi'an, Shaanxi	SXS19-23	Yuncheng, Shanxi
SNS19-02	Xi'an, Shaanxi	SXS19-24	Yuncheng, Shanxi
SNS19-03	Huayin, Shaanxi	SXS19-27	Yuncheng, Shanxi
SNS19-04	Huayin, Shaanxi	SXS19-28	Yuncheng, Shanxi
SNS19-05	Huayin, Shaanxi	SXS19-29	Yuncheng, Shanxi
SNS19-07	Huayin, Shaanxi	SXS19-30	Yuncheng, Shanxi
SNS19-08	Xianyang, Shaanxi	SXS19-31	Yuncheng, Shanxi
SNS19-09	Xianyang, Shaanxi	SXS19-32	Linfen, Shanxi
SNS19-10	Xianyang, Shaanxi	SXS19-33	Linfen, Shanxi
SNS19-11	Xianyang, Shaanxi	SXS19-34	Linfen, Shanxi
SNS19-12	Xianyang, Shaanxi	SXS19-35	Linfen, Shanxi
SNS19-13	Xianyang, Shaanxi	SXS19-36	Yuncheng, Shanxi
SNS19-14	Xianyang, Shaanxi	SXS19-37	Yuncheng, Shanxi
SNS19-15	Xianyang, Shaanxi		

denaturation at 98°C for 5 min, first 8 cycles of denaturation at 98°C for 30 s, annealing at 50°C for 30 s, elongation at 72°C for 30 s, and then 25 cycles of denaturing at 98°C for 30 s, annealing at 66°C for 30 s, elongation at 72°C for 30 s and a final extension at 72°C for 5 min, eventually held at 4°C. The PCR products were fractionated by using gel electrophoresis on 1% (w/v) agarose gels in TBE buffer (Tris, boric acid, EDTA), stained with ethidium bromide (EB) prior to visualization under UV light in order to verify the size of the desired PCR product. The PCR product was then isolated from the gel and recovered using AMPure XP magnetic beads (Beckman Coulter, Brea, CA, United States). The second round of PCR reaction was performed using the first round of PCR product as a template to obtain a sequencing library with

molecular tags. The second round of PCR reaction consisted of DNA (10 ng/μl) 2 μl, universal P7 primer (10 μm) 1 μl, universal P5 primer (10 μm) 1 μl, 2× PCR Ready Mix 15 μl. The condition of PCR reaction: an initial denaturation at 98°C for 3 min, then 5 cycles of denaturing at 94°C for 30 s, annealing at 55°C for 30 s, elongation at 72°C for 30 s and a final extension at 72°C for 5 min. When the second round of PCR reaction was completed, the PCR product was also purified and recovered using AMPure XP magnetic beads. Following the preparation of the DNA sequencing libraries, paired-end sequencing of the library was performed on the HiSeq XTen sequencers (Illumina, San Diego, CA, United States).

Raw reads were filtered and having adaptor sequence removed by Cutadapt (v 1.2.1), and the low-quality bases from reads 3' to 5' (Q<20) removed by PRINSEQ-lite (v 0.20.3). The remaining clean data were then mapped to the reference genome (GenBank assembly accession: GCA\_013390295.1) by BWA (version 0.7.13-r1126) with default parameters. Samtools (version: 0.1.18) was used to calculate each genotype of the target locus. When the read frequency of a major allele was higher than 80%, it was described as homozygous, and those loci with missing data were discarded.

## Data Analysis

The SNP sites with high level of polymorphisms, but without missing data, were selected for further analysis. Arlequin 3.11 (Ma et al., 2020b) was used to analyze gene flow and the genetic variation of the three population, the genetic differentiation coefficient (Fst), and the number of migrants per generation (Nm) satisfied the following equation:  $Nm = (1 - Fst) / 4 Fst$  (Jimenez et al., 1999). Analysis of molecular variance (AMOVA) was also calculated through the Arlequin 3.11 software (Ma et al., 2020b). A range of parameters, including the observed number of alleles (Na), the number of effective alleles (Ne), Nei (1973) gene diversity (h), Shannon's information index (I), Nei's genetic identity (Nei's I) and genetic distance (Nei's D) were all calculated by POPGENE (Yeh et al., 1997). A phylogenetic tree was constructed by using maximum likelihood method with 1,000 bootstrap in Mega7 (Nei and Kumar, 2000; Kumar et al., 2016). Genetic distance (GD) and Principal Coordinate Analysis (PCoA) were analyzed by using GenALEX v6.5 (Peakall and Smouse, 2012). Population genetic structure was assessed by Structure 2.3.4 (Pritchard et al., 2000). Every possible cluster (K) ranges from 1 to 10 with 10 runs at 10,000 burn-in followed by 100,000 Markov Chain Monte Carlo simulations. The optimal value of K was acquired by online tool Structure Harvester<sup>3</sup> after obtaining the result from Structure (Earl, 2012). Eventually, population genetic structure was visualized by CLUMPP and DISTRICT (Rosenberg, 2004; Jakobsson and Rosenberg, 2007).

## Determination of Mating Type

The mating type of 87 isolates was determined by using mating type-specific primers which were designed by Haasbroek et al. (2014). The same reverse primer, MAT\_CommonR

<sup>3</sup><http://taylor0.biology.ucla.edu/structureHarvester/>

(5'-AATGCGGACACGGAATAC -3'), was used in combination with each of the two different forward primers: MAT\_1-1F (5'-CTCGTCCTTGGAGAAGAATATC-3') and MAT\_1-2F (5'-GCTCCTGGACCAAATAATACA-3'). A PCR fragment in the length of 608bp was amplified from the MAT1-1 locus and the other fragment in the length of 393bp was from the MAT1-2 locus. The PCR was performed using a mix containing 12.5µl 2 × PCR Taq Master Mix (Biotechnologies Co., Ltd., Beijing, China), 1µm mixture of three primers, 100ng DNA, with sterile water adding up to 25µl. The program was carried out first with initial denaturation at 94°C for 2min, and then 30cycles of denaturation at 94°C for 30s, annealing at 61°C for 30s, elongation at 72°C for 1min; ending with a final extension at 72°C for 10min, and finally held at 4°C. The PCR products were analyzed on a 1% (w/v) agarose gel electrophoresis running for 30min at 100V, which were then stained with

ethidium bromide (EB) and visualized under UV light. The ratios between MAT1-1 and MAT1-2 from three different geographical population were subjected to a 2 test using IBM SPSS Statistics 20 (International Business Machines Corporation, Armonk, NY, United States; Zeng et al., 2021) to estimate divergence from the expected ratio of 1:1 at the  $p < 0.05$  level.

## RESULTS

### Genetic Diversity Analysis

Among the 77 SNP loci analyzed, 41 loci containing missing data or monomorphisms were discarded, while the remaining 36 SNP loci with high-quality and informative polymorphisms were retained to assess the genetic diversity of 87 *S. turcica* isolates (Table 2).

**TABLE 2** | Characteristics of 36 SNP markers used for genetic diversity of *Setosphaeria turcica* isolates in China.

Locus	Ref	Alt	Allele I <sup>1</sup>	Allele II <sup>2</sup>	Na <sup>3</sup>	Ne <sup>4</sup>	h <sup>5</sup>	I <sup>6</sup>
L01	C	T	0.7356	0.2644	2	1.6365	0.3890	0.5776
L02	A	C	0.6897	0.3103	2	1.7484	0.4281	0.6194
L03	C	T	0.4828	0.5172	2	1.9976	0.4994	0.6926
L04	A	G	0.6092	0.3908	2	1.9090	0.4762	0.6691
L05	G	C	0.4023	0.5977	2	1.9264	0.4809	0.6739
L07	G	A	0.3563	0.6437	2	1.8474	0.4587	0.6513
L08	G	A	0.4138	0.5862	2	1.9423	0.4851	0.6782
L09	C	T	0.5172	0.4828	2	1.9976	0.4994	0.6926
L11	C	G	0.6552	0.3448	2	1.8243	0.4518	0.6442
L13	G	A	0.6437	0.3563	2	1.8474	0.4587	0.6513
L14	G	A	0.1379	0.8621	2	1.3120	0.2378	0.4012
L15	T	C	0.6552	0.3448	2	1.8243	0.4518	0.6442
L18	G	A	0.7241	0.2759	2	1.6653	0.3995	0.5890
L19	G	T	0.1494	0.8506	2	1.3408	0.2542	0.4217
L22	G	A	0.5862	0.4138	2	1.9423	0.4851	0.6782
L26	C	A	0.8391	0.1609	2	1.3700	0.2700	0.4412
L29	G	C	0.6552	0.3448	2	1.8243	0.4518	0.6442
L30	C	T	0.6092	0.3908	2	1.9090	0.4762	0.6691
L33	A	G	0.7471	0.2529	2	1.6073	0.3779	0.5655
L36	T	C	0.5977	0.4023	2	1.9264	0.4809	0.6739
L37	C	T	0.9195	0.0805	2	1.1737	0.1480	0.2799
L39	C	G	0.8621	0.1379	2	1.3120	0.2378	0.4012
L40	G	A	0.4023	0.5977	2	1.9264	0.4809	0.6739
L46	G	A	0.8391	0.1609	2	1.3700	0.2700	0.4412
L48	C	T	0.5977	0.4023	2	1.9264	0.4809	0.6739
L50	T	C	0.0115	0.9885	2	1.0233	0.0227	0.0628
L55	C	A	0.7356	0.2644	2	1.6365	0.3890	0.5776
L58	G	A	0.4138	0.5862	2	1.9423	0.4851	0.6782
L59	G	C	0.7586	0.2414	2	1.5779	0.3662	0.5527
L60	G	A	0.5057	0.4943	2	1.9997	0.4999	0.6931
L61	A	G	0.3333	0.6667	2	1.8000	0.4444	0.6365
L62	G	A	0.4943	0.5057	2	1.9997	0.4999	0.6931
L63	C	T	0.7126	0.2874	2	1.6937	0.4096	0.5998
L64	T	A	0.6437	0.3563	2	1.8474	0.4587	0.6513
L66	G	C	0.2529	0.7471	2	1.6073	0.3779	0.5655
L70	C	T	0.4023	0.5977	2	1.9264	0.4809	0.6739
Mean			0.4023	0.5977	2	1.7267	0.4046	0.5870

<sup>1</sup>Allele I = Ref.

<sup>2</sup>Allele II = Alt.

<sup>3</sup>Na = Observed number of alleles.

<sup>4</sup>Ne = Effective number of alleles.

<sup>5</sup>h = Nei's gene diversity.

<sup>6</sup>I = Shannon's information index.

The genetic diversity information of the 36 SNP loci was provided in **Table 2**. The observed number of alleles ( $N_a$ ) at each SNP locus was 2. The average number of effective alleles ( $N_e$ ) was 1.7267, ranging from 1.0233 to 1.9997. The minor allele frequency (MAF) ranged from 0.0115 to 0.4943, with about 80% of the markers having  $MAF > 0.2$ . Nei's gene diversity ( $h$ ) ranged from 0.0227 to 0.4999, with the mean of 0.4046. Shannon's information index ( $I$ ) varied from 0.0628 to 0.6931, with the mean of 0.587. In particular, the SNP locus L50 showed the lowest value, whereas the polymorphisms at L60 and L62 were the most abundant.

The genetic diversity among the three populations was summarized in **Table 3**. The values of the observed number of alleles ( $N_a$ ) of the three populations ranged from 1.9444 to 1.9722 with the mean of 1.9537. The average number of effective alleles ( $N_e$ ) was 1.6413, varying from 1.5774 to 1.7541. Nei's gene diversity ( $h$ ) ranged from 0.3430 to 0.4053 with the mean of 0.3668. Shannon's information index ( $I$ ) ranged from 0.5140 to 0.5793, with the mean of 0.5387. Although the variations among the three populations were not particularly large, relatively speaking, the genetic diversity of the Henan

population was the lowest, and that of the Shaanxi population was the highest.

## Population Genetic Structure Analysis

The optimum  $K$  value was obtained by analyzing the three populations using the online tool Structure Harvester. When the  $\Delta K$  value was the biggest, there was the optimal number of genetic clusters. We could observe that the best  $K$  value was 2, revealing that all isolates came from two genetic clusters (**Figure 2A**). In these three populations, each population consisted of two genetic clusters, termed as Cluster I and Cluster II, but the main source of clusters in every population was different (**Figure 2B**). The Henan and Shaanxi populations were dominated by Cluster I, while the main constituent of Shaanxi population was Cluster II. Most of the isolates were classified into one of the two clusters, whereas the remaining small proportions of the isolates formed a mixture of the two clusters. It appears clear that although there was no particularly significant correlation between genetic clusters and geographic subgroup, genetic divergence among the Henan, Shaanxi, and Shaanxi populations does exist.

**TABLE 3** | Estimates of genetic diversity of three populations of *Setosphaeria turcica* based on 36 SNP markers.

Pop	$N_a^1$	$N_e^2$	$h^3$	$I^4$
Henan	1.9722	1.5774	0.3430	0.5140
Shaanxi	1.9444	1.7541	0.4053	0.5793
Shanxi	1.9444	1.5924	0.3520	0.5229
Mean	1.9537	1.6413	0.3668	0.5387

<sup>1</sup> $N_a$  = Observed number of alleles.

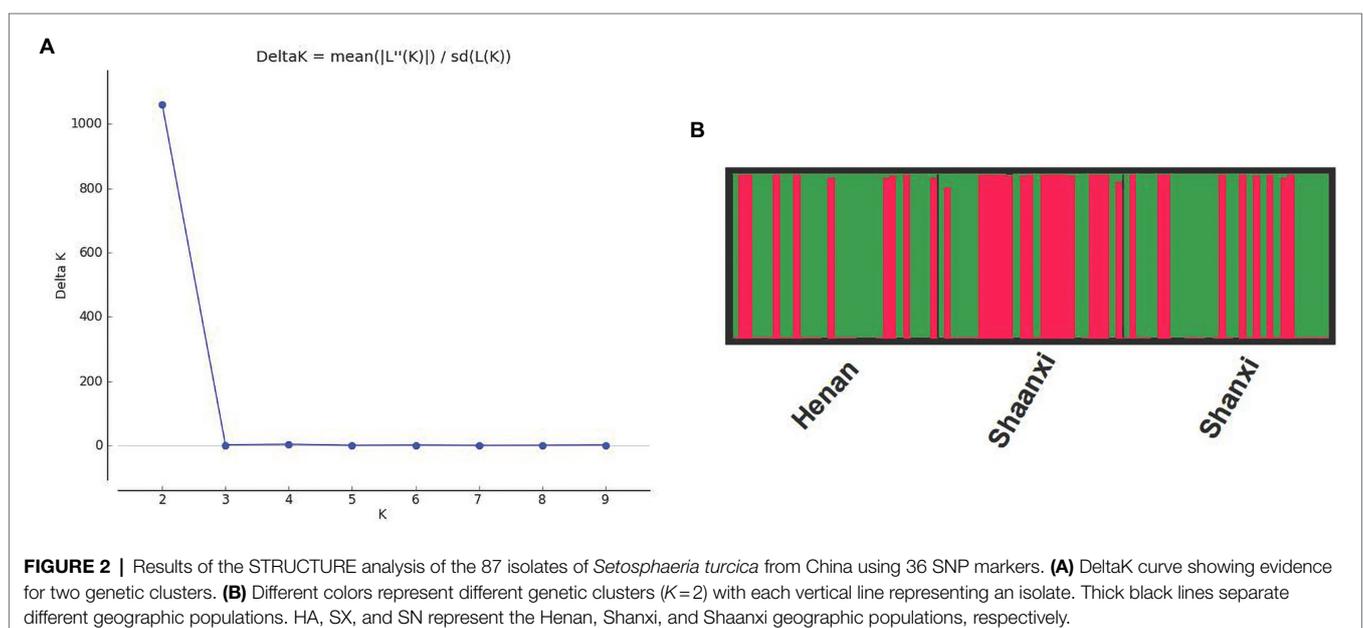
<sup>2</sup> $N_e$  = Effective number of alleles.

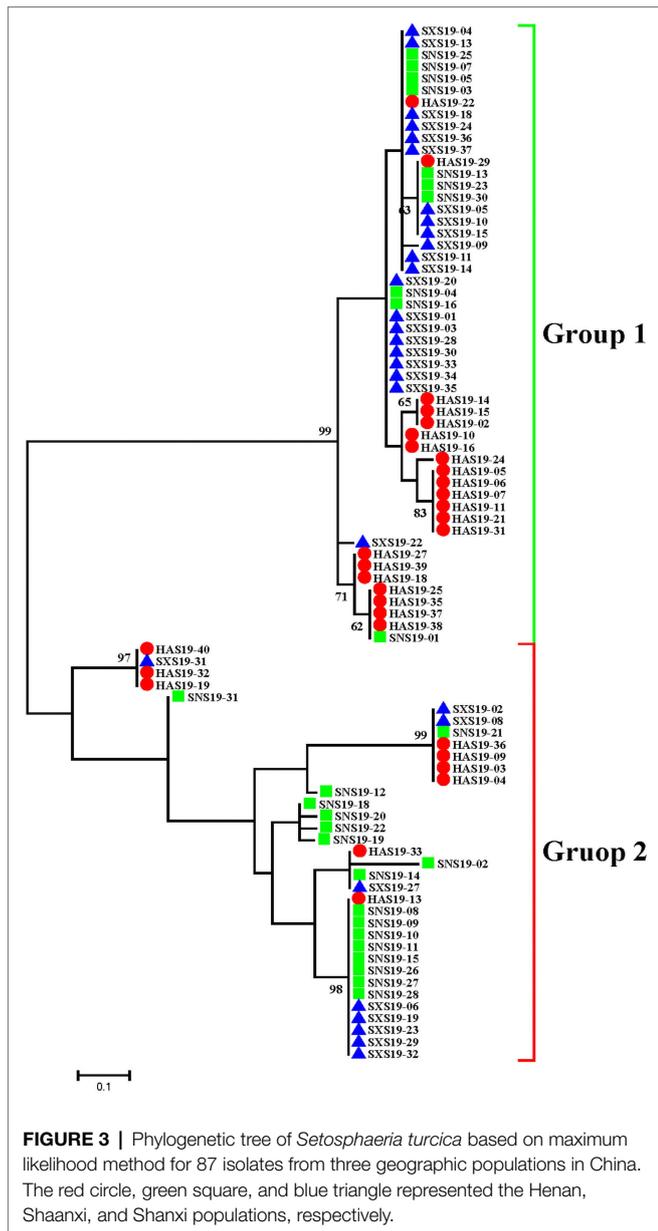
<sup>3</sup> $h$  = Nei's (1973) gene diversity.

<sup>4</sup> $I$  = Shannon's information index.

## Phylogenetic Analysis

The relationship of the three populations was also demonstrated by phylogenetic tree of maximum likelihood means, which corroborated the population genetic STRUCTURE analysis (**Figure 3**). Based on 0.08 genetic distance, 87 isolates were divided into two groups. There were 52 isolates in group 1, including 21 isolates from Henan, 21 isolates from Shaanxi, and 10 isolates from Shaanxi. There were 35 isolates in group 2, including 9 isolates from Henan, 9 isolates from Shaanxi, and 17 isolates from Shaanxi. Though each group





contained three subpopulations, the proportions of the three subpopulations in each group were variable. In Group 1, Henan and Shanxi populations accounted for a larger proportion, and were the main subpopulations, whereas Shaanxi population was the main subpopulation in Group 2. Therefore, although there was no particularly significant correlation between genetic groups and geographic origin, there might be a certain genetic distance between Henan, Shanxi, and Shaanxi populations, together with genetic differentiation to some extent.

## Principal Coordinate Analysis

The analysis methods of maximum likelihood (ML) phylogenetic tree, population STRUCTURE, and principal

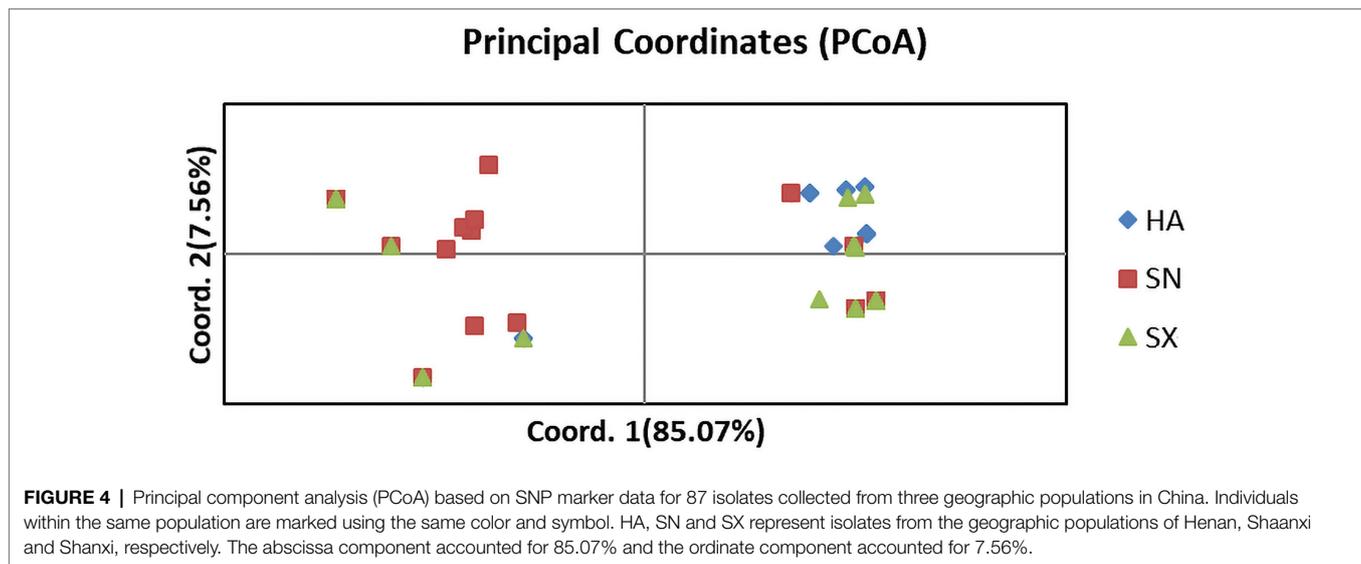
coordinate analysis (PCoA) are complementary in this study. PCoA based on the paired genetic distance matrix also elucidated the population genetic structure. As shown in **Figure 4**, the PCoA analysis primarily contained two components, among which the abscissa component accounted for 85.07% and the ordinate component accounted for 7.56%. The sum of these two components exceeded 90%, which could better explain the population genetic structure. It appears that there are no distinct separations, but rather they are often mixed together to various extent. Such a result is a manifestation that gene exchanges among the three populations are relatively frequent and the genetic diversity is rather limited.

## Analysis of Genetic Differentiation Between Populations

The genetic differentiation coefficient ( $F_{st}$ ) was calculated to compare the three populations. Among Henan/Shaanxi, Henan/Shanxi, and Shaanxi/Shanxi, the  $F_{st}$  values were 0.1581, 0.0385, and 0.1208, respectively; the number of migrants per generation ( $N_m$ ) values was 1.3253, 6.4885, and 1.8356, respectively (**Table 4**). The  $F_{st}$  values ranging from 0 to 0.05, 0.05 to 0.15, 0.15 to 0.25, and more than 0.05 represent low, moderate, high, and great genetic differentiations, respectively (Hartl and Clark, 1997). The value of  $N_m$  greater than 4 indicated that gene exchanges between the two populations are frequent (Jimenez et al., 1999). Because the gene exchanges between the Henan/Shaanxi populations were the most frequent, resulting in the lowest degree of genetic differentiation between these two populations. The genetic differentiation coefficient clearly displayed that genetic divergence was significant among populations Henan/Shaanxi and Shaanxi/Shanxi. Such a result indicated that there was a certain degree of genetic differentiation among the three populations, and gene exchanges were relatively frequent.

## Genetic Distance and Genetic Identity Between Populations

Nei's genetic distance (Nei's  $D$ ) and genetic identity (Nei's  $I$ ) mainly analyzed the genetic relationship between populations, as tabulated in **Table 5**. Nei's  $I$  varied from 0 to 1, Nei's  $D$  ranged from 0 to infinity. The closer the Nei's  $I$  value was to 1, the higher the genetic similarity between the two populations will be; and the larger the Nei's  $D$  value, the greater the genetic distance between the two populations will be (Liu and Mundt, 2020). Among the three groups, Henan/Shanxi had the lowest Nei's  $D$  values, and Henan/Shaanxi had the highest values. On the other hand, Nei's  $I$  value of Henan/Shanxi was the highest, being the closest to 1, while Henan/Shaanxi had the lowest Nei's  $I$  value. Therefore, the genetic similarity between the Henan and Shanxi populations was the highest, and correspondingly the genetic distance was the closest and the communication between them was the most frequent, while the genetic similarity between Henan and Shaanxi, between Shaanxi and Shanxi were relatively lower.



**TABLE 4 |** Pairwise matrices of genetic differentiation coefficient ( $F_{st}$ , below the diagonal) and the number of migrants per generation ( $N_m$ , above the diagonal) among *Setosphaeria turcica* populations in China.

Population	Henan	Shaanxi	Shanxi
Henan		1.3253	6.4885
Shaanxi	0.1581		1.8356
Shanxi	0.0385	0.1208	

**TABLE 5 |** Pairwise matrices of Nei's genetic identify (above the diagonal) and Nei's genetic distance (below the diagonal) among *Setosphaeria turcica* population in China.

Population	Henan	Shaanxi	Shanxi
Henan		0.8635	0.9596
Shaanxi	0.1468		0.8922
Shanxi	0.0413	0.1140	

## Analysis of Molecular Variation

Analysis of molecular variance (AMOVA) revealed that the variation within and between the population accounted for approximately 89.23 and 10.78% of the total variation, respectively (Table 6). Such a result clearly demonstrated that the genetic variation within the populations, rather than between populations, was the major source of total variation.

## Identification of Mating Type and Ratios

The mating types of the 87 *S. turcica* isolates were investigated by using multiplex PCR analysis with mating type-specific primers (Table 7). As a result, 39 isolates were determined as MAT1-1, and the remaining 48 isolates were determined as MAT1-2, accounting for 44.83 and 55.17% of the total isolates, respectively. The frequencies of mating type did not deviate significantly from a 1:1 ratio by the  $\chi^2$  test ( $0.1 \leq p \leq 1.0$ ).

**TABLE 6 |** Analysis of molecular variance (AMOVA) for *Setosphaeria turcica* populations in China.

Source of variation	Degree of freedom	Sum of squares	Variance components	Percent of variation (%)
Among pops	2	61.30	0.82	10.78
Within pops	84	572.26	6.81	89.23
Total	86	633.56	7.64	100.00

**TABLE 7 |** Analysis of molecular variance (AMOVA) for *Setosphaeria turcica* populations in China.

Pop	MAT1-1	MAT1-2	$\chi^2$	P
Henan	15	15	0.000	1.0
Shaanxi	16	11	0.475	0.5
Shanxi	8	22	3.455	0.1
Total	39	48	0.469	0.5

In the Henan population, 15 isolates were MAT1-1 and 15 were MAT1-2, representing a perfect ratio of 1:1. However, this was not the case in the Shaanxi and Shanxi populations, especially in Shanxi population, about 73% of isolates were identified as MAT1-2. Moreover, the isolates of *S. turcica* containing both MAT1-1 and MAT1-2 mating types were not found.

## DISCUSSION

In this study, the population genetic diversity, genetic structure, and distribution of mating type of *S. turcica* isolates sampled from Henan, Shaanxi and Shanxi provinces in China were investigated by using SNP markers and mating type-specific primers. Population genetic diversity is prerequisite for evaluating the status of biological resources. Nei's genetic diversity ( $h$ )

and Shannon's information index (I) are important parameters for measuring biological genetic diversity (Lu et al., 2009). Moderate level of genetic diversity in these three population was found as indicated by Shannon's information index (0.51–0.57), which is somewhat incongruent with previous studies on the genetic diversity of *S. turcica* isolates. Ma et al. (2020a) and Dai et al. (2021), respectively, attained low Shannon's information index of 0.20 to 0.37 and 0.07 to 0.33, indicating low genetic diversity among different populations. Whereas Human et al. (2016) and Nieuwoudt et al. (2018) both achieved high level of genetic diversity with Shannon's information index of 0.87 to 0.90 and 0.66 to 0.94. It is conceivable that genetic diversity may vary among different geographic populations.

The population genetic STRUCTURE and ML phylogenetic tree both separated the three populations into two groups, and the variations among the Henan, Shanxi, and Shaanxi populations is a manifestation that there was no particularly significant connection between genetic groups and the geographic origin of *S. turcica* isolates. Such a finding is consistent with a previous study that SSR groups were not significantly correlated with physiological races or geographic locations (Li et al., 2019). The same was true for studies on the relationship between genetic diversity and geographic regions in *S. turcica* in South Africa (Human et al., 2016), and elsewhere (Gao et al., 2011). However, contrary findings have also been reported (Borchardt et al., 1998). We speculate that the airborne transmission of conidia through wind or other media and the proximity of the three provinces might well be attributable for the observation made in this study that the genetic group is not significantly related to its geographic origin among the *S. turcica* isolates. In addition, the estimation of Nei's genetic identity and genetic distance are reliable parameters for assessing the genetic relationship between populations. The relatively high level of genetic identity (0.9596) between Henan and Shanxi population also illustrated the similarity of the two subgroups, further corroborating the above findings.

Gene flow (Nm) can reveal genetic infiltration between populations, reduce geographic differentiation and cause the homogenization of the genetic characteristics between populations (Rousset, 1997);  $F_{st}$  can be used to clarify the level of genetic variation between populations and is also a vital parameter for understanding the evolutionary history of the population. Hedrick (2005) reported that the larger the value of  $F_{st}$  is, the greater is the level of genetic divergence between populations. In this study, we identified a relatively high level of gene flow (6.2385) of *S. turcica* isolates between Henan and Shanxi, indicating frequent gene exchange between these two populations. The value of  $F_{st}$  between Henan/Shaanxi (0.1573) and Shanxi/Shaanxi (0.1202) was greater than Henan/Shanxi (0.0385), demonstrating relatively moderate genetic divergence between these populations, which was lower than that previously reported by Dai et al. (2021). In addition, the result of the AMOVA analysis demonstrated the genetic variation within populations, rather than those between populations, was the major source of genetic variation, which was consistent with previous reports by Li et al. (2019) and Ma et al. (2020a).

By using type-specific PCR primers, we were able to identify two mating types, including MAT1-1 and MAT1-2, which are present in all the three different geographic populations. Although the distribution of the two mating types in three populations was not exactly equal, especially the Shanxi population had an obviously skewed distribution toward MAT1-2, but on the whole, the total ratio of the MAT1-1 and MAT1-2 was close to 1:1, indicating a positive association between the possibility of sexual reproduction among populations and genetic diversity (Dai et al., 2021). Such a premise seems to be contradictory with the moderate genetic diversity in three *S. turcica* populations as revealed in this study. It is possible that there are limitations in the use of SNP molecular markers for exploring the genetic diversity and population genetic structure of plant pathogens, which mandates cautious interpretations of the findings and warrants further investigation and refinement of the molecular tools used in such study.

There are relatively few studies about the pathogen of *S. turcica* that affects sorghum growth and development, this study represents the first attempt to use SNP molecular markers to analyze genetic diversity and population genetic structure of 87 *S. turcica* isolates from Henan, Shaanxi and Shanxi, the three provinces in China. Our results revealed relatively equal mating type distribution, moderate genetic diversity, and genetic differentiation among the three populations. Although careful interpretation is required for the results, there is no doubt that it will enrich our knowledge about the genetic diversity, population genetic structure of *S. turcica* isolates and provide theoretical bases for the disease management of sorghum NLB in the relevant regions.

## CONCLUSION

In this study, SNP molecular markers were developed to analyze the genetic diversity and population genetic structure of *S. turcica* for the first time. The *S. turcica* populations from three provinces in China are featured with moderate genetic diversity, high levels of gene flow among populations, and relatively equal mating type distribution of MAT1-1 and MAT1-2. The findings have improved our current understanding about the genetic diversity and population genetic structure of *S. turcica* in China.

## DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: NCBI BioProject–PRJNA798196.

## AUTHOR CONTRIBUTIONS

YH designed the study. LC and JD performed all the experiments. JD and LZ analyzed the data and drafted the manuscript. TL

reviewed and edited the manuscript. All authors contributed to the article and approved the submitted version.

## FUNDING

This work was supported by the grant U1804104 of the National Science Foundation of China, the grant CX(21)3022 of the Independent Innovation of Agricultural Sciences in Jiangsu

Province and the grant BK20191236 of Jiangsu Science and Technology Development Program.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2022.853202/full#supplementary-material>

## REFERENCES

- Borchardt, D. S., Welz, H. G., and Geiger, H. H. (1998). Genetic structure of *Setosphaeria turcica* populations in tropical and temperate climates. *Phytopathology* 88, 322–329. doi: 10.1094/PHYTO.1998.88.4.322
- Dai, Y., Gan, L., Lan, C., Lu, X., Yang, X., and Gao, Z. (2021). Genetic differentiation and mixed reproductive strategies in the northern corn leaf blight pathogen *Setosphaeria turcica* from sweet corn in Fujian province. *China. Front. Microbiol.* 12, 1343–1356. doi: 10.3389/fmicb.2021.632575
- Earl, D. A. (2012). Structure Harvester: a website and program for visualizing Structure output and implementing the Evanno method. *Conserv. Genet. Resour.* 4, 359–361. doi: 10.1007/s12686-011-9548-7
- Fan, Y., Ma, J., Gui, X., An, X., Sun, S., and Dong, J. (2007). Distribution of mating types and genetic diversity induced by sexual recombination in *Setosphaeria turcica* in northern China. *Front. Agric. China* 1, 368–376. doi: 10.1007/s11703-007-0062-3
- Ferguson, L. M., and Carson, M. J. P. (2004). Spatial diversity of *Setosphaeria turcica* sampled from the eastern United States. *Phytopathology* 94, 892–900. doi: 10.1094/PHYTO.2004.94.8.892
- Gao, J., Gao, Z., Lv, S., Zhang, X., Zhuang, J., and Zhang, S. (2011). Establishment of UP-PCR system and genetic diversity analysis of *Exserohilum turcicum*. *J. Maize Sci.* 19, 112–117. doi: 10.13597/j.cnki.maize.science.2011.06.002
- Guo, L., Jia, H., Cao, Z., Gu, S., Sun, S., and Dong, J. (2013). Analysis on mating type and parasitic fitness diversity in sexual hybridization offsprings of *Setosphaeria turcica*. *Sci. Agric. Sin.* 46, 4058–4065. doi: 10.3864/j.issn.0578-1752.2013.19.011
- Haasbroek, M. P., Craven, M., Barnes, I., and Crampton, B. G. (2014). Microsatellite and mating type primers for the maize and sorghum pathogen, *Exserohilum turcicum*. *Australas. Plant Pathol.* 43, 577–581. doi: 10.1007/s13313-014-0289-4
- Hartl, D. L., and Clark, A. G. (1997). *Principles of Population Genetics*. 3rd Edn., Sunderland, MA: Sinauer Associates, Inc.
- Hedrick, P. W. (2005). A standardized genetic differentiation measure. *Evolution* 59, 1633–1638. doi: 10.1111/j.0014-3820.2005.tb01814.x
- Hou, X., Fan, Y., Dong, J., and Ma, J. (2006). Identification of physiological races and AFLP analysis in F1 generation of sexual hybridization between isolates of *Setosphaeria turcica*. *Acta Phytotaxonomica Sin.* 28, 1280–1262. doi: 10.1360/yc-006-1280
- Human, M. P., Barnes, I., Craven, M., and Crampton, B. G. (2016). Lack of population structure and mixed reproduction modes in *Exserohilum turcicum* from South Africa. *Phytopathology* 106, 1386–1392. doi: 10.1094/PHYTO-12-15-0311-R
- Jakobsson, M., and Rosenberg, N. A. (2007). CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* 23, 1801–1806. doi: 10.1093/bioinformatics/btm233
- Jimenez, P., Agundez, D., Alia, R., and Gil, L. (1999). Genetic variation in central and marginal populations of *Quercus suber* L. *Silvae Genet.* 48, 278–283.
- Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33, 1870–1874. doi: 10.1093/molbev/msw054
- Leonard, K. J., and Suggs, E. G. (1974). *Setosphaeria proleta*, the ascigerous state of *Exserohilum prolatum*. *Mycologia* 66, 281–297. doi: 10.1080/00275514.1974.12019603
- Li, Y., Jiang, W., Zhang, Q., Ali, E., Ji, P., Pan, H., et al. (2019). Population structure and genetic diversity of *Setosphaeria turcica* from corn in Heilongjiang province. *China. J. Appl. Microbiol.* 127, 1814–1823. doi: 10.1111/jam.14449
- Liu, J., and Mundt, C. C. (2020). Genetic structure and population diversity in the wheat sharp eyespot pathogen *Rhizoctonia cerealis* in the Willamette Valley, Oregon, USA. *Plant Pathol.* 69, 101–111. doi: 10.1111/ppa.13110
- Lu, Q., and Dahlberg, J. A. (2001). Chinese sorghum genetic resources. *Econ. Bot.* 55, 401–425. doi: 10.1007/BF02866563
- Lu, N., Zheng, W., Wang, J., Zhan, G., Huang, L., and Kang, Z. (2009). SSR analysis of population genetic diversity of *Puccinia striiformis* f. sp. *tritici* in Longnan region of Gansu. *China. Chin. Agric. Sci.* 42, 2763–2770. doi: 10.3864/j.issn.0578-1752.2009.08.015
- Luttrell, E. S. (1958). The perfect stage of *Helminthosporium turcicum*. *Phytopathology* 56, 196–287. doi: 10.2307/3756536
- Ma, Z., Liu, B., He, S., and Gao, Z. (2020a). Analysis of physiological races and genetic diversity of *Setosphaeria turcica* (Luttr.) KJ Leonard & Suggs from different regions of China. *Can. J. Plant Pathol.* 42, 396–407. doi: 10.1080/07060661.2019.1679261
- Ma, Y., Liu, T., Liu, B., Gao, L., and Chen, W. (2020b). Population genetic structures of *Puccinia triticina* in five provinces of China. *Eur. J. Plant Pathol.* 156, 1135–1145. doi: 10.1007/s10658-020-01956-4
- Nei, M. (1973). Analysis of gene diversity in subdivided populations. *Proc. Natl. Acad. Sci.* 70, 3321–3323. doi: 10.1073/pnas.70.12.3321
- Nei, M., and Kumar, S. (2000). *Molecular Evolution and Phylogenetics*. England: Oxford university press.
- Nieuwoudt, A., Human, M. P., Craven, M., and Crampton, B. G. (2018). Genetic differentiation in populations of *Exserohilum turcicum* from maize and sorghum in South Africa. *Plant Pathol.* 67, 1483–1491. doi: 10.1111/ppa.12858
- Peakall, R., and Smouse, P. E. (2012). GenAlEx 6.5: genetic analysis in excel. Population genetic software for teaching and research—an update. *Bioinformatics* 28, 2537–2539. doi: 10.1093/bioinformatics/bts460
- Pritchard, J. K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* 155, 945–959. doi: 10.1093/genetics/155.2.945
- Rosenberg, N. A. (2004). DISTRUCT: a program for the graphical display of population structure. *Mol. Ecol. Notes* 4, 137–138. doi: 10.1046/j.1471-8286.2003.00566.x
- Rousset, F. (1997). Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics* 145, 1219–1228. doi: 10.1093/genetics/145.4.1219
- Sree, S. U., Reddy, N., Sabbavarapu, M. M., Madhusudhana, R., Mathur, K., Bhat, V., et al. (2012). Genetic diversity and pathogenic variation in isolates of *Exserohilum turcicum* causing common leaf blight of sorghum. *Indian Phytopathol.* 65, 349–355.
- Tang, L., Gao, Z., Yao, Y., and Liu, X. (2015). Identification and genetic diversity of formae speciales of *Setosphaeria turcica* in China. *Plant Dis.* 99, 482–487. doi: 10.1094/PDIS-06-14-0570-RE
- Tang, L., Xiao, C., and Wang, W. (2012). Research and application progress of SNP markers. *Chin. Agric. Sci. Bull.* 6, 288–293. doi: 10.1007/s11783-011-0280-z
- Yeh, F. C., Yang, R. C., Boyle, T. B., Ye, Z. H., and Mao, J. X. (1997). *Popgene, the User-Friendly Shareware for Population Genetic Analysis*, Canada: Molecular Biology and Biotechnology Centre, University of Alberta.
- Zeng, Y., Xiong, T., Liu, B., Carstens, E., Chen, X., Xu, J., et al. (2021). Genetic diversity and population structure of *Phyllosticta citriarsiana* in China. *Phytopathology* 111, 850–861. doi: 10.1094/PHYTO-12-18-0451-R

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may

be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Cui, Deng, Zhao, Hu and Liu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY).

*The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*