Australian Journal of Crop Science AJCS
AJCS 6(2):342-349 (2012) ISSN:1835-2707

Genetic diversity in the foxtail millet (*Setaria italica*) germplasm as determined by agronomic traits and microsatellite markers

Heng-Sheng Lin, Gwo-Ing Liao, Chih-Yun Chiang, Chang-Sheng Kuoh, Song-Bin Chang*

Department of Life Sciences, Institute of Biodiversity, National Cheng Kung University, Tainan 701, Taiwan (R.O.C.)

***Corresponding author: sbchang@mail.ncku.edu.tw**

Abstract

Setaria italica (L.) P. Beauv. is a model plant that attracts international attention, which is the second most widely cultivated species of millet, especially in East Asia. It has the longest history of cultivation among the millets, having been grown in China since sometime in the sixth millennium BC. It also has been a main crop of the indigenous people of Taiwan for a long time. However, insufficient researches had been conducted about the foxtail millet germplasm in Taiwan. To assess the genetic diversity of millet population, a total of 324 landraces of foxtail millet were collected from around Taiwan, and four years of field researches were conducted for agronomic traits observation. The genetic diversity of the millet population was measured using 33 agronomic traits and 40 microsatellite markers. Average number of alleles (2.4), highly polymorphic information content (PIC) (0.381), observed heterozygosity (0.190) and expected heterozygosity (0.354) were shown. Thirty-five SSR markers showed significant deviation from Hardy-Weinberg equilibrium in 324 landraces and all these markers had low null allele frequencies. Analysis with the Unweighted Pair Group Method with Arithmetic Mean (UPGMA) method and Principal Component Analysis (PCA) revealed that the 324 landraces could be divided into three groups that coincided with the geographical areas, including northern Taiwan, central Taiwan, and sourthern Taiwan, which account for close relationship between crop distribution and human activities. Agronomical analyses with a plant height of 80.6 to 155.2 cm, spike length of 7.5 to 28.9 cm and growth periods of between 141 and 178 days had showed that there was enough variation to create promising lines for breeding programs. This study not only provides a complete foxtail millet germplasm from Taiwan but also demonstrates that the Taiwanese foxtail millets are very diverse and can be useful for selective breeding of specific traits and in enhancing the genetic base of breeding programs in the future.

Keywords: crop, foxtail millet, landraces, plant, SSR, Taiwan.
Abbreviation: SSR- simple sequence repeats, PCR- polymerase chain reaction, RAPD- random amplified polymorphic DNA, AFLP- amplified fragment length polymorphism, RFLP- restriction fragment length polymorphism. SPSS- statistical product and service solutions.

Introduction

Foxtail millet (*Setaria italica* (L.) Beauv.) is an annual and self-pollinating crop of the *Setaria* genus (Leonard and Martin, 1963). Foxtail millet, not just a cereal of the Old World, is also a crop widely used in Eurasia, Americas, Africa and Australia. With a relatively small genome (515 Mb), foxtail millet is a suitable plant for molecular and genetic research (Wang *et al.*, 1998). It has been adopted as a model organism for providing a deeper understanding of plant biology. The importance of this species is increasing since a project to provide a draft version of its genome was conducted by the Joint Genome Institute (JGI) of the US Department of Energy (Doust et al, 2009). As a modle organism, great amount of studies will be conducted. The familiar model systems, such as Arabidopsis (*Arabidopsis thaliana*), rice, and maize, have large numbers of local strains deposited in stock centers that serve as important resources for the research communities. However, insufficient resources of local foxtail millet landraces have been collected and characterized for genetic variation (Li and Wu, 1996; Reddy *et al.*, 2006; Hirano *et al.*, 2011). In Taiwan, foxtail millet has long been a main crop of the indigenous community. The earliest archeological remains of foxtail millet in Taiwan were found in Nan Kuang Li ruin (23 ° 07′04N, 120 ° 16′05E), approximately 4,500 years ago (Tseng, 2007). Due to its greater drought tolerance and shorter growth period compared with rice, the indigenous people cultivated foxtail millet instead. More than 160 landraces were recorded in the 1960's (http://ttdares.coa.gov.tw/). Foxtail millet was not merely food; the ancient indigenous people cultivated it for various purposes, including for use in festivals and marriages. These varying landraces of foxtail millets have resulted from a long-standing selection for various objectives in different indigenous tribes. Assessing the genetic diversity that exists in the foxtail millet germplasm is important not only for providing genetic resources and aiding in conservation but also for practical applications, such as broadening the genetic base and manipulation for heterosis. Expansion of the genetic base is vital in species in which inbreeding has resulted in a decline in genetic diversity. As reported by Deb (2009), maintaining crop diversity can ensure agricultural sustainability and food security.The relative genetic diversity of a species' population can be determined using morphological and molecular markers. Phenotypic traits are of limited importance because they are frequently affected by environmental conditions and the developmental status of the plant (Tatineni *et al.*, 1996).

However, molecular markers, which are based on DNA sequence polymorphism, are independent of environmental conditions and show higher numbers of polymorphisms. Molecular markers are popular for investigating genetic diversity because of their high efficiency, low sample number requirements, and low number of limitations on the growth stage (Bjorklund *et al.*, 2009). Several molecular markers have been used in foxtail millets, including RFLP (Wang *et al.*, 1998; Fukunaga *et al.*, 2002), RAPD (Schontz and Rether, 1999), and AFLP (Le Thierry D'ennequin *et al.*, 1999). Microsatellites, also called simple sequence repeats (SSRs), are a tandem repeating sequence of 1-6 base pairs of DNA. Microsatellites have been useful for their high number of polymorphisms, high level of variation, abundant information and ease of manipulation; because of these features, they are widely employed in many species (Gupta *et al.*, 1996; Kumar et al., 2009; Shahroodian et al., 2011; Zane *et al.*, 2002). SSRs from foxtail millets from China and India were used by Jia *et al.* (2007; 2009) and Gupta *et al.* (2012). These researches mainly focused on the species origin and the genetic map construction. Recently, a correlation between the geographical area and the genetic structure of foxtail millet landraces using transposons in Japan was shown by Hirano *et al.* (2011). Although there is a wide range of *S. italica* landraces in Taiwan, only a limited number of specimens from Taiwan are available in the international stock centers, which were established to preserve the genetic diversity of plants in the world. The herbarium information on this species is fragmented, and the genetic background of the Taiwanese landraces is still a mystery. It is therefore important to investigate the genetic variation of *S. italica* among the populations found in localized habitats in Taiwan to construct a broader picture of the genetic diversity of this species around the world. There were several aims of this study. (1) To construct a collection of the varying landraces of foxtail millets under cultivation in Taiwan; (2) To characterize 33 agronomic traits of foxtail millet over four years of field research; and (3) To assess the genetic diversity of foxtail millet populations using microsatellites. We also investigated the correlation between geographical areas and the genetic variation of foxtail millet.

Results

Characterization of foxtail millet collection

In total, this germplasm collection consists of 324 landraces of foxtail millet from Taiwan. Economically, all the foxtail millet landraces were not competing with commercial cultivars. The landraces collected in this study were cultivated only by the farmers for daily use. Biologically, all of the landraces are representatives of the indigenous tribes. Furthermore, these landraces were developed independently and separated by geographical barriers, such as rivers, valleys and mountains. Geographically, the most elevated collection site was at an altitude of 1582 m (24˚5809′ N, 121˚3027′ E). The habitats of these landraces included diverse climates (from 8˚C to 35˚C) and different types of soil texture (ranging from sand to loam). The foxtail millet landraces collected in this study is diverse and can be used for further understanding of foxtail millet breeding.

Agronomic trait observation and UPGMA clustering analysis

For investigating the agronomic traits, thirty-three quantitative and qualitative traits were recorded and analyzed for 324 foxtail millet landraces (supplementary Table S1). During the agronomic trait investigation, the divergences among 324 foxtail millet landraces and two out-group *Setaria* species were analyzed using the UPGMA method (CHINA and ARES). In these samples, the millets could be divided into three clusters (Fig. 1). In these three clusters, they had different features for classification. In each cluster, the landraces have the quantitative traits with similar pattern. On the other hand, clustering results of qualitative traits show more divergent among three clusters, which were described as the following. In cluster A, three spike types accounted for a different proportion of 3.3% (with many branches), 40% (with branches) and 56.7% (without branches). Most landraces in this cluster tend to have dark-colored spikes such as dark red (6.67%), coffee (20%) and black (40%), with the exception of a small part of the landraces being light colors including yellow (10%), deep yellow (13.3%) and orange (10%). In regards to the waxy degree of millets in this cluster, 38% of the foxtail millet landraces have non-waxy grains and 62% of the landraces have waxy millets. Leaf shape of the landraces in this cluster was all sword shaped. In cluster B, three spike types had accounted for different proportion of 67.24% (with branches) and 32.76% (without branches). The spike colors in these clusters tend to be light color such as yellow (15.51%), deep yellow (6.9%), orange (62.07%), light coral (1.72%) and dark red (13.79%). There were no dark color spikes in this cluster. Speaking to the waxy degree of millets in this cluster, 13.8% landraces had non-waxy millet and 86.2% landraces had waxy millets. Leaf shape of the landraces of this cluster was all sword shaped except 2 short drooping shape landraces. In cluster C, three spike types had accounted for different proportion of 34.45% (with branches) and 65.55% (without branches). The spike colors in this cluster tend to be light color such as yellow (83.19%), deep yellow (3.36%), and orange (12.61%). In regards to the waxy degree of millets in this cluster, only 3 landraces (1.26%) had non-waxy millet and all the other landraces (98.74%) had waxy millets. The leaf shape of the landraces of this cluster was all sword shaped except 2 long drooping shape landraces. Detailed features of every phenotypic trait of the foxtail millet landraces in each cluster are listed in supplementary Table S1.

Genetic diversity and relationships among 324 foxtail millet landraces using microsatellites

Forty previously developed microsatellite markers were used to detect the genetic diversity of the 324 foxtail millet landraces. The average number of alleles per microsatellite locus was 2.4 (Table 1). The PIC values of these 40 microsatellite loci ranged from 0.117 to 0.672, with an average value of 0.381. As listed in Table 1, the HObs ranged from 0.018 (SITM05) to 0.372 (SITM13) and had an average value of 0.190, while the HExp ranged from 0.112 (SITM43) to 0.574 (SITM75) with an average value of 0.354. According to the results of the test for deviation from Hardy-Weinberg equilibrium, 35 of the 40 microsatellite loci (87.5%) showed a significant deviation ($p < 0.05$), except SITM37, SITM38, SITM41, SITM69 and SITM82. In a null allele frequency analysis, all microsatellites loci had null allele frequencies close to zero.

Genetic relationships among populations: UPGMA clustering analysis and PCA using microsatellites

Based on the UPGMA clustering analysis, we were able to divide the 324 foxtail millet landraces from Taiwan into 3 groups (Fig. 2). Most of the group 1 landraces originated from collection sites in northern Taiwan, including Taoyuan, Hsinchu, Taichung, Yilan and northern Hualien. Most of the

Table 1. Genetic diversity analyses using 40 polymorphic microsatellite primers in 324 foxtail millet landraces.

Locus	Na	PIC	HObs	HExp	HW	NF	Locus	Na	PIC	HObs	HExp	HW	NF
SITM04	3	0.515	0.343	0.459	***	0.021	SITM51	3	0.557	0.231	0.492	***	0.033
SITM05	2	0.369	0.018	0.123	***	0.029	SITM54	2	0.322	0.271	0.495	***	0.023
SITM06	3	0.441	0.354	0.451	***	0.027	SITM57	3	0.587	0.354	0.439	***	0.082
SITM07	2	0.356	0.196	0.255	***	0.029	SITM62	2	0.162	0.092	0.132	***	0.021
SITM08	3	0.527	0.186	0.434	***	0.022	SITM63	2	0.327	0.345	0.518	***	0.019
SITM09	2	0.361	0.031	0.172	***	0.016	SITM64	3	0.563	0.233	0.329	***	0.024
SITM13	3	0.592	0.372	0.445	***	0.013	SITM65	3	0.562	0.265	0.472	***	0.019
SITM14	2	0.312	0.034	0.126	***	0.018	SITM67	2	0.274	0.047	0.493	***	0.017
SITM15	2	0.147	0.023	0.328	***	0.011	SITM68	2	0.337	0.053	0.358	***	0.015
SITM37	2	0.371	0.136	0.239	NS	0.113	SITM69	2	0.362	0.135	0.387	NS	0.013
SITM38	2	0.315	0.279	0.247	NS	0.035	SITM70	3	0.576	0.328	0.453	***	0.008
SITM41	2	0.265	0.179	0.335	NS	0.018	SITM71	2	0.319	0.235	0.528	***	0.015
SITM42	2	0.217	0.226	0.398	***	0.024	SITM72	3	0.534	0.328	0.531	***	0.024
SITM43	2	0.273	0.032	0.112	***	0.028	SITM73	4	0.672	0.026	0.216	***	0.023
SITM44	2	0.238	0.264	0.375	***	0.027	SITM75	3	0.469	0.364	0.547	***	0.023
SITM46	3	0.591	0.023	0.198	***	0.031	SITM76	3	0.438	0.219	0.437	***	0.015
SITM47	3	0.499	0.136	0.298	***	0.027	SITM82	2	0.336	0.091	0.165	NS	0.019
SITM48	2	0.344	0.214	0.473	***	0.029	SITM85	2	0.137	0.316	0.476	***	0.047
SITM49	2	0.349	0.037	0.217	***	0.021	SITM86	2	0.314	0.317	0.433	***	0.029
SITM50	2	0.117	0.103	0.243	***	0.029	GS030	2	0.186	0.157	0.349	***	0.011

PIC: polymorphic information content, HObs: observed heterozygosity, HExp: expected heterozygosity, HW: deviation from Hardy-Weinberg equilibrium, ***: $P < 0.01$, NS: Not significant, NF: null allele frequencies.

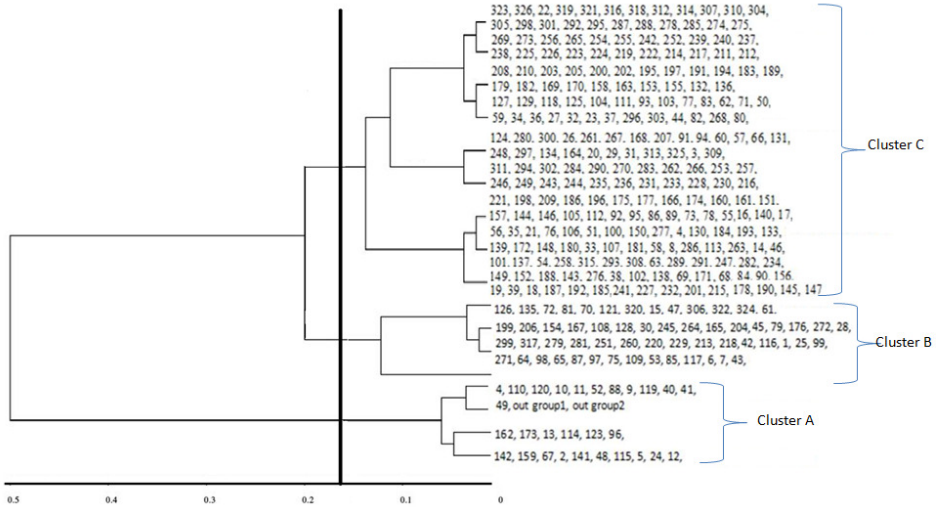

Fig 1. UPGMA phylogenetic analysis of the agronomic traits of foxtail millet landraces in Taiwan. The tree is constructed based on thirty-three agronomic traits. The 324 landraces collected in Taiwan and the two out groups collected from China could be divided into three clusters. The out group 1 is the out group CHINA, and the out group 2 is the out group ARES. The genetic distances of the landrace are demonstrated by the UPGMA dendrogram using Gower's coefficient.

group 2 landraces originated from central Taiwan, including the counties of Nantou, Chiayi, Hualien, Taitung, Taichung, and Kaohsiung. Most of the group 3 landraces originated from southern Taiwan, including Kaohsiung, Pingtung, Taitung, and Lanyu. The out-group *Setaria* landraces were collected from mainland of China. At a genetic similarity coefficient of approximately 0.19, clustering produced three groups. To illustrate the genetic relationships among the 324 landraces, principal component analysis (PCA) was conducted. The results from the PCA analysis based on the microsatellite dataset indicated three major groups (Fig. 3). Three groups of landraces were distinguishable based on the first two PCA components. Foxtail millet landraces in group A were mainly collected from northern Taiwan, including the counties of Taoyuan, Hsinchu, Hualien, and Yilan. The foxtail millet landraces from group B were mainly collected from central Taiwan, including Nantou, Chiayi, Hualien, Taitung, and Taichung, while the foxtail millet landraces in group C were mainly collected from southern Taiwan, including Kaohsiung, Pingtung, Taitung, and Lanyu.

Discussion

Valuable foxtail millet germplasm in Taiwan

The existing foxtail millet collection consists of a limited number of accessions that are derived from every distribution and were chosen to represent the maximum genetic spectrum in the collection (Reddy *et al.*, 2006). To broaden the scope of the available foxtail millet resources, the germplasm should include as much genetic diversity as possible. Foxtail millets were considered to have multiple domestication in many areas worldwide (Doust *et al.*, 2009). In Taiwan, it had been found to be cultivated as early as 4500 years ago (Tseng, 2007). For

research communities worldwide, the foxtail millets in Taiwan were suspected to be close to Chinese ones since the close geographical distributions. However, in the report of Hirano *et al.* (2011), the Taiwanese foxtail millet landraces were clustered together with the landraces of Philippines and Nepal. Additionally, the Taiwanese foxtail millet landraces collected in this study were all cultivated by the indigenous population for a long time without exchanges with the Chinese foxtail millet accessions. Due to the political policies in the past, limited cultural exchanges were happened, leading to the isolation of these landraces. In this study, we provide a collection of foxtail millet germplasm in Taiwan, which has been phenotyped for a large number of traits, including spike color and yield. Although there are varying ranges of quantitative trait in different landrace, the clustering result still showed out similar patterns, which could be seen in that great amount of the landraces in each cluster located in similar quantitative range, including the plant height, spike length, blade length of flag leaf, blade width of flag leaf, stem width, internode length, germination rate, plant survival at maturity, spike weight, grain weight, days to heading, growth periods, percent of threshing, grain number, harvest index, germination period, capacity, ratio of leaf width/length, and flowering period (supplementary Table S1). Although the foxtail millet germplasm showed a comparatively high phenotypic variation for these traits amongst the different landraces, the results of clustering are still obscure. To broaden the scope of clustering, the qualitative traits were observed.

Variable agronomic traits investigation in 324 foxtail millet landraces

For the agronomic trait analysis, the quantitative traits were thought of as features of productivity or performance that could be measured quantitatively. In genetics, a quantitative trait is one in which the inheritance of a phenotypic characteristic can vary in degree and can be affected by the interactions of two or more genes and the environment (Bhattacharjee *et al.*, 2007; Bjorklund *et al.*, 2009). Therefore, the obscure clustering of the quantitative traits in each cluster may be caused by complex genes interaction, which required further study focusing on specific genes. In this study, the qualitative traits were characterized as features affecting performance that cannot be measured quantitatively, including spike type, spike color, seedling color, inflorescence bristle, degree of lodging at maturity, anther color, millet quality, leaf color, blade pubescence, stem pubescence, leaf shape, vein color, insect damage and disease. Based on UPGMA analysis, the clustering results represent a partial correlation to spike color. In cluster A, the landraces tended to have a dark spike color, such as dark red (6.67%), coffee (20%) or black (40%); however, there were some light colored landraces, including yellow (10%), deep yellow (13.3%) and orange (10%). In cluster B, the foxtail millet landraces tended to have medium spike colors, such as yellow (15.51%), deep yellow (6.9%), orange (62.07%), light coral (1.72%) and dark red (13.79%); there were no dark colored spike landraces. In cluster C, the landraces tended to have light spike colors, such as yellow (83.1%), deep yellow (3.36%), orange (12.6%), light coral (0.42%) and dark red (0.42%); few landraces in cluster C had dark colored spikes. Compared to the spike color, the other qualitative traits showed lower levels of variation among the foxtail millet landraces (supplementary Table S1). Therefore, we conclude that the results of UPGMA clustering represent a partial correlation to spike color. According to the reports by Li and Wu (1996), it is possible that the landraces of foxtail millet were formed by natural and man-made selections. They also pointed out that the color of seedlings and grains could be important keys of selection by the farmers due to different extent of difficulties in crop management. For example, purple seedling foxtail millets could made green weeds distinguished and controlled easily in the field (Li and Wu, 1996). In our study, only one landrace in cluster B show out purple seedling color (landrace 30). In fact, phenotypic traits of plants have been reported to be influenced by two major factors: the growth environment and genetics (Bhattacharjee *et al*., 2007; Bjorklund *et al*., 2009). As humidity, temperature, sunshine, and elevation change, the agronomic traits also change. Therefore, molecular evaluation is more suitable for studying the genetic diversity of a species. To further identify the relationship between ethnic populations and foxtail millet, genetic traits must be used.

Genetic diversity using microsatellites in Taiwanese foxtail millets

Although foxtail millet has a limited genetic base, the discovery of large quantity of genomic SSR markers has made it possible to conduct general molecular diversity studies in this important crop to identify genetically diverse germplasm for use in crop improvement. Broadening the genetic base in species is a primary focus because inbreeding has led to a reduction in genetic diversity. Research on the genetic diversity and relationship among foxtail millet landraces is valuable for preserving the germplasm. The number of alleles should be considered when estimating genetic diversity and allow us to compare populations in which the number and distribution of alleles differ greatly. A total of 96 alleles were observed using 40 microsatellite selective markers, and there was an average of 2.4 alleles per locus, suggesting a high number of polymorphism at these microsatellite loci and revealing a high level of genetic diversity in the existing foxtail millet germplasm (Table 1). Comparing to previous studies, Jia *et al.* (2007) had designed 26 EST-SSRs from 1213 EST sequences. They selected twelve foxtail millet accessions and one green foxtail (*Setaria viridis*) accession for polymorphism analysis. The average allele number of four selected SSR markers was 2.5 alleles, which is close to our result. On the other hand, Jia *et al.* (2009) used 37 SSR markers to analyze genetic diversity of 40 foxtail millet accessions. Totally 228 alleles were detected, with an average 6.16 alleles per locus. The value of allele numbers observed in this study located between the results obtained by Jia *et al.* (2007; 2009). These results could possibly be caused by narrower genetic diversity of foxtail millet landraces in Taiwan than the Chinese ones. Comparing to the results of relative species, the average number of alleles found is slightly lower than broomcorn millet (4.91) and rice (3.83) (Cho *et al.*, 2000; Hu *et al.*, 2009). Additionally, in the reports by Jia *et al.* (2009), there is positive correlation between PIC and number of alleles per locus. The PIC value estimates the discriminatory power of a marker. The average PIC value in the 40 polymorphic microsatellites examined for foxtail millet is 0.381 (Table 1). Markers with high PIC values (PIC > 0.5), such as SITM73, SITM13, SITM46, SITM57, SITM70, SITM64, SITM65, SITM51, SITM72, SITM08, and SITM04 can be effectively used in genetic diversity studies in foxtail millet. Among these 40 microsatellite loci, 11 loci (27.5%) showed a high number of polymorphisms (PIC $\geq$ 0.5); 22 loci (55%) had a medium number of polymorphism (0.25 $\leq$ PIC < 0.5), and 7 loci (17.5%) were low polymorphic loci (PIC < 0.25). Compared to the average PIC value in other species, such as pearl millet (0.58) (Kapila *et al.*, 2007), barley (0.50) (Russell *et al.*, 1997), rice (0.55) (Cho *et al.*, 2000) and sorghum (0.56) (Brown *et al.*, 1996), the average PIC value (0.381) in these Taiwanese foxtail millet landraces is lower.

However, it still reveals the high degree of polymorphism in the species in Taiwan. The PIC value detected in other crops may be influenced by cultivar uniformity. The lower PIC value detected in Taiwanese foxtail millet is suspected to be a consequence of the self-pollinating trait of foxtail millets in Taiwan. The landraces in this study were collected from different indigenous tribes around Taiwan. Therefore, low frequencies of gene flow were detected within the foxtail millet species. However, the indigenous cultural interflow and artificial selection in limited geographic areas also increased the degree of variation within the foxtail millet population. In light of this, a high polymorphic content at foxtail millet microsatellites from Taiwanese landraces was detected. Based on the results, the observed heterozygosity (HObs) ranged from 0.018 to 0.372, and the expected heterozygosity (HExp) ranged from 0.112 to 0.574 with average values of 0.190 and 0.354, respectively. The value for HExp was lower than the average heterozygosity value of 126 wild rice individuals (*Oryza glumaepatula*) in Brazil (HExp, 0.67) (De campos *et al.*, 2008). The average HExp value is close to proso millet (0.37) (Cho *et al.*, 2010). The observed heterozygosity (HObs) of Indian foxtail millets ranged from 0.000 to 0.683 with a mean of 0.045 (Gupta *et al.*, 2012). Although foxtail millet is a self-pollinating crop, the heterozygosities detected at some of the loci are higher than the results reported by Gupta *et al.* (2012). This could also be due to man-made selection of outcrossing, isolation by geographical barriers, natural selections of heterosis (Upadhyaya *et al.*, 2008). These results regarding the observed and expected heterozygosity demonstrate the genetic variability of foxtail millet in Taiwan. Based on the results of the Hardy-Weinberg equilibrium deviation test, 87.5% of microsatellite loci showed a significant level of deviation ($p < 0.01$). In fact, the growth environment of foxtail millets collected in Taiwan shows high levels of variation, such as altitude (from 0 m to more than 1500 m) and temperature (from 8°C to 35°C). In addition to the varying growth environments, the collection sites were geographically separated by many different barriers, such as mountains, valleys and rivers. The most important factor to influence the deviation from the Hardy-Weinberg equilibrium in the foxtail millet landraces in this study might be caused by the high self-pollinating rate of foxtail millets. A null allele is any allele that cannot be detected by the assay used to genotype individuals at a particular locus (Pemberton *et al.*, 1995). Generally speaking, when the estimated null allele frequencies are close to zero, the absence of a null allele may occur. In our study, all loci had null allele frequencies below 0.05, revealing that no null allele was detected.

Close correlation between Taiwanese foxtail millets and the indigenous population revealed by UPGMA and PCA

To further understand the relationships among the Taiwanese foxtail millet landraces, we applied UPGMA clustering analysis using microsatellites. A total of 324 foxtail millet landraces in Taiwan were divided into three groups (Fig. 2) that corresponded to the geographical areas from which they were collected: northern, central, and southern Taiwan (Supplementary Fig. S2). Foxtail millet landraces in group 1 were mainly collected from northern Taiwan, including Taoyuan, Hsinchu, Taichung, Yilan, and northern Hualien. The foxtail millet landraces in group 2 were mainly collected from central Taiwan, including Nantou, Chiayi, Hualien, Taitung, Taichung, and Kaohsiung. The foxtail millet landraces in group 3 were mainly collected from southern Taiwan, including Kaohsiung, Pingtung, Taitung, and Lanyu.It is worth to be notice that the results of UPGMA clustering also corresponded to the distribution of the indigenous population. In group 1, the landraces were collected from the northern Atayal (Supplementary Fig. S1). In group 2, the landraces were collected from the Saisiat, southern Atayal, Bunun, Amis, Tsou, Yami and Thao populations. The group 3 landraces were collected from Paiwan, Puyuma, and Ruikai populations (Supplementary Table S5). The northern Atayal mainly populate Taipei and northern Yilan, while the southern Atayal mainly populate southern Yilan, Hualien, Taichung and Nantou. Based on the UPGMA analysis, the foxtail millet landraces collected from the Atayal population were divided into group 1 and 2, which are consistent with their geographical distribution. In group 2, the landraces were collected from the Saisiat, southern Atayal, Bunun, Amis, Tsou, Thao and Yami populations. The genetic relevance of millets collected from southern Atayal is close to the Saisiat millets. According to historical records, the Saisiat had cultural interaction with the Atayal (Masuda, 1942; Tadao, 1946; Hu and Lin, 2003). This finding suggests that cultural exchanges between the Saisiat and the Atayal increased the chances for the dispersal of foxtail millet (Supplementary Fig. S2).The Bunun were the most extenively distributed indigenous group in central Taiwan (Supplementary Fig. S2). Historical data indicate that the Bunun invaded its neighboring indigenous groups, including the southern Atayal, Amis, Tsou and Thao (Utsurikawa, 1935; Ye, 2002). These results coincide with the close relationship among the foxtail millet landraces of the Bunun, Atayal, Amis, Tsou and Thao.According to the UPGMA analysis, the millets collected from Yami were categorized into group 2, which mainly included landraces from central Taiwan. The Yami were distributed on the Lanyu Island, which is separated from the main island of Taiwan. The grouping results from the foxtail millet landraces may be a result of the frequent exchanges between the Yami and Amis through fishing activity (Hsu et al., 2001). In group 3, the landraces were collected from the Paiwan, Puyuma, and Ruikai populations. There was a long standing interaction, migration, and cultural flow between the Paiwan and the other two populations. In fact, the Ruikai and Puyuma were reported to be assimilated into the Paiwan population (Utsurikawa, 1935). The clustering results are consistent with the reports about the three indigenous populations. According to the above analyses, genetic diversity and population structure of foxtail millet landraces in Taiwan had close relationship with the indigenous cultural interflow, which are limited in the three main geographic areas of Taiwan. Although there is high genetic variation among all the landraces of foxtail millet in Taiwan, genetically, the landraces had closer relationship to the landraces within the same group than to the landraces belong to different groups. Using PCA analysis, the foxtail millet landraces in Taiwan were clustered into three groups, which were slightly different than those generated by the UPGMA analysis. The foxtail millet landraces collected from the northern Atayal, southern Atayal and Saisiat were clustered in the same group. In addition, the landraces collected from the Yami were clustered with the landraces from the Paiwan, Puyuma and Ruikai populations. These results are suspected to be caused by different parameters in the UPGMA and PCA analyses. According to the reports by Hirano et al. (2011), they used transposon to divided 425 landraces of foxtail millet worldwide into eight clusters. Also, twenty-one Taiwan foxtail millet landraces in the study were scattered into three clusters. These also demonstrated that the foxtail millet landraces in Taiwan showed out high genetic diversity.

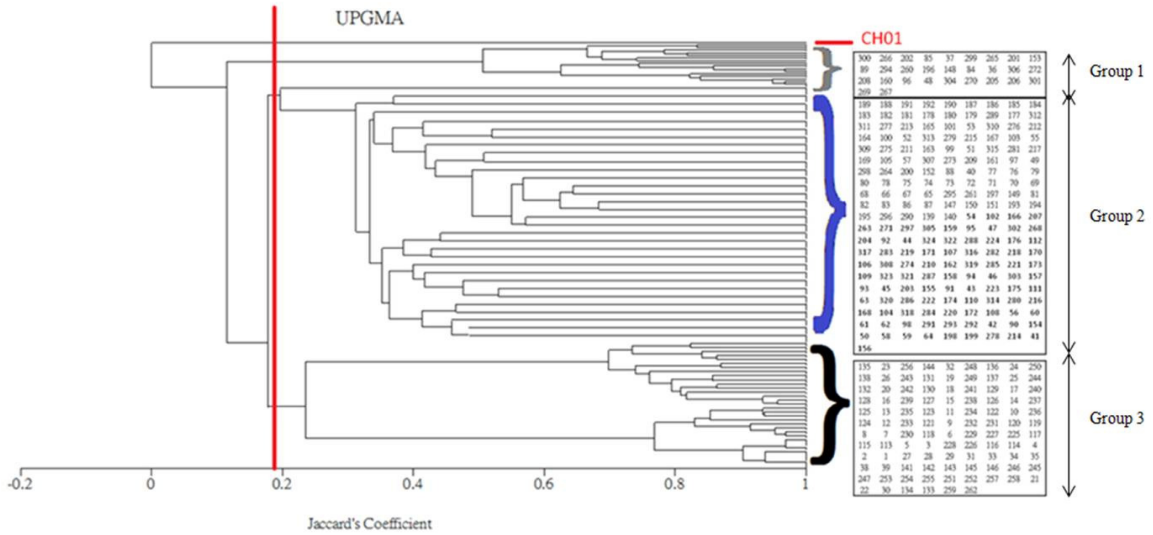

Fig 2. Phylogenetic relationship between the individual foxtail millets in Taiwan based on UPGMA. All 324 foxtail millet landraces were divided into three groups, which coincided with the geographical region in Taiwan (Group1: Northern Taiwan; Group 2: Central Taiwan; Group 3: Sourthern Taiwan). The out group CHO1 represent the out group lndraces "CHINA" and "ARES", which were clustered together. The collection numbers in our laboratory of foxtail millet landraces are shown to the right of the clustering illustration. Clustering could be demonstrated at a Jaccard's coefficient is at 0.19 (vertical line).

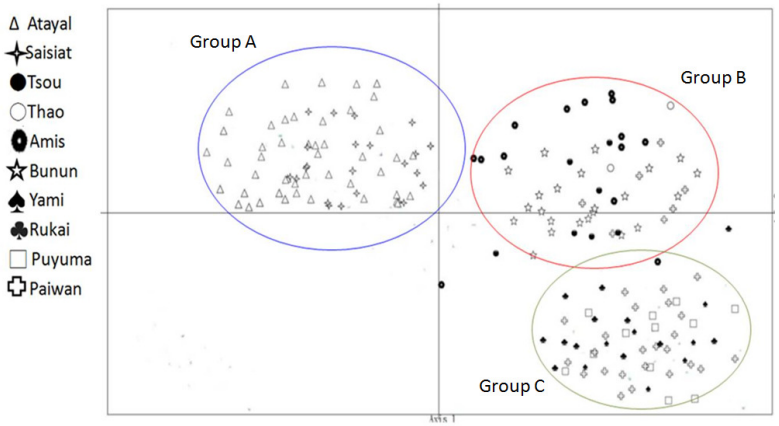

Fig 3. Scatter plot of PCA of 40 microsatellite loci in 324 *Setaria italica* landraces. Scatter plot of the first and second principal components based on the variation of 40 microsatellite loci in 324 *Setaria italica* landraces from around Taiwan that shows three considerable regions that demonstrate the relationship between the foxtail millet landrace distributions. All 324 foxtail millet landraces were divided into three groups, which coincided with the geographical region in Taiwan (Group A: Northern Taiwan; Group B: Central Taiwan; Group C: Sourthern Taiwan). The symbols indicate the populations of the different ethnic indigenous populations.

Materials and methods

Material collection

Different landraces of *Setaria italica* were collected as seed samples in field before harvest from multiple locations in Taiwan. During collection processes, the accessions which exchange occurred through recent past were excluded in our research. All of the vouchers were deposited in the Department of Life Science, National Cheng Kung University, Tainan, Taiwan (Supplementary Table S2). Two out group *Setaria* species include CHINA (*Setaria viridis*) and the ARES (*Setaria verticillata*). A sketch map of the landrace collection sites is provided (Supplementary Fig. S1). Seeds from each landrace were stored dry at -20°C upon collection and were thawed prior to each experiment.

Field experiments and crop management

All 324 collected landraces were cultivated in the fields in Dalin Township, Chaiyi County, Taiwan (23°61′N, 120°49′E). The field research was conducted from January to May in 2007, 2008, 2009 and 2010, with average temperature and humidity of 21.52°C (75.4%), 20.40°C (80.2%), 21.08°C (81.2%) and 21.28°C (76%), respectively. Detailed environment conditions for morphological assessment during the four year experiments were provided (Supplementary Table S3). All of the field experiments were carried out as a randomized complete block design (RCBD) with three replicates. In each plot (3.40 m × 3.40 m), 10 rows were included and 100 seedlings were cultivated at a distance of 30 cm within the row. Border rows with a maize cultivar (*Zea mays* L.) were drill-seeded with a 90

cm row spacing to simulate a more realistic environment around the test plants and to minimize interplot interactions. During the growth process, individual plants were selected randomly, and the agronomic traits were recorded during the growth period (Supplementary Table S4) (Fried and Meister, 1987).

Sampling and measurement methods

During the growth period, 33 agronomic and phenotypic traits were characterized, including plant height, spike length, blade length of flag leaf, blade width of flag leaf, stem width, internode length, germination rate, plant survival at maturity, spike weight, grain weight, days to heading, growth period, percent of threshing, grain number, harvest index, germination period, capacity, ratio of leaf width/length, flowering period, spike type, spike color, seedling color, inflorescence bristle, degree of lodging at maturity, anther color, millet quality, leaf color, blade pubescence, stem pubescence, leaf shape, vein color, insect damage and disease (IBPGR, 1985). The experimental measurements, units and descriptions are characterized in Supplementary Table S4. For each specimen of a foxtail millet landrace in every plot, ten individuals were randomly chosen for agronomic and phenotypic trait observation.

Statistical analysis of agronomic traits

All agronomic traits recorded were characterized quantitatively and reported as the quantitative value (QV) in supplementary Table S1. After standardizing the average of the agronomic traits, we used SAS 9.2 software to conduct cluster analysis (SAS Institute Inc., Cary, NC, USA). A Gower's coefficient was used to calculate the similarity. The genetic distances of the foxtail millet landraces were determined with a dendrogram created using the unweighted pair-group method with the arithmetic means (UPGMA) method from the Gower's coefficient program (NTSYS-PC 2.11, Setauket, New York).

DNA extraction and microsatellite assays

Genomic DNA was isolated from leaf tissue from each individual using a DNeasy Plant Mini Kit (Qiagen, Hilden, Germany). Twenty seeds of each landrace were sown and germinated in sterile plate under 25°C for ten days, and young leaves of ten individuals were cut for DNA isolation. Forty primers were used in the PCR program for the 324 foxtail millet landraces (Lin et al., 2011). The PCR reactions were performed on a MyCyclerTM Thermal Cycler (BIO-RAD, Benicia, California, USA) in 20 µl volumes containing 50 ng of genomic DNA, 200 µM dNTPs, 4 µl of 5X PCR buffer, 1 µM of each forward and reverse primer, and 1 U of Taq DNA polymerase (Promega, Madison, Wisconsin, USA). The PCR temperature profile included an initial denaturation step at 94°C for 3 min, followed by 45 cycles of 1 min at 94°C, 1 min at the annealing temperature specific to each primer (supplementary Table S6), and 2 min at 72°C. A final extension was performed at 72°C for 5 min. The amplified PCR products were resolved on 2% agarose gel and 8% polyacrylamide gel. The gels were then stained with ethidium bromide and visualized under UV light.

Analysis of genetic diversity

Forty microsatellite markers developed in a previous study (supplementary Table S6) were used to analyze the genetic diversity of 324 foxtail millet landraces (Lin et al., 2011). The amplified bands were then scored using a 1/0 (presence/absence) system. The genetic similarities (GS) were estimated using Jaccard's coefficient. Cluster analyses were carried out using the Unweighted Pair Group Method with Arithmetic Mean (UPGMA) clustering algorithm. The similarity matrix and dendrogram were constructed using the NTSYS-PC 2.11 program (Setauket, New York). Principal component analysis (PCA) based on the correlation matrix of the population allelic frequency was performed using SPSS software version 13.1 (SPSS Inc., Chicago, IL, USA).

The number of alleles (Na), expected heterozygosity (HExp), observed heterozygosity (HObs), deviation from the Hardy Weinberg equilibrium (HWE), and the null allele frequency (NF) were quantified using Cervus 3.0 and the online version of GenePop (Kalinowski et al., 2007; Rousset et al., 2008). The parameters used for microsatellite polymorphism evaluation included the number of alleles per locus (Na) and the polymorphic information content (PIC).

$$PIC = 1 - \sum_{i=1}^{n} \mathrm{P}i^2 - 2\left[\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} \mathrm{P}i^2\,\mathrm{P}j^2\right]$$

where Pi is the frequency of the ith allele, and n is the number of alleles) (Botstein, 1980) (Table 1). Based on the amplification profiles for 324 foxtail millet landraces, correlation analyses with the PICs, the number of repeat units, and the number of alleles for the microsatellites were calculated using MVSP software for Windows (MVSP 3.1; Kovach Computing Services, Pentraeth, Wales, UK).

Conclusion

This study collected 324 foxtail millet landraces in Taiwan, which had not been studied completely. Thirty-three agronomic traits were observed during four years of field crops researches. Genetic diversity of these foxtail millet germplasm was performed by forty microsatellite markers and the result had revealed that the foxtail millets in Taiwan are very diverse. In addition, UPGMA and PCA clustering had divided the foxtail millet landraces into three groups which are consistent with three main regions in Taiwan. This study demonstrates that the genetic diversity and population structure of foxtail millet in Taiwan had close relationships with the indigenous migration and cultural interflow. In the future, some landraces with specific traits could be selected for breeders or scientists, which can aid in sustaining agriculture and breeding.

Acknowledgements

We are thankful for all indigenous friends who provide their precious foxtail millet seeds to this study, including Atayal, Saisiat, Bunun, Thao, Tsou, Amis, Puyuma, Rukai and Yami groups in Taiwan. We are also indebted to the anonymous reviewers for their invaluable comments.

Reference

Bhattacharjee R, Khairwal IS, Bramel PJ, Reddy KN (2007) Establishment of a pearl millet [*Pennisetum glaucum* (L.) R. Br.] core collection based on geographical distribution. Euphytica 155: 35-45.

Bjorklund M, Ranta E, Kaitala V, Bach LA, Lundberg P, Stenseth NC (2009) Quantitative Trait Evolution and Environmental Change. PLoS ONE 4: e4521.

Botstein D, White RL, Skolnick M, Davis RW (1980) Construction of a genetic linkage map in man using

restriction fragment length polymorphisms. Am J Hum Genet 32: 314-331.

Brown SM, Hopkins MS, Mitchell SE, Senior ML, Wang TY, Duncan RR, Gonzalez-Candelas F, Kresovich S (1996) Multiple methods for the identification of polymorphic simple sequence repeats (SSRs) in sorghum (*Sorghum bicolor* (L.) Moench). Theor Appl Genet 93: 190-198.

Cho YG, Ishii T, Temnykh S, Chen X, Lipovich L, Mccouch SR, Park WD, Ayres N, Cartinhour S (2000) Diversity of microsatellites derived from genomic libraries and GenBank sequences in rice (*Oryza sativa* L.). Theor Appl Genet 100: 713–722.

Cho YI, Chung JW, Lee GA, Ma KH, Dixit A, Gwag JG, Park YJ (2010) Development and characterization of twenty-five new polymorphic microsatellite markers in proso millet (*Panicum miliaceum* L.). Genes Genom 32: 267-273.

De Campos VAR, De Oliveira Borba TC, Brondani C, Rangel PHN, De Oliveira Camargo GS, De Campos Telles MP, Diniz Filho JAF, Vianello Brondani RP (2008) Genetic analysis of a local population of *Oryza glumaepatula* using SSR markers: implications for management and conservation programs. Genetica 137: 221-231.

Deb D (2009) Valuing Folk Crop Varieties for Agroecology and Food Security, Bioscience Resource Project Commentaries. The Bioscience Resource Project, Inc., USA.

Doust AN, Elizabeth AK, Katrien MD, Jeffery LB (2009) Foxtail Millet: A sequence-driven grass model system. Plant Physiol 149: 137-141.

Fried PM, Meister E (1987) Inheritance of Leaf and Head Resistance of Winter Wheat to Septoria nodorum in a Diallel Cross. Phytopathology 77: 1371-1375.

Fukunaga K, Wang Z, Kato K, Kawase M (2002) Geographical variation of nuclear genome RFLPs and genetic differentiation of foxtail millet, *Setaria italica* (L.) P. Beauv. Genet Resour Crop Ev 49: 95–101.

Gupta PK, Balyan HS, Sharma PC, Ramesh B (1996) Microsatellites in plants: a new class of molecular marker. Curr Sci 70: 45-54.

Gupta S, Kumari K, Sahu PP, Vidapu S, Prasad M (2012) Sequence-based novel genomic microsatellite markers for robust genotyping purposes in foxtail millet [*Setaria italica* (L.) P. Beauv.] Plant Cell Rep 31:323-337.

Hirano R, Naito K, Fukunaga K, Watanabe KN, Ohsawa R, Kawase M (2011). Genetic structure of landraces in foxtail millet (*Setaria italica* (L.) P. Beauv.) revealed with transposon display and interpretation to crop evolution of foxtail millet. Genome 54: 498-506.

Hsu MT, Liao SC, Wu MY (2001). The history of aborigine in Taiwan - The Amis. - The Historical Research Commission of Taiwan Province, Nantou. (In Chinese)

Hu CY, Lin HY (2003). Re-examining the Development of the Nanzhuang Area and the Ethnic Boundary of the Saisiat in Taiwan. Humanitas Taiwanica 59: 177-214. (In Chinese)

Hu X, Wang J, Lu P, Zhang H (2009) Assessment of genetic diversity in broomcorn millet (*Panicum miliaceum* L.) using SSR markers. J Genet Genomics 36: 491-500.

IBPGR (1985) Descriptors for *Setaria italica* and *S. pumila.* International Board for Plant Genetic Resources, Rome, Italy.

Jia X, Shi YS, Song YC, Wang GY, Wang TY, Li Y (2007) Development of EST-SSR in foxtail millet (*Setaria italica*). Genet Resour Crop Ev 54: 233-236.

Jia X, Zhang Z, Liu Y, Zhang C, Shi Y, Song Y, Wang T, Li Y (2009) Development and genetic mapping of SSR markers in foxtail millet [*Setaria italica* (L.) P. Beauv.] Theor Appl Genet 118: 821-829.

Kalinowski ST, Taper ML, Marshall TC (2007) Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. Mol Ecol 16: 1099–1106.

Kapila RK, Yadav RS, Plaha P, Rai KN,Yadav OP, Hash CT, Howarth CJ (2007) Genetic diversity among pearl millet maintainers using microsatellite markers. Plant Breed 127: 33-37.

Kumar P, Gupta VK, Misra AK, Modi DR,Pandey BK (2009) Potential of Molecular Markers in Plant Biotechnology. Plant OMIC 2:141-162.

Le Thierry D'ennequin M, Panaud O, Toupance B (2000) Assessment of genetic relationships between *Setaria italica* and its wild relatives *S. viridis* using AFLP marker. Theor Appl Genet 100: 1061-1066.

Leonard WH, Martin JH (1963) Cereal crops, The Macmillan Company, New York.

Li Y, Wu SZ (1996) Traditional maintenance and multiplication of foxtail millet (*Setaria italica* (L) P. Beauv.) landraces in China. Euphytica 87: 33–38.

Lin HS, Chiang CY, Chang SB, Kuoh CS (2011) Development of Simple Sequence Repeats (SSR) Markers in *Setaria italica* (Poaceae) and Cross-Amplification in Related Species. Int J Mol Sci 12: 7835-7845.

Masuda F (1942) Research of the criminal system of the aboriginal people, SMC Publishing, Taipei.

Pemberton JM, Slate J, Bancroft DR, Barrett JA (1995) Nonamplifying alleles at microsatellite loci: a caution for parentage and population studies. Mol Ecol 7: 249–252.

Reddy VG, Upadhyaya HD, Gowda CLL (2006) Characterization of world's foxtail millet germplasm collections for morphological traits. J SAT Agric Res 2: 1–3.

Rousset F (2008) Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux. Mol Ecol Resour 8: 103-106.

Russell J, Fuller J, Young G, Tomas B, Taramino G, Macaulay M, Waugh R, Powell W (1997) Discriminating between barley genotypes using microsatellite markers. Genome 40: 442–450.

Schontz D, Rether B (1999) Genetic variability in foxtail millet, *Setaria italica* (L.) P. Beauv: identification and classification of lines with RAPD markers. Plant Breed 118: 190-192.

Shahroodian SH, Azadfar D, Soltanloo H, Ramezanpour SS (2011) Genetic variability in natural Iranian populations of *Cupressus sempervirens* var. horizontalis in Caspian Sea coastward assessed by SSR markers. Plant OMIC 4:19-24.

Tadao K (1946) The regional differences of Taiwanese aboriginal by Han's impact. Colonization Forum 3: 331-356.

Tatineni V, Cantrell RG, Davis DD (1996) Genetic diversity in elite cotton germplasm determined by morphological characteristics and RAPDs. Crop Sci 36: 186-192.

Tseng CH (2007) Understanding Taiwan history by archeology, Online paper of Taiwan Monthly, Taiwan Provicial Government, Taiwan. (In Chinese)

Upadhyaya HD, Dwivedi SL, Baum M, Varshney RK, Udupa SM, Gowda CLL (2008). Genetic structure, diversity, and allelic richness in composite collection and reference set in chickpea (*Cicer arietinum* L.) BMC Plant Biol 8: 106.

Utsurikawa N (1935) Taiwan's aboriginal systems, Taihoku Teikoku Daigaku, Taipei. (In Chinese)

Wang ZM, Devos KM, Liu CJ, Wang RQ, Gale MD (1998) Construction of RFLP-based maps of foxtail millet, *Setaria italica* (L.) P. Beauv. Theor Appl Genet 96: 31-36.

Ye JN (2002) Aboriginal history history of the Bunun, Taiwan Historica, SMC publishing, Taipei, Taiwan. (In Chinese)

Zane L, Bargelloni L, Patarnello T (2002) Strategies for microsatellite isolation: a review. Mol Ecol 11: 1–16.