

Article

Analysis of Genetic Diversity and Population Structure of Sesame Accessions from Africa and Asia as Major Centers of Its Cultivation

Komivi Dossa^{1,2}, Xin Wei¹, Yanxin Zhang¹, Daniel Fonceka^{2,3}, Wenjuan Yang¹, Diaga Diouf⁴, Boshou Liao¹, Ndiaga Cissé^{2,*} and Xiurong Zhang^{1,*}

¹ Key Laboratory of Biology and Genetic Improvement of Oil Crops, Ministry of Agriculture, Oil Crops Research Institute of the Chinese Academy of Agricultural Sciences, No. 2 Xudong 2nd Road, 430062 Wuhan, Hubei, China; dossakomivi@gmail.com (K.D.); weixin@caas.cn (X.W.); zhangyanxin@caas.cn (Y.Z.); yangwenjuanqau@163.com (W.Y.); lboshou@hotmail.com (B.L.)

² Centre d'Etudes Régional pour l'Amélioration de l'Adaptation à la Sécheresse (CERAAS), BP 3320 Route de Khombole, Thiès 21000, Senegal; daniel.fonceka@cirad.fr

³ Centre de coopération internationale en recherche agronomique pour le développement (CIRAD), UMR AGAP, F-34398 Montpellier, France

⁴ Laboratoire Campus de Biotechnologies Végétales, Département de Biologie Végétale, Faculté des Sciences et Techniques, Université Cheikh Anta Diop, BP 5005 Dakar-Fann, Dakar 107000, Senegal; diaga.diouf@ucad.edu.sn

* Correspondence: cissendiaga02@hotmail.com (N.C.); zhangxr@oilcrops.cn (X.Z.); Tel.: +221-33-9514993 (N.C.); +86-27-86811836 (X.Z.); Fax: +221-33-9514995 (N.C.); +86-27-868118356 (X.Z.)

Academic Editors: Sean Mayes, Festo Massawe, Prakrit Somta and Wai Kuan Ho
Received: 10 December 2015; Accepted: 30 March 2016; Published: 12 April 2016

Abstract: Sesame is an important oil crop widely cultivated in Africa and Asia. Understanding the genetic diversity of accessions from these continents is critical to designing breeding methods and for additional collection of sesame germplasm. To determine the genetic diversity in relation to geographical regions, 96 sesame accessions collected from 22 countries distributed over six geographic regions in Africa and Asia were genotyped using 33 polymorphic SSR markers. Large genetic variability was found within the germplasm collection. The total number of alleles was 137, averaging 4.15 alleles per locus. The accessions from Asia displayed more diversity than those from Africa. Accessions from Southern Asia (SAs), Eastern Asia (EAs), and Western Africa (Waf) were highly diversified, while those from Western Asia (WAs), Northern Africa (NAf), and Southeastern Africa (SAf) had the lowest diversity. The analysis of molecular variance revealed that more than 44% of the genetic variance was due to diversity among geographic regions. Five subpopulations, including three in Asia and two in Africa, were cross-identified through phylogenetic, PCA, and STRUCTURE analyses. Most accessions clustered in the same population based on their geographical origins. Our results provide technical guidance for efficient management of sesame genetic resources in breeding programs and further collection of sesame germplasm from these different regions.

Keywords: *Sesamum indicum*; genetic diversity; geographic origin; Asia; Africa; SSR

1. Introduction

Sesame (*Sesamum indicum* L.) has been described as one of the oldest oilseed plants used by humans [1,2]. It is a diploid species with $2n = 2x = 26$ chromosomes that belongs to the family of Pedaliaceae, *Sesamum* genus, and is the most commonly cultivated edible oil crop species out of over 30 species in this genus [3,4]. It is predominantly considered as a self-pollinated plant, although low percentage of cross pollination is reported [5].

Sesame seeds have the highest oil content (~55%) among major oilseeds [6]. Beyond the traditional use for direct consumption, sesame seeds are also used as an active ingredient in antiseptics, bactericides, viricides, disinfectants, moth repellants, and antitubercular agents because they contain natural antioxidants such as sesamin, sesamol, and sesamolol [7,8]. The potentially beneficial effects of sesame on human health have recently renewed interest in this ancient crop [9].

Africa and Asia account for more than 96% of the world sesame seed production, with India, Sudan, and China as leading producer countries [10]. Sesame is thus a major crop in these continents, with a large number of varieties [11–14].

The origin of cultivated sesame and the process underlying its domestication have been controversial, though Africa and Asia were identified as probable origins [3,15–20]. Later evidences from inter-specific hybridization, molecular, and lignans analyses suggested the Indian subcontinent as the center of domestication and the probable wild progenitor is believed to be *Sesamum malabaricum* Burm [13,21].

Genetic diversity of crops plays an important role in sustainable development and food security [22], as it allows adaptation to various biotic and abiotic stresses. The African and Asian continents have been projected to be the hardest hit by the negative effects of climate change on crop production [23], which will result in large parts of these continents becoming drier and/or suffering floods. Currently, breeding efforts in sesame have been aimed at the production of white seeded varieties with high market value, shorter duration, and inbuilt tolerance to drought and high temperatures. In addition, attributes such as resistance to phyllody, higher yields, improved plant architecture, and indehiscent capsules have also been incorporated [14].

Knowledge of the genetic diversity and population structure of germplasm collections is an important foundation for crop improvement and a key component of effective conservation and breeding strategies [24]. It is therefore important to assess in detail the genetic diversity of large accessions of cultivated sesame in these continents for proper and efficient utilization in breeding and also to make inferences about the history of its domestication.

Molecular characterization is now the favorite way to quantify variation within germplasm samples. Many studies have dissected the genetic diversity of sesame accessions using different types of molecular markers (Random Amplified Polymorphic DNA (RAPD) [25–27], Inter-simple Sequence Repeats (ISSR) [28], Amplified Fragment Length Polymorphism (AFLP) [9], Sequence-Related Amplified Polymorphism (SRAP) [29], Expressed Sequence Tags-SSR (EST-SSR) [30], Simple Sequence Repeats (SSR) [31–33], or a combination of molecular markers [29,34,35]). However, few studies have been conducted by using large samples of sesame accessions from Africa and Asia, which are the major centers of sesame cultivation. There, is therefore, the need to conduct large-scale genetic diversity analysis among the accessions of sesame from these two continents that have been frequently involved in the exchange of germplasm.

In this study, we examined the genetic diversity of a representative set of sesame accessions from 22 countries of Africa and Asia. The overall goal is to provide a comprehensive insight into sesame diversity between and within these continents for a better management of the genetic resources used in breeding programs.

2. Experimental Section

2.1. Plant Materials

A total of 96 accessions of cultivated sesame (*Sesamum indicum* L., $2n = 2x = 26$), comprised of landraces and modern cultivars representing geographically and phenotypically wide variation, were used in this study. Among these accessions, 48 were collected in nine African countries and 48 in 13 Asian countries (see Table 1 and Table S1 in the Supplementary materials). All samples from Asia and some samples from Africa were obtained from a large collection of sesame accessions preserved at the China National Genebank, Oilcrops Research Institute, Chinese Academy of Agricultural Sciences.

These samples have been purified through self-pollination for generations. The rest of the accessions were personally collected by the first author. The field experiment was conducted at the Oilcrops Research Institute located in Wuhan, Hubei province (China).

Table 1. Summary of accessions used in the study.

Continents	Geographical Regions	Number of Countries	Number of Accessions
Asia	West Asia (WAs)	2	3
	East Asia (EAs)	5	32
	South Asia (SAs)	7	13
	Total	14	48
Africa	West Africa (Waf)	5	40
	North Africa (NAf)	2	3
	Southeast Africa (SAf)	2	5
	Total	9	48

2.2. DNA Extraction

Leaves from 10 bulked three-week-old seedlings per accession were used for DNA isolation according to the Cetyltrimethyl Ammonium Bromide (CTAB) method [34]. Moreover, to reveal the extent of the intra-accession variability of some accessions from West Africa, 288 individuals from 24 accessions were analyzed using a single plant DNA extraction approach. DNA quality and quantity were assessed by spectrophotometry (NanoDrop 2000, Thermo Scientific, Wilmington, DE, USA). DNA samples were stored at $-20\text{ }^{\circ}\text{C}$, for further use.

2.3. PCR and Electrophoresis

SSR polymorphism screening was first performed using 4000 candidate markers and six accessions (three accessions from Africa and three from Asia). A total of 33 highly polymorphic SSR markers [33] providing coverage across all the 16 Linkage Groups (LG) reported in the first draft of the sesame genome [36] were selected to scan for polymorphism between the accessions (Figure 1, Table S2). Polymerase Chain Reaction (PCR) with the SSR primers was performed in a total volume of 15 μL containing 30 ng of DNA, 1 pmol of each primer, 0.2U Taq DNA polymerase, and 2x Reaction Mix (Tiangen Biotech, Beijing, China) supplied together with the dNTPs and MgCl_2 . All PCRs were conducted in 96-well plates in a S-1000 Thermal Cycler (Bio-Rad, Hercules, CA, USA). The PCR cycles were $94\text{ }^{\circ}\text{C}$ (5 min), 35 cycles of $94\text{ }^{\circ}\text{C}$ (30 s), $55\text{ }^{\circ}\text{C}$ (30 s), $72\text{ }^{\circ}\text{C}$ (30 s), followed by the extension step for 5 min at $72\text{ }^{\circ}\text{C}$. The amplified products were separated in 6% denaturing polyacrylamide gel and visualized by silver staining as described by [34].

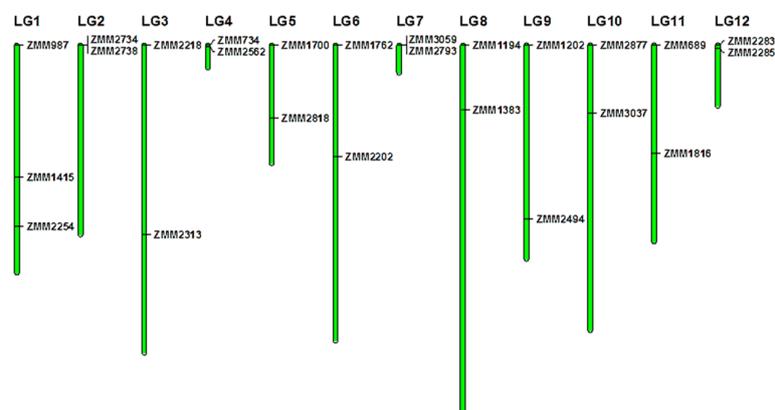


Figure 1. Sesame genome-wide distribution of the 33 polymorphic markers used in this study. LG1–LG16 represent the identified Linkage Groups of the sesame genome [36]. Green bars represent linkage groups. Black lines indicate the locations of primers on linkage groups.

2.4. Scoring and Data Analysis

For each locus across the genotypes, allele scoring was done manually based on the presence of a particular size allele in each of the germplasm samples. Presence was denoted as “1” and absence of an allele as “0.” For variability analysis within accessions, only two individuals from the same accession (K1712) displayed different alleles from the 10 others at one marker. Thus, further analyses focused on the bulked samples of the 96 accessions, which were grouped into six populations according to their geographical origins as indicated in Table 1. The number of alleles (N_a), Effective number of alleles (N_e), Nei’s Gene Diversity (H_e), Observed heterozygosity (H_o), and Shannon’s Information Index (I) were estimated using POPGENE version 1.32 (University of Alberta, Edmonton, Canada) [37]. Major Allele Frequency (MAF), Number of private alleles (N_p), and Polymorphic Information Content (PIC) were calculated with the software PowerMarker version 3.25 (NC State University, Raleigh, NC, USA) [38]. In addition, Analysis of Molecular Variance (AMOVA) was done using GENALEX 6.4 (The Australian National University, Canberra, Australia) [39], in order to estimate the genetic structure between and among geographical regions and continents. Since sesame populations from North Africa, West Asia, and Southeast Africa have very low sample sizes compared to other populations, they were not considered in the AMOVA analysis for geographical regions in order to avoid bias in the analyses. However, all accessions, according to their continent of origin (Asia or Africa), were included in the AMOVA analysis for continents. The significance of variance components was tested by permuting the DNA marker data 999 times. The 33 markers used in this study were mapped onto the 16 Linkage Groups (LGs) of sesame genome according to their physical positions using MapChart 2.3 (Wageningen UR, Wageningen, Netherlands) [40]. To identify the pair-wise genetic relationships between the 96 accessions, a genetic distance matrix was calculated with GENALEX 6.4 [39]. Principal component analysis (PCA) based on genotype data of SSR markers was performed using GENALEX 6.4.

A Neighbor-Joining (NJ) tree based on Nei’s genetic distance [41] was also drawn in MEGA version 6.06 (Temple University, Philadelphia, PA, USA) [42]. Additionally, the population structure was inferred using the Bayesian clustering method implemented in the program STRUCTURE 2.2 (Stanford University, Stanford, CA, USA) [43]. The software was run with the admixture model and correlated allele frequencies. Five runs were performed for each k (1 to 10) representing the number of clusters considered. The burn-in number and iterations for each run were both set to 100,000 and the true k was determined according to the method described by [44]. Sesame accessions with membership probabilities ≥ 0.60 were assigned to the corresponding subgroup and accessions with membership probabilities < 0.60 were assigned to a mixed subgroup. All accessions were mapped according to their geographical coordinates with ArcGIS software version 9.3 (Esri, Redlands, CA, USA). The geographical coordinates were obtained from the China National Genebank and for some accessions from Africa the originating country geographical coordinates were assigned.

3. Results

3.1. Assessment of the Intra-Accession Variability

A pre-test for assessing the variability within 24 accessions from West Africa with a single plant DNA extraction approach was performed. Out of the 35 markers used, only the marker ZMM1522 was polymorphic (2.86%), with two individuals among the 12 from the accession K1712 exhibiting different alleles. This result suggests that the genotyped accessions were relatively homozygous.

Two markers including the marker ZMM1522 from the 35 markers were excluded and the 33 marker-genotype combinations that displayed clear bands were retained to analyze the bulked samples of the 96 accessions.

3.2. SSR Polymorphism in the Sesame Accessions

Thirty-three polymorphic SSRs were used to assess the genetic diversity in a sesame panel, with at least two markers per linkage group. A total of 137 alleles among the 96 sesame accessions were observed (Table S2). Number of alleles observed ranged from two (ZMM294, ZMM2202, ZMM2313, ZMM2321, ZMM2356, ZMM2734, ZMM2738) to 10 (ZMM1762), with an average of 4.15 alleles per locus. Major allele frequency (MAF) average was 0.59 and PIC average was 0.45 (Table 2).

3.3. Allele Variation among Geographical Regions

Diversity indices varied greatly between the six geographical regions. Among them, the highest MAF (0.7) was observed in North Africa (NAf), whereas quite similar MAFs were found in the five other groups (Table 2). The analysis of allelic patterns across the geographical regions revealed that accessions from West Africa (Waf) had the largest allele number (4.909). Africa showed a slightly higher mean number of alleles than Asia (5.2 vs. 4.8). Ne values ranged from 1.265 (West Asia) to 1.458 (South Asia). Values of H_e also varied among the different groups and ranged between 0.343 (North Africa) and 0.598 (South Asia). Asia showed relatively higher H_e values than Africa. The Shannon's Information index (I) values ranged from 0.209 (West Asia) to 0.414 (South Asia).

Out of the 137 alleles, 49 (35.77%) were specific to geographical origin. West Africa and East Asia exhibited the highest private allele numbers (17 and 8, respectively), but no private alleles were found in West Asia, which might be due to the low sample number from this area. In general, the data showed that Africa harbored more private alleles (30) than Asia (21).

South Asia showed the highest PIC value (0.5421), whereas North Africa showed the lowest PIC value (0.2828). Generally, accessions from Asia displayed relatively higher values for H_e , I, and PIC than those from Africa.

3.4. Pattern of Genetic Diversity and Phylogenetic Relationships

The genetic distance matrix generated by GENALEX 6.4 was used for Principal Component Analysis of the 96 sesame accessions (Figure 2). The first and second axis, respectively, explained 36.40% and 18.22% of the variance within the molecular data. Populations from West Africa, South Asia, and East Asia were clearly distinguished by PCA analysis (Figure 2).

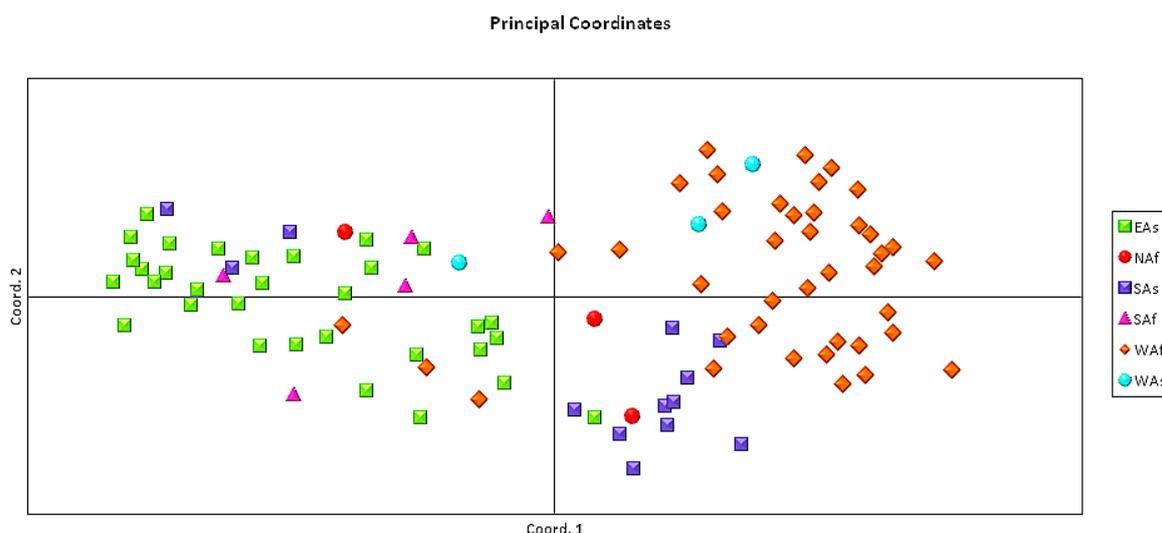


Figure 2. Principal component analysis of the 96 sesame accessions collected from six geographical regions. Each symbol represents one variety from one of the six studied regions.

Table 2. Summary of the genetic diversity of the sesame accessions based on their different geographical regions.

Groups	N	MAF	Na	Ne	He	Ho	I	PIC	Np
Regions									
East Asia	32	0.558 ± 0.138	4.030 ± 2.038	1.439 ± 0.355	0.565 ± 0.130	0	0.396 ± 0.250	0.501 ± 0.146	8
West Asia	3	0.585 ± 0.186	2.242 ± 0.560	1.265 ± 0.382	0.484 ± 0.161	0	0.209 ± 0.301	0.399 ± 0.153	0
South Asia	13	0.505 ± 0.176	3.848 ± 1.543	1.458 ± 0.349	0.598 ± 0.182	0	0.414 ± 0.240	0.542 ± 0.185	5
North Africa	3	0.707 ± 0.246	1.878 ± 0.739	1.385 ± 0.408	0.343 ± 0.261	0.006	0.303 ± 0.320	0.282 ± 0.224	5
Southeastern Africa	5	0.581 ± 0.214	2.787 ± 1.139	1.357 ± 0.377	0.509 ± 0.210	0	0.305 ± 0.297	0.445 ± 0.209	4
West Africa	40	0.578 ± 0.178	4.909 ± 2.350	1.379 ± 0.362	0.550 ± 0.179	0.002	0.346 ± 0.263	0.502 ± 0.178	17
Continents									
Africa	48	0.565 ± 0.184	5.242 ± 2.739	1.407 ± 0.348	0.559 ± 0.181	0.003	0.382 ± 0.235	0.510 ± 0.181	30
Asia	48	0.513 ± 0.145	4.848 ± 2.575	1.4712 ± 0.348	0.604 ± 0.134	0	0.430 ± 0.221	0.543 ± 0.151	21

These statistics are means across the 33 loci.

3.5. Analysis of Molecular Variance

The AMOVA results indicate that 44.66% of the total molecular variation in sesame accessions used in the study was partitioned among geographical groups, and 55.34% was attributed to differentiation within geographical groups (Table 3). In terms of continents, 34.95% of the total molecular variation observed was due to differentiation between Asia and Africa, whereas the rest (65.05%) was due to variance within continents (Table 4).

Table 3. AMOVA between the geographical regions.

Source	df	SS	MS	Est. Var.	%Tv	<i>p</i>
Among geographical regions	2	1103.034	551.517	20.135	44.66	0.001
Within geographical regions	82	2246.107	27.391	24.95	55.34	
Total	84	3349.141		45.085	100	

df = degree of freedom, SS = Sum of Square, MS = Mean Square, Est. Var. = Estimate Variance, %Tv = Percentage of total variation.

Table 4. AMOVA between the continents.

Source	df	SS	MS	Est. Var.	%Tv	<i>p</i>
Among continents	1	615.726	615.726	18.189	34.952	0.001
Within continents	94	3182.003	33.851	33.851	65.048	
Total	95	3797.729		41.999	100	

df = degree of freedom, SS = Sum of Square, MS = Mean Square, Est. Var. = Estimate Variance, %Tv = Percentage of total variation.

3.6. Population Structure

The model-based approach implemented in STRUCTURE was performed to examine the relatedness among the 96 sesame accessions, using the genotypic data for 33 polymorphic SSRs. To facilitate the determination of the exact *k* value corresponding to the genetic groups, the *ad hoc* quantity Δk was used. The highest value of Δk for the 96 sesame accessions was for *k* = 2 (Figure 3).

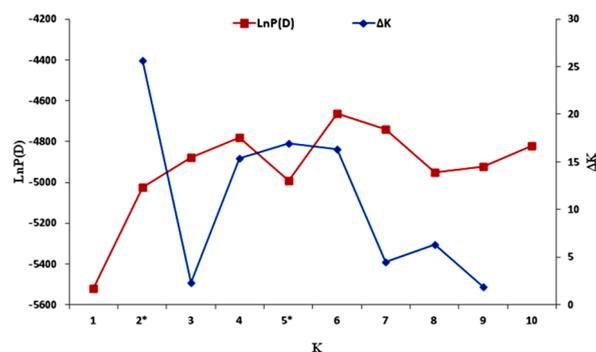


Figure 3. Estimated LnP(D) and Δk of the 96 sesame accessions over five runs for each *K* value.
* Significant *K* values identified.

The results indicated that all the accessions could be classified into two groups, designated as G1 and G2 (Figure 4a). Of the total accessions, 93.75% showed values of probability of membership higher than 0.6 and were therefore classified as members of a particular group, whereas six—representing only 6.25% of accessions—were classified as admixtures with degrees of membership (probability of membership <0.6) shared among the two genetic groups. G1 gathered together 39 accessions, including 31 from Asia, and G2 included in total 51 accessions, with 42 from Africa. The six accessions assigned

to the admixed group include an accession from North Africa, three accessions from South Asia, and two accessions from East Asia (Table S3). It was thus suggested that based on genetic difference, the two major groups might be related to the two continents (Africa and Asia).

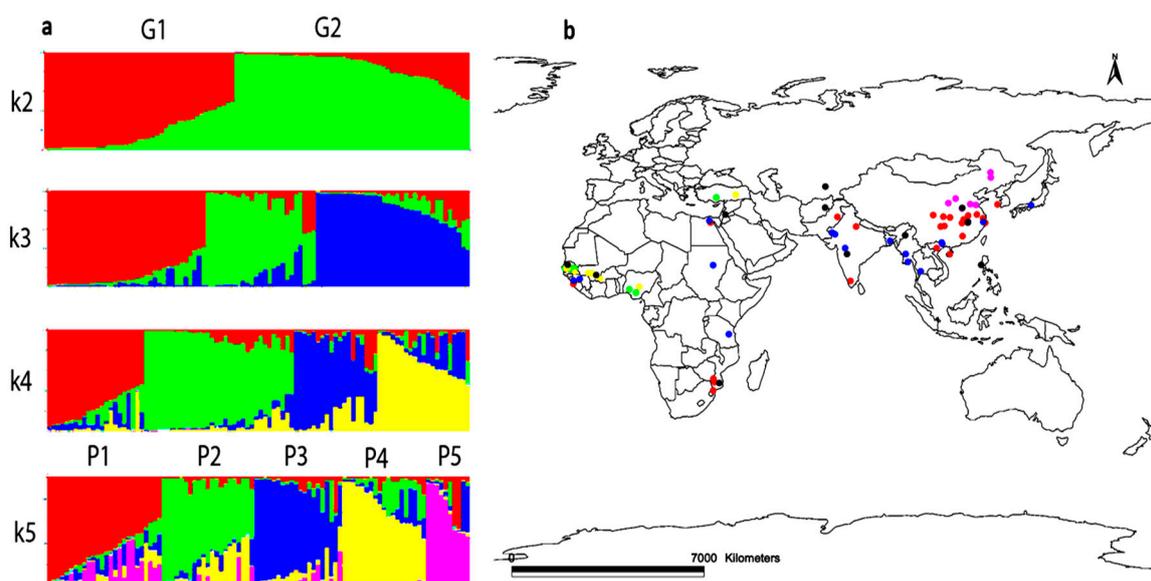


Figure 4. (a) Estimated population structure of sesame accessions assessed by STRUCTURE. Each individual is represented by a thin vertical bar, partitioned into up to k colored segments; (b) mapping the 96 sesame accessions based on their geographical regions. The different colors assigned represent the different groups inferred by STRUCTURE analysis. P1 = Red, P2 = Green, P3 = Blue, P4 = Yellow, P5 = Pink, Pmixed = Black.

In order to better understand the geographical differentiation based on the defined geographical regions, the main groups were further subdivided into five subpopulations (P1–P5) according to STRUCTURE with $k = 5$ the highest after $k = 2$ (Figure 4a). It was observed that two subpopulations, P1 and P5, included most of the accessions from East Asia, whereas the other groups had accessions mainly from South Asia. Most of the West African accessions clustered in subpopulations P2 and P4, while the rest of the accessions from North Africa, Southeastern Africa, and West Asia were clustered among the five subpopulations with no clear pattern. These results might be influenced by the relatively small number of accessions from each of these three regions. Eleven accessions (11.45%) including eight from Asia, were assigned to the admixed population (Pmixed). The classification derived from the STRUCTURE analysis enabled the mapping of all accessions based on their geographical coordinates and their related subpopulations (Figure 4b). Subpopulation P5 was found to be clustered only in North East China, while subpopulation P1 was mostly found to cluster in South East China with some accessions from India, Guinea, Mozambique, and Egypt. Subpopulation P3 was observed to be distributed throughout most South Asian countries and some African countries such as Sudan, Tanzania, Guinea, and Egypt. Both P2 and P4 include accessions mostly from West Africa and Turkey.

4. Discussion

This study analyzed the genetic diversity and population structure of cultivated sesame accessions from Africa and Asia using 33 SSR markers. SSRs have widely been used in genetic diversity and evolutionary analysis in many crops because of their low cost, high polymorphism information content, reproducibility, co-dominant nature, and complete genome coverage [45,46]. The markers used in this study covered the whole genome and can provide a more comprehensive relationship analysis of the samples [33]. The number of alleles detected and the average values of H_e and PIC are higher than

those reported by [32] and lower than reports of [33], where the authors used 150 and 33 worldwide sesame accessions genotyped with 16 and 216 SSR markers, respectively. The differences observed with other studies might be due to the use of different accessions, sampling approaches (bulk *vs.* individuals), and the number of SSR markers. In addition, the bulked sampling approach used in this study may have the tendency to influence the estimates of the diversity indices compared to the individual sampling approach, which is more informative. The genetic diversity observed in African accessions was lower than Asian accessions as shown by He, I PIC, and AMOVA results. These findings are in agreement with the low genetic diversity in African materials previously reported by [25]. In another study [9], sesame accessions from Asia were grouped into four geographical areas but all displayed higher genetic diversity than accessions from Africa. In general, the geographical focus of the higher genetic diversity of a species usually reveals its domestication center [47–49].

The findings from our study are in agreement with the earlier reports of sesame domestication from India [13,21] in the Asian continent.

It was observed in the study that accessions from three regions, namely East Asia, South Asia, and West Africa, had the highest diversity indices, suggesting that these regions contained more genetic diversity of sesame than the other geographical regions. Previous works [25,26] have also reported high genetic diversity of sesame accessions from South Asia and East Asia. However, the southeastern region of Africa was expected to contain more genetic diversity due to its long history of sesame cultivation [9,50,51]. The current study was likely to be influenced by the small sample size from this origin. There is therefore the need to collect more germplasm from Southeastern Africa to confirm whether the genetic diversity of sesame in this region is high or not. Although several genetic diversity studies of worldwide germplasm for sesame have been reported [6,9,25,32], to the best of our knowledge this is the first report on the genetic diversity of a large set of sesame accessions from West Africa, which was largely unexplored. The five subpopulations observed in the current study using both phylogenetic and STRUCTURE analyses are higher than in the report of Cho et al. [32], who found in total three subpopulations including two subpopulations of Korean accessions and one subpopulation comprising worldwide accessions. The higher number of subpopulations observed in the current study might be due to the use of more diverse and representative sesame accessions. Very few studies found a correlation between molecular marker patterns and the geographical origins of sesame, as observed with the accessions from South Asia, East Asia, and West Africa [33,52].

Earlier works in rice and sorghum where the researchers used 1794 worldwide accessions of rice and 3367 accessions of sorghum showed some level of correlation between geographical origins and SSR patterns [53,54].

We observed a close genetic relationship between accessions from East Africa, North Africa, and Guinea in West Africa to the accessions from Asia in the study. This close genetic relationship observed might be due to the introduction of sesame into many countries and material exchange from widely separated locations [28]. For instance, the similarity of Turkish and some West African materials could be explained by the exchange of sesame seeds through the research programs financed by IAEA (International Atomic Energy Agency) for breeding improved sesame cultivars. Moreover, the exchange of plant materials between Asia and East Africa dates back to a long time ago and is still occurring [55], with a steady increase in annual exportation of raw sesame seeds mainly for industrial applications but also for research purposes to China, India, Japan, and other countries. The possibility of crossover events between materials from different locations grown in the same area is high, knowing that cross-pollination in sesame has been reported to occur at a frequency between 5% and 60% [33]. This crossing could explain the similarity of accessions from the eastern part of Africa and Asia. Similar patterns have also been observed by other researchers [9,28,32].

According to [13], sesame was originally domesticated in South India and spread into different areas. In this study, a high proportion of private alleles were observed within each geographical region, with 44% of the variation among accessions being attributed to variation among geographical regions. This suggests that although geographically proximate subpopulations are genetically more similar than

distant ones, differentiation is occurring in each population independently. It is therefore clear that the set of accessions used in this study included diverse accessions and could prove to be a valuable gene pool for allele mining and association mapping for future improvement of the sesame crop.

Finally, contrary to the conclusions of [9], based on the high variation detected within West African, East Asian, and South Asian accessions and their geographical differentiation, the diversity available to breeding programs can be maximized by selecting genotypes from these geographical origins.

A few previous studies have compared sesame's genetic diversity between Asia and Africa, but in those studies the authors did not include a representative set of African materials compared to the Asian samples included. Our results fine-tune the previous knowledge about sesame diversity in both continents using a set of representative accessions. Large sesame germplasm collections are available in the gene banks of many countries such as the United States, South Korea, China, and India [14,34]. However, few sesame germplasm accessions from West Africa have been collected in these gene banks. From this study, there is an indication that new geographically isolated gene pools are evolving in West Africa. Future germplasm collections should focus on West African accessions, which could be exploited in breeding programs mainly oriented towards drought resistance. Moreover, a comparative analysis of the oil content and quality of a large set of sesame from both continents needs to be conducted so as to enable better selection of accessions for breeding purposes.

Supplementary Materials: The following are available online at www.mdpi.com/2073-4425/7/4/14/s1. Table S1: Origin and summary phenotype data of the 96 accessions used in the study; Table S2: Characteristics of the 33 SSR markers used in this study; Table S3: Geographical coordinates and group, subpopulation, and probability of each accession inferred by STRUCTURE.

Acknowledgments: We thank Xiaodong Zhu and Shi-yang Jiang for their laboratory assistance, Roger Bayala for helping to map sesame accessions, and Kyky Ganyo and Bassiaka Ouattara for their helpful suggestions on the manuscript. Finally, we are sincerely grateful to the anonymous reviewers for their helpful comments. This work was funded by the International Advanced Agriculture Science and Technology Introduction program (948 Program, No. 2011-G1 (5)-03), the Agricultural Science and Technology Innovation Project of the Chinese Academy of Agricultural Sciences (CAAS-ASTIP-2013-OCRI), and the *Deutscher Akademischer Austausch Dienst* (DAAD) award.

Author Contributions: Komivi Dossa, Xin Wei, Yanxing Zhang, Ndiaga Cissé, and Xiurong Zhang conceived and designed the experiments; Komivi Dossa, Xin Wei, Yanxing Zhang, Daniel Fonceka, and Yang Wenjuan performed the research and analyzed the data; Komivi Dossa wrote the paper; and Daniel Fonceka, Diaga Diouf, Boshou Liao, and Xiurong Zhang edited and provided critical review of the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Weiss, E.A. Sesame. In *Oilseed Crops*; Longman: London, UK, 1983.
2. Mabberley, D.J. *The Plant-Book: A Portable Dictionary of the Higher Plants*, 2nd ed.; Cambridge University Press: Cambridge, UK, 1997.
3. Nayar, N.M.; Mehra, K.L. Sesame: Its uses, botany, cytogenetics and origin. *Econ. Bot.* **1970**, *24*, 20–31. [[CrossRef](#)]
4. Kobayashi, T.; Kinoshita, M.; Hattori, S.; Ogawa, T.; Tsuboi, Y.; Ishida, M.; Ogawa, S.; Saito, H. Development of the sesame metallic fuel performance code. *Nucl. Technol.* **1990**, *89*, 183–193.
5. Ashri, A. Sesame (*Sesamum indicum* L.). Oilseed Crops. In *Genetics Resources, Chromosome Engineering, and Crop Improvement*; Singh, R.J., Ed.; CRC Press: Boca Raton, FL, USA, 2007; Volume 4, pp. 231–289.
6. Wei, X.; Liu, K.; Zhang, Y.; Feng, Q.; Wang, L.; Zhao, Y.; Li, D.; Zhao, Q.; Zhu, X.; Zhu, X.; et al. Genetic discovery for oil production and quality in sesame. *Nat. Commun.* **2015**, *6*. [[CrossRef](#)] [[PubMed](#)]
7. Bedigian, D. Characterization of Sesame (*Sesamum indicum* L.) Germplasm: A Critique. *Genet. Resour. Crop Evol.* **2010**, *57*, 641–647. [[CrossRef](#)]
8. Bedigian, D. Cultivated sesame and wild relatives in the genus *Sesamum* L. In *Sesame: The Genus Sesamum. Medicinal and Aromatic Plants-Industrial Profiles*; Bedigian, D., Ed.; CRC Press, Taylor and Francis Group: Boca Raton, FL, USA, 2011; pp. 33–77.

9. Laurentin, H.E.; Karlovsky, P. Genetic relationship and diversity in a sesame (*Sesamum indicum* L.) germplasm collection using amplified fragment length polymorphism (AFLP). *BMC Genet.* **2006**, *7*. [[CrossRef](#)] [[PubMed](#)]
10. FAOSTAT. Food and Agriculture Organization statistical databases 2013. Available online: <http://faostat.fao.org/> (accessed on 29 August 2015).
11. Seegler, C.J.P. *Oil Plants in Ethiopia, Their Taxonomy and Agricultural Significance*; PUDOC: Wageningen, The Netherlands, 1983.
12. Burkill, H.M. *The Useful Plants of West Tropical Africa*, 2nd ed.; RBG Kew: Richmond, UK, 1997; Volume 4, pp. 414–422.
13. Bedigian, D. Evolution of sesame revisited: Domestication, diversity and prospects. *Genet. Resour. Crop Evol.* **2003**, *50*, 779–787.
14. Pathak, N.; Rai, A.K.; Kumari, R.; Thapa, A.; Bhat, K.V. Sesame Crop: An Underexploited Oilseed Holds Tremendous Potential for Enhanced Food Value. *Agric. Sci.* **2014**, *5*, 519–529. [[CrossRef](#)]
15. Hiltebrandt, V.M. Sesame (*Sesamum indicum* L.). *Bull. Appl. Bot. Plant Breed.* **1932**, *2*, 1–114.
16. De Candolle, A. *Origin of Cultivated Plants*, 2nd ed.; Noble Offset Printers: New York, NY, USA, 1886.
17. Vavilov, N.I.; Dorofeev, V.F. *Origin and Geography of Cultivated Plants*; Cambridge University Press: Cambridge, UK, 1992.
18. Nayar, N.M. Sesame, *Sesamum indicum* L. (Pedaliaceae). In *Evolution of Crop Plants*, 2nd ed.; Smartt, J., Simmonds, N.W., Eds.; Wiley: New York, NY, USA, 1995; pp. 404–407.
19. Ashri, A. Sesame breeding. *Plant Breed. Rev.* **1998**, *16*, 179–228.
20. Weiss, E.A. *Oilseed Crops*, 2nd ed.; Blackwell Science: Malden, MA, USA, 2000.
21. Bedigian, D. Sesame in Africa: Origin and dispersals. In *Food, Fuel and Fields—Progress in African Archaeobotany, Africa Prae-Historica*; Neumann, K., Butler, A., Kahlheber, S., Eds.; Heinrich-Barth-Institute: Cologne, Germany, 2003; pp. 17–36.
22. Esquinas-Alcazar, J. Protecting crop genetic diversity for food security: Political, ethical and technical challenges. *Nat. Rev. Genet.* **2005**, *6*, 946–953. [[CrossRef](#)] [[PubMed](#)]
23. Lobell, D.B.; Burke, M.B.; Tebaldi, C.; Mastrandrea, M.D.; Falcon, W.P.; Naylor, R.L. Prioritizing climate change adaptation needs for food security in 2030. *Science* **2008**, *319*, 607–610. [[CrossRef](#)] [[PubMed](#)]
24. Thomson, M.J.; Septiningsih, E.M.; Suwardjo, F.; Santoso, T.J.; Silitonga, T.S.; McCouch, S.R. Genetic diversity analysis of traditional and improved Indonesian rice (*Oryza sativa* L.) germplasm using microsatellite markers. *Theor. Appl. Genet.* **2007**, *114*, 559–568. [[CrossRef](#)] [[PubMed](#)]
25. Bhat, K.V.; Babrekar, P.P.; Lakhanpaul, S. Study of genetic diversity in Indian and exotic sesame (*Sesamum indicum* L.) germplasm using random amplified polymorphic DNA (RAPD) markers. *Euphytica* **1999**, *110*, 21–33. [[CrossRef](#)]
26. Ercan, A.G.; Taskin, M.; Turgut, K. Analysis of genetic diversity in Turkish sesame (*Sesamum indicum* L.) populations using RAPD markers. *Genet. Resour. Crop Evol.* **2004**, *51*, 599–607. [[CrossRef](#)]
27. Pham, T.D.; Bui, T.M.; Werlemark, G.; Bui, T.C.; Merker, A.; Carlsson, A.S. A study of genetic diversity of sesame (*Sesamum indicum* L.) in Vietnam and Cambodia estimated by RAPD markers. *Genet. Resour. Crop Evol.* **2009**, *56*, 679–690. [[CrossRef](#)]
28. Kim, D.H.; Zur, G.; Danin-Poleg, Y.; Lee, S.W.; Shim, K.B.; Kang, C.W.; Kashi, Y. Genetic relationships of sesame germplasm collection as revealed by inter-simple sequence repeats. *Plant Breed.* **2002**, *121*, 259–262. [[CrossRef](#)]
29. Zhang, Y.X.; Zhang, X.R.; Hua, W.; Wang, L.H.; Che, Z. Analysis of genetic diversity among indigenous landraces from sesame (*Sesamum indicum* L.) core collection in China as revealed by SRAP and SSR markers. *Genes Genom.* **2010**, *32*, 207–215. [[CrossRef](#)]
30. Wei, L.; Zhang, H.; Zheng, Y.; Guo, W.; Zhang, T. Development and utilization of EST-derived microsatellites in sesame (*Sesamum indicum* L.). *Acta Agron. Sin.* **2008**, *34*, 2077–2084. [[CrossRef](#)]
31. Dixit, A.; Jin, M.H.; Chung, J.W.; Yu, J.W.; Chung, H.K.; Ma, K.H.; Park, Y.J.; Cho, E.G. Development of polymorphic microsatellite markers in sesame (*Sesamum indicum* L.). *Mol. Ecol. Notes* **2005**, *5*, 736–738. [[CrossRef](#)]
32. Cho, Y.I.; Park, J.H.; Lee, C.W.; Ra, W.H.; Chung, J.W.; Lee, J.R.; Ma, K.H.; Lee, S.Y.; Lee, K.S.; Lee, M.C.; et al. Evaluation of the genetic diversity and population structure of sesame (*Sesamum indicum* L.) using microsatellite markers. *Genes Genom.* **2011**, *33*, 187–195. [[CrossRef](#)]

33. Wei, X.; Wang, L.; Zhang, Y.; Qi, X.; Wang, X.; Ding, X.; Zhang, J.; Zhang, X. Development of Simple Sequence Repeat (SSR) Markers of Sesame (*Sesamum indicum*) from a Genome Survey. *Molecules* **2014**, *19*, 5150–5162. [[CrossRef](#)] [[PubMed](#)]
34. Zhang, Y.X.; Zhang, X.R.; Che, Z.; Wang, L.H.; Wei, W.L.; Li, D.H. Genetic diversity assessment of sesame core collection in China by phenotype and molecular markers and extraction of a mini-core collection. *BMC Genet.* **2012**. [[CrossRef](#)] [[PubMed](#)]
35. Wu, K.; Yang, M.; Liu, H.; Tao, Y.; Mei, J.; Zhao, Y. Genetic analysis and molecular characterization of Chinese sesame (*Sesamum indicum* L.) cultivars using Insertion-Deletion (InDel) and Simple Sequence Repeat (SSR) markers. *BMC Genet.* **2014**. [[CrossRef](#)] [[PubMed](#)]
36. Wang, L.; Yu, S.; Tong, C.; Zhao, Y.; Liu, Y.; Song, C.; Zhang, Y.; Zhang, X.; Wang, Y.; Hua, W.; *et al.* Genome sequencing of the high oil crop sesame provides insight into oil biosynthesis. *Genome Biol.* **2014**. [[CrossRef](#)] [[PubMed](#)]
37. Yeh, F.C.; Boyle, T.J.B. Population genetic analysis of co-dominant and dominant markers and quantitative traits. *Belg. J. Bot.* **1997**, *129*, 156–159.
38. Liu, K.J.; Muse, S.V. PowerMarker, an integrated analysis environment for genetic marker analysis. *Bioinformatics* **2005**, *21*, 2128–2129. [[CrossRef](#)] [[PubMed](#)]
39. Peakall, R.; Smouse, P.E. GENALEX 6: Genetic analysis in Excel: Population genetic software for teaching and research. *Mol. Ecol. Notes* **2006**, *6*, 288–295. [[CrossRef](#)]
40. Voorrips, R.E. MapChart: Software for the graphical presentation of linkage maps and QTLs. *J. Hered.* **2002**, *93*, 77–78. [[CrossRef](#)] [[PubMed](#)]
41. Nei, M. Genetic distance between populations. *Am. Nat.* **1972**, *106*, 283–292. [[CrossRef](#)]
42. Tamura, K.; Stecher, G.; Peterson, D.; Filipowski, A.; Kumar, S. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Mol. Biol. Evol.* **2013**, *30*, 2725–2729. [[CrossRef](#)] [[PubMed](#)]
43. Falush, D.; Stephens, M.; Pritchard, J.K. Inference of population structure, extensions to linked loci and correlated allele frequencies. *Genetics* **2003**, *164*, 1567–1587. [[PubMed](#)]
44. Evanno, G.; Regnaut, S.; Goudet, J. Detecting the number of clusters of individuals using the software structure: A simulation study. *Mol. Ecol.* **2005**, *14*, 2611–2620. [[CrossRef](#)] [[PubMed](#)]
45. Huang, X.Q.; Börner, A.; Röder, M.S.; Ganai, M.W. Assessing genetic diversity of wheat (*Triticum aestivum* L.) germplasm using microsatellite markers. *Theor. Appl. Genet.* **2002**, *105*, 699–707. [[PubMed](#)]
46. Khlestkina, E.K.; Huang, X.Q.; Quenum, F.J.-B.; Chebotar, S.; Roder, M.S.; Börner, A. Genetic diversity in cultivated plants—loss or stability? *Theor. Appl. Genet.* **2004**, *108*, 1466–1472. [[PubMed](#)]
47. Paterson, A. Genetics and Genomics of Cotton. In *Plant Genetics and Genomics: Crops and Models*; Springer-Verlag: New York, NY, USA, 2009; Volume 3, p. 509.
48. Aguilar-Meléndez, A.; Morrell, P.L.; Roose, M.L.; Kim, S.C. Genetic diversity and structure in semi wild and domesticated chiles (*Capsicum annuum*; Solanaceae) from Mexico. *Am. J. Bot.* **2009**, *96*, 1190–1202. [[CrossRef](#)] [[PubMed](#)]
49. Wei, X.; Qiao, W.H.; Chen, Y.T.; Wang, R.S.; Cao, L.R.; Zhang, W.X.; Yuan, N.N.; Li, Z.C.; Zeng, H.L.; Yang, Q.W. Domestication and geographic origin of *Oryza sativa* in China: Insights from multilocus analysis of nucleotide variation of *O. sativa* and *O. rufipogon*. *Mol. Ecol.* **2012**, *21*, 5073–5087. [[CrossRef](#)] [[PubMed](#)]
50. Nyongesa, B.; Were, B.; Gudu, S.; Dangasuk, O.; Onkware, A. Genetic diversity in cultivated sesame (*Sesamum indicum* L.) and related wild species in East Africa. *J. Crop. Sci. Biot.* **2013**, *16*, 9–15. [[CrossRef](#)]
51. Hika, G.; Geleta, N.; Jaleta, Z. Genetic Variability, Heritability and Genetic Advance for the Phenotypic Traits in Sesame (*Sesamum indicum* L.) Populations from Ethiopia. *Sci. Technol. Arts Res. J.* **2015**, *4*, 20–26. [[CrossRef](#)]
52. Ali, G.M.; Yasumoto, S.; Seki-Katsuka, M. Assessment of genetic diversity in sesame (*Sesamum indicum* L.) detected by amplified fragment length polymorphism markers. *Electron. J. Biotechnol.* **2007**, *10*, 12–23. [[CrossRef](#)]
53. Yan, W.; Agrama, H.; Jia, M.; Fjellstrom, R.; McClung, A. Geographic Description of Genetic Diversity and Relationships in the USDA Rice World Collection. *Crop Sci.* **2010**, *50*, 2406–2417. [[CrossRef](#)]

54. Billot, C.; Ramu, P.; Bouchet, S.; Chantereau, J.; Deu, M.; Gardes, L.; Noyer, J.-L.; Rami, J.-F.; Rivallan, R.; Li, Y.; *et al.* Massive Sorghum Collection Genotyped with SSR Markers to Enhance Use of Global Genetic Resources. *PLoS ONE* **2013**, *8*, e59714. [[CrossRef](#)] [[PubMed](#)]
55. Zohary, D.; Hopf, M.; Weiss, E. *Domestication of Plants in the Old World: The Origin and Spread of Domesticated Plants in SOUTHWEST Asia, Europe, and the Mediterranean Basin*; Oxford University Press on Demand: Oxford, UK, 2012.



© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).