

Article

Genetic Diversity in *Jatropha curcas* L. Assessed with SSR and SNP Markers

Juan M. Montes ^{1,*}, Frank Technow ², Matthias Martin ¹ and Klaus Becker ³

¹ JatroSelect GmbH, Echterdingerstr. 30, Stuttgart 70599, Germany;

E-Mail: matthias.martin@jatroselect.com

² Institute of plant breeding, seed science and population genetics, University of Hohenheim,

Fruwirthstr. 21, Stuttgart 70599, Germany; E-Mail: Frank.Technow@uni-hohenheim.de

³ JatroSolutions GmbH, Echterdingerstr. 30, Stuttgart 70599, Germany;

E-Mail: kbecker@uni-hohenheim.de

* Author to whom correspondence should be addressed; E-Mail: juan.montes@jatroselect.com;

Tel.: +49-711-459-99771; Fax: +49-711-459-99789.

Received: 5 June 2014; in revised form: 30 July 2014 / Accepted: 31 July 2014 /

Published: 7 August 2014

Abstract: *Jatropha curcas* L. (*jatropha*) is an undomesticated plant that has recently received great attention for its utilization in biofuel production, rehabilitation of wasteland, and rural development. Knowledge of genetic diversity and marker-trait associations is urgently needed for the design of breeding strategies. The main goal of this study was to assess the genetic structure and diversity in *jatropha* germplasm with co-dominant markers (Simple Sequence Repeats (SSR) and Single Nucleotide Polymorphism (SNP) in a diverse, worldwide, germplasm panel of 70 accessions. We found a high level of homozygosity in the germplasm that does not correspond to the purely outcrossing mating system assumed to be present in *jatropha*. We hypothesize that the prevalent mating system of *jatropha* comprise a high level of self-fertilization and that the outcrossing rate is low. Genetic diversity in accessions from Central America and Mexico was higher than in accession from Africa, Asia, and South America. We identified makers associated with the presence of phorbol esters. We think that the utilization of molecular markers in breeding of *jatropha* will significantly accelerate the development of improved cultivars.

Keywords: *jatropha*; genetic diversity; molecular markers; SSR; SNP; breeding

1. Introduction

Jatropha is an important undomesticated plant, which has received great attention in recent years for its utilization in biodiesel production, rehabilitation of wasteland, and rural development [1]. Jatropha is a small tree or large shrub, which belongs to the Euphorbiaceae family, and has a life expectancy of up to 50 years [2]. The plant has a high potential for greening and rehabilitation of wastelands and the seeds have a high oil concentration with excellent quality for conversion into biodiesel [3]. Although various efforts have been made to develop jatropha as an industrial crop [4,5], the absence of improved cultivars and lack of agronomic knowledge represent the main bottleneck that limits the full exploitation of this plant's potential [6].

Understanding the genetic structure (genetic diversity within and among individuals) of jatropha germplasm is of significant importance to establish breeding strategies and design breeding programs. In jatropha, the genetic diversity among individuals has been investigated with molecular markers in various studies [7]. A consensus seems to exist that the highest genetic diversity is present in germplasm originated from Central America and Mexico. The genetic diversity within individuals has been investigated with co-dominant markers [8,9] and the high level of homozygosity in the germplasm investigated differed significantly from the homozygosity level expected in outcrossing plants [10]. These results trigger questions about the mating system of jatropha.

Jatropha is a monoecious plant with inflorescences having separate female and male flowers and pollen transportation by wind is limited because of its weight and adhesiveness [11]. There is evidence that insect play an important role in pollination of jatropha [12]. In addition, the opening time shift between male and female flowers has been reported and all these observations lead to hypothesize that jatropha should have a high level of heterozygosity as observed in other outcrossing crops [9]. The results of various studies reporting a high level of homozygosity in jatropha point to the revision of the hypothesis about the mating system. Recent results indicate that selfings within an inflorescence and/or within a plant is occurring to a much higher frequency than crosses [13], and the behavior of insect pollinators might influence the mating system significantly.

Phorbol esters are the principal chemical compounds that cause toxicity in the oil and seed cake, and they are not heat liable [14]. There exist non-toxic germplasm from Mexico but the field performance in terms of grain yield is generally low in comparison with the toxic germplasm, as observed in multi-location scientific field experiments (personal communication). This indicates a clearly different adaptation pattern. Therefore, it is important to understand the genetic basis of phorbol esters in order to design breeding strategies to develop high yielding non-toxic cultivars with a broader adaptation.

Molecular markers associated with phorbol esters production have been developed [7], but there is scarce information indicating the number of genes controlling the trait and information of gene action (dominance level). To understand the genetic basis of phorbol esters, it will be necessary to cross toxic and non-toxic germplasm and assess the expression of the trait with molecular markers and measure the expression of the trait in the progeny. A recent study has been reported and it indicates the monogenic and dominant character of phorbol esters biosynthesis [15].

The main goal of this study was to assess the genetic structure and diversity of a large panel of jatropha germplasm with co-dominant markers (Simple Sequence Repeats (SSR) and Single Nucleotide Polymorphism (SNP)). Our objectives were to (i) examine the internal genetic structure

within accessions; (ii) assess polymorphism level among accessions; (iii) estimate the level of molecular variance among and within geographical regions (world region and countries); and (iv) evaluate the potential of markers to classify accessions according to their phorbol esters presence.

2. Experimental Section

The plant material comprised 70 accessions of *Jatropha curcas* L. from different geographical origins (Table 1). The number of plants included in the DNA sample of each accession is indicated as plants in pool (Table 1). Plants in pool are the progeny of a single mother tree. The complete material was examined with 54 SSR markers (100% public) selected to cover the eleven *Jatropha* chromosomes. A subset of 16 accessions was examined with 120 SNP (100% proprietary). Genotyping of SSR and SNP markers was performed by Trait Genetics GmbH (Gatersleben, Germany). Phorbol esters determination was performed as described elsewhere [14]. The range of phorbol ester was 0.0–10.3 mg·g⁻¹.

Table 1. Internal genetic structure in accessions of *Jatropha curcas* L. analyzed with 54 SSR and 120 SNP markers. PE: phorbol esters (P = present; A = absent), Mda: percentage of markers with different alleles at a locus, Msa: percentage of markers with the same allele at a locus, Mis: percentage of missing marker data. Plants in pool refer to the number of plants contained in the DNA sample analyzed. SAM: South America; CNAM: Central and North America.

| ID | World Region | Country | PE | Plants in Pool | SSR | | | SNP | | |
|----|--------------|------------|----|----------------|---------|---------|---------|---------|---------|---------|
| | | | | | Mda (%) | Msa (%) | Mis (%) | Mda (%) | Msa (%) | Mis (%) |
| 1 | AFRICA | Gambia | P | 15 | 5.6 | 94.4 | 0.0 | 0.0 | 100.0 | 0.0 |
| 2 | AFRICA | Madagascar | P | 20 | 5.6 | 94.4 | 0.0 | 0.0 | 100.0 | 0.0 |
| 3 | ASIA | Bangladesh | P | 19 | 5.6 | 94.4 | 0.0 | 0.0 | 100.0 | 0.0 |
| 4 | ASIA | Laos | P | 8 | 5.6 | 94.4 | 0.0 | 0.0 | 100.0 | 0.0 |
| 5 | SAM | Paraguay | P | 18 | 5.6 | 94.4 | 0.0 | 0.0 | 100.0 | 0.0 |
| 6 | SAM | Argentina | P | 18 | 9.3 | 88.9 | 1.8 | 0.0 | 100.0 | 0.0 |
| 7 | CNAM | Mexico | A | 12 | 14.8 | 85.2 | 0.0 | 10.0 | 90.0 | 0.0 |
| 8 | CNAM | Mexico | P | 10 | 5.6 | 94.4 | 0.0 | 0.0 | 100.0 | 0.0 |
| 9 | CNAM | Mexico | P | 19 | 35.2 | 63.0 | 1.8 | 30.0 | 60.0 | 10.0 |
| 10 | CNAM | Mexico | A | 19 | 9.3 | 90.7 | 0.0 | 0.0 | 100.0 | 0.0 |
| 11 | CNAM | Mexico | A | 18 | 9.3 | 90.7 | 0.0 | 0.0 | 100.0 | 0.0 |
| 12 | CNAM | Mexico | A | 17 | 9.3 | 90.7 | 0.0 | 0.0 | 100.0 | 0.0 |
| 13 | CNAM | Mexico | A | 18 | 9.3 | 90.7 | 0.0 | 0.0 | 100.0 | 0.0 |
| 14 | CNAM | Mexico | A | 14 | 9.3 | 90.7 | 0.0 | 0.0 | 100.0 | 0.0 |
| 15 | CNAM | Mexico | A | 9 | 7.4 | 92.6 | 0.0 | 0.0 | 100.0 | 0.0 |
| 16 | CNAM | Mexico | A | 19 | 7.4 | 92.6 | 0.0 | 0.0 | 100.0 | 0.0 |
| 17 | AFRICA | Gambia | P | 6 | 5.6 | 94.4 | 0.0 | | | |
| 18 | ASIA | Laos | P | 6 | 5.6 | 94.4 | 0.0 | | | |
| 19 | SAM | Brasil | P | 6 | 5.6 | 92.6 | 1.8 | | | |
| 20 | CNAM | Mexico | P | 6 | 5.6 | 92.6 | 1.8 | | | |

Table 1. Cont.

| ID | World Region | Country | PE | Plants in Pool | SSR | | | SNP | | |
|----|--------------|------------|----|----------------|---------|---------|---------|---------|---------|---------|
| | | | | | Mda (%) | Msa (%) | Mis (%) | Mda (%) | Msa (%) | Mis (%) |
| 21 | CNAM | Mexico | A | 6 | 9.3 | 85.2 | 5.5 | | | |
| 22 | CNAM | Mexico | A | 6 | 7.4 | 92.6 | 0.0 | | | |
| 23 | CNAM | Mexico | A | 6 | 7.4 | 88.9 | 3.7 | | | |
| 24 | CNAM | Colombia | P | 6 | 5.6 | 83.3 | 11.1 | | | |
| 25 | CNAM | Colombia | P | 6 | 5.6 | 87.0 | 7.4 | | | |
| 26 | ASIA | Laos | P | 6 | 5.6 | 92.6 | 1.8 | | | |
| 27 | CNAM | Colombia | P | 1 | 14.8 | 83.3 | 1.9 | | | |
| 28 | CNAM | Colombia | P | 1 | 9.3 | 90.7 | 0.0 | | | |
| 29 | CNAM | Guatemala | A | 1 | 14.8 | 83.3 | 1.9 | | | |
| 30 | CNAM | Guatemala | A | 1 | 20.4 | 77.8 | 1.8 | | | |
| 31 | CNAM | Colombia | P | 1 | 9.3 | 90.7 | 0.0 | | | |
| 32 | CNAM | Colombia | P | 1 | 9.3 | 90.7 | 0.0 | | | |
| 33 | AFRICA | Gambia | P | 4 | 5.6 | 94.4 | 0.0 | | | |
| 34 | AFRICA | Chad | P | 4 | 5.6 | 94.4 | 0.0 | | | |
| 35 | AFRICA | Cameroon | P | 4 | 5.6 | 94.4 | 0.0 | | | |
| 36 | AFRICA | Madagascar | P | 4 | 5.6 | 94.4 | 0.0 | | | |
| 37 | AFRICA | Madagascar | P | 4 | 5.6 | 94.4 | 0.0 | | | |
| 38 | SAM | Argentina | P | 4 | 5.6 | 94.4 | 0.0 | | | |
| 39 | SAM | Brasil | P | 4 | 5.6 | 94.4 | 0.0 | | | |
| 40 | SAM | Brasil | P | 4 | 5.6 | 94.4 | 0.0 | | | |
| 41 | CNAM | Mexico | A | 4 | 9.3 | 88.9 | 1.8 | | | |
| 42 | CNAM | Mexico | A | 4 | 5.6 | 63.0 | 31.4 | | | |
| 43 | CNAM | Mexico | A | 4 | 5.6 | 74.1 | 20.3 | | | |
| 44 | CNAM | Guatemala | P | 4 | 7.4 | 88.9 | 3.7 | | | |
| 45 | SAM | Argentina | P | 4 | 5.6 | 94.4 | 0.0 | | | |
| 46 | CNAM | Mexico | A | 4 | 11.1 | 87.0 | 1.9 | | | |
| 47 | CNAM | Mexico | A | 4 | 16.7 | 81.5 | 1.8 | | | |
| 48 | AFRICA | CapeVerde | P | 4 | 5.6 | 92.6 | 1.8 | | | |
| 49 | AFRICA | Madagascar | P | 1 | 5.6 | 92.6 | 1.8 | | | |
| 50 | AFRICA | Madagascar | P | 1 | 5.6 | 94.4 | 0.0 | | | |
| 51 | AFRICA | Madagascar | P | 1 | 5.6 | 94.4 | 0.0 | | | |
| 52 | AFRICA | Madagascar | P | 1 | 5.6 | 94.4 | 0.0 | | | |
| 53 | AFRICA | Madagascar | P | 1 | 5.6 | 94.4 | 0.0 | | | |
| 54 | AFRICA | Madagascar | P | 1 | 5.6 | 94.4 | 0.0 | | | |
| 55 | AFRICA | Madagascar | P | 1 | 5.6 | 94.4 | 0.0 | | | |
| 56 | AFRICA | Madagascar | P | 1 | 5.6 | 92.6 | 1.8 | | | |
| 57 | AFRICA | Madagascar | P | 1 | 5.6 | 90.7 | 3.7 | | | |
| 58 | AFRICA | Madagascar | P | 1 | 5.6 | 66.7 | 27.7 | | | |
| 59 | AFRICA | Madagascar | P | 1 | 5.6 | 94.4 | 0.0 | | | |
| 60 | AFRICA | Madagascar | P | 1 | 5.6 | 94.4 | 0.0 | | | |
| 61 | AFRICA | Madagascar | P | 1 | 5.6 | 94.4 | 0.0 | | | |

Table 1. Cont.

| ID | World Region | Country | PE | Plants in Pool | SSR | | | SNP | | |
|------|--------------|------------|----|----------------|---------|---------|---------|---------|---------|---------|
| | | | | | Mda (%) | Msa (%) | Mis (%) | Mda (%) | Msa (%) | Mis (%) |
| 62 | AFRICA | Madagascar | P | 1 | 5.6 | 94.4 | 0.0 | | | |
| 63 | AFRICA | Madagascar | P | 1 | 5.6 | 94.4 | 0.0 | | | |
| 64 | AFRICA | Madagascar | P | 1 | 5.6 | 94.4 | 0.0 | | | |
| 65 | AFRICA | Madagascar | P | 1 | 5.6 | 94.4 | 0.0 | | | |
| 66 | AFRICA | Madagascar | P | 1 | 5.6 | 94.4 | 0.0 | | | |
| 67 | AFRICA | Madagascar | P | 1 | 7.4 | 92.6 | 0.0 | | | |
| 68 | AFRICA | Madagascar | P | 1 | 5.6 | 94.4 | 0.0 | | | |
| 69 | AFRICA | Madagascar | P | 1 | 5.6 | 94.4 | 0.0 | | | |
| 70 | AFRICA | Madagascar | P | 1 | 5.6 | 94.4 | 0.0 | | | |
| Mean | | | | | 7.6 | 90.4 | 2.0 | 0.0 | 100.0 | 0.0 |
| Min | | | | | 5.6 | 63.0 | 0.0 | 0.0 | 60.0 | 0.0 |
| Max | | | | | 35.2 | 94.4 | 31.5 | 30.0 | 100.0 | 10.0 |

Statistical Analysis

For each accession we computed the percentage of markers for which (1) only one allele was present; (2) multiple alleles were present and (3) the percentage of missing markers. This was done separately for SSR and SNP markers. The polymorphic information content (PIC) was calculated for each marker as described by Botstein *et al.* [16]. We computed, for each marker, the percentage of accessions where multiple alleles were present. For each marker allele, we computed the percentage of accessions it was present in (Allele proportion), regardless of whether other alleles of the same marker were present in the accession or not.

Based on the Modified Roger Distances (MRD) between the individuals we performed a principal component analysis (PCA), a cluster analysis and an analysis of molecular variance (AMOVA). The MRD was computed separately for SSR and SNP data, according to Reif *et al.* [17]. For samples where multiple alleles were present, we assumed equal allele frequency of all alleles. The AMOVA was used to partition molecular variance into variance between world regions, countries within world regions and accessions within countries. All computations were performed in the R statistical software environment [18]. In particular, the R functions princomp and hclust were used for PCA and cluster analysis, respectively. The AMOVA was performed with the packages vegan [19] and pegas [20]. To facilitate visualization of the cluster, the accessions 49 to 70, which cluster all together and do not provide additional information, were eliminated from the cluster.

3. Results and Discussion

The percentage of markers with multiple alleles at a locus was low (Table 1). For the SSR markers, the highest percentage of markers with multiple alleles at a locus was found in the accession 9 (Mexico) followed by accessions 30 (Guatemala), 47 (Mexico), 7 (Mexico), 27 (Colombia but originated from

Salvador-Nicaragua), 29 (Guatemala) and 46 (Mexico). The SNP markers confirmed the different internal genetic structure of the accessions 9 and 7.

The analysis of SSR markers resulted in an average number of alleles per locus of 2.5 and 38 markers were polymorphic (Table 2). Three markers (M16, M34 and M53) showed multiple alleles in most accessions. However, the overall percentage of markers with multiple alleles was low (7.7%). A total of 25 SSR markers showed PIC values larger than 0.1. Similarly to the results obtained with SSR markers, the average number of accessions with multiple alleles per locus was very low in the SNP analysis (Table 3).

Table 2. Allele proportion in 70 accessions of *Jatropha curcas* L. and polymorphic information content (PIC) of SSR markers. Ama: percentage of accessions with multiple alleles at a locus.

| Marker | Allele Proportion | | | | | | | Ama (%) | PIC |
|--------|-------------------|------|------|------|------|------|------|---------|------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | | |
| M1 | 0.77 | 0.24 | | | | | | 1.5 | 0.29 |
| M2 | 0.96 | 0.04 | 0.01 | | | | | 1.4 | 0.08 |
| M3 | 1.00 | | | | | | | 0.0 | 0.00 |
| M4 | 0.99 | 0.03 | | | | | | 1.4 | 0.04 |
| M5 | 0.99 | 0.02 | 0.02 | | | | | 1.5 | 0.04 |
| M6 | 0.73 | 0.24 | 0.03 | 0.01 | | | | 1.4 | 0.35 |
| M7 | 1.00 | 0.02 | | | | | | 1.5 | 0.01 |
| M8 | 0.74 | 0.10 | 0.10 | 0.04 | 0.03 | 0.02 | | 2.9 | 0.42 |
| M9 | 0.99 | 0.05 | 0.03 | | | | | 6.0 | 0.08 |
| M10 | 0.81 | 0.23 | | | | | | 4.3 | 0.27 |
| M11 | 1.00 | | | | | | | 0.0 | 0.00 |
| M12 | 0.94 | 0.05 | 0.03 | | | | | 1.5 | 0.12 |
| M13 | 0.94 | 0.06 | 0.02 | | | | | 1.5 | 0.11 |
| M14 | 1.00 | | | | | | | 0.0 | 0.00 |
| M15 | 0.90 | 0.14 | 0.01 | | | | | 5.7 | 0.20 |
| M16 | 1.00 | 0.90 | 0.10 | | | | | 100.0 | 0.44 |
| M17 | 0.69 | 0.33 | | | | | | 1.4 | 0.34 |
| M18 | 0.87 | 0.20 | | | | | | 7.1 | 0.24 |
| M19 | 1.00 | 0.02 | 0.02 | | | | | 3.0 | 0.03 |
| M20 | 1.00 | | | | | | | 0.0 | 0.00 |
| M21 | 0.73 | 0.27 | 0.04 | | | | | 4.3 | 0.35 |
| M22 | 1.00 | | | | | | | 0.0 | 0.00 |
| M23 | 0.99 | 0.01 | | | | | | 0.0 | 0.03 |
| M24 | 1.00 | 0.04 | 0.01 | | | | | 5.7 | 0.05 |
| M25 | 0.73 | 0.27 | | | | | | 0.0 | 0.32 |
| M26 | 0.71 | 0.10 | 0.09 | 0.07 | 0.03 | 0.01 | 0.01 | 2.9 | 0.47 |
| M27 | 0.76 | 0.24 | 0.01 | 0.01 | | | | 2.9 | 0.34 |
| M28 | 1.00 | | | | | | | 0.0 | 0.00 |
| M29 | 1.00 | | | | | | | 0.0 | 0.00 |
| M30 | 1.00 | 0.01 | | | | | | 1.4 | 0.01 |
| M31 | 1.00 | | | | | | | 0.0 | 0.00 |

Table 2. Cont.

| Marker | Allele Proportion | | | | | | Ama | PIC |
|--------|-------------------|------|------|------|------|------|-------|------|
| | 1 | 2 | 3 | | 1 | 2 | | |
| M32 | 1.00 | | | | | | 0.0 | 0.00 |
| M33 | 0.87 | 0.09 | 0.04 | 0.03 | | | 2.9 | 0.24 |
| M34 | 0.99 | 0.99 | 0.03 | 0.01 | | | 98.6 | 0.41 |
| M35 | 1.00 | 0.03 | | | | | 2.9 | 0.03 |
| M36 | 0.99 | 0.01 | | | | | 0.0 | 0.03 |
| M37 | 0.83 | 0.09 | 0.06 | 0.03 | 0.01 | | 1.4 | 0.30 |
| M38 | 0.91 | 0.14 | 0.05 | 0.03 | 0.02 | 0.02 | 13.8 | 0.25 |
| M39 | 1.00 | | | | | | 0.0 | 0.00 |
| M40 | 1.00 | | | | | | 0.0 | 0.00 |
| M41 | 1.00 | 0.01 | 0.01 | | | | 2.9 | 0.03 |
| M42 | 0.76 | 0.14 | 0.10 | 0.01 | | | 1.4 | 0.37 |
| M43 | 1.00 | | | | | | 0.0 | 0.00 |
| M44 | 1.00 | | | | | | 0.0 | 0.00 |
| M45 | 1.00 | | | | | | 0.0 | 0.00 |
| M46 | 1.00 | | | | | | 0.0 | 0.00 |
| M47 | 1.00 | | | | | | 0.0 | 0.00 |
| M48 | 0.99 | 0.01 | | | | | 0.0 | 0.03 |
| M49 | 0.77 | 0.13 | 0.10 | 0.04 | 0.03 | | 7.1 | 0.38 |
| M50 | 0.81 | 0.17 | 0.01 | | | | 0.0 | 0.27 |
| M51 | 0.81 | 0.17 | 0.01 | | | | 0.0 | 0.27 |
| M52 | 0.83 | 0.17 | | | | | 0.0 | 0.24 |
| M53 | 1.00 | 0.99 | 0.29 | 0.01 | 0.01 | | 100.0 | 0.50 |
| M54 | 1.00 | 0.24 | | | | | 23.5 | 0.19 |
| Mean | | | | | | | 7.7 | 0.15 |
| Min | | | | | | | 0.0 | 0.00 |
| Max | | | | | | | 100.0 | 0.50 |

Table 3. Allele proportion in 16 accessions of *Jatropha curcas* L. and polymorphic information content (PIC) of SNP markers. Ama: percentage of accessions with multiple alleles at a locus.

| Marker | Allele Proportion | | Ama | PIC |
|--------|-------------------|-----|-----|------|
| | 1 | 2 | | |
| M1 | 0.9 | 0.1 | 0.0 | 0.11 |
| M2 | 0.5 | 0.5 | 0.0 | 0.37 |
| M3 | 0.6 | 0.4 | 6.3 | 0.11 |
| M4 | 0.9 | 0.1 | 0.0 | 0.37 |
| M5 | 0.6 | 0.5 | 6.3 | 0.37 |
| M6 | 0.9 | 0.2 | 6.3 | 0.23 |
| M7 | 0.8 | 0.3 | 0.0 | 0.30 |
| M8 | 0.7 | 0.3 | 0.0 | 0.34 |
| M9 | 0.9 | 0.1 | 0.0 | 0.11 |
| M10 | 0.9 | 0.1 | 0.0 | 0.19 |
| M11 | 0.9 | 0.1 | 0.0 | 0.11 |

Table 3. Cont.

| Marker | Allele Proportion | | Ama | PIC |
|--------|-------------------|-----|------|------|
| | 1 | 2 | | |
| M12 | 0.7 | 0.3 | 0.0 | 0.34 |
| M13 | 0.8 | 0.2 | 0.0 | 0.29 |
| M14 | 0.6 | 0.6 | 12.5 | 0.38 |
| M15 | 0.6 | 0.4 | 6.3 | 0.37 |
| M16 | 0.7 | 0.4 | 12.5 | 0.36 |
| M17 | 0.9 | 0.2 | 6.3 | 0.23 |
| M18 | 0.8 | 0.3 | 0.0 | 0.30 |
| M19 | 0.6 | 0.4 | 0.0 | 0.37 |
| M20 | 0.8 | 0.2 | 0.0 | 0.26 |
| M21 | 0.7 | 0.3 | 0.0 | 0.34 |
| M22 | 0.9 | 0.2 | 12.5 | 0.19 |
| M23 | 0.5 | 0.5 | 0.0 | 0.37 |
| M24 | 0.6 | 0.5 | 6.3 | 0.37 |
| M25 | 0.9 | 0.1 | 0.0 | 0.20 |
| M26 | 0.7 | 0.4 | 6.3 | 0.35 |
| M27 | 0.9 | 0.1 | 6.3 | 0.16 |
| M28 | 0.9 | 0.1 | 6.3 | 0.16 |
| M29 | 0.6 | 0.5 | 6.3 | 0.37 |
| M30 | 0.6 | 0.5 | 6.3 | 0.37 |
| M31 | 0.9 | 0.1 | 0.0 | 0.11 |
| M32 | 0.9 | 0.1 | 0.0 | 0.11 |
| M33 | 0.9 | 0.1 | 0.0 | 0.11 |
| M34 | 0.6 | 0.5 | 6.3 | 0.37 |
| M35 | 0.8 | 0.3 | 0.0 | 0.30 |
| M36 | 0.6 | 0.4 | 0.0 | 0.36 |
| M37 | 0.6 | 0.4 | 0.0 | 0.37 |
| M38 | 0.8 | 0.2 | 0.0 | 0.26 |
| M39 | 0.9 | 0.1 | 0.0 | 0.11 |
| M40 | 0.8 | 0.2 | 0.0 | 0.26 |
| M41 | 0.8 | 0.3 | 6.7 | 0.29 |
| M42 | 0.6 | 0.4 | 0.0 | 0.37 |
| M43 | 0.9 | 0.1 | 0.0 | 0.19 |
| M44 | 0.8 | 0.2 | 0.0 | 0.26 |
| M45 | 0.8 | 0.3 | 12.5 | 0.30 |
| M46 | 0.8 | 0.2 | 0.0 | 0.26 |
| M47 | 0.9 | 0.2 | 6.3 | 0.23 |
| M48 | 0.9 | 0.1 | 0.0 | 0.19 |
| M49 | 0.7 | 0.3 | 0.0 | 0.35 |
| M50 | 0.9 | 0.1 | 0.0 | 0.11 |
| M51 | 0.9 | 0.1 | 0.0 | 0.11 |
| M52 | 0.5 | 0.5 | 0.0 | 0.37 |
| M53 | 0.9 | 0.1 | 0.0 | 0.11 |
| M54 | 0.9 | 0.1 | 0.0 | 0.20 |
| M55 | 0.6 | 0.4 | 0.0 | 0.36 |

Table 3. Cont.

| Marker | Allele Proportion | | Ama | PIC |
|--------|-------------------|-----|-----|------|
| | 1 | 2 | | |
| M56 | 0.9 | 0.1 | 0.0 | 0.11 |
| M57 | 0.6 | 0.4 | 7.1 | 0.36 |
| M58 | 0.7 | 0.4 | 6.3 | 0.35 |
| M59 | 0.6 | 0.5 | 6.3 | 0.37 |
| M60 | 0.7 | 0.4 | 6.3 | 0.35 |
| M61 | 0.6 | 0.5 | 6.3 | 0.37 |
| M62 | 0.6 | 0.4 | 0.0 | 0.37 |
| M63 | 0.6 | 0.5 | 6.3 | 0.37 |
| M64 | 0.9 | 0.1 | 6.3 | 0.16 |
| M65 | 0.6 | 0.4 | 0.0 | 0.35 |
| M66 | 0.8 | 0.2 | 0.0 | 0.26 |
| M67 | 0.9 | 0.1 | 0.0 | 0.20 |
| M68 | 0.9 | 0.1 | 0.0 | 0.19 |
| M69 | 0.9 | 0.1 | 0.0 | 0.11 |
| M70 | 0.9 | 0.1 | 0.0 | 0.11 |
| M71 | 0.6 | 0.4 | 6.3 | 0.37 |
| M72 | 0.8 | 0.2 | 0.0 | 0.24 |
| M73 | 0.8 | 0.3 | 0.0 | 0.30 |
| M74 | 0.8 | 0.3 | 6.3 | 0.32 |
| M75 | 0.5 | 0.5 | 0.0 | 0.38 |
| M76 | 0.8 | 0.3 | 6.3 | 0.28 |
| M77 | 0.8 | 0.3 | 6.3 | 0.32 |
| M78 | 0.6 | 0.5 | 6.3 | 0.37 |
| M79 | 0.8 | 0.2 | 0.0 | 0.26 |
| M80 | 0.7 | 0.3 | 0.0 | 0.34 |
| M81 | 0.6 | 0.5 | 6.3 | 0.37 |
| M82 | 0.6 | 0.4 | 0.0 | 0.37 |
| M83 | 0.8 | 0.2 | 0.0 | 0.23 |
| M84 | 0.9 | 0.1 | 0.0 | 0.12 |
| M85 | 0.8 | 0.2 | 0.0 | 0.26 |
| M86 | 0.5 | 0.5 | 0.0 | 0.38 |
| M87 | 0.9 | 0.1 | 6.7 | 0.16 |
| M88 | 0.6 | 0.5 | 6.3 | 0.37 |
| M89 | 0.6 | 0.4 | 0.0 | 0.37 |
| M90 | 0.6 | 0.5 | 6.3 | 0.37 |
| M91 | 0.7 | 0.4 | 6.3 | 0.35 |
| M92 | 0.9 | 0.1 | 0.0 | 0.11 |
| M93 | 0.7 | 0.3 | 0.0 | 0.34 |
| M94 | 0.6 | 0.5 | 6.3 | 0.37 |
| M95 | 0.7 | 0.3 | 0.0 | 0.35 |
| M96 | 0.9 | 0.1 | 0.0 | 0.11 |
| M97 | 0.8 | 0.2 | 0.0 | 0.26 |
| M98 | 0.6 | 0.4 | 0.0 | 0.37 |
| M99 | 0.5 | 0.5 | 0.0 | 0.38 |

Table 3. Cont.

| Marker | Allele Proportion | | Ama | PIC |
|--------|-------------------|-----|------|------|
| | 1 | 2 | | |
| M100 | 0.7 | 0.3 | 0.0 | 0.34 |
| M101 | 0.8 | 0.2 | 0.0 | 0.26 |
| M102 | 0.9 | 0.1 | 0.0 | 0.11 |
| M103 | 0.8 | 0.2 | 0.0 | 0.28 |
| M104 | 0.5 | 0.5 | 0.0 | 0.38 |
| M105 | 0.6 | 0.4 | 0.0 | 0.36 |
| M106 | 0.8 | 0.2 | 0.0 | 0.26 |
| M107 | 0.7 | 0.3 | 0.0 | 0.34 |
| M108 | 0.9 | 0.1 | 0.0 | 0.11 |
| M109 | 0.8 | 0.3 | 6.3 | 0.32 |
| M110 | 0.9 | 0.1 | 0.0 | 0.19 |
| M111 | 0.9 | 0.1 | 6.3 | 0.16 |
| M112 | 0.5 | 0.5 | 0.0 | 0.38 |
| M113 | 0.8 | 0.3 | 0.0 | 0.30 |
| M114 | 0.8 | 0.2 | 0.0 | 0.26 |
| M115 | 0.8 | 0.3 | 6.3 | 0.28 |
| M116 | 0.8 | 0.2 | 0.0 | 0.26 |
| M117 | 0.9 | 0.2 | 12.5 | 0.19 |
| M118 | 0.6 | 0.4 | 0.0 | 0.37 |
| M119 | 0.9 | 0.1 | 0.0 | 0.11 |
| M120 | 0.7 | 0.3 | 0.0 | 0.34 |
| Mean | | | 2.4 | 0.27 |
| Min | | | 0.0 | 0.11 |
| Max | | | 12.5 | 0.38 |

The variance component analysis revealed that the highest proportion of the genetic variation is present among the world regions and among the accessions within a world region and country (Table 4). The variance components based on SSR and SNP marker data of the same accessions (N = 16) were similar. The variance components based on the complete set of 70 accessions with the SSR data showed, as expected, that most of the variation is in accessions within a world region and country.

The cluster analysis based on SNP marker data resulted in a grouping pattern that clearly distinguishes accessions according to presence or absence of phorbol esters (Figure 1). The cluster analysis based on SSR marker data showed a similar pattern extended to a larger number of accessions (Figure 2). The principal component analysis of the SSR marker data showed a clear grouping pattern to distinguish accessions according to their phorbol ester classification (Figure 3).

Figure 1. Cluster of 16 accessions of *Jatropha curcas* L. based on 120 SNP markers. Accessions's labels combined accession identification number, world region and phorbol presence. SAM: South America; CNAM: Central and North America. P: phorbol present; A: phorbol absent.

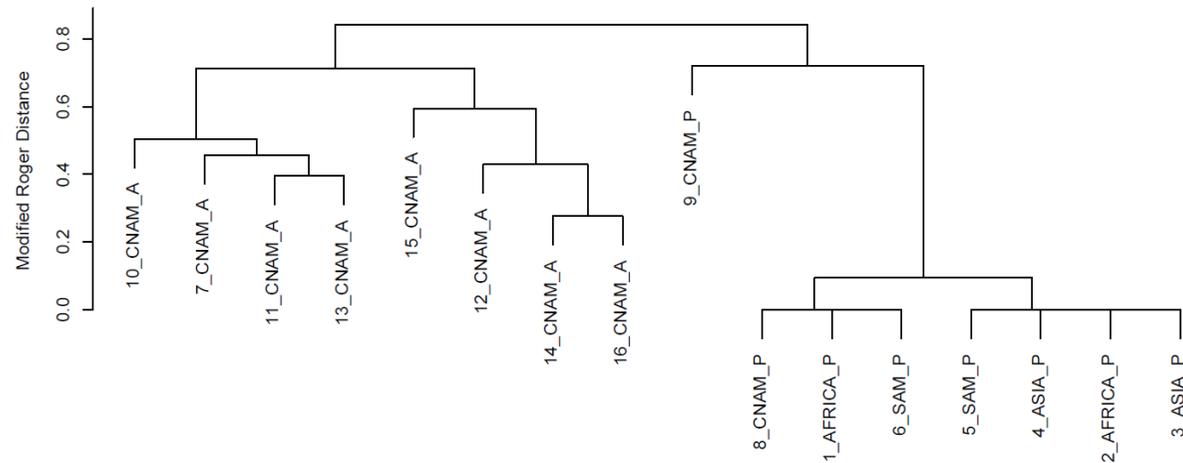


Figure 2. Cluster of 48 accessions of *Jatropha curcas* L. based on 54 SSR markers. Accessions's labels combined accession identification number, world region and phorbol presence. SAM: South America; CNAM: Central and North America. P: phorbol present; A: phorbol absent.

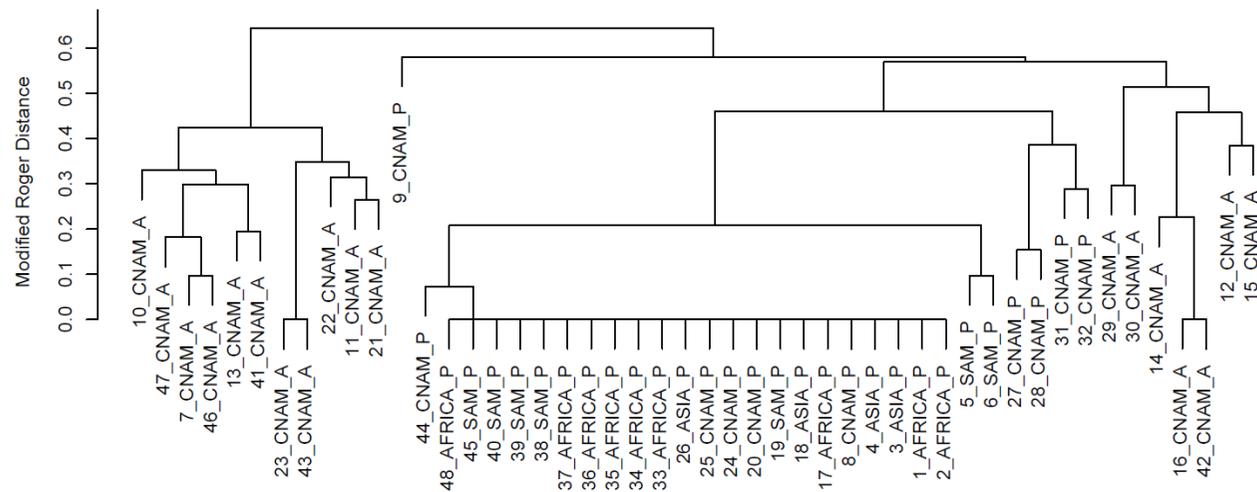


Figure 3. Biplot of the first two principal components for 48 accessions of *Jatropha curcas* L. based on 54 SSR markers. PE: phorbol esters.

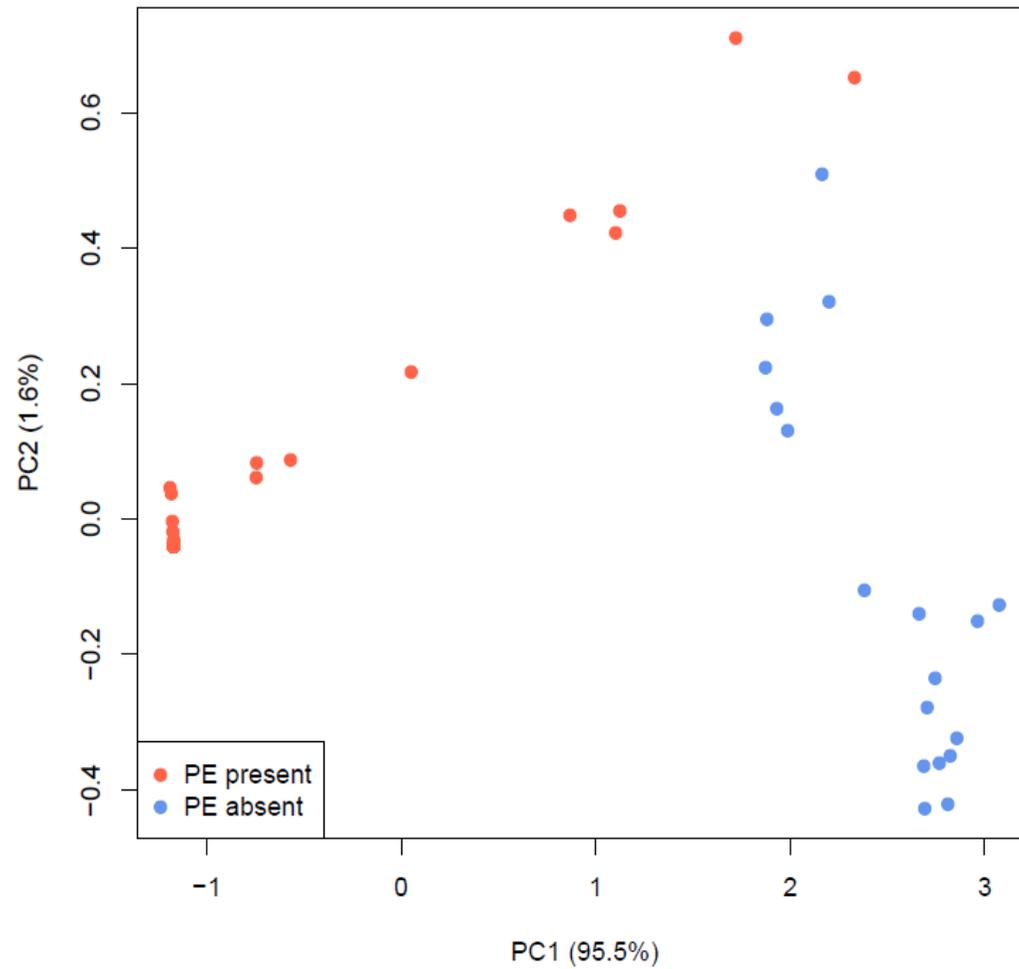


Table 4. Variance components expressed as percentage of the total variance of genetic distances estimated with 54 SSR and 120 SNP markers in a genetic diverse panel of accessions of *Jatropha curcas* L. Df: degree of freedom.

| Source of Variation | Df ^a | Df ^b | SSR ^a | SNP ^a | SSR ^b |
|--|-----------------|-----------------|------------------|------------------|------------------|
| World region | 3 | 3 | 50.0 | 47.5 | 32.6 |
| Countries within world region | 3 | 9 | 0.0 | 0.0 | 19.3 |
| Accessions within countries and world region | 9 | 57 | 50.0 | 52.5 | 48.0 |

^a Reduced panel of accessions (N = 16); ^b Complete panel of accessions (N = 70).

Knowledge of the internal genetic structure within and across accessions is of significant importance in breeding programs. The optimal design of a breeding strategy relies on the accurate information of the internal genetic structure and polymorphism of the germplasm at the breeder's hands. The optimal breeding category (clone, line, population, hybrid) for jatropha will then be defined in a comprehensive manner by combining knowledge of genetic structure, polymorphism, heterosis level and the cost to produce the improved cultivars.

The low number of markers that showed multiple alleles at a locus indicated a high level of homozygosity in this germplasm panel of jatropha. There were only seven accessions with more than 10% of markers with multiple alleles at a locus (Table 1). The high level of homozygosity is present in germplasm from all world regions. This result is not in agreement with the expected level of homozygosity in outcrossing crops and it questions the mating system of jatropha.

The monoecious character of jatropha, the different opening time of male and female flowers and the need of insect pollinators for transport of pollen indicated a potential outcrossing mating system. However, molecular results with co-dominant markers indicate a high level of homozygosity. Considering the low genetic variation in germplasm from Africa, Asia, and South America, one might speculate that the high level of homozygosity is due to the lack of genetic variation available at those regions due to a bottleneck of genetic diversity transfer from Central America and Mexico. However, the high level of homozygosity was also observed in germplasm originated from Mexico where the highest genetic diversity of jatropha is present. This indicates that the mating system of jatropha might result generally in a high level of homozygosity. Results from recent studies are contradictory to the assumed purely outcrossing mating system in jatropha and point out to a mixed mating system [13,21].

The cluster of accessions showed a perfect agreement with our expectation and confirms that the accessions from Central America and Mexico have a higher genetic diversity than the accessions from Africa, Asia, and South America (Figures 1 and 2). This information is of importance to the breeders in order to establish germplasm management concepts and the planning of crosses to exploit heterosis and combine trait of economic importance.

In addition, the cluster analysis showed a clear grouping pattern of the accessions according to their phorbol esters presence. The principal component analysis indicated that it is possible to classify accessions for their phorbol esters presence with a high level of certainty (Figure 3). Therefore, we conclude that the molecular makers in our study are a suitable tool to classify accessions rapidly and cost-effectively. In addition to these markers, additional published makers [15] could be also used to accelerate the introgression of non-toxicity into elite recurrent toxic genotypes.

The selection of the marker system to be implemented in breeding programs depends principally on the costs. As the cost per marker data point is decreasing rapidly, the development of SNP arrays with thousands of marker data points for many samples seems to be the most probable future scenario. However, we have identified 38 polymorphic SSR markers with good potential for the classification of germplasm and four SSR markers showed a perfect association with the presence of phorbol esters. Similarly, 14 SNP markers showed a perfect association with that trait. In addition, the analysis of those 14 SNP markers indicated that the heterozygous marker corresponded to accessions with phorbol esters present. This agrees well with the dominance character of phorbol esters biosynthesis recently published [15].

4. Conclusions

We found a high level of homozygosity in a worldwide panel of accessions and this result is not in accordance to the purely outcrossing mating system assumed to be present in *Jatropha*. We hypothesize that the prevalent mating system of *Jatropha* comprise a high level of self-fertilization and that the outcrossing rate is low. The genetic diversity in accessions from Central America and Mexico was higher than in accession from Africa, Asia, and South America. Markers associated with the presence of phorbol esters were identified.

The increasing knowledge of molecular marker information in *Jatropha* allows breeders to design optimal knowledge-based breeding strategies. A first interspecific genetic map [22] and a first intraspecific genetic map [15] have been published, and the position of a limited number of SNP and SSR are known. However, the full exploitation of the molecular tools for breeding will require a higher density of markers to allow association studies and genomic selection. In parallel to the efforts in increasing the marker density, accurate evaluation of field performance covering a range of environments will be also important. We believe that an optimal breeding strategy in *Jatropha* will combine the utilization of high-density marker data with accurate field performance across a range of environments.

Acknowledgments

This work was funded by JatroSelect GmbH. We thank Ganal and Plieske for helping in the design of the molecular assays.

Author Contributions

Juan M. Montes: conceived and wrote the paper. Frank Technow: conducted statistical analysis, made tables and figures and assisted with revisions. Matthias Martin: assisted with revisions. Klaus Becker: assisted with revisions.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Ceasar, S.A.; Ignacimuthu, S. Applications of biotechnology and biochemical engineering for the development of jatropha and biodiesel: A review. *Renew. Sustain. Energy Rev.* **2011**, *15*, 5176–5185.
2. Divakara, B.N.; Upadhyaya, H.D.; Wani, S.P.; Laxmipathi Gowda, C.L. Biology and genetic improvement of *Jatropha curcas* L.: A review. *Appl. Energy* **2010**, *87*, 732–742.
3. Francis, G.; Edinger, R.; Becker, K. A concept for simultaneous wasteland reclamation, fuel production, and socio-economic development in degraded areas in India. Need, potential and perspectives of *Jatropha* plantations. *Nat. Resour. Forum* **2005**, *29*, 12–24.
4. Fairless, D. The little shrub that could—Maybe. *Nature* **2007**, *449*, 652–655.
5. Sanderson, K. Wonder weed plans fail to flourish. *Nature* **2009**, *461*, 328–329.
6. King, A.J.; He, W.; Cuevas, J.A.; Freudenberger, M.; Ramiarmanana, D.; Graham, I.A. Potential of *Jatropha curcas* as source of renewable oil and animal feed. *J. Exp. Bot.* **2009**, *60*, 2897–2905.
7. Sujatha, M. Genetic diversity, molecular markers and marker assisted breeding in *Jatropha*. In *Jatropha, Challenges for a New Energy Crop*, 1st ed.; Carels, N., Sujatha, M., Bahadur, B., Eds.; Springer: New York, NY, USA, 2012; Volume 2, pp. 395–422.
8. Sun, Q.; Li, L.; Li, Y.; Wu, G.; Ge, X. SSR and AFLP markers reveal low genetic diversity in the biofuel plant *Jatropha curcas* in China. *Crop Sci.* **2008**, *48*, 1865–1971.
9. Rosado, T.B.; Laviola, B.G.; Faria, D.A.; Pappas, M.R.; Bhering, L.L.; Quirino, B.; Grattapaglia, D. Molecular markers reveal limited genetic diversity in a large germplasm collection of the biofuel crop *Jatropha curcas* L. in Brazil. *Crop Sci.* **2010**, *50*, 2372–2382.
10. Ambrosi, D.G.; Gala, G.; Purelli, M.; Barbi, T.; Fabbri, A.; Lucretti, S.; Sharbel, T.F.; Barccacia, G. DNA markers and FCSS analyses shed light on the genetic diversity and reproductive strategy of *Jatropha cruces* L. *Diversity* **2010**, *2*, 810–836.
11. Luo, C.W.; Li, K.; Chen, Y.; Sun, Y.Y. Floral display and breeding system of *Jatropha curcas* L. *For. Stud. China* **2007**, *9*, 114–119.
12. Luo, C.W.; Huang, Z.Y.; Chen, X.M.; Li, K.; Chen, Y.; Sun, Y.Y. Contribution of diurnal and nocturnal insects to the pollination of *Jatropha curcas* (Euphorbiaceae) in southwestern China. *J. Econ. Entomol.* **2011**, *104*, 149–154.
13. Bressan, E.A.; Sebbenn, A.M.; Ferreira, R.R.; Lee, T.S.G.; Figueira, A. *Jatropha curcas* L. (Euphorbiaceae) exhibits a mixed mating system, high correlated mating and apomixes. *Tree Genet. Genomes* **2013**, *9*, 1089–1097.
14. Devappa, R.K.; Makkar, H.P.S.; Becker, K. *Jatropha* toxicity—A Review. *J. Toxicol. Environ. Health* **2010**, *13*, 476–507.
15. King, A.J.; Montes, L.R.; Clarke, J.G.; Affleck, J.; Li, Y.; Witsenboer, H.; van der Vossen, E.; van der Linde, P.; Tripathi, Y.; Tavares, E.; et al. Linkage mapping in the oilseed crop *Jatropha curcas* L. reveals a locus controlling the biosynthesis of phorbol esters which cause seed toxicity. *Plant Biotechnol. J.* **2013**, *11*, 986–996.
16. Botstein, D.; White, R.L.; Skolnick, M.; Davis, R.W. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum. Genet.* **1980**, *32*, 314–331.

17. Reif, J.C.; Melchinger, A.E.; Frisch, M. Genetical and Mathematical Properties of Similarity and Dissimilarity Coefficients Applied in Plant Breeding and Seed Bank Management. *Crop Sci.* **2005**, *45*, 1–7.
18. R Core Team., 2013: R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available online: <http://www.R-project.org/> (accessed on 3 June 2014).
19. Oksanen, J.; Blanchet, F.G.; Kindt, R.; Legendre, P.; Minchin, P.R.; O’Hara, R.B.; Simpson, G.L.; Solymos, P.; Henry, M.; Stevens, H.; *et al.* Vegan: Community Ecology Package. R package version 2.0–7. Available online: <http://CRAN.R-project.org/package=vegan> (accessed on 3 June 2014).
20. Paradis, E. PEGAS: An R package for population genetics with an integrated-modular approach. *Bioinformatics* **2010**, *26*, 419–420.
21. Salvador-Figueroa, M.; Magana-Ramos, J.; Vazquez-Ovando, J.A.; Adriano-Anaya, M.L.; Ovando-Medina, I. Genetic diversity and structure of *Jatropha curcas* L. in its centre of origin. *Plant Genet. Resour.* **2014**, doi:10.1017/S1479262114000550.
22. Wang, C.M.; Liu, P.; Yi, C.; Gu, K.; Sun, F.; Lei, L.; Lo, L.C.; Liu, X.; Feng, F.; Lin, G.; *et al.* A first generation microsatellite- and SNP-based linkage map of *Jatropha*. *PLoS One* **2011**, doi:10.1371/journal.pone.0023632.

© 2014 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 license (<http://creativecommons.org/licenses/by/3.0/>).