

Development of DArT Marker Platforms and Genetic Diversity Assessment of the U.S. Collection of the New Oilseed Crop *Lesquerella* and Related Species

Von Mark V. Cruz^{1,3}, Andrzej Kilian², David A. Dierig^{1*}

1 USDA-ARS National Center for Genetic Resources Preservation, Fort Collins, Colorado, United States of America, **2** Diversity Arrays Technology Pty. Ltd., Yarralumla, ACT, Australia, **3** Department of Bioagricultural Sciences and Pest Mgt., Colorado State University, Fort Collins, Colorado, United States of America

Abstract

The advantages of using molecular markers in modern genebanks are well documented. They are commonly used to understand the distribution of genetic diversity in populations and among species which is crucial for efficient management and effective utilization of germplasm collections. We describe the development of two types of DArT molecular marker platforms for the new oilseed crop *lesquerella* (*Physaria* spp.), a member of the Brassicaceae family, to characterize a collection in the National Plant Germplasm System (NPGS) with relatively little known in regards to the genetic diversity and traits. The two types of platforms were developed using a subset of the germplasm conserved *ex situ* consisting of 87 *Physaria* and 2 *Paysonia* accessions. The microarray DArT revealed a total of 2,833 polymorphic markers with an average genotype call rate of 98.4% and a scoring reproducibility of 99.7%. On the other hand, the DArTseq platform developed for SNP and DArT markers from short sequence reads showed a total of 27,748 high quality markers. Cluster analysis and principal coordinate analysis indicated that the different accessions were successfully classified by both systems based on species, by geographical source, and breeding status. In the germplasm set analyzed, which represented more than 80% of the *P. fendleri* collection, we observed that a substantial amount of variation exists in the species collection. These markers will be valuable in germplasm management studies and *lesquerella* breeding, and augment the microsatellite markers previously developed on the taxa.

Citation: Cruz VMV, Kilian A, Dierig DA (2013) Development of DArT Marker Platforms and Genetic Diversity Assessment of the U.S. Collection of the New Oilseed Crop *Lesquerella* and Related Species. PLoS ONE 8(5): e64062. doi:10.1371/journal.pone.0064062

Editor: Matthew E. Hudson, University of Illinois, United States of America

Received: November 26, 2012; **Accepted:** April 8, 2013; **Published:** May 28, 2013

This is an open-access article, free of all copyright, and may be freely reproduced, distributed, transmitted, modified, built upon, or otherwise used by anyone for any lawful purpose. The work is made available under the Creative Commons CC0 public domain dedication.

Funding: Funding for the study was provided by U.S. Department of Agriculture (USDA) - Agricultural Research Service Project No. 5402-21000-007-00. The acquisition of new germplasm in 2010 was supported by the NPGS Plant Exploration/Exchange Program. No additional external funding was received or used for this study. USDA determines the research priorities of all research that it funds. Publication of research is subject to review of USDA officials. Specific study design, data collection, data analysis, and preparation of the manuscript is under the control of the individual researchers.

Competing Interests: The authors acknowledge the help and technical support from the employees of Diversity Arrays Technology Pty. Ltd. AK is an employee of Diversity Arrays Technology Pty Ltd, which offers genome profiling service using the technologies described in this report. This fact, however, has not interfered whatsoever with the full, objective, transparent and unbiased presentation of the research results described in the manuscript nor alters the authors' adherence to all the PLOS ONE policies on sharing data and materials.

* E-mail: david.dierig@ars.usda.gov

Introduction

Lesquerella (*Physaria* sp.), a member of the Brassicaceae, is a North American genus which has more than 90 member species thriving mostly in dry and arid habitats usually mixed with sparse vegetation [1], [2]. Commonly known as bladderpod, it has been identified by the U.S. Department of Agriculture and the U.S. Department of Energy as very promising species for the production of lubricants, engine oils, waxes, coatings and a source of natural estolides with valuable application in the automobile and biofuel industries [3], [4], [5]. There has been an increasing trend in research activities in the crop since its oil was first characterized for novel properties [6]. The fatty acid profile of *lesquerella* oil has been found to vary. Species found east of the Mississippi have mostly densipolic acid, those in the western U.S. mostly *lesquerolic* acid, and a species from Oklahoma predominantly with *auricollic* acid [7]. *Lesquerella* could be grown successfully as a winter annual in the southwest U.S. producing an average seed yield of 1.7 tons/ha [8]. At present, there are more than eight advanced breeding lines of *P. fendleri* ready for

commercialization and a corresponding substantial germplasm collection has been assembled in the U.S. National Plant Germplasm System [9].

Molecular characterization of germplasm collections supplements phenotypic assessment of diversity and is important in the effective management of genetic resources. Molecular markers have been very useful in efforts to accurately identify gaps and redundancy within and among individual germplasm collections and have helped resolve important genebank management issues [10], [11], [12]. Examples of markers used in specific Brassicaceae collections include microsatellites for examining diversity in *B. napus* [13], [14], [15], *lesquerella* [16] and wild relatives such as *Capsella*, *Crambe* and *Sinapis* [17]; AFLPs in *B. oleracea* [18] and *Lepidium* [19]; and RAPD markers in *Raphanus* [20].

Compared to the previously mentioned molecular marker systems, Diversity Array Technology (DArT) markers are relatively new and were developed only in the early 2000. DArT markers overcame the difficulties in correlating bands on gels with allelic variants by utilizing hybridization-based methods and solid state surfaces [21]. They have been increasingly utilized as

evidenced by the number of publications reporting successful implementation [22]. Among its advantages over other marker systems include high throughput capability allowing rapid germplasm characterization in a single experiment, independence of sequence data, and ability to detect single base changes and indels [23]. To date, DArT markers have been successfully applied in genetic diversity analysis, linkage mapping and in finding out population structure of collections in various crop species [24]. Their application in minor crops are likewise increasing due to their potential to accelerate gene discovery and initiate molecular breeding because of their whole genome coverage without relying on prior sequence data information [25], [26].

In lesquerella, allozymes have previously been used [27] and microsatellite markers have been developed [16]. But the small number of markers that are available presents a limitation in linkage mapping and in the study of genetic resources collections. At present, only fifteen microsatellites have demonstrated utility across different *Physaria* species. We present in this paper the development of two platforms of high density DArT markers for lesquerella and the results obtained after testing them to analyze the genetic diversity of *Physaria* and *Paysonia* national germplasm collection. This new molecular marker system for the new crop species will serve as an additional resource to augment the existing systems to assist crop improvement efforts, germplasm management activities and genetic studies.

Materials and Methods

Tissue Sampling

The lesquerella germplasm set was obtained from collections of the USDA-ARS National Arid Land Plant Genetic Resources Unit, Parlier, CA and the USDA-ARS Arid Land Agricultural Research Center, Maricopa, AZ. The samples used in this study were selected based on each accession's geographic location, sampling within distinct counties in each State when possible. Seven advanced *P. fendleri* breeding lines, WCL-LH1, WCL-LO1, WCL-LO2, WCL-LO4, WCL-LY1, WCL-SL1 and WCL-YS1 were included among the samples. DNA was extracted from fresh leaf tissue obtained from five week old seedlings using Qiagen DNeasy 96 Plant Kits (Qiagen Inc., Valencia, CA). Two DArT platforms were developed for lesquerella as described below. The platform development utilized 86 *Physaria* accessions (see Table 1), representing 11 species and one *Paysonia* accession since the latter is a sister genus of *Physaria* and only recently was there a recognized taxonomic delineation between them [2].

Microarray DArT Platform Development

The microarray DArT was developed by first testing combinations of the rare-cutting restriction enzyme *PstI* with several restriction endonucleases that cut frequently on DNA samples from 8 representative accessions to determine the restriction enzyme combination that provided the best complexity reduction. Final genomic representations were prepared using *PstI*/*BstNI* combinations. Approximately 50 ng of genomic DNA was digested with *PstI*/*BstNI* combinations and the resulting fragments ligated to a *PstI* overhang compatible oligonucleotide adapter. A primer annealing to this adapter was used in PCR reaction to amplify complexity-reduced representation of a sample. Amplification products were either used for cloning a in marker development process or labeled with fluorescent dyes and hybridized to DArT array in the genotyping process. Library construction was subsequently performed using 80 *P. fendleri* accessions and 16 accessions of wild related *Brassica* species. The amplified *PstI* restriction fragments from all accessions were cloned

into pCR2.1-TOPO vector (Invitrogen, Australia) as described by Jaccoud et al. [21] and four libraries were generated. The white colonies containing genomic fragments inserted into pCR2.1-TOPO vector were picked into individual wells of 384-well microtiter plates filled with ampicillin/kanamycin-supplemented freezing medium. A total of 3,456 clones from *Physaria* were obtained –1,920 from three libraries of *P. fendleri* and 1,536 from the related species. Inserts from these clones were amplified using M13F and M13R primers in 384 plate format, a subset of PCR products were assessed for quality (10% of 25 µl PCR reaction) through gel electrophoresis, and all remaining PCR products dried, washed and dissolved in a spotting buffer. A total of 6,144 clones were printed with spot duplication on SuperChip poly-L-lysine slides (Thermo Scientific, Australia) using a MicroGrid arrayer (Genomics Solutions, UK). The microarrays included 1,920 *Brassica* clones and 768 *Arabidopsis* clones in addition to the *Physaria* clones.

Each sample was assayed using methods described above for library construction. Genomic representations were assessed for quality through gel electrophoresis in 1.2% agarose and labeled with fluorescent dyes (Cy3 and Cy5). Labeled targets were then hybridized to printed DArT arrays for 16 hours at 62°C in a water bath. Slides were washed as described by Kilian et al. [28], dried initially by centrifugation at 500 × g for 7 min and later by a desiccator under vacuum for 30 min. The slides were scanned using Tecan LS300 scanner (Tecan Group Ltd, Männedorf, Switzerland) generating three images per array: one image scanned at 488 nm for reference signal measures the amount of DNA within the spot based on hybridization signal of FAM-labeled fragment of a TOPO vector multiple cloning site fragment and two images for “target” signal measurement: one scanned at 543 nm (for Cy3 labeled targets) and one at 633 nm (for Cy5 labeled targets). All DArT techniques applied for this work were recently described in much more detail by Kilian et al. [28].

DArTseq Platform Development

For the sequencing-based DArT genotyping, four complexity reduction methods optimized for several other plant species at DArT P/L were used. *PstI*-RE site specific adapter was tagged with 96 different barcodes enabling encoding a plate of DNA samples to run within a single lane on an Illumina Genome Analyzer Ix (Illumina Inc., San Diego, CA). *PstI* adapter included also a sequencing primer site, so that the tags generated were always reading into the genomic fragments from the *PstI* sites. After the sequencing run, the FASTQ files were quality filtered using the threshold of 90% confidence for at least 50% of the bases and in addition filtered more stringently for barcode sequences. Two lanes of GAIIx were run with all samples providing fully replicated sequencing data. The filtered data were split into their respective target (individual) data using barcode splitting script. Each sample had on average 500,000 counts per replicate. After producing various QC statistics and trimming of the barcode the sequences were aligned against the reference created from the tags identified in the sequence reads generated from all the samples. In addition the short sequence tags were aligned against *Arabidopsis thaliana*'s genome available in Genbank. *Arabidopsis* is a close relative of *Physaria* in the Brassicaceae [29]. The output files from alignment generated using Bowtie software [30] were processed using an analytical pipeline developed by DArT P/L to produce “DArT score” tables and “SNP” tables.

Genotyping

Both DArT on array platform and DArTseq use a set of quality parameters to select markers which are of use for a specific

Table 1. Passport information of accessions used in *Physaria* DArT and DArTseq platform development.

Collection No. and ID	Accession No.	Species	Country	State/Prov.	County	Latitude	Longitude	Elevation (m)
1809	W6 20822	<i>P. fendleri</i>	United States	AZ	Cochise	31.75600	-110.05500	134.15
1817	W6 20823	<i>P. fendleri</i>	United States	AZ	Cochise	31.74100	-109.96800	1493.90
1818	W6 20824	<i>P. fendleri</i>	United States	AZ	Cochise	31.50000	-110.00000	1432.56
1819	W6 20825	<i>P. fendleri</i>	United States	AZ	Santa Cruz	31.72400	-110.52600	1493.90
1826	W6 20856	<i>P. fendleri</i>	United States	AZ	Santa Cruz	31.72300	-110.52400	1463.41
1833	W6 20833	<i>P. gordonii</i>	United States	NM	Hidalgo	31.93333	-108.95000	1024.13
1834	PI 641918	<i>P. fendleri</i>	United States	NM	Hidalgo	31.93333	-108.93333	1292.35
1835	-	<i>P. fendleri</i>	United States	NM	Luna	32.21667	-107.41667	1277.11
1840	W6 20858	<i>P. fendleri</i>	United States	NM	Eddy	32.61667	-104.40000	1042.42
1841	PI 596417	<i>P. fendleri</i>	United States	NM	Eddy	32.31667	-104.25000	1130.81
1842	W6 20860	<i>P. fendleri</i>	United States	NM	Eddy	32.53333	-103.98333	1060.70
1843	PI 596418	<i>P. fendleri</i>	United States	NM	Eddy	32.75000	-104.10000	1082.04
1844	PI 596419	<i>P. fendleri</i>	United States	NM	Lincoln	33.68333	-105.83333	1636.78
1845	PI 596420	<i>P. fendleri</i>	United States	NM	Lincoln	33.68333	-105.91667	1572.77
1848	PI 596421	<i>P. fendleri</i>	United States	AZ	Graham	33.25000	-110.28333	795.53
1851	W6 20826	<i>P. fendleri</i>	United States	AZ	Cochise	31.91667	-109.16667	1493.52
1852	W6 20861	<i>P. fendleri</i>	United States	AZ	Cochise	31.65000	-110.01667	1447.80
1874	PI 596423	<i>P. fendleri</i>	United States	AZ	Apache	34.56667	-109.50000	1755.65
1889	PI 596425	<i>P. fendleri</i>	United States	NM	Socorro	33.88333	-106.40000	1661.16
1904	W6 20865	<i>P. fendleri</i>	United States	AZ	Apache	34.35000	-109.36667	1859.28
1906	PI 596426	<i>P. fendleri</i>	United States	NM	Socorro	34.41667	-106.65000	1645.92
1909	W6 20868	<i>P. fendleri</i>	United States	NM	Guadalupe	33.56667	-105.05000	1697.74
1910	PI 596428	<i>P. fendleri</i>	United States	NM	De Baca	34.46667	-104.30000	1149.10
1912	PI 596429	<i>P. fendleri</i>	United States	NM	Guadalupe	34.73333	-104.46667	1359.41
1919	PI 596430	<i>P. fendleri</i>	United States	NM	San Juan	36.60000	-108.91667	1758.70
1920	PI 596431	<i>P. fendleri</i>	United States	NM	San Juan	36.58333	-108.96667	1932.43
1924	W6 20851	<i>P. rectipes</i>	United States	AZ	Luna	35.86667	-109.60000	1912.00
1932	PI 596433	<i>P. fendleri</i>	United States	AZ	Graham	33.28333	-110.38333	877.82
2212	W6 20863	<i>P. argyrea</i>	United States	TX	Uvalde	29.15000	-99.70000	228.60
2226	W6 20828	<i>P. fendleri</i>	United States	TX	Webb	27.36667	-98.98333	243.84
2232	PI 643174	<i>P. lindheimeri</i>	United States	TX	Nueces	27.06667	-97.06667	107.00
2243	W6 20835	<i>Paysonia grandiflora</i>	United States	TX	Wilson	29.16667	-97.93333	91.44
2250	W6 20854	<i>P. recurvata</i>	United States	TX	Medina	29.53333	-99.21667	434.34
2255	PI 596435	<i>P. fendleri</i>	United States	TX	El Paso	31.83333	-106.01667	1341.12
2256	PI 596436	<i>P. fendleri</i>	United States	TX	Hudspeth	31.70000	-105.41667	1240.54
2260	PI 596439	<i>P. fendleri</i>	United States	TX	Hudspeth	31.81667	-105.13333	1219.20
2262	PI 596441	<i>P. fendleri</i>	United States	TX	Culberson	31.36667	-104.81667	944.88
2269	PI 596447	<i>P. fendleri</i>	United States	TX	Pecos	30.71667	-103.18333	975.36
2273	PI 596450	<i>P. fendleri</i>	United States	TX	Brewster	30.20000	-103.56667	1554.48
2277	PI 596452	<i>P. fendleri</i>	United States	TX	Brewster	29.78333	-103.18333	853.44
2278	PI 596453	<i>P. fendleri</i>	United States	TX	Brewster	30.20000	-103.20000	1310.64
2281	PI 596454	<i>P. fendleri</i>	United States	TX	Crockett	30.93333	-101.86667	640.08
2282	PI 596455	<i>P. fendleri</i>	United States	TX	Pecos	30.96667	-101.96667	563.88
2286	PI 596456	<i>P. fendleri</i>	United States	TX	Andrews	32.31667	-102.50000	807.72
2290	PI 596457	<i>P. fendleri</i>	United States	TX	Winkler	31.78333	-103.26667	853.44
2291	PI 596458	<i>P. fendleri</i>	United States	TX	Reeves	31.11667	-103.71667	944.88
2292	PI 596459	<i>P. fendleri</i>	United States	TX	Reeves	31.01667	-103.73333	1005.84
2294	PI 596461	<i>P. fendleri</i>	United States	TX	Jeff Davis	30.53333	-104.10000	1722.12
2295	PI 596462	<i>P. fendleri</i>	United States	TX	Jeff Davis	30.55000	-103.91667	1676.40

Table 2. Passport information of additional *Physaria* and *Paysonia* accessions genotyped using DArTseq.

Collection No. and ID	Accession No.	Species	Country	State/Prov.	County	Latitude	Longitude	Elevation (m)
715	PI 293027	<i>P. fendleri</i>	United States	TX	Brewster	29.32768	-103.20527	1139.00
720	PI 293028	<i>P. fendleri</i>	United States	TX	Val Verde	29.84683	-101.68152	519.00
Barclay 998	PI 279650	<i>P. fendleri</i>	United States	NM	-	-	-	-
1836	W6 20857	<i>P. fendleri</i>	United States	NM	Luna	32.23333	-107.38333	1264.92
1838	PI 596415	<i>P. fendleri</i>	United States	NM	Chaves	32.86667	-105.03333	1621.54
1839	PI 596416	<i>P. fendleri</i>	United States	NM	Eddy	32.83333	-104.80000	1301.50
1854	W6 23374	<i>P. gordonii</i>	United States	AZ	Cochise	31.85000	-109.96667	1732.79
1855	PI 596422	<i>P. fendleri</i>	United States	AZ	Cochise	31.96667	-109.46667	1548.38
1871	W6 20862	<i>P. fendleri</i>	United States	AZ	Navajo	34.88333	-110.11667	1546.86
1879	W6 20838	<i>P. intermedia</i>	United States	AZ	Navajo	34.35000	-110.16667	1828.80
1880	PI 596424	<i>P. fendleri</i>	United States	AZ	Navajo	34.36667	-110.15000	1825.75
1902		<i>P. kaibabensis</i>	United States	AZ	Coconino	36.73333	-112.11667	2237.23
1907	PI 596427	<i>P. fendleri</i>	United States	NM	Torrance	34.55000	-106.13333	1850.14
1926	PI 596432	<i>P. fendleri</i>	United States	AZ	Navajo	34.75000	-110.16667	1615.44
1934	W6 20869	<i>P. fendleri</i>	United States	AZ	Graham	33.23333	-110.25000	908.30
2097	PI 331165	<i>P. fendleri</i>	United States	AZ	Graham	33.38333	-110.38333	917.00
2202	W6 20819	<i>P. densiflora</i>	United States	TX	Gillespie	30.20000	-98.93333	434.34
2203		<i>P. recurvata</i>	United States	TX	Kerr	30.00000	-99.36667	518.16
2217	PI 643172	<i>Paysonia lasiocarpa</i>	United States	TX	Frio	28.63333	-99.38333	198.12
2228	W6 20840	<i>Paysonia lasiocarpa</i>	United States	TX	Duval	27.58333	-98.56667	137.16
2246	W6 20836	<i>Paysonia grandiflora</i>	United States	TX	Atascosa	28.85000	-98.35000	121.92
2247	W6 20837	<i>Paysonia grandiflora</i>	United States	TX	Frio	29.06667	-98.85000	121.92
2257	PI 596437	<i>P. fendleri</i>	United States	TX	Hudspeth	31.73333	-105.21667	1097.28
2258	W6 20829	<i>P. fendleri</i>	United States	TX	Hudspeth	31.83333	-105.20000	1219.20
2259	PI 596438	<i>P. fendleri</i>	United States	TX	Hudspeth	31.81667	-105.13333	1188.72
2264	W6 20870	<i>P. fendleri</i>	United States	TX	Reeves	30.91667	-103.78333	914.40
2265	PI 596443	<i>P. fendleri</i>	United States	TX	Jeff Davis	30.86667	-103.78333	1127.76
2267	PI 596445	<i>P. fendleri</i>	United States	TX	Jeff Davis	30.68333	-103.78333	1234.44
2268	PI 596446	<i>P. fendleri</i>	United States	TX	Jeff Davis	30.48333	-103.75000	1554.48
2270	PI 596448	<i>P. fendleri</i>	United States	TX	Brewster	30.38333	-103.58333	1402.08
2274		<i>P. fendleri</i>	United States	TX	Brewster	30.03333	-103.56667	1280.16
2276	W6 20871	<i>P. fendleri</i>	United States	TX	Brewster	29.66667	-103.11667	822.96
2279	W6 20844	<i>P. mcvaughiana</i>	United States	TX	Pecos	30.46667	-102.91667	1432.56
2293	PI 596460	<i>P. fendleri</i>	United States	TX	Jeff Davis	31.00000	-104.18333	1310.64
2297	W6 20859	<i>P. fendleri</i>	United States	TX	Presidio	30.28333	-104.01667	1463.04
2401		<i>P. douglasii</i>	United States	WA	Grant	46.63333	-119.73333	89.61
2819	W6 20873	<i>P. fendleri</i>	-	-	-	-	-	-
2894		<i>Paysonia densipila</i>	United States	TN	Williamson	35.82861	-86.69722	208.00
2997	W6 20859	<i>P. fendleri</i>	United States	AZ	Apache	34.48333	-109.51667	1899.00
3029		<i>P. intermedia</i>	United States	UT	Garfield	38.02044	-112.36819	1944.00
3011	W6 20320	<i>Paysonia auriculata</i>	United States	OK	Blaine	35.79325	-98.42029	454.00
3000		<i>Paysonia lyrata</i>	United States	AL	Colbert	-	-	-
3030	PI 337050	<i>P. fendleri</i>	United States	NM	Sierra	33.36355	-107.28041	1500.00
3042	W6 20843	<i>P. ludoviciana</i>	United States	UT	Garfield	37.76753	-112.06660	1791.00
3050		<i>P. acutifolia</i>	United States	UT	San Juan	37.50160	-109.64779	1410.00
3058	PI 345712	<i>Paysonia auriculata</i>	United States	OK	Garfield	36.43716	-97.58647	360.00
3083	PI 355037	<i>P. fendleri</i>	United States	TX	Kinney	29.15197	-100.38664	288.00
3085		<i>P. acutifolia</i>	United States	CO	Mesa	38.83627	-108.56218	2073.00
3091	PI 355041	<i>Paysonia perforata</i>	United States	TN	Wilson	36.23255	-86.38455	182.00

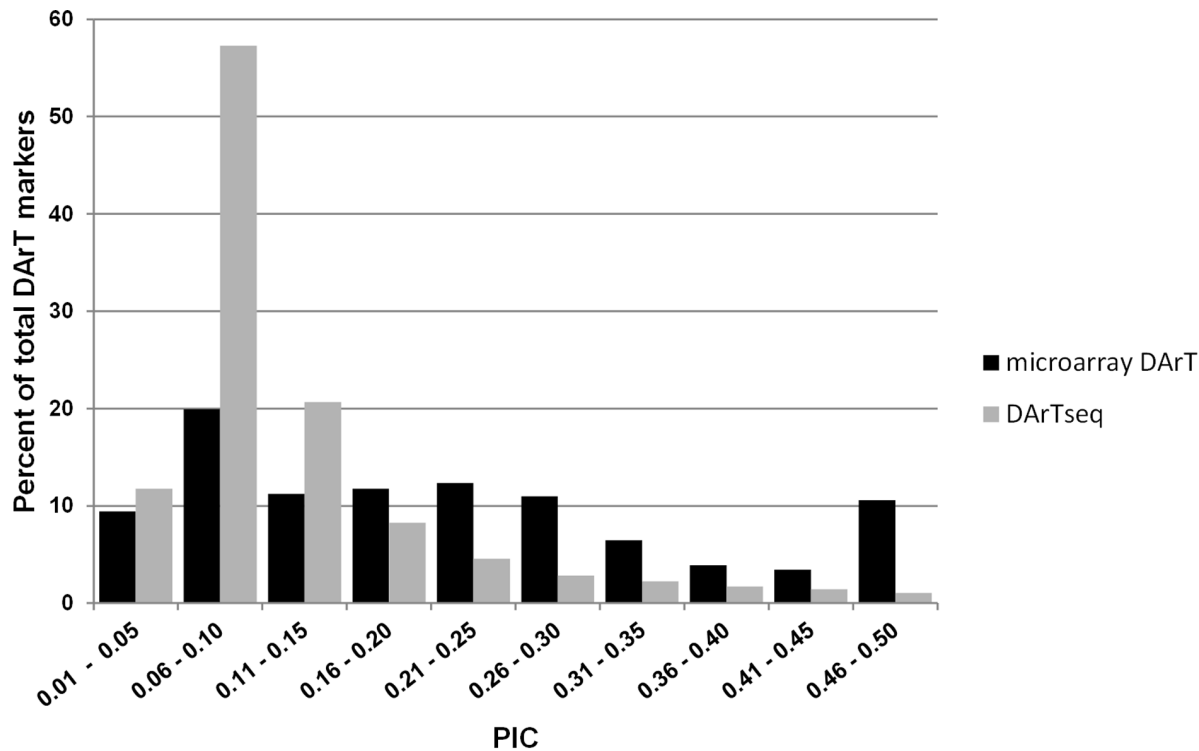


Figure 1. Distribution of marker PIC values by DArT platform.
doi:10.1371/journal.pone.0064062.g001

application. One of these parameters is reproducibility of markers in technical replicates for a subset of samples. In diversity analysis, the reproducibility parameter threshold is set usually at 97% which translates to average reproducibility of the dataset around 99.7%. A total of 87 common accessions were genotyped using both the microarray DArT and DArTseq platforms. A total of 11 *Physaria* species, 1 interspecific hybrid, and 2 *Paysonia* species were analyzed by microarray DArT. Additional accessions were genotyped using the DArTseq platform comprised of 17 *Physaria* and 7 *Paysonia* species. Overall, there was a total of 177 accessions represented by single plant samples genotyped using the two platforms, majority of which are *P. fendleri* (Tables 1 and 2). The marker sequences and genotype data will be stored in the U.S. Germplasm Resources Information Network (GRIN) database (<http://www.ars-grin.gov>) along with the accessions' phenotypic observations and germplasm passport data curated by the U.S. National Plant Germplasm System.

Data Analysis

All the images from DArT platforms were analyzed using DArTsoft v.7.4.7 (DArT P/L, Canberra, Australia). The markers were scored as binary data (1/0), indicating presence or absence of a marker in genomic representation of each sample as described by Wenzl et al. [31]. For quality control, 30% of genotypes were genotyped in full technical replication. Clones with $P > 77\%$, a call rate $> 97\%$ and $> 98\%$ allele-calling consistency across the replicates were selected as markers. P value represents the allelic-states variance of the relative target hybridization intensity as a percentage of the total variance. The informativeness of the DArT markers was determined by calculating the polymorphism information content (PIC) within the panel of diverse accessions according to Anderson et al. [32]. The *P. fendleri* data were used in GenAlEx v.6.41 [33] to determine 2D spatial autocorrelation of

the DArT microarray and Arlequin v.3.5.1.3 [34] to assess the amount of variation among the assigned regional groupings by AMOVA. To summarize the relationships among all examined accessions, cluster analysis was performed on Dice similarity values with the SAHN procedure using the unweighted pair-group method done using NTSYS-pc v. 2.21 m [35]. The Dice coefficient was preferred over simple matching coefficient because DArT is a dominant marker system and there were several non-*Physaria* species in the sample set [36], [37]. In addition, a Bayesian model-based clustering was performed on microarray DArT markers using STRUCTURE v.2.3.4 [38] testing 3 independent runs with K from 1 to 8, each run with a burn-in period of 50,000 iterations and 300,000 Monte Carlo Markov Chain (MCMC) iterations, assuming an admixture model and correlated allele frequencies. The STRUCTURE data was subsequently analyzed by HARVESTER v.06.92 [39]. Mantel tests were made to determine if there are significant correlations between the dendrogram representations and the distance matrices, between the matrices of geographic and genetic distances, and between the distance matrices from the two DArT platforms used in the *P. fendleri*. The missing geographic coordinates of collection sites of ten accessions (PI 293027, PI 293028, PI 337050, PI 345712, PI 355037, PI 355041, PI 355042, PI 275771, PI 283700 and PI 299412) in the GRIN database were estimated using Google Earth v. 6.1.0.5001 based on available locality description and collectors' notes and the information included in the analysis. Map projections of the analyzed *P. fendleri* accessions were made using ArcGIS Explorer v.2.0.0.1750 (ESRI, Redlands, CA) and non-parametric correlation tests done using JMP v.9 (SAS Institute, Cary, NC).

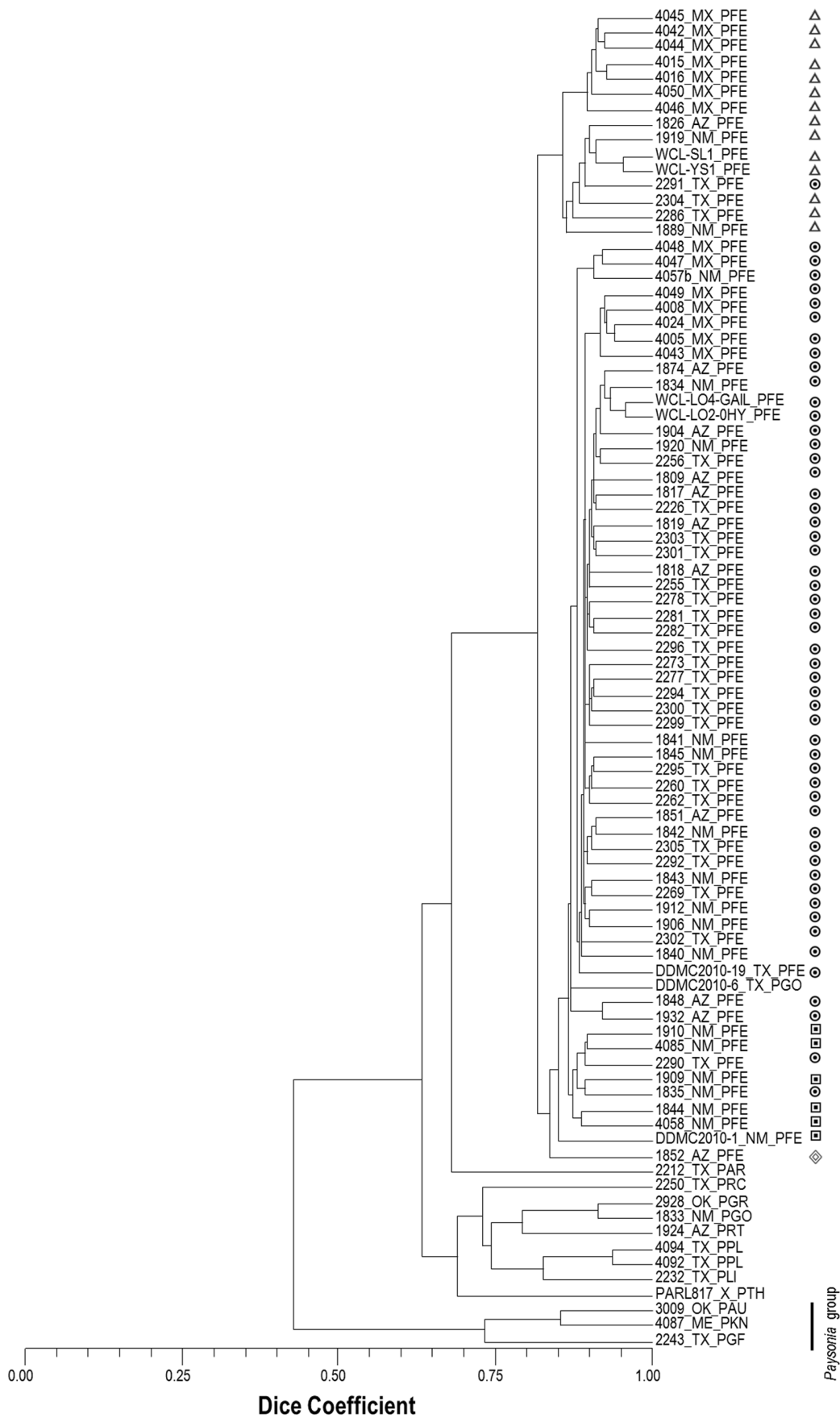


Figure 2. Cluster analysis of *Physaria* and *Paysonia* accessions based on 2,833 DArT markers. The labels denote the germplasm collection numbers and origin. Suffixes indicate respective species (PAR: *P. argyrea*, PAU: *P. auriculata*, PFE: *P. fendleri*, PGO: *P. gordonii*, PGF: *P. grandiflora*, PGR: *P. gracilis*, PKN: *P. 'kathryn'*, PLI: *P. lindheimeri*, PPL: *P. pallida*, PRC: *P. recurvata*, PRT: *P. rectipes*, and PTH: *P. thamnophila*). doi:10.1371/journal.pone.0064062.g002

Results

Relationships among Accessions from Microarray DArT Analysis

A total of 2,833 polymorphic markers were found using microarray DArT, with an average genotype call rate of 98.4% and a scoring reproducibility of 99.7%. The average PIC value was 0.21 and the median 0.19. About 20% of the markers have values in the range of 0.06 to 0.10, and almost an equal proportion of 12% of markers on the following PIC classes - 0.11 to 0.15, 0.16 to 0.20, and 0.21 to 0.25 (Figure 1). Overall, the distribution of PIC values was asymmetrical and skewed towards the lower values.

Cluster analysis and principal coordinate analysis (PC plot not shown) indicated that the different accessions were successfully classified by the marker system based on species, by geographical source, and breeding status (Figure 2), except for one new collection of *P. gordonii*, DDMC2010-6, which clustered with the *P. fendleri* accessions. The cophenetic correlation coefficient between the dendrogram and the distance matrix was highly significant ($r = 0.98$, $t = 10.30$, prob random $Z < \text{obs}$. $Z = 1.00$, 3000 permutations) indicating that the tree is a very good representation of the distance matrix. All *P. fendleri* accessions grouped in a separate cluster from the other species. The main cluster has two subgroups from Mexico, and a subgroup with a majority of accessions from Texas. There was no single group of *P. fendleri* accessions from

Arizona and New Mexico. Most accessions from these States were associated with other accessions from Texas.

The four advanced *P. fendleri* breeding lines were partitioned into two clusters. The breeding lines WCL-SL1 and WCL-YS1 were found to be more genetically distant to WCL-LO2 and WCL-LO4. The last two lines were determined to be more similar to the rest of the *P. fendleri* from North America than WCL-SL1 and WCL-YS1.

Among the different species, the most similar to *P. fendleri* was determined to be *P. argyrea* while the least similar was *P. thamnophila*. The *P. pallida* accessions grouped in one cluster along with *P. lindheimeri* and other accessions representing the species, *P. gracilis*, *P. gordonii*, *P. recurvata*, and *P. rectipes*. The two accessions of *Paysonia auriculata* (3009) and *Paysonia grandiflora* (2243) grouped together in a separate cluster along with the interspecific hybrid swarm 'Kathryn' (4087) from five *Paysonia* species.

The analysis of molecular variance considering *P. fendleri* accessions only, showed that there was a much greater proportion of variation within groups (90%) than among groups (10%) in the species (Table 3). Among the unimproved germplasm set, the pairwise F_{st} values that showed the greatest differentiation was between Arizona accessions and those from New Mexico and the least amount of differentiation between Mexico and Arizona (Table 4). There was a significant correlation found between the computed genetic distance and geographic distance matrices from Mantel test ($r = 0.33$, $t = 5.94$, prob random $Z < \text{obs}$ $z = 1.00$, 3000 permutations), indicating that distant accession pairs are more

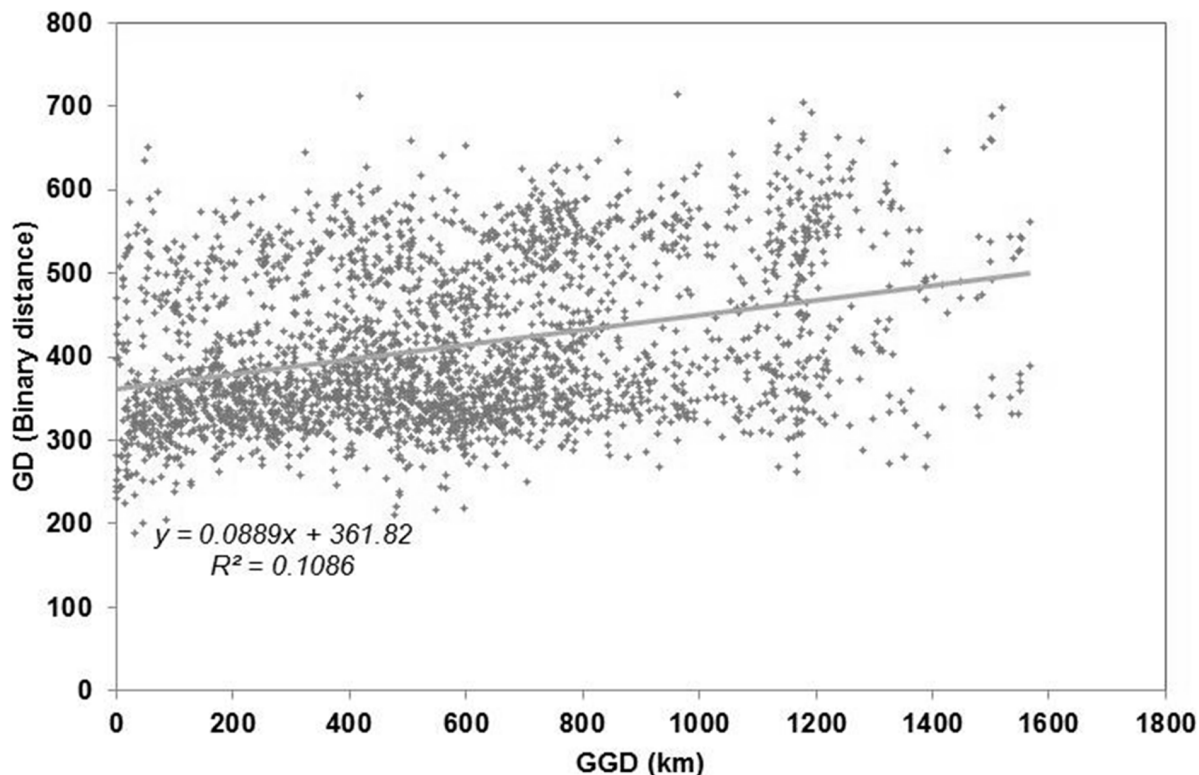


Figure 3. GenAlEx plot of geographic distance (GGD) and genetic distance (GD) from microarray DArT markers. doi:10.1371/journal.pone.0064062.g003

Table 3. Analysis of molecular variance in *P. fendleri* accessions.

Source	df	SS	Variance components	%
A. Microarray DARt				
Among Groups	4	1597.506	17.19956	9.66%
Within Groups	69	11094.561	160.79074	90.34%
Total	73	12692.068	177.99031	100%
B. DARtseq				
Among Groups	4	14299.235	98.59559	7.22%
Within Groups	123	155895.234	1267.44093	92.78%
Total	127	170194.469	1366.03651	100%

doi:10.1371/journal.pone.0064062.t003

different genetically, supporting the previously mentioned result (Figure 3).

Relationships among Accessions from DARtseq Analysis

There was a total of 27,748 markers obtained using the *Physaria* DARtseq platform. The average genotype call rate was 98.8% and a scoring reproducibility of 99.7%. The average PIC value was 0.12 and the median was 0.09. About 57% of markers have PIC values in the range of 0.06 to 0.10, while 20% and 11% on PIC classes 0.11–0.15 and 0.01–0.05, respectively (Figure 1). The distribution of the PIC values of DARtseq markers follows the same skewed pattern as the DARt microarray markers presented earlier.

Cluster analysis using the markers from this platform resulted in a relationship that follows that of the taxonomic groupings based on general morphological affinities presented by Rollins and Shaw [1]. The results did not deviate from those when the first platform with fewer markers used. Four accessions of *P. fendleri* (2258, 2274, 3083, and DDMC2010-8) clustered with the other *Physaria* species – *P. gordonii* and *P. gracilis*, indicating higher genetic similarity to representative accessions of these species than the rest of *P. fendleri*. These four *P. fendleri* accessions will be further examined for misidentification and for oil and morphological trait variation when verified.

The cophenetic correlation coefficient between the dendrogram and the distance matrix was highly significant ($r = 0.94$, $t = 49.20$, prob random $Z < \text{obs}$. $Z = 1.00$, 3,000 permutations) indicating that the tree is a very good representation of the distance matrix. The cluster of *P. fendleri* showed a distinct group of accessions from Mexico and a cluster comprised of all other germplasm from North America (Figure 4). The DARtseq platform indicated two separate clusters for the advanced lines. The breeding lines WCL-LH1, WCL-LO1, and WCL-LY1 are all in a group with greater similarity to accessions from Texas. WCL-LO2, WCL-LO4, WCL-SL1, and WCL-YS1 grouped together in one cluster indicating greater genetic similarity to accessions from Arizona and Mexico. WCL-LO2 was derived from the WCL-LO1 and the remaining three from WCL-LO2. It appears from these data that perhaps more genotypes from the Arizona accession were integrated into WCL-LO2.

Similar to results from the microarray DARt, the accessions of the following species formed distinct groups consistent with the classification by Rollins and Shaw [1]: a) *P. intermedia*, *P. valida*, and *P. rectipes*, b) *P. gordonii*, *P. gracilis*, *P. rectipes*, and *P. lindheimeri*, and c) *P. auriculata*, *P. densipila*, *P. lyrata*, *P. stonensis*, *P. grandiflora*, and *P.*

Table 4. Comparison of population pairwise Fst values using microarray DARt and DARtseq.

	Arizona	Advanced lines	Mexico	New Mexico	Texas
Arizona		0.10717	0.13076	0.03055	0.04049
Advanced lines	0.13854		0.22289	0.10951	0.08193
Mexico	0.00448	0.13043		0.12926	0.10552
New Mexico	0.16823	0.17758	0.16280		0.03056
Texas	0.01547	0.12599	0.02649	0.16289	

*Values in bold are significant at $\alpha = 0.05$. Values below diagonal are from microarray DARt and those above diagonal are from DARtseq.
doi:10.1371/journal.pone.0064062.t004

perforata. The two accessions of *P. lasiocarpa* (2217 and 2228) were most genetically similar to the accessions of *P. grandiflora*. This last group of species comprised of annual, auriculate-leaved types whose taxonomic nomenclature was segregated from *Physaria* and transferred to *Paysonia* based on leaf-trichome morphology, chromosome number, and molecular data from analyzing internal transcribed spacers of nuclear ribosomal DNA [40]. Apart from the accession of *Physaria mexicana* nested within the *Paysonia* cluster, the results of this genetic analysis using both microarray DARt and DARtseq platforms support the previous segregation of *Physaria* from *Paysonia* as proposed by O'Kane and Al-Shehbaz [40] indicating that the group of *Paysonia* species to be the least genetically similar to the *P. fendleri* and other *Physaria* accessions.

Results of the analysis of molecular variance when DARtseq markers were used correspond to that from microarray DARt. A greater proportion of variation within groups (93%) than among groups (7%) in the species was found (Table 3).

The average genetic similarity in the *P. fendleri* group when using microarray DARt was 0.86, while only 0.44 when DARtseq was used. This is in line with the assumption that more differences may be found when more markers are used because of the increased sensitivity and resolution to detect genetic distinctiveness [41]. The genetic similarity matrices of *P. fendleri* obtained using the two platform systems showed a good fit when compared by a Mantel test ($r = 0.48$, $t = 11.28$, prob random $Z < \text{obs}$ $z = 1.00$, 3,000 permutations).

Population Structure Analysis of *P. fendleri*

Using results from microarray DARt, the population structure of the *P. fendleri* samples was determined. The plot of ΔK for each K value is shown in Figure 5a. It was estimated through the method of Evanno et. al. [42] that there are 4 groups contributing significant genetic information in the *P. fendleri* collection. The bar plot of the population assignment test when $K = 4$ is shown in Figure 5b. Three of the four accessions of *P. fendleri* breeding lines are shown to have mixed backgrounds. Of the other *P. fendleri* accessions, twelve (16%) have close to homogeneous genetic background (>98% probability) while 63 accessions (84%) are highly heterogeneous showing intermediate and/or highly mixed composition. A majority of the accessions from Texas are assigned to one cluster, while those from New Mexico have myriad cluster assignments suggesting the greater diversity of *P. fendleri* in this U.S. state. When the cluster assignments of the *P. fendleri* accessions were projected on a map, the segregation among clusters was evident in their geographic location (Figure 6). Further testing for association between the assigned clusters and available site elevation data by computing Spearman's rank correlation

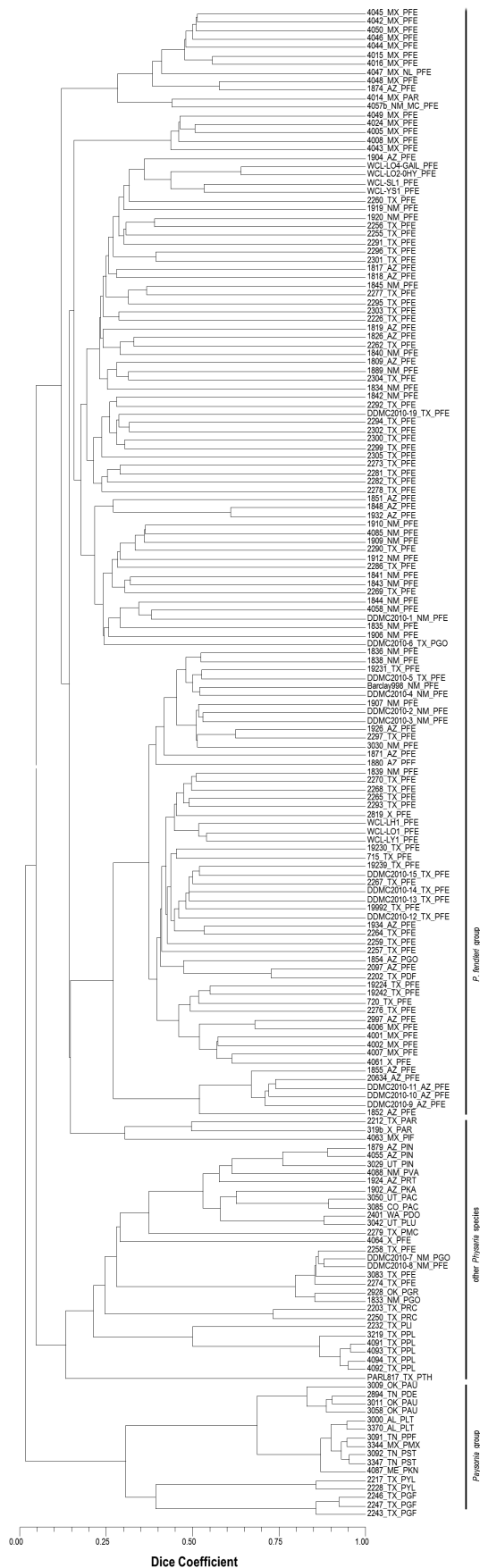


Figure 4. Cluster analysis of *Physaria* and *Paysonia* accessions based on 27,748 DArTseq markers. The labels denote the germplasm collection numbers and origin. Suffixes indicate respective species (PAC: *P. acutifolia*, PAR: *P. argyrea*, PAU: *P. auriculata*, PDE: *P. densipila*, PDF: *P. densiflora*, PDO: *P. douglasii*, PFE: *P. fendleri*, PGO: *P. gordonii*, PGF: *P. grandiflora*, PGR: *P. gracilis*, PIF: *P. inflata*, PIN: *P. intermedia*, PKA: *P. kaibabensis*, PKN: *P. 'kathryn'*, PYL: *P. lasiocarpa*, PLI: *P. lindheimeri*, PLT: *P. lyrata*, PLU: *P. ludoviciana*, PMC: *P. mcvaughiana*, PMX: *P. mexicana*, PPF: *P. perforata*, PPL: *P. pallida*, PRC: *P. recurvata*, PRT: *P. rectipes*, PST: *P. stonensis*, PTH: *P. thamnophila* and PVA: *P. valida*). doi:10.1371/journal.pone.0064062.g004

coefficient showed a very weak positive correlation (Spearman $\rho = 0.11$, $p = 0.33$).

Discussion

The importance of understanding genetic diversity in germplasm collections is critical for the effective management of accessions in genebanks. Molecular characterization supplements morphological evaluation of germplasm and allows measurements to help resolve numerous operational, logistical, and biological questions that face genebank managers and conservation biologists [43], [44]. The *Physaria* collection in the U.S. NPGS has not been well characterized before for genetic diversity, though there have been preliminary studies on subsets of accessions using a limited number of microsatellite markers [16], and an extensive evaluation for diversity in oil characteristics and other morphological characters [45]. In this study, the new DArT platforms for *Physaria* were found both acceptable and provided robust information about the genetic variability of the collection.

The development and utilization of DArT markers allowed us to determine the genetic diversity of the *Physaria* collection. The 2,833 microarray DArT markers were found to be useful in providing a picture of genetic diversity in the *Physaria* germplasm collection using a large set of accessions. Overall, the average PIC of the *Physaria* and *Paysonia* microarray DArT markers was found to be lower than that observed in other species where similar markers were developed, like wheat (0.44) [46], cassava (0.42) [47], and sorghum (0.41) [48], but comparable to that observed in sugar beet (0.28) [49] and *Asplenium* fern (0.21) [50]. The average PIC of the DArTseq markers is much less than that of the microarray DArT. However, the more numerous DArTseq markers may have the capability of providing a better picture of diversity by sampling more points in the genome. The distribution of these almost 30,000 new DArT markers in the *Physaria* and *Paysonia* genome remains to be determined. However, based on the information from a large number of organisms in which DArT system was applied more broadly including genetic mapping and/or sequence-based physical mapping, we can assume that DArT marker from both platforms will be distributed throughout the genome with marker density highly correlated to gene density [28], [51]. Compared to microsatellite markers, DArT markers are very suitable for high-throughput work and previously have been determined to have clear advantages in cost and time aspects of genotyping as demonstrated in other crops [52]. Both microarray DArT and DArTseq platforms have the same development costs. However, the higher number of markers obtained in DArTseq resulted to an overall lower cost per datapoint than microarray DArT. This higher cost effectiveness of DArTseq is in parallel to other sequenced-based genotyping strategies which can provide substantial cost savings compared to microarrays when conducting genetic diversity studies [53], [54]. Importantly, when comparing *effectiveness* of the two platforms one has to keep in mind that it may vary according to specific *application*: in genetic ID and product quality testing (i.e. seed

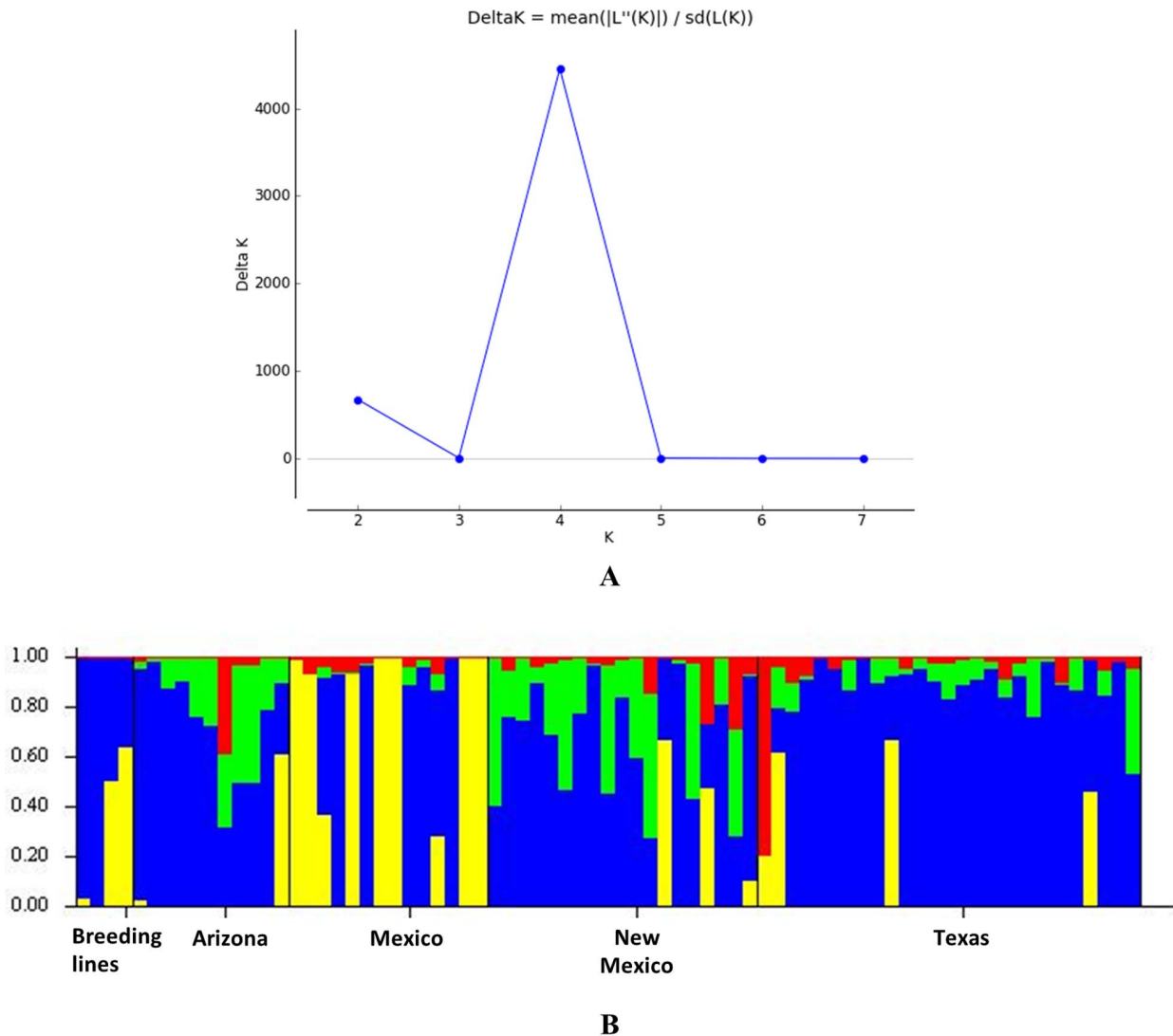


Figure 5. Plot of ΔK from $K=2$ to 7 (a) and the population structure of 75 *P. fendleri* accessions at $K=4$ (b).
doi:10.1371/journal.pone.0064062.g005

purity) modest number of array-based DArT markers may perform as well as DArTseq platform and currently for a better price. A further cost reduction of sequencing may however push even this balance towards DArTseq platform in the future.

The relationships found among the accessions are in line with the previously proposed evolution within the genus. Taxonomists assert that *P. auriculata* is the most primitive species due to its very distinct evolutionarily primitive characters such as large siliques, large number of ovules around the replum, and predominance of simple trichomes [1]. *P. auriculata* has been proposed to be closely related to *P. grandiflora* and this relationship is supported by the results of the molecular marker analysis between the accessions representative of these species (2243 and 3009) showing high genetic similarities. Likewise, *P. gordonii* and *P. gracilis* grouped in the same cluster with *P. rectipes*, *P. recurvata*, and *P. thamnophila* which is in agreement with their previously known taxonomic groupings based on very general morphological similarities. The high genetic similarity between the *P. argyrea* accessions (2212 and 319 b) to the *P. fendleri* group supports phylogenetic subsectional grouping based on pod morphology.

Overall, there was higher genetic similarity found among accessions of the other *Physaria* and *Paysonia* species than among accessions of *P. fendleri*. In particular, species that are included in the federal or state threatened and endangered species list, like *P. pallida* and *P. stonensis* [55], have very low genetic diversity as indicated by results of DArTseq markers. In *P. pallida*, the representative accessions (4091 and 4093) were found to be highly similar. The limited geographic range and the proximity of the collection sites of these two samples suggest that they might have been from just one population. Likewise for *P. stonensis*, there was a very high genetic similarity found on both of the representative accessions (3092 and 3347). Because only a limited number of accessions were included in these other species, a follow up study that includes additional accessions is recommended to validate if this is a general trend.

Physaria 'Kathryn' (4087) is a cultivar from interspecific hybridization developed by allowing five species (*P. densipila*, *P. lescurii*, *P. lyrata*, *P. perforata*, and *P. stonensis*) to intermate for twenty two generations [56]. *P. Kathryn* was determined to be most closely related to *P. auriculata* in our analysis using microarray DArT. However, with the expanded set of accessions included in

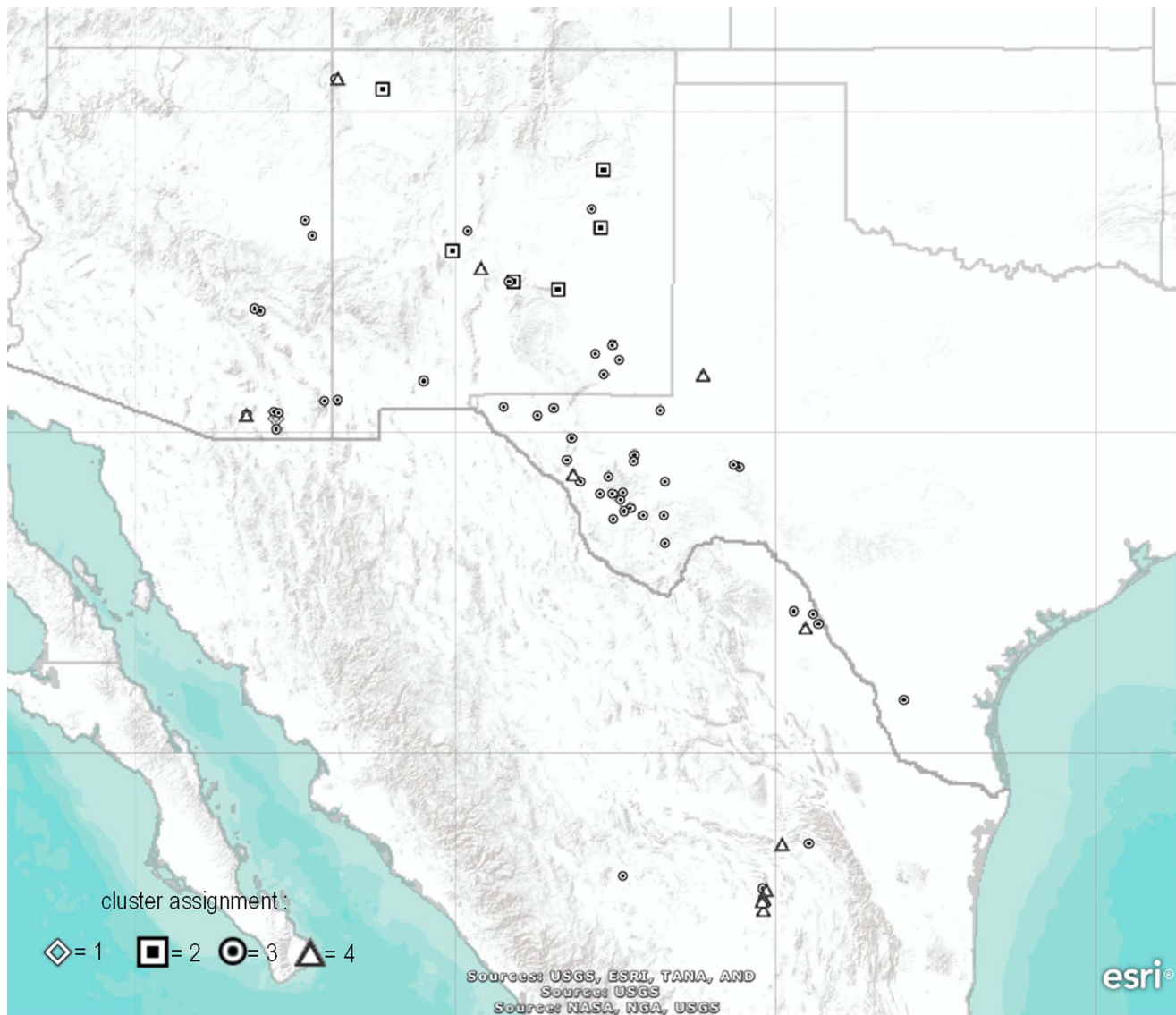


Figure 6. Geographic locations of the *P. fendleri* accessions and their respective cluster assignments (indicated by different icons) based on Bayesian model-based clustering methods.

doi:10.1371/journal.pone.0064062.g006

the DARtseq analysis, it grouped with the representative samples of its parent species – *P. lyrata* (3000 and 3370), *P. perforata* (3091), and *P. stonensis* (3092, 3347). These species were also found most genetically similar to *P. mexicana* (3344) which is the only perennial type in the species cluster. *P. mexicana* is a previously undescribed species in Mexico and is among the more recent species reported by Rollins [57].

DDMC2010-6 was from a more recent germplasm collecting trip and it was entered in the database as *P. gordonii*. However, based on results of DARt markers, it clustered with the *P. fendleri* accessions after using both DARt platforms, indicating the need to review its species assignment. The species identity of this particular accession will again be verified using its plant voucher specimen as well as in the NPGS site handling the germplasm when it is regenerated in the future.

The two clusters of *P. fendleri* breeding lines - WCL-LO2, WCL-LO4, WCL-SL1, and WCL-YS1 in one cluster, while the other set of breeding lines WCL-LO1, WCL-LH1, and WCL-LY1 in

another, indicates the possibility of developing genetically differentiated lines for crop improvement and may have applications in future hybrid development work. The lines WCL-LH1, WCL-LO1 and WCL-LY1 are the first three germplasm lines that were publicly released in 1996. These were developed using recurrent selection on a population made by bulking seeds of one accession that came from Arizona and nine from Texas in 1986 [58]. The bulked seeds were also the starting material for other breeding lines. WCL-LO4 was derived from mass selection from WCL-LO3 and WCL-LY2 (both not included in this study) which has WCL-LY1 as the source population [59], [60]. The other breeding lines were developed through phenotypic selection: WCL-YS1 was a selection from PI 311165, one of the initial accession from Arizona that comprise the original bulk in 1986 [61], while WCL-SL1 came from plants that survived at the highest salinity levels during a salt tolerance screening study when seeds from the original bulked seeds were planted [62].

The accessions of *P. fendleri* from Mexico are genetically similar as indicated by the cluster analyses. The range of *Physaria* species has been reported as limited to the northeastern part of country and concentrated on mountain and high plains of Coahuila, Nuevo León, and Zacatecas [57]. This limited geographic distribution likely prevented their further genetic differentiation. A similar investigation focusing on the other *Physaria* species in this region could confirm this trend.

The array of *P. fendleri* germplasm consisted of 75 accessions analyzed using microarray DArT and 128 accessions using DArTseq. There is ample genetic variability in the *P. fendleri* collection found, as indicated by the cluster analysis as well as analysis of population structure. This is expected for a cross pollinating species that has not been fully domesticated [63]. However, the higher within group variation detected by AMOVA using data from both DArT platforms suggests that there is only a small amount of genetic differentiation among groups in the sample as a whole. Results from the Bayesian clustering approach when comparing the geographical sources of the accessions suggested that there is more variability in New Mexico than the other *P. fendleri* locations and that there is a spatial pattern evident in the microarray DArT results. This pattern of genetic differentiation occurs outward from central Texas, the region identified as the putative origin of *Physaria* as proposed earlier by Payson [64].

An increased population differentiation has been reported in many plant species between source populations and new ones when plants colonize new habitats [65]. The more dynamic nature of *Physaria* populations in distant locations from Texas has been reported by Payson [64] and he attributed part of the process as caused by barriers that separate populations, such as soil properties and moisture availability which are very important to survival of ephemeral populations of the taxon. In *P. fendleri*, Dierig et al. [66] stated that temperature and elevation effects can also account for significant differences in reproductive capacity. Selection patterns have also been investigated by past studies reporting that non-random mating, sexual selection and dormancy characteristics play a significant role in how traits evolved in the species [67], [68]. Seed dormancy characteristics in particular cause a persistent soil seed bank which may prevent genetic

differentiation in the species because certain genotypes are reintroduced back during subsequent seasons [69], [70]. In terms of genetic resources conservation of *P. fendleri* germplasm, the DArT results suggest the existence of more variable genotypes in New Mexico and it is recommended that future collecting missions select this geographic area to possibly expand the genetic diversity in the species collection.

Conclusion

The availability of genetic diversity information of the *P. fendleri* collection will enable better germplasm management and conservation of the species. In this study we report the successful development of two DArT marker platforms that were utilized for genotyping *Physaria* and *Paysonia* accessions. This marker system complements the microsatellite markers developed previously for *Physaria*. The high number of DArT markers allows a greater resolution of genetic differences among accessions and enabled us to examine the extent of variation in the *P. fendleri* collection, as well as provide support to known taxonomic classification and recent nomenclatural changes of certain *Physaria* species to *Paysonia*. We intend to further utilize the DArT markers in developing a linkage map in *Physaria* to assist breeding efforts and for future genetic mapping studies.

Acknowledgments

USDA is an equal opportunity provider and employer. Mention of companies or commercial products does not imply recommendation or endorsement by the USDA over others not mentioned. USDA neither guarantees nor warrants the standard of any product mentioned. Product names are mentioned solely to report factually on available data and to provide specific information. We would like to thank David Ellis and Chris Richards for the valuable suggestions during the conduct of the study.

Author Contributions

Conceived and designed the experiments: VMVC AK DAD. Performed the experiments: VMVC AK DAD. Analyzed the data: VMVC AK. Wrote the paper: VMVC AK DAD. Participated in germplasm collection: VMVC DAD.

References

- Rollins RC, Shaw EA (1973) The Genus *Lesquerella* (Cruciferae) in North America. Harvard Univ. Press, Cambridge, MA. 288 p.
- Al-Shehbaz IA, O'Kane SL Jr (2002) *Lesquerella* is united with *Physaria* (Brassicaceae). *Novon* 12: 319–329.
- Kish S (2008) *Lesquerella*: The next source of biofuel. USDA Cooperative State Research, Education, and Extension Service (CSREES). Available: <http://www.csrees.usda.gov/newsroom/impact/2008/nri/pdf/lequerella.pdf>. Accessed 2012 March 29.
- Moser BR, Cermak SC, Isbell TA (2008) Evaluation of castor and lesquerella oil derivatives as additives in biodiesel and ultralow sulfur diesel fuels. *Energy & Fuels* 22: 1349–1352.
- Dierig DA, Ray DT (2009) New Crops Breeding: *Lesquerella*. In: Vollman J, Rajcan I (eds). *Oil Crops, Handbook of Plant Breeding* 4. Springer Science Media, NY. 507–516.
- Cruz VMV, Dierig DA (2012) Trends in literature on new oilseed crops and related species: Seeking evidence of increasing or waning interest. *Industrial Crops and Products* 37: 141–148.
- Jenderek MM, Dierig DA, Isbell TA (2009) Fatty-acid profile of *Lesquerella* germplasm in the National Plant Germplasm System collection. *Industrial Crops and Products* 29: 154–164.
- Wang GS, McCloskey W, Foster M, Dierig D (2010) *Lesquerella*: A winter oilseed crop for the Southwest. Arizona Cooperative Extension. The University of Arizona, Tucson, AZ. 4 p.
- Cruz VMV, Dierig DA (2010) National conservation activities on *Lesquerella* spp. for ensuring species survival and germplasm availability for new crops research and development. Proc Colorado Native Plant Society Annual Meeting. Sept. 10–12, 2010. Denver, CO.
- Mondini L, Noorani A, Pagnotta MA (2009) Assessing plant genetic diversity by molecular tools. *Diversity* 1: 19–35.
- FAO (2010) The Second Report on the State of the World's Plant Genetic Resources for Food and Agriculture. Commission on Genetic Resources for Food and Agriculture. Food and Agriculture Organization of the United Nations. Rome, Italy. 69–70.
- Börner A, Khlestkina EK, Pshenichnikova TA, Osipova SV, Kobiljski B, et al. (2012) Genetics and genomics of plant genetic resources. *Journal of Stress Physiology & Biochemistry* 8 suppl. vol. 3, S10.
- Cruz VMV, Nason J, Luhman R, Marek LF, Shoemaker RC, et al. (2006) Analysis of bulked and redundant accessions of *Brassica* germplasm using assignment tests of microsatellite markers. *Euphytica* 152: 339–349.
- Cruz VMV, Luhman R, Marek LF, Rife CL, Shoemaker RC, et al. (2007) Characterization of flowering time and SSR marker analysis of spring and winter type *Brassica napus* L. germplasm. *Euphytica* 153: 43–57.
- Hasan M, Seyis F, Badani AG, Pons-Kühnemann J, Friedt W, et al. (2006) Analysis of genetic diversity in the *Brassica napus* L. gene pool using SSR markers. *Genetic Resources and Crop Evolution* 53: 793–802.
- Salywon AM, Dierig DA (2006) Isolation and characterization of microsatellite loci in *Lesquerella fendleri* (Brassicaceae) and cross-species amplification. *Mol Ecol Notes* 6: 382–384.
- Redden R, Vardy M, Edwards D, Raman H, Batley J (2009) Genetic and morphological diversity in the Brassicas and wild relatives. Proc. 16th Australian Research Assembly on Brassicas. Sept. 14–16, 2009. Ballarat, Victoria. 5 p.
- Faltusová Z, Kučera L, Ovesná J (2011) Genetic diversity of *Brassica oleracea* var. *capitata* Gene Bank accessions assessed by AFLP. *Electronic Journal of Biotechnology*. <http://dx.doi.org/10.2225/vol14-issue3-fulltext-4http://dx.doi.org/10.2225/vol14-issue3-fulltext-4>. Accessed 2012 March 29.

19. Toledo J, Dehal P, Jarrin F, Hu J, Hermann M, et al. (1998) Genetic variability of *Lepidium meyenii* and other Andean *Lepidium* species (Brassicaceae) assessed by molecular markers. *Ann Bot* 82: 523–530.
20. Kamel EA, Hassan HZ, El-Nahas AI, Ahmed SM (2004) Molecular characterization of some taxa of the genus *Raphanus* L. (Cruciferae = Brassicaceae). *Cytologia* 69: 249–260.
21. Jaccoud D, Peng K, Feinstein D, Kilian A (2001) Diversity Arrays: a solid state technology for sequence information independent genotyping. *Nucleic Acids Res* 29: e25.
22. Cruz, unpublished data.
23. Jones N, Ougham H, Thomas H, Pasakinskiene I (2009) Markers and mapping revisited: finding your gene. *New Phytologist* 137: 165–177.
24. DArT (2012) Papers about DArT. Diversity Arrays Technology, Pty Ltd. Available: <http://www.diversityarrays.com/publications.html>. Accessed 2012 March 30.
25. Varshney RK, Glaszmann JC, Leung H, Ribaut JM (2010) More genomic resources for less-studied crops. *Trends in Biotechnology* 28: 452–460.
26. Howard EL, Whittock SP, Jakše J, Carling J, Matthews PD, et al. (2011) High-throughput genotyping of hop (*Humulus lupulus* L.) utilising diversity arrays technology (DArT). *Theor Appl Genet* 122: 1265–1280.
27. Cabin RJ, Mitchell RJ, Marshall DL (1998) Do surface plant and soil seed bank populations differ genetically? A multipopulation study of the desert mustard *Lesquerella fendleri* (Brassicaceae). *Am J Bot* 85: 1098–1109.
28. Kilian A, Wenzl P, Huttner E, Carling J, Xia L, et al. (2012) Diversity Arrays Technology: A Generic Genome Profiling Technology on Open Platforms. *Methods Mol Bio* 888: 67–88.
29. O’Kane SL Jr, Al-Shehbaz IA (2003) Phylogenetic position and generic limits of *Arabidopsis* (Brassicaceae) based on sequences of nuclear ribosomal DNA. *Ann. Missouri Bot. Gard.* 90: 603–612.
30. Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 10: R25.
31. Wenzl P, Carling J, Kudrna D, Jaccoud D, Huttner E, et al. (2004) Diversity arrays technology (DArT) for whole genome profiling of barley. *Proc Natl Acad Sci (USA)* 101: 9915–9920.
32. Anderson JA, Churchill GA, Sutriquet JE, Tanksley SD, Sorrells ME (1993) Optimizing parental selection for genetic linkage maps. *Genome* 36: 181–186.
33. Peakall R, Smouse P (2006) GenAlEx 6: Genetic Analysis in Excel. Population genetic software for teaching and research. *Mol Ecol Notes* 6: 288–295.
34. Excoffier L, Laval G, Schneider S (2005) Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* 1: 47–50.
35. Rohlf FJ (2011) NTSYSpc: Numerical Taxonomy System, ver. 2.21 m. Exeter Publishing, Ltd., Setauket, NY.
36. Dalirfateh S, Meyer A, Mirhoseini S (2009) Comparison of similarity coefficients used for cluster analysis with amplified fragment length polymorphism markers in the silkworm, *Bombyx mori*. *Journal of Insect Science* 9: 71.
37. Sesli M, Yegenoglu ED (2010) Comparison of similarity coefficients used for cluster analysis based on RAPD markers in wild olives. *Genet Mol Res* 9: 2248–2253.
38. Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155: 945–959.
39. Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* 4: 359–361.
40. O’Kane SL Jr, Al-Shehbaz IA (2002) *Paysania*, a new genus segregated from *Lesquerella* (Brassicaceae). *Novon* 12: 379–381.
41. Agarwal M, Shrivastava N, Padh H (2008) Advances in molecular marker techniques and their applications in plant sciences. *Plant Cell Rep* 27: 617–631.
42. Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* 14: 2611–2620.
43. Hedrick PW (2000) Applications of population genetics and molecular techniques to conservation biology. In: *Conservation Biology 4. Genetics, Demography and Viability of Fragmented Populations*. Young, A.G, Clarke, G.M. (eds). Cambridge Univ. Press, Cambridge, UK. 113–125.
44. Karp A, Kresovich S, Bhat KV, Ayad WG, Hodgkin T (1997) Molecular tools in plant genetic resources conservation: a guide to the technologies. IPGRI Technical Bulletin No. 2. International Plant Genetic Resources Institute, Rome, Italy. 47 p.
45. Salywon AM, Dierig DA, Rebman JP, Jasso de Rodriguez D (2005) Evaluation of new *Lesquerella* and *Physaria* (Brassicaceae) oilseed germplasm. *Am J Bot* 92: 53–62.
46. Raman H, Stodart BJ, Cavanagh C, Mackay M, Morell M, et al. (2010) Molecular diversity and genetic structure of modern and traditional landrace cultivars of wheat (*Triticum aestivum* L.). *Crop and Pasture Science* 61: 222–229.
47. Xia L, Peng KM, Yang SY, Wenzl P, de Vicente MC, et al. (2005) DArT for high-throughput genotyping of cassava (*Manihot esculenta*) and its wild relatives. *Theor Applied Genet* 110: 1092–1098.
48. Mace ES, Xia L, Jordan DR, Halloran K, Path DK, et al. (2008) DArT markers: diversity analyses and mapping in *Sorghum bicolor*. *BMC Genomics*. 9, 26. doi: 10.1186/1471-2164-9-26. Accessed 2012 March 29.
49. Simko I, Eujayl I, van Hintum TJJ (2012) Empirical evaluation of DArT, SNP, and SSR marker-systems for genotyping, clustering, and assigning sugar beet hybrid varieties into populations. *Plant Science* 184: 54–62.
50. James KE, Schneider H, Ansell SW, Evers M, Robba L, et al. (2008) Diversity Arrays Technology (DArT) for pan-genomic evolutionary studies of non-model organisms. *PLoS ONE* 3: e1682.
51. Petroli CD, Sansaloni CP, Carling J, Steane DA, Vaillancourt RE, et al. (2012) Genomic characterization of DArT markers based on high-density linkage analysis and physical mapping to the *Eucalyptus* genome. *PLoS ONE* 7: e44684.
52. Kilian A, Huttner E, Wenzl P, Jaccoud D, Carling J, et al. (2005) The fast and the cheap: SNP and DArT-based whole genome profiling for crop improvement. In: *Tuberosa R, Phillips RL, Gale M (eds) Proceedings of the international congress in the wake of the double helix: from the green revolution to the gene revolution*. Avenue Media, Bologna, Italy, 27–31 May 2003, 443–461.
53. Illumina (2013) Agrigenomics genotyping decisions reach a crossroads. Application Spotlight: Analyzing Genetic Variation. Available: http://www.illumina.com/Documents/products/appspotlights/app_spotlight_ngg_ag.pdf. Accessed 2013 February 13.
54. Ayling S (2012) Technical appraisal of strategic approaches to large-scale germplasm evaluation. The Genome Analysis Center, Norwich, UK. Available: <http://agro.biodiver.se/wp-content/uploads/2012/12/Technical-appraisal-NGS-for-genebanks-please-comment.pdf>. Accessed 2013 February 13.
55. USDA-NCRS (2012) The PLANTS Database. National Plant Data Team, Greensboro, NC. Available: <http://plants.usda.gov>. Accessed 2012 September 24.
56. Rollins RC (1988) A population of interspecific hybrids of *Lesquerella* (Cruciferae). *Systematic Botany* 13: 60–63.
57. Rollins RC (1958) Notes on *Lesquerella* (Cruciferae) in Mexico. *Boletín de la Sociedad Botánica de México* 23: 42–47.
58. Dierig DA, Thompson AE, Coffelt TA (1998) Registration of three *Lesquerella fendleri* germplasm lines selected for improved oil traits. *Crop Sci* 38: 287.
59. Dierig DA, Dahlquist GH, Coffelt TA, Ray DT, Isbell TA, et al. Registration of WCL-LO4-Gail lesquerella with improved harvest index. *J Plant Reg* “submitted”.
60. Dierig DA, Dahlquist GA, Tomasi PM (2006b) Registration of WCL-LO3 high oil *Lesquerella fendleri* germplasm. *Crop Sci* 46: 1832–1833.
61. Dierig DA, Tomasi PM, Coffelt TA, Rayford WE, Lauer L (2000) Yellow seed coat *Lesquerella*. *Crop Sci* 40: 865.
62. Dierig DA, Shannon MC, Grieve CM (2001) Salt tolerant *Lesquerella*. *Crop Sci* 41: 604.
63. Rauf S, Teixeira da Silva JA, Khan AA, Naveed A (2010) Consequences of plant breeding on genetic diversity. *Intl J Plant Breeding* 4: 1–21.
64. Payson EB (1921) A Monograph of the Genus *Lesquerella*. *Annals of the Missouri Botanical Garden* 8: 103–236.
65. Barrett SCH, Husband BC (1990) The Genetics of Plant Migration and Colonization. In: *Brown AHD, Clegg MT, Kahler AL, Weir BS (eds). Plant population genetics, breeding, and genetic resources*. Sinauer Associates Inc., Sunderland, MA. 254–277.
66. Dierig DA, Adam NR, Mackey NR, Dahlquist GH, Coffelt TA (2006a) Temperature and elevation effects on plant growth, development, and seed production of two *Lesquerella* species. *Industrial Crops and Products* 24: 17–25.
67. Cabin RJ, Evans AS, Mitchell RJ (1997) Do plants derived from seeds that readily germinate differ from plants derived from seeds that require forcing to germinate? A case study of the desert mustard *Lesquerella fendleri*. *Ann Midl Nat* 138: 121–133.
68. Mitchell RJ, Marshall DL (1998) Nonrandom mating and sexual selection in a desert mustard: an experimental approach. *Am J Bot* 85: 48–55.
69. Cabin RJ (1996) Genetic comparisons of seed bank and seedling populations of a perennial desert mustard, *Lesquerella fendleri*. *Evolution* 50: 1830–1841.
70. Freeland J (2005) *Molecular Ecology*. John Wiley & Sons, Ltd., Hoboken, NJ. 109–154.