

RESEARCH ARTICLE

Biodiversity in Ethiopian linseed (*Linum usitatissimum* L.): molecular characterization of landraces and some wild species

Worku Negash Mhired · J. S. Heslop-Harrison

Received: 6 January 2018 / Accepted: 26 March 2018
© The Author(s) 2018

Abstract Molecular characterization of germplasm is important for sustainable exploitation of crops. DNA diversity was measured using inter-retrotransposon-amplified-polymorphism and inter-simple-sequence-repeat markers in 203 Ethiopian landraces and reference varieties of linseed (flax, *Linum usitatissimum*) and wild *Linum* species. Molecular diversity was high (PIC, 0.16; GD, 0.19) compared to other reports from the species. Genotyping separated reference from landrace accessions, and clustered landrace accessions from different altitudes and geographical regions. Collections showed evidence for recent introduction of varieties in some regions. The phylogeny supported *L. bienne* Mill. as the progenitor of domesticated *L. usitatissimum*. Markers developed here will be useful for genetic mapping and selection of breeding lines. The results show the range of characters that can be exploited in breeding lines

appropriate for smallholder and commercial farmers in Ethiopia, producing a sustainable, secure, high-value crop meeting agricultural, economic and cultural needs.

Keywords Biodiversity · Ethiopia · Linseed · *L. usitatissimum* · Landraces · *L. bienne* · Altitude · Region · Diversity

Introduction

Knowledge about the extent and distribution of genetic diversity in crop plants is important for any breeding program (Rahman et al. 2016; Khan et al. 2007). *Linum usitatissimum* L. is a multipurpose crop grown in many environments for fibre, food, industrial, feed and potentially pharmaceutical uses. The bast or phloem fibres have outstanding mechanical properties (Fernández 2016) with strength and flexibility. Fibre orientation and ratio of fibre length-to-diameter are important characters (Sliseris et al. 2016). Linseed oil has high α -linolenic acid (an omega-3 fatty acid) content (Scharadt 2005) and seeds are used whole or crushed in foods. The hardening oil is used for varnishes, linoleum, putty and for leather preparation. It is in local medicines as demulcent, emollient and laxative, and is taken orally in bronchial infection and diarrhoea (Gul et al. 2016). Cellulosic and lignified material remaining after extraction of fibre or oil is

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s10722-018-0636-3>) contains supplementary material, which is available to authorized users.

W. N. Mhired (✉)
Department of Biology, University of Gondar, Gondar,
Ethiopia
e-mail: mhired3@gmail.com;
worku.negash@uog.edu.et

J. S. Heslop-Harrison
Department of Genetics and Genome Biology, University
of Leicester, Leicester LE1 7RH, UK

used for manufacture of paper and straw board, and for animal feed or bedding. As a food, linseed is eaten with *wot* (a stew) particularly in the north Ethiopia and used to make a beverage of fasting days. Linseed is one of the oldest cultivated plants and was an important commercial crop before the invention of petroleum and extensive use of cotton (McHuguen 1990). Because of the diverse uses and sustainability, there is renewed interest in its cultivation. Additional information on the genetic basis of variation is required to enable modern breeding exploiting the biodiversity in the species and its wild relatives (McKenzie et al. 2008; Kurt and Evans 1998). In Ethiopia, a Vavilovian centre of crop diversity (Zohary and Hopf 2000) and a major part of the Horn of Africa endemism hot-spot (Harrison and Noss 2017), linseed is valued for food and is also exported. We have assessed the morphological variation in the highly diverse Ethiopian germplasm (Worku et al. 2015), which is grown over a wide topical and sub-tropical environments from 3 to 15°N and from 1200 to 3500 m.a.s.l. Ethiopia has several agro-ecological zones (Hurni 1998) and other crops such as durum wheat (Mengistu 2016); enset (*Ensete ventricosum* (Welw.) Cheesm.) (Olango et al. 2015); tef (*Eragrostis tef* (Zucc.)) Trotter (Bedane et al. 2015; Ayalew et al. 2015); coffee (Tadele et al. 2014); sesame (Gebremichael and Parzies 2011); and barley (Muhe and Assefa 2011) have many landraces shown to have high genetic diversity in Ethiopia.

DNA-based molecular markers have advanced genetic studies in the last three decades. Reports from different authors (Uysal et al. 2010; Cloutier et al. 2009; Fu et al. 2002; Wiesnerova and Wiesner 2004) showed that RAPDs, AFLPs, SSRs and ISSRs have low diversity in linseed germplasm, although Pali et al. (2015) reported higher variation within 48 Indian accessions. Oil seed cultivars have been considered to have more genetic diversity than fibre cultivars (Fu et al. 2002). As might be expected from genetic bottlenecks during selection, diversity is higher among wild species, and within cultivated germplasm it decreases from landraces to breeding lines and then to cultivars (Smykal et al. 2011; Habibollahi 2015). The wide variation in chromosome number found in the genus *Linum* ($2n = 16$ up to 84) indicates that hybridization and polyploidy have played a role in its evolution (Bolsheva et al. 2015). To contribute to a better understanding of the genetic diversity and

species relationships in this genus, Bolsheva et al. (2015) recommended comparative studies of karyotypes, and the development of suitable molecular markers will complement this approach.

A high proportion of the linseed genome is composed of repetitive DNA sequences with motifs present in tens to millions of copies (Cullis 2005). Primers amplifying DNA between SSRs (inter simple sequence repeat, ISSR), retroelements (inter retroelement amplified polymorphisms, IRAP) or SSRs and retroelements (retrotransposon-microsatellite amplification polymorphism, REMAP) have proved to be valuable as informative and polymorphic markers (Kalendar and Schulman 2006; Teo et al. 2005; Alsayied et al. 2015). Both SSR (evolving by slippage replication or recombination) and retroelements (amplifying through an RNA intermediate and reinserting in the genome) can evolve rapidly and hence show polymorphisms in germplasm with little diversity. Different markers reveal different classes of variation (Powell et al. 1996), and evolve at different rates, so a primer useful in distinguishing varieties of a crop may be too polymorphic to be of use in wild germplasm of the same species. Saeidi et al. (2006, 2008) analysed a collection of wild *Aegilops tauschii* Coss. germplasm from Iran, finding that SSR markers developed for use in the D genome (derived from *Ae. tauschii*) of wheat cultivars were very highly polymorphic. Thus, while useful in the cultivated hexaploid, the SSR markers gave little information about relationships in the wild species because of their fast evolutionary rate compared to the time of diversification of the accessions. In contrast, within the *A. tauschii* subspecies, another marker class, IRAP markers, were able to show phylogeographic evolutionary patterns over the northern Iran region.

Here, we aimed to exploit molecular (ISSR and IRAP) markers to determine the levels and patterns of genetic diversity and polymorphism of Ethiopian linseed accessions comparing with other international lines and wild relatives. This work towards developing a functional genomic map underpins work to identify candidate genes for agronomic or morphological characters; provides data on linseed crop diversity; can help choose parents for crosses in breeding program; and suggests appropriate sites for increased activity in germplasm conservation or collection.

Materials and methods

Plant materials

Linseed germplasm from cultivated and some wild *Linum* species included 203 lines of which 192 were linseed accessions (of these 17 genotypes were segregating landraces which were divided during the study) and 11 (one accession represented by two genotypes) were accessions from seven wild *Linum* relatives (Table S1). The germplasm was acquired from international and in country gene banks, Research Centres, and local farmers' on-farm holdings (Worku et al. 2015). One hundred and four (92 original + 12 segregating) accessions came from the Ethiopian Biodiversity Institute (EBI); 22 (17 + 5 segregating) lines from Ethiopian Agricultural Research Centres; 57 collections from local farmers; six varieties (AC Carnduff, AC Emerson, AC McDuff, AC Watson, CDC Bethune and Macbeth: AC McDuff represented by two genotypes) were from Canadian Crop Development Centres; one Irish origin linseed cultivar; 10 (PI 253971 with two genotypes) accessions of six wild *Linum* species from USDA (PI 650336, PI 650322, PI 650318, PI 650315, PI 522290, PI 253971, PI 253971, PI 231886, PI 650297, PI 440472); and *L. volkensii* Engl. was a new collection from the Ethiopia. Germplasm from Agricultural Research Centres is coded ARCxx and from local farmers is coded WLxxxx. Selection of accessions acquired from EBI took into consideration their distribution to represent the different parts of the country and agro-ecosystems.

For spatial diversity analysis altitude information grouped into eight classes using Agarwal (1996) formula: $I = \left(\frac{L-S}{K}\right)$; where I is class width; L is the largest (3449 m.a.s.l.) and S is the smallest (1410 m.a.s.l.) values from altitude records, respectively; K is number of classes obtained from $K = 1 + 3.322 \log_{10} n$; and n is total number of observations, which is 130. Former administrative regional divisions of Ethiopia were used to study the regional diversity in the germplasm.

DNA isolation, ISSR and IRAP-PCR analysis

Genomic DNA was isolated from freshly harvested leaves of a young, single plant per accession. One

hundred and fifty mg lyophilized dry leaves from 30 to 45-day old seedlings were ground to fine powder using a Silamat plus Vivadent machine in 1.5 mL micro centrifuge tube with 8–10 glass beads. To the finely ground and chilled powdered leaves, a pinch of PVP (polyvinylpyrrolidone) and 1 mL preheated 2% CTAB with β -mercaptoethanol solution was added. Then the homogenized content was transferred to 2 mL tube and incubated for 30 min at 65 °C in a water bath and mixed occasionally 2–3 times by inverting the tubes. Tubes were cooled to ambient temperature and 500 μ L chloroform:isoamyl alcohol (24:1) was added to each tube. The content in the tubes was rocked at 180° with 36 turns/min for 30 min using an orbital shaker, the centrifuged for 10 min at 13,000 rpm and the supernatant transferred to a new 2 mL tube and the chloroform extraction steps repeated for the second time. To the final supernatant 600 μ L ice-cold isopropanol was added, rocked gently and high molecular weight DNA was spooled out. The DNA extracts were washed using 1 mL wash buffer and dried then the dried pellets were again dissolved in 150 to 250 μ L TE pH 8.0 buffer. RNase treatment was conducted and DNA quality and concentration was checked using 0.8% agarose gel electrophoresis and a Nanodrop (Thermo Scientific NanoDrop-1000 Spectrophotometer, USA) for A260/A280 and A260/A230 ratios. Finally, the DNA extracts were stored at – 20 °C until needed for PCR and other experiments.

Two ISSR primers (3PCT1 and 3PCT2) from Wiesner et al. (2001), and a set of six IRAP primers: Sukula, Nikita, 3'/LTR, 5'/LTR, LTR6149 and LTR6150 (Kalendar et al. 2006; Teo et al. 2005; Saeidi et al. 2008) were used (Table S2). After tests with some genotypes and optimization of annealing temperature, two ISSR and six IRAP primers giving clear and polymorphic patterns were used for further analysis of DNA from single plants of the accessions.

PCR amplification reactions were conducted in a reaction volume of 15 μ L using appropriate single or pairs of primers, 8.15–9.15 μ L sigma water, 1.5 μ L 10 \times buffer, 1 μ L of 250 mM MgCl₂, 1.2 μ L of 10 mM dNTP mix, 1 μ L forward primer, 1 μ L reverse primer where needed, 1 μ L of 100 ng template DNA and 0.15 μ L Taq polymerase (KAPA Biosystems). The annealing temperature was optimized using gradient PCR and the annealing temperatures for 3PCT2, Sukula, Nikita and 3'/LTR were 60, 48, 44, and 48 °C, respectively. The PCR programme for ISSR

(3PCT2) primer was set at 94 °C for 4 min for initial denaturation and then for 30 cycles each cycle comprised 1 min denaturation at 94 °C, 45 s annealing at 60 °C temperature, 2 min extension at 72 °C with 7 min final extension at the end of the 30th cycle and pause at 4 °C. The PCR programme for IRAP primers was 95 °C for 2 min for initial denaturation and then for 30 cycles each cycle comprised 1 min denaturation at 95 °C, 45 s ramping + 0.5 Cs⁻¹ to annealing temperature for each primer, 2 min + 3 s extension at 72 °C with 10 min final extension at the end of the 30th cycle, and pause at 4 °C. Fragments were size-separated by electrophoresis in 2% agarose gels, for IRAP primers, using one part “high resolution” agarose and three parts normal agarose.

Analysis

Each marker band was scored for presence (1) or absence (0), and the data were analysed using the genetic analysis package PowerMarker ver. 3.23. Genetic distances were calculated from the data table using Nei et al. (1983) frequencies between each pair of accessions (Lapitan et al. 2007; Yoon et al. 2012). ISSR and IRAP data was analysed using Neighbour-joining with Nei et al. (1983) genetic distances. In the cluster analyses for the associations between each of two individuals or groups of assays were replicated. Based on the number of individuals or groups of individuals, 100 and 1000 replications were made for cluster analyses or for grouping assays into more genetically related groups. Bootstrap percentages from 50 to 70% were used to determine members of a cluster and phylogeny trees were displayed using Geneious (www.Geneious.com).

Results

ISSR and IRAP primers amplified multiple, polymorphic products from genomic DNA of linseed (*L. usitatissimum*) and seven wild species. The total numbers of bands, degree of polymorphism, number of unique alleles, range of band size, gene diversity (GD) and polymorphism information content (PIC) exhibited by ISSR and IRAP primers from linseed and wild genotypes were calculated (Table 1a–d). In total, 203 genotypes studied using four optimized PCR primer amplifications, 435 independent marker bands

were scored in 41 gels; 88,305 band positions were scored for presence or absence of marker bands. From the 203 defined accessions, 11,008 marker bands (with only 4.80% missing data where amplification or gel failed or was not tested) were analysed. Genotypes were split into cultivated and wild species groups and the results from the parameters for each group are indicated in Table 1a–d. Wild species were polymorphic with all primers and had multiple (up to 36 out of 59) unique alleles (Table 1c). Table 1d shows analysis of the Ethiopian flax and linseed accessions excluding the Canadian reference lines; three of the 22 unique alleles were present in the Canadian germplasm only.

The gene diversity (GD) value at individual primer level varied from 0.004 to 0.500 in the assayed genotypes (for the related metric of polymorphism information content (PIC), the maximum value was 0.375). The mean GD value from regional groups of genotypes ranged between 0.209 ± 0.152 from Gondar region and 0.385 ± 0.139 from Kefa region (mean PIC 0.176 ± 0.113 and 0.301 ± 0.106 respectively). The Canadian lines scored 0.210 ± 0.188 mean GD (0.171 ± 0.148 PIC), similar to the least diverse Gondar/Shewa accessions. It was also notable that the lines from the ARC Ethiopian research centres had relatively little diversity (GD 0.242 ± 0.150 and PIC 0.202 ± 0.113).

Germplasm was grouped into its origin in 13 former administrative regions of Ethiopia where these data were available (see Fig. 1, Table S1 and Table 2a). Individual regional groups of accessions (Table 2a) had showed polymorphism levels from 61.90 to 99.65%: the lowest was in Canadian lines and the highest from Gondar region. Ethiopian accessions with collection information were grouped into eight altitude classes (Table 2b). Like regional groups, all altitude classes showed less than 100% polymorphism (91.11 to 96.82%) with the least from altitude 7 (high, around 3000 m) and the most from altitude 2 (low, below 2000 m), almost from the extremes of altitude ranges.

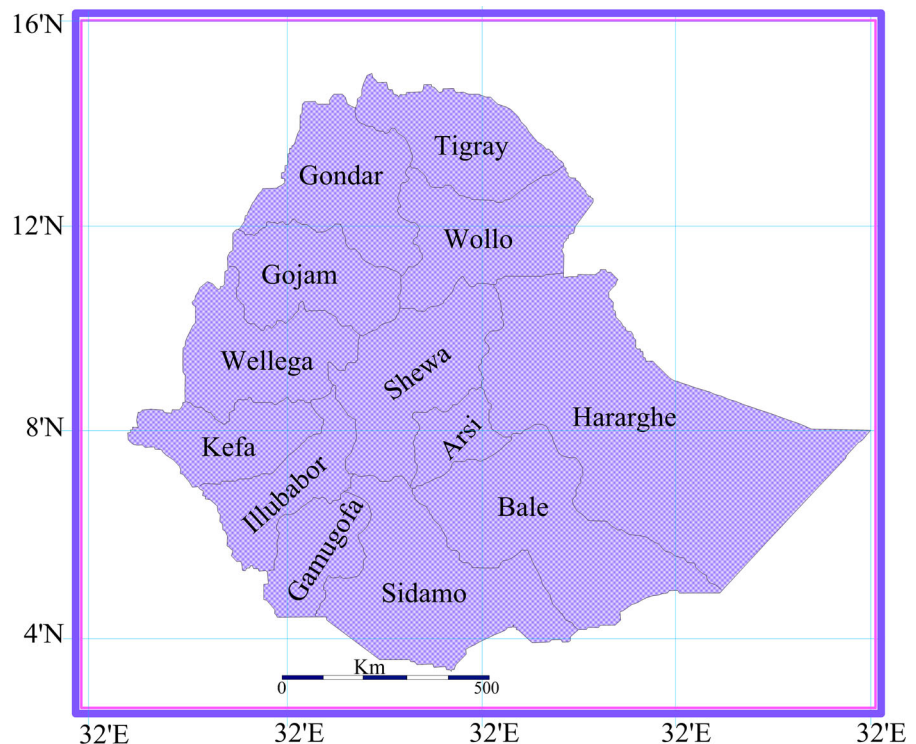
Phylogenetic relationships

The genetic-marker distance data were used for phylogenetic analysis. The overall 203 genotypes were split into different populations based on their species, region and altitude classifications and used to create trees.

Table 1 Polymorphism and gene diversity exhibited by ISSR and IRAP primers: (a) in 193 accessions from cultivars + 11 accessions from seven wild *Linum* spp; (b) in 193 linseed accessions (185 from Ethiopia + seven [one doubled] from Canada + one from Ireland); (c) in 11 accessions [one doubled] from seven wild spp. [one from Ethiopia]; and (d) in 185 Ethiopian linseed accessions

Primer	Ta	No. of observation	Total bands	Polymorphic bands	Degree of polymorphism (%)	No. of unique allele	Band size		Gene diversity	PIC
							Min	Max		
a										
ISSR	3PCT2	60	199	102	100	25	190	3000	0.149 ± 0.157	0.125 ± 0.122
IRAP	Sukula	48	184	114	100	25	228	3169	0.161 ± 0.154	0.136 ± 0.121
	Nikita	44	199	117	100	28	165	3043	0.143 ± 0.145	0.122 ± 0.115
	3'LTR	48	191	155	100	21	264	2858	0.140 ± 0.109	0.125 ± 0.089
							Mean		0.148 ± 0.140	0.127 ± 0.110
b										
ISSR	3PCT2	60	188	70	98.59	2	190	3000	0.208 ± 0.160	0.174 ± 0.121
IRAP	Sukula	48	173	84	100	1	360	3169	0.213 ± 0.152	0.179 ± 0.117
	Nikita	44	188	98	100	13	165	3043	0.168 ± 0.147	0.143 ± 0.115
	3'LTR	48	180	133	100	6	264	2858	0.165 ± 0.108	0.145 ± 0.087
							Mean		0.184 ± 0.140	0.157 ± 0.108
c										
ISSR	3PCT2	60	11	59	100	36	236	3000	0.252 ± 0.119	0.213 ± 0.083
IRAP	Sukula	48	11	63	100	32	228	2754	0.269 ± 0.120	0.226 ± 0.083
	Nikita	44	11	48	100	25	343	3043	0.279 ± 0.132	0.232 ± 0.090
	3'LTR	48	11	44	100	23	335	2500	0.279 ± 0.134	0.232 ± 0.092
							Mean		0.269 ± 0.125	0.225 ± 0.086
d										
ISSR	3PCT2	60	180	70	98.57	1	190	3000	0.208 ± 0.163	0.174 ± 0.123
IRAP	Sukula	48	165	84	100	3	360	3169	0.214 ± 0.154	0.179 ± 0.118
	Nikita	44	180	95	100	13	165	3043	0.170 ± 0.146	0.145 ± 0.114
	3'LTR	48	172	133	100	6	264	2858	0.165 ± 0.110	0.145 ± 0.088
							Mean		0.185 ± 0.141	0.158 ± 0.109

Fig. 1 Map of Ethiopia: with Latitude–Longitude position information (Wikimedia Commons); and its former administrative regions



The 11 wild *Linum* accessions (Table S1) were chosen to represent four sections of the genus and various chromosome numbers ($2n = 16, 18, 20, 28, 30$). *L. hirsutum* L. was used as an outgroup: it is in section Dasylinum, and groups with *L. volkensii* (an old-world yellow-flowered flax) and *L. trigynum* L. (both section Linopsis; see McDill et al. 2009). The two *L. austriacum* L. (sect. Linum), from German and Russia place in different groups: from Russia forms independent cluster and from German groups with *L. narbonense* L. (sect. Linum) and *L. flavum* L. (sect. Syllinum). The four *L. bienne* accessions (sect. Linum) form a well-supported cluster.

When grouped by the 13 former regions (Fig. 1) used for collections, there were five reasonably well-supported clusters with greater than 65% bootstrap (Fig. 2b). Accessions from North Ethiopia (Tigray), grouped into cluster I; accessions from North and northern central parts of Ethiopia (Gondar and Shewa regions) grouped into cluster II; accessions from northwest, southwest and South Ethiopia (Gojam, Wellega, Illubabor, Kefa, Gomugofa and Sidamo) grouped into cluster III; accessions from southeast Ethiopia (Bale and Arsi) grouped into cluster VI; and

accessions from eastern Ethiopia (Wollo and Hararghe) grouped into cluster V.

The eight altitude classes from a range of 1410–3449 m (Table 2b) were grouped into four clusters by the relationship tree: the high altitude class 8 (3195–3449 m); altitude classes 1 (1410–1665 m) and 3 (1920–2174); altitude classes 4, 5, 6 and 7 (2175–3194 m); and altitude class 2 (1665–1919).

Discussion

Genetic diversity and polymorphism

IRAP and ISSR markers were appropriate for assay of genetic diversity in the Ethiopian linseed accessions and the *Linum* genus. Levels of polymorphism and shared bands enabled inferences about genetic structure, diversity and biogeography.

The relationship between the seven *Linum* wild species (Fig. 2a) were broadly consistent with the results based on a combined ITS and chloroplast (*ndhF* + *trnK* intron + *trnL* – F) topology by McDill et al. (2009). Based on EST-SSR primer data Fu (2011), reported *L. bienne* Mill. and *L. decumbens*

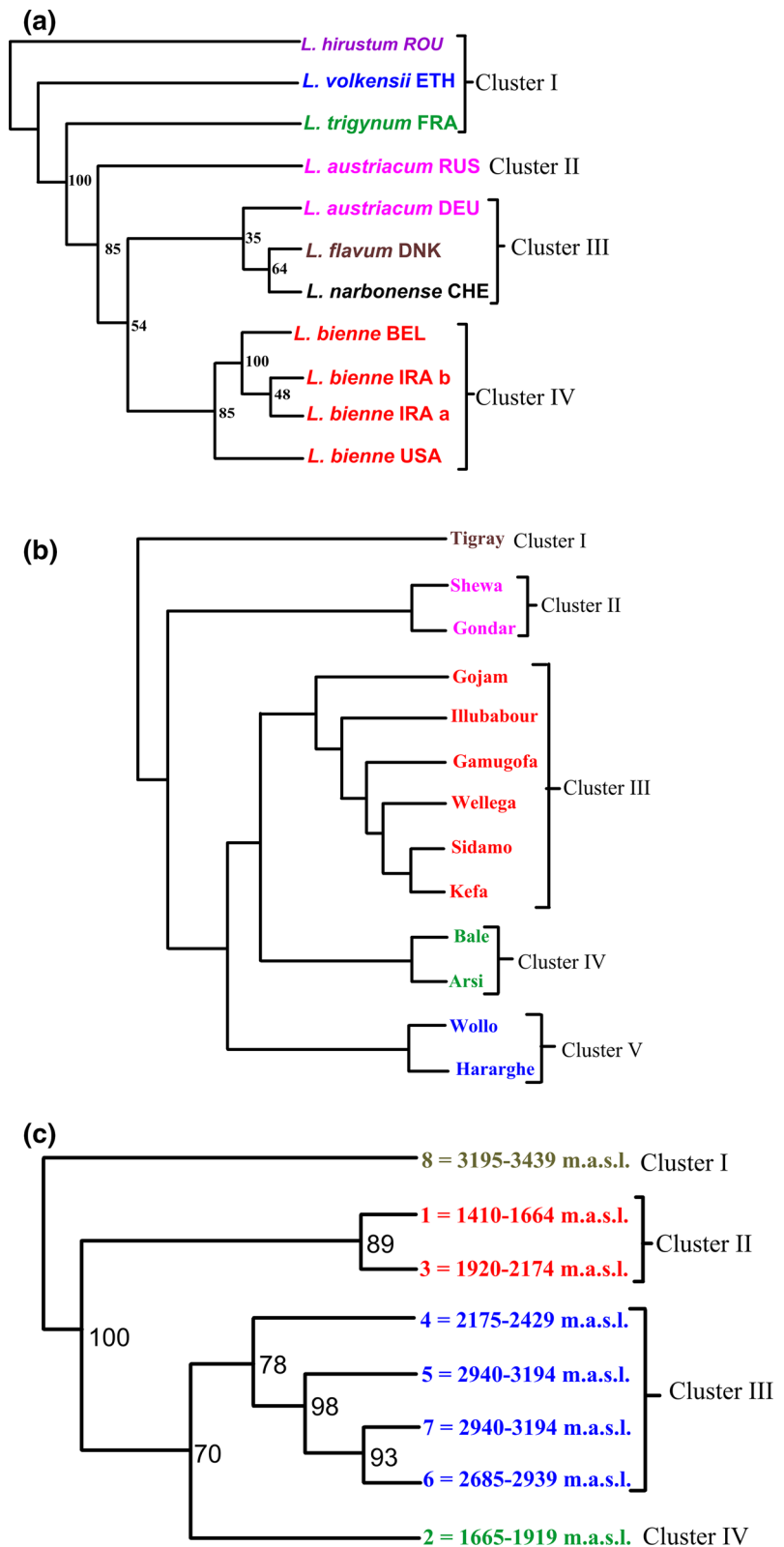
Table 2 Genetic diversity, degree of polymorphism and number of unique alleles: (a) within regional grouping; and (b) within altitude classes

S. no.	Region	No. of assay	Total band no.	Polymorphic band	Degree of polymorphism (%)	Number and % of unique band	GD	PIC
<i>(a) Within a region</i>								
1	ARC	22	167	157	94.01	48(28.74)	0.242 ± 0.150	0.201 ± 0.113
2	Arsi	18	206	199	96.60	37(17.96)	0.295 ± 0.142	0.241 ± 0.103
3	Bale	20	195	187	95.90	26(13.33)	0.296 ± 0.145	0.242 ± 0.105
4	Gamugofa	7	149	137	91.95	66(44.30)	0.330 ± 0.138	0.2660 ± .100
5	Gojam	6	176	164	93.18	108(61.36)	0.329 ± 0.108	0.269 ± 0.076
6	Gondar	42	289	288	99.65	64(22.15)	0.209 ± 0.152	0.176 ± 0.113
7	Hararghe	8	112	97	86.61	46(41.07)	0.249 ± 0.130	0.210 ± 0.100
8	Illubabour	4	158	147	93.04	102(64.56)	0.353 ± 0.116	0.284 ± 0.087
9	Kefa	4	124	112	90.32	75(60.48)	0.385 ± 0.139	0.301 ± 0.106
10	Shewa	12	214	209	97.66	95(44.39)	0.300 ± 0.141	0.245 ± 0.098
11	Sidamo	1	71	–	–	–	–	–
12	Tigray	16	212	203	95.75	29(13.68)	0.302 ± 0.127	0.249 ± 0.091
13	Wellega	10	145	130	89.66	26(17.93)	0.303 ± 0.141	0.247 ± 0.106
14	Wollo	14	195	187	95.90	57(29.23)	0.289 ± 0.146	0.236 ± 0.105
15	Canada	6	84	52	61.90	17(20.24)	0.210 ± 0.188	0.171 ± 0.148
Mean		12.67	173.40	162.07	91.48	57.87(34.29%)	0.292 ± 0.140	0.238 ± 0.104
S. no.	Altitude class	No. of assay	Total band no.	Polymorphic band	Degree of polymorphism (%)	Number and % of unique band	GD	PIC
<i>(b) Within an altitude class</i>								
1	1410–1664	11	169	158	93.49	47(27.81)	0.320 ± 0.153	0.257 ± 0.110
2	1665–1919	21	220	213	96.82	45(20.45)	0.267 ± 0.135	0.222 ± 0.099
3	1920–2174	22	218	211	96.79	31(14.22)	0.284 ± 0.138	0.234 ± 0.102
4	2175–2429	13	235	227	96.60	100(42.55)	0.263 ± 0.139	0.218 ± 0.100
5	2430–2684	17	179	170	94.97	26(14.53)	0.302 ± 0.144	0.246 ± 0.105
6	2685–2939	7	146	134	91.78	74(50.68)	0.298 ± 0.126	0.246 ± 0.093
7	2940–3194	10	135	123	91.11	44(32.59)	0.315 ± 0.167	0.252 ± 0.119
8	3195–3449	6	144	135	93.75	54(37.50)	0.360 ± 0.130	0.287 ± 0.093
Mean		13.375	180.75	171.38	94.81	52.63(29.11%)	0.288 ± 0.129	0.235 ± 0.094

(both sect *Linum*) collected from different countries grouped into different clusters. Linseed and six of the seven *Linum* species did not have common alleles for ISSR and IRAP primers used for the study. Similarity between linseed and all four *L. bienne* Mill. genotypes (Fig. 2a, Table 1) was expected since *L. bienne* Mill. is the progenitor of *L. usitatissimum* (Soto-Cerda 2011); one allele was common to all linseed and *L. bienne* Mill. genotypes but not to other *Linum* species. Melnikova et al. (2014) also reported *L. usitatissimum* and *L. bienne* were grouped under Sect. *Linum*. *L.*

strictum L., *L. keniense* Fries., *L. holstii* Engl. ex R. Wilczek (likely to be a synonym of *L. volkensii* Engl.) and *L. trigynum* L. var. *sieberi* (Planch.) Cuf. are the four wild *Linum* species in Ethiopia (Edwards 1991). Even these more distant germplasm pools may be useful to increase variation by crossing with linseed, since several *Linum* species can produce viable and fertile hybrids (Seetharam 1972; Friedt 1989). Diederichsen and Richards (2003) and Cullis (2011) reported the *L. bienne* Mill. itself is more diverse in morphology than *L. usitatissimum*: in the accessions in

Fig. 2 Neighbour joining trees based on ISSR and IRAP data showing the genetic relationships among linseed and other *Linum* accessions. Numbers represent bootstrap re-sampling values for support at nodes (%). **a** 11 accessions of seven wild *Linum* species; **b** 13 former Ethiopian Administrative Regions for their linseed germplasm genetic diversity based on 163 linseed germplasm; **c** Eight altitude classes for their linseed germplasm genetic diversity based on 107 linseed germplasm Altitude class 8 is used as an outgroup



the present study, the three *L. bienne* Mill. showed variation in their major morphological and maturity characters, with heterogeneity in maturity, plant height, boll size and seed characters (Worku et al. 2015), supported by the high genotypic diversity revealed here (Fig. 2a; Tables 1c and S1). Additional collections and detailed studies on wild relatives may increase the germplasm pool available to breeders for introgression of novel alleles.

The GD and PIC from the present study showed the presence of high GD and PIC in Ethiopian linseed accessions as compared to the results obtained for samples collected from different countries (Soto-Cerda et al. 2012; Žiarovská et al. 2012).” Soto-Cerda et al. (2012) used a core set of linseed germplasm developed from 60 accessions as representative of 16 countries from three continents and obtained 0.34 mean value of GD from 150 microsatellite (85 gSSR and 65 EST-SSR). Žiarovská et al. (2012) reported PIC values from four ISSR and two IRAP primers ranged from 0.12 to 0.37 and from 0.22 to 0.28, respectively from 18 accessions collected from 11 countries. These values were similar to the range of results from the present study. In *Piper betle*, Grewal et al. (2016) reported ISSR polymorphism was between 43% and 100% (PIC 0.17 to 0.45, mean 0.32). IRAPs were also informative with *Crocus*, where the lack of diversity within the saffron crocus and high diversity in other members of the genus were evident (Alsayied 2015). Diederichsen (2001) compared the genetic variation from Canadian linseed with world collections and reported Canada linseed cultivars showed higher rate of loss in genetic variation because of the breeding programs which largely use selected varieties (and reflected in the GD and PIC values of Canadian lines here). The selected accessions from Ethiopian Agricultural Research Centres also showed relatively low GD and PIC values.

Despite the substantial morphological variation in all the flax accessions here (Worku et al. 2015), the molecular genotype markers were able to group accessions by both altitude and region (Fig. 2) indicating lack of gene flow across the country and/or selection for particular genotypes in each environment. Ethiopia is a mountainous country including dry areas (Engels and Hawkes 1991; Friis et al. 2010) and diverse topography (particularly altitude), important for agro-ecological zonation (Hurni 1998). Linseed grows in tropical highlands where mean temperatures

during the growing period correlate with altitude. Selection may be related to agronomic performance and also regional use differences (Worku et al. 2005, 2012). Previously, Ali et al. (2016) showed clusters of variation between the Ethiopian linseed landraces and non-native cultivars for majority of morphological traits. Noormohammadi et al. (2015) reported wild *L. austriacum* L. populations at intermediate altitudes have greater diversity than populations at lower and higher altitudes. The existing diversity of linseed in Ethiopia is the contribution of diversities in agroecological systems, cultural history of the people and farmers knowledge (Engels and Hawkes 1991; Kiros-Meles et al. 2008). In addition to agroecological (Hurni 1998), morphological (Worku et al. 2015) and molecular similarity, and geographical proximity, linseed in the Arsi and Bali regions is cultivated for cash and most growers used selected linseed varieties to match their demand for higher price with the demand of oil pressing factories for higher oil content from the crop (Worku et al. 2012). Some of the research centres and sites, like Sinana, Bekoji, Kulumsa, Arsi-Robe, for linseed and other crops breeding are located in the two regions, so knowledge awareness and seed acquisition are easily accessible to farmers in the regions.

Gondar and Gojam are major linseed producing regions while Kefa and Gamugofa are not. The attention given to linseed in major producing regions (including by the government and local administrators to improve their linseed production) means new and uniform varieties may be more prevalent in these regions: Victory, Concurrant, CI-1525 and CI-1652 are examples of the four selected varieties introduced to Ethiopia and released in 1978 and 1984 (Belayneh 1991), where oil pressing machines are used and linseed is a cash crop. Accessions were collected in different decades by EBI from different regions in Ethiopia. In the cluster analysis (see Figure S2 and Table S3a & b), cultivars collected after 2000 by EBI grouped. This suggested that newly introduced lines are currently substituting the native linseed cultivars in Ethiopia on farm. The introduction of uniform lines would be cause for genetic erosion in Ethiopian linseed cultivars (Belayneh 1991). However, during the expansion years, the Tigray Region was not accessible. Worku et al. (2012) reported only a single variety name, Yehagerie or Yehabesha, meaning cultivar from own country, whereas from other regions

at least two different additional names of introduced 'improved' varieties: Yeferenj (cultivar introduced from abroad) and Yemengist (cultivar given from the state). For breeding and improvement programs, unimproved germplasm (with high diversity and genetic distances) must be included in crosses (as has been successful in wheat; e.g. Ali et al. (2014), particularly with requirement to meet new climatic stress factors. Landrace linseeds and other crops, once identified and integrated into a program, can maintain diversity; while enabling adoption of varieties with improved performance, it is important to mitigate loss of potentially valuable and adapted genotypes.

Conclusions

In Ethiopia linseed is traditionally grown on marginal lands and farmers accept low crop yields since they do not contribute any input for its growth (Worku et al. 2012) and it is drought tolerant (Durant 1976; Seegeler 1983). Nevertheless, there are strong agro-ecological associations between linseed and other grain crops, and a cultural value in the multi-use crop within Ethiopian communities (Mulatu et al. 2002). The presence of different socio-cultural conditions and the diverse tropical and sub-tropical, mountainous topography in Ethiopia support high genetic diversity in linseed both among regional and altitude groups. In contrast, temperate linseed varieties (Canadian here) have lower genetic diversity and polymorphism, expected because of strong selection made by breeders, farmers and researchers.

Exploitation of diversity within linseed, hybridization with *L. bienne*, and backcrossing with other wild *Linum* species, can introgress novel and useful characters to linseed for quality (oil and fibre), yield, biotic and abiotic stress resistances. The regional and altitude diversity revealed here suggests genes are available for improvement, particularly with respect to abiotic stress resistances. Both genetic mapping and perhaps association genetics using a combination of DNA markers and performance or morphological data may be valuable for in country breeding centres, and have the potential to expand linseed producing regions and productivity under developing agroecological conditions.

Acknowledgements Holetta and Adiet Agriculture Research Centres, Ethiopian Biodiversity Institute, and local farmers from Ethiopia, USDA Germplasm Information Centre from USDA, and Canada Crop Development Centre are thanked for their kind and generous provision of linseed germplasm. The Amhara Agricultural Research Centre, Gondar branch provided research fields and are thanked. The University of Gondar and Holetta Agricultural Research Centre were the sources of funds for the research, and fees were waived by the University of Leicester. We also thank the IAEA Coordinated Research Programmes CRP for input to programmes. This paper is dedicated to Professor Annette Cashmore (1958–2016), a source of inspiration for genetic research and training in the UK and globally.

Funding The necessary equipment, apparatuses and consumables for this study were covered by Professor Pat Heslop-Harrison and Dr Trude Schwarzacher lab facilities. During lab work the living allowance for the first author was covered by Genetics Education Networking for Innovation & Excellence (GENIE) project, University of Leicester and miscellaneous sources from Professor Pat Heslop-Harrison and Dr Trude Schwarzacher.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Agarwal BL (1996) Basic statistics 3rd edn. New Age International, New Delhi
- Ali MF, Mekbib F, Wakjira A (2014) Morphological diversity of Ethiopian linseed (*Linum usitatissimum*) landrace accessions and non-native cultivars. J Plant Breed Genet 2:115–124
- Ali N, Heslop-Harrison JP, Ahmad H, Graybosch RA, Hein GL, Schwarzacher T (2016) Introgression of chromosome segments from multiple alien species in wheat breeding lines with wheat streak mosaic virus resistance. Heredity 117:114–123
- Alsayied N, Fernández JA, Schwarzacher T, Heslop-Harrison JS (2015) Diversity and relationships of *Crocus sativus* and its relatives analysed by Inter Retroelement Amplified Polymorphism (IRAP). Ann Bot 116:359–368. <https://doi.org/10.1093/aob/mcv103>
- Ayalew H, Genet T, Dessalegn T, Wondale L (2015) Multi-variate diversity, heritability and genetic advance in Tef

- (*Eragrostis tef* [Zucc.] Trotter) landraces in Ethiopia. *Afri Crop Sci J* 19:201–212
- Bedane GM, Saukuru AM, George DL, Gupta ML (2015) Evaluation of teff lines for agronomic traits in Australia. *Afri Crop Sci J* 9:242–247
- Belayneh H (1991) Oil Crops Germplasm a Vital Resource for Plant Breeders. In: Engels JMM, Hawkes JG, Worede M (eds) *Plant Genetic resources of Ethiopia*. Cambridge University Press, Cambridge, pp 344–353
- Bolsheva NL, Zelenin AV, Nosova IN, Amosova AV, Samatadze TE, Yurkevich OY et al (2015) The diversity of karyotypes and genomes within section *Syllinum* of the genus *Linum* (Linaceae) revealed by molecular cytogenetic markers and RAPD analysis. *PLoS ONE*. <https://doi.org/10.1371/journal.pone.0122015>
- Cloutier S, Niu Z, Datla R, Duguid S (2009) Development and analysis of EST-SSRs for flax (*Linum usitatissimum*). *Theor Appl Genet* 119:53–63
- Cullis CA (2005) Mechanisms and control of rapid genomic changes in flax. *Ann Bot* 95:201–206
- Cullis CA (2011) *Linum*. In: Kole C (ed) *Wild crop relatives: genomic and breeding resources—oil seeds*. Berlin, Springer, pp 177–189
- Diederichsen A (2001) Comparison of genetic diversity of flax (*Linum usitatissimum*) between Canadian cultivars and a world collection. *Plant Breed* 120:360–362
- Diederichsen A, Richards K (2003) Cultivated flax and the genus *Linum* L.: taxonomy and germplasm conservation. In: Muir AD, Westcott ND (eds) *Flax: The Genus Linum*. CRC Press, New York, pp 22–54
- Durant A (1976) Flax and linseed: *Linum usitatissimum* (Linaceae). In: Simmonds NW (ed) *Evolution of Crop Plants*. Longman, New York, pp 190–193
- Edwards SB (1991) Crops with wild relatives found in Ethiopia. In: Engels JMM, Hawkes JG, Worede M (eds) *Plant genetic resources of Ethiopia*. Cambridge University Press, Cambridge, pp 42–74
- Engels JMM, Hawkes JG (1991) The Ethiopian gene centre and its genetic diversity. In: Engels JMM, Hawkes JG, Worede M (eds) *Plant Genetic Resources of Ethiopia*. Cambridge University Press, Cambridge, pp 23–41
- Fernández JA, Le Moigne N, Caro-Bretelle AS, El Hagea R, Le Ducc A, Lozachmeur M, Bono P, Bergeret A (2016) Role of flax cell wall components on the microstructure and transverse mechanical behaviour of flax fabrics reinforced epoxy biocomposites. *Ind Crops Prod* 85:93–108
- Friedt W, Nichterlein K, Nickel M (1989) Biotechnology in breeding of flax (*Linum usitatissimum*) – the present status and future prospects. In: Marshal G (ed) *Flax–Breeding and Utilization*. Kluwer Acad Publ, Dordrecht, pp 5–13
- Friis Ib, Demissew S, van Breugel P (2010) Atlas of the potential vegetation of Ethiopia, vol 58 of Biologiske skrifter. Royal Danish Academy of Sciences and Letters, Botany, P 306
- Fu YB (2011) Genetic evidence for early flax domestication with capsular dehiscence. *Genet Resour Crop Evol* 58:1119–1129
- Fu YB, Diederichsen A, Richards KW, Peterson G (2002) Genetic diversity within a range of cultivars and landraces of flax (*Linum usitatissimum*) as revealed by RAPDs. *Genet Resour Crop Evol* 49:167–174
- Gebremichael DE, Parzies HK (2011) Genetic variability among landraces of sesame in Ethiopia. *Afri Crop Sci J* 19:1–13
- Grewal A, Singh D, Katiyar RS, Tewari SK, Nainwal RC, Bindu KH (2016) Evaluation of genetic diversity of *Piper betle* cultivars using ISSR markers Shalini Goyat. *Int J Adv Res* 4:571–579
- Gul S, Rajper AA, Kalhor FA, Kalhor SA, Ali A, Shah FA, Ahmed S, Ahmed M (2016) Screening selected linseed (*Linum usitatissimum*) genotypes for yield performance in Sindh, Pakistan. *Nat Sci* 8:53–65
- Habibollahi H, Noormohammadi Z, Sheidai M, Farahani F (2015) Genetic structure of cultivated flax (*Linum usitatissimum*) based on Retrotransposon-based markers. *Genetika* 47:1111–1122
- Harrison S, Noss R (2017) Endemism hotspots are linked to stable climatic refugia. *Ann Bot* 119:207–214. <https://doi.org/10.1093/aob/mcw248>
- Hurni H (1998) Agroecological maps of Ethiopia: Explanatory notes on three maps at a scale of 1:1,000,000. Centre for Development and Environment University of Bern, Switzerland in association with The Ministry of Agriculture, Ethiopia. p 43
- Kalendar R, Schulman AH (2006) IRAP and REMAP for retrotransposon-based genotyping and fingerprinting. *Nat Protocols* 1:2478–2484.s
- Khan MA, Mirza MY, Akmal M, Ali N, Khan I (2007) Genetic parameters and their implications for yield improvement in Sesame. *Sarhad J Agric* 23:623–627
- Kiros-Meles A, Mathew M, Abang MM (2008) Farmers' knowledge of crop diseases and control strategies in the Regional State of Tigray, northern Ethiopia: implications for farmer–researcher collaboration in disease management. *Agric Hum Values* 25:433–452. <https://doi.org/10.1007/s10460-007-9109-6>
- McKenzie R, Roach M, Hotte N, DePauw M, Hobson N, Koziel S et al. (2008) Genetic and genomics resources for flax fibre feedstock improvement. In: International Conference on Flax and Other Bast Plants. p 13
- Kurt O, Evans GM (1998) Genetic basis of variation in linseed (*Linum usitatissimum*) cultivars. *Tr J Agric Forestry* 22:373–379
- Lapitan VC, Brar DS, Abe T, Redona ED (2007) Assessment of genetic diversity Philippine rice cultivars carrying good quality traits SSR markers. *Breeding Sci* 57:263–270
- McDill J, Repplinger M, Simpson BB, Kadereit JW (2009) The phylogeny of *Linum* and Linaceae Subfamily Linoideae, with implications for their systematics, biogeography, and evolution of heterostyly. *Syst Bot* 34:386–405
- McHughen A (1990) Flax (*Linum usitatissimum*): in vitro studies. In: Bajaj YPS (ed) *Biootechnology in agriculture and forest*, vol 10, Legumes and oilseed crops I. Springer, Berlin, pp 502–514
- Melnikova NV, Kudryavtseva AV, Zelenin AV, Lakunina VA, Yurkevich OY, Speranskaya AS et al (2014) Retrotransposon-based molecular markers for analysis of genetic diversity within the Genus *Linum*. *BioMed Res Int*. <https://doi.org/10.1155/2014/231589>
- Mengistu DK, Kidane YG, Fadda C, Enrico Pè M (2016) Genetic diversity in Ethiopian durum wheat (*Triticum turgidum* var *.durum*) inferred from phenotypic variations.

- Plant Genet Resour, 1–11 <https://doi.org/10.1017/s1479262116000393>
- Muhe K, Assefa A (2011) Diversity and agronomic potential of barley (*Hordeum vulgare*) landraces under variable production system in Ethiopia. *Int Res J Plant Sci* 2:305–310
- Mulatu G, Zemed A, Endeshaw B, Awegechew T (2002) Edible oil crops and their integration with the major cereals in North Shewa and South Welo, Central Highlands of Ethiopia: an ethnobotanical perspective. *Hereditas* 137:29–40
- Nei M, Tajima F, Tateno Y (1983) Accuracy of estimated phylogenetic trees from molecular data: II Gene frequency data. *J Mol Evol* 19:153–170
- Noormohammadi Z, Shafaf T, Farahani F, Sheidai M, Talebi SM, Hasheminejad-Ahangarani-Farahani Y (2015) Within and among-genetic variation in Asian flax *Linum austriacum* (Linaceae) in response to latitude changes: cytogenetic and molecular analyses. *Biodiversitas* 16:145–150
- Olango TM, Tesfaye B, Pagnotta MA, Pè ME, Catellani M (2015) Development of SSR markers and genetic diversity analysis in enset (*Ensete ventricosum* (Welw.) Cheesman), an orphan food security crop from Southern Ethiopia. *BMC Genet* 98:1–16
- Pali V, Mehta N, Verulkar SB, Xalxo MS, Saxena RR (2015) Molecular diversity in flax (*Linum usitatissimum*) as revealed by DNA based markers. *Vegetos-Int J Plant Res* 28:157–165
- Powell W, Morgante M, Andre C, Hanafey M, Voge J, Tingey S et al (1996) The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. *Mol Breeding* 2:225–238
- Rahman MA, Thomson MJ, Shah-E-Alam M, de Ocampo M, Egdane J, Ismail AM (2016) Exploring novel genetic sources of salinity tolerance in rice through molecular and physiological characterization. *Ann Bot* 117:1083–1097. <https://doi.org/10.1093/aob/mcw030>
- Saeidi H, Rahiminejad MR, Vallian S, Heslop-Harrison JS (2006) Biodiversity of diploid D-genome *Aegilops tauschii* Coss. in Iran measured using microsatellites. *Genet Resour Crop Evol* 53:1477–1484. <https://doi.org/10.1007/s10722-005-7110-8>
- Saeidi H, Rahiminejad MR, Heslop-Harrison JS. (2008) Retroelement Insertional Polymorphisms, Diversity and Phylogeography within Diploid, D-genome *Aegilops tauschii* (Triticeae, Poaceae) Sub-taxa in Iran. *Ann Bot* 101: 855–861. <https://doi.org/10.1093/aob/mcn042>, www.aob.oxfordjournals.org
- Schardt D (2005) Just the flax. Nutrition action health letter, December p 7–9
- Seegeler CJP (1983) *Linum usitatissimum*: oil plants in ethiopia, their taxonomy and agricultural significance. Center for Agricultural Publishing and Documentation, Wageningen, pp 151–197
- Seetharam A (1972) Interspecific hybridization in *Linum*. *Euphatica* 21:489–950
- Sliseris J, Yan L, Kasal B (2016) Numerical modelling of flax short fibre reinforced and flax fibre fabric reinforced polymer composites. *Compos B* 89:143–154
- Smykal P, Bačova'-Kerteszova N, Kalendar R, Corander J, Schulman AH, Pavelek M (2011) Genetic diversity of cultivated flax (*Linum usitatissimum*) germplasm assessed by retrotransposon-based markers. *Theor Appl Genet* 122:1385–1397
- Soto-Cerda BJ, Saavedra UH, Navarro C, Mora OP (2011) Characterization of novel genetic SSR markers in *Linum usitatissimum* and their transferability across eleven *Linum* species. *J Biotechnol*. <https://doi.org/10.2225/vol14-issue2-fulltext-6>
- Soto-Cerda BJ, Maureira-Butler I, Munoz G, Rupayan A, Cloutier S (2012) SSR-based population structure, molecular diversity and linkage disequilibrium analysis of a collection of flax (*Linum usitatissimum*) varying for mucilage seed-coat Content. *Mol Breeding* 30:875–888
- Tadele S, Mekbib F, Tsfaye K (2014) Genetic diversity of coffee (*Coffea arabica* L.) landraces from Southern Ethiopia as revealed by inter simple sequence repeat marker. *Global Adv Res J Agri Sci* 3:24–34
- Teo CH, Tan SH, Ho CL, Faridah QZ, Othman YR, Heslop-Harrison JS et al. (2005) Genome constitution and classification using retrotransposon-based markers in the orphan crop banana. *J Plant Biol* 48:96–105
- Uysal H, Fu YB, Kurt O, Peterson GW, Diederichsen A, Kusters P (2010) Genetic diversity of cultivated flax (*Linum usitatissimum*) and its wild progenitor pale flax (*Linum bienne* Mill.) as revealed by ISSR markers. *Genet Resour Crop Evol* 57:1109–1119. <https://doi.org/10.1007/s10722-010-9551-y>
- Wiesner I, Wiesnerova D, Tejklova E (2001) Effect of anchor and core sequence in microsatellite primers on flax fingerprinting patterns. *J Agric Sci* 137:37–44
- Wiesnerova D, Wiesner I (2004) ISSR-Based clustering of cultivated flax germplasm is statistically correlated to thousand seed mass. *Mol Biotechnol* 26:207–214
- Worku N, Zemed A, Haileselassie Y (2005) Variation and association analyses on morphological characters of Linseed (*Linum usitatissimum*) in Ethnobotany. *SINET: J Sci* 28:129–140
- Worku N, Zemed A, Haileselassie Y (2012) Linseed (*Linum usitatissimum*) ethnobotany and its cultivation status in Ethiopia. *Int J Agric Appl Sci* 4:48–57
- Worku N, Heslop-Harrison JS, Aduagna W (2015) Diversity in 198 Ethiopian linseed (*Linum usitatissimum*) accessions based on morphological characterization and seed oil characteristics. *Genet Resour Crop Evol* 62:1037–1053
- Yoon MY, Moe KT, Kim DY, Rho IR, Kim S, Kim KT, Won MY, Chung JW and Park YJ (2012) Genetic diversity and population structure analysis of strawberry (*Fragaria xananassa* Duch.) using SSR markers. *J Biotechnol* ISSN: 0717-3458: <https://doi.org/10.2225/vol15-issue2-fulltext-5>
- Žiarovská J, Ražná K, Senková S, Štefánová Bežo M (2012) Variability of *Linum usitatissimum* based on usitatissimum molecular markers. *ARP J Agric Biol Sci* 7:50–58
- Zohary D, Hopf M (2000) Domestication of plants in the Old World: the origin and spread of cultivated plants in West Asia, Europe and the Nile Valley, 3rd edn. Oxford University Press, Oxford, p 328