Molecular biology and genetics/Biologie et génétique moléculaires

SSR and morphological trait based population structure analysis of 130 diverse flax (*Linum usitatissimum* L.) accessions

Shashi Bhushan Choudhary [a,*], Hariom Kumar Sharma [a], Arroju Anil Kumar [a], Rangappa Thimmaiah Maruthi [a], Jiban Mitra [a], Isholeena Chowdhury [a], Binay Kumar Singh [b], Pran Gobinda Karmakar [a]

[a] *ICAR – Central Research Institute for Jute and Allied Fibres, Barrackpore, Kolkata 700120, West Bengal, India*
[b] *ICAR – Indian Institute of Agricultural Biotechnology, Ranchi, Jharkhand, India*

A R T I C L E   I N F O

A B S T R A C T

A total of 130 flax accessions of diverse morphotypes and worldwide origin were assessed for genetic diversity and population structure using 11 morphological traits and microsatellite markers (15 gSSRs and 7 EST–SSRs). Analysis performed after classifying these accessions on the basis of plant height, branching pattern, seed size, Indian/foreign origin into six categories called sub-populations *viz.* fibre type exotic, fibre type indigenous, intermediate type exotic, intermediate type indigenous, linseed type exotic and linseed type indigenous. The study assessed different diversity indices, AMOVA, population structure and included a principal coordinate analysis based on different marker systems. The highest diversity was exhibited by gSSR markers ($SI = 0.46$; $He = 0.31$; $P = 85.11$). AMOVA based on all markers explained significant difference among fibre type, intermediate type and linseed type populations of flax. In terms of variation explained by different markers, EST-SSR markers (12%) better differentiated flax populations compared to morphological (9%) and gSSR (6%) markers at $P = 0.01$. The maximum Nei's unbiased genetic distance ($D = 0.11$) was observed between fibre type and linseed type exotic sub-populations based on EST-SSR markers. The combined structure analysis by using all markers grouped Indian fibre type accessions (63.4%) in a separate cluster along with the Indian intermediate type (48.7%), whereas Indian accessions (82.16%) of linseed type constituted an independent cluster. These findings were supported by the results of the principal coordinate analysis. Morphological markers employed in the study found complementary with microsatellite based markers in deciphering genetic diversity and population structure of the flax germplasm.

© 2016 Académie des sciences. Published by Elsevier Masson SAS. All rights reserved.

1. Introduction

Flax (*Linum usitatissimum* L.) is one of the oldest domesticated crops [1] with many applications. The long bast fiber of the species produces one of the oldest textile fabrics with extraordinary comfort due to unique mechanical properties and chemical composition [2,3], while short fibres together with wooden straw provide raw material for automobile, construction, package and paper industry (making cigarettes, currency notes, artwork) [4,5]. Flax seed is valued for its health-promoting properties due to the presence of bioactive agents like omega-3 fatty acids, lignans, mucilage [6–10] and is increasingly used in human food products [11,12]. Further, seed oil of the species is

traditionally used in a range of industrial applications such as linoleum, paint, varnish, soap, printer ink, etc. To capitalise the expanding horizons of flax seed oil applications in the food industry, a number of low linolenic acid-containing varieties were developed [13,14]. Even seed cake, a byproduct of oil extraction used as a nutritious feed for livestock due to protein richness.

Cultivated flax along with another 199 species belongs to the genus *Linum* (Linaceae) [15,16] and has a chromosome number $2n = 30$ [17]. The self-pollinating species is supposed to originate in either the Middle East or Indian regions, possibly from *L. angustifolium* (*L. bienne*) [17–20]. Molecular marker based studies comprehensively established its monophylogenic origin with oil type as a predecessor of fibre-type flax [21,22]. The oldest archaeological evidence for the wild species came from Tell Abu Hureyra in Northern Syria (11,200–10,500 years ago) [23] and by the eighth millennium BC the species was found distributed throughout the Near East [1]. The first cultivated form of flax archeologically was traced at Tell Ramad in Syria 9000 years ago and spread to Europe as well as the Nile valley as winter oil seed. However, in Eastern Europe, summer fibre-type flax evolved and gradually spread to central Europe [17,24]. On the other hand, the species was introduced into Western Europe (the Netherlands, Northern France, Belgium and Switzerland) between 5000 and 3000 BC by semi-nomads of the Middle East [25]. The available literature is silent about whether the domestication processes of fibre type flax at eastern Europe and middle East were independent events or not. Arguably, it is suggested that all modern fibre varieties in use today originated from Eastern Europe [24]. From there, the species travelled to America along with colonial labourers. During the trans-continental domestication process, the species evolved into different morphotypes depending on man-made differential preference either for fibre or seed oil. The century-long directed selection pressure led to the evolution of distinct morphotypes, particularly in the western and eastern regions of Eurasia with dominance of fibre and oilseed type, respectively [26]. Both morphotypes belong to the same species with distinct morphological, anatomical, physiological, and agronomical characteristics [17]. Presently Canada, India, China, the United States of America and Argentina are leading linseed-producing countries having cultivars characterised by shorter and much branched stem bearing large seeded pods, whereas, in the temperate regions of China, the Russian Federation and Western Europe, tall plants with much less branching were cultivated as fibre flax [27]. India is the second largest linseed producing country [28] with negligible area under fibre flax cultivation. As a result, this country heavily depends on imports to meet its needs in terms of industrial fibres. Diversified agro-climatic conditions and rich plant genetic resources together with escalating linen import cost have inspired researchers to develop locally adapted fibre flax cultivars in India. The Indian Council of Agricultural Research–Central Research Institute for Jute and Allied Fibres (ICAR-CRIJAF), Barrackpore, Kolkata, India maintains 130 accessions of flax from Asian, European, Australian and American countries. Partial utilisation of these genetic resources paved the way to develop and release the first fibre flax cultivar of India namely "*Tiara*" during the 11th annual workshop of All India Network Project on Jute and Allied Fibres in 2014 [29]. However, the lack of quantified information about the diversity and the genetic structure of the flax genetic resource seriously limits prospects towards their conservation and enhanced utilization [30].

To this end, initial success in flax diversity assessment by employing both morphological [31–34] and biochemical [35,36] attributes were encouraging. Subsequently, these studies were complemented well by molecular markers like randomly amplified polymorphic DNA (RAPD), amplified fragment length polymorphism (AFLP), inter simple sequence repeat (ISSR), simple sequence repeat (SSR) and inter-retrotransposon amplified polymorphism (IRAP) [18,37–39]. Presently, marker-based population structure analysis employs methods like F_{ST} [40], analysis of molecular variance [41], cluster analysis, principal component analysis and Bayesian analysis [42] *etc.* across kingdoms. Although some of these methods have powerful algorithms, they all require a suitable marker system reflecting neutral genetic variation to estimate genetic relationships. Microsatellite markers like genomic simple sequence repeats (gSSRs) and expressed sequence tag simple sequence repeats (EST–SSRs) with their ability to enhance the genome wide resolution of genetic variation [43] are one of the most powerful marker systems for population structure analysis of most plant genetic resources [44]. Large numbers of gSSRs [45–49] and EST–SSRs [49–51] have been developed in flax, which can be effectively employed to understand evolutionary and domestication dynamics across morphotypes in the crop.

The literature holds very little information about the population structure of flax and publications are mainly based on the analysis of a few morphological and/or molecular markers across accessions with restricted geographical origin. Therefore, additional research in this line by employing both morphological as well as molecular markers in combination using accessions of diverse origin and morphotypes is highly warranted for a more comprehensive understanding of the evolutionary and domestication processes in the crop. In the present study, 11 morphological traits along with 22 microsatellite-based markers (gSSRs, EST–SSRs) were employed to understand the population structure of 130 flax accessions with diverse growth habits and origins.

2. Materials and methods

2.1. Plant material and layout

A total of 130 accessions (exotic: 75 and indigenous: 55) with diverse growth habits were obtained from the National Active Germplasm Site for Jute and Allied Fibres, ICAR-CRIJAF, Barrackpore, Kolkata, India. These accessions were included in the study after two successive generations of self-pollination by following singleseed descent method to ensure homogeneity of seed lots. The details of these accessions are presented in Table 1. From the

Table 1
Details of flax accessions used in present study.

Fibre type			Intermediate type									Linseed type					
Accession	Source	SPC	Accession	Source	SPC	Accession	Source	SPC	Accession	Source	SPC	Accession	Source	SPC	Accession	Source	SPC
LEX001	Sweden	FTE	LEX003	Australia	ITE	LEX066	Australia	ITE	LIN043	India	ITI	LEX035	Australia	LTE	LIN028	India	LTI
LEX002	Sweden	FTE	LEX007	Canada	ITE	LEX067	USSR	ITE	LIN044	India	ITI	LEX036	Australia	LTE	LIN029	India	LTI
LEX004	Australia	FTE	LEX008	USSR	ITE	LEX068	USSR	ITE	LIN045	India	ITI	LEX037	Canada	LTE	LIN052	India	LTI
LEX005	Australia	FTE	LEX009	USSR	ITE	LEX070	Canada	ITE	LIN046	India	ITI	LEX038	Canada	LTE			
LEX006	Australia	FTE	LEX010	USSR	ITE	LEX071	Canada	ITE	LIN047	India	ITI	LEX039	Canada	LTE			
LEX011	USSR	FTE	LEX012	USSR	ITE	LEX072	Canada	ITE	LIN048	India	ITI	LEX040	USSR	LTE			
LEX013	USSR	FTE	LEX014	USSR	ITE	LEX073	Canada	ITE	LIN049	India	ITI	LEX041	USSR	LTE			
LEX015	USSR	FTE	LEX017	USSR	ITE	LIN001	India	ITI	LIN050	India	ITI	LEX042	USSR	LTE			
LEX016	Canada	FTE	LEX018	USSR	ITE	LIN002	India	ITI	LIN051	India	ITI	LEX043	Canada	LTE			
LEX023	USSR	FTE	LEX019	USSR	ITE	LIN003	India	ITI	LIN053	India	ITI	LEX045	Canada	LTE			
LEX032	Belgium	FTE	LEX020	USSR	ITE	LIN004	India	ITI	LIN054	India	ITI	LEX046	Canada	LTE			
LEX033	Belgium	FTE	LEX021	USSR	ITE	LIN005	India	ITI				LEX048	Canada	LTE			
LEX034	Holland	FTE	LEX022	USSR	ITE	LIN007	India	ITI				LEX049	Australia	LTE			
LEX062	USSR	FTE	LEX024	USSR	ITE	LIN008	India	ITI				LEX050	Belgium	LTE			
LEX063	USSR	FTE	LEX025	USSR	ITE	LIN011	India	ITI				LEX054	Belgium	LTE			
LEX075	USSR	FTE	LEX026	USSR	ITE	LIN012	India	ITI				LEX056	Argentina	LTE			
LEX076	USSR	FTE	LEX027	USSR	ITE	LIN014	India	ITI				LEX057	Sweden	LTE			
LEX064	USSR	FTE	LEX028	USSR	ITE	LIN015	India	ITI				LEX058	Argentina	LTE			
LIM001	India	FTI	LEX029	Argentina	ITE	LIN016	India	ITI				LEX059	Belgium	LTE			
LIM002	India	FTI	LEX030	Argentina	ITE	LIN017	India	ITI				LEX074	Argentina	LTE			
LIM003	India	FTI	LEX044	Australia	ITE	LIN018	India	ITI				LIN009	India	LTI			
LIM004	India	FTI	LEX047	Australia	ITE	LIN023	India	ITI				LIN010	India	LTI			
LIN030	India	FTI	LEX051	Belgium	ITE	LIN024	India	ITI				LIN019	India	LTI			
LIN031	India	FTI	LEX052	Belgium	ITE	LIN037	India	ITI				LIN020	India	LTI			
LIN032	India	FTI	LEX053	Belgium	ITE	LIN038	India	ITI				LIN021	India	LTI			
LIN033	India	FTI	LEX055	Argentina	ITE	LIN039	India	ITI				LIN022	India	LTI			
LIN034	India	FTI	LEX060	Belgium	ITE	LIN040	India	ITI				LIN025	India	LTI			
LIN035	India	FTI	LEX061	Belgium	ITE	LIN041	India	ITI				LIN026	India	LTI			
LIN036	India	FTI	LEX065	USSR	ITE	LIN042	India	ITI				LIN027	India	LTI			

SPC: sub-population code; FTE: fibre type exotic sub-population; FTI: fibre type indigenous sub-population; ITE: intermediate type exotic sub-population; ITI: intermediate type indigenous sub-population; LTE: linseed type exotic sub-population; LTI: linseed type indigenous sub-population.

homogeneous seed lot of each accession, a part was kept for molecular evaluation and the rest were sown in field for morphological evaluation over two growth seasons. Sowing was done over two seasons (2013–2015) in the experimental fields of ICAR-CRIJAF, Barrackpore, Kolkata, India (22.45°N, 88.26°E; 3.14 above msl) following the standard package of practices to avoid abiotic and biotic stresses that may affect plant growth. During field experiment, each accession was represented by two rows of 3 m, each row at a distance of 30 cm. The plant-to-plant spacing within row was maintained at 5–7 cm by thinning out plants at 35 days after sowing (DAS). All the accessions were sown during the last week of November and harvested at 100 DAS and 125 DAS for fibre and seed quality attributes, respectively.

2.2. Morphological evaluation

For morphological evaluation, data were collected following the descriptor of the International Union for the Protection of New Varieties of Plants (UPOV), Geneva on five randomly selected plants in each accession for 11 traits namely: stem colour (green, red, dark red, green with basal red), flower opening (open, partial open), petal arrangement (free, intermediate, overlapping), flower size (small, medium, large), flower petal colour (white, blue, blue violet), petal shape (circular, circular to pentagonal, pentagonal), colour of distal part of filament (white, blue), anther colour (grey, blue, yellow), boll size (small, medium, large), pod cilia (absent, present), seed coat colour (yellow, brown, deep brown). In the case of qualitative traits, all relevant states of expression were recorded. For recording expression of quasi-quantitative traits an abbreviated scale was used by following UPOV, Geneva descriptors.

2.3. DNA extraction

Genomic DNA was isolated from cotyledonary leaves of 7 days old seedlings using a modified CTAB method [52]. The quality and quantity of extracted DNA were determined using agarose gel electrophoresis (0.8%) and a spectrophotometer (Eppendorf), respectively.

2.4. gSSR and EST-SSR analysis

A total of 21 gSSRs [47,49] and 10 EST-SSRs [50] loci linked to the cell wall biochemical synthesis pathway were selected for the present study. Finally, 15 gSSRs and 7 EST-SSR markers exhibiting reliable amplification were used for further analysis (Table 2). PCR reactions were performed in 25 μl reaction mixture containing 0.2 mM of dNTPs, 0.5 μM each of forward and reverse primers, 0.8 unit of Taq DNA polymerase (Invitrogen) and 0.5 μl MgCl$_2$ (2.5 mM) in 1X PCR buffer (10 mM Tris–HCl buffer, p^H 8.0, 0.5 mM KCl). PCR amplification was carried out with a hot start for 5 min at 95 °C followed by 35 cycles at 95 °C for 30 s, 1 min at the annealing temperature and 72 °C for 1 min with a final extension at 72 °C for 10 min in SimpliAmp^{TM} Thermal Cycler (Applied Biosystems, USA). The PCR amplification products were stained with ethidium bromide and ran in 2.0% agarose gel in 1X TBE

buffer (1.5–2.0 h at 3.82 V/cm). The size of the amplicon was determined using a 100 bp DNA ladder (product ranged from 100 to 3000 bp) (Genaxy, India) and visualized as well as photographed under a gel documentation unit (Bio-Rad, USA).

2.5. Data analysis

The expression of morphological traits (a total of 28 alleles) was binary-coded as 1 for presence or 0 for absence in each accession [53,54]. Owing to the self-pollinating reproductive behaviour of flax, any heterozygosity was not expected for these traits. For molecular marker (gSSR and EST-SSR) profiles, each band was scored as for the presence (1) or for the absence (0). Only the distinct, reproducible and well resolved bands were scored. Before analyzing the data, the flax accessions were classified into three broad groups (based on morphotypes and technical height) and each group into two sub-groups (based on origin): fibre type (flax; convar. *elongatum*) (29 accessions: 18 exotic; 11 indigenous), intermediate type (convar. *usitatissimum*) (69 accessions: 37 exotic; 32 indigenous) and oilseed type (linseed; convar. *mediterraneum*) (32 accessions:20 exotic; 12 indigenous) by adopting Kulpa and Danert [55]. Thus, a total of six sub-populations (fibre type exotic, fibre type indigenous, intermediate type exotic, intermediate type indigenous, linseed type exotic and linseed type indigenous) were constituted a *priori* and subsequent analyses were performed considering this classification. Further, genetic diversity parameters namely; Nei's unbiased genetic identity (I) and genetic distance (D), expected heterozygosity (He), Shannon information index (SI), number of effective alleles (Ne), and percentage of polymorphism (P) [56,57] were determined for each sub-population using GenAlEx v. 6.1 software [58]. The software helped to partition the molecular variance among and within sub-populations and test their statistical significance through the analysis of molecular (both morphological and molecular) variance (AMOVA) [41] using 1000 permutations. Furthermore, GenAlEx v. 6.1 software was also used for principal coordinate analysis (PCoA) based on different makers. Assuming an Hardy–Weinberg equilibrium, POP-GENE version 1.31 [59] software was used to estimate total genetic diversity (Ht), intrasub-population genetic diversity (Hs). Intersub-population genetic diversity (Dst), genetic differentiation coefficient (Gst), were calculated according to de Vicente et al. [60] as follows; $Dst = Ht–Hs$ and $Gst = Dst/Ht$.

Population structure analysis based on different markers (morphological traits, gSSRs, EST–SSRs and combined) was done by a Bayesian model-based clustering algorithm of STRUCTURE 2.3 [42,61] software. The STRUCTURE 2.3 software was run using putative population origin with parameters as follows: burn-in periods = 5000, MCMC replicates = 50,000, admixture ancestry model, allele frequency correlated ($K = 1$ to 10, iterations = 10). To determine the optimum value of K, the results of STRUCTURE 2.3 software were used as input in Structure Harvester online software [62]. Based on Evanno et al.'s delta K method [63], the optimum number of clusters was

identified for different markers. After that, the STRUCTURE 2.3 software was run at desired K values for 100,000 burn-in periods and 100,000 MCMC replicates. Thereafter, a structure diagram was prepared in the same software.

3. Results

3.1. Genetic diversity of population

Different diversity indices were estimated for three populations (fibre type, intermediate type and linseed type) and six sub-populations of flax using morphological traits (11), gSSR (15) and EST-SSR (7) markers that produced 28, 47 and 13 loci, respectively (Table 3). At the species level, the highest diversity was exhibited by gSSR markers (SI = 0.46; He = 0.31; P = 85.11%). However, morphological markers recorded the lowest number of alleles (1.6), but higher polymorphic loci (76.79%) than EST-SSR (73.08%).

At the morphotype level, intermediate type flax showed the highest value for all diversity parameters by combining morphological traits with gSSR and EST-SSR markers (Table 3). gSSR markers and morphological traits revealing maximum and minimum diversity across parameters, respectively. However, maximum Shannon's information index (SI) and expected heterozygosity (He) expressed by gSSR in all the three populations namely fibre type(SI = 0.45, He = 0.30), intermediate type (SI = 0.54, He = 0.37) and linseed type (SI = 0.53, He = 0.36). Among all

sub-populations across all markers the highest number of alleles, Shannon's information index and polymorphic loci were recorded by gSSR in intermediate exotic population (Na = 1.96, SI = 0.53 and P = 97.87%, respectively). The lowest value for these parameters was recorded for fibre type exotic accessions for EST-SSR markers (Na = 1.23, SI = 0.11 and P = 38.46%, respectively). The maximum number of alleles in fibre type exotic (Na = 1.86) and intermediate type indigenous (Na = 1.93) sub-population were recorded by morphological traits. In fibre type indigenous, intermediate type exotic, linseed type exotic and indigenous sub-populations gSSR markers recorded the highest number of observed alleles (Na = 1.81, 1.96, 1.85 and 1.72, respectively). The percentage polymorphic loci across sub-populations followed the same trend, except for linseed-type indigenous accessions that recorded highest value for EST-SSR markers (P = 76.92%). By taking all markers into account, intermediate exotic population represent the sub-population with maximum number of total alleles (Na = 1.92), effective alleles (Ne = 1.52) and Shannons' information index (SI = 0.46) with considerably high expected heterozygosity (He = 0.30).

AMOVA analysis based on morphological, EST-SSR and gSSR markers underlined significant difference among fibre type, intermediate type and linseed type (Table 4). In comparison to morphological and gSSR (ϕPT = 0.09 and 0.06, respectively at p = 0.01) EST-SSR markers (ϕPT = 0.12, p = 0.01) better differentiate flax populations. All the three

Table 2
Primer sequence of gSSR and EST-SSR markers used in the study.

Marker	gSSR primers (5′-3′)	Marker	EST-SSR primers (5′-3′)
LU2	F- TCCGGACCCTTTCAATATCA	COMT	F- TATCCAAGCACCCTTCCATT
	R- AACTACCGCCGGTGATGA		R- GTTCCTCGTCGCCAGGCTCGTGT
LU3	F- GCTCGTGATCTCCTTCATCC	CCoAOMT1	F- TCGGCGTCTACACGGGTTACTCC
	R- AAAACCACGTCCAGATGCTC		R- TGTTGTCGTACCCGATCACTC
LU4	F- TTATTTCCGGACCCTTTCAA	4CL2	F- GGGATACGGAATGACAGAGG
	R- AAACTACCGCCGGTGATGAT		R- CGGTGTGTAACCACCCTTGTTTGTC
LU5	F- GTCACTGGGTGTGTGTTTGC	ACTIN	F- CGACTGCTGAAGGGAAATTGT
	R- AGCAGAAGAAGATGGCGAAA		R- CTGCTGGAAGGTGCTGAGGGAA
LU6	F- CCCATTTCTACCATCTCCTT	F5H1	F- CGTCGGAGAGCTTATTTTCA
	R- CAACAGCGGAACTGATGAAA		R- AGCTTTGATGTTGTCCTTGG
LU7	F- CATCCAACAAAGGGTGGTG	F5H2	F- CCACCAAGATGAATGACCAG
	R- GGAACAAAGGTAGCCATGA		R- TGTAGTGGTGGTGGAGTCAGCTGC
LU9	F- TTGCGTGATTATCTGCTTCG	F5H3	F- CGCTGGCCGGGCGTTGAACG
	R- ATGGCAGGTTCTGCTGTTTC		R- ATGTCGGAGATAGGATCGTC
LU10	F- GCCTAAAGCTGATGCGTTTC		
	R- TGTCAGGCTCCTTCTTTTGC		
LU11	F- ATGGCAGGTTCTGCTGTTTC		
	R- TTGCGTGATTATCTGCTTCG		
LU16	F- TTATTCTTGCCTGCCAATCG		
	R- TCCAGCTCTTGCTCGTTCTT		
LU17	F-GTTTGGTGTTTGGGTGCTTT		
	R-TTCGGTTGTAGCATCCATCC		
LU18	F- AGAGGCGGAGGGCATTAC		
	R- TTGGAGAGTTGGAATCGAGA		
LU27	F- GTTTGAGAAGAGGGCATCCA		
	R- GTTGGGGGTGAAGAGGAACAA		
LU32	F- ACGCGTAAACTTTCCGTTTC		
	R- ATAATGTCGGCTGCTTCTGC		
LU35	F- CCAACGGATCATCCTCTAGC		
	R- GGAACAGAAAGGGGAAAGGAA		

LU: *Linum usitatisisum*, COMT: catechol-*O*-methyltransferase, CCoAOMT1: caffeoyl-CoA 3-*O*-methyltransferase 1, 4CL2: 4-coumarate-CoA ligase 2, F5H: *ferulate-5-hydroxylase*.

Table 3
Diversity parameters estimated based on molecular (EST-SSR, gSSR) and morphological traits in *Flax*.

Population	EST-SSR					gSSR					Morphological					Combined				
	Na	Ne	SI	He	P (%)	Na	Ne	SI	He	P (%)	Na	Ne	SI	He	P (%)	Na	Ne	SI	He	P (%)
Population level																				
Fibre type	1.85	1.38	0.38	0.24	84.62	1.89	1.52	0.45	0.30	89.36	1.93	1.47	0.43	0.28	92.86	1.90	1.48	0.43	0.29	89.77
Intermediate type	2.00	1.46	0.45	0.29	100.00	2.00	1.64	0.54	0.37	100.00	1.93	1.38	0.38	0.24	96.43	1.99	1.53	0.47	0.31	98.86
Linseed type	1.92	1.55	0.49	0.33	92.31	1.94	1.63	0.53	0.36	93.62	1.86	1.33	0.33	0.21	85.71	1.91	1.52	0.46	0.31	90.91
Sub-population level																				
FTE	1.23	1.08	0.11	0.06	38.46	1.68	1.42	0.38	0.25	76.60	1.86	1.48	0.44	0.29	92.86	1.76	1.39	0.36	0.23	76.14
FTI	1.62	1.61	0.47	0.33	76.92	1.81	1.49	0.43	0.29	80.85	1.43	1.43	0.37	0.25	67.86	1.76	1.49	0.41	0.28	76.14
ITE	1.77	1.42	0.40	0.26	84.62	1.96	1.62	0.53	0.36	97.87	1.75	1.40	0.37	0.24	85.71	1.92	1.52	0.46	0.30	92.05
ITI	1.84	1.40	0.39	0.25	84.62	1.89	1.62	0.52	0.35	93.62	1.93	1.34	0.35	0.22	96.43	1.92	1.50	0.45	0.29	93.18
LTE	1.77	1.49	0.42	0.28	76.92	1.85	1.60	0.49	0.34	87.23	1.39	1.29	0.28	0.18	64.29	1.78	1.48	0.42	0.28	78.41
LTI	1.69	1.50	0.42	0.29	76.92	1.72	1.48	0.41	0.27	74.47	1.25	1.28	0.25	0.16	53.57	1.68	1.42	0.36	0.24	68.18
Species level	1.65	1.42	0.37	0.25	73.08	1.82	1.54	0.46	0.31	85.11	1.6	1.37	0.34	0.22	76.79	2.00	1.52	0.47	0.31	100.00

Na: observed number of alleles; *Ne*: effective number of alleles; *SI*: Shannon's information index; *He*: expected heterozygosity; *P*: percentage of polymorphic loci; *FTE*: fibre type exotic; *FTI*: fibre type indigenous; *ITE*: intermediate type exotic; *ITI*: intermediate type indigenous; *LTE*: linseed type exotic; *LTI*: linseed type indigenous.

Table 4
Analysis of molecular variance (AMOVA) within/among sub-populations of flax.

Marker	Source	df	Sum of square	Estimated variation	Percentage of total variation	ϕPT[a]	P-value
Morphological	Among sub-population	5	53.578	0.350	9	0.09	0.01
	Within sub-population	124	426.145	3.437	91		
EST-SSR	Among sub-population	5	26.047	0.186	12	0.12	0.01
	Within sub-population	124	166.453	1.342	88		
gSSR	Among sub-population	5	80.069	0.422	6	0.06	0.01
	Within sub-population	124	896.700	7.231	94		

[a] ϕPT, population differentiation: proportion of the variance among the sub-populations relative to the total variance. The probability of obtaining an equal or lower ф*PT* value was determined by 999 random permutations.

marker system revealed higher intrasub-population variations than intersub-population variations. Nei's unbiased measures of genetic distance (*D*) and genetic identity (*I*) of six sub-populations for morphological, EST-SSR and gSSR markers were calculated (Table 5). The analysis indicates a maximum genetic distance (*D* = 0.11) between fibre-type exotic sub-populations and linseed-type exotic sub-populations compared to other sub-populations based on EST–SSR markers. The lowest genetic distance was recorded between intermediate exotic type and indigenous sub-populations for gSSR (*D* = 0.02) and morphological markers (*D* = 0.006). However, for EST-SSR markers, linseed-type indigenous and intermediate-type exotic sub-populations revealed almost no genetic distance (Fig. 1).

3.2. Population structure

For structure morphological, gSSR, EST-SSR+gSSR markers were used (Fig. 1). The Δ*K* values for the optimum number of clusters in the flax population were determined at *K* = 3 and 2 for morphological markers and gSSR/EST-SSR+gSSR markers, respectively by following the method of Evanno et al. [63]. Within cluster accessions were represented by unique colour while accessions with two different colours indicating admixed forms. The results of the EST-SSR marker-based independent analysis are not presented here, due to the limited number of loci. Morphological markers grouped the accessions into three clusters with Δ*K* = 3. The second cluster with 66 accessions

was the largest and mainly comprises intermediate-type accessions from Indian and exotic origin (25.8% each). Cluster-I (17 accessions) are mainly constituted by fibre-type exotic accessions from West European countries. The remaining accessions grouped into cluster-II (47 accessions). gSSR marker-based clustering grouped fibre-type exotic (20%) and intermediate type (Indian = 20.0%, exotic = 38.2%) together in cluster-I. Most of the linseed type exotic (17.3%) and intermediate type (Indian = 28.0%, exotic = 21.3%) belonged to cluster-II. Further, EST-SSR+gSSR markers grouped accessions into two clusters. Cluster-II was dominated by exotic accessions of intermediate and fibre type (~50% and ~25%, respectively), while cluster-I represented comparatively a mixed group with intermediate (Indian~31%, exotic~23%) and linseed-type exotic (~21%) flax.

The combined analysis of morphological, EST-SSR and gSSR markers grouped exotic accessions into 2 and 3 clusters at *K* = 2 and *K* = 3, respectively (Fig. 2). Initially, with two clusters populations (*K* = 2), exotic accessions of fibre and intermediate type (88.9% and 64%, respectively) were distinguished from the rest of the flax genotypes and grouped into a separate cluster. When one more subgroup was allowed (*K* = 3), Indian accessions of fibres were (63.4%) distinguished and constituted a separate cluster with the Indian intermediate type (48.7%). Cluster-II was predominantly constituted by Indian accessions of linseed type (82.16%). However, Indian intermediate type and exotic linseed type accessions were distributed among the above two clusters.
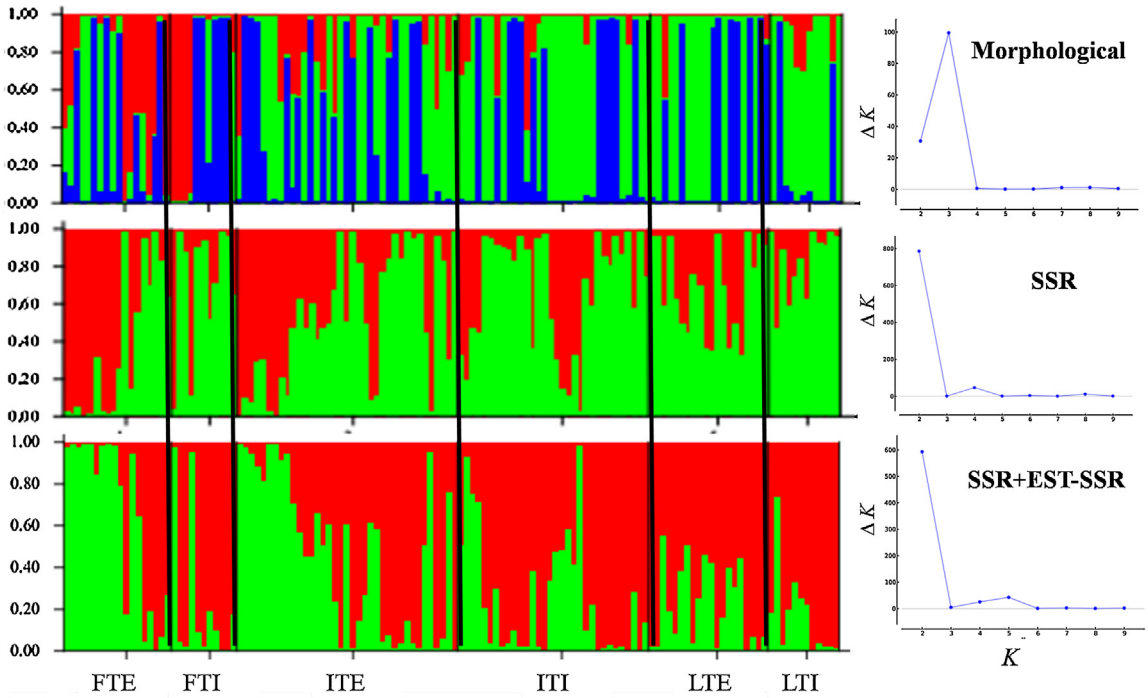
Fig. 1. Population structure flax germplasm accessions based on different markers. Please refer to Table 1 for abbreviation details.
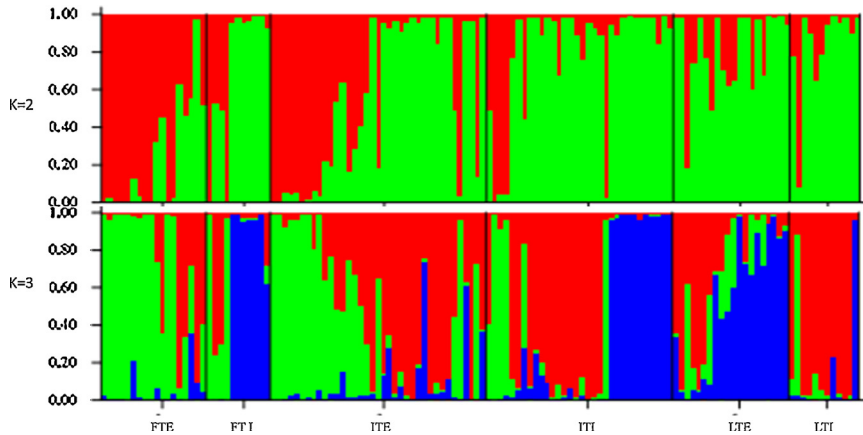

Fig. 2. Population structure of flax germplasm accessions based on combined markers. Please refer to Table 1 for abbreviation details.

Table 5
Nei's unbiased measures of genetic distance (*D*) (upper diagonal) and genetic identity (*I*) (lower diagonal) for six flax sub-populations based on different markers.

Sub-population	Morphological traits						EST–SSR						gSSR					
	FTE	FTI	ITE	ITI	LTE	LTI	FTE	FTI	ITE	ITI	LTE	LTI	FTE	FTI	ITE	ITI	LTE	LTI
FTE	–	0.02	0.03	0.04	0.06	0.06	–	0.10	0.03	0.05	0.11	0.05	–	0.08	0.03	0.04	0.04	0.11
FTI	0.985	–	0.05	0.07	0.10	0.06	0.90	–	0.05	0.03	0.06	0.04	0.92	–	0.07	0.03	0.05	0.07
ITE	0.97	0.96	–	0.006	0.03	0.02	0.97	0.95	–	0.03	0.07	0.00	0.975	0.94	–	0.02	0.03	0.07
ITI	0.96	0.94	0.994	–	0.009	0.02	0.95	0.97	0.97	–	0.02	0.02	0.96	0.97	0.976	–	0.05	0.05
LTE	0.94	0.91	0.97	0.991	–	0.05	0.89	0.94	0.93	0.98	–	0.05	0.96	0.95	0.97	0.96	–	0.08
LTI	0.94	0.94	0.98	0.98	0.95	–	0.96	0.96	0.99	0.98	0.95	–	0.90	0.93	0.93	0.95	0.92	–

FTE: fibre type exotic sub-population; FTI: fibre type indigenous sub-population; ITE: intermediate type exotic sub-population; ITI: intermediate type indigenous sub-population; LTE: linseed type exotic sub-population; LTI: linseed type indigenous sub-population.
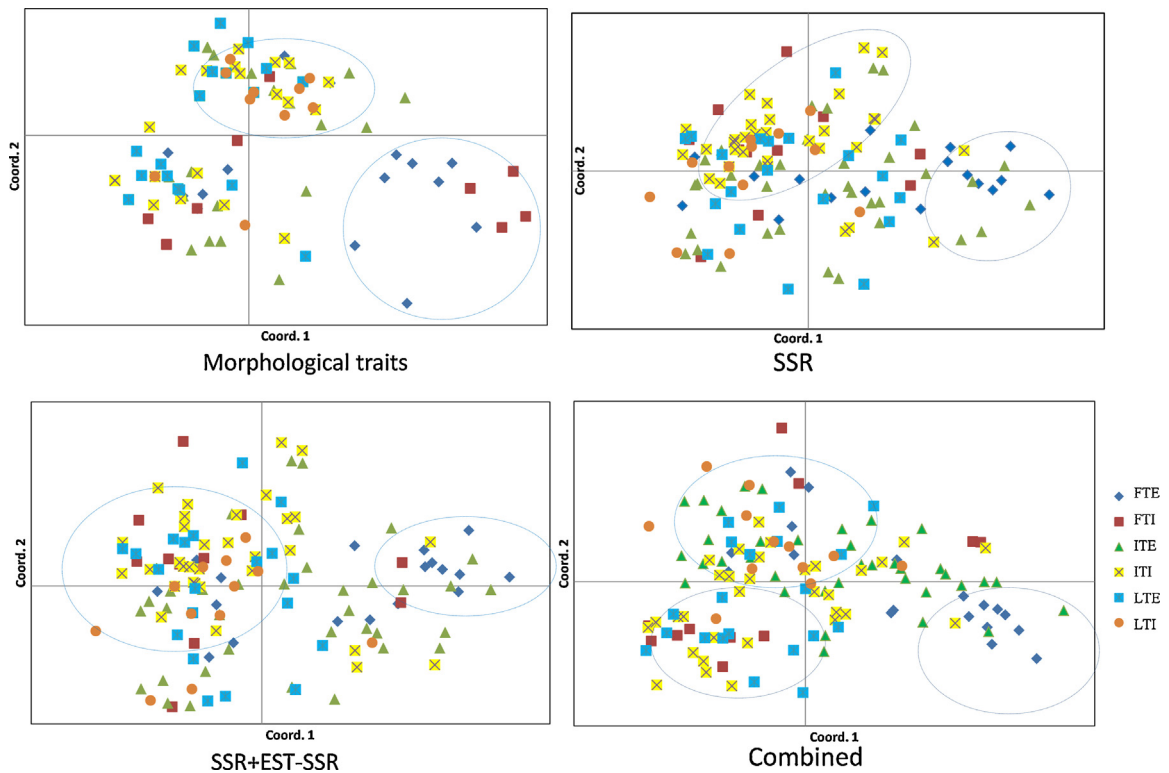
Fig. 3. Principal coordinate analysis (PCoA) plot of flax germplasm populations. Please refer to Table 1 for the abbreviation details.
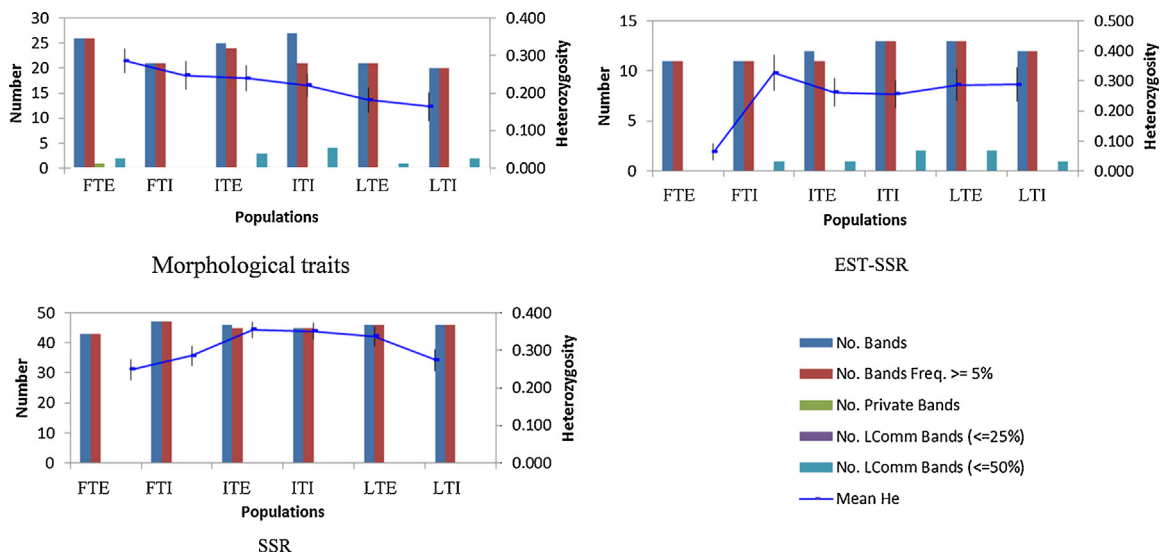

Fig. 4. Heterozygosity in different populations of flax. Please refer to Table 1 for abbreviation details.

The principal coordinate analysis (PCoA) plots for morphological, gSSR, EST-SSR+gSSR and combined (morphological + EST-SSR + gSSR) markers well complement findings of the structure analysis (Fig. 3). In all cases, exotic fibre type and Indian linseed type grouped distinctly. Morphological and EST-SSR+gSSR distinguished exotic fibre type from Indian linseed type. gSSR based

analysis genetically distinguish Indian intermediate type, Indian linseed type and exotic fibre-type flax accessions. However, the combined marker based PCoA plot clearly depicts three clusters. The first axis differentiated fibre-type from linseed-type accessions, except very few outliers, while the second axis separated Indian accessions of fibre type from exotic one. Exotic fibre type accessions

constituted a cluster together with the exotic intermediate type. In contrast, Indian fibre type accessions form a separate cluster with Indian intermediate type accessions. Part of the exotic and Indian accessions of intermediate type exhibited an overlap indicating continuity between fibre and linseed type.

4. Discussion

In the present study, morphological traits were employed to complement microsatellite markers for deciphering genetic diversity and population structure as they were reported as being equally effective for the cause [53,54]. The result highlighted the efficiency of these morphological traits in distinguishing morphotypes namely; fibre type, intermediate and linseed type in flax. In spite of the successful use of morphological traits for morphotypes discrimination, consistent lower values of polymorphic loci percentage (76.79%) were observed in comparison to the employed gSSR primers (85.11%). However, they performed almost *at par* with EST-SSR primers (73.08%). Hence, the efficiency of a given trait/ primer depends on the number of states/fragments it generates as well as on their frequency. The higher number of loci (47) revealed by gSSR markers in comparison to the number of states/loci found for the morphological traits and EST-SSR markers (28 and 13, respectively) explains the observed differential polymorphic loci percentage of the three markers in the study. Similar estimates of genetic relationships among flax sub-populations were obtained for morphological traits, gSSR and EST-SSR markers in the present study. The limitation of morphological traits and EST-SSR markers in revealing genetic diversity and genealogy perhaps stems from the fact that they are functional markers exploring the genetic diversity of a population for genes that may be of agronomic importance [65].

Thus the study suggests that, in spite of the powerfulness of microsatellite markers in depicting genetic relationships, they should not be seen as a substitute of morphological traits [66] as each marker system measures different aspects of genetic variability [67]. In the present study, morphological traits reveal the presence of certain private (rare) alleles that are neither observed in gSSR nor in EST-SSR markers (Fig. 4). Other findings highlighted implications of domestication on the loss of rare alleles in a number of crops, including soybean [68]. Therefore, exotic fibre type accessions with such rare alleles are an asset for Indian flax breeding programmes. Over all, in the light of present results and earlier evidences, a combination of marker systems is recommended to be employed in a complementary manner in order to provide relatively complete information of the genetic relationships in plants.

Morphological markers primarily differentiate fibre-type exotic accessions from the rest of the accessions, while gSSR could differentiate fibre exotic and linseed-type exotic accessions. However, EST-SSR and gSSR markers put together most of the exotic fibre and intermediate type. The relative efficiency of microsatellite markers over morphological markers may be attributed to the neutral nature of DNA markers [69–71]. To further enhance the resolution of population structure combined analysis using morphological, EST-SSR and gSSR markers, particularly at $K = 3$, grouped flax accessions in agreement with their broad morphotype classification and geographical origin. The result highlights the strength of functional markers like morphological traits and EST-SSR markers in augmenting neutral markers (gSSR) during phylogenetic studies of crop-like flax having complicated domestication history [64]. These reports are of special significance for morphological traits given their cost effectiveness and time efficiency while handling large populations. Already their effectiveness in genealogy analysis and genetic diversity assessment are being recognized across crops [53,54].

The PCoA analysis was performed by using genotypic data of morphological traits, gSSR, EST-SSR+gSSR and morphological+gSSR+EST-SSR. Finding that all the markers employed in isolation or in combination successfully differentiated exotic fibre-type flax accessions supports the view that all modern exotic fibre type originated from Eastern Europe and may have a common ancestor [21,24]. However, grouping of Indian linseed type in a separate cluster may be due to small sample size as they represent a few accessions maintained together with fibre type accessions at ICAR-CRIJAF, Barrackpore, Kolkata, India. Hence, they reflect only a limited spectrum of Indian linseed genetic diversity. Finally, based on combined (morphological+gSSR+EST-SSR) markers entire flax sub-populations can be broadly groped into two group namely fibre type and linseed type. This is in conformity with findings of Allaby et al. [21]. Further, the linseed type with presence in more than one cluster appears genetically more diverse than the fibre type (Fig. 3). This is perhaps the outcome of restricted geographical distribution of fibre flax in certain pockets, particularly since the 17th century when cotton replaced flax as the prime fibre crop [1]. In contrast, the popularity of flax for seed oil continued to increase over years, particularly after the industrial revolution [21] that led to widespread cultivation and distribution across continents. The Indian fibre flax appears genetically distinct from exotic fibre type accessions and might have originated from independent domestication events or years of reproductive isolation coupled with evolutionary forces like mutations, selection, recombination, genetic drift and human interference. A similar phenomenon was recorded in flax accessions from West and South East Asia [72]. The finding need to be supplemented with further studies involving more fibre flax accessions from diverse origins, particularly from Asian countries, so as to draw precise conclusions.

5. Future perspective

The genus *Linum* requires end use specific breeding programme to meet diversified applications. To this aim, the information derived from the present study will prove helpful for developing precise selection indices for the identification of superior lines in the flax improvement programme. The identified groups based on both morphological and molecular markers can be an ideal base

material to develop a broad spectrum breeding populations suitable for diversified end uses.

Disclosure of interest

Acknowledgement

References

[1] D. Zohary, M. Hopf, Domestication of plants in the Old World, 3rd ed., Oxford University Press, Oxford, UK, 2000.

[2] C. Morvan, C. Andème-Onzighi, R. Girault, C.D. Himmelsbach, Building flax fibers: more than one brick in the wall, Plant Physiol. Biochem. 41 (2003) 935–944. http://dx.doi.org/10.1016/j.plaphy.2003.07.001.

[3] A. Day, K. Ruel, G. Neutelings, D. Cronier, H. David, S. Hawkins, B. Chabbert, Lignification in the flax stem: evidence for an unusual lignin in bast fibers, Planta 222 (2005) 234–245. http://dx.doi.org/10.1007/s00425-005-1537-1.

[4] M. Mackiewicz-Talarczyk, J. Barriga-Bedoya, J. Mankowski, I. Pniewska, Global flax market situation, in: ID#97, International Conference on Flax and Other Bast Plants, 2008, pp. 408–412 (ISBN #978-0-9809664-0-4).

[5] G.G. Rowland, Growing flax: Production, management and diagnostic guide, Flax Council of Canada and Saskatchewan Flax Development Commission, 1998.

[6] F.S. Hosseinian, G.G. Rowland, P.R. Bhirud, J.H. Dyck, R.T. Tyler, Chemical composition and physicochemical and hydrogenation characteristics of high-palmitic acid solin (low-linolenic acid flaxseed) oil, J. Am. Oil Chem. Soc. 81 (2004) 185–188. http://dx.doi.org/10.1007/s11746-004-0879-6.

[7] H.S. El-Beltagi, Z.A. Salama, D.M. El-Hariri, Evaluation of fatty acids profile and the content of some secondary metabolites in seeds of different flax cultivars (Linum usitatissimum L.), Gen. Appl. Plant Physiol. 33 (2007) 187–202.

[8] T.J. Schmidt, M. Klaes, J. Sendker, Lignans in seeds of Linum species, Phytochemistry 82 (2012) 89–99. http://dx.doi.org/10.1016/j.phytochem.2012.07.004.

[9] M. Bjelková, J. Nôžková, K. Fatrcová-Šramková, E. Tejklová, Comparison of linseed (Linum usitatissimum L.) genotypes with respect to the content of polyunasaturated fatty acids, Chem. Pap. 66 (10) (2012) 972–976. http://dx.doi.org/10.2478/s11696-012-0209-4.

[10] C.M.C. Bassett, D. Rodriguez-Leyva, G.N. Peierce, Experimental and clinical research findings on the cardiovascular benefits of consuming flaxseed, Appl. Physiol. Nutr. Metab. 34 (2009) 965–974. http://dx.doi.org/10.1139/H09-087.

[11] M. Nykter, H.R. Kymäläinen, F. Gates, A.M. Sjöberg, Quality characteristics of edible linseed oil, Agric. Food Sci. 15 (4) (2006) 402–413. http://dx.doi.org/10.2137/145960606780061443.

[12] A.J. Jhala, L.M. Hall, Flax (Linum usitatissimum L.): current uses and future applications, Aust. J. Basic Appl. Sci. 4 (9) (2010) 4304–4312.

[13] J.C.P. Dribnenki, A.G. Green, Linola '947' low linolenic flax, Can. J. Plant Sci. 75 (1995) 201–202.

[14] J.C.P. Dribnenki, A.G. Green, G.N. Atlin, Linola '989' low linolenic flax, Can. J. Plant Sci. 76 (1996) 329–331.

[15] M. Hickey, 100 families of flowering plants, Cambridge University Press, Cambridge, UK, 1988.

[16] J. McDill, M. Repplinger, B.B. Simpson, J.W. Kadereit, The phylogeny of Linum and Linaceae subfamily Linoideae, with implications for their

[17] A. Diederichsen, K. Hammer, Variation of cultivated flax (Linum usitatissimum L. subsp. usitatissimum) and its wild progenitor pale flax (subsp. angustifolium (Huds.) Thell.), Gen. Resour. Crop Evol. 42 (1995) 262–272. http://dx.doi.org/10.1007/BF02431261.

[18] Y.B. Fu, A. Diederichsen, K.W. Richards, G. Peterson, Genetic diversity within a range of cultivars and landraces of flax (Linum usitatissimum L) as revealed by RAPDs, Gen. Resour. Crop Evol. 49 (2) (2002) 167–174. http://dx.doi.org/10.1023/A:1014716031095.

[19] K.S. Gill, D.M. Yermanos, Cytogenetic studies on the genus Linum, Crop Sci. 7 (1967) 623–631.

[20] H. Uysal, Y.B. Fu, O. Kurt, G.W. Peterson, A. Diederichsen, P. Kusters, Genetic diversity of cultivated flax (Linum usitatissimum L.) and its wild progenitor pale flax (Linum bienne Mill.) as revealed by ISSR markers, Gen. Resour. Crop Evol. 57 (2010) 1109–1119. http://dx.doi.org/10.1007/s10722-010-9551-y.

[21] R.G. Allaby, G.W. Peterson, A. Merriwether, Y.B. Fu, Evidence of the domestication history of flax (Linum usitatissimum) from genetic diversity of the sad2 locus, Theor. Appl. Genet. 112 (2005) 58–65. http://dx.doi.org/10.1007/s00122-005-0103-3.

[22] Y.B. Fu, R.G. Allaby, Phylogenetic network of Linum species as revealed by non-coding chloroplast DNA sequences, Gen. Resour. Crop Evol. 57 (2010) 667–677. http://dx.doi.org/10.1007/s10722-009-9502-7.

[23] G. Hillman, The plant remains from Tell Abu Hureyra: a preliminary report, Proc. Prehist. Soc. 41 (1975) 70–73.

[24] H. Helbaek, Domestication of food plants in the Old World, Science 130 (1959) 365–372.

[25] B. Dewilde, 20 Eeuwen vlas in Vlaanderen, Tielt-Bussum, Lannoo, 1983 (439 p).

[26] K.S. Gill, Linseed, Indian Council of Agricultural Research, New Delhi, India, 1987, p. 386. .

[27] A.G. Green, Y. Chen, S.P. Singh, J.C.P. Dribnenki, Flax, in: C. Kole, T.C. Hall (Eds.), Compendium of transgenic crop plants, Blackwell Publishing Ltd, Oxford, UK, 2008, pp. 199–226.

[28] I.A. Chandrawat, R. Maurya, P.K. Singh, S.A. Ranade, H.K. Yadav, Diversity analysis in Indian genotypes of linseed (Linum usitatissimum L.) using AFLP markers, Gene 549 (1) (2014) 171–178. http://dx.doi.org/10.1016/j.gene.2014.07.067.

[29] B. Chaudhary, M.K. Tripathi, S.K. Pandey, H.R. Bhandari, D.R. Meena, S.P. Prajapati, TIARA (JRF 2) – A flax fibre (linen) strain for high fibre yield, ICAR News 21 (3) (2015) 17. .

[30] T.L. Odong, J. van Heerwaarden, J. Jansen, T.J.L. van Hintum, F.A. van Eeuwijk, Determination of genetic structure of germplasm collections: are traditional hierarchical clustering methods appropriate for molecular marker data, Theor. Appl. Genet. 123 (2011) 195–205. http://dx.doi.org/10.1007/s00122-011-1576-x.

[31] K.W. Von, S. Danert, Zur Systematik von Linum usitatissimum L, Kulturpflanze 3 (1962) 341–388.

[32] A. Diederichsen, Comparison of genetic diversity of flax (Linum usitatissimum L.) between Canadian cultivars and a world collection, Plant Breed. 120 (4) (2001) 360–362. http://dx.doi.org/10.1046/j.1439-0523.2001.00616.x.

[33] A. Diederichsen, T.A. Rozhmina, A.A. Zhuchenko, K.W. Richards, Screening for broad adaptation in 96 flax (Linum usitatissimum L.) accessions under dry and warm conditions in Canada and Russia, Plant Genet. Resour. Newslett. 146 (2006) 9–16.

[34] A. Diederichsen, J.P. Raney, Seed colour, seed weight and seed oil content in Linum usitatissimum accessions held by Plant Gene Resources of Canada, Plant Breed. 125 (4) (2006) 372–377. http://dx.doi.org/10.1111/j.1439-0523.2006.01231.x.

[35] H. Tyson, M.A. Fieldes, C. Cheung, J. Starobin, Isozyme relative mobility (Rm) changes related to leaf position; apparently smooth Rm trends and some implications, Biochem. Genet. 23 (9–10) (1985) 641–654.

[36] E. Mansby, O. Diaz, R. von Bothmer, Preliminary study of genetic diversity in Swedish flax (Linum usitatissimum), Genet. Resour. Crop Evol. 47 (4) (2000) 417–424. http://dx.doi.org/10.1023/A:1008721528588.

[37] W. Spielmeyer, A.G. Green, D. Bittisnish, N. Mendham, E.S. Lagudah, Identification of quantitative trait loci contributing to Fusarium wilt resistance on an AFLP linkage map of flax (Linum usitatissimum), Theor. Appl. Genet. 97 (4) (1998) 633–641. http://dx.doi.org/10.1007/s001220050939.

[38] I. Everaert, R.J. De, L.M. De, W.J. Van, B.E. Van, Most similar variety grouping for distinctness evaluation of flax and linseed (Linum usitatissimum L.) varieties by means of AFLP and morphological data, Plant Var. Seed 14 (2) (2001) 69–87.

[39] D. Wiesnerová, I. Wiesner, ISSR-based clustering of cultivated flax germplasm is statistically correlated to thousand seed mass, Mol.

Biotechnol. 26 (3) (2004) 207–213. http://dx.doi.org/10.1385/MB:26:3:207.

[40] S. Wright, The genetical structure of populations, Ann. Eugen. 15 (1951) 323–354.

[41] L. Excoffier, P. Smouse, J. Quattro, Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data, Genetics 131 (1992) 479–491.

[42] J.K. Pritchard, M. Stevens, P. Donnelly, Inference of population structure using multilocus genotype data, Genetics 155 (2000) 945–959.

[43] B.A. Payseur, P. Jing, R.J. Haasl, A genomic portrait of human microsatellite variation, Mol. Biol. Evol. 28 (2011) 303–312. http://dx.doi.org/10.1093/molbev/msq198.

[44] M. Wen, H. Wang, Z. Xia, M. Zou, C. Lu, W. Wang, Development of EST-SSR and genomic-SSR markers to assess genetic diversity in *Jatropha curcas* L, BMC Res. Notes 3 (2010) 42. http://dx.doi.org/10.1186/1756-0500-3-42.

[45] I. Wiesner, D. Wiesnerova, E. Tejklova, Effect of anchor and core sequence in microsatellite primers on flax fingerprinting patterns, J. Agric. Sci. 137 (2001) 37–44. http://dx.doi.org/10.1017/S0021859601001162.

[46] C. Roose-Amsaleg, P.E. Cariou, D. Vautrin, R. Tavernier, M. Solignac, Polymorphic microsatellite loci in *Linum usitatissimum*, Mol. Ecol. Notes 6 (2006) 796–799. http://dx.doi.org/10.1111/j.1471-8286.2006.01348.x.

[47] X. Deng, S. Long, D. He, X. Li, Y. Wang, J. Liu, X. Chen, Development and characterization of polymorphic microsatellite markers in *Linum usitatissimum*, J. Plant Res. 123 (2010) 119–123. http://dx.doi.org/10.1007/s10265-009-0271-3.

[48] B.J. Soto-Cerda, R.A. Carrasco, G.A. Aravena, H.A. Urbina, C.S. Navarro, Identifying novel polymorphic microsatellites from cultivated flax (*Linum usitatissimum* L.) following data mining, Plant Mol. Biol. Rep. 29 (2011) 753–759. http://dx.doi.org/10.1007/s11105-010-0270-5.

[49] S. Cloutier, E. Miranda, K. Ward, N. Radovanovic, E. Reimer, A. Walichnowski, R. Datla, G. Rowland, S. Duguid, R. Ragupathy, Simple sequence repeat marker development from bacterial artificial chromosome end sequences and expressed sequence tags of flax (*Linum usitatissimum* L.), Theor. Appl. Genet. 125 (2012) 685–694. http://dx.doi.org/10.1007/s00122-012-1860-4.

[50] S. Cloutier, Z. Niu, R. Datla, S. Duguid, Development and analysis of EST–SSRs for flax (*Linum usitatissimum* L.), Theor. Appl. Genet. 119 (2009) 53–63. http://dx.doi.org/10.1007/s00122-009-1016-3.

[51] B.J. Soto-Cerda, S.H. Urbina, C. Navarro, O.P. Mora, Characterization of novel genic SSR markers in *Linum usitatissimum* (L.) and their transferability across eleven *Linum* species, Electron. J. Biotechnol. 14 (2011) 6. http://dx.doi.org/10.2225/vol14-issue2-fulltext-6.

[52] A. Kundu, D. Sarkar, A. Bhattacharjee, N. Topdar, M.K. Sinha, B.S. Mahapatra, A simple ethanol wash of the tissue homogenates recovers high quality genomic DNA from *Corchorus* species characterized by highly acidic and proteinaceous mucilages, Electron. J. Biotechnol. 14 (2011)http://dx.doi.org/10.2225/vol14-issue1-fulltext-4.

[53] H.K. Sharma, M. Sarkar, S.B. Choudhary, A.A. Kumar, R.T. Maruthi, J. Mitra, P.G. Karmakar, Diversity analysis based on agro-morphological traits and microsatellite based markers in global germplasm collections of roselle (*Hibiscus sabdariffa* L.), Ind. Crop. Prod. 89 (2016) 303–315. http://dx.doi.org/10.1016/j.indcrop.2016.05.027.

[54] S. Hegay, M. Geleta, T. Bryngelsson, A. Asanaliev, L. Garkava-Gustavsson, H. Persson Hovmalm, R. Ortiz, Genetic diversity analysis in *Phaseolus vulgaris* L. using morphological traits, Gen. Resour. Crop Evol. 61 (2014) 555–566. http://dx.doi.org/10.1007/s10722-013-0056-3.

[55] W. Kulpa, S. Danert, Zur Systematik von *Linum usitatissimum* L, Kulturpflanze 3 (1962) 341–388.

[56] K. Weising, H. Nybom, K. Wolff, G. Kahl, DNA fingerprinting in plants. Principles, methods, and applications, 2nd ed., CRC Press, Boca Raton, FL, USA, 2005, p. 472. .

[57] J.R. Freeland, H. Kirk, S.D. Peterson, Molecular ecology, 2nd ed., Wiley, London, 2011, p. 449. .

[58] R. Peakall, P.E. Smouse, GENALEX 6 genetic analysis in Excel. Population genetic software for teaching and research, Mol. Ecol. Notes 6 (2006) 288–295. http://dx.doi.org/10.1111/j.1471-8286.2005.01155.x.

[59] F.C. Yeh, T.J.B. Boyle, Population genetic analysis of codominant and dominant markers and quantitative traits, Belg. J. Bot. 129 (1997) 157. .

[60] M.C. de Vicente, C. López, T. Fulton, Genetic Diversity Analysis with Molecular Marker Data: Learning Module, International Plant Genetic Resources Institute (IPGRI), Rome, Italy, 2004.

[61] J.K. Pritchard, X. Wen, D. Falush, Documentation for structure software: Version 2. 3, 2010 http://www.pritch.bsd.uchicago.edu/software/structure.v.2.3.1/documentation.pdf.

[62] D.A. Earl, M. vonHoldt, Structure harvester: a website and program for visualizing structure output and implementing the Evanno method, Conserv. Genet. Resour. 4 (2) (2012) 359–361. http://dx.doi.org/10.1007/s12686-011-9548-7 http://www.taylor0.biology.ucla.edu/structureHarvester/.

[63] G. Evanno, S. Regnaut, J. Goudet, Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study, Mol. Ecol. 14 (2005) 2611–2620. http://dx.doi.org/10.1111/j.1365-294X.2005.02553.x.

[64] Yong-Bi. Fu, Population-based resequencing revealed an ancestral winter group of cultivated flax: implication for flax domestication processes, Ecol. Evol. 2 (3) (2012) 622–635. http://dx.doi.org/10.1002/ece3.101.

[65] B.M. Mulato, M. Möller, M.I. Zucchi, V. Quecini, J.B. Pinheiro, Genetic diversity in soybean germplasm identified by SSR and EST-SSR markers, Pesq. Agropec. Bras. 45 (3) (2010) 276–283. http://dx.doi.org/10.1590/S0100-204X2010000300007.

[66] A. Karp, S. Kresovich, K.V. Bhat, W.G. Ayad, T. Hodgkin, Molecular Tools in Plant Genetic Resources Conservation: A Guide to the Technologies, IPGRI Technical Bulletin No. 2, International Plant Genetic Resources Institute, Rome, Italy, 1997.

[67] A. Belaj, L. León, Z. Satovic, R. de la Rosa, Variability of wild olives (*Olea europaea* subsp. *europaea var. sylvestris*) analysed by agro-morphological traits and SSR markers, Sci. Hortic. [TD.JX] 129 (2011) 561–569. http://dx.doi.org/10.1016/j.scienta.2011.04.025.

[68] D.L. Hyten, Q. Song, Y. Zhu, I.Y. Choi, R.L. Nelson, J.M. Costa, J.E. Specht, R.C. Shoemaker, P.B. Cregan, Impacts of genetic bottlenecks on soybean genome diversity, Proc. Natl. Acad. Sci. U S A 103 (2006) 16666–16671. http://dx.doi.org/10.1073/pnas.0604379103.

[69] M. Hagidimitirou, A. Katsiotis, G. Menexes, C. Pontikis, M. Loukas, Genetic diversity of major Greek olive cultivars using molecular (AFLPs and RAPDs) markers and morphological traits, J. Am. Soc. Hortic. Sci. 130 (2005) 211–217.

[70] G. Corrado, M. La Mura, O. Ambrosino, G. Pugliano, P. Varricchio, R. Rao, Relationships of Campanian olive cultivars: comparative analysis of molecular and phenotypic data, Genome 52 (2009) 692–700. http://dx.doi.org/10.1139/g09-044.

[71] R. Rao, M. La Mura, G. Corrado, O. Ambrosino, I. Foroni, E. Perri, G. Pugliano, Molecular diversity and genetic relationships of southern Italian olive cultivars as depicted by AFLP and morphological traits, J. Hortic. Sci. Biotechnol. 84 (2009) 261–266. http://dx.doi.org/10.1080/14620316.2009.11512514.

[72] Y.B. Fu, G.G. Rowland, S.D. Duguid, K.W. Richards, RAPD analysis of 54 North American flax cultivars, Crop Sci. 43 (4) (2005) 1510–1515. http://dx.doi.org/10.2135/cropsci2003.1510.