

# Genetic identity of common buckwheat (*Fagopyrum esculentum* Moench) landraces locally cultivated in the Alps

G. Barcaccia · M. Volpato · R. Gentili · T. Abeli ·  
G. Galla · S. Orsenigo · S. Citterio ·  
S. Sgorbati · G. Rossi

Received: 12 January 2015 / Accepted: 2 June 2015 / Published online: 18 June 2015  
© Springer Science+Business Media Dordrecht 2015

**Abstract** The conservation of landraces in Europe is challenging because very often they have already disappeared or cannot be properly identified, which in turn prevents any possibility for their utilization. This work deals with the collection of molecular and historical data to identify and study the original landraces of common buckwheat (*Fagopyrum esculentum* Moench), locally cultivated in Northern Italy (Valtellina) and to date surviving among other commercial varieties, recently introduced in the same areas of the Alps. As plant materials of *F. esculentum*, we

analyzed a number of Italian accessions along with two foreign accessions from Poland and Nepal, for a total of 174 individuals. Molecular investigations were based on a set of eight nuclear SSR marker loci. The mean observed heterozygosity over all accessions was equal to  $H_o = 0.466$ , being significantly lower than the expected heterozygosity ( $H_e = 0.764$ ). A major finding was the recognition of a marked inbreeding rate ( $F_{it} = 0.387$ ) and a reduced fixation index ( $F_{st} = 0.061$ ), indicating that most genetic variation is found within populations. A significant overall gene flow among accessions was found ( $N_m = 3.846$ ). Results indicated that only two of the examined accessions, the so-called “Nustran” and “Curunin”, could be considered, authentic Valtellina landraces. On the basis of results, we successfully developed a multi-locus marker system and identified a number of co-dominant marker alleles suitable for genetic traceability and authenticity certification of a “Nustran” and a “Curunin” autochthonous landraces of Valtellina and its food derivatives (i.e., Pizzoccheri, Polenta taragna).

**Electronic supplementary material** The online version of this article (doi:10.1007/s10722-015-0273-z) contains supplementary material, which is available to authorized users.

G. Barcaccia · M. Volpato · G. Galla  
Laboratory of Plant Genetics and Breeding, DAFNAE,  
University of Padova, Campus of Agripolis,  
35020 Legnaro, Italy

R. Gentili (✉) · S. Citterio · S. Sgorbati  
Department of Earth and Environmental Sciences,  
University of Milano - Bicocca, Piazza della Scienza 1,  
20126 Milan, Italy  
e-mail: rodolfo.gentili@unimib.it

T. Abeli · G. Rossi  
Department of Earth and Environmental Sciences,  
University of Pavia, Via S. Epifanio 14, 27100 Pavia, Italy

S. Orsenigo  
Department of Agricultural and Environmental  
Sciences - Production, Landscape, Agroenergy,  
University of Milan, Via Celoria 2, 20122 Milan, Italy

**Keywords** Agrobiodiversity conservation ·  
Autochthonous landraces · *Fagopyrum esculentum* ·  
Food traceability · Genetic diversity · Microsatellites

## Introduction

A landrace, according to Camacho et al. (2005), is “a dynamic population of cultivated plants that has

historical origin and distinct identity, and lacks formal crop improvement, as well as often being genetically diverse, locally adapted and associated with traditional farming systems”. Landraces represent not only a valuable source of potentially useful traits, but also an irreplaceable bank of highly co-adapted genotypes (Brush 1999). Despite their importance, the cultivation or use of landraces is challenging because in many cases they have already disappeared or cannot be properly identified (de Carvalho et al. 2013; Jaradat 2014), which in turn, prevent the possibility for their conservation and valorization.

In this context, the importance of landraces of species of interest for agriculture has rapidly increased and the concept of biodiversity conservation has been extended from wild species to crops and livestock (Jarvis et al. 2008; Pessina et al. 2011). The importance of landraces has increased for several reasons, the most important of which are: (1) the awareness of customers looking for organic products related to local chains of production, (2) the fact that landraces are sources of genetic diversity that find utility to improve standard crops in a changing climate (Esquinas-Alcazar 1993; Fideghelli and Engel 2009), (3) their potential economic value at different market scale. Such an interest actively stimulates the study, rescue and valorization of landraces (e.g. Janick et al. 2007; Cavagna et al. 2012). Moreover the political entities implemented their legislation to assure the conservation of landraces threatened by genetic erosion, especially in the most developed countries (e.g. EU Directives 2008/62/EC and 2009/145/EC “providing for certain derogations, for acceptance of vegetable landraces and varieties which have been traditionally grown in particular localities and regions and are threatened by genetic erosion”). Furthermore, worldwide initiatives for the conservation and valorization of landraces are currently active (e.g. Bioversity International and more recently the initiatives of the Global Crop Diversity Trust, The Global Seed Vault, etc.).

The major problems related to the use of landraces are to understand whether historically documented landraces yet exist in their original locations, on one side, and to distinguish them within the plethora of similar commercial varieties, on the other. Knowledge of germplasm diversity among local populations and breeding stocks is expected to have a significant impact on the improvement of crop plants. It can be

obtained by surveying both qualitative and quantitative morphological traits, or using molecular markers for investigating polymorphisms at the DNA sequence level (Barcaccia et al. 1999). In fact, besides linkage mapping, gene targeting and assisted breeding, the plant DNA polymorphism assays are powerful tools for characterizing and investigating germplasm resources and genetic relatedness. Among the PCR-derived techniques, a co-dominant marker system widely exploited for population genetics is that based on microsatellites or simple sequence repeat (SSR) markers (Morgante and Olivieri 1993). In the last few decades, molecular markers have been successfully used to characterize and preserve many landraces across the most important crop species.

Among cultivated species, an interesting case is represented by landraces of common buckwheat (*Fagopyrum esculentum* Moench), a plant of the Polygonaceae family largely distributed and used for human feeding. The traditional use of buckwheat is common in East Asia, where the species is likely native (Ohnishi 1994, 1998), and also in North-East Europe, where buckwheat was probably introduced since the Neolithic era (Alenius et al. 2013) with documents attesting its cultivation in the Alps since XVI century (Giacomini 1954; Ferranti et al. 2002). The European populations of common buckwheat are likely to derive from China through the silk-road (Ohnishi 1993). In Europe, and in the Alpine region, the species *F. tataricum* is also present, being its occurrence highly dependent on the cultivation of *F. esculentum* (Hammer et al. 1999). Common buckwheat is an open-pollinated plant species characterized by allogamy due to heterostyly system of incompatibility, as a consequence individuals within a stylar type are self-incompatible as well as cross-incompatible (Ohnishi and Asano 1999; Zeller and Hsam 2001; Cawoy et al. 2009). Previous studies based on molecular markers in *F. esculentum* showed that local populations of this species are represented by collections of highly heterozygous and heterogeneous plants, although certain landraces from Southern Europe often revealed a genetic variation lower than that found in Asiatic and Northern Europe landraces, and a greater genetic differentiation of landraces that often show differences in seed husk color (Ohnishi 1993). As a consequence, European populations of cultivated *F. esculentum* are of special interest to address general issues concerning the

conservation of landraces in general. Differently from South-Eastern Asian countries, like for instance Tibet, Nepal and China, the use of landraces of buckwheat in the Alps declined since sixties, with the risk of extinction of landraces (Ohnishi 1993). This is a common fate for many landraces in most developed countries, where the use of standard products has already brought to their loss and genetic erosion (Negri et al. 2009; Raggi et al. 2013). Recently common buckwheat is returning popular as a consequence of increasing gluten-free product market. However, in Italy, most of the landraces of *F. esculentum* have been largely replaced by commercial seeds, imported from Central Europe (see also Laghetti et al. 1993). Nevertheless, few local populations supposed to be well adapted to the agro-environmental conditions of the historical sites of cultivation are still cultivated in the Central Alps (Valtellina).

In this work we put together molecular and historical data with the aim to implement a multi-locus genotyping reference system, based on co-dominant microsatellite markers, to identify, if yet existing, and to characterize, the genetic structure of these relict landraces of common buckwheat, historically cultivated in the Italian Alps. Genetic variation within farmer's populations, often indicated with vernacular names, and their genetic relationships with commercial varieties, currently cultivated in Valtellina and Trentino areas are here investigated. The description of the genetic traits of the common buckwheat landraces, historically cultivated in the Italian Alps, allows the conservation of the germplasm (i.e., seed banking), favors its traditional utilization and, in a broader perspective, supports discussion on general issues related to the identification of landraces by means of molecular analyses.

## Materials and methods

### Plant materials

Plant material of *Fagopyrum esculentum* ( $2n = 2x = 16$ ) was represented by 9 Italian accessions (Table 1; Fig. 1) collected in the Valtellina valley, Lombardy (VARA, VARB, VARC, RS4, RS5, RS6 and RS10) and in Trento province (TNC and TNT) during 2013. Moreover, 2 foreign accessions from Poland

(POL1) and Nepal (NEP2), even collected in 2013, were used for comparisons in all molecular analyses. The gene pool of each population was represented by a subsample of 12–19 randomly chosen plants (see Table 1). According to information of local farmers based on tradition, in Valtellina the most ancient germplasm is called “Nustran”, whereas a French variety was introduced later and called “Curunin” (or “Furest”, meaning “coming from abroad”). The “Nustran” and “Curunin” selected accessions are easily distinguished on the basis of seed morphology and colour: “Nustran” is cultivated up to 850 m a.s.l. and shows bigger seeds than “Curunin” which is cultivated up to 1200 m a.s.l.; moreover “Nustran” has brown seeds, whilst “Curunin” has grey seeds (see Fig. 1).

### Genomic DNA isolation

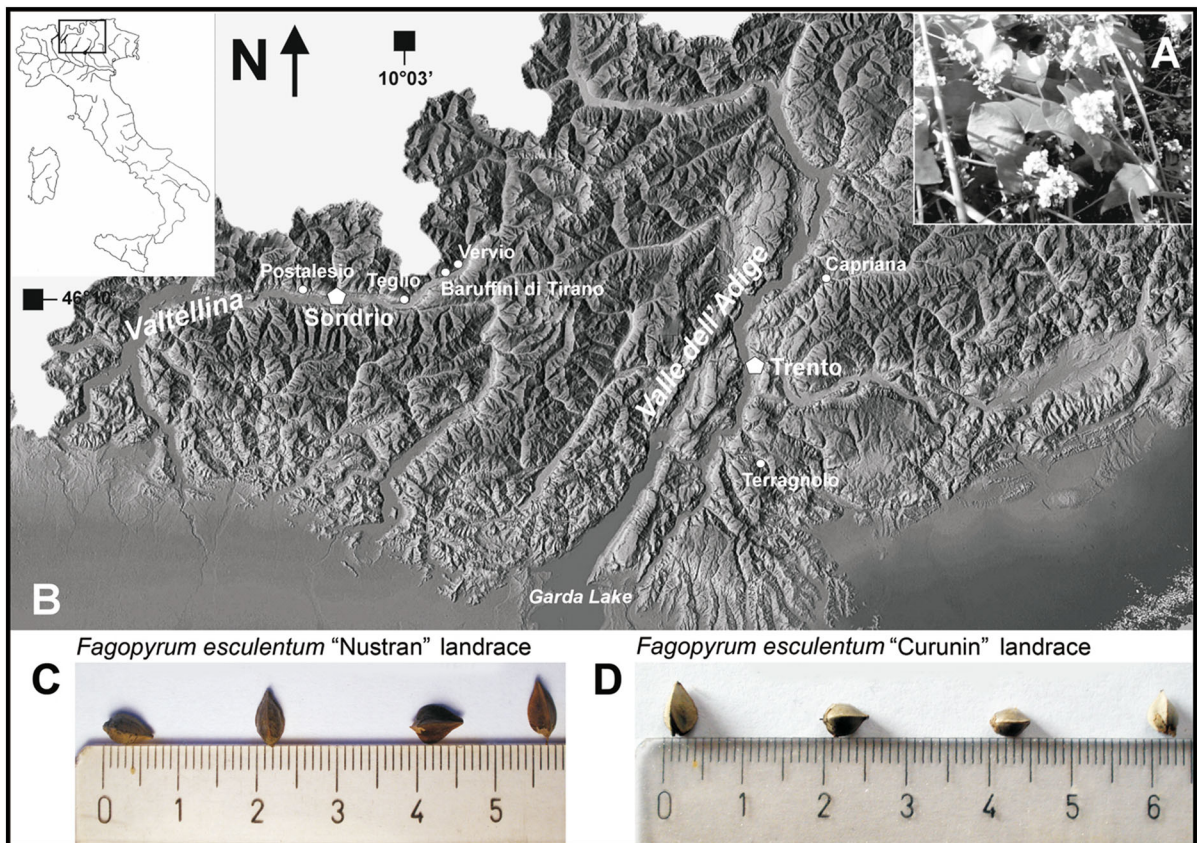
Genomic DNA was isolated from 100 mg of fresh leaf tissue using a CTAB-based protocol for DNA purification, modified from Doyle and Doyle (1987). The integrity of extracted DNA samples was estimated by electrophoresis on a 0.8 % agarose/1 × TAE gel containing 1 × Sybr Safe DNA stain (Life Technologies). Both purity and quantity of DNA extracts were assessed by means of the NanoDrop 2000c UV–Vis Spectrophotometer (Thermo Scientific).

### Analysis of SSR markers

PCR amplifications and microsatellite (SSR) marker analyses for the genotyping of the 216 common buckwheat plants were performed as described by Barkley et al. (2007) using a M13 labeled primer (5'-TTGTAAAACGACGGCCAGT-3'), in combination with a specific SSR-targeting forward primer with a 5'-M13 tail and a specific SSR-targeting reverse primer (see supplementary materials, Table S-1). The set of SSR marker loci investigated in this study was obtained from Konishi et al. (2006) and Iwata et al. (2005). Each PCR reaction consisted of a 10 µl final volume containing 1 × Platinum Master Mix (Life Technologies), 10 % PCR Enhancer, 0.05 mM forward primer with a 5'-M13 tail, 0.15 mM reverse primer, 0.10 mM M13-labelled primer (Life Technologies), 10 ng DNA and distilled water. All PCR reactions were performed using a thermal cycler with 96-well plate (Applied Biosystems 9600). All SSR marker loci were assayed using the same cycling

**Table 1** Information on common buckwheat accessions analyzed in this study

Accessions (ID code and region)	Location/origin (country and province)	Sample size	Vernacular name
Italian accessions			
Subgroup Valtellina (Lombardy region, Italy)			
VARA	Teglio (SO)	16	Nustran
RS4	Postalesio (SO)	18	Nustran
RS5	Vervio (SO)	19	Nustran
VARB	Teglio (SO)	18	Curunin
VARC	Teglio (SO)	15	Curunin
RS6	Baruffini di Tirano (SO)	17	Curunin
RS10	Teglio (SO)	17	Curunin
Subgroup Trentino (Trentino region, Italy)			
TNC	Capriana (TN)	15	n.a.
TNT	Terragnolo (TN)	15	n.a.
Foreign accessions			
POL1	n.d. (Poland)	12	–
NEP2	Chaurikarka (Nepal)	12	–

**Fig. 1** A Sample of *Fagopyrum esculentum*; B Area of provenience of the Italian accessions of the species. Different seed shape and colour of “Nustran” (C) and “Curunin” (D) biotypes. (Color figure online)

conditions. The program consisted of 1 cycle at 95 °C for 2 min for the initial denaturing step, 5 cycles of 95 °C for 30 s, 60 °C for 30 s and 68 °C for 45 s, 35 cycles of 95 °C for 30 s, 58 °C for 30 s and 68 °C for 45 s, 1 cycle of 68 °C for 60 min for final extension and a 4 °C hold for temporal storage. The PCR products were then diluted in formamide and subjected to capillary electrophoresis with the ABI PRISM 3130xl Genetic Analyzer. The LIZ500 was adopted as molecular weight standard.

### Marker data analysis

Polymorphisms highlighted within and between the different common buckwheat accessions were used to estimate the level of genetic diversity and differentiation existing within and among the considered accessions. Descriptive genetic diversity and differentiation statistics, as well as inbreeding coefficients, were calculated using the POPGENE software package version 1.32 (Yeh et al. 1997). The average number of alleles observed per locus ( $n_a$ ) was computed as the arithmetic mean of the total number of alleles observed at each locus. The effective number of alleles per locus was computed as  $n_e = 1 / \sum p_i^2$ , where  $p_i$  is the frequency of the  $i$ th allele (Kimura and Crow 1964). For each marker locus and over all loci, the genetic diversity was computed as  $H = 1 - \sum p_i^2$ , corresponding to the expected heterozygosity of Nei (1973). Genetic diversity values were reported not only as observed heterozygosity ( $H_o$ ), but also as expected heterozygosity ( $H_e$ ) and average heterozygosity ( $H_a$ ). Moreover, the unbiased heterozygosity of Nei (1978) was calculated using the algorithm of Levene (1949). These parameters were also calculated for all buckwheat accessions. The phenotypic diversity of marker allele profiles was estimated using the Shannon's information index,  $I = - \sum p_i^2 \ln p_i^2$ , as reported by Lewontin (1974). A hierarchical analysis of variance with estimation of F-statistics (Wright 1965) was also performed. Estimates of heterozygosity within subpopulations ( $F_{is}$ ) and between subpopulations ( $F_{it}$ ) were determined, as was the fixation index ( $F_{st}$ ) according to Wright (1978). In particular, the inbreeding coefficients  $F_{is}$  and  $F_{it}$  were computed for single SSR loci to measure the deficiency (positive values) or excess (negative values) of heterozygotes at each locus. Moreover,  $F_{st}$  measures the genetic effect

of total population subdivision as the proportional reduction in overall heterozygosity owing to variation in SSR allele frequencies among different subpopulations (i.e., accessions). Values of  $F_{st}$  were averaged across populations and over all loci. Gene flow was estimated from  $F_{ST}$  as follows:  $N_m = 0.25(1 - F_{st})/F_{st}$  (McDermott and McDonald 1993).

The cluster analysis was performed according to the unweighted pair-group arithmetic average method (UPGMA), and the dendrogram and centroids of all populations were constructed from the symmetrical mean genetic similarity matrix. All calculations and analyses were conducted using the Numerical Taxonomy and Multivariate Analysis System (NTSYS-pc) version 2.21q (Rohlf 1993). The proportion of genetic similarity (GS) in all pair-wise comparisons of individuals was calculated by applying the coefficient of simple matching (Rohlf 1993). The principal coordinates analysis (PCA) technique was then applied to compute the first two principal components from the qualitative data matrix. Neighbor-joining (NJ) trees were computed on the basis of the symmetrical matrix of pair-wise genetic similarity estimates for the whole population of buckwheat. A bootstrap statistical analysis was conducted to measure the stability of the computed branches with 1000 resampling replicates.

The estimation of molecular variance (AMOVA) existing within and between common buckwheat populations was performed using the software GenAlEx 6.501 (Peakall and Smouse 2006).

The overall genetic structure of the buckwheat accessions (174 individuals) was also investigated using the model-based clustering algorithm implemented in the STRUCTURE software (Falush et al. 2003), which groups the individuals according to marker allele combination and distribution. All simulations were executed with no a priori population information. Calculations were performed with 500,000 iterations, a burning period equal to 500,000 and by adopting the following setting: (1) the admixture ancestry model; (2) allele frequencies correlated among populations. Computations were made with 20 replicated runs, with each run exploring a range of K spanning from 1 to 12. The most likely value of K was estimated using  $\Delta K$ , as reported in other studies (Evanno et al. 2005).

## Results

Genomic loci of SSR markers scored from a minimum of 7 to a maximum of 22 marker alleles (15 on average), with values of polymorphism information content (PIC) that ranged from 0.69 to 0.95, hence underlying high levels of genetic variability not only within but also among buckwheat accessions (see Supplementary materials, Table S-1).

### Descriptive statistics of SSR marker loci in common buckwheat accessions

Descriptive statistics over all SSR loci along with information on the amount of genetic diversity found across molecular markers and plant accessions are reported in Table 2. All the exploited marker loci were polymorphic among individuals and accessions, with frequency of the most common marker allele that was equal to 0.684 per SSR locus (Table 2). The mean number of observed alleles per SSR locus was found as high as 15.000 over all common buckwheat accessions (Table 2). The mean number of observed alleles per SSR locus was equal to 14.000 (subgroup Valtellina) and 7.625 (subgroup Trentino) in the local accessions, whereas it varied from 5.750 to 8.000 in the foreign (Nepal and Poland, respectively) accessions (Table 2). Moreover, the effective number of alleles per SSR locus was equal to 4.850 (Table 2). It is worth mentioning that these descriptive statistics were comparable for both “Nustran” and “Curunin” biotypes, being the number of observed and effective alleles per SSR locus equal to 11.875 and 11.250, and to 5.253 and 4.540, respectively (see Table 2).

Estimates of both the unbiased Nei’s genetic diversity ( $H$ ) and the Shannon’s information index of phenotypic diversity ( $I$ ) were used as parameters for the characterization of the gene pools of local and introduced buckwheat accessions (Table 2). The mean genetic diversity computed over all marker loci and plant accessions was equal to 0.762 ( $SD = 0.120$ ), and ranged from a minimum of 0.483 (locus FES1346) to a maximum of 0.860 (locus FES1303). The measures of molecular genetic diversity computed for the subgroups of plant accessions were relatively similar, varying from 0.673 (Trentino accession TNC) to 0.764 (Valtellina accession VARA). Moreover, the two main Valtellina accessions scored values of genetic diversity equal to 0.776 for “Nustran” and to

0.743 for “Curunin” accessions (Table 2). The Shannon’s information index computed over all marker loci and plant accessions was equal to 1.875 ( $SD = 0.427$ ), and ranged from a minimum of 0.970 (locus FES1346) to a maximum of 2.285 (locus FES1303). This information index of marker phenotypic diversity was higher in the Valtellina accessions with respect to Trentino accessions (1.876 vs. 1.519, respectively), showing the lowest estimate in the Nepal accession (1.438). Moreover, the two main Valtellina accessions scored measures of marker phenotypic diversity equal to 1.890 for “Nustran” and to 1.762 for “Curunin” accessions (Table 2).

Descriptive statistics for Nei’s genetic diversity ( $H$  statistics) and Wright’s inbreeding coefficients ( $F$  statistics) for single marker locus and population accessions were also computed (Table 3). The mean observed heterozygosity ( $H_o$ ) was as low as 0.466 in the population accessions, ranging from 0.423 to 0.531, and being equal to 0.452 for the subgroup Valtellina and to 0.458 for the subgroup Trentino. The “Nustran” and “Curunin” accessions were shown to have an observed heterozygosity of  $H_o = 0.429$  and  $H_o = 0.471$ , respectively (Table 3). An important finding is that the expected heterozygosity ( $H_e$ ) over all the plant accessions scored a mean value equal to 0.764, ranging from 0.698 to 0.790, being significantly higher than the observed heterozygosity. For instance, the estimates of expected heterozygosity for the “Nustran” and “Curunin” accessions were as high as  $H_e = 0.784$  and  $H_e = 0.750$ , respectively (Table 3). As a consequence, Wright’s inbreeding coefficients  $F_{is}$  and  $F_{it}$  scored positive values, revealing a marked deficiency of heterozygotes at each locus and across individual accessions and accession subgroups.

The value of  $F_{is}$  calculated for each single locus and over all accessions underlined a strong defect of heterozygosity that, on average, was higher for the Italian (Valtellina and Trentino) accessions than for the Poland and Nepal entries, with a mean value equal to  $F_{is} = 0.347$  (Table 3). In particular, this inbreeding coefficient for plant accessions ranged from 0.252 up to 0.446, being on average equal to 0.422 and 0.331 for “Nustran” and “Curunin” accessions, respectively (0.371 for the Valtellina subgroup) and equal to 0.333 for the Trentino subgroup (Table 3). Values of the fixation index  $F_{st}$  computed for each locus and over all accessions are reported on Table 3. This parameter

**Table 2** Descriptive statistics of genetic diversity calculated across markers and buckwheat accessions, including allele sample size of individual genotypes (S), frequency of the most common marker allele ( $p_i$ ), average number of observed alleles ( $n_a$ ) and effective number of alleles ( $n_e$ ) per marker locus, unbiased Nei's molecular genetic diversity (H), and estimates of Shannon's information index of marker phenotypic diversity (I)

SSR loci	S	$p_i$	$n_a$	$n_e$	H	I
FES1407	343	0.406	22.000	4.688	0.787	2.063
FES1346	347	0.684	8.000	1.934	0.483	0.970
FES2802	341	0.284	13.000	5.421	0.816	1.922
FES1286	331	0.469	18.000	3.939	0.746	1.965
FES1303	327	0.258	19.000	7.149	0.860	2.285
FES1094	272	0.272	18.000	6.469	0.845	2.220
FES3331	343	0.377	7.000	4.078	0.755	1.551
FES3508	275	0.378	14.000	5.118	0.805	2.024
Nustran accessions						
VARA	31	0.562	7.250	4.930	0.764	1.696
RS4	36	0.650	8.500	4.690	0.750	1.712
RS5	38	0.625	7.375	3.962	0.709	1.542
Overall Nustran	105	0.614	11.875	5.253	0.776	1.890
Curunin accessions						
VARB	35	0.684	7.125	4.014	0.717	1.570
VARC	29	0.750	6.500	3.832	0.707	1.504
RS6	33	0.684	6.500	3.908	0.710	1.501
RS10	33	0.736	7.375	4.230	0.705	1.581
Overall Curunin	130	0.711	11.250	4.540	0.743	1.762
Subgroup Valtellina	235	0.668	14.000	4.963	0.763	1.876
Trentino accessions						
TNC	29	0.767	6.000	3.830	0.673	1.408
TNT	30	0.625	6.250	3.790	0.699	1.452
Subgroup Trentino	59	0.694	7.625	3.995	0.703	1.519
Foreign accessions						
NEP2 (Nepal)	22	0.750	5.750	3.742	0.696	1.438
POL1 (Poland)	33	0.735	8.000	4.431	0.737	1.666
Overall buckwheat accessions	348	0.684	15.000	4.850	0.762	1.875
SD		0.246	5.398	1.615	0.120	0.427

The overall values and standard deviations are also reported for each of the parameters

ranged from 0.034 to 0.116 across marker loci and varied from 0.026 to 0.094 across plant accessions, with an average value equal to 0.061 (Table 3). Such a fixation index indicates that most of the genetic variation (i.e., about 94 %) is occurring within accessions and that genetic differentiation among local populations and foreign introductions is very low (i.e., only 6 % of the total genetic variation was found among common buckwheat accessions). A supplementary classification of the Italian accessions underlined a narrow range of variation for the  $F_{st}$  estimates, as this fixation index was low for both the Trentino subgroup (0.024) and the Valtellina subgroup (with similar values for “Nustran” and “Curunin” accessions, 0.050 and 0.048, respectively).

Estimates of gene flow ( $N_m$ ) were also computed for each locus and across accessions (Table 3). The calculated  $N_m$  values were  $>1$  in all assayed marker loci, ranging from 2.4 to 9.5 over all accessions (on average,  $N_m = 3.846$ ), and hence supporting a high genetic introgression and a little genetic differentiation between autochthonous populations locally cultivated in the Alps and foreign accessions introduced in Northern Italy.

Genetic distances among common buckwheat accessions and genetic structure analysis

The level and gradient of genetic variation existing among buckwheat individual plants within as well as

**Table 3** Summary of H-statistics, F-statistics and gene flow ( $N_m$ ) estimates for individual SSR marker loci and buckwheat accessions. Nei's genetic diversity values were calculated as observed heterozygosity ( $H_o$ ), expected heterozygosity ( $H_e$ ) and average heterozygosity ( $H_a$ )

	H-statistics			F-statistics			Gene flow
	$H_o$	$H_e$	$H_a$	$F_{is}$	$F_{it}$	$F_{st}$	$N_m$
Locus							
FES1407	0.441	0.789	0.737	0.396	0.430	0.057	4.122
FES1346	0.314	0.484	0.463	0.328	0.351	0.034	7.074
FES2802	0.627	0.818	0.784	0.194	0.225	0.039	6.181
FES1286	0.594	0.748	0.716	0.159	0.202	0.051	4.701
FES1303	0.579	0.863	0.812	0.287	0.325	0.054	4.420
FES1094	0.235	0.849	0.774	0.715	0.736	0.074	3.133
FES3331	0.723	0.757	0.714	-0.009	0.044	0.052	4.523
FES3508	0.217	0.807	0.720	0.684	0.721	0.116	1.909
Nustran accessions							
VARA	0.423	0.790	0.764	0.446	0.460	0.026	9.550
RS4	0.425	0.771	0.750	0.433	0.458	0.043	5.515
RS5	0.437	0.729	0.710	0.385	0.443	0.094	2.399
Overall Nustran	0.429	0.784	0.741	0.422	0.450	0.050	4.865
Curunin accessions							
VARB	0.447	0.738	0.717	0.377	0.404	0.044	5.432
VARC	0.531	0.733	0.710	0.252	0.292	0.053	4.438
RS6	0.489	0.735	0.710	0.311	0.348	0.053	4.438
RS10	0.430	0.727	0.710	0.394	0.427	0.053	4.438
Overall Curunin	0.471	0.750	0.710	0.331	0.363	0.048	4.978
Subgroup Valtellina	0.452	0.767	0.723	0.371	0.406	0.055	4.250
Trentino accessions							
TNC	0.459	0.698	0.672	0.317	0.359	0.061	3.818
TNT	0.456	0.724	0.670	0.319	0.363	0.064	3.641
Subgroup Trentino	0.458	0.716	0.686	0.333	0.349	0.024	10.117
Foreign accessions							
NEP2	0.512	0.729	0.737	0.305	0.330	0.035	6.824
POL1	0.531	0.761	0.696	0.237	0.305	0.089	2.559
Overall accessions	0.466	0.764	0.715	0.347	0.387	0.061	3.846
SD	0.193	0.120	0.108	0.251	0.244	0.026	1.614

Wright's inbreeding coefficients  $F_{is}$ ,  $F_{it}$  and  $F_{st}$  are also reported as measures of genetic differentiation within and between buckwheat accessions

between accessions was primarily investigated by calculating genetic similarity estimates in all possible pair-wise comparisons among the 174 samples using the whole set of marker alleles scored at all genomic loci. In particular, the individual plants from different accessions exhibited a very high level of genetic similarity that varied between 85 and 98 %. The overall population of common buckwheat did not show any segmentation or division in subpopulations, in fact none of the ordination methods (i.e., UPGMA trees and PCA centroids) based on genetic similarity estimates revealed any relevant formation of subgroups of

individuals or subclusters of accessions (see Supplementary materials, Figure S-1, panels A and B).

An additional cluster analysis was based on marker allele frequencies for computing genetic distances in all possible pair-wise comparisons among the 11 accessions. The NJ tree of all the accessions was then constructed revealing a slightly structured distribution of common buckwheat accessions (Fig. 2). In fact, the Italian accessions from Valtellina and Trentino areas were clustered together in a single well-defined subgroup, supported by a bootstrap value equal to 50 %, with the exception of a "Nustran" accession

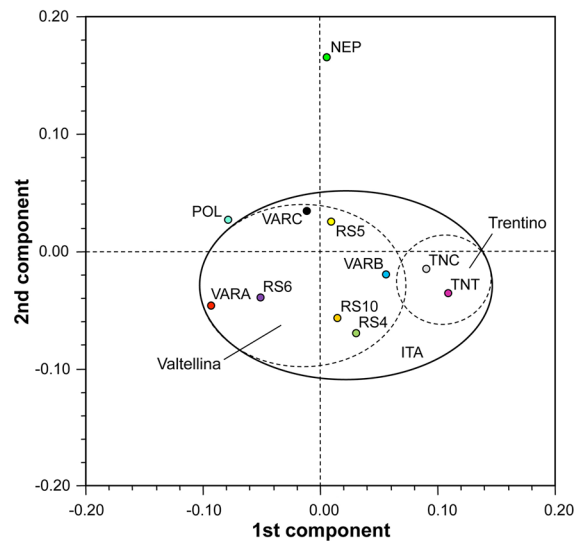


(VARA) and a “Curunin” accession (RS6), that were clustered apart, with a bootstrap value as high as 94 %. It is worth nothing that the accession from Nepal was split from that introduced from Poland, being the former clearly differentiated from all Italian accessions (bootstrap equal to 71 %) and the latter closely related to Italian accessions (bootstrap equal to 27 %).

If it is true that “Nustran” and “Curunin” accession subgroups are not genetically distinguishable each other, it is also true that two of the Italian accessions largely cultivated in the past in the Valtellina valley have shown marker allele compositions and frequencies that make them genetically traceable from the rest of accessions, including foreign ones (Fig. 2).

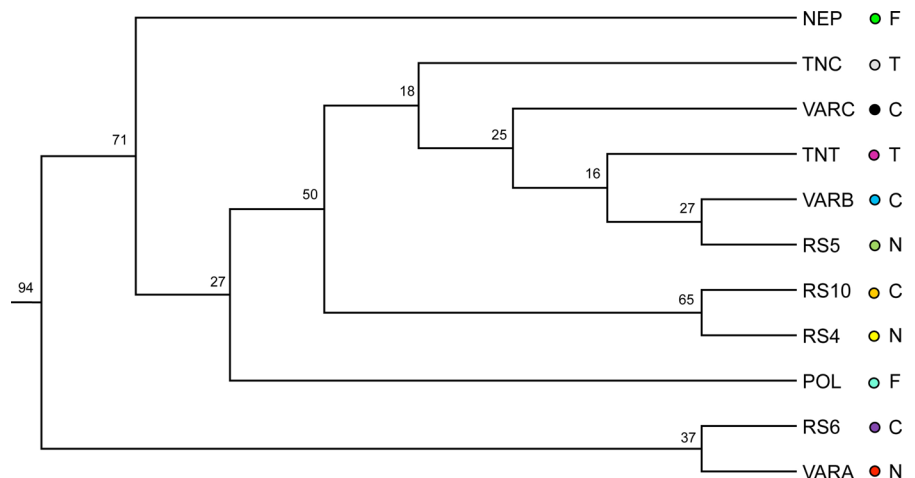
Principal coordinate analysis allowed the definition of centroids for all the 11 common buckwheat accessions (Fig. 3). Most of the accessions could be discriminated from each other on the basis of their marker alleles and genotypes, even though the Italian accessions and subgroups were closely plotted in the central part of the four main quadrants (Fig. 3). The first two principal components were able to explain 36 % of the total genetic variation found within the population as a whole. In particular, the first component, which explains about 21 % of the total diversity, was positively associated with the Trentino accessions and negatively associated with one “Nustran” and two “Curunin” accessions from Valtellina and the Poland accession. The second component, which explains 15 % of the total diversity, was clearly able to discriminate the Nepal accession (Fig. 3).

On the basis of all SSR marker alleles, the genetic structure of the common buckwheat accessions was investigated using the STRUCTURE software. Each plant accession is represented by a vertical histogram portioned into  $K = 3$  colored segments that represent the estimated membership of each hypothesized ancestral genotype (Fig. 4). Single plant accessions were sorted by population types and by membership coefficients: the clustering of genotypes according to the former data revealed the occurrence of highly



**Fig. 3** 2D centroids derived from the genetic similarity estimates computed among plant accessions in all possible pair-wise comparisons using the whole SSR marker data set

**Fig. 2** NJ tree based on the whole SSR marker data set, displaying the genetic distances among pair-wise comparisons of plant accessions, with branches and nodes supported by bootstrap values



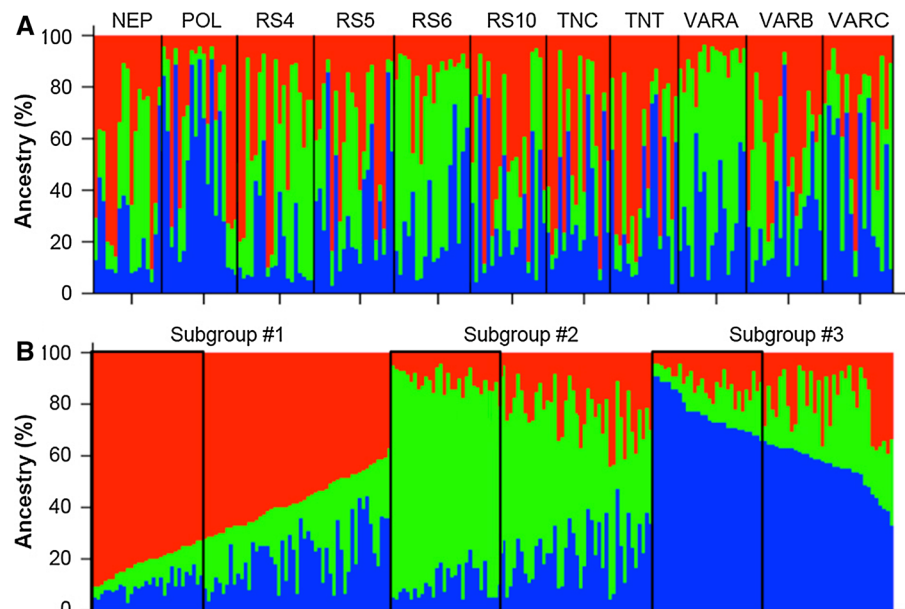
admixed genotypes (Fig. 4, upper panel A), whereas the grouping of genotypes according to the latter data showed three distinct subgroups each with putatively pure genotypes (Fig. 4, lower panel B). Based on this analysis, the two Italian accessions VARA (“Nustran” biotype) and RS6 (“Curunin” biotype) both from Valtellina proved to be the most homogeneous populations in terms of high ancestry assignment and low proportion of admixed genotypes (Fig. 4).

Contingency tests were performed in order to underline the most discriminant markers ( $p < 0.05$ ) useful for genetic identification of the “Nustran” and “Curunin” accessions, VARA and RS6, respectively (see Supplementary materials, Tables S-2 and S-3). Contingency tests were performed for each of the marker alleles over all accessions and led to the identification of significant differences in terms of relative frequencies among the “Nustran” VARA, the “Curunin” RS6, and the remaining accessions analyzed in this study. In particular, several marker alleles scored a frequency significantly different in these two accessions compared to the other accessions (Tables S-2 and S-3). Interestingly, 5 and 3 marker alleles belonging to three marker loci/linkage groups displayed values of  $p < 0.01$ , indicating highly significant differences and highly discriminant frequencies for the “Nustran” accession VARA (Tables S-2) and the “Curunin” accession RS6 (Table S-3).

## Discussion

Local crops are important sources of genetic variation as they are associated with a number of favorable ecological factors ensuring robustness, adaptation to local conditions and peculiar organoleptic and health characters (Veteläinen et al. 2009). The present study is the first that compares the genetic population diversity of common buckwheat (*F. esculentum*) in the Alps. Germplasm was collected on the basis of historical indications given by local people that, hand down and cultivate since centuries this old germplasm resource (i.e., autochthonous landraces), that nowadays is on the edge of extinction. Supposed landraces cultivated in two regions of the Italian Alps were compared with commercial materials from Europe (Poland) and also with materials cultivated in the area of origin of common buckwheat (Himalaya, Nepal) in order to assess their intra-population and inter-population genetic variation and differentiation, the degree of genetic structuration of the population as a whole, and in turn the definition of genetic features of the alpine landraces. In particular, 9 common buckwheat accessions deriving from different farmer’s populations of Valtellina valley and Trento province and two reference populations from Poland and Nepal were characterized in great detail using mapped microsatellite markers. Detecting genetic variability by means of

**Fig. 4** Population structure of the buckwheat population as a whole as estimated with STRUCTURE software. Each plant accession is represented by a vertical histogram portioned into  $K = 3$  colored segments (red, green and blue) that represent the estimated membership of each hypothesized ancestral genotype. Single plant accessions were sorted by population information (A) and membership coefficients (B). (Color figure online)



marker allele compositions and proportions, and of multi-locus genotypes in different populations of alpine agro-ecosystems is a crucial phase towards sustainable management, use (consumption and improvement) and conservation of genetic resources of the yet surviving autochthonous landraces of *F. esculentum*.

### Genetic variability

Results showed that the selected microsatellites were very informative, having a medium–high variation in the buckwheat populations (e.g. PIC estimates varying from 0.69 to 0.95). Nei's genetic diversity, ranging from 0.673 to 0.764 was comparable to that reported for both natural and cultivated populations by Konishi and Ohnishi (2007), using microsatellite markers (i.e.,  $H$  values ranging from 0.65 and 0.83). As a general rule, the outcrossing breeding system of self-incompatible species, along with the annual life cycle and large geographic ranges may have a central role in influencing the conservation of good levels of genetic diversity in common buckwheat populations. Honeybees have been reported to be an effective pollinator favoring gene flow among *F. esculentum* populations in Western Europe (Cawoy et al. 2009). Another factor favoring diversity can be ascribed to the high proportion of interspecific hybridization observed with the infesting weed *Fagopyrum tataricum* Gaertn. Mender-Drienyovszki et al. (2013) and Chen et al. (2004) found that intercrossing between *F. esculentum* and *F. tataricum* can be as high as 37 %. These features probably contribute to the high within population genetic diversity documented in our samples.

The Nei's genetic diversity and the Shannon's information index of the different samples proved to be higher in the "Nustran" and "Curunin" accessions of the subgroup Valtellina with respect to Trentino and other foreign accessions. A possible explanation could be the intense and everywhere diffuse cultivation in the past of common buckwheat in Valtellina with respect to other parts of the Alps. Furthermore, the highest inbreeding coefficients ( $F_{is}$  and  $F_{st}$ ) found in the Valtellina accessions could be explained with the longer isolation of Valtellina in the past compared to other regions of the Alps. However, despite a high allelic variability, the estimates of observed heterozygosity ( $H_o = 0.466$ , on average) were much lower than the ones of expected heterozygosity ( $H_e = 0.764$ ,

on average) across all analyzed populations. A comparable pattern was detected by Song et al. (2011) for the Korean landraces: values of observed heterozygosity ( $H_o = 0.42$ , on average) were in agreement with our findings, whereas the expected heterozygosity accounted slightly lower values ( $H_e = 0.53$ ) compared with our findings. Heterozygosity deficiency, with a robust and significant deviation than predicted by Hardy–Weinberg equilibrium, is a common observation in varieties of cultivated plants, as farmers use to renew cultivations, year after year, with seeds taken from a reduced gene pool, determining a human induced bottleneck and artificial selection (Doebley et al. 2006).

The concomitant recognition of a marked inbreeding rate ( $F_{is} = 0.347$ , on average), a reduced fixation index ( $F_{st} = 0.061$ ) and a high gene flow ( $N_m = 3.846$ ) between over all accessions indicate that most allele diversity and genotype variation is found within accessions and that genetic differentiation among accessions is very low. Such a low genetic differentiation can be attributed to a high gene flow among populations, due to pollen dispersal among adjacent fields and seed exchange among local farmers, once more consistent with the finding of high levels of genetic diversity preserved within single populations.

### Population structure

The main evolutionary driving force that determine genetic structure in cultivated plant populations are gene flow, farmer selection activity and environmental heterogeneity associated to random genetic drift (Neal 2004). Our study revealed that the amount of genetic diversity does not change significantly among regional areas and plant populations (data from AMOVA analysis). Such inter-regional genetic relatedness in cultivated populations of common buckwheat most likely would originate from seed exchanges among farmers of different countries over years. The clustering method implemented in the STRUCTURE analysis suggested low structuration of populations and the existence of admixed ancestry based on three distinct gene pools. However, The PCA and UPGMA clustering analyses clearly suggested that the accession named VARA of the "Nustran" biotype and, to a minor extent, the accession named RS6 of the "Curunin" biotype are populations of the Valtellina valley which have gene pools clearly distinct from the

other accessions, hence clustering apart from the core of accessions.

#### Implications for the identification of local alpine *F. esculentum* landraces

Our results demonstrated that a selected panel of microsatellite markers is useful for the genetic identification and characterization of landraces of cultivated plants. In particular, the DNA genotyping method applied in this research allowed us to identify the landraces that most likely represent the historical gene pool of common buckwheat cultivated in Valtellina among different local accessions and foreign materials. Other studies reported that different molecular techniques can be used for characterization of the landraces and investigation on their origin, like DNA barcoding (Nicolò et al. 2011) and DNA fingerprinting (Gwanama et al. 2000; Ferriol et al. 2004). In general, these tools are of special interest because they provide basic information useful, if not essential, for the preservation and valorization (even commercial) of abandoned landraces.

In the specific case the common buckwheat accessions VARA (provided by Dr. P. Roccatagliata) and RS6 (provided by Dr. M. Deplaz) will require further conservation efforts (i.e., gene bank, started at Pavia University) and should be preferred to other accessions for the commercial exploitation of this crop plant in the Valtellina valley.

Additionally, on the basis of our molecular data, we successfully identified a number of marker alleles across mapped loci suitable for genetic identification, traceability and authenticity certification (Protected Geographic Indication) of given “Nustran” and “Curunin” biotypes and their food derivatives (i.e., flour, pizzoccheri, chat, polenta taragna).

In conclusion, we are confident that population genetics of *F. esculentum* landraces is a crucial step for the identification, conservation and valorization of local genetic resources in this species in their original agro-ecosystems.

**Acknowledgments** This research was partially funded by the project “Produzione di Potenziali Varietà da Conservazione ortive ed agrarie Lombarde Registrate” (V.C.L.R.) funded by Regione Lombardia PSR 2007-2013, Misura 124. Authors thank Dr. Pietro Roccatagliata and Mr. Patrizio Mazzucchelli for seed provision and important information on the history of buckwheat cultivation in Valtellina valley and Dr. Giorgio

Perini (La Pimpinella, Pergine Valsugana, Trento) for seed provision of the Trento province.

#### References

- Alenius T, Mökkönen T, Lahelma A (2013) Early farming in the northern boreal zone: reassessing the history of land use in southeastern Finland through high-resolution pollen analysis. *Geoarchaeology* 28:1–24
- Barcaccia G, Albertini E, Tavoletti S, Falcinelli M, Veronesi F (1999) AFLP fingerprinting in *Medicago* spp.: its development and application in linkage mapping. *Plant Breed* 118:335–340
- Barkley NA, Dean RE, Pittman RN, Wang ML, Holbrook CC, Pederson GA (2007) Genetic diversity of cultivated and wild-type peanuts evaluated with M13-tailed SSR markers and sequencing. *Genet Res* 89:93–106
- Brush SB (1999) Genes in the field: on farm conservation of crop diversity. IPGRI, IDRC Lewis, Boca Raton
- Camacho V, Taina C, Maxted N, Scholten M, Ford-Lloyd B (2005) Defining and identifying crop landraces. *Plant Genet Res* 3:373–384
- Cavagna P, Camerini G, Fibiani M, Andreani L, Cella R, Concia L, Lo Scalzo R (2012) Characterization of the rescued ‘Voghera’ sweet pepper landrace grown in northern Italy. *Span J Agric Res* 10:1059–1069
- Cawoy V, Ledent J-F, Kinet J-M, Jaquemart A-M (2009) Floral biology of Common buckwheat (*Fagopyrum esculentum* Moench). *Eur J Plant Sci Biotechnol* 3:1–9
- Chen QF, Hsam SLK, Zeller FJ (2004) A study of cytology, isozyme, and interspecific hybridization on the big-achene group of buckwheat species (*Fagopyrum*, Polygonaceae). *Crop Sci* 44:1511–1518
- de Carvalho MAAP, Bebeli PJ, Bettencourt E, Costa G, Dias S, Dos Santos TMM, Slaski JJ (2013) Cereal landraces genetic resources in worldwide GeneBanks: a review. *Agron Sustain Dev* 33:177–203
- Doebley JF, Gaut BS, Smith BD (2006) The molecular genetics of crop domestication. *Cell* 127:1309–1321
- Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bull* 19:11–15
- Esquinas-Alcazar JT (1993) Plant Genetic Resources. In: Hayward MD, Bosemark NO, Romagosa I (eds) *Plant breeding: principles and prospects*. Chapman & Hall, London, pp 33–51
- Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* 14:2611–2620
- Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* 164:1567–1587
- Ferranti R, Pirola A, Penati F (2002) Il paesaggio vegetale della Provincia di Sondrio, Suppl. Il Naturalista Valtellinese, Atti del Museo Civico di Storia Naturale di Morbegno 13:38
- Ferriol M, Picó B, Nuez F (2004) Morphological and molecular diversity of a collection of *Cucurbita maxima* landraces. *J Am Soc Hortic Sci* 129:60–69

- Fideghelli C, Engel P (2009) Biodiversity and local genetic resources: from knowledge to exploitation. *Acta Hort* 817:295–310
- Giacomini V (1954) Il grano siberiano (*Fagopyrum tataricum*) in Valtellina. Ramponi, Sondrio
- Gwanama C, Labuschagne MT, Botha AM (2000) Analysis of genetic variation in *Cucurbita moschata* by random amplified polymorphic DNA (RAPD) markers. *Euphytica* 113:19–24
- Hammer K, Knüpfner H, Laghetti G, Perrino P (1999) Seeds from the Past. A catalogue of crop germplasm in Central and North Italy. IdG, Bari
- Iwata H, Imon K, Tsumura Y, Ohsawa R (2005) Genetic diversity among Japanese indigenous Common buckwheat (*Fagopyrum esculentum*) cultivars as determined from AFLP and SSR markers and quantitative agronomic traits. *Genome* 48:367–377
- Janick J, Paris HS, Parrish DC (2007) The Cucurbits of mediterranean antiquity: identification of taxa from ancient images and descriptions. *Ann Bot* 100:1441–1457
- Jaradat AA (2014) The vanishing wheat landraces of the Fertile Crescent. *Emir J Food Agric* 26:203–217
- Jarvis DI, Brown AHD, Cuong PH et al (2008) A global perspective of the richness and evenness of traditional crop-variety diversity maintained by farming communities. *Proc Natl Acad Sci USA* 105:5326–5331
- Kimura M, Crow J (1964) The number of alleles that can be maintained in a finite population. *Genetics* 49:725–738
- Konishi T, Ohnishi O (2007) Close genetic relationship between cultivated and natural populations of Common buckwheat in the Sanjiang area is not due to recent gene flow between them—An analysis using microsatellites markers. *Genes Genet Syst* 82:53–64
- Konishi T, Iwata H, Yashiro K, Tsumura U, Ohsawa R, Yasui Y, Ohnishi O (2006) Development and characterization of microsatellite markers for common buckwheat. *Breed Sci* 56:277–285
- Laghetti G, Hammer K, Perrino P (1993) Collecting in north-west Italy. *FAO/IBPGR Plant Genetic Resources Newsletter* 91(92):23
- Levene H (1949) On a matching problem in genetics. *Ann Math Stat* 20:91–94
- Lewontin RC (1974) The genetic basis of evolutionary change. Columbia University Press, New York
- McDermott M, McDonald BA (1993) Gene flow in plant pathosystems. *Annu Rev Phytopathol* 31:353–373
- Mendler-Drienyovszki N, Cal AJ, Dobránszki J (2013) Progress and prospects for interspecific hybridization in buckwheat and the genus *Fagopyrum*. *Biotechnol Adv* 31:1768–1775
- Morgante M, Olivieri AM (1993) PCR-amplified microsatellite as markers in plant genetics. *Plant J* 3:175–182
- Neal D (2004) Introduction to population biology. Cambridge University Press, Cambridge
- Negri V, Maxted N, Veteläinen M (2009) European Landrace Conservation: an introduction. In: Veteläinen M, Negri V, Maxted N (eds) European landraces: on-farm conservation management and use. Biodiversity technical bulletin 15. Biodiversity International, Rome
- Nei M (1973) Analysis of gene diversity in subdivided populations. *Proc Natl Acad Sci USA* 70:3321–3323
- Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 89:583–590
- Nicolè S, Erickson DL, Ambrosi D, Bellucci B, Lucchin M, Papa R, Kress WJ, Barcaccia G (2011) Biodiversity studies in *Phaseolus* species by DNA barcoding. *Genome* 54:529–545
- Ohnishi O (1993) Population genetics of cultivated Common buckwheat, *Fagopyrum esculentum* Moench. VIII. Local differentiation of land races in Europe and the silk road. *Jap J Genet* 68:303–316
- Ohnishi O (1994) Buckwheat in Karakoram and the Hindukush. *Fagopyrum* 14:17–25
- Ohnishi D (1998) Search for the wild ancestor of buckwheat I. Description of new *Fagopyrum* species and their distribution in China. *Fagopyrum* 15:18–28
- Ohnishi O, Asano N (1999) Genetic diversity of *Fagopyrum homotropicum*, a wild species related to Common Buckwheat. *Genet Resour Crop Evol* 46:389–398
- Peakall R, Smouse PE (2006) GenAIEx v. 6: genetic analysis in Excel. Population genetic software for teaching and research. *Mol Ecol Notes* 6:288–295
- Pessina D, Gentili R, Barcaccia G, Nicolè S, Rossi G, Barbisti S, Sgorbati S (2011) DNA content, morphometric and molecular marker analyses of *Citrus limonimedica* Lush., *C. limon* (L.) Burm. and *C. medica* L. for the determination of their variability and genetic relationships within the genus *Citrus*. *Sci Hort* 129:663–673
- Raggi R, Tiranti B, Negri V (2013) Italian common bean landraces: diversity and population structure. *Genet Resour Crop Evol* 60:1515–1530
- Rohlf EJ (1993) NTSYS-pc: numerical taxonomy and multivariate analysis system, version 1.80. Applied Biostatistics Inc, New York
- Song JY, Lee G-A, Yoon M-S, Ma K-H, Choi Y-M, Lee J-R, Jung Y, Park H-J, Kim C-K, Lee M-C (2011) Analysis of genetic diversity and population structure of Buckwheat (*Fagopyrum esculentum* Moench) landraces of Korea using SSR markers. *Korean J Plant Res* 24:702–711
- Veteläinen M, Negri V, Maxted N (2009) European landraces: on-farm conservation management and use. Biodiversity technical bulletin 15. Biodiversity International, Rome
- Wright S (1965) The interpretation of population structure by F-statistics with special regard to systems of mating. *Evolution* 19:395–420
- Wright S (1978) Evolution and the genetics of populations, variability within and among natural populations, vol 4. University of Chicago Press, Chicago
- Yeh FC, Yang RC, Boyle T, Ye ZH, Mao JX (1997) POPGENE: the user-friendly shareware for population genetic analysis. University of Alberta, Molecular Biology and Biotechnology Centre, Edmonton
- Zeller FJ, Hsam SLK (2001) Genetic analysis of morphological characteristics in Common Buckwheat (*Fagopyrum esculentum* Moench). *Proceedings of the 8th ISB*, pp 214–217