Albanian j. agric. sci. 2013;12 (4): 741-746 *Agricultural University of Tirana*

RESEARCH ARTICLE **(Open Access)**

Diversity of albanian plant genetic resources inventory assesed by eurisco passport descriptors

*BELUL GIXHARI[1], HAIRI ISMAILI[1], FABJAN LASHI[2], ALBAN IBRALIU[3], SONIA DIAS[4]

[1] Albanian Gene Bank, Agricultural University of Tirana, Tirana, Albania

[2] Agri-Bussines and Management Department, Agricultural University of Tirana, Tirana, Albania

[3] Plant Production Departments, Agricultural University of Tirana, Tirana, Albania

[4]EURISCO Coordinator (Bioversity International), Roma, Italy.

Abstract

Assessment of diversity of Albanian National Inventory is carried out using import statistics by passport descriptors from Plant Genetic Resources database. Flora of Albania identified more than 3 250 species of plants, but only 2% of them are included in the National Inventory (NI) of Albania in EURISCO catalogue. Albanian National Inventory of Plant Genetic Resources in EURISCO (2% of Albanian Flora) includes 33 genera, 62 species and 2111 accession, where 54% of them are collected genetic materials. Detailed analysis shows that collected genetic material is the principal source of diversity of Albanian NI in EURISCO. Principal component analysis and comparisons of diversity indices show the descriptors as TAXON, CROPNAME, COLLGEM, GEODATA, NOGEODATA, COLLDATE > Y.2000, COLLDATA < Y.2000, COLLDATE > Y.2000, COLLCODE-GB, SAMPSTAT-100, COLLSRC-10, COLLSRC-20 and COLLSRC-40 present higher range of diversity and were the principal source of variation that contribute more than other on diversity degree of Albanian NI of PGR in EURISCO. Multivariate correlation analysis show very high positive correlation among COLLGEM and GEODATA, NO-GEODATA, COLLDATE, COLLCODE, SAMPSTAT-100, and COLLSRC-10 passport descriptors. There were higher relationships between collected genetic materials and GEODATA descriptors (latitude, longitude, elevation), COLLDATA, COLLCODE and COLLSRC.

Key words: National Inventory, passport descriptors, genetic diversity.

1. Introduction

Genetic diversity allows crops to evolve and adapt and is a major resource for plant breeders to use and meet the challenges in maintaining food security and environmental stability. Within the ongoing challenges posed by climate change on crop adaptation worldwide, the long-term conservation of plant genetic resources in genebanks plays a crucial role for the sustainable use of agro-biodiversity for present and future generations [14, 15].

Genebanks maintain collections of living material of genetic resources of crop plants and their wild relatives, with the aim of collecting, preserving, evaluating, and documenting this material, for utilization by users, such as breeders or researchers [29]. Hence, plant varieties stored in genebanks allow the agricultural communities access to genetic diversity to develop the most suitable and cost-effective crop variety for their specific needs (i.e. more reliable yields, resistance to pests and diseases, tolerance to abiotic stresses) in order to address climate change challenges and issues about food security and quality [15, 27, 9].

European genebanks play a major role in conservation of Plant Genetic Resources (PGR) as they hold about one quarter (about 1,725,315 accessions) of more than 7.4 million estimated accessions of *ex situ* Plant Genetic Resources for Food and Agriculture (PGRFA) currently maintained globally [15]. Therefore, accessibility to comprehensive crop collections is increasingly important [19, 29] and efforts have been made to disseminate information by creating web-based portals as easy links between users and most genebanks and, thus, make PGR available worldwide [6].

European Plant Genetic Resources Search Catalogue (EURISCO), a web-based catalogue (available on-line at http://EURISCO.ecpgr.org/) [11] of PGR of *ex situ* collections maintained in Europe, publicly launched in 2003, currently represents one of the largest directly searchable multi-crop databases in Europe, providing direct access to passport data on almost 1.1 million plant germplasm accessions of crop diversity related to 43 NIs and 332 holding institutes

Correspondence: Belul Gixhari; Agricultural University of Tirana, Albania; Email: gixharibelul@ubt.edu.al
(Accepted for publication 25 November 2013)
***ISSN**: 2218-2020, © Agricultural University of Tirana*

representing more than 5,500 genera and 36,000 species conserved in 43 European countries [11]. EURISCO currently defines standard data according to Multi Crop Passport Descriptor (MCPD) list [10, 16]. The standardization codes ensure a unique identification of germplasm accessions [7, 5] and facilitate data exchange between holding institutes as well as with other on-line sharing catalogues, created to collect PGR data information [7]. There are three web-sources (EURISCO, WIEWS, and ECCDBs) that gather information of germplasm accessions maintained in Europe. These web-sources have different policies, “focal points” and updating mechanisms. Therefore, data gathering mechanisms are not always synchronized and different updated information may have generated discrepancies between the web-sources contents [7].

Albania is very rich in plant biodiversity, including wild relatives of cultivated crops. Albanian Flora identified more than 3 250 species in the country (or 29.5%), from near 11 000 types of plants that are present in Europe [1, 2, 4, 17, 18, 28]. Only 62 species (2%) of them is included in the National Inventory (NI) of Albania in EURISCO catalogue. A comparative analysis between EURISCO and WIEWS and ECCDBs [11, 12, 13, 19, 20] show there were differences or “gaps” between data reported in the EURISCO [11], WIEWS [26], and ECCDBs [12] databases for several European countries. Albania has reported 3654 accession in WIEWS, 2111 accessions in EURISCO and 1494 accessions in ECCDBs database [19, 20].

At present the Albania Genebank (AGB) maintains ex situ near 3000 accessions of cultivated and wild plants and fruit tree species conserved in field collection. The majority of the accessions stored are represented by landraces and improved cultivars. To increase usage value of the plant genetic resources stored in *ex situ* status of conservation, it is necessary to have all information in a well-organized database and the database must be available for everyone interested. Good database of PGR leads to enhanced utilization of germplasm by farmers and ensure that "tomorrow's plant breeders will have today's genetic resources for use in their plant breeding programs [8]. The purpose of this study was a) the evaluation of genetic diversity of plant germplasm stored in Albania genebank, and b) the analysis of genetic diversity, the quality, and amount of data of Albania NI recorded in EURISCO.

2. Material and method

Diversity of NI in EURISCO is carried out using collection of data from AGB database and an on-line search data of Albanian PGR database downloaded in EURISCO website. The comparative analysis for number of individuals (accessions), genus, species per genus, was done based on EURISCO standard MCPD list [10, 16].

Passport descriptors used were: National Inventory code (NICODE), Institute code (INSTCODE), accession number (ACCENUMB), collecting number (COLLNUMB), collecting institute code (COLLCODE) (collected by genebank (GB), and collected by research institutes (RI), genus (GENUS), species (SPECIES/TAXON), common crop name (CROPNAME), accession name (ACCENAME), acquisition data (ACQDATE), location of collecting site (COLLSITE), latitude (LATITUDE) , longitude (LONGITUDE) and elevation (ELEVATION) of collecting site, accessions with or no geographic data (GEODATA) or (NO-GEODATA), collecting data of sample (COLLDATE), collected before year 2000 (< Y.2000), collected after year 2000 (>Y.2000), Biological status of sample (SAMPSTAT): (wild = 100, weedy = 200, landraces = 300, breeding material = 400, advanced cultivar = 500), collecting source (COLLSRC): (wild habitat = 10, cultivated habitat = 20, market = 30; RI or GB = 40), and type of germplasm storage (STORAGE) (seed collection = 10, field collection = 20).

Multivariate correlation and Principal Component Analysis (PCA) on correlation was used to identify similarity and distances between variables of passport data. The number of principal components (PC) to be retained in the analysis was determined using the minimum eigenvalue criterion [23].

Simpson’s; Shannon's; Brillouin and Alpha diversity indices assessed by passport descriptors were calculated based on number of genus, the respective proportions of each accession per genus and species per genus using SPSS 12 software [25]. The “gaps” proportion of each variable is classified into 5 degrees: “no gap” = ≥ 0%, low = from -1% to -19%, medium from -20% to -49%, high from -50% to -9%, and very high from -80% to -100% [19]. All statistics data were calculated employing the SAS JMP Statistical Discovery [24].

3. Results and discussion

Albanian NI of Plant Genetic Resources in EURISCO catalogue present 78.5% of total number of accessions stored in AGB, and it includes 2111 accessions, 33 genera, and 62 species. According to *biological origin status* of PGR from 2111 accession 46% (976 accessions) were breeding genetic materials (BREEDGEM) composed by breeding lines or advanced cultivars and 54% (1135 accessions) were collected genetic materials (COLLGEM) including wild species, semi-wild, landraces or old traditional cultivars (**Table 1).**

The mandatory fields NICODE, INSTCODE, ACCENUMB, and GENUS having the same values were not source of variability on the diversity of NI in EURISCO. Comparison of data show the higher differences of Albanian NI in EURISCO derived from COLLGEM that contribute 54% of total variability of NI in EURISCO. Detailed analysis shows the descriptors COLLCODE, COLLDATE, GEODATA, SAMPSTAT and COLLSRC were the variables with significant contributions on total variability of NI in EURISCO.

There were COLLSITEs with GEODATA, and with NOGEODATA two variables with high contribution on total variance of NI. Study results show that only 61% of accessions were recorded in database with GEODATA each collecting site, and 39% of accessions were with NOGEODATA, 48% (541 accessions) were collected before year 2000, and 52% of them were collected after year 2000.

Table 1. Diversity of Albania NI in EURISCO according to Passport descriptor data

Diversity of NI in EURISCO according to Passport data	*Accessions*	*Percentage*
Albania Genebank	2689	100%
NI in EURISCO	2111	78.5%
Ratio GENEBANK / NI in EURISCO (medium gap**)	578	21.5%
Breeding Genetic Material (BREEDGEM)	976	46%
Collected Genetic Material (COLLGEM)	1135	54%
Diversity of NI in EURISCO assessed by Collected Genetic Material Passport data		
COLLCODE-GB	431	38%
COLLCODE-RI	704	62%
CROPNAME (low gap*)	1132	98%
ACCENAME (medium gap**)	737	65%
COLLSITE	1031	90.8%
GEODATA	689	61%
NOGEODATA (medium gap**)	446	39%
Collected date of sample, before year 2000 (COLLDATE<Y.2000)	541	48%
Collected date of sample, after year 2000 (COLLDATE>Y.2000)	594	52%
Biological sample status (SAMPSTAT) wild = SAMPSTAT 100	392	35%
Biological sample status (SAMPSTAT) landraces = SAMPSTAT 300	743	65%
Collecting source (COLLSRC)	1135	100%
Collecting source (COLLSRC), wild habitat = COLLSRC 10	378	33%
Collecting source (COLLSRC), cultivated habitat = COLLSRC 20	335	30%
Collecting source (COLLSRC), RI or GB = COLLSRC 40	422	37%
Type of germplasm storage (STORAGE)	2111	100%
Type of germplasm storage, Seed collection (SEEDCOLL) = 10	1800	85%
Type of germplasm storage, Field collection (FIELDCOLL) = 20	311	15%

Low gap* = from -1% to -19%, medium gap** = from -20% to -49%.

The most part of accessions (736 accessions or 65%), have biological sample status '*landraces*', and were collected by RI on farms or cultivated habitats before year 2000, and 35% with biological sample status '*wild* ore *semi-wild'* were collected especially in wild habitats by GB after year 2000. PCA identified the variances of the principal components (PC) and the proportion of the total variance each factor accounts for, and three PC that account for 89.5% of the total variance are retained for further analysis. The fourth PC components with eigenvalue > 1.0 account for only 7.6% on the total variance is not retained in our study. Variance of 89.5% > 75% is acceptable for this kind of studies [3, 21, 22]. The percentages of total variation accounted for by each of the first three PCs are 45.7%, 28.7%, and 15.1%, respectively (Table 2). For PC1 there were TAXON, CROPNAME, COLLGEM, GEODATA, COLLDATA<Y.2000, COLLCODE-RI and COLLSRC-40 variables (descriptors) with significant weighting on PC1 variance. Variation in PC2 was mainly result of differences between COLLDATE>Y.2000, COLLCODE-GB, SAMPSTAT-100 and COLLSRC-10. In PC3 there were NOGEODATA, SAMPSTAT-300, and COLLSRC-20 variables that account for 15.1%, of the total of variation (Table 2).

Three-dimensional scaling for relationships of variables (descriptors) that accounts for the larger proportion of that accounts for the larger proportion of the total variance in PC1, PC2 and PC3 revealed by

PCA and cluster analyses show there were eight variables with more significant weighting for PC1 that account for 56.5% of PC1 variance and included into the first cluster group. For PC2 (with 28.7% of total variation), there were four variables with more significant weighting for PC2 that account for 38.1% of PC2 variance and included in the second cluster group (Figure 1).

Table 2. Matrix of eigenvalues and vectors of three PC results of NI in EURISCO

Principal Components / Factor Analysis

PC Number	Eigenvalue	% Variance	Cum Percent	ChiSquare	DF	Prob>ChiSq
1	**6.8560**	45.706	45.706	3565.61	100.508	<.0001*
2	**4.3051**	28.701	74.407	3328.13	98.56	<.0001*
3	**2.2646**	15.098	89.505	3068.42	91.512	<.0001*
4	**1.1364**	7.576	97.081	2795.47	81.043	<.0001
5	0.3047	2.031	99.112	2452.94	70.235	<.0001

Eigenevectors

Variables	***X***	***Y***	***Z***
TAXON	**0.33873**	-0.00406	-0.17077
CROPNAME	**0.34393**	-0.00298	-0.16358
BREEDGEM	0.23111	-0.07498	-0.29841
COLLGEM	**0.35923**	0.11016	0.10724
GEODATA	**0.32741**	0.21298	-0.02979
NOGEODATA	0.28718	-0.14541	**0.35022**
COLLDATE <Y.2000	**0.36200**	-0.08845	0.02611
COLLDATE >Y.2000	0.02228	**0.44217**	0.18551
COLLCODE-GB	0.03079	**0.47254**	0.09461
COLLCODE-RI	**0.36101**	-0.09764	0.06934
SAMPSTAT-100	0.03150	**0.47663**	0.01670
SAMPSTAT-300	0.04764	-0.10260	**0.55914**
COLLSRC-10	0.02516	**0.47600**	0.01833
COLLSRC-20	0.02708	-0.10320	**0.59171**
COLLSRC-40	**0.35694**	-0.06797	-0.08812

In bold eigenvalues > 1.00 and eigenvectors > 0.300

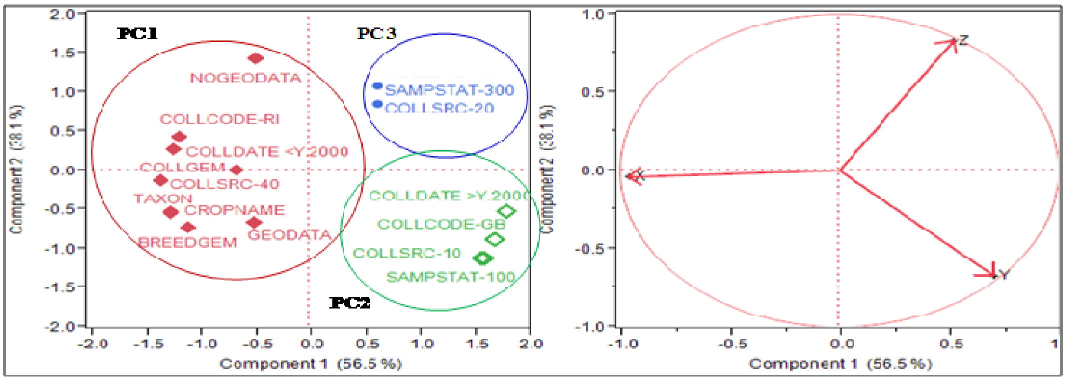

Figure 1. Relationships of descriptors that accounts for the larger proportion of variance in PC1, PC2, and PC3 revealed by PCA and cluster analyses

Diversity indices of NI in EURISCO

Analysis of diversity indices (Simpson 1-D, Shannon, Brillouin and Alpha) of ALB NI in EURISCO demonstrate the presence of high diversity level between data recorded for each descriptor. Simpson diversity index show presence of diversity among the variables analyzed. Simpson's Index of Diversity (1 – D) values range from 0.28 to 0.89. There were variables (descriptors) as COLLDATE>Y.2000, NOGEODATA, COLLGEM and COLLSCR-20 (with index value ≥ 0.85) that present higher diversity level. Shannon and Brillouin indices, with similar significantly correlated values, show the variables as COLLDATE>Y.2000, COOLGEM, TAXON, CROPNAME and NOGEODATA were the most important variables that contribute on the total variation of NI in EURISCO. Comparisons of means for each diversity indices show the descriptors as TAXONS, CROPNAME, COLLGEM, GEODATA,

NOGEODATA, COLLDATE>Y.2000, COLLCODE-GB, COLLSRC-20 and FILEDCOLL-20 were the principal source factors of variation that contribute more than other on diversity of NI in EURISCO (Table 4).

Table 4. Diversity indices of NI in EURISCO assessed by passport descriptors.

Indices	TAXON	CROPNAME	BREEDGEM	COLLGEM	GEODATA	NOGEODATA	COLLDATE<Y.2000	COLLDATE>Y.2000	COLLCODE-GB	COLLCODE-RI	SAMPSTAT 100	SAMPSTAT 300	COLLSRC 10	COLLSRC 20	COLLSRC 40	SEEDCOLL-10	FIELDCOLL-20
Individuals	2111	2069	976	1135	689	446	541	594	431	704	392	743	378	335	422	1800	311
Mean	64.0	62.7	29.6	34.4	20.9	13.5	16.4	18.0	13.1	21.3	11.9	22.5	11.5	10.2	12.8	100	20.7
SE	24.2	23.7	17.9	11.3	8.73	4.21	10.9	5.00	4.88	10.8	4.90	10.9	4.92	3.50	10.8	42.8	6.78
Taxa_S	33	33	6	31	16	19	7	26	13	20	11	21	9	19	3	18	15
Simpson_1-D	0.83	0.83	0.61	0.86	0.80	0.88	0.54	0.89	0.83	0.72	0.80	0.74	0.79	0.85	0.28	0.77	0.83
Shannon_H	2.32	2.32	1.14	2.54	1.99	2.36	1.12	2.58	2.05	1.88	1.85	1.96	1.74	2.29	0.50	1.86	2.11
Brillouin	2.28	2.28	1.13	2.48	1.94	2.28	1.09	2.50	1.99	1.83	1.79	1.90	1.69	2.19	0.49	1.84	2.02
Fisher_alpha	5.55	5.58	0.85	5.89	2.93	4.03	1.14	5.55	2.53	3.83	2.10	4.02	1.66	4.36	0.44	2.78	3.29

4. Conclusions:

Collected Genetic Materials is the principal source of diversity of Albanian NI in EURISCO. Principal component analysis and diversity indices show the descriptors as TAXON, CROPNAME, COLLGEM, GEODATA, COLLDATE > Y.2000, COLLDATA < Y.2000, COLLDATE > Y.2000, COLLCODE-GB, SAMPSTAT-100, COLLSRC-10, COLLSRC-20 and COLLSRC-40 were the principal source of variation that contribute more than other on diversity level of Albanian NI in EURISCO.

There were very high positive correlation among COLLGEM and GEODATA, NO-GEODATA, COLLDATE, COLLCODE, SAMPSTAT-100 and COLLSRC-10 (r range from 0.73 to 0.98). Study results demonstrate the important relationships between collected genetic materials and descriptors as GEODATA (latitude, longitude, elevation), COLLDATA, COLLCODE and COLLSRC., which means each sample or accession collected must have necessarily mandatory fields (NICODE, INSTCODE, ACCENUMB and GENUS) and the most important diversity variables as GEODATA, COLLDATA, COLLCODE and COLLSRC.

5. References

Journals citation:

1. Albanian Flora: **Edition of the Academy of Sciences of the Republic of Albania**,1988, **1**(4): 34-123.
2. Cadima J.F.C.L, Jolliffe I.T: **Variable selection and the interpretation of principal subspaces.** J. Agri. Biol. Environ. Statist.: 2001, **(6):** 62–79.
3. Dias S, Dulloo M.E, Arnaud E: **The role of EURISCO in promoting use of agricultural biodiversity.** 2001, (**2):** 1-34
4. Dulloo M.E, Hunter D, Borelli T: ***Ex situ* and *in situ* conservation of agricultural biodiversity: major advances and research needs**. Notulae Botanicae Horti AgrobotaniciCluj-Napoca. 2010, 38 (**2**): Special Issue 123 – 135.
5. Endresen D.T.F, Knüpffer H: **The Darwin Core extension for genebanks opens up new opportunities for sharing genebank datasets**. Biodiversity Informatics, 2012, (**8**): 12 – 29.
6. Engels J.M.M. Visser L: **A guide to effective management of germplasm collections**. IPGRI, Rome. 2003, (**3**): 32-44.
7. Esquinas-Alcazar J: **Protecting crop genetic diversity for food security: political, ethical and technical challenges**. Nature Reviews Genetics 2005, (**6**): 946-952.
8. Hammer K, Laghetti G, Perrino P, Xhuveli L: **Collecting in the Albanian mountains.** Plant genetic Resources Nwesletter. 1996, (**107**): 37-39.
9. Hausmann A, Dias S: **Comparative analysis between EURISCO and WIEWS and ECCDBs.** Bioversity International, Rome, 2013, (**2**): 24-76.
10. Haussmann B.I.G, Parzies H. K, Presterl T, Susˇic Z, Miedaner T: **Plant genetic resources in crop improvement.** Plant Genetic Resources. 2004**, 2**(1): 3 – 21.
11. Jolliffe I.T: **Principal Component Analysis,** Second edition, p.cm-Springer Series in Statistics. UK,2002, (**3**): p.143-180.

12. Jolliffe I.T: **Discarding variables in a principal component analysis II:** Real data. *Appl. Statist.*, 1973, (**22**): 21–31.
13. Kaiser H.F: **The application of electronic computers to factor analysis.** Educational and Psychological Measurement, 1960, (**20**), 141-151.
14. SAS: **JMP Statistical Discovery.** SAS Institute Inc. 2012, (**10**): 1-76.
15. Toledo A,Burlingame B: **Biodiversity and nutrition: A common path toward global food security and sustainable development**. Journal of Food Composition and Analysis, 2006, (**19**): 477 – 483.
16. Walters C: **Principles of preserving germplasm in gene banks**. In: Strategies for survival. (E. Guerrant, K. Havens and M. Maunder, eds.). Island press, Covelo, CA, USA. 2003, Pp. 113-138
17. Wright B.D: **Crop genetic resource policy: the role of *ex situ* genebanks.** The Australian Journal of Agricultural and Resource Economics, 1997, **41**(1): 81 – 115.