DOI: http://dx.doi.org/10.1590/1678-992X-2018-0091

ISSN 1678-992X

SCIENTIA AGRICOLA

Research Article

Old tools as new support for *on farm* conservation of different types of maize

Rafael Vidal[1], Natália Carolina de Almeida Silva[3], Juliana Bernardi Ogliari[2]*

[1]Universidad de La República/Facultad de Agronomía – Depto. de Biología Vegetal, Av. Garzón, 780 – 12900 – Montevideo – Uruguay.
[2]Universidade Federal de Santa Catarina/Centro de Ciências Agrárias – Depto. de Fitotecnia – Núcleo de Estudos em Agrobiodiversidade, Rod. Admar Gonzaga, 1346 – 88034-001 – Florianópolis, SC – Brasil.
[3]Universidade de Sao Paulo/ESALQ – Depto. de Genética, Av. Pádua Dias, 11 – 13418-900 – Piracicaba, SP – Brasil
*Corresponding author <juliana.bernardi@ufsc.br>

Edited by: Roberto Fritsche Neto

Received March 27, 2018
Accepted August 07, 2018

ABSTRACT: A germplasm collection should represent the diversity of the target species and the gene pools associated with it. However, it is critical to establish collection plans which ensure such representativeness. At times it is difficult to identify the best strategy for collecting domesticated species that are conserved *in situ/on farm*, since, in general, the magnitude of the diversity existing in a geographical area was hitherto unknown. The Diversity Census methodology was developed for previous diagnosis of the diversity of *Zea mays* subsp. *mays* L., conserved by farmers in two municipalities in the far western region of the state of Santa Catarina, southern Brazil. The Diversity Census database allowed for the identification of the best strategy to collect different types of maize landraces. Thus, tests were carried out using two methods described for Core Collections (Modified Random Sampling and Maximization) and a third statistical method for random sampling, stratified by farm area. The Maximization method enabled the capture of all the morphological variation of the traits evaluated in the Diversity Census from the smallest sample size. The relevance of this result is the feasibility of adapting the Core Collection strategy in order to plan more efficient expeditions to collect maize landraces conserved in microregions. Such planning allows for organizing the collection work efficiently, reducing costs, simplifying the work of characterization and helping to plan integrated strategies of *in situ/on fam* conservation.
Keywords: *Zea mays* L., genetic diversity, morphological characters, use values

Introduction

Appreciating landraces as genetic resources, identifying the processes of their loss and establishing complementary conservation strategies require collection and characterization of the diversity conserved *in situ/on farm*. Germplasm collection is a set of activities whose aim is to obtain living physical units, represented by samples that contain the genetic composition of the population of a given species of interest with reproductive ability (Dawson and Were, 1997).

Every collection is normally defined by three key elements: (a) area of coverage (*where to collect*?); (b) target species, its relatives and the populations of interest (*what to collect*?); and (c) sampling strategy (*which and how many units to collect*?). As regards the strategy, because there is no a priori information available in most cases, Marshall and Brown (1975) suggest the collection of 50 to 100 different plants per population, for the cultivated species, in order to include the greatest possible amount of areas, in locations with wide environmental variation. However, the difficulty of defining how many and which populations to include in the sampling continues still unresolved.

The distribution of diversity of landraces is not random and presents a framework along multiple axes – geographic, genetic (Brown, 1989) and cultural (Orozco-Ramírez et al., 2016; Leclerc and Coppens, 2012) – and systematic samplings, considering all aspects of diversity, could thus improve the strategy effectiveness, in comparison with random samplings.

A Core Collection (CC) is a small sample of the original collection, which is included in a spectrum of the genetic variability. The design of a CC seeks to ensure the retention of genes or gene combinations that are present in an *ex situ* collection. This research study proposes the CC tool developed for *ex situ* collections to collect the diversity conserved *in situ/on farm*, once the potential diversity of collection is known and, after that, information is used to guide the activity of collecting the local germplasm. Defining "which populations to conserve" and "to characterize" is crucial when working with limited financial and human resources. Clearly, establishing the smallest number of samples that best represents the diversity of landraces conserved *in situ/on farm* is a key component in the design of more efficient strategies for collection and subsequent conservation of *ex situ* collections. Thus, with the aim of identifying the best strategy for collection of maize landraces conserved *in situ/on farm*, in a microregion in southern Brazil, three sampling methods were compared.

Materials and Methods

Sample universe: *in situ/on farm* Base Collection

The *in situ/on farm* Base Collection (BC-ISOF), as defined for the present study, included 1,513 landraces of *Zea mays* subsp. *mays* L. grown within a geographical area of 558 km^2, bounded by the municipalities of Anchieta (latitude 26°53′ South and longitude 53°33′ West; altitude 745 m.a.s.l.) and Guaraciaba (latitude 26°35′ South and longitude 53°31′ West; altitude 720 m.a.s.l.) (Miranda, 2005), in the far western region of the state of Santa Catarina (FWSC), Brazil (Figure 1A-D). The region is located in the Uruguay River Basin, and native vegetation belongs to the Atlantic Forest biome, one of the 25 biodiversity 'hotspots' in the world (Myers et al., 2000). The BC-ISOF was defined a priori as result of the Diver-

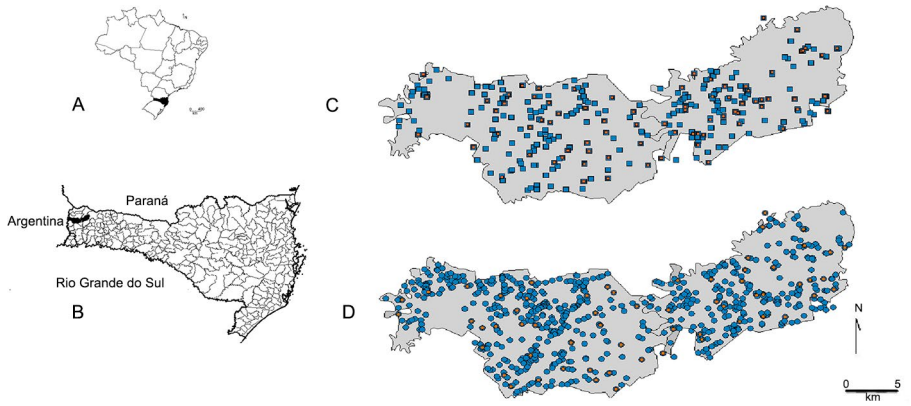

Figure 1 – A) Location of the state of Santa Catarina, Brazil. B) Location of the municipalities of Anchieta and Guaraciaba, in the far western region of the state of Santa Catarina. C) and D) Distribution of *in situ/on farm* Base Collection (in blue color) and Core Collection (in orange color). Squares represent common maize landraces (C) and circles represent popcorn landraces (D). Map contour constructed from spatial data retrieved from http://www.diva-gis.org/gdata.

sity Census methodology. Through interviewing peasants, this methodology allowed for identifying, mapping and characterizing all the landraces of *Zea mays* subsp. *mays* L. Thus, the BC-ISOF database included the following information: identification of the landraces (*name*, *cultivation time* and *risk factors of loss*), morphological characteristics of the grain (*type of endosperm*, *size* and *color*), geographical location (*municipality*, *community*, *longitude* and *latitude*), and *use values* and *conservation*, organized into 13 categories (agronomic, gastronomic, animal feed, adaptive, aesthetic, economic, health, cultural, ornamental, crafts, conservation of diversity, nutritional and medicinal) defined by Costa et al. (2016).

Considering the richness of BC-ISOF by type of endosperm, the landraces were distributed into two subgroups: Common Maize (CM) and Popcorn (PM). The first subgroup corresponded to all landraces with dent, semident, flint and semi-flint grains (endosperm types), and whose maize types were common, flour or sweet. The second subgroup corresponded to maize type characterized as popcorn. This division was based on the names and use values listed by farmers during the Diversity Census (Costa et al., 2016). In the case of PM, less diverse in terms of use values, they were excluded from the *agronomic* and *animal nutrition* use classes. The trait *name* was excluded from the PM subgroup, since 30 % of the landraces did not have names and among those which had names, 68 % were associated with grain color and therefore the addition of this characteristic would be redundant information. In the CM subgroup, 18 % of the landraces did not have names while 21 % of names associated with the trait grain color and 43 low-frequency names (< 1 % of landraces) were identified. Low-frequency names are associated with adjectives for use values or specific uses of farmers (e.g., sweet, big, old) or the origin of the variety (e.g. "*Élcio*", "*Festa*", "*Camponesas*"). These low-frequency landraces would not be used in a classification that only considered color and type of grain, and thus the trait *name* has been included for subgroup CM.

The traits associated with location of the landraces (*municipalities*; *communities*; *latitude* and *longitude*) are complementary and allow for covering the different networks and forms of communication between farmers, which may be associated with: (a) commercial networks frequented by farmers, which are, in general, centered in the municipalities, (b) social networks and the neighborhood, established in the communities, and (c) the flow of pollen among local populations of maize, according to geographical distance (*latitude* and *longitude*). Therefore, the three traits associated with the location of landraces were included in this analysis.

The field studies were carried out on private lands and the owners of the lands gave their individual permission to conduct the studies on these sites. Written consent was given by local organizations representing the farmers, and individually by 1,688 interviewed farmers, as prescribed by the Law of Biodiversity (Law 13,123 of 20 May 2015) (BRASIL, 2015) in compliance with ethical standards. The database obtained from the Census of Diversity was analyzed anonymously.

Sampling procedures

The present study tested three sampling strategies to guide the work of collection of landraces and search for the organization of a germplasm bank at a university. Each of the subgroups of landraces (Figure 2) was tested as follows: (a) stratified random sampling (E), considering the area of each farm and making a random sampling among the farms of a same stratum of area; adaptations of the CC sampling, based on the (b) Maximization (M) and (c) Modified Random (R) methods.

In the case of stratified random sampling (E), sample size was defined by the optimal sharing strategy (Neyman, 1938), considering the total number of farm-

ers (1,688) who conserve the CM and PM landraces. The area class was used as a stratification variable, because it is believed that conservation and management strategies employed by families could vary according to their socioeconomic level. Data on size of the areas was collected from the Diversity Census. Sample size (Table 1) was composed of 141 CM farmers and 244 PM farmers; the sample was determined using an error margin of 5 %. After the sample size had been defined, the selection of farmers within each of the strata was established at random, by means of a random draw, carried out with the aid of the SPSS 2.2 software (Statistical Package for the Social Sciences, Version 18.0). As the variable of stratification corresponded to property, all the landraces of each farm were included in the sampling.

The Maximization strategy (M) proposed by Kim et al. (2007) has two stages; during the first stage, all variables are transformed into nominal ones and during the second stage, CC is sampled. For the first stage, the qualitative variables do not need to be transformed and all classes are considered; for continuous variables, the values are classified into *k* classes, according to the rule of Sturges (1926), where *N* is equal to the number of accessions within the variable. The sampling of CC is generated by a sequence of instructions (algorithm) that can be repeated (iterations) seeking the shortest path to maximize the number of classes detected (heuristic).

The Modified Random sampling method (R) was based on the same variables of the M strategy. The difference is that one accession from each class is chosen at random and, once all the classes of the variable are represented, the procedure is repeated for all the variables until all the classes of the base set have been completed. In cases where the class has one accession only, it is sampled directly. Samplings M and R were determined with the aid of the Power Core software application (Kim et al., 2007), designed to develop CC.

The variables considered for sampling were organized according to subgroups (Table 2). The quantitative variables *time*, *longitude* and *latitude* were divided into 10, 11 and 11 classes for CM, and 11, 12 and 12 classes for PM, respectively, according to Sturges' rule (1926). The variable *grain size* of PM was divided into three classes (large, medium and small), according to farmers' responses; 51 *use value* subcategories were identified for CM and 38 for PM (Table 3). Each of the *use value* subcategories was considered as binary (presence/absence), since a single variety may have more than one use value, simultaneously.

Risk of loss was estimated by the Participatory Four-Cell Analysis method (Oyarzun et al., 2013), a tool which aims to understand the dynamics of landraces, through the analysis of the number of farming families that grow them, and the area occupied. The authors of this method understand that few families and reduced area increase the risk of loss of a landrace. To adapt the Four-Cell Analysis tool to data from the Diversity Census, the frequency of the landraces and their distribution in the municipalities were used. The landraces of the same name with frequency (number of mentions) below 5 % (considered as rare) and unique to a single municipality were included in group IV and considered to be at high risk of loss;

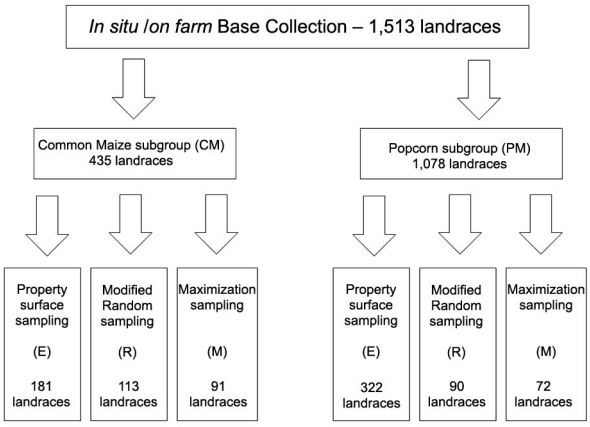

Figure 2 – *In situ/on farm* Base Collection, subgroups (Common and Popcorn) and sampling strategies for maize landraces.

Table 1 – Number of farmers by area class obtained from the Diversity Census and optimum sampling by strata for Common Maize and Popcorn.

Strata (in hectares)	Common Maize subgroup (CM)			Popcorn subgroup (PM)		
	Diversity Census	Sampling	%	Diversity Census	Sampling	%
≤ 5.0	54	22	41	94	33	35
5.1 a 10.0	63	26	41	140	48	34
10.1 a 15.0	74	34	46	175	61	35
15.1 a 20.0	55	25	45	100	35	35
20.1 a 30.0	43	21	49	108	32	30
≥ 30.1	28	13	46	99	35	35
Total	317	141	44	716	244	34

Table 2 – Base Collection Variables used in Random (R) and Maximization (M) sampling and number of classes for Common Maize and Popcorn.

Characteristic	Variable	Type	Classes	
			Common Maize	Popcorn
Location	Municipalities	Nominal	2	2
	Communities	Nominal	63	69
	Latitude	Nominal	11	12
	Longitude	Nominal	11	12
Grain	Color	Nominal	6	5
	Type	Nominal	3	3
	Size	Nominal	0	3
Identification	Name	Nominal	56	0
	Time	Nominal	10	11
	Loss risk	Nominal	4	4
Use Value	Subcategories	Binary	2	2

Table 3 – Categories and subcategories of use values used as variables for Common Maize and Popcorn (PM).

Category	Subcategories Common Maize	Subcategories Popcorn
Agronomics	Productivity	Productivity
	Early	Early
	Easy to thresh	Easy to thresh
	Resistance to lodging	Resistance to lodging
	Late planting	Late planting
	Casing	Prolificacy
	Thin cob	Fast dry
	Rooting	-
	Easy to meal	-
	Hard grain	-
	Tall plant	-
	Low plant	-
	Medium plant	-
Gastronomic	Cooking point	-
	Flavor	Flavor
	Softness	Softness
	Sweet	Sweet
	Flour	Crunchy
	Green corn	*Pé de Moleque*
	Pamonha	Dry
	Cooked corn meal	No shell
	Conserve	Volume
	Porridge	Pops well
	Bread	Small
	-	White
	-	Others
Adaptive	Diseases resistance	Diseases resistance
	Weevil resistance	Weevil resistance
	Pest resistance	Amplitude of adaptation
	Drought resistance	Resistance to contamination
	Rain resistance	-
	Early	-
	Late	-
	Amplitude of adaptation	-
Animal feed	Split grain	-
	Silage	-
Cultural	Tradition	Tradition
	Recreation	Recreation
	Affective	Affective
Aesthetics	Ear	Ear
	Grain	Grain
	Plant	Plant
Economic	Seeds autonomy	Seeds autonomy
	Costs	Costs
	Seed Sale	Seed Sale
Health	Pesticide free	Pesticide free
	Transgenic free	Transgenic free
	Healthy food	Healthy Food
Conservation	Conservation of diversity	Conservation of diversity
Nutritional	Nutritional	Nutritional
Ornamental	Ornamental	Ornamental
Hand craft	Hand craft	-
Medicinal	Medicinal	-

the landraces with frequencies below 5 %, present in two municipalities (Group III), and the landraces with frequencies above or equal to 5 %, exclusive of a municipality (Group II), were considered to have average risk of loss; the landraces with frequencies higher than or equal to 5 % and present in two municipalities were included in Group I and were considered to be at low risk of loss.

Validation of sampling strategies

The richness (S) of each of the samplings (E, R and M) was estimated by counting the total number of classes of each variable. The differences in the values of S between the three strategies (E, R and M) were tested by the Chi square (χ^2) test.

The diversity of the different samplings was estimated by means of the Gini-Simpson indexes (Simpson, 1949) and Shannon (Shannon, 1948). The Gini-Simpson (H_{GS}) index is given by $H_{GS} = 1 - \sum_i p_i^2$, with p_i equal to the frequency of the nth class of the variable or the number of accessions of each class. Shannon's Index (H') is given by $H' = -\sum_i p_i ln p_i$ (Shannon, 1948), where p_i is also the frequency of the nth class. For the Gini-Simpson and Shannon indexes, the significance of the differences between the three sampling methods (E, R, and M), in relation to BC-ISOF, was estimated by Bootstraping, with 9,999 repetitions and a 95 % confidence interval.

To estimate the representativeness of the samplings, a comparison was made between the classes of variables of each sample in relation to BC-ISOF, by determining the percentages of amplitudes retained in the samplings and the characteristics with low representativeness, according to the formula proposed by Diwan et al. (1995), given by

$$\%RR = \frac{\sum_i^{R_{CAIJ}} / R_{CB_{IJ}}}{t} \times 100$$

where R_{CA} represents the i classes of j variables in the sampling (E, R or M), R_{CB} the i classes of j variables in BC-ISOF and t the total number of variables compared. According to the criteria proposed by Diwan et al. (1995), the samplings are considered as representative, if the percentage of retention is greater than 70 % for at least 70 % of the variables and a maximum of 30 % of traits of the samples are significantly different ($p < 0.05$) than BC-ISOF. For Hu et al. (2000), the average value of %RR should be higher for at least 80 % of the traits and not more than 20 % of the variables can contain significantly different means ($p < 0.05$) than the averages for BC-ISOF. To assess significant differences of samplings E, R and M with BC-ISOF, the frequencies of the variables in use were compared with the χ^2 test.

The H' and H_{GS} indexes were estimated with the aid of the PAST software application (Paleontological statistics software package for education and data analysis version 3.16) to identify representativeness and the χ^2 tests were performed by R Core Team models (2018).

Results

Table 4 shows the results by the three sampling methods and maize types. For the subgroup CM, sampling E consisted of 181 landraces, which corresponded to 42 % of the whole subgroup (landraces with endosperm dent, semi-dent, flint, semi-flint, whose maize types are common, flour and sweet), while for PM, the sampling had 322 landraces, equivalent to 31 % of this subgroup. Through the R method, 113 landraces in subgroup CM (26 %) and 90 landraces in subgroup PM (8 %) were sampled, and using the M method, samplings of 91 landraces of CM (21 %) and 72 landraces of the subgroup PM (6 %) were identified.

The χ^2 test for richness (*S*) showed significant differences between sampling E and subgroup CM (Table 5). For PM, the richness of none of the samplings was significantly different from the subgroup (Table 6). The average value of the H' index estimated for subgroup CM was 1.88, while for samplings E, R and M, the mean values of H' were 1.83, 1.99 and 2.02, respectively. For subgroup PM, the mean values of H' were 1.42, and 1.41, 1.50 and 1.67 for sampling E, R and M, respectively. For the diversity index H_{GS}, the mean values found for subgroup CM, E, R and M were 0.57; 0.57; 0.59 and 0.60, respectively, while for subgroup PM, the values of the index H_{GS} were 0.62, 0.61 for sampling E and R, and 0.70 for sampling M. For both indexes (H' and H_{GS}), the values were higher for sampling M. For PM, the H' and H_{GS} indexes of sampling R were significantly different from subgroup by Bootstrapping, with respect to the variables *community* and *risk of loss*. Sampling M was also significantly higher than subgroup, for the H' index, estimated for the variables *community*, *grain color*, *cultivation time* and *risk of loss*, while the H_{GS} indexes, in the same sampling (M), were higher and significantly different from PM for the variables *community*, *cultivation time* and *risk of loss*.

The average of %RR for CM (Table 7) was 89 %, for sampling E, and 99 % for both samplings R and M. For subgroup PM (Table 8), the mean of %RR was 96 %, for sampling E, and they were 100 % for sampling R and M. Analyzed individually, the values of %RR of the study variables ranged from 52 to 100 % for CM and from 78 to 100 % for PM, taking into account all sampling methods. Sample R of subgroup CM (Table 9) was significantly different by the χ^2 test for the variable *risk of loss*, which represents 12 % of the variables.

For PM (Table 10), sampling R had significant differences by the χ^2 test for the variable *use value*, as compared with the entire subgroup, while sampling E showed significant difference for the variable *risk of loss*, and sampling M for the variables *grain type*, *use values* and *risk of loss*, which represent 37 % of the variables.

Table 4 – Number of sampling landraces by the Strata, Modified Random and Maximization methods for the subgroups Common Maize (CM), Popcorn (PM) and percentage of the number of landraces of each subgroup.

Sampling	CM		PM	
Method	Number	%	Number	%
Strata	181	42	332	31
Modified Random	113	26	90	8
Maximization	91	21	72	7
All	435	100	1078	100

Discussion

The collection of landraces is frequently a "forgotten" theme, because it is believed that there are enough landraces conserved *ex situ*, and thus no new collection activities are required. The diversity in the farmers' fields is not static, so much that the landraces are continuously adapting to changes (Camacho-Villa et al., 2005; Mercer and Perales, 2010) promoted by biotic and abiotic factors. Thus, one of the goals of collection in the context of *in situ*/*on farm* conservation is to update the *ex situ* collections. The collection of landraces can provide valuable contributions after being characterized, because they can enhance the local germplasm and encourage the development of public policies that ensure continued conservation of landraces by farmers, as has been the practice

Table 5 – Richness (*S*) and diversity according to the Shannon (*H'*) and Gini-Simpsom index (H_{GS}) of each variable for the Common Maize (CM) subgroup and Strata (E), Modified Random (R) Maximization (M) sampling methods.

Sampling	CM			E			R			M		
Variable	*S*	*H'*	H_{GS}	S	*H'*	H_{GS}	S	*H'*	H_{GS}	S	*H'*	H_{GS}
Municipality	2	0.69	0.50	2	0.68	0.49	2	0.69	0.50	2	0.69	0.50
Community	63	3.73	0.97	54	3.65	0.97	63	3.91	0.97	63	3.97*	0.98
Grain color	6	1.26	0.64	6	1.23	0.61	6	1.29	0.63	6	1.29	0.64
Grain type	3	0.73	0.41	3	0.74	0.42	3	0.74	0.41	3	0.87	0.50
Name	58	3.14	0.93	30	2.97	0.93	56	3.65*	0.96*	56	3.69*	0.96*
Time	8	0.99	0.45	4	0.91	0.43	8	1.05	0.49	8	1.08	0.51
Loss risk	4	1.10	0.63	4	1.10	0.64	4	1.11	0.66	4	1.15	0.67
Use value	52	3.39	0.06	45	3.36	0.06	52	3.54	0.06	52	3.47	0.07
Average	-	1.88	0.57	-	1.83	0.57	-	1.93	0.59	-	2.02	0.60
X^2				15.8			0.07			0.07		

*Differences are significant at 5 % with an CI of 95 % and 9,999 Bootstraps. Critical value of $X^2_{(7;\ 0.05)} = 14.07$.

Table 6 – Richness (S) and diversity according to the Shannon (*H'*) and Gini-Simpsom index (H_{GS}) of each variable for the Popcorn (PM) subgroup and Strata (E), Modified Random (R) Maximization (M) sampling method.

Sampling	PM			E			R			M		
Variable	S	*H'*	H_{GS}	S	*H'*	H_{GS}	S	*H'*	H_{GS}	*S*	H'	H_{GS}
Municipality	2	0.68	0.49	2	0.69	0.49	2	0.68	0.48	2	0.68	0.49
Community	69	3.91	0.97	64	3.82	0.97	69	4.14*	0.98*	69	4.22	0.99*
Grain color	5	1.35	0.71	5	1.37	0.72	5	1.24	0.67	5	1.50	0.77
Grain type	3	0.79	0.52	3	0.79	0.52	3	0.77	0.51	3	0.90	0.56
Grain size	3	0.87	0.55	3	0.88	0.56	3	0.82	0.53	3	0.92	0.57
Time	9	1.26	0.61	9	1.25	0.61	9	1.38	0.62	9	1.34	0.72*
Use value	36	1.96	0.77	28	1.92	0.71	36	2.88	0.84	36	2.88	0.91
Loss risk	3	0.56	0.30	3	0.53	0.29	3	0.94*	0.23*	3	0.94	0.56*
Average	-	1.42	0.62	-	1.41	0.61	-	1.50	0.61	-	1.67	0.70
X^2				2.14			0			0		

*Differences are significant at 5 % with an CI of 95 % and 9,999 Bootstraps. Critical value of $X^2_{(7;\ 0.05)}$ = 14.07.

Table 7 – Amplitude of the variables in the Common Maize (CM) subgroup and percentages of retained amplitudes (% RR) of the sampling by Strata (E) Modified Random (R) and Maximization (M).

Variable	Amplitude CM	%RR		
		E	R	M
Municipality	2	100	100	100
Community	63	86	100	100
Grain color	6	100	100	100
Grain type	3	100	100	100
Name	58	52	97	97
Time	8	88	100	100
Use value	52	87	100	100
Loss risk	4	100	100	100
Average	-	89	99	99

Table 8 – Amplitude of the variables in the Popcorn (PM) subgroup and percentages of retained amplitudes (%RR) of the sampling by Strata (E) Modified Random (R) and Maximization (M) methods.

Variable	Amplitude PM	%RR		
		E	R	M
Municipality	2	100	100	100
Community	69	93	100	100
Grain color	6	100	100	100
Grain type	3	100	100	100
Grain size	56	100	100	100
Time	9	100	100	100
Use value	36	78	100	100
Loss risk	3	100	100	100
Average	-	96	100	100

Table 9 – *p* values for χ^2 test between the Common Maize subgroup class frequencies distributions and Modified Random (R) Strata (E) and Maximization (M) methods.

	Degrees of freedom	R	E	M
Municipality	1	0.848	0.421	0.894
Community	62	0.998	0.999	0.999
Grain color	5	0.726	0.921	0.625
Grain type	2	0.944	0.861	0.253
Name	55	0.872	0.999	0.536
Time	7	0.809	0.999	0.999
Use value	50	0.997	0.999	0.999
Loss risk	3	0.015	0.616	0.069

Table 10 – *p* values for χ^2 test between the Popcorn subgroup class frequencies distributions and Modified Random (R) Strata (E) and Maximization (M) methods.

	Degrees of freedom	R	E	M
Municipality	1	0.85	0.812	0.99
Community	68	0.82	0.961	0.76
Grain color	4	0.22	0.216	0.84
Grain type	2	0.82	0.651	$2.00E^{-4}$
Grain size	4	0.87	0.989	0.70
Time	7	0.16	0.909	0.04
Use values	35	0.01	0.999	$5.27E^{-7}$
Loss risk	2	0.28	$3.29E^{-13}$	$3.63E^{-3}$

for thousands of years. In this context, the role of *ex situ* conservation should be complementary and integrated into the systems of local germplasm conservation, and there should never be any intention of replacing them. A complementary *ex situ* strategy needs to return every time to collect landraces, and in this respect, the size of each sampling is critical.

Thus, the best sampling strategy should represent the greatest diversity with the least number of individuals. We tested three different sampling strategies for two different maize types and identified samples differing between 72 and 322 landraces for popcorn and 91 and 181 for common maize. This difference can represent considerably less cost and work in the collection process, especially when looking to maximize the richness and diversity in the collection.

Richness (*S*) is a more sensitive parameter for low frequencies of classes (Jost et al., 2010). Our results are indicative of a loss of representativeness of classes in sampling E for CM.

Because the *H'* and H_{GS} indexes are of first and second order, they allow for the evaluation of both common

and very common traits, respectively (Hill, 1973). The H' index minimizes redundancies and the samples keep the diversity of BC-ISOF. In a smaller set of individuals, increases are desirable in the values of diversity indexes of the samples, compared with the value of BC-ISOF (Odong et al., 2013).

For the H' and H_{GS} indexes, there were significantly higher differences between the samples R and M, in relation to subgroup CM, for the variable *name*. Between sample M and subgroup CM, there was also significantly higher difference of index H' for the variable *community*. The H' and H_{GS} values estimated for sample M, its significance compared to subgroup CM and the smaller sample suggest that this method M better represents the diversity of landraces of CM, because of a reduction in redundancies.

For the PM landraces, the H' index estimated for sample M and its significance as regards this subgroup suggests that the M sample best represents the diversity and richness of *grain colors*, *cultivation time on farm* and *risks of loss* throughout the geographic area. In fact, grain color has been an attribute used by farmers in the region to recognize the majority of these landraces (Silva et al., 2017).

These results are consistent with the work of Li et al. (2004), who assessed the diversity of 14 characteristics of the maize core collection from a germplasm bank in China. For all the characteristics, the values of H' were significantly higher than those in the base collection. Mahajan et al. (2007) have identified 14 of the 15 descriptors of a core collection of *Sesamum indicum* L. with higher H' index than the original collection. For the evaluation of the core collection of *Pennisetum glaucum*, Bhattacharjee et al. (2007) compared the H' indexes of 11 agronomic descriptors, six of which were higher in the core collection.

A sampling strategy is considered efficient when it retains, on average, at least 80 % of the amplitude of the BC (Frankel and Brown, 1984). Considering the criterium proposed by Diwan et al. (1995), which establishes 70 % as the minimum threshold of representativeness of the variables in the sample, for both subgroups, only the characteristic *name* of sampling E in subgroup MC presented a lower %RR value (52 %). The R and M methods were remarkably superior as regards retention of the amplitude of all variables. By comparison, Malosetti and Abadie (2001), evaluating different sampling strategies (random, constant, proportional and logarithmic) for a core collection of 10 % of the *ex situ* maize collection from Uruguay, found values of RR% between 69 and 91 % for 17 phenotypic descriptors.

According to Wang et al. (2013), the values of %RR are affected by the sample size and the number of classes and, therefore, variables with the maximum number of classes in certain strategies require larger samples. The value below the critical limit found for sampling E in the present study indicated that sample size was insufficient for an adequate representation of the feature *name*. The 12 % percentage of the variables of sample R of CM containing significantly different means ($p < 0.05$) from the averages for the subgroup is lower than the critical levels of 30 % and 20 % proposed by Diwan et al. (1995) and Hu et al. (2000), respectively. Therefore, all the samplings performed for subgroup CM can be considered as representative.

Sampling M from subgroup PM, with 37 % of the variables with significant differences, surpasses the limits of 30 % and 20 % significance established by Diwan et al. (1995) and Hu et al. (2000), respectively. The significant differences in sampling M, compared with the entire subgroup PM, may be explained by an increase in low-frequency landraces. The landraces with *intermediate* grain type, in sampling M, have a 7 % frequency, while in samples R and E landraces with *intermediate* grain type have a 2 % frequency. For the characteristic *use values*, of 37 total classes, 32 had higher frequencies in sample M, with averages of 2 %, in relation to PM subgroup (1 %), while in the samples E and R the mean frequencies were 1 % and 1 %, respectively. For the variable *risk of loss*, group IV (rare landraces, unique to a single municipality) had a frequency of 4 % in the subgroup, 2 % in sampling E, 5 % in sampling R, and, in the case of sampling M, frequency increased to 14 %.

In order to verify whether the increase in low-frequency classes influenced the representativeness of sampling, only high-frequency classes were evaluated by χ^2 (data not shown) and no significant differences were identified in any of the cases. These results confirm that differences in sampling M were explained by an increase in diversity.

Germplasm collections are usually planned to prioritize geographical axes with information on weather, altitude, relief and soil type, which is effective in large areas (Charmet and Balfourier, 1995). Based on this perspective, microregions with reduced environmental variation are represented by a few collection points only, and, in such cases, a large part of the diversity of local landraces may not be captured. The FWSC microregion has a humid moderate subtropical climate with a variation from 1800 to 2200 mm of rainfall (Fick and Hijmans, 2017) and a difference of 500 m in altitude between high and low areas, according to DIVA-GIS V 7.5. This relatively small variation in the climate of the more distant geographic areas may account for the almost zero representation of local landraces in this area, in *ex situ* collections in Brazil and also the shortage of collection expeditions. Recent studies have demonstrated the gaps that exist in the collections and the importance of studying the dynamics of local landraces on a scale of microregions (Perales et al., 2003; Pressoir and Berthaud, 2004; Latournerie et al., 2006; Bracco et al., 2012).

Article 9 of the Convention on Biological Diversity of 1992 establishes that *ex situ* conservation is intended to complement *in situ* conservation. However, most efforts continue to be intended for *ex situ* conservation (Nilsen et al., 2014). Less than one-third of the accessions of 30 species conserved *ex situ* in germplasm banks around the

world corresponds to landraces or old improved cultivars (Hammer et al., 2003). The latest collections of landraces of maize, in Brazil, were carried out in 1980 (Silva et al., 2017). Nowadays, the collection of maize from the germplasm bank of Embrapa (Empresa Brasileira de Pesquisa Agropecuária) has only 12 accessions of maize landraces from the FWSC (EMBRAPA, 2018). In this long period of cultivation in the region, many landraces of this species may have developed new variations and adaptations not covered in previous collections. In the face of shortcomings of *ex situ* strategies, *in situ/on farm* conservation is a dynamic solution that ensures continuous adaptation of landraces to changes in the environment, and it is based on the human and biological components of the ecosystem (Galluzzi et al., 2010).

Generally, in *ex situ* collections, information about accessions is quite limited or inaccurate. Passport information rarely includes characteristics described by farmers or refers to the ecological conditions that are the source of the material (Williams, 1989). Our results showed that the landraces conserved *in situ/on farm* in microregions have ethnobotanical information. The variables *names* or *use values* by means of a Diversity Census, and association of them with sampling maximization strategy (M), enables the inclusion of the ethnobotanical axis of diversity in the collections, as well as significantly reducing collection size and establishing collection routes.

Much more than representativeness of geographical extent, morphological attributes, identity of landraces and use values, the association of Diversity Census and the M sampling allow for incorporating the ethnobotanical axis into collections of the microregions, firstly by accumulating the highest richness and diversity of *names* and *use values* of the samplings and, secondly, because the information is available at the beginning of the collection.

The M method proposed in this research, as a sampling strategy of maize landraces, can also be used to support the development of community seed banks, to the extent that farmers' use values can be included among the selection variables, in order to apply this tool and thus facilitate the identification of priority landraces for conservation. Finally, the small and representative collections facilitate the characterization of landraces of the microregion, which favors the development and use of relevant landraces for *in situ/on farm* conservation. The sampling maximization strategy also solves one of the greatest concerns of *ex situ* collections, which is the significant increase in germplasm collection, more than characterization studies and improvements in the infrastructure of *ex situ* conservation.

Future studies should consider other crops, reproductive systems, regions and cultural contexts to verify whether the M sampling continues to be the best strategy with respect to the representativeness of diversity and number of landraces. Clearly, in this study, the M method established the smallest number of samples that best represents the diversity of maize landraces conserved *in situ/on farm* in FWSC. The design of a more efficient strategy for collection and subsequent conservation of *ex situ* collections in the gene bank at a university will be a key component within a participatory and integrated approach in maize conservation and use. The minimization of genetic erosion in the field, the evolutionary dynamics of diversity conserved *in situ/on farm* renewing the *ex situ* collections and the enrichment and improvement of local germplasm are amongst the benefits accruing from these two complementary approaches of conservation.

Acknowledgments

The authors would like to thank farmers and their organizations, in Anchieta-SC and Guaraciaba-SC, for the information provided about the germplasm. This research was supported by the Fundação de Amparo à Pesquisa e Inovação do Estado de Santa Catarina (FAPESC) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), whereas the scholarships were granted to Natália Carolina de Almeida Silva and Rafael Vidal by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

Authors' Contributions

Project: Ogliari, J.B. Conceptualization: Ogliari, J.B.; Vidal, R. Data acquisition: Silva, N.C.A. Data analysis: Vidal, R. Design of Methodology: Vidal, R.; Silva, N.C.A.; Ogliari, J.B. Writing and editing: Vidal, R.; Ogliari, J.B.; Silva, N.C.A.

References

Bhattacharjee, R.; Khairwal, I.S.; Bramel, P.J.; Reddy, K.N. 2007. Establishment of a pearl millet [*Pennisetum glaucum* (L.) R. Br.] core collection based on geographical distribution and quantitative traits. Euphytica 155: 35-45.

Bracco, M.; Lia, V.V.; Hernández, J.C.; Poggio, L.; Gottlieb, A.M. 2012. Genetic diversity of maize landraces from lowland and highland agro-ecosystems of southern South America: implications for the conservation of native resources. Annals of Applied Biology 160: 308-321.

BRASIL. 2015. Presidency of the Republic Civil House = Presidência da República Casa Civil Available at: http://www.planalto.gov.br/CCIVIL_03/_Ato2015-2018/2015/Lei/L13123.htm [Accessed June 21, 2018] (in Portuguese)

Brown, A.H.D. 1989. The case for core collections. p. 136-156. In: Brown, A.H.D.; Frankel, O.H.; Marshall, D.R.; Williams, J.T., eds. The use of plant genetic resources. Cambridge University Press, Cambridge, UK.

Camacho-Villa, T.C.; Maxted, N.; Scholten, M.; Ford-Lloyd, B. 2005. Defining and identifying crop landraces. Plant Genetic Resources 3: 373-384.

Charmet, G.; Balfourier, F. 1995. The use of geostatistics for sampling a core collection of perennial ryegrass populations. Genetic Resources and Crop Evolution 42: 303-309.

Costa, F.M.; Silva, N.C.A.; Ogliari, J.B. 2016. Maize diversity in southern Brazil: indication of a microcenter of *Zea mays* L. Genetic Resources and Crop Evolution 64: 681-700.

Dawson, I.; Were, J. 1997. Collecting germplasm from trees-some guidelines. Agrofor Today 9: 6-9.

Diwan, N.; McIntosh, M.S.; Bauchan, G.R. 1995. Methods of developing a core collection of annual Medicago species. Theoretical Applied Genetics 90: 755-761.

Empresa Brasileira de Pesquisa Agropecuária [EMBRAPA]. 2018. Passport data, statistics of characterization and assessment in germplasm banks = Dados de passaporte, estatísticas de caracterização e avaliação em bancos de germoplasma Available at: https://www.embrapa.br/alelo [Accessed Feb 21, 2018] (in Portuguese).

Fick, S.E.; Hijmans, R.J. 2017. Worldclim 2: new 1-km spatial resolution climate surfaces for global land areas. International Journal of Climatology 37: 4302-4315. DOI:10.1002/joc.5086

Frankel, O.H.; Brown, A.H.D. 1984. Plant genetic resources today: a critical appraisal. p. 249-257. In: Holden, J.H.W.; Williams, J.T., eds. Crop genetic resources: conservation and evaluation. Allen and Unwin, Winchester, UK.

Galluzzi, G.; Eyzaguirre, P.; Negri, V. 2010. Home gardens: neglected hotspots of agro-biodiversity and cultural diversity. Biodiversity and Conservation 19: 3635-3654.

Hammer, K.; Arrowsmith, N.; Gladis, T. 2003. Agrobiodiversity with emphasis on plant genetic resources. Naturwissenschaften 90: 241-250.

Hill, M.O. 1973. Diversity and evenness: a unifying notation and its consequences. Ecology 54: 427-432.

Hu, J.; Zhu, J.; Xu, H.M. 2000. Methods of constructing core collections by stepwise clustering with three sampling strategies based on the genotypic values of crops. Theoretical Applied Genetics 101: 264-268.

Jost, L.; De Vries, P.; Walla, T.; Greeney, H.; Chao, A.; Ricotta, C. 2010. Partitioning diversity for conservation analyses. Diversity and Distributions 16: 65-76.

Kim, K.W.; Chung, H.K.; Cho, G.T.; Ma, K.H.; Chandrabalan, D.; Gwag, J.G.; Park, Y.J. 2007. PowerCore: a program applying the advanced M strategy with a heuristic search for establishing core sets. Bioinformatics 23: 2155-2162.

Latournerie, L.; Tuxill, J.; Yupit, E.; Arias, L.; Cristobal, J.; Jarvis, D.I. 2006. Traditional maize storage methods of Mayan farmers in Yucatan, Mexico: implications for seed selection and crop diversity. Biodiversity and Conservation 15: 1771-1795.

Leclerc, C.; Coppens, D.E.G. 2012. Social organization of crop genetic diversity: The G × E × S interaction model. Diversity 4: 1-32.

Li, Y.; Shi, Y.; Cao, Y.; Wang, T. 2004. Establishment of a core collection for maize germplasm preserved in Chinese National Genebank using geographic distribution and characterization data. Genetic Resources and Crop Evolution 151: 845-852.

Mahajan, R.K.; Bisht, I.S.; Dhillon, B.S. 2007. Establishment of a core collection of world sesame (*Sesamum indicum* L.) germplasm accessions. Sabrao Journal of Breeding and Genetics 39: 53-64.

Malosetti, M.; Abadie, T. 2001. Sampling strategy to develop a core collection of Uruguayan maize landraces based on morphological traits. Genetic Resources and Crop Evolution 48: 381-390.

Marshall, D.R.; Brown, A.H.D. 1975. Optimum sampling strategies in genetic conservation. p. 53-80. In: Frankel, O.H.; Hawkes, J.G., eds. Genetic resources for today and tomorrow. Cambridge University Press, Cambridge, UK.

Miranda, E.E. 2005. Brazil in relief = Brasil em Relevo. Campinas: Embrapa Satellite Monitoring = Embrapa Monitoramento por Satélite. Available at: http://www.relevobr.cnpm.embrapa.br [Accessed Oct 30, 2018] (in Portuguese).

Mercer, K.L.; Perales, H.R. 2010. Evolutionary response of landraces to climate change in centers of crop diversity. Evolutionary Applications 3: 480-493.

Neyman, J. 1938. Contribution to the theory of sampling human populations. Journal of the American Statistical Association 33: 101-116.

Nilsen, L.; Subedi, A.; Dulloo, M.; Ghosh, K.; Chavez-Tafur, J.; Blundo Canto, G.; De Boef, W. 2014. The relationship between national plant genetic resources programmes and practitioners promoting on-farm management: results from a global survey. Plant Genetic Resources 12: 143-146.

Odong, T.L.; Jansen, J.; van Eeuwijk, F.A.; van Hintum, T.J. 2013. Quality of core collections for effective utilization of genetic resources review, discussion and interpretation. Theoretical Applied Genetics 126: 289-305.

Orozco-Ramírez, Q.; Ross-Ibarra, J.; Santacruz-Varela, A.; Brush, S. 2016. Maize diversity associated with social origin and environmental variation in southern Mexico. Heredity 116: 477-484.

Oyarzun, P.J.; Borja, R.M.; Sherwood, S.; Parra, V. 2013. Making sense of agrobiodiversity, diet, and intensification of smallholder family farming in the highland Andes of Ecuador. Ecology of Food and Nutrition 52: 515-541.

Perales, R.H.; Brush, S.B.; Qualset, C.O. 2003. Dynamic management of maize landraces in central Mexico. Economic Botany 7: 21-34.

Pressoir, G.; Berthaud, J. 2004. Patterns of population structure in maize landraces from the central valleys of Oaxaca in Mexico. Heredity 92: 88-94.

Shannon, C.E. 1948. A note on the concept of entropy. Bell System Technology Journal 27: 379-423.

Silva, N.C.A.; Vidal, R.; Ogliari, J.B. 2017. New popcorn races in a diversity microcenter of *Zea mays* L. in the far west of Santa Catarina, southern Brazil. Genetic Resources and Crop Evolution 64: 1191-1204.

Simpson, E.H. 1949. Measurement of diversity. Nature 163: 688.

Sturges, H. 1926. The choice of a class-interval. Journal of the American Statistical Association 21: 65-66

Wang, J.C.; Hu, J.; Guan, Y.J.; Zhu, Y.F. 2013. Effect of the scale of quantitative trait data on the representativeness of a cotton germplasm sub-core collection. Journal of Zheijang University Science B - Biomedicine and Biotechnology 14: 162-170.

Williams, J.T. 1989. Practical considerations' relevant for effective evaluation. p. 235- 244. In: Brown, A.H.D.; Frankel, O.H.; Marshall, D.R.; Williams, J.T., eds. The use of plant genetic resources. Cambridge University Press, Cambridge, UK.