Title: Development of a cost-effective diversity-maximizing decision-support tool for in situ crop genetic resources conservation: The case of Cacao.

Aurelia F. Samuel, Adam G. Drucker, Sven B. Andersen, Henner Simianer, Maarten van Zonneveld

List of authors

Aurelia F. Samuel
Department of Plant and Environmental Sciences, Faculty of Science - University of Copenhagen
Thorvaldsensvej 40, 1871 Frederiksberg, Denmark.

Adam G. Drucker
Bioversity International
Via dei Tre Denari 472/a, 00057 Maccarese (Rome), Italy.

Sven B. Andersen
Department of Plant and Environmental Sciences, Faculty of Science - University of Copenhagen
Thorvaldsensvej 40, 1871 Frederiksberg, Denmark.

Henner Simianer
Department of Animal Sciences, Georg-August-University Goettingen
Albrecht-Thaer-Weg 3, 37075 Goettingen, Germany.

Maarten van Zonneveld
Bioversity International
Regional Office for the Americas, Recta km 17 - CIAT, Cali, Colombia.
Abstract

This paper undertakes an exploration of the potential application of the Weitzman prioritisation approach to the conservation of plant genetic resources (PGR). Using a case study approach and assuming limited conservation resources, this approach is applied to calculate the maximum diversity which can be conserved based on a clustering of cacao species (*Theobroma cacao* L.). A conservation budget allocation model applied across a set of ten clusters and nine subclusters of cacao, together with the use of alternative diversity and risk measures, allowed for an evaluation of a range of potential conservation outcomes. Alternative risk measures generally resulted in the allocation of conservation funds to the same priority clusters of cacao (Criollo and Curaray). However, the use of the number of locally common alleles as an alternative to the original Weitzman diversity measure produced a markedly different result, prioritizing the Purús cluster in Western Amazonia. The Curaray cluster was highly prioritized under both diversity measures, indicating its distinctiveness relative to both the Criollo and Purús clusters. We conclude that the Weitzman approach can indeed allow diversity comparisons to be made between the outcomes of different PGR conservation strategies, although subjective decisions regarding the choice of diversity and risk measures remain.

Keywords  
1. Introduction

The diversity of domesticated species and their wild relatives continues to form the basis of global food supply and contributes to the capacity of agricultural systems to cope with increasing human demands and global change. The UN Food and Agriculture Organization (FAO) reports on the State of the World’s Plant Genetic Resources (FAO 2010) and Animal Genetic Resources (FAO 2007) for Food and Agriculture, amongst others, depict an unprecedented loss across the globe of plant and animal genetic resources (including their wild relatives), as well as associated traditional knowledge and cultural practices. Causes of the loss of such diversity include institutional, market and environmental factors.

The scale of the challenge and the limited resources available suggest that conservation interventions need to be prioritized. Given that not all remaining biodiversity can be saved, a useful decision-support tool that takes into account the degree of extinction risk is the prioritization tool described by Weitzman (1992, 1993, 1998). The Weitzman approach supports decisions aimed at answering the “Noah’s Ark” question i.e. “which species to take on board Noah's Ark?” The suggestion was that Noah should take species on board “in the order of their gains in utility plus diversity, weighted by the increase in their probability of survival, per dollar of cost (Metrick and Weitzman 1998, Weitzman 1998). The Weitzman approach thus combines measures of diversity, current risk status and conservation costs so as to permit the identification of a cost-effective diversity-maximizing conservation portfolio. Hence, for any given quantity of conservation funding available, it is possible to identify a priority conservation portfolio that maximizes the diversity which can be conserved.

Nevertheless, despite the conceptual basis developed and the strong appeal due to its rigorous mathematical justification, as well as the possibility to derive optimum conservation decisions with well-defined properties, the application of the approach remains largely limited to animals. While the original study focused on wild species of crane, Weitzman approaches have also been successfully developed for case study analyses of the in situ conservation of livestock breeds (Reist-Marti et al. 2003, Simianer et al. 2003, Pinent et al. 2005, Zerabruk et al. 2007, Zander et al. 2009). Yet the approach has not only not been widely adopted in the context of informing genetic resources conservation programme design and implementation (Drucker 2010), but also remains to be adapted and tested for application to plant genetic resources (PGR). Such adaptation might be expected to be particularly useful in supporting concrete PGR conservation policy design and implementation.
The main aim of this paper is to undertake an innovative application of the Weitzman approach to a specific crop gene pool in order to assess its broader applicability to plant genetic resources conservation in general. An assessment is carried out through a case study approach that permits an exploration of issues related to the impacts on diversity conservation outcomes resulting from the use of different diversity measures, risk measures (including climatic factors, accessibility of the resource and inbreeding) and alternative conservation funding allocation strategies.

Cacao (*Theobroma cacao*) was selected for this case study approach as it provides a useful illustrative example of a commodity crop, which like many other plant genetic resources (PGR) has been increasingly subject to genetic erosion despite its important contribution to smallholder farmer livelihoods. Threats to the conservation of cacao infra-specific and wild relative diversity include the loss of remaining forest cover, the simplification of cacao shade canopies and the conversion of cacao agroforestry systems to other agricultural land uses (Schroth and Harvey, 2007). Although the choice of risk measures related to the above threats for our case study crop is in part motivated by the availability and abundance of data on cacao, we also consider that this case study approach is appropriate for shedding light on the usefulness of a PGR-focused Weitzman approach in general.

2. Conservation Strategies

2.1 Morphological and Molecular data

Diversity measures may be derived from morphological or molecular data. Morphological markers (or phenotypic markers) are observable traits that were among the earliest markers used in germplasm management (Stanton et al. 1994). Morphological traits however, are normally affected by the production environment, which may create challenges for genetic characterization, and they may also be subject to convergent development, so that livestock breeds or plant cultivars with widely different genetic origin become similar in appearance. Common garden experiments aim to evaluate traits of plant varieties under similar conditions but are still constrained by environment. The use of molecular markers for diversity analysis on the other hand, may require more investment in technology, but have become increasingly popular and are less subject to convergent selection by humans (Toro and Caballero 2005). Most genetic characterization studies use neutral molecular markers, which do not directly measure the diversity of functional traits for current and future use. Instead they are used to measure kinship and genetic distinctiveness between identified varieties and can be used in optimization approaches of relevance to the *in situ* conservation of crop genetic resources (Eaton et al. 2006). Over recent years several methods have been developed to conserve
genetic diversity in *ex situ* core collections based on such genetic characterization data (Kim et al. 2007, Thachuk et al. 2009). However, PGR *in situ* conservation optimization approaches should, in addition to genetic distinctiveness, also consider socio-economic and environmental factors that stimulate or threaten conservation and use of genetic resources on farm and in wild populations (Willemen et al. 2007).

2.2 *Diversity Measures*

The more familiar measures of neutral molecular diversity applied to plant genetic resources in the literature include: the allelic richness or the number of alleles per locus corrected by rarefaction (El Mousadik and Petit 1996, Petit et al. 1998, Widmer and Lexer 2001, Hao et al. 2006), number of private or unique alleles (Petit et al. 1998, Motamayor et al. 2008), the number of locally common alleles\(^1\) (van Zonneveld et al. 2012) and the gene diversity - defined by Nei (1973) as the probability that two alleles chosen at random from the population are different. Genetic distances are also commonly used in assessing genetic diversity. These include; Goldstein et al’s (1995) average squared distance, the shared allele distance (Chakraborty and Jin 1993) and Nei’s standard genetic distance (Nei 1972).\(^2\) Given this wide range of existing literature on diversity indices, it is interesting to consider, as we do below, the degree to which such measures may be useful in informing a Weitzman type of analysis focussed on PGR.

The Weitzman approach uses a pair-wise distance matrix that can be used to obtain a measure of the optimum genetic diversity between operational taxonomic units (OTUs) such as species, populations, varieties or genetic cluster groups, among others. These genetic distances are subsequently weighted by their extinction probabilities to calculate the “expected diversity”, which combines a measure of both the distinctiveness of the genetic resources and their risk of becoming extinct. The extinction probability is a risk measure - whose estimation is based on genetic resource-related

\(^1\) “Locally common alleles” are those that are frequent only in one to several populations (Maguire et al. 2002) or in a limited area and may indicate a long history of selection and local adaptation to specific environments (Frankel et al. 1995). Such locally common alleles have previously been applied to prioritize areas of *in situ* conservation of an outcrossing fruit tree species cherimoya (*Annona cherimola* Mill.) (van Zonneveld et al. 2012).

\(^2\) Although it is beyond the aim of this study to analyse the choice of genetic distance measure, since it depends on the type of morphological and molecular markers under consideration (see Staub et al. 1996, Motamayor et al. 2002), an overview can be found in Laval et al (2002) for animal genetic resources and in Lowe et al. (2004), de Vicente et al (2004) and Kumar et al. (2009) for general explanations about different markers.
and external socio-economic factors (and in some cases, within-population information such as the degree of inbreeding) used to help prioritize the genetic resources for conservation. This risk measure reveals how the diversity of a set of populations is expected to change in the context of biodiversity management programs (García et al. 2005) and is discussed in further detail below. In the context of the prioritization of conservation of domesticated animal breeds, the within-breed coefficient of inbreeding has often been used as the major criterion for determining the extinction risk. It is then assumed that breed-level (equivalent to varieties or cultivars in plants) conservation efforts and funding might be allocated in such a way as to reduce the inbreeding rate such that the conserved diversity between breeds is maximized (Simianer et al. 2003). García et al. (2005) have sought to address within-population diversity issues (Caballero and Toro 2002, Eding et al. 2002), as well as issues related to the subjective choice of extinction probability indicators, by defining a close relationship between the rate of inbreeding and the probability of extinction in animals. This relationship was subsequently used to calculate genetic extinction probabilities based on allelic frequencies.

2.3 Risk Factors

A major consideration in the use of the Weitzman approach to prioritize funding for the conservation of genetic resources is the choice of risk function, used to represent the risk of extinction of each OTU within a certain time horizon. The risk measure of a genetic resource is expressed as its extinction probability (Tapio et al. 2006). A high risk of extinction combined with high genetic distinctiveness would indicate that active intervention may be required to reduce the risk of losing genetic diversity. Such interventions could involve in situ and/or ex situ conservation schemes (Pattison et al. 2007). While in this study, we need to apply risk measures appropriate for PGR conservation issues and for which data is available, we note that in the case of prioritization for the conservation of domesticated animal breeds, the intra-breed inbreeding coefficient has often been used as the major criterion determining the risk of breed extinction (Thaon d'Arnoldi et al. 1998, García et al. 2005, Pinent et al. 2005, Tapio et al. 2006, Zerabruk et al. 2007, Zander et al. 2009), often in combination with other risk factors. This is because high degrees of inbreeding and low effective population sizes are considered to indicate vulnerability and risk of loss in highly out-breeding species like animals and tree species. Consequently, the FAO defines an AnGR risk measure largely based on population size and trends (FAO 1998). However, other risk measures have also been derived from socio-economic factors such as the present value of the breed, the existence or not of a conservation scheme or breed organizations, the spatial
distribution of the breed, the degree of market access and its socio-cultural importance (see Zander et al. 2009, Cuc et al. 2011).

Given the issues of risk measure appropriateness for PGR, as well as data availability, in addition to considering a measure of inbreeding, we also take into account accessibility and environmental factors. The accessibility risk measure allows us to account for the distance of the PGR in question to the nearest human settlement. This threat measure has been identified as being particularly important for natural ecosystems in South America (Jarvis et al. 2010). Consequently, we explore the potential utility of this risk measure within the Weitzman approach as it is hypothesized that it is positively correlated with increases in the likelihood of replacement of current genotypes by new commercial materials and hence can be associated with threats related to changing land use. Such “accessibility-driven” replacement in crops has, for example, already been used to explain genetic erosion of cassava in the Peruvian Amazon (Willemen et al. 2007) and for cacao in Nicaragua (Trognitz et al. 2011).

A third risk measure considered in this study is related to suitability to projected climatic changes, given the long-term nature of this threat. Changes in climatic factors are increasingly being identified as potential threats to the conservation of PGR (Jarvis et al. 2009) and in the context of cacao may be considered to relate to changes in forest cover and hence shade canopies important in cacao production. To date, climatic factors have not been considered within studies that have applied the Weitzman approach.

Materials and methods

Good classification based on genetic diversity measures may be expected to greatly facilitate the development of efficient strategies for conservation. Consequently, this study has chosen to focus on a major crop that is both important to small-scale farmers’ livelihoods as well as being well classified in terms of its diversity.

Cacao is a monoecious species, pre-dominantly outcrossing although some trees also self-fertilize with insect-mediated pollination (Silva et al. 2011, MAPFORGEN 2012). Varieties can be propagated sexually or clonally (Eskes 2011). In Peru, for example, 63% of the total cacao production of 80,000 hectares still consists of traditional varieties that are seed-propagated (Eskes 2011).
For *Theobroma cacao*, Motamayor et al. (2008) introduced a new classification system based on molecular markers. They identified 10 genetic clusters: Marañon, Curaray, Criollo, Iquitos, Nanay, Contamana, Amelonado, Purús, Nacional and Guiana, as opposed to the traditionally recognized varieties “Criollo” and “Forastero” or “Trinitario” (consisting of “Criollo” x “Forastero” hybrids). The data set analysed in this study consists of 952 *Theobroma cacao* accessions previously described by Motamayor et al (see referenced link of data in Motamayor et al. 2008:p.7). The original microsatellite data, DNA extraction from the samples, microsatellite markers and Polymerase Chain Reaction (PCR) amplification are all reported in Motamayor et al. (2008). An analysis of the geographic coordinates of the 952 accessions led to twelve accessions being removed as they were found to be located in Africa (focus of this study is on samples from South and Central America) or in the ocean (erroneous points). Among the remaining 940 geo-referenced individuals of *Theobroma cacao* analysed with 96 microsatellite markers in Motamayor et al. (2008), 723 were retained for this study, as they had a coefficient of membership to their most likely cluster equal to or higher than 0.70 (as identified by Motamayor et al. 2008) based on a Bayesian model-based clustering algorithm (Pritchard et al. 2000).

*The Weitzman algorithm*

The Weitzman approach starts by calculating the total diversity of a set $D(S)$ from an N x N matrix with a recursive algorithm (i.e. an algorithm that first solves a smaller version of the problem) as suggested by Weitzman (1992). This is defined as the solution to a recursion:

$$ D(S) = \max_{i \in S} \left[ D(S_{\bar{i}\bar{S}}) + d(i, S_{i\bar{S}}) \right] $$

(1)

where $D(S)$ denotes the diversity function of a set $S$ of $N$ clusters, $S_{i\bar{S}}$ represents a subset without cluster $i$ and $d(i, S_{i\bar{S}})$ denotes the distance between cluster $i$ and the set without $i$ (see further explanation on related theoretical and empirical procedures in Thaon d’Arnoldi et al. 1998). From the above, the partial contribution of cluster $i$ can be calculated as:

$$ PC_i = \left[ D(S) - D(S_{i\bar{S}}) \right] D(S) $$

(2)

Furthermore, the expected diversity at the end of the assumed time period (which reflects the amount of diversity remaining) is computed in Simianer et al (2003) as:
\[ E(D) = \sum_{K} P(K)D_K \]  

(3)

where \( K \) is a vector of size \( N \) which contains the indicator variables \( k_i, i=1,\ldots, N \), where \( k_i = 1 \) if cluster \( i \) exists at time \( t \) (where \( t \) is the assumed planned horizon, which is assumed to be 50 years in our model as per Reist-Marti et al. 2003 and Simianer et al. 2003) and \( k_i = 0 \) if cluster \( i \) is extinct.

Let \( 0 \leq z_i \leq 1 \) be the extinction probability, reflecting the chance that cluster \( i \) becomes extinct before time \( t \) is reached. The probability \( P \) that a given situation arises based on the specified vector \( K \) can be computed as:

\[ P(K) = \prod_i (k_i + (-1)^i z_i) \]  

(4)

So if \( D_K \) denotes the diversity of the set of clusters not extinct at time \( t \) in vector \( K \), then the expected diversity at the end of the time period is as given in equation (3) above. Hence, the relative change in expected diversity if the extinction probability of the respective cluster is increased by one unit (expressed as a fraction, decimal or percentage) represents the marginal diversity (a negative value) and can be calculated as:

\[ D_i' = \frac{\partial E(D)}{\partial z_i} \]  

(5)

From the above equations, the most useful criterion in establishing conservation priorities is derived; this is termed the conservation potential \( CP_i \). Thus, the conservation potential of cluster \( i \) is:

\[ CP_i = -z_i \times D_i' \]  

(6)

The conservation potential thus reflects the possible increase in expected diversity, if an endangered cluster \( i \) was made completely safe.
**Optimum allocation of conservation funds**

According to Simianer et al. (2003)\(^3\), in the context of AnGR, the efficient use of conservation funds should result in a distribution of the total available funds among a set/subset of all endangered breeds such that the conserved diversity between breeds is maximized. Simianer et al. (2003) proposed three different models for the optimum allocation of conservation funds and demonstrated how the allocation of such funds under the different models may result in different effective population sizes and inbreeding levels (Simianer et al., 2003:pp.380-382). Model A assumes that the effect of a conservation scheme is to increase effective population size by an additive increment which is proportional to the funds spent per animal in this population. By contrast, model B assumes a multiplicative effect on effective population size which is proportional to the amount of money spent per individual animal in the population. Model C assumes that the effect of a conservation scheme on effective population size is multiplicative by a factor that is only dependent on the amount of conservation funding spent on the overall population, irrespective of its (effective) size.

However, in considering which model (A, B or C) would be most appropriate for PGR applications, two major issues arise given that the PGR context differs from the AnGR one. Firstly, we note that the animal models link an increase in effective population size to a reduction in inbreeding due to conservation interventions, where funding is spent per individual animal. Secondly, for many plant species the effect of a conservation scheme is expected to be mostly proportional to the funding spent per (farm) unit area of the endangered crop, rather than per individual plant, as crop species are generally sown or planted en masse. However, for some plant species and specific purposes, it may also be desirable to spend conservation funds on individual trees or on a sub-population basis per unit area. Such a case might arise, for example, for trees with specific characteristics maintained in backyards (e.g. Criollo cacao varieties).

Taking the above PGR conservation issues into account, an adapted version of model A appears to be the most appropriate for plants. Under such a model, it is assumed that the effect of a conservation scheme is to increase the population size by an additive increment, which is proportional to the funds spent per (farm) unit area or per single tree in a population. The actual total budget required for conservation will then be a function of the desired population size goal which will need to be established *a priori* by a decision-maker.

---

\(^3\) See detailed comparison of these models in Simianer et al (2003).
The specifications of the models in Simianer et al. (2003) linked conservation expenditure to a reduction in extinction probability between breeds. Hence, the adapted version of model A starts with the basic assumption that the value of the measure of risk is inversely proportional to the population size.

\[ z_i = \gamma \frac{1}{2N_i} \]  

(7)

Where \( N_i \) is the population size and, \( 0 < \gamma \leq 1 \) is a constant.

The allocation mechanism developed implies that a certain budget \( b \) spent on a population will increase population size and hence the value of the risk measure of that population will reduce to \( z_i^* \). This reduction in risk is represented as \( \Delta z_i = z_i^* - z_i \). A functional relationship between \( b, z_i \) and \( \Delta z_i \) can then be derived as \( \Delta z_i = f(z_i, b) \). The reduction in risk is subsequently calculated under the assumption of model A, which implies that the conservation scheme/cost will have an additive effect on the population size (i.e. \( \frac{b}{N_i} \)) in this population. Therefore, \( N_i \) is increased to \( N_i^* \) where \( N_i^* = N + \lambda \frac{b}{N_i} \) and \( \lambda > 0 \). The substitution and recalculation of these functions results in equation (7) which gives the reduction in the value of risk for model A as

\[ \Delta z_i = -\frac{4\lambda b z_i^2}{\gamma^2 + 4\lambda b z_i^2} z_i \]  

(8)

Application

Based on the classification of Motamayor et al. (2008), the Weitzman approach was applied to the 10 identified clusters of sampled trees from different parts of tropical America presented in Table 1. Although these are not biological populations of the type applied in existing Weitzman AnGR analyses, this botanical classification and dataset is suitable for the purposes of the current analysis and allows for conservation recommendations at a regional scale to be generated. All calculations of Weitzman diversity between these clusters/sub-clusters were made using a Weitzpro program (Derban et al. 2002), which not only computes the diversity, but also the percentage contribution to overall diversity made by each OTU. Thus, given the extinction risk measure and diversity values, the marginal diversity and the conservation potential of each OTU can be derived. Nei’s standard distance (1972) was applied using the gene marker software PowerMarker v 3.25 (Liu and Muse 2005). The locally common alleles (LCA) measure was
calculated as the number of locally common alleles occurring with a frequency higher than 5% and found in 25% or fewer clusters using GenAlEx 6.4 (Peakall and Smouse 2006).

As previously noted, the threat measures selected for analysis in this study are; accessibility, future climate suitability and inbreeding, as described in Table 1. In our model, accessibility increases the likelihood of replacement of current genotypes by new commercial materials or other crops. Accessibility values were obtained from an accessibility map developed by Nelson (2008) who carried out a combined analysis of sites and transport networks to quantify the time from any site to the nearest populated place of more than 50,000 habitants, in order to provide an index that captures accessibility. The assumption is that the further the individual trees are from such human population centres, the less threatened by replacement with other varieties or other crops they become.

The climatic change adaptability risk factor was calculated using Maxent climate envelope modelling (Phillips et al. 2006) to identify which cacao genotypes are most likely to be affected by climate change by the 2050s (Emission scenario A2a, average of 19 GCM models downscaled to 2-5 minutes resolution [http://ccafs-climate.org/]). Since it may be assumed that individuals adapted to different climatic regions will respond differently to climate change, we modelled the impact of climate change for groups of accessions from different climatic zones. Climatic clustering of the 940 accessions were performed in R based on extracted values from 19 bioclimatic layers of Worldclim at a resolution of 2-5 minutes (Hijmans et al. 2005) and the optimal number of 7 climate clusters was identified in R with a Kelley-Gardner-Sutcliffe penalty function following Grum and Atieno (2007). We assigned a suitability value of occurrence in the 2050s for each plant individual per climate cluster. This was done with Maxent climate envelope modelling of cacao distribution per climate cluster under the expected climate conditions by that time. These suitability values are between 0 and 1. The lower the suitability value the higher the threat of climate change.

Lastly, this study builds upon García et al.’s (2005) notion of genetic extinction and within-group information by including an extinction risk measure for inbreeding. A relative homozygosity index was calculated using PowerMarkerv3.25 (Liu and Muse, 2005) and used to construct an extinction risk measure based on inbreeding. The software calculates the frequency of observed homozygous loci divided by the expected frequency of homozygosity, assuming Hardy-Weinberg equilibrium. This measure of inbreeding ranges between a value of 0 (complete heterozygosity) to a value of 1 (complete inbreeding) if the population is in Hardy-Weinberg equilibrium. In other
words, this is the relative surplus of homozygous loci observed to the expected with Hardy-Weinberg. As mentioned above, we did not apply this model to biological populations and hence homozygosity measures should not be interpreted under the original Hardy-Weinberg equilibrium assumptions. Instead, we applied this approach to genetic clusters of geographically scattered tree individuals at a continental level, resulting in the consideration of the calculated levels of homozygosity as an estimate of potential inbreeding effects in each genetic group.
Table 1

Risk measures of *Theobroma cacao* L. clusters.

<table>
<thead>
<tr>
<th>Variables (abbreviation)</th>
<th>Criterion and value</th>
</tr>
</thead>
</table>
| Accessibility (*a*)      | 1= higher accessibility/high threat  
                          | 0= lower accessibility / low threat |
| Predicted change in climate by the 2050s (*c*) | 1= high degree of climate change /high threat  
                          | 0=low degree of climate change/low threat |
| Inbreeding (*m*)         | 1=inbreeding/homozygosity/high threat  
                          | 0=heterozygosity/ low threat |
Although it is acknowledged that definitive risk measures would be difficult to develop, as it is unlikely that all factors of influence can ever be taken into account and unexpected future changes can always occur (Reist-Marti et al. 2003), it is nonetheless worthwhile to consider a range of risk measures as their varying conservation strategy implications permit us to obtain insights into their appropriateness for different types of conservation programs (i.e. crop types) and objectives. Weitzman-based studies (such as Zander et al. 2009; and Cuc et al. 2011) have applied the sum of a number of different types of risk factors and rescaled them using the formula presented in Reist-Marti et al (2003). We have consequently adopted this method, rescaling to a value between 0.2 and 0.8 to rule out the possibility that a group may be considered completely safe due to the existence of a conservation program or entirely doomed to extinction after investment in its conservation. Thus, the values of our three selected risk variables, inbreeding, accessibility and climatic factor, are calculated as follows

\[ z_i = \frac{0.6}{x_d} \sum_{a=1}^{3} z_{ia} + 0.2 \]

Where \( x_d \) is the difference between the maximum and minimum values in cluster \( i \), \( z_{ia} \) is the value of variable \( a \) of cluster \( i \). The risk measures assume that if no conservation efforts are made, the genetic diversity of cluster \( i \) will be lost within the next 50 years.

Maps of the geographic distribution of potential funding allocation based on our modelling were drawn up in DIVA-GIS 7.3 (www.diva-gis.org) with a resolution of 10 minutes (approximately 18 km at the equator) and using a circular neighbourhood of one diameter (approximately 111 km at the equator). In each grid cell, we summed the allocated funds of each cacao cluster with respect to their area of origin. Final maps were edited in ArcMap 10.

For our adapted allocation model, 100 currency units each with a nominal value of 500 were assumed to be made available for allocation, leading to a total budget of 50,000 (i.e.100 x 500) units. All the parameters were calculated as described in Simianer et al (2003).

3. Results

3.1 Weitzman diversity, risk measures and conservation potential

A Weitzman algorithm, using four different measures of risk based on accessibility, future climate change, inbreeding and a combined measure represented by their average (\( tz \)) was used to estimate marginal diversities and conservation
potentials (Table 2 and 3) of the ten major clusters of Cacao. The Weitzman total diversity $D$ was 8.05 (Table 2). When each of the ten clusters were individually excluded from the set, exclusion of the Criollo cluster led to the biggest loss in total diversity, down to 5.71 which is equivalent to a 29.1% decline (Table 2). Exclusion of the remaining clusters led to much smaller losses in diversity, down to only 7.17 from 7.72, which is equivalent to only a 4.0% to 10.9% decline (Table 2). The maximum likelihood tree (see Fig.1) showed that the Criollo cluster was distinct with its own branch, while the remaining nine clusters were more closely related. The Criollo cluster also had the highest pair-wise distances relative to the remaining cacao clusters, including when other genetic distance measures such as Goldstein et al.’s (1995) average squared distance, the shared allele distance (Chakraborty and Jin, 1993) and Reynolds distance (1983) were applied to these data (results not shown). The Criollo cluster also showed the highest marginal diversities ($D'$) for all three risk measures in Table 2.

Furthermore, the Criollo group ranked first in conservation potential ($CP$), with all four risk measures indicating the highest total genetic diversity conserved if this cluster is made safe (see Tables 2 and 3). The remaining cacao groups ranked differently in terms of conservation potential depending on the risk measures. For example, the Amelonado group was ranked second with regard to conservation potential when using the accessibility and inbreeding risk measures, but fell to seventh under the future climate change risk measure. Similarly, the Nanay, Guiana and Nacional groups of cacao changed their relative conservation potential rankings under the three different risk measures (Table 2). When the combined risk measure $tz$ was used (Table 3), the Criollo group again remained the cluster with the highest conservation potential ranking. The combined risk measure also ranked the remaining cacao groups quite similarly except for the lowest three ranking clusters (Iquitos, Marañon and Purus), while the Marañon cluster remained ranked 10th three times out of four (Table 3).

No exceptional differences in the contribution to diversity between the groups could be identified even in a reduced set (excluding Criollo) when compared with the results in Table 2 (see Appendix A). For example, the contribution of the Contamana and Curaray groups continued to be higher than that of the remaining groups, as in Table 3. However, contributions from both the Purús (10.86%) and the Nacional (12.54%) groups are higher than the Amelonado (9.6%) group in Table 2.
Table 2
Weitzman results based on different measures of risk for 10 clusters of cacao (notations are described below).

<table>
<thead>
<tr>
<th>Clusters</th>
<th>Cont. %</th>
<th>Accessibility ((a))</th>
<th>Climate change ((c))</th>
<th>Inbreeding measure ((in))</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(D)</td>
<td>(z(a)) (D'(a))</td>
<td>(CP(a))ranking (z(c)) (D'(c))</td>
<td>(CP(c))ranking (z(in)) (D'(in))</td>
</tr>
<tr>
<td>Amelonado</td>
<td>7.276</td>
<td>9.6 0.774 0.969</td>
<td>0.750(2\textsuperscript{nd}) 0.352 0.908</td>
<td>0.320(7\textsuperscript{th}) 0.689 0.685</td>
</tr>
<tr>
<td>Contamana</td>
<td>7.223</td>
<td>10.2 0.619 0.852</td>
<td>0.527(6\textsuperscript{th}) 0.435 0.806</td>
<td>0.351(6\textsuperscript{th}) 0.456 0.857</td>
</tr>
<tr>
<td>Criollo</td>
<td>5.706</td>
<td>29.1 0.456 2.034</td>
<td>0.927(1\textsuperscript{st}) 0.476 2.288</td>
<td>1.090(1\textsuperscript{st}) 0.800 2.145</td>
</tr>
<tr>
<td>Curaray</td>
<td>7.174</td>
<td>10.9 0.551 1.095</td>
<td>0.603(4\textsuperscript{th}) 0.630 0.968</td>
<td>0.610(2\textsuperscript{nd}) 0.436 1.084</td>
</tr>
<tr>
<td>Guiana</td>
<td>7.553</td>
<td>6.14 0.599 0.876</td>
<td>0.524(7\textsuperscript{th}) 0.591 0.670</td>
<td>0.396(3\textsuperscript{rd}) 0.526 0.732</td>
</tr>
<tr>
<td>Iquitos</td>
<td>7.659</td>
<td>4.8  0.756 0.571</td>
<td>0.432(8\textsuperscript{th}) 0.746 0.483</td>
<td>0.361(5\textsuperscript{th}) 0.200 0.469</td>
</tr>
<tr>
<td>Marañon</td>
<td>7.724</td>
<td>4.0  0.356 0.538</td>
<td>0.192(10\textsuperscript{th}) 0.359 0.415</td>
<td>0.149(10\textsuperscript{th}) 0.358 0.415</td>
</tr>
<tr>
<td>Nacional</td>
<td>7.331</td>
<td>8.9  0.544 0.979</td>
<td>0.533(5\textsuperscript{th}) 0.246 1.028</td>
<td>0.253(9\textsuperscript{th}) 0.487 0.915</td>
</tr>
<tr>
<td>Nanay</td>
<td>7.562</td>
<td>6.0  0.795 0.843</td>
<td>0.670(3\textsuperscript{rd}) 0.547 0.718</td>
<td>0.393(4\textsuperscript{th}) 0.406 0.660</td>
</tr>
<tr>
<td>Purús</td>
<td>7.422</td>
<td>7.8  0.597 0.603</td>
<td>0.360(9\textsuperscript{th}) 0.503 0.611</td>
<td>0.308(8\textsuperscript{th}) 0.436 0.568</td>
</tr>
<tr>
<td>Total Diversity</td>
<td>8.046</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: \(D\) is the Overall Weitzman Diversity (in the absence of each respective cluster), \(z\) is the risk value for each threat measure, \(D'\) denotes the marginal diversities and \(CP\) is the conservation potential. Cont. % is the percentage contribution of a cluster is equal to the amount of diversity reduced arising from the exclusion of that cluster.
Table 3

Weitzman results based on different combined measures of risk for 10 clusters of cacao.

<table>
<thead>
<tr>
<th>Clusters</th>
<th>$z$ ($tz$)</th>
<th>$D$ ($tz$)</th>
<th>$CP(tz)$ ranking</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amelonado</td>
<td>0.605</td>
<td>0.840</td>
<td>0.508 (3rd)</td>
</tr>
<tr>
<td>Contamana</td>
<td>0.503</td>
<td>0.841</td>
<td>0.423 (6th)</td>
</tr>
<tr>
<td>Criollo</td>
<td>0.577</td>
<td>2.160</td>
<td>1.247 (1st)</td>
</tr>
<tr>
<td>Curaray</td>
<td>0.539</td>
<td>1.052</td>
<td>0.567 (2nd)</td>
</tr>
<tr>
<td>Guiana</td>
<td>0.572</td>
<td>0.754</td>
<td>0.431 (5th)</td>
</tr>
<tr>
<td>Iquitos</td>
<td>0.567</td>
<td>0.505</td>
<td>0.287 (9th)</td>
</tr>
<tr>
<td>Marañon</td>
<td>0.358</td>
<td>0.447</td>
<td>0.160 (10th)</td>
</tr>
<tr>
<td>Nacional</td>
<td>0.426</td>
<td>0.974</td>
<td>0.415 (7th)</td>
</tr>
<tr>
<td>Nanay</td>
<td>0.583</td>
<td>0.742</td>
<td>0.432 (4th)</td>
</tr>
<tr>
<td>Purús</td>
<td>0.512</td>
<td>0.591</td>
<td>0.303 (8th)</td>
</tr>
</tbody>
</table>

Note: Unless otherwise stated, notations are as detailed in Table 2.
3.2 Optimum allocation of conservation funds and expected diversity

The optimal allocation of resources to the ten different clusters of cacao using the adapted allocation model A was calculated for the total average risk measure \((tz)\) and the risk measure based on inbreeding \((in)\) as seen in Fig. 1. Under both of these risk measures, the largest share of the funds was allocated to the Criollo cluster. Using \(tz\) and \(in\), the allocation model suggests spending 51\% and 68\% of the total funds, respectively, on the Criollo cluster. These two risk measures \((tz\) and \(in)\) also assigned 17\% and 24\%, respectively, of the funds to the Amelonado cluster; and 7\% and 3\%, respectively, to the Guiana cluster. The geographical distribution of the allocated funding using only the inbreeding risk measure \((in)\) is indicated in Fig. 2. The area with highest allocation is Central America and the northern Caribbean, an analytical finding that is expected to result from the fact that it is in these locations that the Criollo cacao accessions had been collected. As further shown in Fig. 1, using the total risk measure, the clusters that ranked first to third in Table 3, were all candidates for conservation funding, while the last four clusters (Iquitos, Maraño, Nacional & Purús) were not allocated any funds at all. In the case of Nacional, this is despite the fact that, it had a similar \(CP(tz)\) value (0.42) with both Nanay (0.43) and Guiana (0.43). Under the inbreeding \((in)\) risk measure (Table 2), funds were allocated to relatively fewer of the cacao clusters (Fig.1).

Without the Criollo cluster, the use of the total \((tz)\) and inbreeding \((in)\) risk measures still demonstrated different optimum allocations of funds (see Appendix B). Using \(tz\), all the clusters with the exception of Maraño were allocated funds, while the use of the inbreeding \((in)\) risk measure allocated funds to only five clusters.
Fig. 1. Maximum likelihood tree from the Weitzman algorithm (1992) and the allocation of funds for the conservation of 10 cacao clusters based on total risk ($tz$) and inbreeding ($in$) measures respectively.

Fig. 2. Geographical distribution of the allocated funding of cacao based on the inbreeding risk measure ($in$) and model A.
3.3 Contrasting diversity implications of conservation outcomes

The expected diversity of cacao was calculated using the four different risk measures (Fig. 3). The results indicated that most expected diversity is conserved when the risk measure is based only on inbreeding (ED\text{in}: 35%). Using accessibility as the risk measure generates similar results, while the least expected diversity is conserved when using the climate change risk measure (ED\text{c}: 17%).

By contrast, for the cacao clusters without the Criollo group, the results in Fig. 3 indicate that the most expected diversity is conserved when the accessibility risk measure is used (ED\text{a}: 39%). Using the inbreeding measure (\text{in}) results in lower expected diversity conserved (ED\text{in}: 21%), while, once again, the climate change measure results in the least conserved expected diversity (ED\text{c}: 18%).

![Figure 3](image)

Fig. 3. Percentage increase in expected diversity of 10 cacao clusters (with Criollo) and 9 cacao clusters (without Criollo) under conservation fund allocation model A (Risk notations \text{c}, \text{a}, \text{in} and \text{tz} are as explained in Tables 2 and 3).
3.4 Locally common alleles (LCA) diversity measure

The marginal diversities ($D'$), for the 10 clusters under the inbreeding risk measure (Table 4, column B) and the combined risk measures (column C) are approximately the same as their actual respective LCA diversity values presented in column A. Given that the conservation potential is simply the marginal diversity multiplied by the risk measure (see equation 6); the actual LCA diversity values can easily be used as a substitute for the original Weitzman marginal diversity measure in all applications that follow the above format. The results reveal that prioritization based on the LCA is in contrast to those derived using the original Weitzman diversity measure (Tables 2 and 3). The former lower ranked Purús and Contamana clusters (in Tables 2 and 3) are now ranked highest (first and second, respectively), while the Criollo and Amelonado clusters now rank among the lowest (tenth and eighth, respectively in column C in Table 4). By contrast, the Curaray cluster was highly ranked under both the original Weitzman diversity measure and the LCA measure, with CP ($tz$) values ranking, respectively second and third (Tables 3 and 4).
Table 4
Locally Common Alleles as a Weitzman marginal diversity \((D')\) measure based on alternative risk measures for 10 clusters of cacao.

<table>
<thead>
<tr>
<th>Clusters</th>
<th>COLUMN A</th>
<th>COLUMN B</th>
<th>COLUMN C</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LCA(ranking)</td>
<td>(D) (in)</td>
<td>(CP) (in) (ranking)</td>
</tr>
<tr>
<td>Amelonado</td>
<td>0.448(8\textsuperscript{th})</td>
<td>0.449</td>
<td>0.309(6\textsuperscript{th})</td>
</tr>
<tr>
<td>Contamana</td>
<td>1.260(2\textsuperscript{nd})</td>
<td>1.261</td>
<td>0.574(2\textsuperscript{nd})</td>
</tr>
<tr>
<td>Criollo</td>
<td>0.302(10\textsuperscript{th})</td>
<td>0.302</td>
<td>0.242(8\textsuperscript{th})</td>
</tr>
<tr>
<td>Curaray</td>
<td>1.052(3\textsuperscript{rd})</td>
<td>1.052</td>
<td>0.458(3\textsuperscript{rd})</td>
</tr>
<tr>
<td>Guiana</td>
<td>0.333(9\textsuperscript{th})</td>
<td>0.334</td>
<td>0.176(9\textsuperscript{th})</td>
</tr>
<tr>
<td>Iquitos</td>
<td>0.875(5\textsuperscript{th})</td>
<td>0.875</td>
<td>0.175(10\textsuperscript{th})</td>
</tr>
<tr>
<td>Marañon</td>
<td>0.990(4\textsuperscript{th})</td>
<td>0.991</td>
<td>0.355(4\textsuperscript{th})</td>
</tr>
<tr>
<td>Nacional</td>
<td>0.646(6\textsuperscript{th})</td>
<td>0.646</td>
<td>0.315(5\textsuperscript{th})</td>
</tr>
<tr>
<td>Nanay</td>
<td>0.625(7\textsuperscript{th})</td>
<td>0.625</td>
<td>0.254(7\textsuperscript{th})</td>
</tr>
<tr>
<td>Purús</td>
<td>1.365(1\textsuperscript{st})</td>
<td>1.364</td>
<td>0.595(1\textsuperscript{st})</td>
</tr>
</tbody>
</table>

Note: Unless otherwise stated, notations are as detailed in Tables 2 and 3.
4. Discussion

The results indicate different optimal prioritization strategies depending on the choice of risk measure. As there is no way to determine, which risk measure will best predict future outcomes, the onus is on the decision-maker to judge which of the four different measures (climate change, accessibility, inbreeding and a combined measure) detailed in this study is most appropriate for the specific PGR under their consideration. However, the methods presented in this study can nonetheless assist in reducing the complexity of the decision to be made, while differences in the outcomes predicted by the different risk models help raise a number of important questions that need to be addressed if the goal is to define a cost-efficient diversity-maximizing conservation strategy.

For example, the results clearly show that, in the case of cacao, the Criollo cluster is inbred and that the most expected diversity would be conserved when interventions that take into account the inbreeding measure as the only risk factor are used. This inbreeding of the Criollo cluster may be the result of selection (for flavour and quality traits) during cultivation/domestication (Motamayor et al. 2002), combined with self-compatibility (Warren et al. 1995) and habitat fragmentation (Aguilar et al. 2008). Although special characteristics may be maintained through inbreeding, it may also result in a reduction of adaptive capacity and evolutionary potential, making tropical tree species populations [such as those of cacao] more vulnerable to changes in climate, as well as the impacts of pest and diseases (Dawson et al. 2009). Hence, there are a number of arguments to support the use of an inbreeding risk measure when this cacao cluster is prioritized in its entirety.

For the remaining cacao clusters the discriminating effect of the inbreeding risk measure is much less prominent, possibly because differences in inbreeding are less pronounced. In particular, the Iquitos cluster from the Peruvian Amazon shows an excess of heterozygosity (relative inbreeding = -0.112), which could be due to the vegetative propagation of heterozygote plant individuals, as has been similarly observed in Peruvian cherimoya cultivation around Lima (see van Zonneveld et al. 2012). However, given that in Peru, more than 60% of the cacao trees are sexually propagated (Eskes 2011), alternative explanations of excess heterozygosity could be related to selection during the life cycle of the most heterozygous individuals who have more fitness than more homozygous individuals, self-incompatibility in small populations of outcrossing species, or a higher likelihood than random of reproduction between individuals that are genetically dissimilar to each other (Stoeckel et al. 2006). Such clusters may be severely threatened under fragmented conditions that would lead to a reduction of mating partners. Therefore, improved
understanding of the reasons for such increased heterozygosity may be required in order to better inform risk measure choice.

Both the accessibility and future climate change risk measures further support the notion that Criollo should be highly prioritized. Criollo trees are already grown in areas with high climate and precipitation seasonality, very distinct from the climates of the trees from the other cluster groups (Thomas et al. 2012). Therefore, Criollo germplasm may be of interest in supporting adaptation of other cacao populations to the more extreme climates that are expected to occur more often under progressive climate change (Hansen et al. 2012). Furthermore, the different rankings that these measures generated for some of the remaining cacao clusters reveal a range of potential prioritization strategies. For instance, the Guiana cluster is highly prioritized when future climatic change is taken into account, with the results suggesting that the accessibility and inbreeding measures are perhaps not very relevant risk measures for that particular cluster. By contrast, the accessibility and inbreeding risk measures led to a higher prioritization of the Amelonado cluster than when using the future climate change measure.

Hence, this leads to the question of which risk measure is the most appropriate for determining an appropriate conservation strategy. Temporal dimensions may also need to be taken into consideration, given that accessibility and inbreeding threats may imply the need for relatively urgent conservation interventions, whereas climate change is a longer-term risk which may instead initially require the need for monitoring the way in which existing varieties adapt to the changing environment.

Additionally, if we are to accept the notion that risk measures should include within-population information and/or capture as many risk factors as possible, it may be argued that prioritization should be based on the results that are generated using the combined risk measures (although these have less discriminatory power due to their combination of several single measures\(^4\)). As identified in the literature, the commonly used risk measures are the inbreeding and combined risk measures. Unfortunately, we cannot order these risk measures according to their predicted outcome on diversity, as the expected conserved diversity is calculated on the condition that the risk measure is a true prediction of the future. The widely different priority rankings generated by the different risk measures hence underscores the

\(^4\) At least as currently based on equally weighted individual measures. Alternative weightings might however be worth exploring in future research.
importance of making an appropriate choice between them. Under such circumstances, it may be that those risk measures that generate significantly different conservation priorities may be worth careful consideration; with their choice of application for prioritization tools being based on additional knowledge regarding biology, genetics and cultivation of the material in question.

The analysis of cacao based on LCA diversity measures resulted in the prioritisation of the Purús cluster, with the Criollo cluster changing from highest priority when using the original Weitzman diversity measure, to having the lowest priority when using the LCA measure. The Purus cluster, as identified by Motamayor et al. (2008), is also closer to wild cacao populations in the centre of cacao origin from which all the other cacao varieties are supposed to be derived (Thomas et al. 2012). If this is true, then the Purus cacao cluster may indeed have been subjected to a long history of human and natural selection, and have a high likelihood of containing many unknown traits, which make this cluster interesting for conservation as well as characterization for agronomic and commercially interesting traits.

By contrast, the Curaray cluster was ranked highly under both measures revealing that, it is both genetically distinct as well as having a high number of locally common alleles. This indicates that the Curaray cluster may well merit a high priority for conservation. Furthermore, these results imply that the choice of application of either the original or LCA diversity measures is likely to be dependent on the conservation goal. While the Weitzman original diversity measure prioritizes based on the genetic distinctness of specific allelic combinations (such as, in the present case, Criollo cacao trees), the LCA measure instead prioritizes genetic groups with unique high allele frequencies which may be associated with a long history of local human and/or natural selection of trees to specific local environmental conditions (see Frankel et al. 1995; van de Wouw et al. 2010).

In assessing the overall potential applicability of the Weitzman approach to PGR, we also note that in terms of actual implementation, conservation goals may be achieved by developing incentive mechanisms that are capable of compensating growers of Criollo cacao and other prioritised clusters for any opportunity costs they may incur from cultivating cacao varieties with lower yields and vigour (Cheesman 1944) relative to newly released ones. One such incentive mechanism could involve the promotion of high-value differentiation (Trognitz et al. 2011), although emerging payments for agrobiodiversity conservation services (PACS) schemes could complement such approaches. Such schemes have also been shown to be capable of taking cost-effectiveness vs. social equity trade-offs into account, for example by including consideration of participating farmer numbers and types (e.g. by gender or
livelihood status) as part of a multidimensional conservation goal (Narloch et al. 2011a, Narloch et al. 2011b), rather than just unit areas or plant populations as has been the case in the conservation fund allocation model used in this study. However, such allocation modelling is beyond the scope of the current study.

5. Conclusions and Future Research

This study has shown that the Weitzman approach, with an optimum conservation budget allocation mechanism, can indeed allow diversity comparisons to be made between the outcomes of different conservation strategies across plant genetic resource populations especially if appropriate measures of risk that are relevant to a range of phenomena and time-scales can be identified and adequate data exists to quantify such risks.

Nevertheless, the innovative application of the Weitzman approach to PGR diversity issues undertaken here raises a number of important questions regarding dimensionality, as for plants the choice of appropriate risk measure appears to remain less clear than that for animals, where inbreeding has been widely used as the key risk factor. The difference in results of our prioritization exercise using Weitzman and LCA diversity measures also raises questions regarding which diversity indicator to use. Choice of an appropriate diversity indicator may vary depending on what kind of diversity it is considered desirable to conserve, e.g. genetically unique varieties vs. “unknown” crop diversity, to name just two possibilities. Therefore, the choice of diversity and risk measures still leave a number of subjective decisions to be resolved regarding their appropriateness for the crop genetic resource and conservation objective in question. Nevertheless, from an economics point of view it would appear that a cost-effective conservation strategy can indeed by identified through the application of the Weitzman approach to PGR issues once the use of such non-economic diversity and risk measures have been agreed upon a priori. Thus, despite the dimensionality challenges, this simplification of the prioritization task considerably reduces the complexity of the decision-making procedure and therefore can be considered to facilitate conservation strategy design.

Nonetheless, the algorithm used in the Weitzman approach still faces challenges in analysing the large data sets typically associated with crop genetic resources, which tend to be much larger than for the animal genetic resources to which this approach has previously been applied. The main reason for this is that as the algorithm is recursive and the required storage capacity and computing time for a particular sample size \(N\) is approximately proportional to \(2^N\) (2 to the power of \(N\) represents the survival or extinction patterns that may occur), resulting in the Weitzman diversity
metric $D$ being demanding to calculate for sample sizes larger than 25 (Thaon d’Arnoldi et al, 1998). In other words, the size of the optimization problem can become prohibitive. As stated in the available Weitzmanpro package (Derban et al, 2002) used in this study, the algorithm is unable to compute the diversity for matrices with dimensions larger than 30 on account of the calculation time required. A matrix dimension of 30 generates a total set of $2^{30}$ or 1.07 billion combinations. In this study, the combinations for computation would have been $2^{940}$ in the absence of the clustered data sets. These are extremely large numbers of combinations and so far the highest sample size analysed in the literature has been 35 in Tapio et al (2006). Further technical details on how this limitation may be overcome are beyond the scope of this study but can clearly be identified as an area requiring further research.

An additional challenge is related to the fact that this study employed genetic ex-situ data which had been collected a long time ago and may not correctly reflect the current diversity at the collection sites. Hence, caution needs to be exercised with regard to case study conclusions based on the use of such data. Nevertheless, despite the fact that the analysis was carried out at a cluster level, the results presented here do provide further insights into how optimum conservation programs may be implemented for crop genetic resources.

We also note that although in this present cacao study we have used statistically well-defined genetic clusters that represent cultivars or cultivar groups to test the Weitzman approach, the methods used may also be applicable for clusters based on geographical origin. Such adaptability of the approach is important from an uptake perspective as for many in situ conservation contexts it may be more useful to prioritize specific geographic areas (e.g. countries, provinces, municipalities, ecological zones) or geographically restricted populations, based on overall crop diversity in those areas or populations, rather than prioritizing crop varieties per se for in situ conservation. This is because for several crops in the tropics, especially those that are underutilized, no registered varieties exist and crop production may include many local landraces and semi-domesticated populations, which makes it difficult to prioritize on the basis of different varieties.

Despite such challenges, it is apparent that even under the current limitations, the Weitzman approach can indeed help to provide practical recommendations for in situ conservation when the locations of the PGR clusters, species or varieties under study are known. Furthermore, given the sensitivity of the Weitzman approach to the scale of analysis (i.e. prioritization results will differ depending on whether the analysis has been carried out at local, national, regional
or global scales), a national level analysis, as opposed to the regional one carried out here, may well be worth undertaking in order to compare and contrast it with the actual conservation activities and strategies that currently form national biodiversity strategies and action plans (NBSAPs).
Acknowledgements

This publication and associated research was made possible through project funding associated with Bioversity International’s (CGIAR) Vavilov-Frankel Fellowship scheme (VFF). Maarten van Zonneveld is also grateful to the CGIAR Forest, Agroforestry and Tree Research program for financial support. We are also grateful to cacao expert Dapeng Zhang at the USDA Agricultural Research Service, Sustainable Perennial Crops Lab for comments on an earlier version of the paper.
References


Jarvis, A., J. Ramirez, J. Hanson, and C. Leibing. 2009. Interdependence in the face of climate change. in S. Fujisaka, Williams, D. and Halewood, M., editor. The impact of climate change on countries’ interdependence on genetic resources for food and agriculture. CGRFA.


MAPFORGEN. 2012.


Appendix A. Weitzman diversity and the relative contribution to diversity (when excluding the Criollo cluster)

<table>
<thead>
<tr>
<th>Clusters</th>
<th>Weitzman Diversity excluding each cluster ($D$)</th>
<th>Contribution of cluster in %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amelonado</td>
<td>5.156</td>
<td>9.64</td>
</tr>
<tr>
<td>Contamana</td>
<td>4.882</td>
<td>14.44</td>
</tr>
<tr>
<td>Curaray</td>
<td>4.772</td>
<td>16.36</td>
</tr>
<tr>
<td>Guiana</td>
<td>5.050</td>
<td>11.49</td>
</tr>
<tr>
<td>Iquitos</td>
<td>5.318</td>
<td>6.79</td>
</tr>
<tr>
<td>Marañon</td>
<td>5.383</td>
<td>5.66</td>
</tr>
<tr>
<td>Nacional</td>
<td>4.990</td>
<td>12.54</td>
</tr>
<tr>
<td>Nanay</td>
<td>5.140</td>
<td>9.92</td>
</tr>
<tr>
<td>Purús</td>
<td>5.086</td>
<td>10.86</td>
</tr>
<tr>
<td><strong>Total D</strong></td>
<td><strong>5.706</strong></td>
<td></td>
</tr>
</tbody>
</table>
Appendix B. Allocation of resources for conservation of cacao when excluding the Criollo cluster based on total extinction measure \((tz)\) and inbreeding \((in)\), respectively.