Combining molecular and landscape tools for targeting evolutionary processes in reserve design: An approach for islands

Raquel Vasconcelos1,2*, Orly Razgour3, Pedro Tarroso1,2, Mauro Fasola4, Salvador Carranza2, Paulo Célio Alves1,5

1 CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, InBIO Laboratório Associado, Universidade do Porto, Campus Agrário de Vairão, Vairão, Portugal, 2 Institute of Evolutionary Biology (CSIC-UPF, Consejo Superior de Investigaciones Científicas- Universitat Pompeu Fabra), Passeig Marítim de la Barceloneta, Barcelona, Spain, 3 Biological Sciences, University of Southampton, Life Sciences Building, Southampton, United Kingdom, 4 Dipartimento Scienze della Terra e dell’Ambiente, Università degli studi di Pavia, Pavia, Italy, 5 Departamento de Biologia da Faculdade de Ciências, Universidade do Porto, Porto, Portugal

Abstract

The importance of targeting ecological and evolutionary processes in reserve design has been widely acknowledged in the literature but rarely implemented on islands. Using Socotran reptiles as models, we aim to relate richness of widespread and restricted-range species directly with landscape variables and to compare the impact of setting conservation targets for lineages versus species. Socotra Island is a UNESCO Natural World Heritage Site, containing high levels of endemism in relation to its area, especially of reptiles, the vertebrates with the most comprehensive available genetic data. We predicted the occurrences of reptile species using distribution models and used a novel approach to interpolate maps of spatial phylogenetic patterns. Patterns of intra and interspecific diversity and differences between spatial outputs of lineage and species richness were related to eco-geographic variables. We evaluated differences in target achievement for each conservation unit within protected areas (PAs) under the current Zoning Plan (ZP) using gap and reserve design analyses. Although intraspecific richness was strongly correlated with interspecific richness, differences in their spatial distribution reached ~30% in some areas. Differences were more pronounced for wide-ranging than restricted-range taxa. Gap analysis indicates that most conservation units are under-represented in sanctuaries and that intra and interspecific richness were significantly higher outside PAs. This work will guide local-scale conservation planning as the ZP is due to be re-evaluated. This is one of the few studies on islands using genetic data from an entire class of vertebrates to incorporate lineage diversity in reserve design. This study provides an alternative methodological framework for supporting the use of landscape and genetic tools in reserve design, circumventing the use of phylogenetic distances and deterministic spatial interpolation of lineage diversity that can be widely applied to other systems.
Introduction

In face of extensive anthropogenic land cover changes, habitat loss, climate change and accelerating extinction rates, conservation planning needs new and improved tools to prioritize the allocation of scarce resources to maximize successful conservation efforts [1]. Prioritization of areas for conservation to date has mainly focused on representing biodiversity, i.e. including targeted portion of each conservation unit within a Protected Area, rather than enabling biodiversity persistence [2]. Representativeness has been achieved by spatial analysis of quantitative distribution data to design reserve networks that offer a comprehensive representation of biodiversity features (commonly species as conservation units).

The entire range of biodiversity cannot be sampled, and much remains unknown. Therefore, better known taxa or landscape variables are commonly used in reserve design to assess how well a planning unit (i.e. conservation area or grid cell) represents the biodiversity of a wider area [3,4]. Sampling all planning units is time consuming and costly, especially in remote areas. Species distribution models (SDMs), through relating location data to environmental variables, offer an important tool for overcoming the problem of geographically sparse and spatially biased distribution data, allowing to calculate representativeness more precisely [5]. To maximize persistence, the second aim of reserve design, knowledge on genetic diversity and geophysical variables should be taken into account to preserve the evolutionary potential of populations [6]. Genetic diversity is the basis for populations to adapt to environmental changes, disease outbreaks and other extinction risks [7]. Recognizing diversity below species level and incorporating it into modelling frameworks will enable conservation planners to identify areas where populations would potentially persist [6]. Incorporating knowledge on geophysical variables is important for quantifying landscape barriers to gene flow and meta-populations dynamics [6]. Understanding how landscape characteristics structure populations is essential for managing the genetic diversity of species and populations under threat through identifying mechanisms of isolation [8]. Contemporary landscape variables (e.g. land use) can play a role in shaping patterns of genetic diversity at fine geographic and temporal scales, while geological features and historical landscape changes (e.g. rock types) affect genetic diversity across broader spatial and temporal scales [9]. Genetic diversity acts as a reservoir for future adaptations and neutral evolution, and therefore preserving it will protect some of the processes that maintain and generate biodiversity [10]. On islands, the probability of loss of genetic diversity through drift and other neutral processes is generally higher, due to greater isolation, exposure to stochastic catastrophic events, and generally smaller population sizes [7,11]. Thus, identifying spatial patterns of genetic diversity and levels of differentiation among populations is important for guiding the assignment of conservation units and the planning of protected areas, particularly on islands [7]. However, identification of spatial patterns of intra-specific diversity is generally addressed through projecting the pairwise genetic distances between sampled individuals into mid-points between sample locations to interpolate maps of phylogenetic patterns [12]. The practice of mid-point projection is inadequate [13], because it relies on deterministic interpolation methods and assumes that pairwise genetic information is located halfway between the samples, which is unlikely to be the case especially on islands where habitat variables can vary substantially over short distances.

Genetic data can shape conservation priorities, and even predictions of protected areas (PAs) under climate change [14]. Yet, in practice, genetic diversity is often overlooked in biodiversity conservation planning [2], and reserve design studies accounting for intraspecific genetic variability in terrestrial systems are scarce, especially on island systems (e.g. [15,16]). Yet, islands host the highest levels of endemic diversity and extinction rates [11], and genetic variation is most unique, especially on large or remote islands [17]. Consequently, islands
represent a large proportion of the global biodiversity hotspots [18]. Some studies have addressed included intraspecific genetic variability through using measures of phylogenetic diversity (e.g. [19]), but this approach has been questioned in case of unresolved phylogenetic relationships of taxa, and because the link between extinction probability and functional diversity, as well as with phylogenetic diversity, is controversial [20]. On islands, it is common for very closely-related radiating taxa to occupy distinct functions in the ecosystem [21], exposing the inadequacies of this framework for island systems. Globally, the distribution of intraspecific genetic diversity follows similar patterns to species richness, with increased diversity in the tropics [22], and global biodiversity hotspots offer good representation of species evolutionary history, both at the clade and species level [23]. However, at finer spatial scales, areas of high species richness are not always concordant with hotspots of genetic diversity. Therefore setting conservation priorities to maximize species representation may not be adequate for preserving genetic diversity within species [24].

The goal of the present work is to address prioritization in reserve design by combining molecular and spatial analyses circumventing the use of phylogenetic diversity and avoiding the caveats of deterministic methods in the interpolation of phylogenetic patterns. This approach was tested on the reptiles of Socotra Island, a highly suitable model system due to the characteristics of the geographic location and taxonomic group. The Socotra Archipelago (a Yemeni governorate; Socotra hereafter) is one of the most inaccessible and biodiversity rich areas worldwide [25]. It comprises four islands, of which Socotra Island is the main one, with maximum length of approximately 130 km, width of 40 km, and an area of circa 3800 km$^2$ (Fig 1). Located in the Horn of Africa biodiversity hotspot, it is a UNESCO World Heritage Natural site since 2008 and considered one of the world’s most irreplaceable protected areas for conservation [26]. Yet, it is also a key under-surveyed area, with limited coverage of genetic data both in terms of data availability and taxonomic coverage [22]. Its conservation Zoning Plan (ZP) classified some areas as sanctuaries (Fig 1; 1.9% of the island’s surface) following the Biosphere Reserve Concept, and allocated other areas (26.7%) where exploitation of resources is fully or partially permitted [27]. Studies have called for the ZP to be revised to take into consideration new distribution and systematic data [28,29]. A UNEP-GEP project carried out in collaboration with the local government, has as one of its main goals “improved design and management effectiveness of the existing network of PAs”. Therefore this study will serve as the basis for selecting the best new areas to protect reptile diversity on Socotra Island. In this study we aim to: 1) identify and compare spatial patterns of species and lineage richness, determining the effect of landscape variables on these patterns; 2)

![Fig 1. Study area. Map of the study area showing the geographic location of the Socotra Archipelago, (a, b) and the type of protection in the Zoning Plan designations over an elevation map (c). Geographic Coordinate System, Datum WGS84, Units in decimal Degrees.](https://doi.org/10.1371/journal.pone.0200830.g001)
assess the adequacy of the existing ZP for conserving reptile diversity; and 3) compare reserve designs based on species and lineage diversity. With this novel framework we show the importance of incorporating intraespecific diversity in reserve design.

Methods

Study area

Socotra Island (12°30'; N 54° E) is located in the Arabian Sea, about 250 km east from the Horn of Africa (Fig 1). Its climate is governed by the Intertropical Convergence Zone and related monsoon cycles. Its vegetation is mainly associated to arid ecosystems, with some differences between the limestone and granitic areas [30].

Socotra has a continental origin from a Gondwana fragment that split from Oman as a result of the opening of the Gulf of Aden 18 million years ago [31]. The granitic complex on Socotra Island, that includes the Haggeher mountains, was always above sea-level (Fig 1). But the limestone plateaus were formed during the Paleo-Eocene and were below sea-level during interglacial transgressions on the western side of the island [32]. Reptiles constitute the most important vertebrate fauna of Socotra. While there are no amphibians, probably only one endemic mammal, and 10 endemic bird species, all 29 native reptile species are endemic to the archipelago, 23 to Socotra Island. The reptile community of Socotra was recently barcoded (using the mitochondrial cytochrome c oxidase 1, COI, region), providing data on their intra-specific genetic diversity [28].

Sampling

Genetic samples and distribution data of the 23 reptile species were collected during expeditions to Socotra Island between 2007 and 2014 (see Vasconcelos et al. 2016 [28] for further details). Using a 10x10 km square grid covering Socotra island, 103 stations (detailed in S1 Table in Supporting Information) were sampled with transects gathering 1299 genetic samples (tail tips) and 2597 presence records (one sample/record per individual). All samples used in this study were collected with appropriate permits from local authorities.

Phylogenetic analyses

Phylogenetic analyses were carried out to determine the number of lineages present in each species based on Sanger sequencing data. Lineages were defined as a sequence or group of sequences within a species which optimized set of nodes defined the transition to inter-specific processes based on the Bayesian implementation of the General Mixed Yule Coalescent Model (bGMYC), detailed below in the lineage delimitation analyses. The datasets used for this analysis consisted of 23 alignments (one per species) of 663 base pairs of the COI gene (number of sequences per species detailed in S2 Table), downloaded from GenBank (accessions KU567304–KU567683; details in Vasconcelos et al. 2016 [28]). This marker was shown to detect intra-specific diversity and to have a good phylogenetic signal despite its reduced length [28]. Phylogenetic analyses could not be performed on three of the species [Myriopholis filiformis (Boulenger 1899), Pachyclamalus brevis Günther 1881, and Xerotypilops socotranus (Boulenger 1889); see [33] for all naming authorities] because they were represented by a single haplotype. After removing identical sequences, alignments were performed with the online application of MAFFT v.7 (http://mafft.cbrc.jp/alignment/server/). Removal of identical haplotypes is necessary because they may cause problems with downstream bGMYC analyses [34]. Best-fit models of sequence evolution were inferred in jModeltest v.2.1.3 [35] with the Akaike Information Criterion (AIC) for model selection.
An ultrametric tree was produced for each alignment using Bayesian Inference implemented in BEAST v1.8.0 [36] to feed the bGMYC analyses, because it allowed to accommodate the uncertainty associated with the estimation of a phylogeny [37]. The bGMYC R package [37] was subsequently used to determine the number of deep mitochondrial lineages per species using an R-library version 3.3 [38]. Two individual runs of 5x10^7 generations were carried out, sampling at intervals of 10,000 generations. Preliminary tests were performed to study models and priors. Models and prior specifications applied were as follows (otherwise by default): model of sequence evolution as calculated by jModeltest [35]; constant size coalescent tree prior; uniform strict clock prior (0–100); random starting tree; base substitution prior Uniform (0,100); alpha prior Uniform (0,10) [39]. Posterior trace plots and effective sample sizes (ESS) of the runs were monitored in Tracer v1.5 [40] to ensure convergence. The results of the individual runs were combined in LogCombiner discarding 10% of the samples and the maximum clade credibility (MCC) ultrametric tree was produced with TreeAnnotator (both provided with the BEAST package). All information related to each species alignment including model selected, number of variable and parsimony-informative positions, and the results of bGMYC are presented in S2 Table.

Lineage delimitation analysis

Deep mitochondrial lineages within Socotran reptiles were objectively identified and delimited using the General Mixed Yule-Coalescent Model implemented in a Bayesian framework (bGMYC) [41,42]. This model detects the most likely tree depth (in the single-threshold approach) at which the pattern of tree branching shifts between a Yule process (reflecting inter-specific phylogenetic structure) and a coalescent process (reflecting intraspecific structure). Because inference of lineages relies on point estimates of the topology and branch lengths, the associated phylogenetic error could decrease the accuracy of delimitation results. Therefore, we assessed uncertainty in phylogenetic tree estimation and model parameters with a Bayesian implementation of the GMYC model, which integrates over these potential sources of error via MCMC simulation [37]. We used the R-package bGMYC to calculate marginal posterior probabilities of lineage limits from the posterior distribution of ultrametric trees reconstructed with BEAST. A post-burn-in sample of 250 trees resampled from that posterior was used to calculate the posterior distribution of the GMYC model. We ran the bGMYC analysis for 100,000 generations with a burn-in of 10,000 generations.

Spatial modelling and analysis

SDMs were generated for 20 of the 23 reptile species of Socotra Island, using Maxent v3.3.3 [43]. The initial dataset included a total of 1774 presence records and 14 landscape variables (LVs; S3 Table). Presence location records were obtained from bibliographic sources and fieldwork observations (detailed in Vasconcelos et al. 2016 [28]), and were mapped in ArcGIS v10 (ESRI). LVs were extracted from previous studies [30,32,44], freely available websites (http://srtm.csi.cgiar.org) or derived from other LVs (detailed in S3 Table). LVs data came from Digital Elevation Models, local weather stations, geological maps, and habitat cover data at a spatial resolution of approximately 100 m (0.0009°) to match the resolution of the location records. All habitat variables are presented as percent cover in each cell, except for linear water features (wadis), for which the Euclidean distance of each cell from the closest wadi was calculated using the ‘Euclidean Distance’ tool in ArcGIS.

A total of 10 model replicates were run with random seed which allows a different random training/testing data partition in each run. For species with <20 location records, we ran either eight (for N<20) or four (N<10) replicates. In each run, 30% of the data were set aside for
model testing using bootstrapping. After testing test different values of regularization and number of features, models were run with the default settings and features. The area under the curve (AUC) of the receiver-operating characteristics (ROC) plot was used to measure model fit (good if >0.75). Standard deviation between individual model replicates was used as an indication of prediction uncertainty. The final models included eight LVs with the lowest absolute correlation scores (R≤0.75) and 657 presence records (available at figshare DOI: 10.6084/m9.figshare.4978223) after attempting to minimize the autocorrelation by removing locations that were clustered in the tridimensional LVs PCA space using the ‘near 3D’ function in ArcGIS. SDMs could not be generated for three of the 23 species (Myriopholis filiformis, M. wilsoni (Hahn 1978), and Xerotyphlops socotranus) because they had too few number of location records (≤5), therefore only grids with location records were used.

The minimum training threshold was used to transform probability of occurrences into binary maps, representing the predicted suitable range of each species. This thresholding uses all training points and is preferred when data quality is high. Moreover, when designing PAs, it is more appropriate to use a less restrictive threshold [45]. We calculated the true skill statistics (TSS) as an alternative measure of accuracy of the binary models. TSS ranges from -1 to 1 and if values are higher than 0 indicates a performance better than random [46]. To determine the suitable range of each lineage for species with more than one lineage, PHYLIN R-package [13] was used to spatially interpolate genetic distances from phylogeographic data with kriging [13], circumventing the projection of pairwise genetic distances into mid-points between sample locations [13]. The output of PHYLIN was transformed into binary data, using 0.5 probability as threshold, and intersected with the binary SDMs of the respective species to keep only areas with high probability of occurrence in lineage binary models. In total, 45 lineage suitability maps were obtained. Those, together with the species binary SDMs, comprise the dataset for 57 biodiversity features. SDM outputs were overlaid in ArcGIS to generate richness maps for all species, all lineages, as well as wide-range and restricted-range biodiversity features (that occupy less than 30% of the Socotra Island area, following recommendations in [47]), resulting in a total of six layers of relative richness. Differences between species and lineage richness output maps (all species/lineages, wide-range and restricted-range species/lineages) were calculated in ArcGIS. Relative richness maps were reclassified into binary SDMs using as threshold the upper quartiles values (75th percentiles). To identify landscape variables associated with high levels of species/lineage diversity, the diversity values of each grid cell in the richness layers for total, wide-range and restricted-range features were directly correlated against the values of the same grid cells for the 14 available LVs, as well as latitude and longitude using the Pearson correlation coefficient (ρ) in R [38].

**Gap analysis**

A gap analysis was carried out to assess how much the different categories of the ZP encompass biodiversity features and richness [48,49]. The ZP map [27] was georeferenced and the four categories were digitized into polygons. These categories present increasing levels of protection: 0 = general use, allowing a significant level of habitat modification; 1 = resource use reserve, resource use hereafter, allowing a moderate level of sustainable habitat modification; 2 = national park, allowing sustainable use of natural resources; 3 = nature sanctuaries, sanctuaries hereafter, requiring special permits to enter (Fig 1). Category 1 to 3 were considered as Protected Areas. This layer was converted into a grid with the same spatial resolution as the SDMs (~100 m) and intersected with the 57 binary SDMs of all biodiversity features in ArcGIS to calculate the number of cells that fall within each ZP category. Percent inclusion of each species/lineage suitable range within each category was calculated through dividing the number of
cells within each category by the total number of grid cells in the binary SDM. This analysis was repeated for all species, lineages and restricted-range biodiversity features richness SDMs. Richness of wide-range biodiversity features was not included in the gap analysis because of their lower conservation value.

Representation targets (percentage of suitable range that should fall within PAs) were calculated according to biodiversity feature attributes, including threatened status, level of protection, and distribution ranges. Following the approach used in previous studies [50–52], Critical Endangered (CR) species should be entirely within a PA and higher proportions of representation should be achieved for restricted-range biodiversity features. A minimum target of 10% was set for the feature with the widest range and 100% for the narrowest, then targets of features with intermediate geographic ranges were log-linearly interpolated between these extremes. Only one of the study species is threatened, *Hemidactylus dracaenacolus* Rösler & Wranik 1999, and classified as Critically Endangered (CR), so a 100% target was set for this species. Four additional restricted-range species are considered Data Deficient (DD), and for those a plus of 30% of representation target was also set.

We did not consider phylogenetic diversity because most of the phylogenetic relationships of Socotran reptiles with other taxa are still unresolved. Additionally, examples of niche divergence and ecological diversification were already detected among closely-related reptiles in Socotra [53,54], suggesting a detach between branch lengths and functional diversity that would hamper a phylogenetic diversity approach. Biodiversity features were considered to be protected if percent cover within PAs was higher than the set representation target. Representation targets for relative richness of total species/lineages and restricted-range features were set as 100% of the 75th percentiles, and were considered to fail target achievement if below this threshold. Target achievement for each biodiversity feature was calculated as a percentage of the sum of the proportion of cells within PAs in respect to the representation target.

### Reserve design

Two reserve design analyses were carried out separately for species and lineages, using Zonation v4.0 [55], a software designed for spatial conservation prioritization that is frequently used in reserve design studies [56]. This software uses a gradient-like iterative heuristic to produce a sequential removal of grid cells throughout the ZP for achieving optimal solutions [55]. Target-based planning was chosen as removing rule because the goal was to find the best solution in which the maximum number of lineages met different conservation targets for both runs. In order to generate spatial aggregation within the solution, the rule ‘only remove from edges’ was selected for both runs. The Boundary Length Penalty, which devalues reserve structures with a high extent of edge, was chosen as the method for inducing reserve network aggregation, because some biodiversity features have fragmented distributions and it is one of the best methods for generating aggregation for management purposes [55]. The warp factor was set to one with an aggregation level of 0.4 based on several runs with different values. A cost layer was added to both analyses, so that grid cells with main roads and urban areas (with a buffer radius of 112 m or 1 km, respectively) were given a high cost when selected (100), secondary roads and resource use areas a middle cost (50), national parks a cost of 25 and sanctuaries a minimum cost (1).

Grid cells with highest rank have highest conservation value. The minimum set of grid cells with higher rank in the final solutions, which assured that all species or lineages were represented and achieved the targets, were selected for reclassifying the solution into binary files.
Results

Phylogenetic and spatial analyses

In total, we identified 45 lineages (mean = 2.0 ± 1.3 lineages/species; range: 1–6). The species with highest number of lineages (6) was *Hemidactylus pumilio* (Boulenger 1899), while 11 of the 23 species only had one lineage (S2 Table). The number of lineages identified per species with bGMYC was always within the range of number of entities identified using different distance and tree-based techniques in Vasconcelos et al. (2016) [28]. Theoretical variogram models fitted heuristically to the empirical variogram obtained from PHYLIN showed the best representation of the genetic structuring in space.

SDMs performed well and had high discrimination and predictive abilities (mean AUC\(_{\text{train}}\) = 0.90 ± 0.06 [range: 0.73–0.99], mean AUC\(_{\text{test}}\) = 0.83 ± 0.006 [range: 0.73–0.98]) and low standard deviations (Table 1). The TSS mean values (S4 Table) are also good (mean TSS = 0.40), even for restricted-range taxa (mean TSS = 0.48). The LVs with highest contributions to all species’ SDMs were shrubland, slope and woodland (Table 1). The LVs with higher contributions to the SDM of the CR gecko *H. dracaenacolus* are temperature, total solar radiation and slope, and to the DD trogonophid *P. brevis* are distance to wadis and garrigue cover. The LVs with higher contributions to the SDMs of the remaining species are highlighted in Table 1, and LVs correlated with richness are detailed in Table 2 (see following section).

Table 1. Spatial modelling results. Total number of presence data of endemic reptiles of Socotra Island used in the species distribution models (SDMs), including the training (train) and test datasets, average and standard deviation (SD) of AUC for SDM replicates, and average percent contribution of each landscape variable (LV) to SDMs. LVs contributing >20% to SDMs are highlighted in bold.

| Species Code       | Presence data | AUC | LV contribution |
|--------------------|---------------|-----|-----------------|
|                  | Train | Test | Train | Test | SD | gar | mea | shr | slo | tem |trs | wad | woo |
| CHmo               | 18    | 13   | 9.1   | 0.9  | 0.09 | 2.7 | 10.4 | 5.3 | 7.1  | 3.1 | 21.1 | 34.5 | 27.0 | 4.2 | 1.4 |
| Divi               | 14    | 10   | 9.1   | 0.9  | 0.09 | 2.7 | 10.4 | 5.3 | 7.1  | 3.1 | 21.1 | 34.5 | 27.0 | 4.2 | 1.4 |
| HAr1               | 18    | 13   | 9.1   | 0.9  | 0.09 | 2.7 | 10.4 | 5.3 | 7.1  | 3.1 | 21.1 | 34.5 | 27.0 | 4.2 | 1.4 |
| HAt1               | 14    | 10   | 9.1   | 0.9  | 0.09 | 2.7 | 10.4 | 5.3 | 7.1  | 3.1 | 21.1 | 34.5 | 27.0 | 4.2 | 1.4 |
| Hksi              | 34    | 24   | 9.1   | 0.9  | 0.09 | 2.7 | 10.4 | 5.3 | 7.1  | 3.1 | 21.1 | 34.5 | 27.0 | 4.2 | 1.4 |
| HMso              | 14    | 10   | 9.1   | 0.9  | 0.09 | 2.7 | 10.4 | 5.3 | 7.1  | 3.1 | 21.1 | 34.5 | 27.0 | 4.2 | 1.4 |
| HEdr1             | 18    | 13   | 9.1   | 0.9  | 0.09 | 2.7 | 10.4 | 5.3 | 7.1  | 3.1 | 21.1 | 34.5 | 27.0 | 4.2 | 1.4 |
| HEgr1             | 14    | 10   | 9.1   | 0.9  | 0.09 | 2.7 | 10.4 | 5.3 | 7.1  | 3.1 | 21.1 | 34.5 | 27.0 | 4.2 | 1.4 |
| HEho              | 34    | 24   | 9.1   | 0.9  | 0.09 | 2.7 | 10.4 | 5.3 | 7.1  | 3.1 | 21.1 | 34.5 | 27.0 | 4.2 | 1.4 |
| HEin1             | 14    | 10   | 9.1   | 0.9  | 0.09 | 2.7 | 10.4 | 5.3 | 7.1  | 3.1 | 21.1 | 34.5 | 27.0 | 4.2 | 1.4 |
| HEpu1             | 14    | 10   | 9.1   | 0.9  | 0.09 | 2.7 | 10.4 | 5.3 | 7.1  | 3.1 | 21.1 | 34.5 | 27.0 | 4.2 | 1.4 |
| MEba1             | 34    | 24   | 9.1   | 0.9  | 0.09 | 2.7 | 10.4 | 5.3 | 7.1  | 3.1 | 21.1 | 34.5 | 27.0 | 4.2 | 1.4 |
| MYFi1             | 14    | 10   | 9.1   | 0.9  | 0.09 | 2.7 | 10.4 | 5.3 | 7.1  | 3.1 | 21.1 | 34.5 | 27.0 | 4.2 | 1.4 |
| MMYma1            | 14    | 10   | 9.1   | 0.9  | 0.09 | 2.7 | 10.4 | 5.3 | 7.1  | 3.1 | 21.1 | 34.5 | 27.0 | 4.2 | 1.4 |
| MYWi1             | 14    | 10   | 9.1   | 0.9  | 0.09 | 2.7 | 10.4 | 5.3 | 7.1  | 3.1 | 21.1 | 34.5 | 27.0 | 4.2 | 1.4 |
| PAbri1            | 14    | 10   | 9.1   | 0.9  | 0.09 | 2.7 | 10.4 | 5.3 | 7.1  | 3.1 | 21.1 | 34.5 | 27.0 | 4.2 | 1.4 |
| PRgu1             | 14    | 10   | 9.1   | 0.9  | 0.09 | 2.7 | 10.4 | 5.3 | 7.1  | 3.1 | 21.1 | 34.5 | 27.0 | 4.2 | 1.4 |
| PRin1             | 14    | 10   | 9.1   | 0.9  | 0.09 | 2.7 | 10.4 | 5.3 | 7.1  | 3.1 | 21.1 | 34.5 | 27.0 | 4.2 | 1.4 |
| PPrd1             | 14    | 10   | 9.1   | 0.9  | 0.09 | 2.7 | 10.4 | 5.3 | 7.1  | 3.1 | 21.1 | 34.5 | 27.0 | 4.2 | 1.4 |
| PRob1             | 14    | 10   | 9.1   | 0.9  | 0.09 | 2.7 | 10.4 | 5.3 | 7.1  | 3.1 | 21.1 | 34.5 | 27.0 | 4.2 | 1.4 |
| PReso1            | 97    | 68   | 9.1   | 0.9  | 0.09 | 2.7 | 10.4 | 5.3 | 7.1  | 3.1 | 21.1 | 34.5 | 27.0 | 4.2 | 1.4 |
| TRso1             | 66    | 47   | 9.1   | 0.9  | 0.09 | 2.7 | 10.4 | 5.3 | 7.1  | 3.1 | 21.1 | 34.5 | 27.0 | 4.2 | 1.4 |

*Non-modelled taxa; gar = garrigue; mea = meadow; shr = shrubland; slo = slope; tem = temperature; trs = total solar radiation; wad = distance to wadis; woo = woodland (S3 Table).*

https://doi.org/10.1371/journal.pone.0200830.t001
On average 37.0±26.5% (range: 0.7–89.2%) of Socotra Island was predicted to be suitable for the different reptile species and lineages (Table 3). The species with the smallest predicted suitable range is *H. dracaenacolus*, while *Mesalina balfouri* (Blanford, 1881) has the largest predicted range (Table 3).

### Spatial pattern of species and lineage diversity

Even though total species richness was strongly positively correlated with lineage richness (*ρ* = 0.93, *P*<0.0001), differences between outputs reached circa 30% in some areas (Fig 2). The area with the highest overall species richness was located in the Diksam Mountains, whilst highest levels of overall lineage richness were found in the northeast and east of the island (Fig 2). Slope was most frequently the highest correlated LV with levels of richness, followed by ruggedness (highly correlated with slope) and total solar radiation (Table 1). The richest areas for both total species and lineages were positively associated with solar radiation (*ρ* = 0.32 and 0.22, respectively, *P*<0.0001 for both), and negatively associated with garrigue-like habitat (basal alluvial grasslands and shrublands with *Dactylolentium robecchii*, *Justicia rigida* and *Croton socotranus*, including mangroves; garrigue hereafter), ruggedness, and slope (*ρ* = -0.17 and -0.26, -0.34 and -0.24, -0.40 and -0.29, respectively, *P*<0.0001 for all; Table 1). The areas with the highest differences between species and lineage richness maps were located around the Zahr Basin, in the central plateaus, where relative species richness was considerably higher than lineage richness, and in the extreme east tip of the island, where relative lineage richness was higher (Fig 2). Differences in outputs were strongly negatively associated with longitude (*ρ* = -0.51, *P*<0.0001), and, unlike outputs for total species and lineages, also to shrubland (*ρ* = -0.27, *P*<0.0001), and positively associated with geology (*ρ* = 0.28, *P*<0.0001) and evapotranspiration (*ρ* = 0.26, *P*<0.0001; Table 1).

Spatial differences were much stronger when considering wide-range features than when considering restricted-range features (maximum of 43% versus 17%, respectively; Fig 2). While wide-range features showed similar spatial patterns as total features, restricted-range features showed the greatest differences between species and lineage richness in the Haggeher and Diksam Mountains, where relative species richness was considerably higher, and on the eastern part of the island, where the opposite pattern was evident. Restricted-range diversity (for species and lineages, respectively) was 5 to 7 folds more positively correlated with moisture and precipitation, and negatively correlated with temperature, shrubland and distance to wadis (for the different reptile species and lineages; Table 3).

#### Table 2. Correlation tests

Correlation values (*ρ*) of richness of total, wide-range and restricted-range species/lineages, and of the differences between species and lineage richness, with latitude (lat), longitude (lon) and 14 landscape variables (described in S3 Table). Absolute *ρ*>0.20 are highlighted in bold. All correlations were significant (*P*<0.05).

| Richness              | lat  | long | alt  | ept  | gar  | geo  | mea  | mix  | pre  | rug  | shr  | slo  | tem  |trs  | wad  | woo  |
|-----------------------|------|------|------|------|------|------|------|------|------|------|------|------|------|-----|------|------|
| Total species         | 0.02 | 0.02 | 0.07 | 0.18 | -0.17 | 0.03 | 0.16 | 0.00 | 0.04 | -0.34 | -0.09 | -0.40 | -0.03 | 0.32 | 0.04 | 0.11 |
| Total lineages        | 0.11 | 0.22 | 0.06 | 0.09 | -0.26 | -0.07 | 0.12 | 0.04 | 0.08 | -0.24 | 0.01  | -0.29 | -0.07 | 0.22 | 0.07 | 0.17 |
| Total difference      | -0.23 | -0.51 | 0.04 | 0.26 | 0.22 | 0.28 | 0.12 | -0.12 | -0.08 | -0.32 | -0.27 | -0.33 | 0.08  | 0.29 | -0.08 | -0.14 |
| Wide-range species    | 0.01 | -0.01 | 0.03 | 0.19 | -0.17 | 0.02 | 0.16 | -0.06 | -0.01 | -0.29 | -0.06 | -0.36 | 0.02  | 0.26 | 0.10 | 0.07 |
| Wide-range lineages   | 0.11 | 0.22 | 0.03 | 0.09 | -0.27 | -0.10 | 0.12 | 0.01 | 0.04 | -0.19 | 0.04  | -0.25 | -0.03 | 0.17 | 0.11 | 0.15 |
| Wide-range difference | -0.20 | -0.47 | 0.02 | 0.30 | 0.13 | 0.24 | 0.15 | -0.16 | -0.11 | -0.34 | -0.24 | -0.37 | 0.12  | 0.30 | 0.01 | -0.13 |
| Restricted-range species | 0.04 | 0.11 | 0.16 | 0.03 | -0.05 | 0.07 | 0.09 | 0.18 | 0.21 | -0.32 | -0.15 | -0.33 | -0.22 | 0.34 | -0.19 | 0.19 |
| Restricted-range lineages | 0.05 | 0.09 | 0.16 | 0.05 | -0.06 | 0.08 | 0.09 | 0.16 | 0.19 | -0.32 | -0.16 | -0.32 | -0.20 | 0.34 | -0.18 | 0.18 |
| Restricted-range difference | -0.01 | 0.11 | 0.03 | -0.06 | 0.00 | -0.03 | -0.02 | 0.12 | 0.10 | -0.08 | 0.01  | -0.07 | -0.11 | 0.06 | -0.09 | 0.09 |

gar = garrigue; mea = meadow; shr = shrubland; slo = slope; temp = temperature; tsr = total solar radiation; wad = distance to wadis; woo = woodland

https://doi.org/10.1371/journal.pone.0200830.t002
## Table 3. Gap analysis results.

The number (nr) of grid cells of species (sp) and lineages (ln), in total, in comparison with the island area (range), and within each category of the Zoning Plan, as well as species conservation status (status), representation targets and target achievement (≥100% if above target) of the 57 biodiversity features and the fourth quantile of richesses is showed. Species codes follow Table 1 and lineages are identified by a number. Restricted-range species and their lineages (restr.) are shaded in grey.

| Biodiversity | feature | ln | status | total (%) | target 0 = general use (%) | target 1 = resource use (%) | target 2 = national park (%) | target 3 = sanctuaries (%) | target achievement (nr) (%) |
|--------------|---------|----|--------|-----------|--------------------------|-----------------------------|-----------------------------|-----------------------------|---------------------------|
| CHmo         | 1 1     | NT | 292535 | 85.3      | 10.3                     | 2947                        | 73689                       | 25.2                        | 210830                    | 5069 1.7                  | 289588 961.9              |
| Dlvi         | 1 0     | LC | 136499 | 39.8      | 15.9                     | 1434                       | 22567                       | 16.5                        | 109478                    | 80.2                      | 3020 2.2                  | 135065 623.4              |
| Dlvi_1       | 0 1     |    | 89221  | 26.0      | 19.5                     | 1372                       | 13703                       | 15.4                        | 71175                     | 79.8                      | 2971 3.3                  | 87849 504.6               |
| Dlvi_2       | 0 1     |    | 104923 | 30.6      | 18.1                     | 1434                       | 17771                       | 16.9                        | 83747                     | 79.8                      | 1971 1.9                  | 103489 545.8              |
| HArı         | 1 0     | LC | 276256 | 80.6      | 10.7                     | 2834                       | 60715                       | 22.0                        | 207751                    | 75.2                      | 4956 1.8                  | 273422 927.0              |
| HArı_1       | 0 1     |    | 201130 | 58.7      | 12.9                     | 2834                       | 54848                       | 27.3                        | 139228                    | 69.2                      | 4220 2.1                  | 198296 764.1              |
| HArı_2       | 0 1     |    | 209015 | 61.0      | 12.6                     | 2455                       | 14725                       | 21.7                        | 53721                     | 75.9                      | 2287 3.2                  | 70733 460.7               |
| HEdr         | 1 1     | CR | 2275   | 0.7       | 100.0                    |                            |                             |                             |                            |                            |                            | 2275 100.0                |
| HEdr         | 1 1     | NT | 4623   | 1.3       | 90.4                     |                            |                             |                             |                            |                            |                            | 4623 110.6                |
| HEdr         | 1 0     | NE | 276100 | 80.5      | 10.7                     | 2915                       | 66317                       | 24.0                        | 201666                    | 73.0                      | 5202 1.9                  | 273185 926.4              |
| HEdr         | 1 0     | LC | 136499 | 39.8      | 15.9                     | 1434                       | 22567                       | 16.5                        | 109478                    | 80.2                      | 3020 2.2                  | 135065 623.4              |
| HEdr         | 1 0     | LC | 136499 | 39.8      | 15.9                     | 1434                       | 22567                       | 16.5                        | 109478                    | 80.2                      | 3020 2.2                  | 135065 623.4              |

(Continued)
wadis than wide-range diversity (Table 1). On the other hand, wide-range diversity was more strongly negatively correlated with garrigue than restricted-range diversity (Table 1).

Table 3. (Continued)

| Biodiversity feature | sp total | ln total | sp restr. | ln restr. | target range | target achievement |
|----------------------|----------|----------|-----------|-----------|---------------|--------------------|
|                      | total (%) | total (%) | total (%) | total (%) | 0 = general use | 1 = resource use | 2 = national park | 3 = sanctuaries | 0 = general use | 1 = resource use | 2 = national park | 3 = sanctuaries |
| PRid                 | 4329     | 97.8     | 0         | 0         | 2647          | 61.1              | 1682               | 38.9           | 4329          | 102.3          |
| PRin                 | 250861   | 73.2     | 11.3      | 1992      | 5022          | 20.1              | 193202             | 77.0           | 5145          | 214              | 248869 875.2   |
| PRin_1               | 220379   | 64.3     | 12.2      | 1992      | 50245         | 22.8              | 163958             | 74.4           | 4184          | 1.9            | 218387 809.3   |
| PRin_2               | 75194    | 21.9     | 21.1      | 224       | 15328         | 20.4              | 5763               | 76.6           | 209          | 2.7            | 74970 472.3    |
| PRin_3               | 132040   | 38.5     | 16.1      | 219       | 31207         | 23.6              | 98746              | 74.8           | 1868          | 1.4            | 131821 618.5   |
| PRob                 | 140257   | 40.9     | 15.7      | 2704      | 60278         | 43.0              | 75177              | 53.6           | 2098          | 1.5            | 137553 626.5   |
| PRso                 | 268762   | 78.4     | 10.9      | 2800      | 65578         | 24.4              | 194858             | 72.5           | 5526          | 2.1            | 265962 911.0   |
| PRso_1               | 86988    | 25.4     | 19.7      | 2470      | 21573         | 24.8              | 59679              | 68.6           | 3264          | 3.8            | 84518 492.1    |
| PRso_2               | 160917   | 46.9     | 14.6      | 2266      | 35332         | 22.0              | 119640             | 74.3           | 3679          | 2.3            | 158651 676.4   |
| PRso_3               | 159275   | 46.5     | 14.7      | 330       | 47533         | 29.9              | 109545             | 68.8           | 1847          | 1.2            | 158945 681.0   |
| TRso                 | 248971   | 72.6     | 11.4      | 2865      | 67633         | 27.2              | 173417             | 69.7           | 5056          | 2.0            | 246106 868.0   |
| TRso_1               | 236304   | 68.9     | 11.8      | 1733      | 67301         | 28.5              | 165445             | 69.2           | 3825          | 1.6            | 234571 844.7   |
| TRso_2               | 180283   | 52.6     | 13.7      | 2621      | 36674         | 20.3              | 136050             | 75.5           | 4938          | 2.7            | 177662 718.7   |
| XEsso *              | 5        | 0.0      | 100.0     | 0         | 1.0           | 20.0              | 4                  | 80.0           | 0            | 0.0            | 5 100.0         |

* Non-modelled taxa. LC, Least Concern; CR, Critically Endangered; NT, Near Threatened; NE, Not Evaluated; DD, Data Deficient.

https://doi.org/10.1371/journal.pone.0200830.t003

Fig 2. Patterns of biodiversity features. Spatial patterns of reptile diversity on Socotra Island, including maps of relative percent of species and lineage richness for the total number of biodiversity features (maximum per grid cell = 18 and 30, respectively), wide-range (maximum per grid cell = 13 and 26, respectively) and restricted-range taxa (maximum per grid cell = 5 and 6, respectively), as well as the percent differences between the species and lineage richness outputs. Geographic Coordinate System, Datum WGS84, Units in decimal Degrees.

https://doi.org/10.1371/journal.pone.0200830.g002
Differences in wide-range outputs followed the same association trends with LVs as differences between total species and lineages, while differences in restricted-range outputs were not significantly associated with either one of the LVs (Table 1).

**Gap analysis and reserve design**

Almost all biodiversity features, except for one restricted-range species, *Pristurus insignoides* (Arnold 1986) and two species with higher levels of threat, the CR *H. dracaenacolus*, and the Near Threatened (NT) *Hemidactylus granti* (Boulenger 1899)–had very low (below 5%) representation within sanctuaries (Table 2). Twelve biodiversity features were better represented outside PAs than in sanctuaries [lineage 1 and 5 of *Hemidactylus inintellectus* (Sindaco, Zilian, Razzetti, Pupin, Grieco 2009), lineage 1 to 3 of *Hemidactylus pumilio*, lineage 2 of *Hakaria simonyi* (Steindachner 1889), *Hemorophis socotrae* (Günther 1881), *Myriopholis macrura* (Boulenger 1899) and its lineages, *Pachycalamus brevis*, and *Pristurus obsti* (Rössler & Wranik 1999)]. One Data Deficient species, *Pachycalamus brevis*, did not reach its conservation targets within PAs.

For overall and restricted-range diversity, species and lineage richness was significantly higher outside than inside PAs (Generalized Linear Models: all P-values <0.0001; Fig 3), i.e., within category 0 (absolute richness means [95% Confidence Intervals]: species 12.17 [12.07–12.27]; lineages 19.36 [19.17–19.54]; restricted-range species 1.77 [1.73–1.81]; restricted-range

Fig 3. Protection of different biodiversity features comparison. Box-plots depicting the percent richness of the different biodiversity features falling within the different categories of Socotra’s Zoning Plan (rectangular colors match zoning plan category colors in Fig 1), including mean (thick line), first and third quartile (box), 1.5 interquartile range (whiskers) and outliers (dots). Categories include species richness (sp, in red) and lineage richness (ln, in blue) for total (black thick lines) and restricted-range biodiversity features (restr., grey thick lines).

https://doi.org/10.1371/journal.pone.0200830.g003
lineages 2.04 [1.99–2.09]) than within categories 1 to 3 (species 9.36 [9.35–9.37]; lineage: 13.53 [13.51–13.55]; restricted-range species: 0.76 [0.75–0.76]; restricted-range lineages: 0.83 [0.82–0.83]) There was almost no difference in the representation of richness within categories 1 to 3 of the ZP (Fig 3). Nevertheless, upper quantiles are higher in sanctuaries for restricted-range species than in general use areas (Fig 3 and Table 2). Except for the latter, upper quantiles of richness representation within category 0 and 3 were similar and always below 7% of target achievement. In fact, target achievements for binary richness were all below 98% in any ZP category, and therefore did not reach assigned targets (Fig 3 and Table 2).

An optimal solution for designing new sanctuaries to protect a targeted proportion of all species and lineage distributions was achieved using the reserve design software (Fig 4). The binary outputs for species and lineages encompassed 12.94 and 18.62% of the island area, of which 13.22 and 8.10% were already included in sanctuaries, and 47.00 and 65.92% in national parks, respectively. On the other hand, 37.35 and 24.97%, and 2.43 and 1.03% of the species and lineages outputs intersected resource or general use areas, respectively. These values were significantly different between outputs (Chi-square = 6871.25, d.f. = 3, P<0.001).

Discussion

The importance of considering intraspecific diversity for characterizing spatial patterns

We identified and compared spatial patterns of species and lineage richness, showing that, even though total species and lineage richness are strongly positively correlated, differences between outputs reach circa 30% in some areas of Socotra Island, especially for wide-ranging biodiversity features. This highlights the importance of including lineages in conservation planning. Similarly, other studies found that conservation priorities for rainforest vertebrates in northeast Australia and large African mammals set to maximize species representation do not adequately represent patterns of intraspecific genetic diversity [14,24].

We determined the effect of environmental heterogeneity on inter and intraspecific patterns of diversity. We found that, overall, areas with highest reptile species and lineage richness are characterized by low habitat ruggedness, gentle slopes and high total solar radiation. Reptiles are ectotherms and depend on exposure to the sun for thermoregulation. In places with very steep slopes and high ruggedness sun exposure is limited. In addition, only reptiles adapted to climbing vertical surfaces could occupy those areas. Therefore observed patterns of
diversity follow the expected habitat preferences of reptiles based on their ecological and physiological requirements [57]. Detailed studies on thermal and habitat ecology of the reptile species of Socotra should follow to learn more about the possible ecological and physiological drivers and constraints to intra- and interspecific diversification, similarly to what was found in other island reptiles [58]. In this study, high levels of richness are mainly associated with mountainous landscapes and areas that remained environmentally stable across geological times. Theory proclaims that species should show higher levels of genetic diversity in stable areas where suitable conditions were maintained across glaciation cycles due to long-term persistence of populations and consequent opportunities for diversification and population substructuring [59]. As such these refugia areas, with their higher levels of species and genetic diversity, are important for maximizing the persistence of biodiversity, particularly under global environmental change [60], and for guiding conservation planning. The pattern of increased species richness in mountain regions has been described for other islands and island taxa, including Cabo Verde plants [61], Madeiran beetles [62], and Cabo Verdean reptiles [15]. The elevational gradient on islands, associated with steep changes in climatic and habitat variables, is a strong selective pressure and an ecological barrier that promotes long-term diversification. Mountains act as islands-within-islands, with greater isolation by distance over smaller areas than lowland ecosystems, which results in higher speciation rates [61]. For example, in Socotra, the two sister taxa pairs P. insignis (Blanford 1881) and P. insignoides, and Hemidactylus dracaenacolus and Hemidactylus granti [63] occupy different elevational ranges [34]. This phylogeographic pattern was also recovered for those species pairs when using only the COI marker [28]. The same pattern was also detected in closely-related reptile taxa in adjacent mainland Arabia [39, 64].

Areas of highest lineage richness are also located in the northeast and east of the island, and largely coincide with the oldest rocks of the basement complex and Jurassic-Triassic, dated at around 476 and 145–252 million years ago (Mya), respectively [65]. Lineage richness is associated with similar landscape variables as species richness. However, lineage richness is also strongly negatively correlated with garrigue around the Zahr Basin, a shrubland habitat type found on limestone areas associated with dry climates and oceanic influence with high tolerance to salt. The distribution of this habitat on the island generally coincides with the most recent Quaternary rocks on Socotra Island that were submerged under water during interglacial periods [32]. These submersions likely resulted in population bottlenecks [66], reducing genetic diversity, that were followed by more recent population expansions, which are still observed in the star-shaped-patterns of the genetic network of most reptile species in Socotra [67, 68]. Thus, older and geologically more stables areas of the island appear to have allowed the diversification and longer maintenance of lineage diversity in Socotran reptiles, including recent diversification, as seen for Hemidactylus [69]. Hence, the areas with the highest differences between species and lineage richness outputs were located around the Zahr Basin, where relative species richness is considerably higher than lineage richness, and in the extreme east tip of the island, where the opposite pattern is detected. Because the geological ages of areas of the island that are associated with high lineage richness, but not species richness, have a broad west-east pattern, longitude is, in turn, strongly negatively associated with differences between species and lineage richness.

Spatial differences are much stronger when considering wide-range than restricted-range biodiversity features. In general, wide-range diversity richness patterns correspond to overall diversity richness patterns, while restricted-range diversity shows a different pattern and landscape variable associations. Most LVs with important dissimilar associations with wide-range versus restricted-range diversity are water related (high moisture and precipitation, low distance to wadis and avoidance of dry shrublands). This highlights that many restricted-range
species and lineages are specialists, adapted to the rare conditions of higher water availability on this arid island, as well as the importance of temperature for the diversification of reptiles [70]. Similar patterns were observed in other taxa from desert environments, such as in birds and bats [71,72].

The effectiveness of representing intraspecific diversity

Several biodiversity features have reduced predicted suitable ranges, potentially occupying less than 10% of the island’s area, including the CR H. dracaenacolus, the NT H. granti and the DD P. brevis. This highlights the need to assess the adequacy of the existing ZP for conserving reptile diversity. Even though most biodiversity features are included within PAs, we found that additional sanctuaries (the only areas offering high levels of protection) need to be designed in order to protect a higher percentage of lineage distributions and to cover areas with highest species and lineage richness of reptiles (the main vertebrate group on Socotra). Contrary to the patterns shown in studies performed on other islands across the world [15,49,56], all individual species and lineages are well-represented within at least one of the categories of PAs in Socotra Island. Nevertheless, almost all biodiversity features have very low representation within sanctuaries, and twelve of them are better represented outside PAs than in sanctuaries. Moreover, although the suitable range of all species and lineages fall within PAs, substantial parts of the ranges of some species (e.g. P. brevis) are located partially outside national parks or sanctuary areas. Consequently, construction of infrastructures allowed in the resource use area could affect their protection in the future, especially taking into account human population growth and increasing demand for natural resources. In fact, a similar proportion of areas with the highest values of total and restricted-range biodiversity features is represented in general use areas and sanctuaries, even though the ideal situation would be for sanctuaries to encompass the richest areas, especially of restricted-range lineages. Similar to patterns observed in other parts of the world, under the existing reserve system, both average species and lineage richness is higher outside than inside PAs [73,74]. Higher richness outside PAs cannot be attributed to a low coverage of PAs on Socotra Island because this area is exceptional in that national parks and sanctuaries occupy 73% of its surface, but instead to the fact that the planning units chosen to be protected were not optimized for vertebrates.

Our reserve design study shows that using intraspecific diversity improves the results relative to the use of species diversity. Achieving targets for species requires similar number of grid cells than for lineages (12.94 versus 18.62% of the island area), however the overlap of proposed reserve areas for species with present sanctuary and natural park areas is proportionally 19% lower than with lineages, while overlap with general use areas is 2.4 fold higher. We therefore recommend the allocation of new sanctuaries for fauna, additional to the existing ones, according to our reserve design results based on lineage richness patterns (blue areas in Fig 4) because they maximize representativeness of all biodiversity features for reptiles and largely coincide with many important bird areas [29]. Lineage sanctuary areas identified in our study are generally located in close proximity to existing sanctuaries, with several new proposed areas partially falling within the boundaries of existing sanctuaries (Fig 4). Proximity to existing PAs is an important consideration because it will reduce PAs implementation and management costs, as well the renegotiation efforts with tribal chiefs of those regions. The next step will be to integrate data from other groups that could present different patterns from vertebrate taxa (e.g. invertebrates), and to use genetic markers with faster mutation rates (e.g. microsatellites) to assess connectivity (in terms of gene flow patterns) between proposed PAs.

In conclusion, we show that genetic diversity, here measured as number of lineages, represents distinct spatial patterns from the commonly used measure of species richness, which
disregards evolutionary potential. Our approach to reserve design, which is based on spatial patterns of lineage diversity, enables the inclusion of the evolutionary processes of diversification and speciation. Moreover, we highlight the importance of considering restricted-range species/lineages because their spatial patterns of richness and associations with landscape variables vary from those of wide-ranging biodiversity features and overall species/lineage richness patterns. Our proposed approach to reserve design can be readily applied to other archipelagos and terrestrial systems with available intraspecific-level genetic information.

Supporting information

S1 Table. Details on sampled stations. The codes, region name and coordinates (latitude and longitude in WGS 84) are given for each station.

S2 Table. Phylogenetic analyses details. Information on species codes, number of sequences (N), models of sequence evolution selected by jModelTest, number of variable, singleton and parsimony informative sites for each reptile species, and number of lineages detected using bGMYC.

S3 Table. Spatial modelling and analysis details. Landscape variables (LVs) used in the spatial analyses, including codes, description, average and standard deviation (SD) values, units and original source. The eight LVs with the lowest correlation scores selected for modeling the distribution of reptiles in Socotra are marked underlined.

S4 Table. True skill statistics values (TSS). TSS was calculated with the presences points of each species plus 10 000 random background points. The selected threshold was the same as used for the SDMs (minimum training presence).

Acknowledgments

We are grateful to Socotra’s Environmental Protection Authority, Ministry of Water and Environment of Yemen, Socotra branch for the permits and logistic aid. All necessary permits were obtained for the described study, which complied with all relevant regulations. Observation data were partly collected within the initiatives by the United Nations Development Program and by the Italian Cooperation for ’Biodiversity conservation and sustainable development of Socotra’ led by M.F.. Especially we thank X Santos, G Llorente, E Pujol-Buxó, E Razzetti, C Carugati, U Ziliani, R Sindaco, F Pella, Ahmed Saeed Suleiman, Salem Saeed Balagahar, Abu-bakar Salim, Mohamed Uthman, Yahya Salah Saeed for help in data collection. We are also like to express our sincere gratitude to the anonymous reviewer that provided very detailed constructive comments that contributed to improving the manuscript.

Author Contributions

Conceptualization: Raquel Vasconcelos, Salvador Carranza.

Formal analysis: Raquel Vasconcelos, Orly Razgour, Pedro Tarroso, Salvador Carranza.

Funding acquisition: Raquel Vasconcelos, Mauro Fasola, Salvador Carranza.

Investigation: Raquel Vasconcelos.
Methodology: Raquel Vasconcelos, Orly Razgour, Salvador Carranza.

Validation: Raquel Vasconcelos, Pedro Tarroso, Mauro Fasola, Salvador Carranza, Paulo Célio Alves.

Writing – original draft: Raquel Vasconcelos, Orly Razgour.

Writing – review & editing: Raquel Vasconcelos, Orly Razgour, Pedro Tarroso, Mauro Fasola, Salvador Carranza, Paulo Célio Alves.

References

1. Bottrill MC, Joseph LN, Carwardine J, Bode M, Cook C, Game ET, et al. Is conservation triage just smart decision making? Trends Ecol Evol. 2008; 23: 649–654. https://doi.org/10.1016/j.tree.2008.07.007 PMID: 18848367

2. Laikre L. Genetic diversity is overlooked in international conservation policy implementation. Conserv Genet. 2010; 11: 349–354. https://doi.org/10.1007/s10592-009-0037-4

3. Margules CR, Pressey RL. Systematic conservation planning. Nature. 2000; 405: 243–253. https://doi.org/10.1038/s00124-0005485-3

4. Rodrigues ASL, Brooks TM. Shortcuts for Biodiversity Conservation Planning: The Effectiveness of Surrogates. Annu Rev Ecol Evol Syst. 2007; 38: 713–737. https://doi.org/10.1146/annurev.ecolsys.38.091206.095737

5. Elith J, Leathwick JR. Species Distribution Models: Ecological Explanation and Prediction Across Space and Time. Annu Rev Ecol Evol Syst. 2009; 40: 677–697. https://doi.org/10.1146/annurev.ecolsys.110308.120159

6. Scoble J, Lowe AJ. A case for incorporating phylogeography and landscape genetics into species distribution modelling approaches to improve climate adaptation and conservation planning. Divers Distrib. 2010; 16: 343–353. https://doi.org/10.1111/j.1472-4642.2010.00658.x

7. Höglund J. Evolutionary Conservation Genetics. New York: Oxford University Press; 2009.

8. Manel S, Schwartz MK, Luikart G, Taberlet P. Landscape genetics: combining landscape ecology and population genetics. Trends Ecol Evol. 2003; 18: 189–197. https://doi.org/10.1016/S0169-5347(03)0008-9

9. Wang IU. Recognizing the temporal distinctions between landscapes and phylogeography. Mol Ecol. 2010; 19: 2605–2608. https://doi.org/10.1111/j.1365-294X.2010.04715.x PMID: 20561197

10. Hughes AR, Inouye BD, Johnson MTJ, Underwood N, Vellend M. Ecological consequences of genetic diversity. Ecol Lett. 2008; 11: 609–623. https://doi.org/10.1111/j.1461-0248.2008.01179.x PMID: 18400018

11. Whititaker RJ, Fernández-Palacios JM. Island Biogeography—Ecology, evolution, and conservation. second. Oxford, New York: Oxford University Press; 2007.

12. Thomassen HA, Cheviron Z a, Freedman AH, Harrigan RJ, Wayne RK, Smith TB. Spatial modelling and landscape-level approaches for visualizing intra-specific variation. Mol Ecol. 2010; 19: 3532–3548. https://doi.org/10.1111/j.1365-294X.2010.04737.x PMID: 20723053

13. Tarroso P, Velo-Antón G, Carvalho SB. PHYLIN: an R package for phylogeographic interpolation. Mol Ecol Resour. 2015; 15: 349–357. https://doi.org/10.1111/1755-0998.12512 PMID: 25123598

14. D’Amen M, Zimmermann NE, Pearman PB. Conservation of phylogeographic lineages under climate change. Glob Ecol Biogeogr. 2013; 22: 93–104. https://doi.org/10.1111/geb.12091 x

15. Vasconcelos R, Brito JC, Carvalho SB, Carranza S, James Harris D. Identifying priority areas for island endemics using genetic versus specific diversity—The case of terrestrial reptiles of the Cape Verde Islands. Biol Conserv. Elsevier Ltd; 2012; 153: 276–286. https://doi.org/10.1016/j.biocon.2012.04.020

16. Thomassen HA, Fuller T, Buermann W, Milá B, Kiesewetter CM, Jarrín V, P, et al. Mapping evolutionary process: a multi-taxa approach to conservation prioritization. Evol Appl. 2011; 4: 397–413. https://doi.org/10.1111/j.1752-4571.2010.00172.x PMID: 25567981

17. Wilson A, Arcese P, Keller LF, Pruett CL, Winker K, Patten M a., et al. The contribution of island populations to in situ genetic conservation. Conserv Genet. 2009; 10: 419–430. https://doi.org/10.1007/s10592-008-9612-3

18. Myers N. Biodiversity Hotspots Revisited. Bioscience. 2003; 53: 916–917.
19. Pollock LJ, Rosauer DF, Thomhill AH, Kujala H, Crisp MD, Miller JT, et al. Phylogenetic diversity meets conservation policy: small areas are key to preserving eucalypt lineages. Philos Trans R Soc B Biol Sci. 2015; 370: 20140007–20140007. https://doi.org/10.1098/rstb.2014.0007 PMID: 25561668

20. Bennett JR, Elliott G, Mellish B, Joseph LN, Tulloch AIT, Probert WJM, et al. Balancing phylogenetic diversity and species numbers in conservation prioritization, using a case study of threatened species in New Zealand. Biol Conserv. 2014; 174: 47–54. https://doi.org/10.1016/j.biocon.2014.03.013

21. Losos JB, Ricklefs RE. Adaptation and diversification on islands. Nature. 2009; 457: 830–836. https://doi.org/10.1038/nature07893 PMID: 19212401

22. Miraldo A, Li S, Borregaard MK, Florez-Rodriguez A, Gopalakrishnan S, Rizvanovic M, et al. An Anthropocene map of genetic diversity. Science (80-). 2016; 353: 1532–1535. https://doi.org/10.1126/science.aaf4381 PMID: 27708102

23. Sechrest W, Brooks TM, da Fonseca G a B, Konstant WR, Mittermeier R a, Purvis A, et al. Hotspots and the conservation of evolutionary history. Proc Natl Acad Sci U S A. 2002; 99: 2067–2071. https://doi.org/10.1073/pnas.251680798 PMID: 11854502

24. Moritz C. Strategies to protect biological diversity and the evolutionary processes that sustain it. Syst Biol. 2002; 51: 238–54. https://doi.org/10.1080/10635150252899752 PMID: 12028731

25. Kier G, Kreft H, Lee TM, Jetz W, Ibisch PL, Nowicki C, et al. A global assessment of endemism and species richness across island and mainland regions. Proc Natl Acad Sci U S A. 2009; 106: 9322–9327. https://doi.org/10.1073/pnas.0810306106 PMID: 19470638

26. Le Saout S, Hoffmann M, Shi Y, Hughes A, Bernard C, Brooks TM, et al. Protected Areas and Effective Biodiversity Conservation. Science (80-). 2013; 342: 803–805. https://doi.org/10.1126/science.1239268

27. Scholte P, Al-Okaishi A, Suleyman AS. When conservation precedes development: a case study of the opening up of the Socotra archipelago, Yemen. Oryx. 2011; 45: 401–410. https://doi.org/10.1017/S0030605310001535

28. Vasconcelos R, Montero-Mendieta S, Simó-Riudalbas M, Sindaco R, Santos X, Fasola M, et al. Unexpectedly High Levels of Cryptic Diversity Uncovered by a Complete DNA Barcoding of Reptiles of the Socotra Archipelago. PLoS One. 2016; 11: e0149985. https://doi.org/10.1371/journal.pone.0149985 PMID: 26930572

29. Porter RF, Suleiman AS. The Important Bird and Biodiversity Areas of the Socotra archipelago, Yemen. Sandgrouse. 2016; 38: 169–191.

30. De Sanctis M, Adeeb A, Farcomeni A, Patriarca C, Saeed A, Attorre F. Classification and distribution patterns of plant communities on Socotra Island, Yemen. Appl Veg Sci. 2013; 16: 148–165. https://doi.org/10.1111/j.1654-109X.2012.01212.x

31. Autin J, Bellahsen N, Leroy S, Husson L, Beslier M-O, D’Acremont E. The role of structural inheritance in oblique rifting: Insights from analogue models and application to the Gulf of Aden. Tectonophysics. 2013; 607: 51–64. https://doi.org/10.1016/j.tecto.2013.05.041

32. Beydoun ZR, Bichan HR. The geology of Socotra Island, Gulf of Aden. Q J Geol Soc London. 1970; 125: 413–446.

33. Razzetti E, Sindaco R, Grieco C, Pella F, Ziliani U, Pupin F, et al. Annotated checklist and distribution of the Socotran Archipelago Herpetofauna (Reptilia). Zootaxa. 2011; 44: 1–44.

34. Fujisawa T, Barralou T, Delimiting species using single-locus data and the Generalized Mixed Yule Coalescent approach: a revised method and evaluation on simulated data sets. Syst Biol. 2013; 62: 707–24. https://doi.org/10.1093/sysbio/syt033 PMID: 23681854

35. Rambaut A, Drummond AJ. Tracer v1.6. In: Available from: http://tree.bio.ed.ac.uk/software/tracer/. 2013.

36. Reid NM, Carstens BC. Phylogenetic estimation error can decrease the accuracy of species delimitation: a Bayesian implementation of the general mixed Yule-coalescent model. BMC Evol Biol. 2012; 12: 196. https://doi.org/10.1186/1471-2148-12-196 PMID: 23031350

37. Garcia-Porta J, Simó-Riudalbas M, Robinson M, Carranza S. Diversification in arid mountains: biogeography and cryptic diversity of Pristurus rupestris rupestris in Arabia. J Biogeogr. 2017; in press: 1–11. https://doi.org/10.1111/jbi.129229

38. R Core Team. A language and environment for statistical computing [Internet]. Vienna, Austria: R Foundation for Statistical Computing; 2015. Available: http://www.r-project.org/

39. Rambaut A, Drummond AJ. Tracer v1.6. In: Available from: http://tree.bio.ed.ac.uk/software/tracer/. 2013.
41. Pons J, Barraclough TG, Gomez-Zurita J, Cardoso A, Duran DP, Hazell S, et al. Sequence-Based Species Delimitation for the DNA Taxonomy of Undescribed Insects. Syst Biol. 2006; 55: 595–609. https://doi.org/10.1086/506008 PMID: 16967577

42. Fujisawa T, Barraclough TG. Delimiting species using single-locus data and the generalized mixed yule coalescent approach: A revised method and evaluation on simulated data sets. Syst Biol. 2013; 62: 707–724. https://doi.org/10.1093/sysbio/syt033 PMID: 23681854

43. Phillips SJ, Anderson RP, Schapire RE. Maximum entropy modeling of species geographic distributions. Ecol Model. 2006; 190: 231–259. https://doi.org/10.1016/j.ecolmodel.2005.03.026

44. Malatesta L, Attorre F, Altabelli A, Adeeb A, De Sanctis M, Taleb NM, et al. Vegetation mapping from high-resolution satellite images in the heterogeneous arid environments of Socotra Island (Yemen). J Appl Remote Sens. 2013; 7: 073527. https://doi.org/10.1117/1.JRS.7.073527

45. Liu C, Berry PM, Dawson TP, Pearson RG. Selecting thresholds of occurrence in the prediction of species distributions. Ecoscropy (Cop). 2005; 3: 385–393.

46. Allouche O, Tsoar A, Kadmon R. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). J Appl Ecol. 2006; 43: 1223–1232. https://doi.org/10.1111/j.1365-2664.2006.01214.x

47. Hammill E, Tulloch AIT, Possingham HP, Strange N, Wilson KA. Factoring attitudes towards armed conflict risk into selection of protected areas for conservation. Nat Commun. 2016; 7: 11042. https://doi.org/10.1038/ncomms11042 PMID: 27025894

48. Chan Y, Zhang J, Jiang J, Nielsen SE, He F. Assessing the effectiveness of China’s protected areas to conserve current and future amphibian diversity. Divers Distrib. 2017; 23: 146–157. https://doi.org/10.1111/ddi.12508

49. Bosso L, Mucedda M, Fichera G, Kiefer A, Russo D. A gap analysis for threatened bats populations on Sardinia. Hystrix. 2016; 27. https://doi.org/10.4404/hystrix-27.2–11788

50. Rodrigues ASL, Akçakaya HR, Andelman SJ, Bakarr MI, Boitani L, Brooks TM, et al. Global Gap Analysis: Priority Regions for Expanding the Global Protected-Area Network. Bioscience. 2004; 54: 1092–1100.

51. Venter O, Fuller RA, Segan DB, Carwardine J, Brooks T, Butchart SHM, et al. Targeting Global Protected Area Expansion for Imperiled Biodiversity. Moritz C, editor. PLoS Biol. 2014; 12: e1001891. https://doi.org/10.1371/journal.pbio.1001891 PMID: 24960185

52. Zhang Y, Liu Y, Fu J, Phillips N, Zhang M. Bridging the “gap” in systematic conservation planning. J Nat Conserv. Elsevier GmbH.; 2016; 31: 43–50. https://doi.org/10.1016/j.jnc.2016.03.003

53. Garcia-Porta J, Šmíd J, Sol D, Fasola M, Carranza S. Testing the island effect on phenotypic diversification: insights from the Hemidactylus geckos of the Socotra Archipelago. Sci Rep. Nature Publishing Group; 2016; 6: 23729. https://doi.org/10.1038/srep23729 PMID: 27071837

54. Martín N, Martínez S, Pujol-Buxó E, Viñolas A, Llorente GA, Sanpere C, et al. Stable isotopes and diet uncover trophic-niche divergence and ecological diversification processes of endemic reptiles on Socotra Island. Zool Anzeiger—A J Comp Zool. 2017; 267: 69–81. https://doi.org/10.1016/j.jcz.2017.01.005

55. Molianen A, Pouzols FM, Meller L, Veach V, Arponen A, Leppänen J, et al. Spatial conservation planning methods and software zonation User Manual. Version 4. In: C-BIG Conservation Biology Informatics Group, Department of Biosciences, University of Helsinki, Finland [Internet]. 2014 [cited 4 Jan 2016] p. 288. Available: http://cbig.it.helsinki.fi/files/zonation/zonation_manual_v4_0.pdf

56. Taylor C, Cadenhead N, Lindenmayer DB, Wintle BA. Improving the design of a conservation reserve for a critically endangered species. PLoS One. 2017; 12: 1–14. https://doi.org/10.1371/journal.pone.0169629 PMID: 28121984

57. Currie DJ. Energy and Large-Scale Patterns of Animal- and Plant-Species Richness. Am Nat. 1991; 137: 27–49. Available: http://www.jstor.org/stable/2462155

58. Hertz PE, Arima Y, Harrison A, Huey RB, Losos JB, Glor RE. Asynchronous evolution of physiology and morphology in Anolis lizards. Evolution (N Y). 2013; 67: 2101–2113. https://doi.org/10.1111/evo.12072 PMID: 23815663

59. Qu Y, Ericson PG, Quan Q, Song G, Zhang R, Gao B, et al. Long-term isolation and stability explain high genetic diversity in the Eastern Himalaya. Mol Ecol. 2014; 23: 705–720. PMID: 24600707

60. Razgour O, Juste J, Ibáñez C, Kiefer A, Rebelo H, Puechmaille SJ, et al. The shaping of genetic variation in edge-of-range populations under past and future climate change. Ecol Lett. 2013; 16: 1258–66. https://doi.org/10.1111/ele.12158 PMID: 23689048

61. Steinbauer MJ, Field R, Grytnes J, Trigas P, Ah-Peng C, Attorre F, et al. Topography-driven isolation, speciation and a global increase of endemism with elevation. Glob Ecol Biogeogr. 2016; 1–11. https://doi.org/10.1111/geb.12469
62. Boiero M, Carvalho JC, Cardoso P, Aguiar CAS, Rego C de Faria e Silva I, et al. Spatial factors play a major role as determinants of endemic ground beetle Beta diversity of Madeira Island laurisilva. PLoS One. 2013; 8: e64591. https://doi.org/10.1371/journal.pone.0064591 PMID: 23724065

63. García-Porta J, Morales HE, Gómez-Díaz E, Sindaco R, Carranza S. Patterns of diversification in islands: A comparative study across three gecko genera in the Socotra Archipelago. Mol Phylogenet Evol. Elsevier Inc.; 2016; 98: 288–299. https://doi.org/10.1016/j.ympev.2016.02.007 PMID: 26911521

64. Šmíd J, Shobrak M, Wilms T, Jogor U, Carranza S. Endemic diversification in the mountains: genetic, morphological, and geographical differentiation of the Hemidactylus geckos in southwestern Arabia. Org Divers Evol. 2016; https://doi.org/10.1007/s13127-016-0293-3

65. Ahmed A, Leroy S, Keir D, Korostelev F, Khanbari K, Rolandone F, et al. Crustal structure of the Gulf of Aden southern margin: Evidence from receiver functions on Socotra Island (Yemen). Tectonophysics. 2014; 637: 251–267. https://doi.org/10.1016/j.tecto.2014.10.014

66. Rijsdijk KF, Hengl T, Norder SJ, Otto R, Emerson BC, Ávila SP, et al. Quantifying surface-area changes of volcanic islands driven by Pleistocene sea-level cycles: Biogeographical implications for the Macaronesian archipelagos. J Biogeogr. 2014; 41: 1242–1254. https://doi.org/10.1111/jbi.12336

67. Vasconcelos R, Carranza S. Systematics and biogeography of Hemidactylus homoeolpis Blanford, 1881 (Squamata: Gekkonidae), with the description of a new species from Arabia. Zootaxa. 2014; 3835: 501–527. https://doi.org/10.11646/zootaxa.3835.4.4 PMID: 25081467

68. Badiane A, García-Porta J, Červenka J, Kratochvíl L, Sindaco R, Robinson MD, et al. Phylogenetic relationships of Semaphore geckos (Squamata: Sphaerodactylidae: Pristurus) with an assessment of the taxonomy of Pristurus rupestris. Zootaxa. 2014; 3835: 33–58. https://doi.org/10.11646/zootaxa.3835.1.2 PMID: 25081434

69. Gómez-Díaz E, Sindaco R, Pupin F, Fasola M, Carranza S. Origin and in situ diversification in Hemidactylus geckos of the Socotra Archipelago. Mol Ecol. 2012; 21: 4074–4092. https://doi.org/10.1111/j.1365-294X.2012.05672.x PMID: 22738330

70. Gillman LN, Wright SD. Species richness and evolutionary speed: The influence of temperature, water and area. J Biogeogr. 2014; 41: 39–51. https://doi.org/10.1111/jbi.12173

71. Tielenman BI, Williams JB. The adjustment of avian metabolic rates and water fluxes to desert environments. Physiol Biochem Zool. 2000; 73: 461–79. https://doi.org/10.1086/317740 PMID: 11009400

72. Razgour O, Korine C, Saltz D. Does interspecific competition drive patterns of habitat use in desert bat communities? Oecologia. 2011; 167: 493–502. https://doi.org/10.1007/s00442-011-1995-z PMID: 21519884

73. Gray CL, Hill SLL, Newbold T, Hudson LN, Börger L, Contu S, et al. Local biodiversity is higher inside than outside terrestrial protected areas worldwide. Nat Commun. 2016; 7: 12306. https://doi.org/10.1038/ncomms12306 PMID: 27465407

74. Oliveira U, Soares-Filho BS, Paglia AP, Brescovit AD, De Carvalho CJB, Silva DP, et al. Biodiversity conservation gaps in the Brazilian protected areas. Sci Rep. 2017; 7. https://doi.org/10.1038/s41598-017-08707-2 PMID: 28831073