Supplementary material for

Terrestrial protected areas maintain freshwater ecosystem resilience to costly aquatic invasive species in the Panama Canal

Jorge Salgado, María I. Vélez, Catalina González-Arango, Aaron O'Dea

Materials and Methods

Regional data of dams, tNPAs and macrophyte invasive species

To estimate for the number of large lowland (<1000 masl) reservoirs in the American Tropics we follow[1] by using the location of point data of 1293 existing dams across Latin American and the Caribbean from the GRanD database[2]; as well as 1436 hydropower dams from the FHReD database[3] that are planned or under construction. These dam datasets include dams across the regions with more than 0.1 km^3 storage capacity and only hydroelectric dams with a power capacity >1 MW for FHReD. The identified dam locations were then overlaid with polygons from the World Database of Protected Areas (WDPA)[4] considering terrestrial natural protected areas (tNPAs) with national, regional, and international designations. Protected dams were defined here as those having a direct adjacent influence from a tNPA (i.e., any lake littoral area falling within a tNPA). In this sense, other forms of river protection such as headwaters inclusions within PNAs were excluded.

To explore the extent of macrophyte IAS occurrences in protected reservoirs, eleven well-known nuisance macrophyte species (native and non-native) for Latin America and the Caribbean were selected. These were: *Alternanthera philoxeroides*, *Egeria densa*, *Egeria najas*, *Eichhornia crassipes*, *Eichhornia azureus*, *Myriophyllum aquaticum*, *Myriophyllum*

spicatum, *Pistia stratiotes*, *Salvinia auriculata*, *Typha angustifoli*a, and *Urochloa subquadripara.* These species were selected based on the high level of invasiveness reported by[5,6]. Macrophyte invasive species occurrences across the region were obtained from global biodiversity information facilities, and supplemented records from literature, commercial reports, photographs and newspapers. The "Spocc'' package in R[7] was used to download freely available occurrence data from the Global Biodiversity Information Facility (http://www.gbif.org), Atlas of Living Australia (http://www.ala.org.au), Biodiversity Information Serving Our Nation (http://www.bison.usgs.gov), iNaturalist (http://www.inaturalist.org), Ecoengine Interface (http://www.github.com/ropensci/ecoengine), iDigBio (http://www.idigbio.org), and Ocean Biographic Information System (http://www.iobis.org) (accessed March 2021). A limit of 10,000 records per species was used for each search and any species data without georeferencing were excluded. Identified species records were then overlaid with the selected lowland protected point data of the GRanD and (FHReD) databases.

Gatun Lake dynamics

Macrophyte communities were sampled across a series of different lake sectors adjacent to tNPAs (BCNM and Península Gigante) and to sectors outside the influence of NPAs (La Represa and Bahia Trinidad; see Fig. 1 in the main text). tNPAs encompass large areas of secondary forest, whereas not protected areas are characterised by forest plantations, agricultural land, grasslands and urban centres. The main criteria for lake sectors selection were based on depicting the differences between land-uses while avoiding any other extrinsic influence from the Canal daily operations. Three extra sectors in the Chagres River

46 were further studied to assess the extent to which lake macrophyte communities may
47 compare to the parental river.

48 Submerged, floating-leaved and emergent macrophyte (angiosperms, and charophytes)
49 species abundance data were sampled at each selected lake sector using a modification of
50 the UK Standard Joint Nature Conservation Committee (JNCC) protocols for site monitoring[8].
51 In summary, this methodology allows for the characterization of macrophyte communities
52 within each sector, based on 100m transects at both littoral and deeper waters, to give good
53 spatial lake coverage[9]. At each 100m transect, macrophyte data is collected at four depths
54 ranging between 25 and 75 cm at intervals of every 20m (20 points in total per transect).
55 When inner littoral areas were inaccessible due to dense floating mats, shoreline
56 macrophyte-sampling points at the depths of 25-50 cm were assessed from a boat starting
57 from the margins of the floating mats. Macrophytes in deeper water (depths >75cm) were
58 also surveyed using a boat along a transect starting at the midpoint of each 100m transect
59 and running towards the centre of the lake sector. Macrophytes were sampled at every 5m
60 until no macrophyte was recorded or until a maximum of 10 sampling points were achieved.
61 At each sampling point, we used a combination of bathyscope and grapnel sampling, and all
62 aquatic macrophyte species occurring within a $1m^2$ area were recorded using a percentage
63 plant cover score. Representation of the main macrophytes present in each lake sector was
64 the basis for selecting the 100m transects. Seven lake sectors adjacent to NPAs were
65 selected, with two transects per sector surveyed (n = 329 sampling points). Four lake sectors
66 not adjacent to tNPAs were chosen with a total of nine transects surveyed (n = 212 sampling
67 points). For the Chagres River, three sectors were selected with one 100m transect surveyed
68 per sector (n = 44 sampling points). While this sampling approach may have missed some

69 macrophyte species known to be present in the lake, the methodology nevertheless
70 provides a useful representation of variation in macrophyte distributions and abundances
71 for the majority of occurring species[9,10].

72 Physical-chemical data from the sampling lake sectors and the Chagres River, were derived
73 from available monitoring data from 2012 and 2013 collected at five adjacent sampling
74 stations. These data were obtained from "Autoridad del Canal de Panama–ACP''[11] and
75 included dissolved oxygen (DO), conductivity, pH, chlorophyll-a, nitrate (NO_3), and secchi
76 depth; variables that have shown to influence macrophyte distributions in the Gatun Lake[12].
77 A direct measure of water clarity variation at each macrophyte sampling point, was further
78 derived through a secchi depth reading at the deepest point of each transect and divided by
79 the water depth at each sampling point[13].

80 **Data analysis**

81 The statistical analyses focused on two complementary aspects of macrophyte community
82 structure in the Gatun Lake[14]: *turnover* the directional change in assemblage composition
83 from one sampling unit to another; and *community heterogeneity* the variation in species
84 composition arising from shifts in species identities and abundances among groups of
85 sampling units over time. By linking these two measures of beta diversity, the underlying
86 nature of patterns in beta diversity that arise simultaneously from presence/absence data
87 and relative abundance information can be better revealed[14]. We used a combined
88 multivariate analysis approach of Homogeneity Analysis of Multivariate Dispersions (HMD[15])
89 and Permutational Analysis of Variance (PerMANOVA[16]). HMD analysis is a non-parametric
90 method that compares variability of mean distance to a centroid (dispersion) within
91 predefined groups (in our case lake and river sectors), to variability in this distance between

92 the predetermined groups in a PCoA[15]. Macrophyte community heterogeneity was defined
93 in our study as the distance to the spatial median of the dissimilarities in macrophyte
94 species relative abundances among sampling points, grouped respectively, within the three
95 conservation lake sectors categories. A sector with higher values of mean distance to the
96 group median was assumed to be characterized by greater multivariate dispersion in
97 macrophyte species abundance between the sampling points, and hence, by greater
98 community heterogeneity[10,15]). Conversely, low multivariate dispersion (lower mean
99 distance to the group median) indicates a more homogenous community structure. The
100 significance ($p<0.05$) of each HMD analysis was assessed via ANOVA and differences
101 between pairs of lake and river sectors groups were then tested *post hoc* using the Tukey
102 HSD. HMD analyses were performed using Bray-Curtis dissimilarities on *log+1* transformed
103 macrophyte data using the *betadisper* function in the "vegan" package[17] and results were
104 plotted using boxplots and PCoA plots. Given that we compared groups with unequal
105 numbers of samples, we used the "setting bias.adjust=TRUE" in betadisper to impose a
106 sqrt(n/(n-1)) correction[18].

107 PerMANOVA is a non-parametric method for multivariate analysis of variance that
108 compares the variability of average dissimilarity within groups, versus the variability among
109 groups, using the ratio of the F-statistic through permutational tests. PerMANOVA enabled
110 thus assessments of the significance of the community compositional heterogeneity
111 attributed to variation in the identity of species (turnover) between the different lake and
112 river sectors.

113 To assess how the surveyed macrophyte communities at the different study areas related to
114 historical plant communities, we compared the survey data against a previously published

plant-macrofossil abundance data1[2]. The plant-macrofossil data was derived from a dated littoral sediment core (LGAT1) taken in a lake sector outside the influence of NPAs in the southwest zone of the lake (see Fig.1in the main text). We run a combined HMD analysis using Bray-Curtis dissimilarities on *log+1* transformed contemporary and plant macrofossil data. Prior to analysis, the plant macrofossil data was clustered into two temporal groups representing pre-canal times (i.e., pre-1913); and Gatun Lake times (i.e., 1915-present).

The main gradients of variation in macrophyte species and physical-chemical parameters between the different study lake sectors was explored independently via principal component analysis (PCA; "FactoMiner" R package[19]). The variation in water clarity values at each macrophyte sampling point was then assessed via HMD and Tukey HSD. The HMD analysis was run on Euclidean distances on *log+1* transformed water clarity data.

References

1. Thieme, M. L., et al. *Conserv. Letters* **13**, e12719 (2020).
2. Lehner, B., C. et al. *Front. Ecol Environ*. **9**, 494–502 (2011).
3. Zarfl, C., et al. *Aquat. Sci.* **77,** 161–170.
4. UNEP-WCMC. *Protected areas map of the world*. www.protectedplanet.net (accessed in February 2021).
5. Lozano, V. & Brundu, G. *Hydrobiologia*, **812**, 115–130 (2018).
6. Lozano, V. *Ecologies*, 2, 27–42 (2021).
7. Chamberlain, S. et al. *Spocc: R Interface to Many Species Occurrence Data Sources. R package version 1.1.0. 2020*. Available online: https://CRAN.R-project.org/package=spocc (accessed in February 2021).

8. JNCC. *Draft common standards monitoring guidance for standing waters*. Joint Nature Conservation Committee, Peterborough, UK, (2009).
9. Gunn, I. D. et al. *Hydrobiologia* **656**, 87–97 (2010).
10. Salgado, J. et al. *Ecosphere* **9**, e02406 (2018).
11. ACP- *Autoridad del Canal de Panamá. Informes de calidad del agua, Canal de Panamá, 2003–2013. Autoridad del Canal de Panamá, Departamento de Ambiente, Agua y Energía.* https://micanaldepanama.com/nosotros/cuenca-hidrografica/ (Accessed Decmeber 2020).
12. Salgado et al. *Sci Total Environ*. **729**, 138444.
13. Salgado et al. Riv. Reser. Appl. (accepted)
14. Anderson, M. J., et al. *Ecol. Letters*, **14**, 19–28 (2011).
15. Anderson, M. J., et al. *Ecol. Letters*, **9**, 683–693 (2006).
16. Anderson, M. J. *Wiley statsref: statistics reference online*, 1-15
17. Oksanen, J. *The vegan package. Community ecology package*, **10**, 719 (2019).
18. Stier, A.C. et al. *Ecology* **94**, 1057–1068 (2013).
19. Husson, F. et al. *Package "FactoMineR". An R package* **96**, 698 (2016)

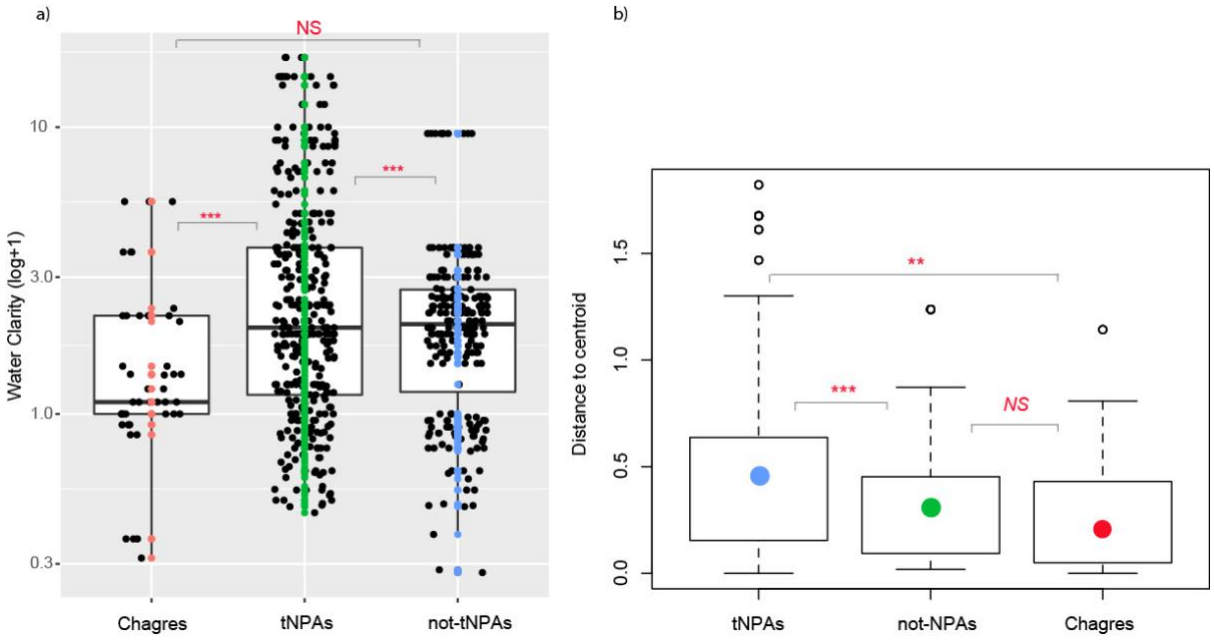

Figure 1 Boxplots showing (a) the mean variation of water clarity at each macrophyte sampling point in three study areas of the Gatun Reservoir: the Chagres River, lake areas adjacent to tNPAs, and lake areas not adjacent to tNPAs (not-tNPAs); and (b) the degree of water clarity heterogeneity of each study area measured via homogeneity multivariate dispersion test (HMD). The differences in means and distance to mean of each study area was assessed via post hoc pairwise comparison via Tukey Honest test under a significance level of $p \leq 0.05$. *$p \leq 0.05$; ** $p \leq 0.01$; ***$p \leq 0.001$; NS=not significant.