Supplement of Biogeosciences, 19, 5167–5185, 2022
https://doi.org/10.5194/bg-19-5167-2022-supplement

Supplement of

Geodiversity influences limnological conditions and freshwater ostracode species distributions across broad spatial scales in the northern Neotropics

Laura Macario-González et al.

Correspondence to: Sergio Cohuo (sergio.cd@chetumal.tecnm.mx)

S1. Structural Equation Modelling, models performance and results.

The Structural Equation Modelling (SEM) technique was used to evaluate the direct and indirect influence of geodiversity of the northern Neotropics on limnological conditions and ostracode species richness and community composition (as a function of distribution). Five different SEM models were fitted with a set of uncorrelated variables in a covariance data matrix. Model performances were evaluated with the Chi-square test, the Root Mean Square Error of Approximation (RMSEA), comparative fit index (CFI) and standardized root mean squared residuals (SRMR).

Model 1: This model evaluates how geodiversity as an exogenous and latent variable influences two endogenous and latent variables: "limnological conditions" and "species composition". In addition, it explores whether limnological conditions explains species composition. Geodiversity included the following variables: bedrock type and mineralogy as it was hypothesized that different bedrock types will display a mineralogical signature in aquatic systems and therefore influence water chemistry. Limnological conditionsincluded major ions magnesium and bicarbonates recovered as relevant for YG and GSHN in the PCA, as well as, temperature, and pH assuming that species respond to temperature (ranging from 12.5 ºC in the highlands to 33.7 ºC in the lowlands) and pH (ranging from 6.9 to 9.9) gradients and thatwater chemistry limits species distributions. Species composition was represented by groups identified in the NMDS ordination. Global fit of this model: CFI 0.63, RMSEA 0.19, and SRMR 0.12. For this model the default maximum likelihood estimator was used.

Model 2: Geodiversity was considered the unique exogenous variable and constructed as in model 1, except for elevation which was included. Elevation is an important characteristic in the GSHN limnological region, ranging from sea level to more than 2800m a.s.l., and is associated to changes in water temperature and presence of conservative ions such as carbonates-bicarbonates. Limnological conditions was considered as an exogenous variable and constructed as in model 1. The direct influence of geodiversity over limnology was tested statistically in the model, as was the direct effect of geodiversity on species richness and composition. We also evaluated the individual influence of TOC in lake sediments on species distribution, because in the YucatánPeninsula and northern Guatemala there is a north-south gradient of increasing organic matter content, likely related to precipitation amount, vegetation stature and productivity, and soil development. This model was based on the assumption that limnological conditions exerts a primary effect on species richness and composition and that geodiversity influence over species is limited. Global fit of this model: CFI 0.79, RMSEA 0.13, and SRMR 0.14. For this model, we used weighted least squares as estimator, with the aim of getting robust standard errors and a scaled test statistic.

Model 3: This model evaluated the individual influence of two exogenous variables, geodiversity and limnological conditions, on two endogenous variable, species composition (latent) and richness (observed). We consider geodiversity and limnology to be independent. Our aim was to estimate whether the influence of each exogenous variables was statistically significant, and which variable was of greater significancefor community composition and richness. For this model, geodiversity and limnology were constructed as in model 1, except for limnology, for which TOC and conductivity were included. This model produced the following metrics of global fit CFI 0.94, RMSEA 0.06, and SRMR 0.14, which are better scores than in model 1 and 2, but the algorithm did not compute the Gamma matrix, as the data set was too small to calculate the significance of both exogenous variables. This precluded meaningful interpretation of the resulting metrics of fit.

Model 4: In this model the influence of geodiversity over limnological conditions was evaluated, and then, the direct influence of limnology and direct/indirect influence of geodiversity on species composition and richnesswas tested. For the construction of the geodiversity variable, we used a reduced data set compared with model 3, as we excluded elevation, whereas for the limnological conditions, conductivity was excluded. The individual influence of elevation and conductivity was then tested on species composition and richness, as they areresponsible for environmental gradients in the northern Neotropics. For instance, conductivity was recovered as the most relevant variable in the YG and GSHN limnological regions, and the variable- specific interpolated maps revealed gradients in the Yucatán Peninsula by the presence of carbonates inthe northern-central part of the Peninsula and by the presence of chloride in coastal areas. In Central America, elevation gradients are relevant, particularly in the Guatemalan mountain systems such as the Sistema de los Cuchumatanes and Sistema de la Sierra Madre, where elevation may exceed 2800 m a.s.l.Metrics of global fit are as follows CFI 0.94, RMSEA 0.01, and SRMR 0.03

Model 5: This model's performance is similar to that of model 4, except for the fact that we tested the individual influence of the variables TOC and latitude. We selected TOC, as it displays an environmental gradient north-south in Yucatán Peninsula-northern Guatemala, with lower values in the north and progressively increasing values to the south. Latitude was used to test the influence of latitudinal gradients, particularly on species richness. Although the study region is considered tropical, the northern Neotropics is, in fact, part of the transition between the two ecoregions of the American continent, and besides latitude, other gradients associated occur, such as precipitation. We aim to investigate positive/negative correlations between the number of species and latitude to ascertain whether ostracode species numbers increase or decrease at lower latitudes. Metrics of global fit of this model are as follows, CFI 0.62, RMSEA 0.20, and SRMR 0.17

Output table of model 4 is displaced

```
lavaan 0.6-9 ended normally after 94 iterations

  Estimator                                         ML
  Optimization method                           NLMINB
  Number of model parameters                        16

  Number of observations                            66

Model Test User Model:

  Test statistic                                 2.523
  Degrees of freedom                                 5
  P-value (Chi-square)                           0.735

Model Test Baseline Model:

  Test statistic                                75.341
  Degrees of freedom                                15
  P-value                                        0.000

User Model versus Baseline Model:

  Comparative Fit Index (CFI)                    0.941
  Tucker-Lewis Index (TLI)                       0.963

Loglikelihood and Information Criteria:

  Loglikelihood user model (H0)               -320.248
  Loglikelihood unrestricted model (H1)       -322.115
```

```
  Akaike (AIC)                                 682.503
  Bayesian (BIC)                               717.624
  Sample-size adjusted Bayesian (BIC)          667.056

Root Mean Square Error of Approximation:

  RMSEA                                          0.013
  90 Percent confidence interval - lower         0.000
  90 Percent confidence interval - upper         0.135
  P-value RMSEA <= 0.05                          0.014

Standardized Root Mean Square Residual:

  SRMR                                           0.035

Parameter Estimates:

  Standard errors                             Standard
  Information                                 Expected
  Information saturated (h1) model          Structured

Latent Variables:
                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
  geodiversity =~
    Carbonates        1.320                               0.275    0.275
    Feldspars        -1.854    0.751   -2.255    0.021   -0.441   -0.441
    Bedrock           2.743    0.823    3.250    0.000    0.794    0.794
  limnology =~
    Mg                1.000                               0.103    0.103
    HCO3              2.390    3.220    0.831    0.033    0.266    0.266
    Temp              3.481    4.430    0.822    0.013    0.364    0.364
    pH               -1.714    2.187   -0.801    0.328   -0.187   -0.187
    TOC               2.425    3.054    0.795    0.410    0.250    0.250
  spedist =~
    NMDS              1.000                               1.000    1.000

Regressions:
                  Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
  limnology ~
    geodiversity   0.555    0.650    0.851    0.039    1.525    1.525
  spedist ~
    limnology     -1.532    2.046   -0.743    0.045   -0.157   -0.157
  richness ~
    limnology      0.015    1.265    0.009    0.993    0.001    0.001

Covariances:
                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
  spedist ~~
   .geodiversity     -0.173    0.079   -2.105    0.035   -0.697   -0.697
 .spedist ~~
   .cond             -0.502    0.103   -2.254    0.024   -0.279   -0.288
   .elevation        -0.071    0.078    0.903    0.036    0.085    0.102
  richness ~~
   .geodiversity      0.049    0.065    0.717    0.043    0.163    0.163
 .richness ~~
   .cond              0.022    0.139    0.156    0.876    0.022    0.022
 .richness ~~
   .elevation        -0.200    0.119   -1.682    0.093   -0.200   -0.241
 .cond ~~
   .limnology         2.475    0.141    0.672    0.033    0.035    0.148
 .elevation ~~
   .geodiversity      0.889    0.062    0.715    0.041    0.163    0.163


Variances:
                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
   .Carbonates        0.905    0.432    2.126    0.031    0.919    0.919
   .Feldspars         0.772    0.354    2.217    0.025    0.784    0.784
   .Bedrock           0.313    0.133    2.302    0.022    0.307    0.307
   .Elevation         0.687    0.215    3.194    0.001    0.687    0.687
   .Mg                0.989    0.166    5.951    0.000    0.989    0.989
   .HCO3              0.929    0.164    5.682    0.000    0.929    0.929
   .Temp              0.868    0.144    6.038    0.000    0.868    0.868
   .pH                0.935    0.142    6.796    0.001    0.965    0.965
   .TOC               0.921    0.236    3.979    0.000    0.938    0.938
   .cond              0.936    0.358    2.613    0.009    0.936    0.936
```

```
   .NMDS               0.000                                      0.000    0.000
   .richness           0.999    0.145    6.916    0.000    0.999    0.999
    geodiversity       0.081    0.047    1.718    0.046    1.000    1.000
   .limnology         -0.014    0.033   -0.431    0.046   -1.325   -1.325
   .spedist            0.697    0.192    3.632    0.000    0.697    0.697

R-Square:
                   Estimate
    Carbonates        0.081
    Feldspars         0.216
    Bedrock           0.693
    Altitude          0.313
    Mg                0.011
    HCO3              0.071
    Temp              0.132
    pH                0.035
    TOC               0.062
    cond              0.064
    NMDS              1.000
    richness          0.001
    limnology            NA
    spedist           0.303
```

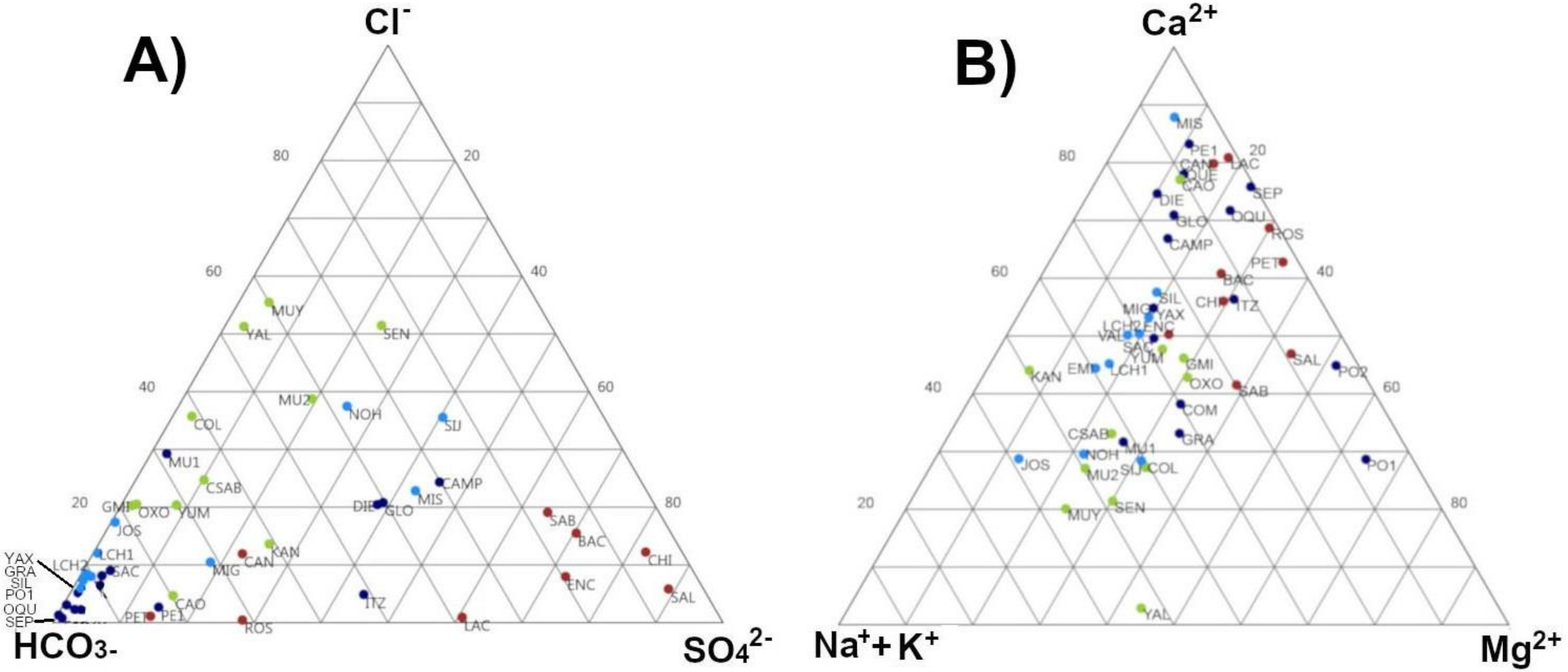

Figure S1. Ternary plots showing major cations and anion proportions [%] of 76 aquatic systems of the northern Neotropical region. Major anions (A) and major cations (B) from YG limnological group; Abbreviations correspond to those in Table S1, colors representing limnological subregions, correspond to the cluster analysis dendrogram (Fig. 2).

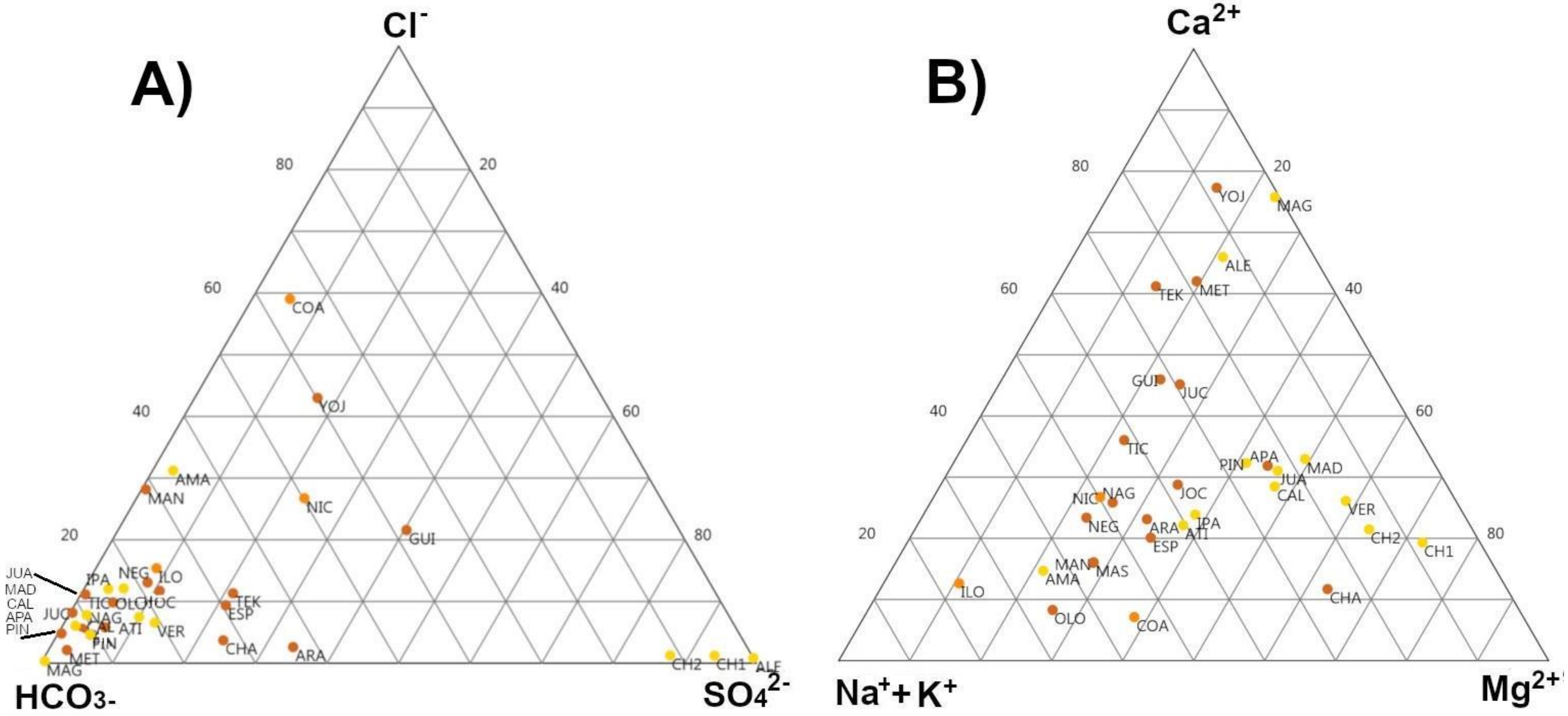

Figure S2. Ternary plots showing major cations and anion proportions [%] of 76 aquatic systems of the northern Neotropical region. A) major anion from GSHN limnological group; d) major cation from GSHN limnological group. Abbreviations correspond to those in Table S1 and colors, representing limnological subregions, correspond to the cluster analysis dendrogram (Fig.

Table S2.1. Loading values of the 13 variables used for principal component analysis of the YG region ($P < 0.05$). Significant scores for components 1 and 2 in bold.

Variable	**PC1 (23.47 %)**	**PC2 (14.63 %)**
Temperature	0.02	0.35
DO	0.06	0.45
pH	0.01	**0.73**
Conductivity	**0.77**	0.04
HCO_3^-	0.13	0.05
Cl^-	**0.64**	0.06
Na+	**0.72**	0.04
Ca	0.51	0.11
Mg2+	**0.65**	0.04
TOC	0.10	0.13
Altitude	0.40	0.12
SO_4^{2-}	0.47	0.12
Age	0.43	0.47

Table S2.2. Loading values of the 13 variables used for principal component analysis of the GSHN limnological region ($P < 0.05$). Significant scores for components 1 and 2 in bold.

Variable	**PC1 (26.83%)**	**PC2 (22.89%)**
Conductivity	0.57	0
HCO_3^-	**0.70**	0.11
Cl^-	**0.62**	0.08
Na+	**0.72**	0.07
Ca	0.33	0.15
Mg2+	0.48	0.05
TN	0.01	0.62
TOC	0.01	**0.79**
Phyllosilicates	0.11	0.20
Altitude	0.11	0.09
CO_3^-	**0.56**	0.17
Bedrock	0.12	**0.75**
Age	0.23	**0.80**

Table S3. List of ostracode species found in our study.

#	Species name	Species code	#	Species name	Species code
1	*Alicenula serricaudata* (Klie, 1935a)	ASE	36	*Hemicypris* sp.1	*Hemicypris* sp.
2	*Alicenula yucatanensis* sp. nov.	AYU	37	*Heterocypris nicaraguensis* Hartmann 1959	HNI
3	*Candona* sp. 1	*Candona* sp.	38	*Heterocypris punctata* (Keyser, 1975)	HPU
4	*Cypretta campechensis* Cohuo-Durán *et al*., 2013	CCA	39	*Keysercypria* sp. 1	*Keysercypria* sp. 1
5	*Cypretta* cf. *campechensis* Cohuo-Durán et al., 2013	CCAf	40	*Keysercypria* sp. 2	*Keysercypria* sp. 2
6	*Chlamydotheca* cf. *colombiensis* Roessler 1985	CCOf	41	*Keysercypria* sp. 3	*Keysercypria* sp. 3
7	*Cypretta elongata* sp. nov.	CEL	42	*Keysercypria* sp. 4	*Keysercypria* sp. 4
8	*Cypretta* cf. *elongata*	CELf	43	*Keysercypria* sp. 5	*Keysercypria* sp. 5
9	*Cypria gibbera* Furtos, 1936[a]	CGI	44	*Keysercypria granadae* (Hartmann 1959)	KGR
10	*Chlamydotheca* sp. 1	*Chlamydotheca* sp.	45	*Limnocythere floridensis* Keyser, 1975	LFL
11	*Cytheridella ilosvayi* Daday, 1905	CIL	46	*Limnocythere* sp. 1	*Limnocythere* sp.
12	*Cypretta maya* Cohuo-Durán *et al*., 2013	CMA	47	*Limnocytherina royi* Hartmann 1959	LRO
13	*Cypria petenensis* Ferguson et al., 1964	CPE	48	*Limnocythere* cf. *stationis* Vávra, 1891	LST
14	*Cypretta spinosa* Cohuo-Durán *et al*., 2013	CSP	49	*Neocypridopsis* sp. 1	*Neocypridopsis* sp.
15	*Cyprinotus unispinifera* Furtos, 1936b	CUN	50	*Pseudocandona antilliana* Broodbakker, 1983c	PAN
16	*Chlamydotheca unispinosa* (Baird, 1862)	CUNI	51	*Peryssocytheridea* cf. *cribrosa* Klie, 1933a	PCRf
17	*Cypridopsis* sp. [Ca 1]	*CVI sp 1*	52	*Penthasilenula* sp. 1	*Penthasilenula* sp.
18	*Cypridopsis* sp. [Ca 2]	*CVI sp 2*	53	*Potamocypris islagrandensis* Hoff 1943b	PIS
19	*Cypridopsis* sp. [Ca 3]	*CVI sp 3*	54	*Pericythere marginata* Hartmann, 1959	PMA
20	*Cypridopsis* sp. 4	*CVI sp 4*	55	*Paracythereis opesta* (Brehm, 1939)	POP
21	*Cypridopsis* sp. 5	*CVI sp 5*	56	*Potamocypris* sp.1	*Potamocypris* sp. 1
22	*Cypridopsis* sp. 6	*CVI sp 6*	57	*Potamocypris* sp. 2	*Potamocypris* sp. 2
23	*Cypridopsis* sp. 7	*CVI sp 7*	58	*Potamocypris* sp. 3	*Potamocypris* sp. 3
24	Cypricercinae sp. 1	CYP 1	59	*Pseudocandona* sp. 1	*Pseudocandona* sp. 1
25	Cypricercinae sp. 2	CYP 2	60	*Pseudocandona* sp. 2	*Pseudocandona* sp. 2
26	Cypricercinae sp. 3	CYP 3	61	*Pseudostrandesia* sp. 1	*Pseudostrandesia* sp. 1
27	*Cyprinotinae* sp. 1	CYP 4	62	*Pseudostrandesia* sp. 2	*Pseudostrandesia* sp. 2
28	*Cypria* sp. 1	*Cypria* sp. 1	63	*Strandesia bicuspis* (Claus 1892)	SBI
29	*Cypria* sp. 4	*Cypria* sp. 4	64	*Stenocypris cylindrical major* (Baird, 1859b)	SCY
30	*Cypria* sp. 5	*Cypria* sp. 5	65	*Strandesia intrepida* Furtos, 1936b	SIN
31	*Cyprideis* cf. *salebrosa*	*Cyprideis* sp.	66	*Strandesia* sp. 1	*Strandesia* sp.
32	Cyprididae sp. [Ca 1]	Cyprididae sp. 1	67	*Tanycypris* sp. 1	*Tanycypris* sp.
33	*Desc. 1*	Cyprididae sp.2	68	*Thalassocypria* sp.1	*Thalassocypria* sp.
34	*Diaphanocypris meridana* (Furtos, 1936b)	DME	69	*Vestalenula* sp.1	*Vestalenula* sp.
35	*Darwinula stevensoni* (Brady & Robertson, 1885)	DST	70	*Vestalenula pagliolii* (Pinto and Kotzian 1961)	VPA