11-5-2012

Biodiversity, water chemistry, physical characteristics, and anthropogenic disturbance gradients of sandstone springs on the Colorado Plateau

Rebecca H. Weissinger  
*Western State College, Gunnison, CO, rebecca_weissinger@nps.gov*

Dustin W. Perkins  
*National Park Service, Moab, UT, dustin_w_perkins@nps.gov*

Eric C. Dinger  
*Utah State University, Logan, UT, eric_dinger@nps.gov*

---

Follow this and additional works at: https://scholarsarchive.byu.edu/wnan

Part of the Anatomy Commons, Botany Commons, Physiology Commons, and the Zoology Commons

---

Recommended Citation

Available at: https://scholarsarchive.byu.edu/wnan/vol72/iss3/13
ABSTRACT.—Springs located on the Colorado Plateau are highly threatened and represent a small percentage of the landscape; yet they are disproportionately important to diverse native flora and fauna. The relationships between anthropogenic disturbance, aquatic macroinvertebrate species composition, and environmental variables at these springs have received little study. We selectively visited 40 sandstone springs in southeastern Utah and southwestern Colorado to span a range of impacts. We classified the springs into impact categories based on a spring impact score, and we measured biodiversity (aquatic macroinvertebrates), water chemistry (nutrients, dissolved O$_2$, pH, specific conductivity, temperature, turbidity, coliform bacteria [Escherichia coli]), physical characters (solar radiation, substrate, vegetation cover, bank stability, discharge), and presence of anthropogenic disturbance. Escherichia coli abundance was higher in high impact categories, and turbidity increased with increasing disturbance. No differences in total N, total P, specific conductivity, flow, dissolved O$_2$, pH, or substrate were found among the impact categories. Vegetation cover was higher in low impact categories than in moderate and high impact categories, while potential annual and growing-season solar radiation was lower in low impact categories than in high impact categories. Global and subsequent multiple response permutation procedure (MRPP) comparisons suggested strong differences in aquatic macroinvertebrates between low and high impact springs and no difference at moderate impact springs. Mean taxa richness (α-diversity), total taxa richness (γ-diversity), and percent of taxa richness composed of shredders peaked at moderate disturbance levels. The percentage of non-insect taxa richness was reduced in high impact categories, and Odonata (dragonflies and damselflies) were higher in low impact categories than in high impact categories. All high impact springs had both livestock use and vehicle use (roads or off-highway vehicles), and our data suggest that disturbances caused by one or both of these uses alter the aquatic macroinvertebrate assemblage. We suggest that disturbance may increase macroinvertebrate richness, where a mix of tolerant and intolerant species co-occur, until macroinvertebrate richness reaches a threshold; after surpassing this threshold, macroinvertebrate diversity decreases.

RESUMEN.—Los manantiales ubicados en Colorado Plateau están sumamente amenazados y representan un pequeño porcentaje del paisaje. Sin embargo, son desproporcionadamente importantes para una gran variedad de flora y fauna nativas. La relación entre la perturbación antropogénica, la composición de las especies de macroinvertebrados acuáticos y las variables ambientales en los manantiales de Colorado Plateau han sido objeto de poco estudio. De manera selectiva, visitamos 40 manantiales de piedra arenisca en el suroeste de Utah y el suroeste de Colorado para abarcar una serie de impactos, y los dividimos en categorías de impacto, según una calificación de Impacto en Manantiales. Medimos la biodiversidad (macroinvertebrados acuáticos), la química del agua (nutrientes, oxígeno disuelto, pH, conductividad específica, temperatura, turbidez, bacterias coliformes [Escherichia coli]), características físicas (radiación solar, sustratos, cubierta vegetal, estabilidad de los bancos, descargas) y la presencia de perturbación antropogénica. La abundancia de Escherichia coli fue mayor en las categorías de Gran Impacto y la turbidez aumentó con el aumento de la perturbación. No se encontraron diferencias en el nitrógeno total, fósforo total, conductividad específica, caudal, oxígeno disuelto, pH o sustratos entre las diferentes categorías de impacto. La cubierta vegetal fue mayor en las categorías de Bajo Impacto que en las categorías de Moderado y Alto Impacto, mientras que la posible radiación solar anual y en la temporada de crecimiento fue menor en las categorías de Bajo Impacto que en las categorías de Alto Impacto. Las comparaciones realizadas a través del procedimiento de permutación de respuestas múltiples (multiple response permutation procedure, MRPP) tanto globales como subsecuentes sugirieron una marcada diferencia en los macroinvertebrados acuáticos entre los manantiales de Alto y Bajo Impacto, y ninguna diferencia en los manantiales de Moderado Impacto. La riqueza promedio de taxa (diversidad α), la riqueza total de taxa (diversidad γ) y el porcentaje de la riqueza de taxa compuesta de trituradores alcanzaron el máximo en niveles moderados de perturbación. El porcentaje de riqueza de taxa de insectos se redujo en las categorías de Alto Impacto y la presencia de Odonatos (libélulas) fue más elevada en las categorías de Bajo Impacto que en las categorías de Alto Impacto. Todos los manantiales de Alto Impacto tuvieron la presencia de ganado y de vehículos (vehículos de carretera o fuera de carretera), y nuestros datos sugieren que uno o ambos de estos impactos alteran el ensamblaje de macroinvertebrados acuáticos. Sugerimos que la perturbación puede aumentar la riqueza de macroinvertebrados, donde un conjunto de especies tolerantemente y no tolerantemente coocurren, hasta que alcanza cierto nivel; luego de alcanzar este nivel, la diversidad de los macroinvertebrados disminuye.
In arid and semiarid regions, springs occupy a small fraction of the total landscape, yet they support disproportionately high levels of productivity, endemism, and biodiversity (Erman 2002, Hershler and Sada 2002, Sada et al. 2005). At broader spatial scales, springs play key roles in arid ecosystems by providing refugia for migratory birds (Skagen et al. 1998), reptiles, and amphibians. This capacity supports high rates of primary and secondary production, and springs often provide the primary water source for larger habitats (streams and ponds).

As springs are a source of water in arid and semiarid regions, human activities are often concentrated around them, making springs highly susceptible to anthropogenic impacts (Shepard 1993). Specifically, surface and groundwater diversions, livestock use, and recreational activities are disproportionately high in these areas and represent primary threats to spring ecosystems. A study in northern Arizona classified more than 93% of springs on non-National Park Service (NPS) federal lands as ecologically impaired or functioning at risk (Grand Canyon Wildlands Council 2002). Another study of 505 springs in northern Nevada found that 85% of the springs were disturbed by livestock and diversions and 95% of the springs were affected by human disturbances (Sada et al. 1992). Overall, the condition of springs on the Colorado Plateau, as well as other western arid and semiarid areas, is poorly known.

Aquatic macroinvertebrate assemblages and metrics have been widely used as indicators of water quality in wadeable stream and riverine ecosystems (Cairns and Pratt 1993, Lenat and Barbour 1994, Barbour et al. 1999), and they have been suggested for use as bioindicators in spring ecosystems (Cantonati et al. 2006, Keleher and Rader 2008). Aquatic macroinvertebrates are useful as bioindicators because they (1) occupy critical positions in the food web; (2) encompass multiple feeding strategies, relying on both autochthonous and allochthonous production; (3) have multiple life history strategies, including being short-lived (i.e., rapidly responding); (4) have multiple habitat specializations; and (5) are relatively easy to sample and identify to functional taxonomic levels (Merritt et al. 2008). Furthermore, the susceptibility of macroinvertebrates to pollution functions to retain a biological memory of events that other water quality sampling methods may not detect.

Indicators that are used to assess the biotic integrity of water bodies in other ecoregions may not be applicable in the Southwest. For example, use of tolerance values, where individual taxa are assigned a value of 0 (intolerant) to 10 (tolerant) based on empirical evidence, is a common method of aquatic bioassessment (Hilsenhoff 1987). However, there is limited knowledge of both tolerance values and general information about aquatic macroinvertebrates in the American Southwest (Blinn and Ruiter 2006), and we know of none for springs on the Colorado Plateau. Furthermore, some diversity and abundance indicators (e.g., EPT—Ephemeroptera, Plecoptera, and Tricoptera [mayflies, stoneflies, and caddisflies]) that have been used in many regions are underrepresented in Colorado Plateau springs. This underrepresentation is especially true in smaller, isolated, low-flow springs that are subject to naturally high levels of disturbance, such as fluctuations in temperature and water availability. Current metrics for the Colorado Plateau have had limited testing in arid and semiarid springs, and these metrics are in need of development and validation to determine their effectiveness in bioassessment of anthropogenic impacts.

Our objective was to explore the use of aquatic macroinvertebrates as bioindicators for monitoring ecosystem impairment in arid and semiarid spring ecosystems, so we conducted integrated assessments of the biological, chemical, and physical condition of selected springs in southeastern Utah and southwestern Colorado. Specific questions we addressed included the following: Do water chemistry and physical characteristics show a response to a gradient of increasing anthropogenic activities? And do aquatic macroinvertebrate assemblages show a response to a gradient of increasing anthropogenic activities?

**Methods**

**Study Area and Sampling Sites**

The Colorado Plateau is a semiarid geographic province bounded by the Uinta and Wasatch Mountains of Utah to the north, the Rocky Mountains of Colorado to the east, the Mogollon Rim of Arizona to the south, and the Great Basin of Nevada and Utah to the west. The region ranges between 1200 and 3350 m
in elevation and receives 15 to 40 cm of rain annually. This study focuses more narrowly on southeastern Utah and southwestern Colorado (Fig. 1).

Forty springs were visited between May and September 2008 and in May 2009. These springs were selected by NPS and Bureau of Land Management (BLM) resource specialists
to represent a range of anthropogenic disturbances. Sites were selected expressly for obtaining near-equal sampling effort across impact categories. Probabilistic sampling would have resulted in too many springs being (1) inaccessible, (2) dry, and (3) predominantly moderate and high impact sites. Furthermore, it was not our objective to extrapolate status to a larger population. Sampling was not restricted to any single spring type (e.g., limnocrene, discharge amount, etc); however, we aimed to apply our results to any spring in the region. Springs tended to be mostly coldwater and of low discharge. Sampled springs emerged from sandstone aquifers and were located between 1200 and 2000 m in elevation.

Spring Impact Categories

We created an overall index of spring impact based on the combination of 2 metrics: (1) presence/absence of anthropogenic disturbance and (2) physical degradation.

Presence/absence of near-field anthropogenic disturbance (within 0.2 km) was based on counting the occurrence of 5 different types of activities: roads and off-highway vehicles [OHV], hiking trails, recreational use (camping, bathing, etc.), livestock use (grazing, trampling, fecal matter), and flow modifications (diversions). Thus, the possible range of disturbances was 0–5; however, springs sampled had a range of 0–4 disturbances. Scores were based on presence/absence to create a rapid scoring process, and these scores do not account for degree of disturbance.

Physical degradation was based on the average of 2 separate environmental parameters: (1) vertical bank stability and (2) riparian soil disturbance. Vertical bank stability was estimated around the main spring pool, where the cut banks were nearly vertical and actively eroding because they lacked significant root stabilization (Stacey et al. 2007). Bank stability was scored on extent of erosion: <5% = 1, 5%–30% = 2, 31%–60% = 3, 61%–90% = 4, and >90% = 5. Riparian soil disturbance surrounding the spring was estimated as soil displacement and compaction due to humans, livestock, and native ungulates. Similar to bank stability, scores were assigned as follows: <1% = 1, 1%–5% = 2, 5%–15% = 3, 16%–25% = 4, and >25% = 5. Overall physical degradation (low, moderate, and high) was assigned based on the average of the 2 measures: <2 = low, 2–3 = moderate, and >3 = high.

The final spring impact score was then calculated as the sum of anthropogenic disturbances and physical degradation scores (low = 1, moderate = 2, high = 3). Possible scores ranged from 1 to 7. Spring condition was then categorized as low impact (2 or lower), moderate impact (3 or 4), and high impact (5 or higher).

Chemical and Physical Variables

We estimated instantaneous discharge of the spring output by funneling water through a pipe at the most convenient location (e.g., minimal disturbance to the ecosystem). We then measured how long it took to fill a container of known volume (e.g., 20 s to fill 500 mL) and then converted that time to standard discharge units (m$^3$ · s$^{-1}$). Six readings were averaged for the final discharge estimate.

We measured dissolved oxygen, pH, specific conductivity, and temperature using a calibrated In-Situ TROLL 9500 multiparameter probe. Also, we took a 500-mL sample to measure total nitrogen and total phosphorous at a certified water quality laboratory. A bottle was filled either by dipping in the deepest section of the spring or by siphoning off water from the pipe if funneled for discharge measurements. We measured turbidity with a calibrated Hach 2100P turbidimeter. As part of regular water quality sampling, we also tested for fecal coliform (Escherichia coli). A 100-mL coliform sample was collected and immediately placed on ice. We processed samples within 8 h of the collection time using the Idexx Colilert system and incubated the samples for 24 h at 35 °C before taking a final reading.

Potential solar radiation was recorded as near to the spring’s orifice as possible using a Solar Pathfinder™, which estimates the available solar radiation based on the horizon, vegetation, and other topographic features. We then calculated an average solar radiation for both the entire year and the growing season (May–September). We visually estimated the percentage of the main pool (when the spring outflow formed a pool) or channel’s bank (when the spring outflow formed a stream) that was overhung with live vegetation, using gross categories (1 = no overhanging vegetation, 2 = <10% of banks with overhanging cover, 3 = 10%–25%, 4 = 26%–50%, 5 = >50%; Stacey et al. 2007).
presence or absence of live vegetation was recorded at 2-m intervals along the entire riparian perimeter or along a transect of up to 100 m at large sites (Stacey et al. 2007). We sampled substrate using a modified Wolman Pebble count of 50 haphazardly selected particles. Pebbles were then binned into size categories of fines (<2 mm), gravel (2–45 mm), and cobble (46–180 mm).

Biological Variables

The objective of the aquatic macroinvertebrate sampling at each site was to collect as many different kinds of invertebrates as possible while minimizing research impacts to the site. We either sampled the entirety of the habitat at smaller springs or sampled for up to 3 h at larger springs. We collected samples with a rectangular kick net and hand net, each with 500-micron mesh, and then handpicked invertebrates from vegetation and substrate. All spring-associated microhabitats (e.g., depositional zones, springhead, outflows, aquatic macrophytes, gravel, etc.) were sampled, including standing water at the spring site, outflows, and small water pockets surrounding each spring. Although handpicking invertebrates is generally qualitative, by collecting invertebrates in proportion to their occurrence (i.e., the more individuals of a particular taxa present, the more they were added to the collection), we were able to use numbers collected as a semi-quantitative measure of relative abundance. Constant field personnel utilized equal effort in handpicking, and this equal effort assured that relative abundance comparisons were valid. Samples collected at all microhabitats were composited to form a single sample for each site. Invertebrates were identified and stored in 70% ethanol. Ambiguous taxa (instances of lower taxonomic resolution and with individuals of differing resolution) were resolved by removing parents (individuals of lower taxonomic resolution) or by merging children (individuals with higher taxonomic resolution) with parents based on abundances at different taxonomic levels (Cuffney et al. 2007). Macroinvertebrate metrics, including percent richness, composition of various taxonomic groups, functional feeding groups, and tolerance groups, were calculated using the Invertebrate Data Analysis System (IDAS v 3.0; Cuffney 2003). For macroinvertebrate data, we also calculated mean taxa richness (α-diversity), total taxa richness (γ-diversity), and taxa turnover ([β-diversity, defined as total species richness/mean species richness – 1]) for macroinvertebrate data (Whittaker 1960).

Data Analysis

Differences in univariate response variables (biological, chemical, and physical) between impact categories (low, moderate, and high) were tested either with a one-way ANOVA when normality and variance assumptions were met or with the nonparametric Kruskal–Wallis test when not. Normality was assessed using the Shapiro–Wilk’s test for normality, and equality of variance was assessed using Levene’s test (Sokal and Rohlf 1995). When appropriate, proportion data were arcsine square-root transformed to meet normality assumptions. All univariate analyses were completed in R 2.10.0 (R Development Core Team 2009). Posthoc pairwise comparisons with a Holm adjustment were done using pairwise $t$ tests for parametric data and Wilcoxon’s signed-ranks tests for nonparametric data. Statistical significance was assessed at the $\alpha = 0.05$ level.

Prior to the community analyses, a series of data-screening procedures and transformations were done. We deleted rare species (defined here as taxa found at only one spring), because they increase the signal-to-noise ratio and obscure patterns in the data (Marchant 2002, Van Sickle et al. 2007). Relative abundance data were log transformed [$\log_{10}(x + 1)$] to reduce the influence of hyperabundant taxa.

Nonmetric multidimensional scaling (NMDS) is an indirect gradient analysis technique that uses rank community dissimilarities to iteratively search for the optimal arrangement of sample objects in as few dimensions as possible (McCune and Grace 2002). NMDS was run with Sørensen’s distance in PC-ORD version 5 (McCune and Mefford 1999). We assessed dimensionality and adequacy of the final configuration by evaluating the relationship of final stress versus the number of dimensions; in addition, a Monte Carlo test with 250 runs of the randomized data quantified the probability of observing a stress as low as or lower than that observed through chance alone.

We tested for differences in assemblage composition among groups by using a multiple response permutation procedure (MRPP) with Sørensen’s distance (Mielke and Berry 2001).
**TABLE 1.** Chemical and physical variable correlations with impact categories. Bold type indicates a significant difference in that variable ($\alpha = 0.05$). Kruskal–Wallis test results are presented as medians with a $\chi^2$ test statistic and df = 2. ANOVA test results are presented as means with an F test statistic and df = 2, 37. Wilcoxon’s signed-ranks test results are presented for nonparametric pairwise comparisons with a W test statistic and Holm-adjusted P value. T test results are presented for parametric pairwise comparisons with a t test statistic and Holm-adjusted P value. NA = not applicable.

<table>
<thead>
<tr>
<th>Impact category</th>
<th>Low</th>
<th>Moderate</th>
<th>High</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Response variable (SE) (range)</strong></td>
<td><strong>Mean</strong></td>
<td><strong>Median</strong></td>
<td><strong>Mean</strong></td>
<td><strong>Median</strong></td>
</tr>
<tr>
<td><em>Escherichia coli</em> (MPN · 100 mL$^{-1}$)</td>
<td>NA</td>
<td>&lt;0.01</td>
<td>NA</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Turbidity (ntu)</td>
<td>NA</td>
<td>0.74 (0.01–124)</td>
<td>NA</td>
<td>2.88 (0.93–11.2)</td>
</tr>
<tr>
<td>Specific conductivity (μS · cm$^{-1}$)</td>
<td>NA</td>
<td>565.29 (197.76–1175.24)</td>
<td>NA</td>
<td>613.66 (255.07–3381)</td>
</tr>
<tr>
<td>Dissolved oxygen (% saturation)</td>
<td>96.39 (7.61)</td>
<td>81.26 (8.21)</td>
<td>NA</td>
<td>76.91 (12.82)</td>
</tr>
<tr>
<td>pH</td>
<td>8.02 (0.11)</td>
<td>7.93 (0.11)</td>
<td>NA</td>
<td>7.94 (0.17)</td>
</tr>
<tr>
<td>Total nitrogen (mg · L$^{-1}$)</td>
<td>NA</td>
<td>0.02 (&lt;0.01–11)</td>
<td>NA</td>
<td>0.006 (&lt;0.01–6.4)</td>
</tr>
<tr>
<td>Total phosphorous (mg · L$^{-1}$)</td>
<td>NA</td>
<td>&lt;0.01 (&lt;0.01–0.58)</td>
<td>NA</td>
<td>&lt;0.01 (&lt;0.01–0.17)</td>
</tr>
<tr>
<td>Vegetation cover (%)</td>
<td>34.54 (2.03)</td>
<td>25.5 (3.31)</td>
<td>NA</td>
<td>18.6 (3.08)</td>
</tr>
<tr>
<td>Overbank vegetation cover (%)</td>
<td>NA</td>
<td>75 (1–75)</td>
<td>NA</td>
<td>38 (1–75)</td>
</tr>
<tr>
<td>Potential annual solar radiation (%)</td>
<td>51.4 (6.38)</td>
<td>68.36 (17.65)</td>
<td>NA</td>
<td>85.2 (4.83)</td>
</tr>
<tr>
<td>Potential growing season solar radiation (%)</td>
<td>NA</td>
<td>66 (0–97.6)</td>
<td>NA</td>
<td>83.6 (4–100)</td>
</tr>
</tbody>
</table>
MRPP is a nonparametric permutation procedure that tests for differences among 2 or more a priori groups. A $P$ value assesses the probability of observed group differences between low, moderate, and high impact categories under the null hypothesis, while an $A$ statistic quantifies effect size and within-group homogeneity (McCune and Grace 2002). We used both global MRPP (testing for differences among all groups) and pairwise MRPP (testing for difference between just 2 groups). MRPP was also used as a follow-up analysis to understand the specific effects of anthropogenic disturbances on the macroinvertebrate assemblages. Indicator species analysis (ISA) (Dufrene and Legendre 1997) was further used to determine taxa with high fidelity to the different groupings. Both ISA and MRPP were analyzed using PC-ORD 5.

**RESULTS**

Sample sizes for the spring impact categories were 15 for low impact, 15 for moderate impact, and 10 for high impact. Trails and recreational use (camping, bathing, human trampling) were the only disturbances associated with low impact springs. Springs in the moderate impact category often had one additional disturbance to the presence of trails, such as vehicle use (road or OHV), livestock use, or a flow modification. High impact springs had both livestock use and vehicle use present, in addition to other disturbances. Allocation of spring impact categories was unrelated to elevation and spring discharge, 2 potentially confounding variables.

**Chemical and Physical Variables**

Turbidity, vegetation cover, *E. coli*, and potential solar radiation varied significantly between our impact categories (Table 1). Concentrations of fecal coliform bacteria (*E. coli*) were higher at springs in the high impact category than at springs in the low and moderate impact categories. Turbidity increased with increasing disturbance (Table 1). Vegetation cover decreased as disturbance increased. Potential annual and growing-season solar radiation were lower at low impact than at high impact springs (Table 1). Substrate did not differ between categories. Also, overhanging vegetation cover, dissolved oxygen, total nitrogen (nitrate/nitrite), total phosphorous, specific con-
ductivity, and pH showed no significant relationship to impact categories.

Biological Variables

The total number of macroinvertebrate taxa encountered from all 40 springs was 76, but 21 occurred only at a single spring and were omitted from the final multivariate dataset, making a total of 55 taxa. Mean taxa richness (α-diversity) and total taxa richness (γ-diversity) were highest in moderate impact springs (Table 2), while taxa turnover (β-diversity) was similar between all the impact categories.

A 3-dimensional NMDS ordination was found to be the most parsimonious solution (stress = 17.52, instability = 0.0001, iterations = 165). The 3 axes represent 79.2% of the variation in the dataset (Axis 1 = 17.0, Axis 2 = 40.2, Axis 3 = 22.0). For ease of visualization, the axes representing the most variation (Axes 2 and 3) are presented in Figure 2. The axes do not have inherent meaning; rather the configuration of points in space indicates similarities between aquatic macroinvertebrate communities, with similar sites lying closer together and dissimilar sites lying farther apart.

Visual interpretation of the ordination suggests that macroinvertebrate assemblages in the same impact category generally lie closer together than sites in different impact categories, although assemblages in moderate impact springs overlap those in both low and high impact springs. Global MRPP confirms that aquatic macroinvertebrate assemblages differ between impact categories but with a relatively small effect size (A = 0.0143, P = 0.0281). Follow-up pairwise MRPP comparisons suggested a strong difference between low and high impact springs (A = 0.0231, P = 0.0045) and no difference between moderate impact springs and either low or high impact springs (low vs. moderate: A = 0.0039, P = 0.2643; moderate vs. high: A = 0.0077, P = 0.1553).

The presence of trails, nonmotorized recreational use, flow modifications, and visibility from a nearby road or trail did not significantly affect aquatic macroinvertebrate communities (Table 3). However, both livestock use and motorized traffic (roads and OHV use) showed significant differences in macroinvertebrate assemblages, although most sites with livestock use also had vehicular use and the effects of these 2 disturbances could not be separated using our data.

![Fig. 2. Nonmetric multidimensional scaling ordination of aquatic macroinvertebrate assemblage data classified by impact category. Sites with similar assemblages lie closer together.](image)

**Table 2.** Macroinvertebrate diversity measures for each impact category. n = sample size, definitions of α- and β-diversity as in Magurran (2004), γ-diversity from Whittaker (1960).

<table>
<thead>
<tr>
<th>Impact category</th>
<th>n</th>
<th>Mean taxa richness (α-diversity)</th>
<th>Taxa turnover (β-diversity)</th>
<th>Total taxa richness (γ-diversity)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>15</td>
<td>11.1</td>
<td>3.86</td>
<td>54</td>
</tr>
<tr>
<td>Moderate</td>
<td>15</td>
<td>14.1</td>
<td>3.4</td>
<td>62</td>
</tr>
<tr>
<td>High</td>
<td>10</td>
<td>8.5</td>
<td>3.59</td>
<td>39</td>
</tr>
</tbody>
</table>

**Table 3.** Pairwise MRPP tests of differences in macroinvertebrate assemblages based on presence or absence of impact type. N = number of sites with that impact present out of 40. A = measure of effect size. Bold text indicates significant differences at α = 0.05.

<table>
<thead>
<tr>
<th>Impact type</th>
<th>N</th>
<th>A</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Presence of trails</td>
<td>29</td>
<td>0.002</td>
<td>0.299</td>
</tr>
<tr>
<td>Recreation use (e.g., camping)</td>
<td>14</td>
<td>0.001</td>
<td>0.343</td>
</tr>
<tr>
<td>Flow modifications</td>
<td>7</td>
<td>0.001</td>
<td>0.343</td>
</tr>
<tr>
<td>Visibility from road or trail</td>
<td>35</td>
<td>0.007</td>
<td>0.079</td>
</tr>
<tr>
<td>Livestock use</td>
<td>16</td>
<td>0.0217</td>
<td>0.004</td>
</tr>
<tr>
<td>Motorized traffic (road and OHV)</td>
<td>14</td>
<td>0.0172</td>
<td>0.02</td>
</tr>
</tbody>
</table>
MACROINVERTEBRATE METRICS.—Of the 18 macroinvertebrate metrics, only 4 showed significant differences among spring impact categories (Table 4). Springs with moderate impact levels had significantly higher taxa richness than springs with high impact levels. Percent taxa richness of non-insects (Gastropoda, Crustacea, Oligochaeta, etc.) were significantly reduced in high impact springs. The percent richness comprising Odonata (dragonflies and damselflies) was higher at springs with low impact than those with high impact. The percent richness composed of shredders was higher in springs with moderate impact than in those with high impact. There were no differences detected among other metrics of taxa groups (e.g., percent richness of Ephemeroptera), functional feeding groups (e.g., percent richness of predators [from Merritt et al. 2008]), and tolerance values (e.g., percent richness of tolerant organisms; Table 4).

INDICATOR TAXA.—Only 2 taxa were statistically significant indicators of our impact categories. The whirligig beetle, Gyrinus sp. (Coleoptera: Gyrinidae), was an indicator of moderate impact sites (IV = 33.8, \( P = 0.0440 \)), while the mosquito, Culex sp. (Diptera: Culicidae), was an indicator of high impact sites (IV = 25.1, \( P = 0.0308 \)). It should be noted that 2 significant results from the testing of 55 taxa is near the number that might be expected from simple chance. This finding, combined with low indicator values (IV), suggests that there is no biological significance to the ISA analysis.

DISCUSSION

Our data suggest that aquatic macroinvertebrate assemblages respond differently across impact categories, so some utility exists in using macroinvertebrates as bioindicators of ecological integrity. In particular, aquatic macroinvertebrate assemblages at springs with high levels of anthropogenic impact are less diverse compared to springs of low and moderate impact. Diversity increased with moderate levels of disturbance but then declined with high levels of disturbance. However, very few of the macroinvertebrate metrics demonstrated patterns across impact categories. We suspect that the general lack of detectable macroinvertebrate differences between low and moderate disturbances, combined with higher levels of diversity in moderately disturbed springs, is due to these assemblages evolving with moderate to high levels of natural disturbances, including native ungulate grazing, aboriginal use of springs, flashfloods, and the possibility of drying during extensive droughts. Some aspects of disturbance (e.g., increased nutrient inputs from livestock; Hubbard et al. 2004) could increase macroinvertebrate diversity due to increased rates of primary and secondary production.

There appears to be a disturbance threshold between where most invertebrate species are unaffected and where diversity actually increases with some level of disturbance. Similar to our study, Keleher and Rader (2008) found that macroinvertebrates in Great Basin springs did not respond to the effects of moderate grazing, and it was difficult to distinguish the macroinvertebrate community between moderately and minimally impacted sites. Also, Sada et al. (2005) found that species richness in springs of Nevada’s White Mountains was greatest at intermediate levels of disturbance. Kodric-Brown and Brown (2007) emphasized that springs in arid lands have historically experienced major disturbances from Pleistocene large mammals to aboriginal peoples and livestock grazing.

We hypothesize that livestock use and/or OHV use (or their interaction) observed in this study are above ambient levels of disturbance, and thus observed disturbance is higher than these species can naturally tolerate. All high impact springs had both grazing and vehicle use present. We had only 3 springs with only livestock use and 2 springs with only OHV use, making the replications too few to differentiate the effects of these 2 disturbance types. However, no other type of disturbance (hiking trails, flow modification, visibility from the road, nonmotorized recreational use) detected in this study had noticeable effects on the aquatic macroinvertebrate community. In addition to the effects observed in the macroinvertebrate assemblages, springs in the high impact categories also had lower vegetative cover and soil stability, higher solar radiation, and higher levels of E. coli.

Our observations are not dissimilar to other researchers. Similar impacts to those seen in our study, including soil disturbance, bank instability, and a decrease in vegetative cover, have been reported for livestock grazing (Kauffman and Krueger 1984, Fleischner
<table>
<thead>
<tr>
<th>Macroinvertebrate metric</th>
<th>Impact category</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low</td>
<td>Median (Range)</td>
</tr>
<tr>
<td></td>
<td>Mean (SE)</td>
<td></td>
</tr>
<tr>
<td>Taxa richness</td>
<td>11.13 (1.05)</td>
<td>NA</td>
</tr>
<tr>
<td>Non-insects (%) richness</td>
<td>NA</td>
<td>11.1 (0–33.3)</td>
</tr>
<tr>
<td>Odonata (%) richness</td>
<td>NA</td>
<td>18.2 (0–27.3)</td>
</tr>
<tr>
<td>Coleoptera (%) richness</td>
<td>NA</td>
<td>25 (16.67–63.64)</td>
</tr>
<tr>
<td>Diptera (%) richness</td>
<td>NA</td>
<td>9.09 (0–23)</td>
</tr>
<tr>
<td>Non-chironomid Diptera (%) richness</td>
<td>NA</td>
<td>9.09 (0–23)</td>
</tr>
<tr>
<td>Non-chironomid Diptera and non-insects (%) richness</td>
<td>NA</td>
<td>31.25 (0–41.67)</td>
</tr>
<tr>
<td>Ephemeroptera (%) richness</td>
<td>NA</td>
<td>9.09 (0–20)</td>
</tr>
<tr>
<td>EPT (%) richness</td>
<td>NA</td>
<td>11.11 (0–22.22)</td>
</tr>
<tr>
<td>Gastropoda (%) richness</td>
<td>NA</td>
<td>0 (0–16.67)</td>
</tr>
<tr>
<td>Shredders (%) richness</td>
<td>NA</td>
<td>0 (0–25)</td>
</tr>
<tr>
<td>Filterer-collectors (%) richness</td>
<td>NA</td>
<td>0 (0–8.33)</td>
</tr>
</tbody>
</table>
1994, Foote and Hornung 2005) and for OHV vehicle use in arid and semiarid riparian areas (Herbst 2004, Ouren et al. 2007). Grazing impacts to macroinvertebrates were documented by Myers and Resh (1999), who observed both the extirpation of a caddisfly population in a Great Basin spring when cattle were allowed to graze and the subsequent recovery of the population when cattle grazing was eliminated.

It is conceivable that some level of livestock grazing or vehicle use could mimic moderate types of disturbance. Kodric-Brown and Brown (2007) provide examples where livestock grazing played an important role in reducing aquatic vegetation (thus maintaining open water habitat and high oxygen concentrations) and preventing the extirpation of native fish in desert springs of Nevada and Australia. Additional work is needed to separate the effects of these 2 disturbance types and to directly quantify different levels of their impacts on aquatic macroinvertebrate communities. Such work would benefit aridland managers.

Macroinvertebrate indicators of anthropogenic disturbance of Colorado Plateau springs appear to be different than the typical indicators found in other regions. We found no difference in many traditional indicators (EPT taxa, Diptera, National Tolerance Groupings, or functional feeding groups) of aquatic integrity among the different impact categories despite differences in soil stability, vegetation cover, solar radiation, and E. coli. Low and moderate impact springs were typified by a higher percentage of species richness composed of non-insect taxa (worms, spiders, amphipods, crustaceans, amphibians, mollusks, and flatworms), while low impact sites had more Odonata species. High impact sites were typified by having low numbers of Odonata and high numbers of Culex. Indicators of less-disturbed springs (in both the low and the moderate impact categories) in our study are generally considered indicators of moderate to poor water quality in more mesic areas or in lotic systems.

Regional disturbances that were not addressed in our study include aquifer depletion and impacts of agricultural, industrial, and urbanization effects. It is possible that we could have found more macroinvertebrate indicators of ecosystem degradation of Colorado Plateau springs if we had extended our study to these
impacts. Boyle and Strand (2003) found that macroinvertebrate assemblages of the East Fork of the Virgin River were more associated with nitrogen-related variables influenced by agricultural and urban activities. In the North Fork of the Virgin River, assemblages were associated with natural conditions, despite both streams being similar in topography, elevation, and terrestrial vegetation cover. Macroinvertebrates are also responsive to disturbances at different scales of impact and are good indicators for use with small (Blinn and Ruiter 2006), medium (Boyle and Strand 2003), and large (Blinn et al. 1995, Vinson 2001) rivers on the Colorado Plateau.

An important qualifier of our work is that our sampling design was not a probabilistic random sample of springs on the Colorado Plateau. Sites were selected expressly for obtaining near equal sampling efforts across impact categories. Thus, our results should not be taken as evidence that 25% of Colorado Plateau springs are highly impacted. This percentage merely expresses the sampling effort targeted towards this impact category. The proportion of springs heavily impacted on public lands, both BLM and NPS, can be assessed only with a random, probabilistic sample.

A combination of habitat heterogeneity and the natural isolation of springs in arid systems has complicated the development of bioindicators of anthropogenic disturbances. Because traditional univariate bioindicators using macroinvertebrates largely did not respond to impact levels, but multivariate analysis did show differences in assemblage structure, future research should explore the possibility of developing multivariate predictive models of biodiversity. These models, commonly called O/E (observed/expected) models, use general predictor variables (e.g., geology, discharge, latitude, longitude, elevation) to predict the composition of invertebrate assemblages when a site is unimpaired. A ratio of the number of observed taxa to the number predicted (or expected) can then be used to assess ecosystem health.

Conclusions

Springs on the Colorado Plateau remain important biodiversity hotspots due to their relative rarity, use by disproportionate numbers of species based on the area occupied, and high levels of plant endemism. Our study highlights the importance of creating habitat-specific metrics when using aquatic invertebrates for bioassessment. We conclude that, while the macroinvertebrate community of Colorado Plateau springs is likely adapted to some levels of disturbance, the levels of combined grazing and vehicle use found in this study affected some biotic (aquatic macroinvertebrates, vegetative cover, E. coli) and physical (turbidity and solar radiation) components of these springs. Our study suggests that several invertebrate metrics correlated to livestock grazing and vehicle use at springs on the Colorado Plateau. However, we did not quantify the amount of grazing and vehicle use at each site, and the 2 impacts are confounded in our sample.

Acknowledgments

We thank Mary Moran and Jeff Troutman from the National Park Service and Anne Marie Aubrey from the Bureau of Land Management for logistical support and discussion of potential study sites. The Bureau of Land Management/Utah State University National Aquatic Monitoring Center, in particular Mark Vinson and Joe Kotynek, led all invertebrate fieldwork. Clayton Bliss, Elsa Jensen, Katrina Lund, and Kelly Lawrence conducted water quality and environmental fieldwork. Scott Miller, Anne Brasher, and Matt Miller were helpful in giving analytical advice and reviewing this manuscript. The Canyonlands Natural History Association, the Rocky Mountain Cooperative Ecosystems Studies Unit (NPS), and the Northern Colorado Plateau Inventory and Monitoring Network (NPS) funded this study.

Literature Cited


Received 9 June 2011

Accepted 17 May 2012