Environmental change explains cichlid adaptive radiation at Lake Malawi over the past 1.2 million years

Sarah J. Ivory, Margaret W. Blome, John W. King, Michael M. McGlue, Julia E. Cole, and Andrew S. Cohen

*Department of Earth, Environmental, and Planetary Sciences, Brown University, Providence, RI 02912; †Department of Earth and Environmental Sciences, University of Kentucky, Lexington, KY 40506; ‡Department of Geosciences, University of Arizona, Tucson, AZ 85721

Edited by John P. Smol, Queen’s University, Kingston, Canada, and accepted by Editorial Board Member David Jablonski August 12, 2016 (received for review July 6, 2016)

Long paleoecological records are critical for understanding evolutionary responses to environmental forcing and unparalleled tools for elucidating the mechanisms that lead to the development of regions of high biodiversity. We use a 1.2-My record from Lake Malawi, a textbook example of biological diversification, to document how climate and tectonics have driven ecosystem and evolutionary dynamics. Before ∼800 ka, Lake Malawi was much shallower than today, with higher frequency but much lower amplitude water-level and oxygenation changes. Since ∼800 ka, the lake has experienced much larger environmental fluctuations, best explained by a punctuated, tectonically driven rise in its outlet location and level. Following the reorganization of the basin, a change in the pacing of hydroclimatic variability associated with the Mid-Pleistocene Transition resulted in hydrologic change dominated by precession rather than the high-latitude teleconnections recorded elsewhere. During this time, extended, deep lake phases have abruptly alternated with times of extreme aridity and ecosystem variability. Repeated crossings of hydroclimatic thresholds within the lake system were critical for establishing the rhythm of diversification, hybridization, and extinction that dominate the modern system. The chronology of these changes closely matches both the timing and pattern of phylogenetic history inferred independently for the lake’s extraordinary array of cichlid fish species, suggesting a direct link between environmental and evolutionary dynamics.

Significance

Tropical African lakes are well-known to house exceptionally biodiverse assemblages of fish and other aquatic fauna, which are thought to be at risk in the future. Although the modern assemblages are well-studied, direct evidence of the origin of this incredible wealth of species and the mechanisms that drive speciation are virtually unknown. We use a long sedimentary record from Lake Malawi to show that over the last 1.2 My both large-scale climatic and tectonic changes resulted in wet-dry transitions that led to extraordinary habitat variability and rapid diversification events. This work allows us to understand the environmental context of aquatic evolution in the most biodiverse tropical lake.

Outcrops provide only intermittent snapshots, and it is difficult to tease apart the influence of climatic versus geomorphic or tectonic processes. Furthermore, to integrate climate history in understanding ecological and evolutionary records, distant sources of climate information, such as marine sediment cores, are of limited value for interpreting in situ processes (6, 7). Thus, long continuous records that preserve ecological, climatic, and sedimentological indicators from within lakes are crucial to disentangling sensitivity thresholds and observing environmental and evolutionary change.

Here we use a suite of paleoecological and mineralogical records from a 380-m drill core (MAL05-1B) from Lake Malawi (Fig. 1; 590-m water depth; 11°18′S, 34°26′E), in southeastern Africa, to reconstruct its ecosystem and limnological history over the last 1.2 My (SI Appendix, Figs. S8–S10 and Table S2). Lake Malawi houses more fish species than any other lake on Earth, including ∼800 cichlid species, and is an iconic “laboratory” for testing questions about rates, modes, and drivers of evolution (8). The lake is deep (706 m), old (>5 My), and permanently stratified with a large watershed (94,000 km²), and preserves an excellent long sedimentological record (9). Furthermore, its location at the southernmost seasonal reach of the Intertropical Convergence Zone (ITCZ) and its effective moisture sensitivity make it a promising archive for reconstructing regional hydroclimate and paleoecology. Previous
paleoecological studies have demonstrated that the lake’s biota has undergone dramatic changes in the past 140 ky, which are strongly correlated with water chemistry, availability of clear-water, rocky, and littoral habitats, and water-column stratification changes (4, 10). Here we extend that record for fossil ostracode crustaceans, fish, silicified green algae, lake flies, charred particles, and various mineralogical indicators over the ∼1.2-My core record.

Results and Discussion

Prior studies of the MAL05-1B core from Lake Malawi lacked age control over large portions of the record (5). In this study, we present data with a refined age model that uses paleomagnetic excursion and intensity records, thus allowing us to attribute timing to events and evaluate the frequency of hydroclimate variability (SI Appendix, Figs. S1–S7 and Table S1).

A principal component (PC) analysis of all hydrologically sensitive fossil and mineralogical indicators shows a first-order change in the frequency and amplitude associated with lake state starting about 800 ka (Fig. 2 and SI Appendix, Figs. S8–S10). Before that time, ostracode assemblages dominated by the freshwater *Sclerocypris jenkinae* and marsh/riverine *Ilyocypris* (mostly *I. alta*) are observed. These ostracodes, in combination with abundant freshwater mollusks, sedge pollen, and frambooidal pyrite (typical of sediment reduction in areas of high biological oxygen demand), are correlated with negative PC1 values. This assemblage indicates direct river influence with through-flowing,
hydrologically open, marsh or shallow lake conditions (11, 12). The intervening wetter/stratified lake phases were relatively brief, with a mean duration of 12.2 ± 3.2 ky for the 10 longest sustained intervals of bottom-water anoxia before the lake state shift at 800 ka.

Following a transitional period from 800 to 700 ka, Lake Malawi switched to a very different type of basin, with the disappearance of local riverine influence, freshwater marshes, and indications of open hydrology. In particular, lowstands were characterized by extended periods of saline/alkaline fauna and mineralogical indicators (e.g., monospecific assemblages of Limnothyra ostracodes and calcium carbonate nodules). These lowstands alternated with much more persistent, very deep, stratified lake conditions when deep-water anoxia prevailed over much longer intervals and benthic fauna was absent. These periods had a mean duration of 24.1 ± 16.0 ky for the 10 longest sustained intervals. Stone et al. (10) previously showed that these deep lake phases were stratified, typified by “blue water,” much like highly transparent lake conditions observed in Lake Malawi today. In contrast, the lowstand lakes were well-mixed and highly eutrophic. The three most prominent and longest deep lake phases conspicuously straddle insolation minima on an ~400-ky beat following the deepening of the basin after 800 ka.

The major patterns of change can be delineated into five paleoenvironmental basin states to better understand the tempo (Fig. 2). This overall pattern of ecosystem and lake-level change is very closely mirrored by the physical properties and sediment characteristics described by Lyons et al. (5) when cast on the same age model (Fig. 2). However, our data conclusively demonstrate that neither dataset can be simply correlated with lake level nor treated as a stationary record for reconstruction of climate (13) throughout the 1.2-My duration of the record, as the pre–800-ka lake was very different from the modern lake and experienced a much smaller dynamic range of water-level variability.

The change after 800 ka cannot be explained by climate variability alone. Lake conditions before 800 ka suggest a near-continuously open hydrologic system, implying that Lake Malawi had an outlet despite persistent shallower depths, even during lowstands. Following the transition that ended at ~700 ka, although the lake was much deeper during wet periods, it was frequently a closed basin with water levels that did not reach the modern outlet height, through the Shire River and into the Zambezi River system. The sill elevation of the Shire outlet precludes an open basin at a much lower elevation; similarly, it would be extremely difficult to maintain an open basin with a lake surface elevation hundreds of meters above the current Shire outlet, as is implied in the quantitative paleolake-level reconstruction shown in Lyons et al. (5). It is much more likely that the state change after 800 ka was linked to a secondary overprint of watershed processes of tectonic origin, potentially outside of the current lake margins. A likely candidate is a significant rise in lake surface base level driven by a shift in the outlet position and subsequent increased threshold elevation of the lake. If an outlet were blocked by progressive uplift, this blockage would cause the lake to deepen during wetter climates until it encountered the next-lowest threshold [e.g., the Shire River outlet at ~477 m asl (meters above sea level)]. The most likely candidate for an earlier outlet is the current inlet and rift-antecedent Ruhuhu River (Fig. 1). Uplift at or near the current drainage divide separating the Ruhuhu River from eastward-directed watersheds flowing to the Indian Ocean would explain the incised meanders and eastward-directed tributaries observed within the modern Ruhuhu watershed (SI Appendix, Figs. S12 and S13).

As a result of our improved age model, our record demonstrates a change in hydroclimate variability after 900 ka, suggesting a similar timing to the Mid-Pleistocene Transition (MPT) (14). This frequency change preceded the basin reconfiguration by ~100 ky. An analysis of evaporative spectrum as well as spectral analysis conducted using the multitaper method on the PC1 hydroclimate reconstruction indicates a change from suborbital hydroclimate variability (~10 ky) to precessional and suborbital frequencies (~19–23 ky, 10 ky) across the MPT (Figs. 2 and 3 and SI Appendix, Fig. S11).

This is in contrast with high-latitude and North African paleoclimate records that show a transition from 41-ky frequencies associated with obliquity to 100-ky frequencies related to eccentricity across the MPT (Fig. 2) (7, 15). In the high latitudes, changes in obliquity result in large summer insolation changes (3 to 9%) and therefore snow melt and ice-sheet growth. However, the obliquity influence reduces toward the tropics and, at the equator, increased obliquity results in a negligible decrease in local summer insolation (~0.4%), likely not sufficient to force hydrological change (16). Furthermore, although there is no indication from the multitaper spectral analysis of a direct 41-ky frequency before the MPT or 100-ky frequency following the MPT as in high-latitude studies, we observe some power associated with combination tones of Milankovitch periods (SI Appendix, Fig. S11). In particular, Quaternary equatorial marine core and deep time records suggest that variability associated with eccentricity or obliquity may be modulated by the strong influence of tropical precession, producing mixed-frequency peaks (17, 18). These frequencies are observed in our record and vary across the MPT, suggesting a modulation by precession of obliquity before the MPT (13.5 ky) and of eccentricity after the MPT (70 ky). This observation implies that teleconnections from high latitudes were overprinted on the strong local insolation control before and after the MPT.

Suborbital frequencies such as those observed at Lake Malawi before the MPT are in theory important climate forcing factors in the tropics (19). Notably, a 40-ky record from Afrotropical Lake Challa illustrates a half-precession frequency, which represents the response of equatorial hydrology to insolation at locations with a double rainy season (20). Although today Lake Malawi at ~15°S is located outside of the modern equatorial zone and rainfall is unimodal, the observance of half-precession suggests that before the MPT, Malawi received rainfall twice yearly. We suggest that the double rainy season as far south as 15°S before the MPT could result from a wider southern-hemisphere Hadley cell. Observations over the last few decades have recorded such Hadley expansion, suggesting that this phenomenon results from a warmer atmosphere (21, 22). The transition at Malawi following the MPT from suborbital to precession/suborbital variability in hydrologic responses is likely the reverse effect. The MPT in marine records corresponds to a shift in mean δ18O, related to a mean increase in Antarctic ice volume across that boundary (23). Lower Antarctic ice volume before the MPT could result in a more southerly extension of the southern-hemisphere Hadley cell, resulting in a double rainy season in a wider area in the southern tropics relative to today, and thus a more symmetrical Hadley circulation about the equator.
In addition to clarifying the roles of watershed tectonics and pre- versus post-MPT climate variability in the African subtropics, our refined chronology allows the dynamics of the Lake Malawi cichlid fish radiation to be understood in a new context. The strong correspondence between lake-level history, basin configuration, and the rhythm of insolation variability, especially the 400-ky eccentricity modulation of the magnitude of precession cycles, has imposed a first-order structure on the Lake Malawi ecosystems over the course of our record. Shallower base level before 800 ka, probably associated with a lower outlet, prevented deep lake formation even at times of high insolation. An easterly directed outlet, perhaps via the Ruhuhu River to either the Rovuma or Rufiji River systems, would explain the strong phylogenetic relationship of the Malawi cichlids with rivers in this region as opposed to the more southerly connections that exist today (24, 25). After 800 ka with the formation of a deeper topographic basin, insolation thresholds became much more important in dictating lake persistence.

We hypothesize that thresholds exist for lake persistence related to insolation magnitude that can also be applied to other East African lake systems based on basin morphometry. Lakes with high surface area-to-volume ratios may be unable to persist when eccentricity is low, and precessional maxima are therefore lower than during high-eccentricity intervals. In other words, there may be a threshold effect imposed by the long-term (400-ky) eccentricity cycles that effectively make lakes insensitive to precessional forcing at the lowest amplitude intervals. Lake basins in this state would remain either low or dry during these periods, with deep lakes only developing during extreme precession maxima, as is the case in the Eastern Rift Valley today (26). However, in steep-sided rifts with low surface area-to-volume such as Lake Malawi, basin geometry requires 70% of the water volume to be evaporated before significant lake-level regression occurs (5). In this case, the lake remains higher unless a much larger change in forcing drives lake-level changes. Thus, the threshold is only crossed during eccentricity maxima when summer insolation reaches sufficiently low values to cause very severe precipitation reductions. Such lakes only dry out during extreme insolation lows, resulting in the Lake Malawi megadroughts of the early Late Pleistocene (27). This would result in an insensitivity to summer...
Insolation variability during eccentricity minima; however, in contrast, here lakes remain deep during these times, as is the case for the past ∼70 ky, and at ∼400-ky intervals before today. The onset of ∼400-ky cycles at 800 ka generated alternating intervals in Lake Malawi with vastly different lake characteristics (Fig. 4). During times of deep lake persistence, blue lake phases at Malawi were fresh, and strongly stratified with low surface water nutrients and phytoplankton and extensive littoral rocky habitat (4, 10). In contrast, green lake phases were characterized by well-mixed, turbid, saline, shallow lakes, with little submerged rocky substrate (5).

We hypothesize that maintaining blue lake conditions for long periods of time (on the order of >10^6 y) was critical to allowing rocky habitat-based ecosystems to develop and diversification based on assortative mating to evolve. Molecular genetic data for Lake Malawi cichlids are consistent with this interpretation. We propose that this green-blue lake phasing generated diversification and extinction cycles (28–30) as a first-order control on diversification and extinction within littoral habitats, and is consistent with hypothesized phases of diversification (8) and timing from molecular genetic data (24). An initial burst of diversification around 749 ka (31) corresponds to the earliest evidence of sustained blue lake conditions in Lake Malawi over the last ∼1.2 My, which would have created the rock–sand habitat differentiation thought to have generated initial within-lake adaptations (Fig. 4). Before that, lake-level and other environmental fluctuations appear to have been too frequent to allow diversification, and rocky littoral habitats were uncommon. A subsequent blue lake phase at ∼400 ka may mark the timing of the evolution of distinct trophic morphologies, also consistent with the branch topology of mtDNA trees (24). Finally, most populations of rock-dwelling (mbuna) cichlids expanded during the last 75,000 y, the most recent blue lake phase of the lake (32), consistent with the third phase of male color diversification of Kocher (8). That study also found support for dispersal within regions with genetic diversity accumulating locally following lake-level rise, implying that assortative mating evolved on similar timescales. Conversely, based on comparison with modern lakes, especially those undergoing rapid eutrophication (33, 34), the transition to green lake phases would have generated coordinated episodes of extinction, hybridization, and reduced biodiversity. Haplotype sharing and hybridization were probably the norm during these phases, when assortative mating using visual cues would have been suppressed by more turbid waters (33, 34). Molecular clocks indicate the onset of mbuna-dominated diversification within the time frame of the penultimate blue lake phase of Malawi (35), consistent with our model outlined in Fig. 4.

Diversification could have occurred during predominant green phases, but it is probably manifest in Malawian cichlid phylogeny in very different ways. Bursts of diversification almost certainly occurred during brief blue lake interludes in what are otherwise green lake cycles. However, these were probably aborted as the lake quickly reverted to its green lake state. More importantly, introgressive hybridization among lineages in which assortative mating breaks down under turbid green lake conditions may have been an important source of evolutionary novelty. For example, molecular clock data suggest that during an early green lake phase, some deep benthic cichlid clades evolved through mbuna hybridization (36). These taxa today are specialized for light-feeding, similar to what would have been the shallow water norm during a green lake interval. Other groups that underwent minor radiations during green lake phases are open-water, pelagic predators such as Diplotaenidodon, which today show little genetic differentiation around the lake and would not be subject to the most severe effects of green lake habitat loss experienced by the mbuna cichlids (25). More generally, green lake phases may be critical for “priming the pump” of diversification by creating a wide variety of novel morphologic/behavioral phenotypes during hybridization episodes (37) that could be exploited during subsequent blue lake radiations. Our causative model would also help explain why Lake Tanganyika does not appear to show the three-phase evolutionary history proposed by Kocher (8), as that lake, because of its much greater depth, probably persisted as a blue lake with rocky littoral habitats through the 400-ky cycles that so strongly affected Lake Malawi (38).

**Methods**

**Age Model.** Our age model uses all ^14^C, Ar/Ar, and paleomagnetic reversal and excursion data previously published in Lyons et al. (5). Additionally, a series of new age constraints based on paleomagnetic excursion and intensity data has been added (SI Appendix, Figs. S2–S7 and Table S1). The age model itself was created by fitting the data with a monotonically polynomial spline (SI Appendix, Fig. S1).

**Paleoecology.** Sediment samples (n = 1,945) were taken at 16-cm intervals (average intersample resolution ∼570 y). Samples were split into two aliquots; the first was weighed and then disaggregated using deionized water and freeze–thaw treatment (often requiring multiple cycles), and finally washed through 63-μm mesh stainless steel sieves to recover all material sand-sized and greater. The second aliquot was weighed wet, oven-dried at 60 °C for at least 5 d, and reweighed after cooling. This calculated water content was then used to determine the dry weight of the sieved aliquot. Wet-sieved residues were counted for the total number of ostracodes per sample (normalized to abundance per g dry weight) and for taphonomic condition (% broken, whole carapace, adult, carbonate-coated, and oxidation/reduction-stained), and each valve was identified to the species level where possible. Additionally, the wet-sieved residues were counted for absolute abundance of other common fossil remains in this same size range (chaoborid fragments, charred particles, fish (including bones, teeth, and scales), algae (Pediastrum and Botryococcus), and Polypodiaceae sori) and the relative abundance of mollusk fragments and both monomineralic and lithic grains in the sand fraction (specifically quartz, mica, pyrite, siderite, vivianite, aragonite needles, and ooids). Because of the variety of data types in the dataset, we used a principal component analysis to reduce dimensionality. Indicators included in this analysis were the relative abundance of the key ostracode taxa Cypriodaphnia, Limnothamnium, Candonopsis, ilicypris, and Stephanodiscus. The taphonomic condition (broken, whole carapace, adult, carbonate-coated, and log-transformed concentrations of chaoborid fragments, charred particles, fish (including bones, teeth, and scales), algae (Pediastrum and Botryococcus), and Polypodiaceae sori, and the relative abundance of mollusk fragments and authigenic and terrigenous mineral grains in the sand fraction (quartz, mica, pyrite, siderite, and ooids). Spectral Analysis and Evolutionary Spectrum. For both the multitaper evolutionary spectral analysis conducted on PC1 and presented in the main text of the paper and the spectral analysis in SI Appendix, Fig. S11, multitaper mean spectral analysis with the AR1 (autoregressive) noise spectrum was conducted to look in detail at the relative influence of different frequencies across the Mid-Pleistocene Transition. For this analysis, PC1 was smoothed (five points), divided into pre- and post-MPT time series (0–700 and 800–1,200 ka), and run separately. Data pretreatment to a uniform time step involved interpolating to a 250-y resolution and then resampling to 1,000 y, which is closer to the average resolution before pretreatment.

**Acknowledgments.** We thank all participants involved in the Lake Malawi Drilling Project, in particular Chris Scholz, Tom Johnson, Robert Lyons, and Eric Brown. We also thank Walter Salzburger for valuable discussion of cichlid evolution and Sylvia Dee for spectral analysis code. Initial core processing and sampling were conducted at LacCore at the University of Minnesota. Sample processing was conducted by Devin Gaugler, Chris Johnson, Jeanine Ash, and Claire DeCelles. Funding came from the US National Science Foundation–Earth System History Program (EAR-0602350); International Continental Scientific Drilling Program, American Chemical Society–PRF (54376-DN18); and Smithsonian Institution. Partial sample preparation and analysis support was received from DoseEC, Chevrons and BP summer scholarships, and UA SAGUARD.


