Appendix S2. Supplementary methods and tables.

Plant- and plot-level nutrient predictions

We developed our plant nutrient database with calibrations derived from partial least squares regression (PLS) that relate near infrared spectra (NIRS) to C, N, P, and K. We did not have an a priori expectation whether NIRS PLS calibration for these nutrients would perform better with a global model using all data or if local models fit for data subsets would be preferable. We therefore constructed calibration models for individual continents, individual plant functional types, and for the entire dataset. Splitting by continents did not produce models with greater explanatory power for any of our calibrations. However, removing Bryophytes (a restricted set of samples) did improve our models, as did splitting litter and live biomass samples into separate calibrations (see Appendix S2: Table S1, below).

To optimize each individual calibration model, we considered different spectral regions and spectral preprocessing steps. Following the implementation in the R package ‘leaf.spec’ we considered all combination of five common spectral regions. We also considered seven preprocessing steps, including the first and second derivatives (D1f and D2f), vector normalization (SNV), constant offset elimination (COE), multiplicative scattering (MSC), straight line subtraction (SLS), and min-max normalization (MMN) as well as D1f combined with either SLS, SNV, or MSC (Appendix S2: Table S1, below). Models were preferred if they had low Mean Squared Error of Prediction (RMSEP). Prior to analysis, we also inspected the scans for outliers and removed one deviant scan.

Model validation is essential to producing reliable PLS calibrations. We compared four methods for selecting a representative validation dataset. These methods were the duplex/Kennard-Stone (KS) algorithm (Kennard & Stone, 1969; Snee, 1977), a KS with a PCA of the NIR scans, a modified KS algorithm that incorporates variation in both the spectra and the Si values (“SPXY”; Snee, 1977; Saptoro et al., 2012), a further modification that uses the Mahalanobis distance (MD) instead of Euclidean distances. Different selection methods work best for each element; interestingly, SPXY worked best for C and P, whereas MDKS work best for N and KS work best for K. The C, N, P, and K models validated well against their test sets and N performed particularly well (Appendix S2: Fig. S1, below). Therefore, the final models were used to predict the entire dataset based on their NIR scans. Predicted values were flagged as outliers if either their scans were outside the maximum MD of the calibration set or if the predicted value itself was greater than three time the standard deviation. To form the final nutrient dataset, the wet-lab chemistry data from the calibration samples were preferred over the predicted values.

Imputation of Missing Soil Data

Three sites, kibber.in, marc.ar and serengti.tz (Appendix S1: Table S1), were unable to provide post-treatment soil samples. Consequently, we used the command ‘mice’ in the R-package mice (van Buuren and Groothuis-Oudshoorn 2011) to impute the missing soil N data using Bayesian regression. To accomplish this, we used the strong relationship between mean annual precipitation (MAP) and Soil %N and assumed a normal distribution (method = ‘norm’).
produce a single imputed dataset (m=1) of Soil %N for each plot within each of these sites. Prior to imputation the Soil %N ~ MAP regression coefficient relationship was: mean ± SE = 0.386*10^{-4} ± 3.12*10^{-05} (t_{139} = 12.38, R^2 = 0.52, P < 0.001), whereas after data imputation it was: mean ± SE = 0.58*10^{-4} ± 4.82*10^{-05} (t_{179} = 12.0, R^2 = 0.45, P < 0.001)

Grazing index used in the SEM

In addition to the binomial variable which accounted for the fence treatment, we developed a grazing index which was intended to account for variation in the diversity and abundance of herbivores across sites. For each site, we asked the PI to list the species of herbivores > 2 kg (approximately rabbit-sized or larger) that regularly consume aboveground biomass at their site and assign an importance value, from 1 (very low impact and frequency) to 5 (high impact high frequency). While imperfect, we judged that an importance value will provide expert opinion across sites where other metrics (large herbivore units, etc.) are likely to be biased due to ecological or evolutionary differences across sites. The grazing index for each site was calculated as the sum of the importance values for all common herbivores > 2 kg in body mass and yielded a discrete variable that ranged from 0 (sites with no significant herbivores) to 27 (Serengeti National Park).

Hierarchical structural equation modeling with derived variables

We assessed the system-level direct and indirect effects of environmental variation, eutrophication and herbivory on plant nutrient content using structural equation modeling (SEM). SEM is conceptually related to path analysis but with more advanced means of matching data to theory and evaluating model fit (e.g., Grace 2006). We started with an a priori conceptual model, or meta-model (Appendix S3: Fig. S1; sensu Grace et al. 2010), in which the putative drivers of plant nutrients were hypothesized to affect total standing nutrient content via effects mediated by the amount of grass biomass present (%grass), the total amount of biomass present in each plot (plot biomass) and the nutrient content of key plant nutrients, which in our case was the weighted sum of N + P + K in the aboveground phytomass (plant NPK) of each plant functional type. The total content of N, P and K (total standing NPK; g m^{-2}) in each plot was considered the final response in our model. As the total standing NPK in each plot is the mathematical product of plant NPK and total plot biomass, the path coefficients pointing to this variable were computed rather than estimated (Grace unpublished). Due to the hierarchical nature of our model, each response variable (% grass, plot biomass and plant NPK) was analyzed in a piece-wise fashion and separately for those predictors which were measured at the plot-level and those which existed only at the site-level. Models were analyzed in a piecewise fashion in R using a hierarchical approach (e.g., page 270, Gelman and Hill 2007).

We began by analyzing the plot-level responses using the ‘lmer’ command in the R-package lme4 with NPK, FENCE and soil % N as predictors and sites as random effects (1|SITE). We explored model which allowed interaction effects between the two treatment effects, NPK and FENCE, and those between treatments and both soil % N and plot % grass. We then isolated the random conditional modes for each site using the ‘ranef’ command in R from the lme4 package.
and used them as responses in a subsequent analysis which included site-level predictors (climate and the grazing index). For each response variable and model analyzed at the plot- or site-level we began with a full and a suite of simplified candidate model from which we used forward and backward stepwise selection procedure based on model AIC values (using the ‘stepAIC’ and commend and the option ‘direction = “both”’ in R).

In the case of an interactive effect between variables, such as between NPK fertilization and plot soil %N, for example, we incorporated the combined effect of the component effects into the model using the composite variable approach of Grace and Bollen (2008). This method creates a new ‘composite’ predictor that represents the combined effect of a set of coefficients that make up a nonlinear effect on a particular response variable. We then determined the strength of the standardized path coefficient for each composite variable following the methods of Grace and Bollen (2008).

References
R Development Core Team (2012) *R: A Language and Environment for Statistical Computing*, R Development Core Team.
Table S1. NIR calibration model results for components C, N, P, and K. R^2 of Validation is the coefficient of determination for the calibration model applied to the representative data subset. Rank is the number of latent vectors used by the PLS regression. Optimal spectral ranges and preprocessing for each model are also provided.

<table>
<thead>
<tr>
<th>Material</th>
<th>Component</th>
<th>R^2 of Validation</th>
<th>Rank</th>
<th>RMSEP</th>
<th>Spectral range (cm-1)</th>
<th>Preprocessing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Live C</td>
<td>0.81</td>
<td>10</td>
<td>1.63</td>
<td>9400 to 4600</td>
<td>D1f and SNV</td>
<td></td>
</tr>
<tr>
<td>Live N</td>
<td>0.95</td>
<td>8</td>
<td>0.17</td>
<td>9400 to 6100 and 5450 to 4250</td>
<td>D1f and MSC</td>
<td></td>
</tr>
<tr>
<td>Live P</td>
<td>0.77</td>
<td>10</td>
<td>0.09</td>
<td>9400 to 4600</td>
<td>D1f</td>
<td></td>
</tr>
<tr>
<td>Live K</td>
<td>0.9</td>
<td>8</td>
<td>0.48</td>
<td>7500 to 5450</td>
<td>D1f and MSC</td>
<td></td>
</tr>
<tr>
<td>Litter C</td>
<td>0.8</td>
<td>8</td>
<td>2.09</td>
<td>7500 to 6100</td>
<td>D2f</td>
<td></td>
</tr>
<tr>
<td>Litter N</td>
<td>0.95</td>
<td>8</td>
<td>0.08</td>
<td>7500 to 6100 and 4600 to 4250</td>
<td>D1f and MSC</td>
<td></td>
</tr>
<tr>
<td>Litter P</td>
<td>0.86</td>
<td>10</td>
<td>0.03</td>
<td>6100 to 4600</td>
<td>D2f</td>
<td></td>
</tr>
<tr>
<td>Litter K</td>
<td>0.92</td>
<td>3</td>
<td>0.08</td>
<td>9400 to 6100</td>
<td>D1f and SNV</td>
<td></td>
</tr>
</tbody>
</table>

**Figure S1.** Validation diagnostic plots for NIR C, N, P, and K calibrations showing model predictions plotted against measured wet-lab chemistry.