Influences on plant nutritional variation and their potential effects on hominin diet selection

Amanda G. Henry [1,2]

Antje Hutschenreuther [2]

Oliver C. C. Paine [3]

Jennifer Leichleiter [4]

Daryl Codron [5,6]

Jacqui Codron [5]

James Loudon [7]

Stephanie Adolph [8]

Matt Sponheimer [3]

1 Faculty of Archaeology, Leiden University, Einsteinweg 2, 2333CC Leiden, The Netherlands

2 Plant Foods in Hominin Dietary Ecology Research Group, Max Planck Institute for Evolutionary Anthropology, Deutscher Platz 6, 04103 Leipzig, Germany

3 Department of Anthropology, University of Colorado Boulder, Boulder, CO 80309, USA

4 Institut für Geowissenschaften, AG für Angewandte und Analytische Paläontologie, Johannes Gutenberg–Universität Mainz, 55128 Mainz, Germany

5 Florisbad Quaternary Research Department, National Museum, PO Box 266, Bloemfontein, 9300, South Africa

6 Centre for Environmental Management, University of the Free State, PO Box 339, Bloemfontein, 9300, South Africa

7 Department of Anthropology, East Carolina University, 231 Flanagan Building, East Fifth Street, Greenville, NC 27858-4353 USA

8 Institut für Tierernährung, Ernährungsschäden und Diätetik, Veterinärmedizinische Fakultät, Universität Leipzig, An den Tierkliniken 9, 04103 Leipzig, Germany.

Keywords:
protein; fiber; antifeedants; phytoliths; South Africa; Cradle of Humankind

Abstract:

The selection of foods in any environment depends on a variety of factors, including the nutrient availability and antifeedant loads in the component habitats. How these nutritional properties vary and covary in time and space is not well known, particularly among wild plant species. We collected plant samples from several habitats within the Cradle of Humankind World Heritage Site in South Africa, and measured their macronutrient and antifeedant properties in order to explore how season, habitat, plant type, and plant organ affected the quality of these potential plant foods. Our results have implications for early hominin use of similar habitats.

Introduction

Every bite an animal makes represents a decision to consume a particular food. Over time this series of decisions can have life or death consequences for the consumer. A large body of research has developed to ask why certain plant foods are chosen over others, and how this may relate to the habitats

in which the consumer lives (e.g., Hughes, 1990; Johnson et al., 2006; Manly et al., 2007). The foods that each consumer chooses, and the habitats from which they select these foods, are determined in part by the animal's intrinsic digestive capacity, in part by the selectivity of the individual animal, and in part by the kinds of plants that are available at each particular point in time in each habitat (Duncan and Gordon, 1999). It is generally accepted that mammals select their foods due to the nutritional qualities of these foods (Lindroth, 1989). However, these nutritional qualities can vary significantly both among habitats and between seasons, and it is likely that mammals must be flexible in their food choices in response to this variation (Lindroth, 1989).

The human fossil record indicates that early hominin species made a variety of different dietary choices. Several key sites (e.g., Sterkfontein, Swartkrans, Drimolen, Kromdraai, Gladysvale, Coopers, Malapa, Rising Star) in the Cradle of Humankind World Heritage Site of South Africa provide an abundant fossil record of some of the earliest hominins, including *Australopithecus africanus*, *Paranthropus robustus*, *Australopithecus sediba* and early *Homo*. Some hominin taxa, such as *P. robustus*, show high within-individual variation in stable carbon isotope ratios, suggesting seasonal or interannual dietary changes (Sponheimer et al., 2006). However, the same taxon has much lower between-individual variation than, for example, *Au. africanus*, in which the the stable carbon isotope compositions indicate high levels of inter-individual dietary variation (Ungar and Sponheimer, 2011; van der Merwe et al., 2003; Ungar and Sponheimer, 2011). Some of these hominin species preserve isotope, microwear, and phytolith evidence suggesting they consumed foods predominantly from forested areas, despite environmental markers indicating they lived in open landscapes (Henry et al., 2012). Clearly, these hominins were choosing different foods and different habitats in which to feed. This variety of behaviors is likely due in part to the kinds of nutrients available across the landscape.

Reconstructions of the environments surrounding the cave sites in the Cradle emphasize the patchy and mosaic nature of the paleolandscape. In their recent paper, Peterson and colleagues (2018)

summarize the paleoecological data from several of the major fossil-bearing sites. At Makapansgat, analyses of the habitat preferences of bovids and other large mammals (Reed, 1997; Vrba, 1980), small mammal isotopes and microwear (Hopley et al., 2006), and preserved pollen (Cadman and Rayner, 1989) all indicate a closed, wooded environment, though grasslands, edaphic grasslands and bushlands were also present. Sterkfontein was slightly more open, with the large mammal communities (Reed, 1997; Vrba, 1980) and isotopic compositions (Lee-Thorp et al., 2007) suggesting medium density woodland in the immediate vicinity of the cave and grassier and bushier habitats in the broader region. Nevertheless, dense temperate woodland is represented by pollen and fossil wood (Bamford, 1999). Swartkrans was even more open, with dry grasslands and edaphic grasslands represented by the large mammals (Reed, 1997; Vrba, 1980), though the micromammals suggested the presence of riverine woodland as well (Avery, 2001; Denys, 1992). Interestingly, the hominin fossils from the site seem to be preferentially associated with closed habitat fauna (de Ruiter et al., 2008). Clearly the hominins had access to a wide range of habitats, likely in the form of ecotones from grassland to forest and edaphic grasslands along river margins (Reed, 1997). It is important to note, however, that reconstructions of the paleolandscape based on material from the cave sites is inherently biased toward dryer periods, during which the caves were open and bones could accumulate (Pickering, 2018). This bias may also overemphasize the importance of open habitats, but would not negate the overall mixed nature of this region.

Broader climate reconstructions based on both terrestrial and marine sedimentary records indicate alternations of wetter and drier periods across Africa in the Plio-Pleistocene, with evidence for increased variability and aridity during several distinct periods. The first two of these, one from c. 2.9 – 2.4 Ma and the other peaking at 1.8 Ma are directly relevant to the fossil material preserved in the Cradle of Humankind (deMenocal, 2004; Pickering and Kramers, 2010). These dryer periods would have meant more open environments, and more deposition in cave sites.

Some have questioned whether the grassland component of these habitats would have been as prevalent as they are today, given that human activity including fire management, grazing, and tree-clearing may contribute to the relatively treeless aspect of the current-day landscape (Mucina and Rutherford, 2006). A variety of data suggests that these grassland biomes are in fact ancient, likely due to naturally occurring fires (Bond et al., 2003), climate regimes (particularly rainfall under <1000mm) (Staver et al., 2011), and/or soil nutrient availability (Milewski and Mills, 2015). Even if wildfires were not as prevalent in the past, the cold climate and low rainfall that characterizes this region is likely sufficient for the presence of open grassy biomes, as suggested by the fossil grassland-adapted fauna. Grasslands may have been less dominant than today, but it is clear that hominins in this region had access to a wide variety of habitat types.

To better understand how and why hominins may have chosen the foods they did from this diversity of habitats, we need to be able to identify how the inherent nutrients of their food plants varied across the landscapes. Most studies focus on seasonal changes within a habitat, but with a few notable exceptions (e.g., Codron et al., 2006; Sept, 2001) there has been little focus on nutrient variation across habitats. These landscape-level variations may have been a strong influence of hominin food choices.

Several factors are known to influence plant nutritional values. Plants vary in protein, fiber and other content between seasons, and as they age (Lindroth et al., 1986). The part of the plant also determines its quality as a food, with tree leaves generally being a better resource than the stems of those plants. Broad categories of plants, such as grasses, trees, and forbs also predictably differ in terms of their nutritional properties, to the point foraging models incorporating such information in an

attempt to explain herbivore food choices often match well to observed diets in the field (Belovsky, 1990; Owen-Smith and Cooper, 1987; 1989). Given that habitats are made up of distinct communities of plants with different proportions of each plant type, as well as different primary species, the nutrition available to a consumer will vary from habitat to habitat (as noted in Stoner, 1996). And finally, while the intrinsic properties of each plant taxon likely differ from one to the other, it is possible that some taxa or types of plants are more variable in their nutritional properties than others either within or between habitats.

There has been little consideration of the pattern of covariation of nutritional properties in which all of these potential influences are simultaneously considered (e.g. habitat, season, plant part and plant type), and the relative importance of each influence compared. Sept's (2001) study is a notable exception. Using the appearance and frequency of certain plants across modern East African landscapes, she modeled where a variety of hominins (with different dietary and behavioral constraints) could have met their nutritional needs. While a pioneering study, it was limited by a lack of direct nutritional and antifeedant properties of the individual plants in each habitat, and it did not consider the within-taxon variability in nutritional qualities over time and space. Only by examining all of these factors at the same time can we have a meaningful comprehension of how the nutritional landscape may vary, and which of the influences should be most carefully considered. Such an understanding is vital for making informed predictions about where a species relying primarily on plant foods, such as a hominin, would have foraged on a landscape. By collecting information about the nutritional properties of plants among habitats within a constrained landscape, we can both better understand the factors that influence the variation among these specimens, and also ask several important questions about the variation among these plants:

Is the nutritional variation among habitats greater than that between seasons?

Both habitat variation and seasonal variation are known to influence an animal's feeding behavior (Romanowski and Żmihorski, 2009). However, the relative importance of these two factors might determine, for example, if an animal chooses to migrate seasonally across a landscape or if they stay within the same general area and make use of a variety of microhabitats within the same constrained environment (Alerstam et al., 2003). These data may help us understand the within-individual variation in isotopes seen among *P. robustus* - were these hominins staying in place and tracking local seasonal changes, or did they travel among habitats?

Which habitats are most similar to each other? Which are most different? What drives this variation?

An animal's patch choice depends on whether their minimum nutritional requirements are met. If sufficient nutrients cannot be obtained in one patch the animal must make a decision to leave and forage elsewhere. In many cases, animals choose a mixture of foods from different habitats in order to meet nutrient requirements. For example, moose must consume aquatic plants in order to meet minimum sodium requirements, but must also forage on forbs and tree leaves to meet protein and energy requirements (Belovsky, 1978). This leads to the moose dividing its time between beaver ponds and woodlands. Hominins likewise had to meet nutritional requirements, but if a variety of habitats all provide the required nutrition, then this might lead to the high between-individual variation such as seen in *Au. africanus*.

Which habitat has the most calories? Which has the most protein?

Recent studies have indicated that primate species will focus on a narrow ratio of protein to non-protein energy (NPE) (e.g., Takahashi, 2018), and that the ideal ratio broadly depends on the dietary category (i.e., frugivore, folivore, etc.) to which that species belongs. However, in periods of dietary constraint such as a seasonal change in food availability, some species, including humans and

some frugivorous primates, will prioritize meeting protein requirements and will even consume excess NPE to meet the protein goal (Felton et al., 2009; Righini et al., 2014; Takahashi, 2018). Other species, such as gorillas, prioritize NPE during periods when fewer fruits are available, and will over-consume protein (Rothman et al., 2011). Clearly, the availability of protein energy and NPE can drive how animals use their landscape. By exploring how these limited nutrients vary among habitats we can make predictions about how different kinds of herbivores may have used the landscape.

Where is the antifeedant load the least?

Mammalian herbivores actively select against plant physical and chemical defenses, such as phytoliths and phenols (Bryant and Kuropat, 1980). Plants which are heavily browsed or grazed also increase the production of these antifeedants (Karban and Myers, 1989; Massey et al., 2007). We can look across a landscape to see which habitats, plant types, and plant parts are least protected, and if this protection correlates with nutritional properties in those plants. While less-protected plant parts may seem like more likely targets for consumers, it is also possible that plants put less energy into protecting lower nutrient parts, making them less desirable foods.

Where is there the most bio-available nutrition? Are certain plants better choices than others?

Given the preference for protein and calories and the aversion to antifeedants among most mammalian herbivores, those habitats, plant types, plant parts, and seasons in which the sought-after nutrients are least protected might have been key resources across this landscape. We can explore whether preferred nutritional qualities co-vary with antifeedants, and how this co-variation could influence which plant types, plant parts, or habitats in which hominins foraged. Several authors have suggested that wetland and near-water habitats would have been preferred, as they potentially contained many nutrient-rich foods (Reed, 1997; Wrangham et al., 2009). Among food types, the underground storage organs of biennial or perennial plants have been proposed to be key resources

(Dominy et al., 2008; Macho, 2014). We therefore explored whether wetlands were the most nutritionally valuable, and how the roots of USO-producing plants compared to other plant parts.

Which plants and habitats are the most reliable? Which are the least?

When faced with the potential for high antifeedant loads or low nutritional qualities, mammalian herbivores are challenged to determine what will constitute a good food source. Many of these factors are not apparent upon visual inspection of the plant item, and the herbivore must have other means of predicting and/or assessing food quality (Illius and Gordon, 1990). The predictability of habitats, that is, an animal's ability to know beforehand whether a habitat will have food at any given time, has been implicated in the evolutionary development of migration among birds (Alerstam et al., 2003). It has also been suggested as a reason why different troops of Howler monkeys choose different foods - given low predictability of antifeedants (Glander, 1982), these primates consume only those individual plants which they know are safe (Stoner, 1996). Within the Cradle, we can examine the within-taxon and within-habitat variability to see if certain plants or habitats were more reliable or predictable sources of nutrients.

Finally, a combined analysis of the variation within plant nutritional properties across a landscape may reveal unexpected patterns about the potential importance of certain categories of plants or particular habitats.

Methods:

In order to explore how plant nutritional qualities and antifeedant properties varied across a landscape relevant to human evolution, we collected plants from five distinct habitats within the Cradle Game Reserve. Permission to perform this research and collect samples was given by The Gauteng

Provincial Government Department of Agriculture and Rural Development Directorate of Conservation Permits Office (permit to collect and convey protected plants, number 0204; and permit for the exportation of a protected plant, number 1582). This reserve lies within the Cradle of Humankind World Heritage Site, roughly 50 km northwest of Johannesburg, South Africa. While the areas within the Cradle Game Reserve likely do not directly match the exact habitats in which early hominins lived, they provide an analogous range of habitats, including open grassland, open woodland, and both river and wetland habitats. The predominant local vegetation type is the Rocky Highveld Grassland (Low and Rebelo, 1996), which is a fire-maintained grassland with patches of woodier vegetation. Open grasslands are found generally on exposed hills and ridges, while woodlands are generally found in valleys or on somewhat protected hillsides. The summer (October to March) is hot and rainy (650-700 mm), while the winter (April to September) is usually cold and dry (close to 0 mm), often dipping below freezing, and subject to frequent wildfires.

Previous work in the game reserve by our group (Codron et al., 2015) has identified several habitats within the Game Park, which differ from each other in terms of water availability, substrate, and predominant plant taxa. We chose five of these areas, each with differing characteristics. These included one open grassland (Bloedveld or BV), one riverine habitat (Tick River or TR), one marshy habitat (Peter's Vlei or PV), and two wooded habitats, one more open (Dolomite Open Woodland or DOW) and one more dense (Kudu Hill or KH) (Figure 1).

Within each habitat, we used Modified-Whittaker sampling plots measuring 20x50m in order to establish the abundance of the major plant taxa (Stohlgren et al., 1995). Within each sampling plot, we then chose four sampling sites, hereafter called 'replicates', at roughly the corners of a 20x20m square in the middle of the sampling plot. At each replicate site, we collected material from the two most common trees and the two most common grasses in that sampling plot, and from any forbs and sedges

that were found within the sampling plot. In many cases, this involved moving away from the exact corner of the square to the closest representative of that species. We chose only the two most common grasses and trees for several reasons. First, these taxa represented the majority of the biomass available to consumers from each habitat, and were therefore more relevant for describing the nutritional profile of each habitat. Second, less-common taxa were sufficiently rare that it would have been impossible to acquire four replicates. Finally, we were limited in our ability to process larger number of specimens so had to be selective in the plants we sampled.

From each taxon, we collected all of the seasonally-available plant parts, aiming for representative samples of leaves, stems, fruits, flowers, and roots. When sampling trees it was usually possible to collect enough material from a single individual; however for grasses and herbaceous species it was occasionally necessary to collect material from several individual plants at the replicate site in order to have enough material for sampling. After collecting the major taxa, we then looked for species that showed signs of having actually been consumed. We collected samples for nutritional analysis in one wet season (January 2013) and one dry season (July 2013).

Each plant part was manually separated from the whole plant. Fruits were processed entire (e.g. not separated into seed and flesh), as the rare fruits found in this landscape are small (< 3cm) and fleshy, and are often consumed whole by baboons (Peters, 1993). Each sample (one plant part of one replicate) was weighed, and then placed into a food dehydrator (Sedona food dehydrator SD-9000 series). he sample was allowed to dry at 35 ºC until the weight no longer changed. Samples were brought back to Germany where the nutrients were measured.

We measured dry matter, crude ash, crude protein, crude fiber, neutral detergent fiber (NDF), acid detergent fiber (ADF), and lignin at laboratories of the Institute of Animal Nutrition, Nutritional

Diseases and Dietetics of the Leipzig University Faculty of Veterinary Medicine. These nutritional qualities were measured according to standard methods, which are described in Appendix A. Phytoliths, tannins, polyphenols and lipids were measured in the laboratories of the Plant Foods in Hominin Dietary Ecology research group at the Max Planck Institute for Evolutionary Anthropology. We had to adapt several methodologies for analyzing these nutritional qualities and antifeedants, which are also described in Appendix A. Non-structural carbohydrates (NSC) were calculated by subtraction of lignin, crude protein, total ash and NDF from total dry matter. NDF is the only fiber fraction containing hemicelluloses in addition to cellulose and lignin. Thus, NDF represents best index of energy available from total insoluble fiber.

From these values, we then calculated gross (GE) and metabolizable energy (ME) values for each sample. Gross energy represents the intrinsic caloric value of the food, equivalent to the values one would achieve by bomb calorimetry. We did not do direct calorimetry and instead calculated GE based on Kamphues (2014) with modifications as suggested by Pagan (1999). These publications suggest the following energetic values per component: 17.5 kJ/gram NSC, 23.9 kJ/gram protein, 39.8 kJ / gram lipid, and 20.1 kJ/gram NDF. Metabolizable energy is the amount of energy that is available to the consumer when the constraints of digestion are accounted for, and thus is a slightly more relevant measure of the energy value of the plant. We estimated ME by using standard conversion factors (Atwater factors) for NSC, protein and lipids (16.7 kJ/gram, 16.7 kJ/gram and 37.6 kJ/gram respectively), and adding an additional 6.7 kJ for each gram of NDF. This latter value was calculated by Conklin-Brittain and colleagues (2006) as the energy available to a chimpanzee from the digestion of fiber. We felt this value, rather than others developed for ruminants for example, was appropriate for a study exploring the nutrition available to early hominins. However, in the same study Conklin-Brittain and colleagues do note that foods with higher lignin values will have lower digestibility of NDF fraction, suggesting that our metabolizable energy values will be slightly too high for the foods

with a greater percentage of lignin (e.g. tree leaves). Furthermore, we recognize that the Atwater factors may indicate a slightly too high value for proteins from plants (Schakel et al., 1997). We have used the standard values, however, in order to avoid errors in estimating the correct values and to make our results more directly comparable to previous work (e.g. Conklin-Brittain et al., 2006).

A large number of plant specimens were collected, only about half were nutritionally analyzed. Some plant parts were avoided, because we judged them to be unlikely food sources for most mammalian feeders (e.g. grass roots, tree stems). Some species did not provide enough material for nutritional analysis. Finally, we focused on taxa that were found in multiple habitats in order to better explore within-species, between-habitat variation. In total, we performed nutritional analyses on 277 samples, from 33 taxa. Our raw data are available on Figshare (10.6084/m9.figshare.6825596).

We were interested in exploring how the intrinsic properties of the plants available on the landscape would affect their nutritional properties. As explained in the introduction above, several studies have suggested that these nutritional properties vary within taxa among habitats, among different parts of the plants, and between seasons. Furthermore, broad categories of plants (e.g. forbs vs. grasses) have been shown to be relevant determinants of diets, rather than individual species (Belovsky, 1984, 1999). We chose to test how these main potential drivers of plant nutritional properties (habitat, season, plant part, and plant type) might variously affect the plants in our study. We measured and calculated more than ten nutritional properties, however, making a simple relationship difficult to test. We therefore first performed a factor analysis on all of the combined nutritional data to reduce these ten variables into a smaller number of factors which explain some level of co-variation among our nutritional data.

Details of the Statistical analysis:

We examined the effects of season (dry or wet), habitat (factor with levels BV, DOW, KH, PV, and TR), plant part (flower, fruit, leaf, root, stem) and plant type (forb, grass, sedge, tree) on the nutritional qualities of the plants in three ways (see Table 1 for a list of our variables and their abbreviations in the model). To avoid problems of highly correlated variables in the PCA (see statistical analyses, below), we used values for hemicellulose and cellulose instead of ADF and NDF in the final analysis. Hemicellulose was calculated by subtracting ADF from NDF, and cellulose was calculated by subtracting lignins from ADF. For the first approach we first conducted a factor analysis of all of the nutritional variables (comprising ash, protein, hemicellulose, cellulose, lignin, fat, phenols, tannins, total energy, and metabolizable energy but not phytolith content) and modeled how the derived factors were affected by season, habitat, plant type, and plant part. Secondly, we fitted individual models for each nutritional component (this time including phytoliths), to explore how they were affected by season, habitat, plant type, and plant part. Phytoliths were excluded from the factor analysis because we had both nutritional and phytolith data for too few specimens to meet the sample size requirements of the factor analysis. Finally, we performed some simple comparisons of individual nutritional properties to see how they varied across the landscape.

Prior to conducting the analyses we inspected the distributions of the variables and transformed them where required to achieve a more symmetric distribution. We natural log transformed our variables true_ash, percent_lipid_dry, percent_phenol, and percent_tannin, and we square root transformed calc_met_energy (after subtracting its minimum) and true_lignin. The data were

appropriate for a factor analysis as indicated by Bartlett's test of sphericity (χ=1906.2, df=45, P<0.001) and the Kaiser-Meyer-Olkin measure of sampling adequacy (0.708). The first three principal components in a PCA analysis explained 75% of the cumulative variance, therefore we determined

three factors with the factor analysis. The first three factors accounted for 66.0% of the variance in the data. Factor 1 correlated positively with percent_tannin, percent_phenol, and percent_lignin and negatively with cellulose and hemicellulose, and accounted for 33.4% of the variance. Factor 2 correlated positively with true_protein and percent_lipid_dry, and accounted for 18.4% of the variance; and factor 3 correlated positively with calc_gross_energy and negatively with true_ash and accounted for 14.2% of the variance (Table 2). Metabolizable energy loaded nearly equally on all three factors (33 on factors 1 and 3, 29 on factor 2).

For the three factors, we then fitted linear mixed models (LMM, see Baayen, 2008) to explore how the plant and site variables explained the pattern seen in each factor (Supplementary Information Models 1-3). As fixed effects we included site, season, plant type, and plant organ, while the individual and the species of plant were included as random intercepts. To keep type I error rate at the nominal level of 5% we included a random slope of season (manually dummy coded and then centered) within plant species (Schielzeth and Forstmeier, 2009; Barr et al., 2013). We could not include any interactions among the fixed effects because none of the four fixed effects was fully crossed with any other. See Table 3 for the fixed and random effects used in all of our models.

To explore the variability or reliability of nutritional components within seasons, sites, plant types and plant organs, we then fitted linear mixed models using the same fixed and random effects, but instead using the absolute residuals from Models 1 through 3 as the response variable (Supplementary Information Models 4-6). This test is analogous to a Levene's Test for an ANOVA. Furthermore, because some of our predictions involved analysis of specific nutrients (e.g. protein, phytoliths), we fitted linear models with the same structure (same fixed and random effects) with each of the individual nutrient variables as the response (Supplementary Models 7-17).

For all models except 4-6, we then tested the significance of the four fixed effects as a whole by comparing each full model with a respective null model lacking the fixed effects but having the same random effects structure (Forstmeier and Schielzeth, 2011). For Models 4-6 we tested the individual fixed effects by re-running the model excluding each fixed effect one by one. We inspected QQ-plots of the residuals and residuals plotted against fitted values to check whether the assumptions of normally distributed and homogeneous residuals were fulfilled. These plots indicated no severe violations of these for the tests of the factors nor for the tests of the individual nutrient variables. For each model we evaluated its stability by excluding the levels of the random effects, one at a time. These indicated that the models were largely robust even when individual plants or species were removed. Collinearity was not an issue for any of the models as indicated by a maximum Generalized Variance Inflation Factor of 1.28 (Fox and Monette, 1992; Quinn and Keough, 2002). We fitted the models in R (version 3.3.1; R Core Team, 2016) using the function *lmer* of the R package *lme4* (version 1.1-12; Bates et al., 2015), and Variance Inflation Factors were determined using the function *vif* of the package *car* (version 2.1-2; Fox and Weisberg, 2011). For models 1-6, the sample size was the same: 225 measures conducted for 58 individual plants of 31 species. The sample sizes for the tests of the individual nutrients ranged from 237 to 259 measures of 62 to 104 individual plants of 31 to 60 species for the models of the individual nutrients (some models included more taxa than the factor analyses, depending on the number of samples we had processed for each nutrient or antifeedant).

Results

Approach 1: Factor analysis of nutritional properties.

This test explored which of our main proposed drivers of variation - plant type, plant organ, habitat (site), and season - most strongly influenced the overall variation in nutritional properties, as

represented by factor analysis. Each factor represented a combination of several of our nutritional properties.

For each of the three factors, the full model was clearly significant as compared to the respective null model (range of χ: 94.33 to 156.94 , all df=12, all $P<0.001$), indicating that fixed effects in the model (i.e., season, site, plant type, and plant organ) explained a significant amount of variation in the factors. More specifically, organ appeared to have a significant influence on all three factors (range of χ: 46.67 to 117.46, all df=4, all $P<0.001$), and the same was the case for site (range of χ: 15.69 to 32.25, all df=4, all $P<0.001$). Plant type had a significant effect on factor 1 (χ=30.97, df=3, $P<0.001$), but not on factor 2 (χ=5.94, df=3, P=0.114) or factor 3 (χ=5.68, df=3, P=0.128). Finally, seasonal was a significant influence only on factor 2 (χ=9.515, df=1, P=0.002), and not for factor 1 (χ=2.26, df=1, P=0.136) or factor 3 (χ=0.013, df=1, P=0.910). (Supplemental Tables)

Interpretation of factor 1:

As noted above, factor 1 correlated positively with percentage of lignin, tannins, and phenols, and negatively with cellulose and hemicellulose. Lignin, tannins, and phenols are strong antifeedants that can reduce the positive nutritional qualities of food by binding or sequestering proteins (Rhoades, 1979, and citations therein). Cellulose and hemicellulose are the structural fibers found in plant materials, and they require specialized gut adaptations and microbial associations in order to be digested. Plant type, plant organ, and habitat all influenced the variation in factor 1, but season did not. Unsurprisingly, grasses had significantly lower values for factor 1, forbs and sedges moderate values,

and trees were higher (Figure 2a, Figure 3). The quantities of these antifeedants and fibers did not seem to vary between seasons, and the differences among plant parts, while significant, were not large, with leaves having higher values and fruits and stems having low values. A bit surprising was the variation among habitats. Despite the preponderance of grasses in the Bloedveld open grassland habitat, this was the area with the highest values on factor 1, driven in part by the high values of factor 1 for the two trees sampled in this habitat (*Diospyros lycioides* and *Searsia lancia*). The trees alone do not explain this pattern, however, since among the grasses, those from Bloedveld also had higher loadings on factor 1 than those from other sites, indicating both greater antifeedants and lower cellulose and hemicellulose (Figure 2b). Kudu Hill, in contrast, had the lowest values for factor 1, despite being a more heavily forested area, indicating both fewer antifeedants and higher fiber content. Both grasses and trees in this habitat had lower values for factor 1. Figure 3b reveals a bimodal distribution for factor 1 values in Bloedveld, DOW, Kudu Hill, and to some extent, Tick River, reflecting differences between grasses and trees. An herbivore seeking to minimize antifeedants would therefore do well in Kudu Hill, regardless of whether they were a browser or grazer, though the higher cellulose and hemicellulose contents of these foods may deter non-ruminants who are unable to digest these resistant fibers.

Model 4 tested whether there was significant variation within any of our fixed effects (season, habitat, plant type, and organ), and allowed us to assess which of the states for these fixed effects was most variable. Of the fixed effects, plant type showed significant influence on the absolute residuals

(χ=9.557, df=3, P=0.023). Sedges have generally high values, reflecting their increased absolute residuals on Model 1 (Figure 4a). This indicates that they are a particularly unreliable food source, with some individuals having higher than expected values of factor 1, and others lower than expected values. An herbivore choosing to consume sedges could not predict prior to the first bite whether that individual sedge was strongly protected by lignin, tannins, and phenols, nor how much fiber it contained. The roots of sedges (including storage organs) are no exception to this pattern, with some

falling quite high on factor 1 and others rather lower (see Figure 1a).

Interpretation of factor 2:

Factor 2 was positively correlated with protein and lipid, the most energy-dense nutritional components we measured. This factor varies significantly with three of the fixed effects (site, season, plant organ) but not with plant type, possibly due to the high variability in protein values within each plant type (Figure 5). Among plant organs, flowers contain the highest values, followed by leaves. Fruits and roots have intermediate values while stems are very low, indicating very low protein and fat values. Factor 2 is the only factor which was strongly influenced by season. Perhaps unsurprisingly, values were higher in the wet season than in the dry, indicating more protein and lipids in the wet season. The very forested Kudu Hill had by far the highest values for factor 2, driven both by high values for the trees and for the grasses (Figure 1b). While trees represented the greatest biomass, the understory grasses were abundant and would have been a good source of nutrition. Our previous study of the mechanical properties of these grasses also suggests they may have been valuable foods for early hominins (Paine et al., 2018). This should be tempered, however, by our observation that the very-protein rich grasses found in Kudu Hill, notably the *Setaria verticillata*, were fiercely protected in the dry season by spiky panicles that clung together and would stick to any available surface. Though likely difficult to consume, we note that this grass is favored by baboons (Altmann, 1998), and the mechanical protection is possibly a response to high foraging pressure on this taxon as well as an efficient means of seed dispersal. Peter's Vlei, the marshy area, had low values on factor 2, suggesting this habitat had little to offer in the way of protein and lipids.

Model 5 assessed the variation within each of our fixed effects for Model 2. None of the four fixed effects appeared to have higher than expected variation in protein and lipid content.

Interpretation of factor 3

Factor 3 was correlated positively with gross energy and negatively with ash. We found it quite interesting that metabolizable energy did not follow the same loading pattern as gross energy, but note that the loading of metabolizable energy was not particularly strong on any factor (Table 2). Because the calculation of metabolizable energy depends on the herbivore chosen as the model consumer, the measure of gross energy might better represent the energy available to all consumers in a habitat. Factor 3 was strongly influenced only by site and organ. Roots had lowest values (lots of ash, little energy) and leaves were highly variable (Figure 6). We were surprised by the low value of roots on factor 3, given that the roots in our analysis include only those we specifically thought might be of the highest nutritional value, including *Typha*, *Juncus*, and several sedges (*Kyllinga*, *Mariscus*). These roots are often mentioned as good sources of non-structural carbohydrates (sugars and starches; NSC) (Schoeninger et al., 2001; Laden and Wrangham, 2005; Marlowe and Berbesque, 2009; Schoeninger et al., 2001). Gross energy was the only variable that accounted for NSC yet it was the variable on which these roots had the lowest values. Despite having potentially higher NSC values, roots did not contain enough of the other nutritional components (proteins, lipids, NDF) to achieve high energy value. In contrast to the pattern observed for factor 2, Bloedveld had the highest loadings on factor 3, indicating high gross energy and low ash, distantly followed by Kudu Hill and DOW (Figure 7). Peter's Vlei had the lowest scores, while Tick River was not much higher. This is likely due to the higher structural fiber values in the plants at Bloedveld. Gross energy is calculated using proteins and lipids but also NSC and NDF, the latter of which was much higher at Bloedveld than at Kudu Hill. For animals who are able to access the energy available in NDF, Bloedveld would provide more calories than Kudu Hill.

The examination of the variability within loadings on factor 3 (Model 6) revealed that site and plant organ had differing levels of variation (Figure 4b and c). Leaves were much less variable than the

other plant organs, indicating these would be rather reliable foods in terms of their gross energy and ash content. Possibly due to the abundance of grasses, Bloedveld had the least variation within factor 3, again making it a predictable habitat. Kudu Hill was the most variable, and therefore least predictable in terms of ash and gross energy content.

Approach 2: LMMs of individual nutritional qualities as a factor of site, season, plant part and plant type:

We fitted LMMs to each individual nutritional variable, and comparison to the full null models clearly revealed significance (range of χ: 24.73 to 219.39, all df=12, all $P \leq 0.02$). The patterns among the individual variables followed broadly the patterns for the factors on which they most strongly loaded during the analysis of Approach 1. We briefly discuss each individual nutritional variable below, in the same order as they loaded on the factor analysis. Detailed results are presented in the Supplemental Tables Models 7-17.

Tannins were strongly influenced by season, organ, and type. Grasses had lower tannin content than forbs and sedges, while trees had quite high values. Leaves had relatively high tannin content, but roots were surprisingly high. The roots in our analysis were those specifically chosen for their potential edibility, including the rhizomes and corms from the aquatic plants in Peter's Vlei, so this high value was unexpected. Fruits and stems had generally low values of tannins. Tannins were higher in the wet season.

Phenols follow a very similar pattern to tannins, in that season, organ, and type, and to a lesser extent site were strong drivers of variation. Trees had the highest phenol content, while grasses had the lowest. Stems, and to a lesser extent fruits, had less phenol than leaves and roots. Bloedveld had the greatest amount of phenols and Kudu Hill the least but the difference was not as striking as among the other chemical antifeedants, and the other sites had intermediate values. Wet season samples had more

phenol.

Lignins were strongly influenced by plant type and site. Organ was not a significant driver of variation in lignin. Trees stand out as having exceptionally high lignin content. Kudu Hill and Bloedveld are the two endpoints on the range of lignin values among sites, but Kudu Hill is the clear outlier among the sites in having very low lignin values.

The two main non-lignin fiber components, *cellulose* and *hemicellulose*, were both strongly affected by type and organ, but only cellulose was also influenced by site. Grasses stand out as having drastically higher cellulose and hemicellulose content, followed by sedges. Forbs and trees had generally much lower values (up to 13% less). Sedges seem to have more hemicellulose than cellulose. The pattern of cellulose and hemicellulose differs among plant organs. Cellulose is lowest in roots and slightly higher in fruits. Leaves and flowers have intermediate values while stems have the highest values. In contrast, hemicellulose is lowest in fruits, followed by roots. Leaves and flowers are again intermediate and stems highest. Cellulose is highest in Peter's Vlei, while the other sites cluster more closely, with Tick River having the lowest values, followed closely by Kudu Hill and DOW, and Bloedveld being somewhat higher.

Protein varied significantly according to plant organ, season, and site. Flowers had the highest protein content, while fruits and stems were lowest. Leaves and roots had moderate to high values. Wet season samples had more protein than dry. Differences among sites were the most striking, however, with Kudu Hill having the highest protein content, with a large gap between this habitat and all others (Figure 8). While values in Bloedveld were low, Peter's Vlei had the lowest values.

Despite loading with protein on factor 2, *lipid* values varied significantly only with plant type and plant organ. Trees had the highest content of lipids, and sedges the lowest, while grasses and forbs had moderate values. Fruits were relatively lipid-rich, while roots and stems had the lowest lipid values.

Calculated gross energy varied significantly only with plant organ, and was highest in fruits,

likely driven by the high percentage of lipids in these plant organs. Flowers were next highest, followed by stems, leaves, and finally roots. Again the low values of roots was a surprise, given our choice of those roots thought to be the most potentially valuable for a hominin consumer.

A significant amount of variation in percentage of *ash* is explained by plant organ and by site. Roots and leaves have more ash than fruits do, while stems and flowers appear to be more moderate. Bloedveld has the lowest ash content while Tick River had the highest. Surprisingly, given the generally high percentages of protein and gross energy, Kudu Hill also had high percentages of ash. While ash does not provide calories, it could provide salts or other vital minerals. Given the low phytolith content of the plants in Kudu Hill (see below), this ash is not likely to be dominated by silica.

Of all of our models exploring the effect of season, site, type, and organ on the individual nutritional components, the one for *calculated metabolizable energy* was least strongly supported, though still significant (P=0.020). It varied significantly only with plant organ, with fruits having the highest available metabolizable energy, while all of the other plant organs had relatively low levels. Roots have the lowest values, again suggesting these food resources would not have been of great value to hominin foragers.

A significant amount of the variation within *phytolith* content was explained by each of our fixed effects (plant type, plant organ, season, and site), though season was only significant at the $p < 0.05$ level. As expected given the known distribution of phytoliths across plant types (Piperno, 2006), grasses and sedges had significantly more phytoliths than did trees or forbs. Somewhat more surprisingly, roots and inflorescences had the highest values of phytoliths, while leaves had moderate values. Stems and especially fruits had low values. Dry season plants had more phytoliths, but this difference was less pronounced than that between types or organs. The two woodland habitats, DOW and Kudu Hill, had overall low phytolith contents while the two wet habitats, Tick River and Peter's Vlei had higher values. Bloedveld, the grassland, had a surprisingly moderate phytolith content.

The analyses of individual nutrients revealed a stronger effect of season than was present in the

factor analyses. Proteins, lignins, tannins, and phytoliths all had strongly seasonal variation. Overall, however, site, plant type, and plant organ more strongly influenced the nutritional qualities of the plants we sampled, suggesting these are more important to consider when predicting how nutrients vary across a landscape.

Approach 3: Correlations among individual nutritional properties:

Protein vs. non-protein energy

As presented in the introduction above, previous studies of primates and other animals have emphasized that animals prefer a fixed ratio of protein energy to non-protein energy (NPE) in the long term, but during food stress some will prioritize protein energy while others will prioritize NPE depending on the foods available in their environment. We calculated the protein energy (as 0.239 x g protein in 100g of dry plant) and non-protein energy (0.398 * % lipid + 0.201 * % NDF + 0.175 * % NSC) in kJ per gram of plant material. NSC stands for non-structural carbohydrates, and as is calculated as 100 - (% protein + % NDF + % lipid + % ash). The plot of protein vs. non-protein energy reveals that the grasses and trees from Kudu Hill are exceptionally high in protein, though relatively lacking in non-protein energy (Figure 9). While we were unable to sample many fruits, these clearly stood out from the other plants in having very high non-protein energy. The specimen with the highest non-protein energy was the fruit of *Ximenia caffra* from Kudu Hill. The tree is commonly known as the large sourplum and is a member of the olive family. The fruit pulp is regularly consumed by humans and baboons (Peters et al., 1992). The other plant organs with high non-protein energy values (above 20 kJ per 100g) were the stems of grasses and forbs, though these were usually low in protein.

Protein vs. antifeedants

While the protein and NPE energy content are important, antifeedants such as lignins, phenols and phytoliths may increase the costs of digestion. Therefore in Figure 10 we compared protein vs.

quantities of phenols and tannins, and for a reduced data set for which we had phytolith data, we plotted protein vs. phytoliths (Figure 11).

Previous studies have shown that plants produce increased chemical protection (tannins and phenols) in response to damage by herbivores (Karban and Myers, 1989). If protein-rich plants were preferentially consumed, then we might expect them to have higher percentages of lignins and phenols. There was a significant relationship between protein and the combined percentage of lignins and phenols in grasses only (Spearman's ρ -0.479, $P< 0.000$), but not in the other plant types. However, the direction of the correlation in the grasses was opposite to what we expected - protein-rich grasses had a lower percentage of lignins and phenols than protein-poor plants. In particular, the high-protein plants in Kudu Hill were distinctly low in lignins and phenols, further emphasizing their nutritional value.

Plants increase phytolith production in response to grazing (Massey et al., 2007). As with tannins and phenols, if protein-rich plants are preferentially sought out by consumers, then it is possible that protein-rich plants would have high phytolith contents. However, we found no significant correlation between protein and phytoliths, even when we considered the four plant types separately (Spearman's ρ ranged from -0.03 to 0.25). This lack of relationship between protein and phytoliths may suggest that high protein plants are not preferentially foraged in this landscape, perhaps due to the plants using other means of protection against foraging, or due to other restrictions (e.g., unfavorable fiber contents).

Discussion:

The factor analysis, GLMs of the factors, and models for each individual nutrient have provided us the ability to answer the discrete questions posed in the introduction to this paper.

Which habitat has the most calories? Which has the most protein?

Protein clearly and significantly varied among habitats, with Kudu Hill having the most protein in both grass and tree leaves. Calories had a less clear pattern, with both gross and metabolizable energy showing no consistent significant variation among habitats. However, fruits were always the most calorie-rich plant organs, supporting the idea they were important hominin foods (e.g, Dunbar, 1976; Peters and Vogel, 2005), and suggesting that habitats with more fruits may have been preferred, at least during periods when fruits were available.

Protein and lipid content, two of the major components of energy for non-ruminant consumers, loaded together on factor 2. The LMM of factor 2 again indicated that Kudu Hill was potentially the most desirable habitat, though the assessment of variability in factor 3 suggested that Kudu Hill was perhaps the least predictable habitat, at least in terms of gross energy and ash content. This variability might be driven by the relatively high ash content of the Kudu Hill plants.

Where is the antifeedant load the least?

Phenols, tannins and lignins were most common in trees and in leaves. Bloedveld (the open grassland) stood out as having the highest non-silica antifeedant load, but surprisingly had only moderate phytolith content. Kudu Hill had by far the lowest concentrations of all antifeedants, making it an attractive habitat.

Where is there the most bio-available nutrition? Are certain plants better choices than others?

Calculated metabolizable energy was the least well-supported of our individual nutritional properties, and was found to be highest in fruits. The roots included in our study, which comprised mostly those potentially valuable USOs from wetland taxa, were very low in calculated metabolizable energy. Other means of looking at bio-available nutrition, namely looking at protein vs. antifeedant content, indicated quite strongly that the plants available in Kudu Hill provided the highest protein with

the lowest costs. Surprisingly, the wetland and near-water habitats in our analysis (Tick River and Peter's Vlei) were never characterized by high amounts of any of the preferred nutritional qualities, suggesting these habitats would not be particularly valuable to hominins. Furthermore, the sedge USOs included in our studies had quite high cellulose and hemicellulose values, and were unreliable in terms of antifeedant content. This strongly suggests that the sedge USOs available in the Cradle region were not typically valuable foods.

Which habitats are most similar to each other? Which are most different? What drives this variation?

In most cases, Bloedveld and Kudu Hill stood out as the most different to each other, though not always in the anticipated ways. We expected that Bloedveld would have had moderate protein values (and higher in the wet season when new growth was available) and lower tannin, phenol and lignin values, given that these antifeedants are often associated with trees. We had further expected a high phytolith content in Bloedveld, given the abundance of grass taxa. We observed, however, that the protein values were rather low and the chemical antifeedants were quite high in this grassland habitat. Phytolith content was high, but lower than that of the two wetland habitats.

The two water-associated habitats, Peter's Vlei and Tick River, were similar in some of our analyses, probably because of the presence of sedges in both habitats. Sedges were quite poor foods, and generally stood apart from grasses and trees in their nutritional properties.

Is the nutritional variation among habitats greater than that between seasons?

The variation in all three of our factors was strongly influenced by site, while season strongly influenced the variation only in factor 2. We could interpret this to mean that habitat was perhaps more important than season in driving nutritional variation. However, in the absence of strong patterns in metabolizable energy, factor 2 represents the most calorically valuable qualities of plants (proteins and lipids). These vary between seasons, suggesting that protein might be restricted in most of these

habitats on a seasonal basis.

Which plants and habitats are the most reliable? Which are the least?

Sites were not clearly more or less variable for factor 1 and 2, but for factor 3 (which is correlated positively with gross energy and negatively with ash) Kudu Hill stood out as a variable habitat. Though protein and lipid values in this habitat are high, the low fiber and high ash content of the plants in this habitat could have driven this variability in factor 3. Sedges were more variable on factor 1, suggesting that their antifeedant and fiber content could not be reliably predicted.

Conclusions:

All of our driving factors (season, habitat, plant type and plant organ) strongly influenced the various plant nutritional and antifeedant properties, both in combination and individually. We were unable to assess any interactions among our variables, and therefore could not assess, for example, the combined effect of season and habitat. However, our results allowed us to explore the patterns of variation among these plants, which were occasionally surprising.

Despite strong temperature and rainfall variation between seasons in the Cradle region, most nutritional properties remained relatively constant between the wet and dry seasons. Only tannins, phenols and protein differed among seasons, and all three were higher in the wet season. The amount of increase was rather low: on average across all of the samples, the protein value increased by about 0.7% in the wet season, while the tannin and phenol contents increased by about 0.1%.

Habitats had a stronger effect on the nutritional value, suggesting that patch choice models may be appropriate for exploring hominin feeding behaviors. Woodland habitats, particularly that at Kudu Hill, were more nutritionally valuable than we had anticipated. Grasses from these wooded environments were generally good resources with relatively high protein contents and low chemical and mechanical antifeedant contents. These plants may be particularly important in the dry season,

when protein content among all of the sampled plants was lower. However, these grasses do not appear to have much in the way of non-protein energy, so would not be good for an animal that leverages NPE.

Sedges, USOs, and most other plants available from wetland or river-edge habitats in our study area were surprisingly low in protein, NPE, and calculated metabolizable energy. These poor nutritional values suggest that wetlands and river edges would have been marginal habitats for hominins, at least within the Highveld floristic region that characterizes the Cradle today.

The Highveld region in which we sampled is unlikely to be representative of most potential hominin landscapes across Africa. Similarly elevated and therefore cold semi-arid plains or plateaus are limited to the central provinces of South Africa, the Ethiopian highlands and some regions in the Maghreb. Plant communities in areas exposed to freezing temperatures are different than those in warmer climates, and are generally reduced in the number of taxa and may have different seasonal patterns. Future projects exploring nutritional and antifeedant variation in more representative habitats may help us to better understand how hominins may have used their landscapes in different climatic, tectonic and environmental contexts.

Acknowledgments:

We thank Jack, Karen, and the staff at Kenjara Lodge for their hospitality. This work was supported by The Leakey Foundation, the Wenner-Gren Foundation (grant #8965), the National Science Foundation (grant # 1613421), the Max Planck Society, and the University of Colorado Boulder. It was also funded in part by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program under grant agreement number STG–677576 (“HARVEST”). R. Mundry greatly assisted with the statistical analysis, but any mistakes should be attributed to AGH.

References:

Alerstam, T., Hedenström, A., Åkesson, S., 2003. Long-distance migration: evolution and determinants. Oikos. 103, 247–260.

Avery, D.M., 2001. The Plio-Pleistocene vegetation and climate of Sterkfontein and Swartkrans, South Africa, based on micromammals. Journal of Human Evolution. 41, 113–132.

Baayen, R.H., 2008. Analyzing Linguistic Data: A Practical Introduction to Statistics using R. Cambridge University Press.

Bamford, M., 1999. Pliocene fossil woods from an early hominid cave deposit, Sterkfontein, South Africa. South African Journal of Science. 95, 231–237.

Barr, D.J., Levy, R., Scheepers, C., Tily, H.J., 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. Journal of Memory and Language. 68, 255–278.

Bates, D., Mächler, M., Bolker, B., Walker, S., 2015. Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software. 67, 1–48.

Belovsky, G.E., 1978. Diet optimization in a generalist herbivore: The moose. Theoretical Population Biology. 14, 105–134.

Belovsky, G.E., 1984. Herbivore Optimal Foraging: A Comparative Test of Three Models. The American Naturalist. 124, 97–115.

Belovsky, G.E., 1990. How important are nutrient constraints in optimal foraging models or are spatial/temporal factors more important? In: Hughes, R.N. (Ed.), Behavioural Mechanisms of Food Selection, NATO ASI Series. Springer-Verlag, Berlin, pp. 255–280.

Bond, W., Midgley, G., Woodward, F., 2003. What controls South African vegetation—climate or fire? South African Journal of Botany. 69, 79–91.

Bryant, J.P., Kuropat, P.J., 1980. Selection of winter forage by subarctic browsing vertebrates: the role of plant chemistry. Annual Review of Ecology and Systematics. 11, 261–285.

Cadman, A., Rayner, R.J., 1989. Climatic change and the appearance of Australopithecus africanus in

the Makapansgat sediments. Journal of Human Evolution. 18, 107–113.

Codron, D., Lee-Thorp, J.A., Sponheimer, M., Codron, J., 2006. Nutritional content of savanna plant foods: implications for browser/grazer models of ungulate diversification. European Journal of Wildlife Research. 53, 100–111.

Codron, J., Duffy, K.J., Avenant, N.L., Sponheimer, M., Leichliter, J., Paine, O., Sandberg, P., Codron, D., 2015. Stable isotope evidence for trophic niche partitioning in a South African savanna rodent community. Current Zoology. 61, 397–411.

Conklin-Brittain, N.L., Knott, C.D., Wrangham, R.W., 2006. Energy intake by wild chimpanzees and orangutans: methodological considerations and a preliminary comparison. In: Hohemann, G., Robbins, M.M., Boesch, C. (Eds.), Feeding Ecology in Apes and Other Primates: Ecological, Physiological and Behavioural Aspects. Cambridge Univeristy Press, pp. 445–472.

de Ruiter, D.J., Sponheimer, M., Lee-Thorp, J.A., 2008. Indications of habitat association of Australopithecus robustus in the Bloubank Valley, South Africa. Journal of Human Evolution. 55, 1015–1030.

deMenocal, P.B., 2004. African climate change and faunal evolution during the Pliocene–Pleistocene. Earth and Planetary Science Letters. 220, 3–24.

Denys, C., 1992. Les analyses multivariées: Une aide al'interprétation des paléoenvironnements. l'exemple des rongeurs Plio-Pléistocènes d'Afrique Australe. Geobios. 25, 209–217.

Dominy, N.J., Vogel, E.R., Yeakel, J.D., Constantino, P., Lucas, P.W., 2008. Mechanical properties of plant underground storage organs and implications for dietary models of early hominins. Evolutionary Biology. 35, 159–175.

Dunbar, R.I.M., 1976. Australopithecine diet based on a baboon analogy. Journal of Human Evolution 5, 161–167. https://doi.org/10.1016/0047-2484(76)90019-1

Duncan, A.J., Gordon, I.J., 1999. Habitat selection according to the ability of animals to eat, digest and detoxify foods. Proceedings of the Nutrition Society. 58, 799–805.

Felton, A.M., Felton, A., Raubenheimer, D., Simpson, S.J., Foley, W.J., Wood, J.T., Wallis, I.R., Lindenmayer, D.B., 2009. Protein content of diets dictates the daily energy intake of a free-ranging primate. Behavioral Ecology. arp021.

Forstmeier, W., Schielzeth, H., 2011. Cryptic multiple hypotheses testing in linear models: overestimated effect sizes and the winner's curse. Behavioral Ecology and Sociobiology. 65, 47–55.

Fox, J., Monette, G., 1992. Generalized Collinearity Diagnostics. Journal of the American Statistical Association. 87, 178–183.

Fox, J., Weisberg, S., 2011. An R Companion to Applied Regression, Second. ed. Sage Publications, Thousand Oaks CA.

Glander, K.E., 1982. The impact of plant secondary compounds on primate feeding behavior. American Journal of Physical Anthropology. 25, 1–18.

Henry, A.G., Ungar, P.S., Passey, B.H., Sponheimer, M., Rossouw, L., Bamford, M., Sandberg, P., de Ruiter, D.J., Berger, L., 2012. The diet of Australopithecus sediba. Nature. 487, 90–93.

Hopley, P., Latham, A., Marshall, J., 2006. Palaeoenvironments and palaeodiets of mid-Pliocene micromammals from Makapansgat Limeworks, South Africa: A stable isotope and dental microwear approach. Palaeogeography, Palaeoclimatology, Palaeoecology. 233, 235–251.

Hughes, R.N. (Ed.), 1990. Behavioral Mechanisms of Food Selection, NATO ASI series. Springer-Verlag, Berlin.

Illius, A.W., Gordon, I.J., 1990. Constraints of diet selection and foraging behavior in mammalian herbivores. In: Hughes, R.N. (Ed.), Behavioural Mechanisms of Food Selection. Springer-Verlag, Berlin.

Johnson, C.J., Nielsen, S.E., Merrill, E.H., Mcdonald, T.L., Boyce, M.S., 2006. Resource Selection Functions Based on Use–Availability Data: Theoretical Motivation and Evaluation Methods. Journal of Wildlife Management. 70, 347–357.

Kamphues, P.D.J., 2014. Supplemente zur Tierernährung für Studium und Praxis. Schlütersche.

Karban, R., Myers, J.H., 1989. Induced plant responses to herbivory. Annual Review of Ecology and Systematics. 20, 331–348.

Laden, G., Wrangham, R., 2005. The rise of the hominids as an adaptive shift in fallback foods: Plant underground storage organs (USOs) and australopith origins. J. Hum. Evol. 49, 482–498.

Lee-Thorp, J.A., Sponheimer, M., Luyt, J., 2007. Tracking changing environments using stable carbon isotopes in fossil tooth enamel: an example from the South African hominin sites. Journal of Human Evolution. 53, 595–601.

Lindroth, R.L., 1989. Mammalian herbivore-plant interactions. In: Abrahamson, W.G. (Ed.), Plant-Animal Interactions. MCGraw Hill, New York, pp. 163–206.

Lindroth, R.L., Batzli, G.O., Seigler, D.S., 1986. Patterns in the phytochemistry of three prairie plants. Biochemical Systematics and Ecology. 14, 597–602.

Low, B., Rebelo, A., 1996. Vegetation of Southern Africa, Lesotho and Swaziland: a companion to the vegetation map of South Africa, Lesotho and Swaziland. Department of Environmental Affairs and Tourism, Pretoria, South Africa.

Macho, G.A., 2014. Baboon Feeding Ecology Informs the Dietary Niche of Paranthropus boisei. PLoS ONE. 9, e84942.

Manly, B.F.L., McDonald, L., Thomas, D.L., McDonald, T.L., Erickson, W.P., 2007. Resource selection by animals: statistical design and analysis for field studies, Second edition. ed. Springer Science & Business Media.

Marlowe, F.W., Berbesque, J.C., 2009. Tubers as fallback foods and their impact on Hadza hunter-gatherers. American Journal of Physical Anthropology. 140, 751–758.

Massey, F.P., Ennos, A.R., Hartley, S.E., 2007. Herbivore specific induction of silica-based plant defences. Oecologia. 152, 677–683.

Milewski, A.V., Mills, A.J., 2015. Why was the Highveld treeless? Looking laterally to the Pampas for

global edaphic principles beyond biogeographical accidents. South African Journal of Botany, Biome Boundaries in South Africa. 101, 98–106.

Mucina, L., Rutherford, M.C., 2006. The vegetation of South Africa, Lesotho and Swaziland.

Owen-Smith, N., Cooper, S.M., 1987. Palatability of Woody Plants to Browsing Ruminants in a South African Savanna. Ecology 68, 319–331. https://doi.org/10.2307/1939263

Owen-Smith, N., Cooper, S.M., 1989. Nutritional ecology of a browsing ruminant, the kudu (Tragelaphus strepsiceros), through the seasonal cycle. Journal of Zoology 219, 29–43. https://doi.org/10.1111/j.1469-7998.1989.tb02563.x

Pagan, J.D., 1999. Carbohydrates in equine nutrition. Feed mix. 29–42.

Paine, O.C.C., Koppa, A., Henry, A.G., Leichliter, J.N., Codron, D., Codron, J., Lambert, J.E., Sponheimer, M., 2018. Grass leaves as potential hominin dietary resources. Journal of Human Evolution. 117, 44–52.

Peters, C.R., O'Brien, E.M., Drummond, R.B., 1992. Edible wild plants of Sub-Saharan Africa. Royal Botanic Gardens, Kew.

Peters, C.R., Vogel, J.C., 2005. Africa's wild C4 plant foods and possible early hominid diets. Journal of Human Evolution 48, 219–236. https://doi.org/10.1016/j.jhevol.2004.11.003

Peterson, A., Abella, E.F., Grine, F.E., Teaford, M.F., Ungar, P.S., 2018. Microwear textures of Australopithecus africanus and Paranthropus robustus molars in relation to paleoenvironment and diet. Journal of Human Evolution. 119, 42–63.

Pickering, R., 2018. Geology and uranium-lead (U-Pb) dating of the South African Paranthropus-bearing cave deposits. American Journal of Physical Anthropology. 165, 208.

Pickering, R., Kramers, J.D., 2010. Re-appraisal of the stratigraphy and determination of new U-Pb dates for the Sterkfontein hominin site, South Africa. Journal of Human Evolution. 59, 70–86.

Piperno, D.R., 2006. Phytoliths: A comprehensive guide for archaeologists and paleoecologists. AltaMira Press, Lanham, Maryland.

Quinn, G.P., Keough, M.J., 2002. Experimental Design and Data Analysis for Biologists. Cambridge University Press.

R Core Team, 2016. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Reed, K.E., 1997. Early hominid evolution and ecological change through the African Plio-Pleistocene. Journal of Human Evolution. 32, 289–322.

Rhoades, D.F., 1979. Evolution of plant chemical defense against herbivores. In: Rosenthal, G.A., Janzen, D.H. (Eds.), Herbivores: Their Interaction with Secondary Plant Metabolites. Academic Press, New York, pp. 4–55.

Righini, N., Fernandez, A., Garber, P.A., Rothman, J.M., 2014. Howler monkey nutritional geometry: What we can learn from an interspecific approach. In: Program of the Thirty-Seventh Meeting of The American Society of Primatologists. p. 98.

Romanowski, J., Żmihorski, M., 2009. Seasonal and habitat variation in the diet of the tawny owl (Strix aluco) in central Poland during unusually warm years. Biologia. 64, 365–369.

Rothman, J.M., Raubenheimer, D., Chapman, C.A., 2011. Nutritional geometry: gorillas prioritize non-protein energy while consuming surplus protein. Biology Letters. rsbl20110321.

Schakel, S.F., Buzzard, I.M., Gebhardt, S.E., 1997. Procedures for estimating nutrient values for food composition databases. Journal of Food Composition and Analysis 10, 102–114.

Schielzeth, H., Forstmeier, W., 2009. Conclusions beyond support: overconfident estimates in mixed models. Behavioral Ecology. 20, 416–420.

Schoeninger, M.J., Bunn, H.T., Murray, S.S., Marlett, J.A., 2001. Composition of Tubers Used by Hadza Foragers of Tanzania. Journal of Food Composition and Analysis. 14, 15–25.

Sept, J., 2001. Modeling the edible landscape. In: Stanford, C.B., Bunn, H.T. (Eds.), Meat-Eating and Human Evolution. Oxford University Press, New York, pp. 73–98.

Sponheimer, M., Passey, B.H., Ruiter, D.J. de, Guatelli-Steinberg, D., Cerling, T.E., Lee-Thorp, J.A.,

2006. Isotopic Evidence for Dietary Variability in the Early Hominin Paranthropus robustus. Science. 314, 980–982.

Staver, A.C., Archibald, S., Levin, S., 2011. Tree cover in sub-Saharan Africa: Rainfall and fire constrain forest and savanna as alternative stable states. Ecology. 92, 1063–1072.

Stohlgren, T.J., Falkner, M.B., Schell, L.D., 1995. A Modified-Whittaker nested vegetation sampling method. Vegetatio. 117, 113–121.

Stoner, K.E., 1996. Habitat selection and seasonal patterns of activity and foraging of mantled howling monkeys (Alouatta Palliata) in Northeastern Costa Rica. International Journal of Primatology. 17, 1–30.

Takahashi, M.Q., 2018. The Nutritional Ecology of Adult Female Blue Monkeys, Cercopithecus mitis, in the Kakamenga Forest, Kenya. Columbia University, New York.

Ungar, P.S., Sponheimer, M., 2011. The Diets of Early Hominins. Science. 334, 190–193.

van der Merwe, N.J., Thackeray, J.F., Lee-Thorp, J.A., Luyt, J., 2003. The carbon isotope ecology and diet of Australopithecus africanus at Sterkfontein, South Africa. Journal of Human Evolution. 44, 581–597.

Vrba, E.S., 1980. The significance of bovid remains as indicators of environment and predation patterns. In: Fossils in the Making: Vertebrate Taphonomy and Paleoecology. University of Chicago Press, Chicago, pp. 247–272.

Wrangham, R., Cheney, D., Seyfarth, R., Sarmiento, E., 2009. Shallow-water habitats as sources of fallback foods for hominins. American Journal of Physical Anthropology. 140, 630–642.

890

Table 1. Tested nutritional and antifeedant properties and their abbreviations used in the statistical analyses.

Plant property	Abbreviation	Description
crude ash	true_ash	Percent of ash in dry matter
crude protein	true_protein	Percent of protein in dry matter
cellulose	cellulose	Percent of cellulose in of dry matter, calculated by subtracting acid detergent fiber (ADF) from neutral detergent fiber (NDF)
hemicellulose	true_ADF	Percent of hemicellulose in dry matter, calculated by subtracting lignin from ADF
lignin	true_lignin	Percent of lignin in dry matter
tannin	percent_tannin	Percent of tannins in dry matter
phenols	percent_phenol	Percent of phenols in dry matter
crude lipids	percent_lipid_dry	Percent of lipids in dry matter
phytoliths	percent_phyt_dry	Percent of phytoliths in dry matter
total energy	calc_gross_energy	Total energy in mJ/kg, based on Kamphues (2014) and Pagan (1999)
metabolizable energy	calc_met_energy	Bioavailable energy in mJ/kg, based on Conklin-Brittain et al. (2006)

895

Table 2. Results of the factor analysis. Indicated are the loadings of the variables on the three factors (largest loading per variable in bold), Eigenvalues of the factors and the percent of variance in the data each explained.

variable	Factor 1	Factor 2	Factor 3
log.percent_tannin	**0.97**	0.21	-0.06
log.percent_phenol	**0.96**	0.26	0.02
sqrt.true_lignin	**0.57**	0.15	0.17
hemicellulose	**-0.60**	-0.38	0.15
cellulose	**-0.63**	-0.47	0.16

true_protein	0.25	**0.88**	-0.09
log.percent_lipid_dry	0.45	**0.64**	0.21
log.true_ash	0.14	0.29	**-0.94**
calc_gross_energy	0.03	0.1	**0.54**
sqrt.calc_met_energy	**0.33**	**0.29**	**0.33**
Eigenvalue	2.184	1.331	0.9960
% variance explained	33.452	18.472	14.201

900

Table 3: Fixed and random effects used in the modeling.

Fixed effects:		
variable name	levels	explanation
season	wet	plants collected in January 2013
	dry	plants collected in June 2013
site	BV	Bloedveld open grassland
	DOW	Dolomite open woodland
	KH	Kudu Hill woodland
	PV	Peter's Vlei wetland
	TR	Tick River river margin
simple_organ	flower	inflorescence
	fruit	reproductive part, including seeds, drupes, nuts
	leaf	for trees, leaves were stripped from branches and petioles were included. for grasses leaves including the sheathes were separated from the culm
	root	any underground part, including storage organs where available
	stem	for trees the stems were usually only lateral shoots or branches, for grasses the stems were usually the culm with a minimal amount of leaf sheathes
plant_type	forb	low-growing herbaceous dicotyledonous plants that did not fall into any of the other plant categories
	grass	monocotyledonous plants in the family Poaceae
	sedge	monocotyledonous plants in the family Cyperaceae
	tree	tall (>3m) woody dicotyledonous plants

Random effects		
plant_id	(multiple)	a unique number given to each plant individual replicate. Multiple plant parts are represented from each replicate.
species	(multiple)	the unique species of plant, of which there were multiple replicates
season.code	(multiple)	a random slope of season within species

Figures:

Figure 1)

Representative photos of each of the five habitats included in our study. The top row shows each habitat during the dry season (January 2014) and the bottom row during the wet season (July 2014). From left to right the five habitats are grassland (Bloedveld or BV, 25°54'37.65"S, 27°51'45.57"E), river margin (Tick River or TR, 25°55'1.69"S, 27°52'36.25"E), marsh (Peter's Vlei or PV, 25°55'10.31"S, 27°51'52.83"E), open woodland (Dolomite Open Woodland or DOW, 25°54'40.64"S, 27°50'38.06"E) and more dense woodland (Kudu Hill or KH, 25°55'13.57"S, 27°50'8.47"E). The upper left photo includes members of the team during the sampling.

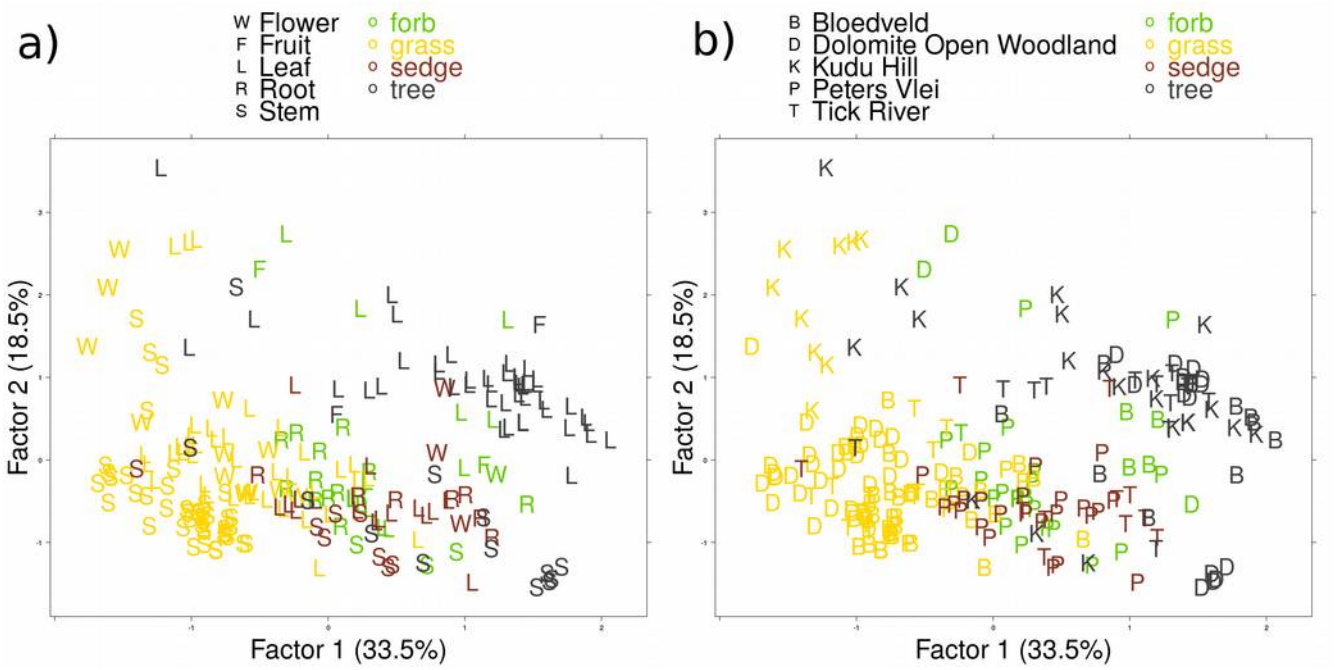


Figure 2)

Each sample included in our analysis is indicate by their loadings on factor 1 and factor 2. In a), the organ to which each sample belongs is indicated by the letter, while the plant type is indicated by the color. In b) the site to which each sample belongs is indicated by the letter. Visibly, grasses separate from trees, while plants from Kudu Hill cluster at the margin of the range of other habitats.

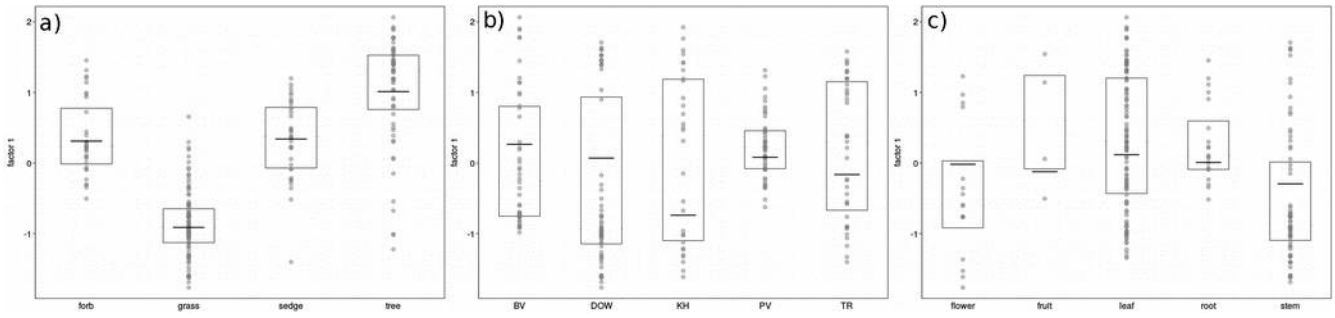


Figure 3)

Factor 1 varied significantly according to type, site and organ. Loadings for all samples on factor 1 separated by a) plant type; b) site; and c) organ.

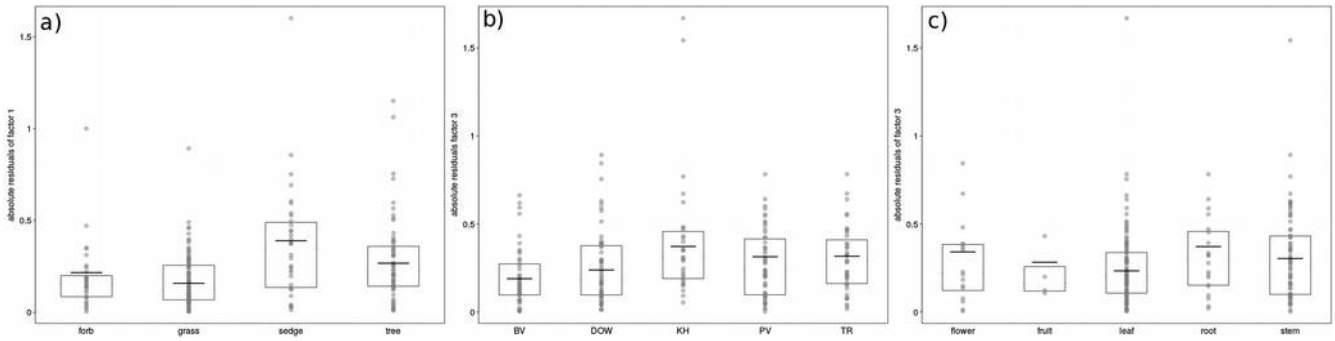


Figure 4)

The models of the absolute residuals from Model 1 and Model 3 indicated that some groups of plants were more variable (less reliable) than others. a) Absolute residuals from Model 1 plotted according to plant type. Sedges have higher values, indicating higher variability in Model 1. b) Absolute residuals from Model 3 plotted according to site. Kudu Hill has slightly but significantly higher values than the other sites. c) Absolute residuals from Model 3 plotted according to organ. Roots have slightly but significantly higher values than the other plant types, while leaves were particularly low.

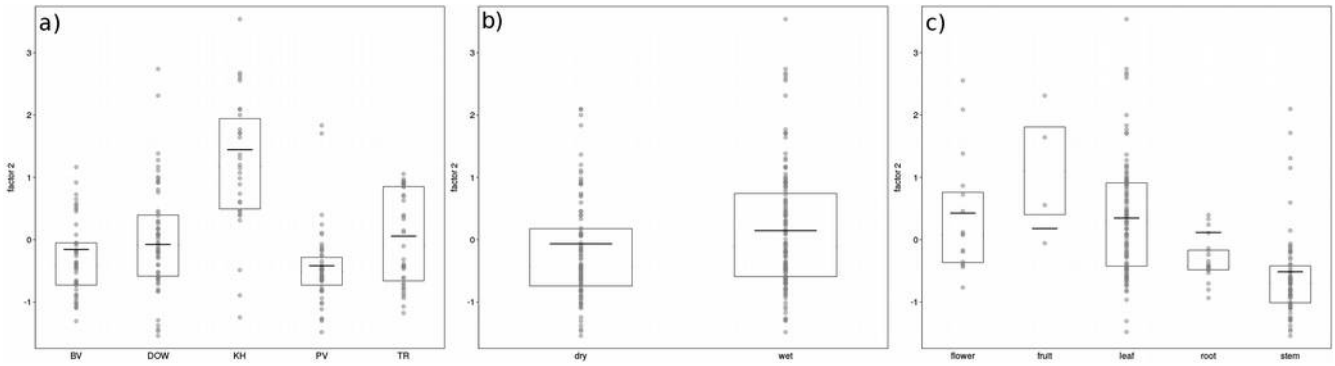


Figure 5)

Factor 2 varied significantly according to site, season and plant type. Loadings for all samples on factor 1 separated by a) site, b) season, and c) plant type.

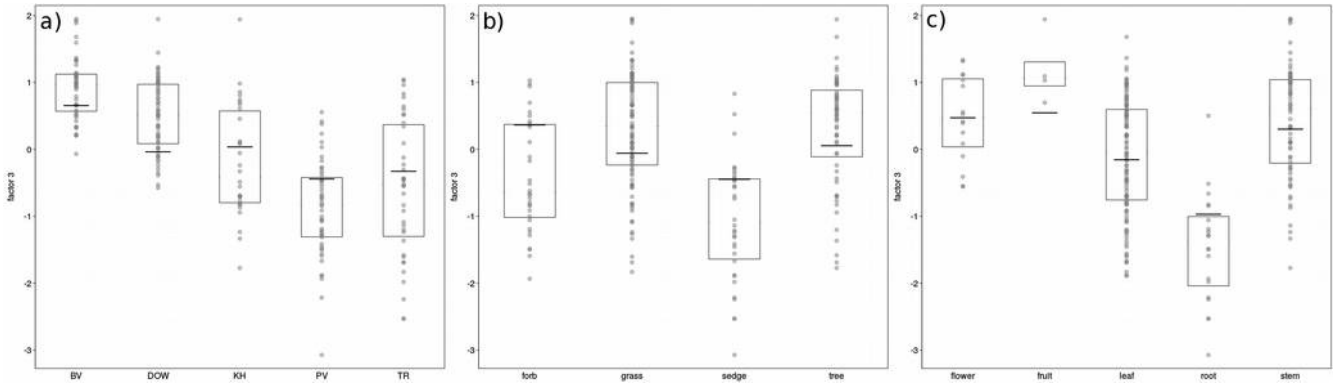
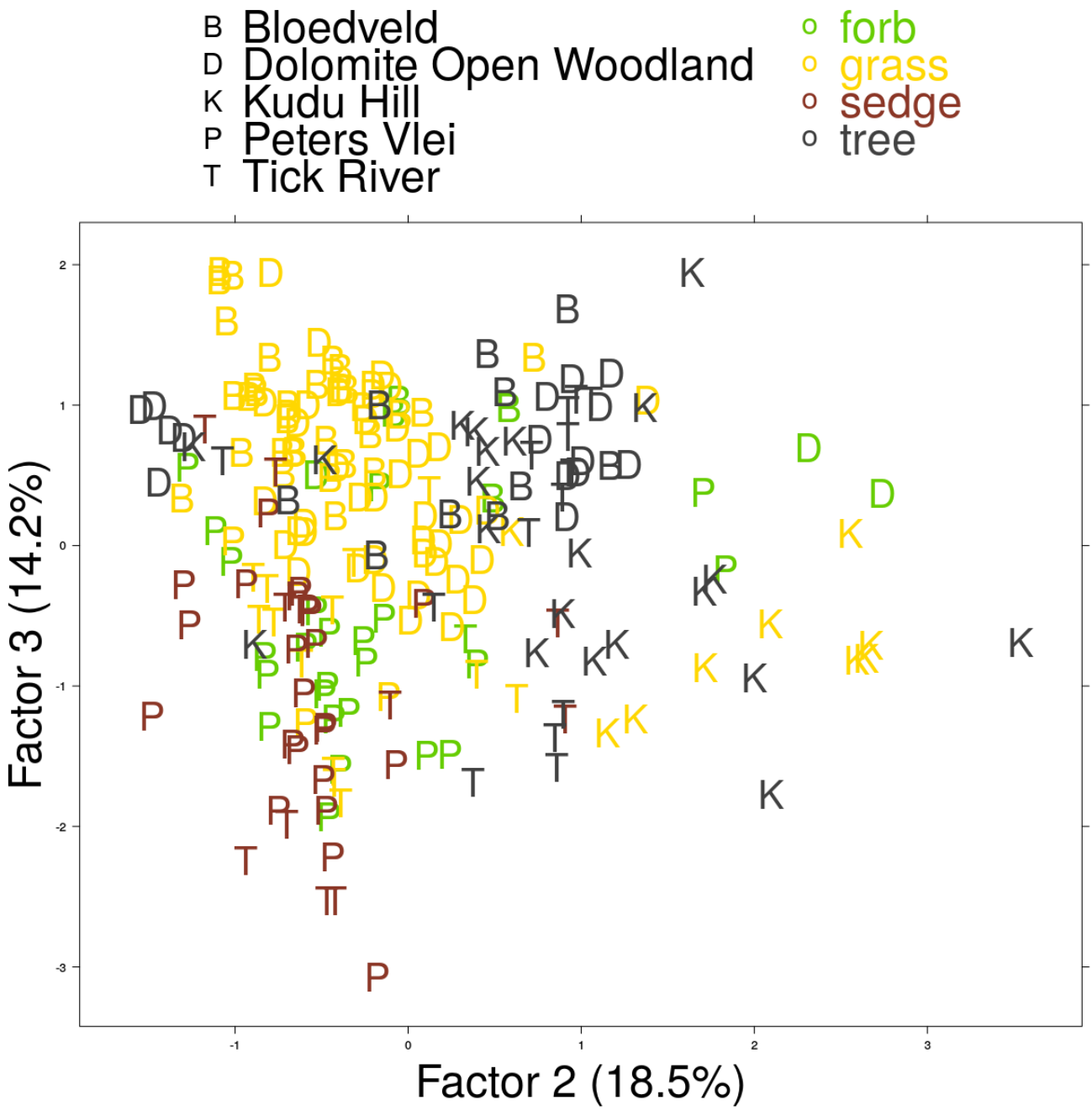


Figure 6)

Factor 2 varied significantly according to site, type and organ. Loadings for all samples on factor 1 separated by a) site, b) type, and c) organ.



950 Figure 7)

As in figure 2, each sample included in our analysis is plotted according to their loadings on factor 2 and 3. The sites are indicated by letter and the plant type by color.

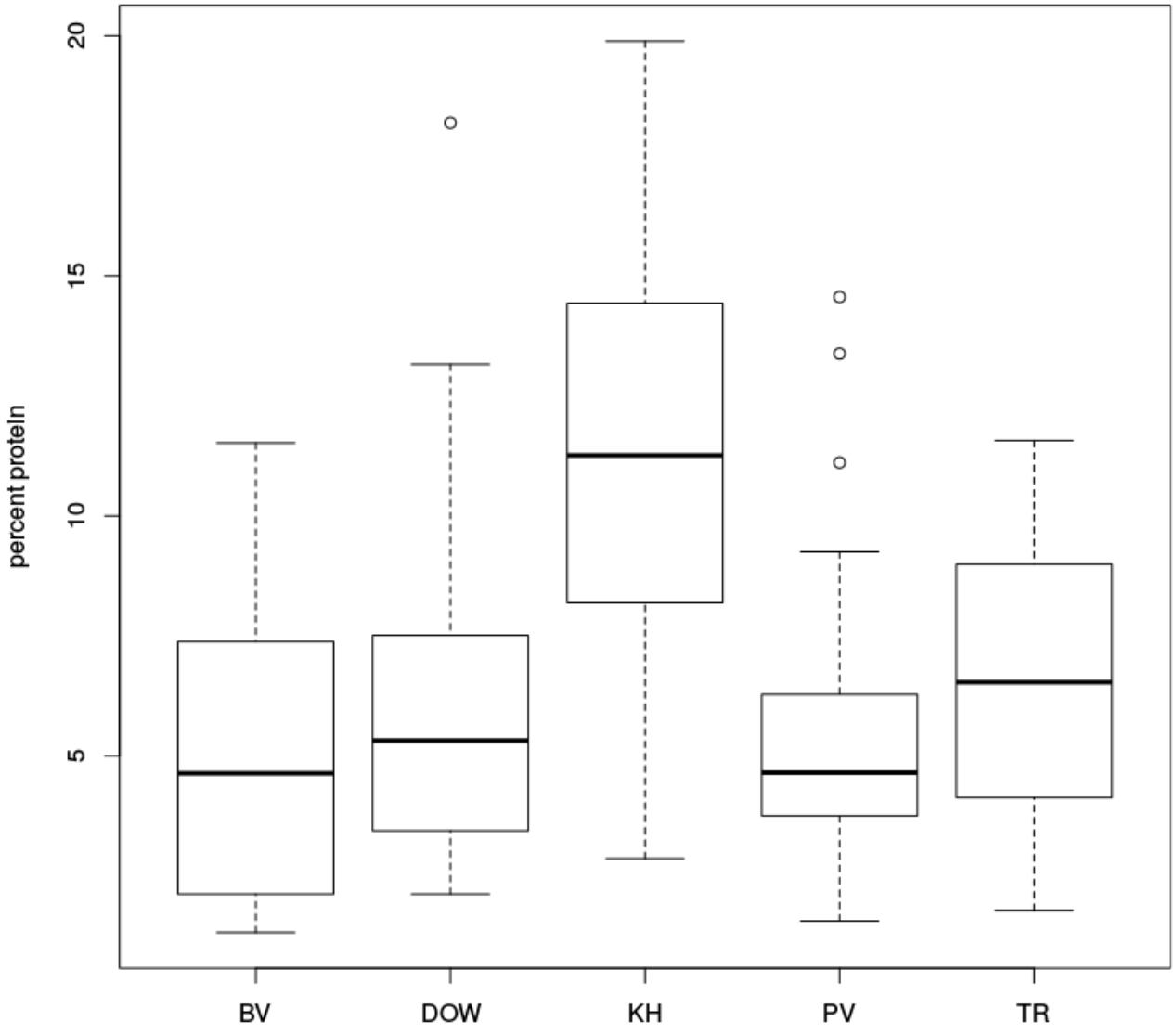
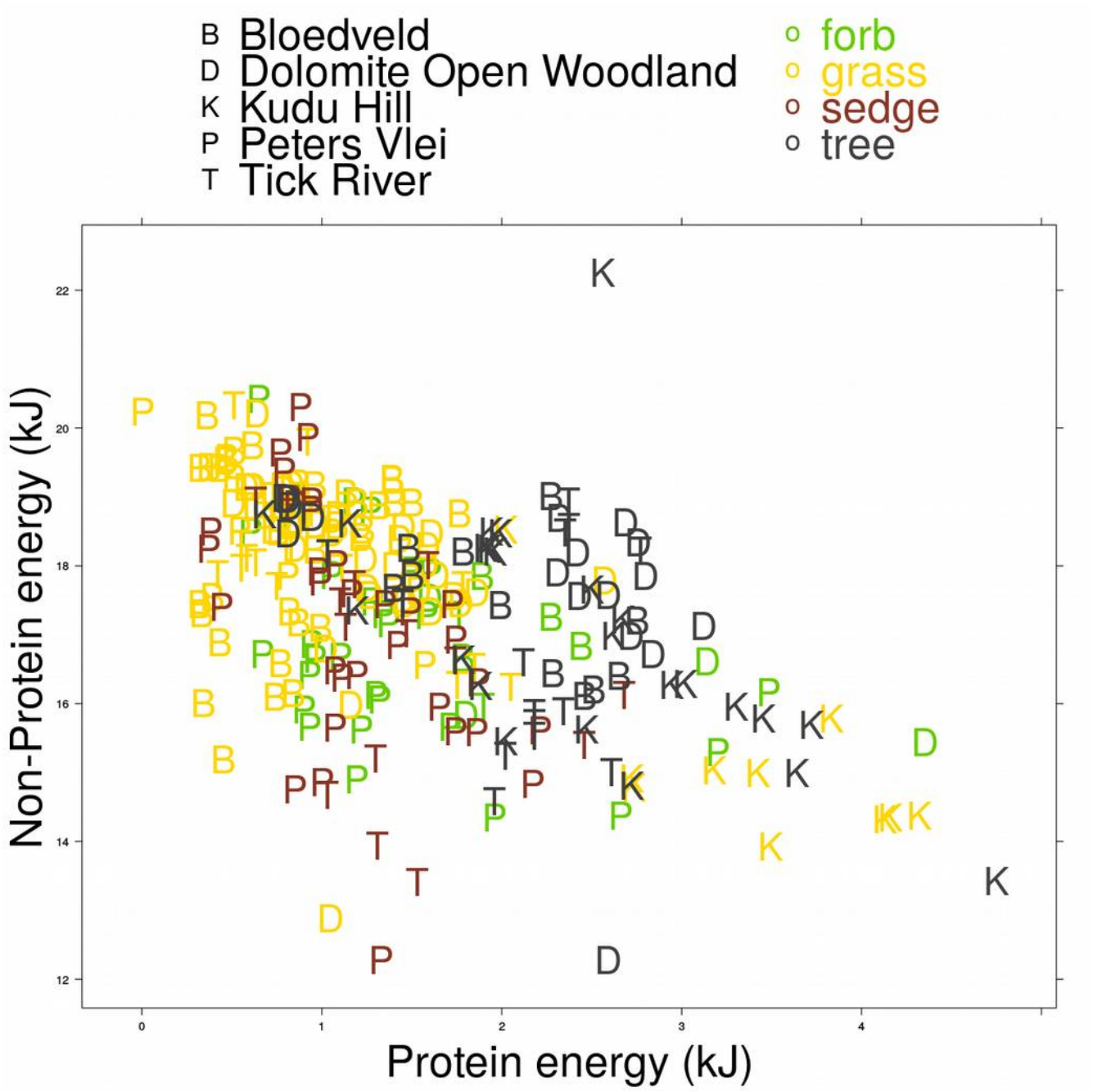


Figure 8)

The percentage of protein within all plants from each site. Kudu Hill plants have much higher protein values than other sites.



960 Figure 9)

The energetic value of plants is given as a ratio of the kilojoules derived from protein versus those derived from non-protein sources (lipids and carbohydrates). The plant types are indicated by colors and the sites are indicated by letter. Grasses generally have low protein and high non-protein energy (NPE), with the exception of several samples from Kudu Hill that show the opposite pattern.

965

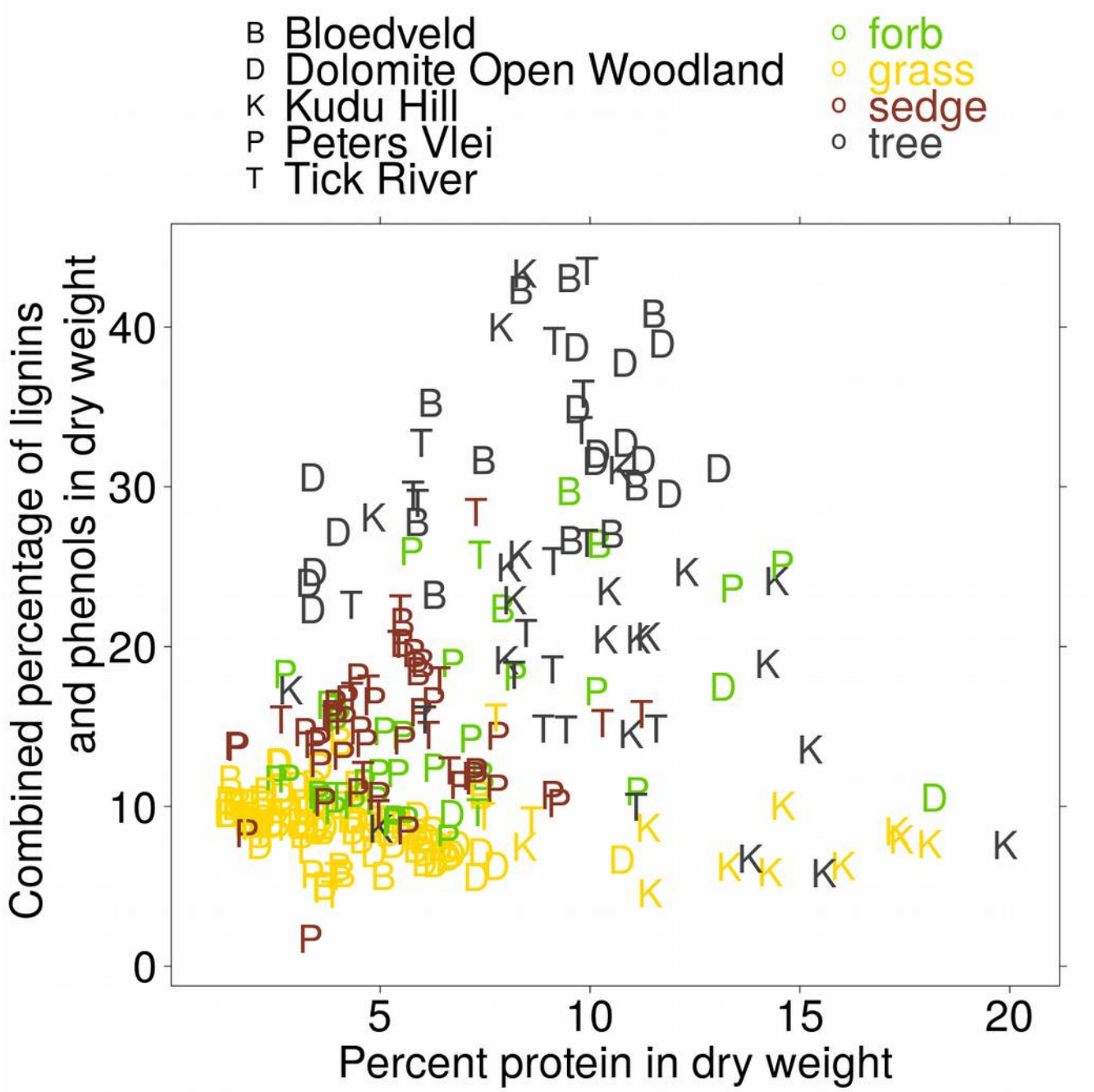


Figure 10)

The combined percentage of lignins and phenols is plotted against the percentage of protein in all samples. The plant types are indicated by colors and the sites are indicated by letter.

970

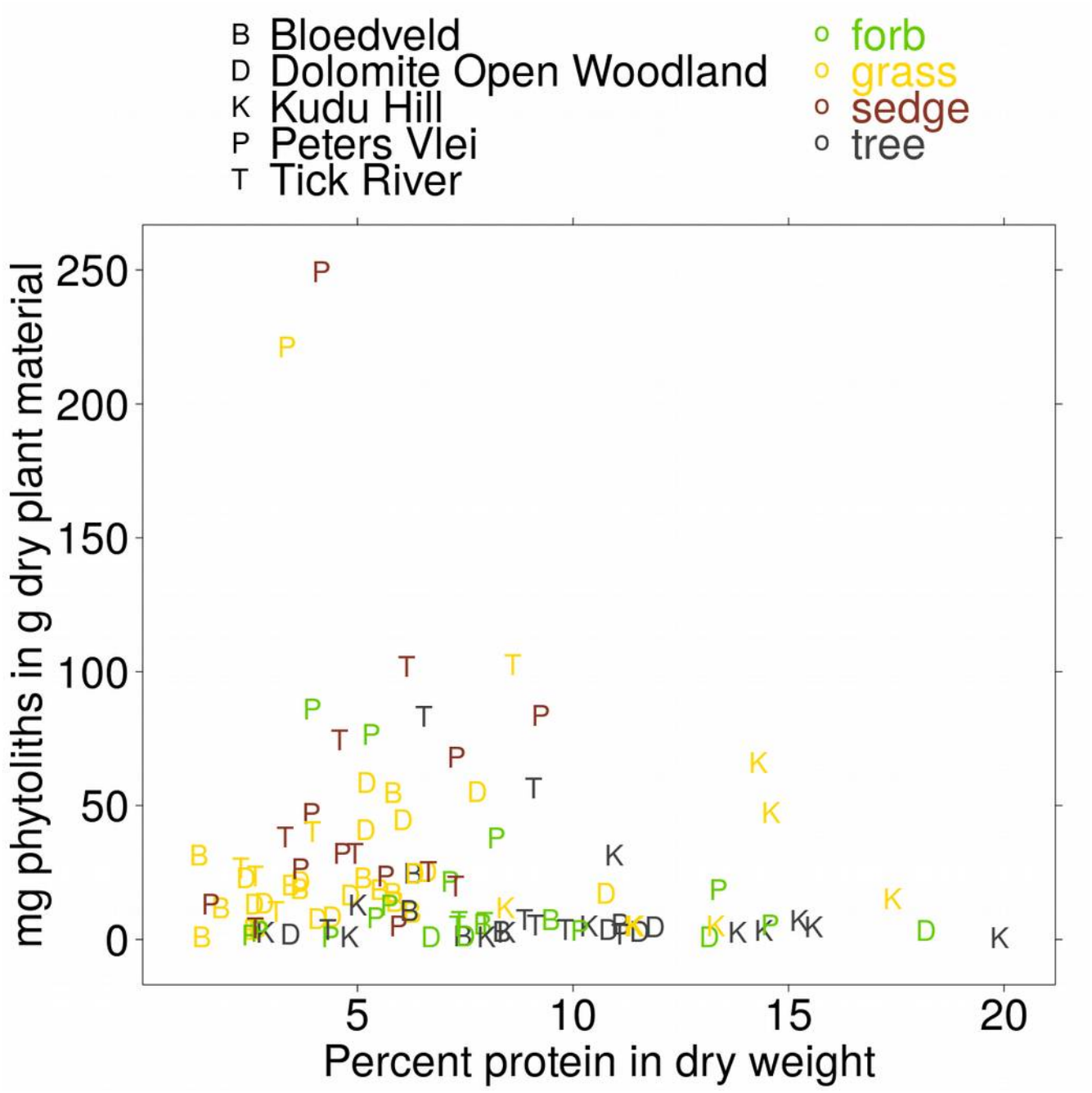


Figure 11)

The percentage of phytoliths is plotted against the percentage of protein in all samples. The plant types are indicated by colors and the sites are indicated by letter.

Supplementary Material for Henry et al. "Influences on plant nutritional variation and their potential effects on hominin diet selection"

Part 1: Protocols for Nutritional Analyses.

After sampling, the plant replicates were processed in a grinding mill (Cyclotec 1093, FOSS TECATOR AB, Höganäs, Sweden) with a 1 mm screen sieve and analyzed by Standard chemical methods according to descriptions of the VDLUFA (Naumann and Bassler, 1993) to determine nutrient content.

Samples were assayed for the contents of dry matter (DM), crude ash (Cash), crude protein (CP), crude fat (Cfa), crude fiber (Cfe), neutral detergent fiber (NDF), acid detergent fiber (ADF), and acid detergent lignin (ADL), and other constituents as described below. All measurements were performed in duplicate and chemical values are reported as percentage of dry matter. The evaluation of the total energy content is based on conventional estimates of the energy values of different nutrients. All chemicals described were used either in technical grade or in HPLC grade for the chromatographic analysis. Compounds were obtained from ROTH (Karlsruhe, Germany) unless noted otherwise.

Dry matter was determined after drying the samples at 105°C for 4 h, and crude ash after 7 hours of combustion at 600°C in a muffle furnace.

Crude fiber (Rfe) content of the different plant materials was analyzed as described in the AOAC (1984, p. 7.074-7.077, 162-163) (Naumann and Bassler, 1993) using the Fibertec™ system (Hot Extractor 1020; FOSS TECATOR AB, Höganäs, Sweden). Crude fiber was obtained from the loss in weight on incineration of dried residue remaining after the consecutive treatment of the probes with both 1.25 % sulfuric acid and 1.25% potassium under specific conditions. Samples with more than 2 % fat content were extracted with petroleum ether before crude fiber determination.

For determining the ADF, ADL and NDF content of the plant samples according to the methods of van Soest et al. (1991), a modified filter bag technique by ANKOM technology (ANKOM Technology Corporation, NY, USA, method 5 and 6 4/13/111) was utilized. For this, 0.45-0.55 g of ground probe material was weighed into nitrogen and ash free filter bags and then extracted in the ANKOM 200 fiber Analyzer. After extraction, bags were washed with hot distilled water and acetone, dried and weighed back. ADL was received by the acidic treatment (72 % H_2SO_4) of the filter bags after performing ADF determination of sample material. Hemicellulose was determined as the difference of NDF and ADF, whereas cellulose is calculated as the difference of ADF-ADL (Rinne et al., 1997)

Total nitrogen content was estimated via a standard semimicro Kjeldahl method (Thiex et al., 2002) on a Kejeltec 1030 auto analyzer (FOSS TECATOR AB, Höganäs, Schweden). All samples (0.2 g) were digested for 1.5 h in a solvent mixture consisting of 4 ml concentrated sulfuric acid (96%) and one copper containing catalyst tablet (31.5 g K_2SO_4 , 0.15 g $CuSO_4$ *$5H_2O$, FOSS, Rellingen) to oxidize the organic substance. Subsequently, the solvents were diluted with 15 ml of distilled water, distillated with a small amount of sodium hydroxid (32%) into 30 ml of 1% boric acid, and finally back titrated with 0.1 mol/l HCl. Crude protein content was calculated as N*6.25.

To measure phytolith content, a known amount of dried plant material was digested in heated nitric acid until the organic material was removed, following the wet ashing protocol in (Piperno, 2006) The remaining sample contained only the organic silica, and the weight was recorded. The difference between starting and ending weights provided the percent phytolith content.

Lipids were measured using a protocol provided by Dr. Martin Kainz, Groupleader Ecotoxicology and Lipid Biochemistry of The WasserCluster Lunz GmbH. This method is based on a chloroform and methanol extraction of small aliquots of the dried ground plant material. 50mg of the sample is placed in 2ml chloroform and kept overnight at -80 C. The following day 1ml of methanol was added, followed by 1ml of 2:1 chloroform:methanol, then 0.8ml of a 0.9% w/v sodium chloride

solution. The tubes were then sonicated for 10 minutes, vortexed, and finally centrifuged for three minutes at 3000rpm. The bottom layer contained organics and was transferred to a pre-weighed glass tube. This material was then dried under a stream of nitrogen gas, and the final weight of the tube measured. This method was faster than a traditional Soxhlet extraction. The lipid values this method provides include all of the soluble lipids within a plant, including those bound up in other more complex molecules, and therefore tends to provide a slightly higher lipid percentage than the Soxhlet extraction.

Tannins and phenols were measured using a modification of the Spectranomics Protocol for Total Phenol and Tannin Determination published by the Carnegie Institute for Science (Carnegie Institution for Science, 2011). The methodology was adapted for our samples (instead of the frozen leaf disks as described in the protocol) by using c. 20mg of the plant material which had been ground using the grinding mill (above) as the starting amount. The samples were extracted using methanol for phenols, and methanol and polyvinylpyrrolidone for tannins. They were then assayed using a Folin-Ciocalteau Phenol reagent, and compared to a gallic acid standard, using a microplate reader with a 735nm filter.

AOAC, 1984. Official methods of analysis of the Official Association of Analytical Chemists, 14th edition.

Carnegie Institution for Science, 2011. Spectranomics Protocol: Total Phenol and Tannin Determination.

Naumann, C., Bassler, R., 1993. Chemical analyses of feedstuff. VDLUFA-Press Darmstadt Ger.

Piperno, D.R., 2006. Phytoliths: A comprehensive guide for archaeologists and paleoecologists. AltaMira Press, Lanham, Maryland.

Rinne, M., Jaakkola, S., Huhtanen, P., 1997. Grass maturity effects on cattle fed silage-based diets. 1. Organic matter digestion, rumen fermentation and nitrogen utilization. Anim. Feed Sci. Technol. 67, 1–17. https://doi.org/10.1016/S0377-8401(96)01141-8

Thiex, N.J., Manson, H., Anderson, S., Persson, J.-Å., 2002. Determination of Crude Protein in Animal Feed, Forage, Grain, and Oilseeds by Using Block Digestion with a Copper Catalyst and Steam Distillation into Boric Acid: Collaborative Study. J. AOAC Int. 85, 309–317.

Van Soest, P.J., Robertson, J.B., Lewis, B.A., 1991. Methods for Dietary Fiber, Neutral Detergent Fiber, and Nonstarch Polysaccharides in Relation to Animal Nutrition. J. Dairy Sci. 74, 3583–3597. https://doi.org/10.3168/jds.S0022-0302(91)78551-2

Part 2: Individual models

Model 1: Effect of plant type, plant organ, season and habitat on Factor 1.

Full model: Factor 1 ~ plant_type + simple_organ + season + site + (1+season.code||species) + (1|plant.id)

Null Model: Factor1 ~ 1 + (1+season.code||species) + (1|plant.id)

Test of the LMM for Factor 1

	Df	AIC	BIC	LogLik	deviance	Chisq	DF	Pr(>Chisq)
null.f1	5	345.132	362.213	-167.566	355.132	NA	NA	NA
full.f1	17	274.800	332.874	-120.400	240.800	94.332	12	0.000

Influence of each of the fixed effects on the LMM for Factor 1

	Df	AIC	LRT	Pr(Chi)
<none>	NA	274.800	NA	NA
plant_type	3	299.766	30.996	0.000
simple_organ	4	313.475	46.674	0.000
season	1	275.026	2.226	0.136
site	4	282.490	15.690	0.003

Coefficients of each of the states of each fixed effects in the model for Factor 1.

	Estimate	Std. Error	t value
(Intercept)	0.576	0.253	2.276
plant_typegrass	-1.221	0.285	-4.284
plant_typesedge	-0.026	0.309	-0.083
plant_typetree	0.734	0.285	2.577
simple_organfruit	-0.115	0.263	-0.438
simple_organleaf	0.134	0.101	1.330
simple_organroot	0.029	0.142	0.203
simple_organstem	-0.280	0.101	-2.781
Seasonwet	0.107	0.068	1.570
siteDOW	-0.189	0.131	-1.446
siteKH	-1.029	0.238	-4.316
sitePV	-0.210	0.225	-0.931
siteTR	-0.428	0.162	-2.636

Model 2: Effect of plant type, plant organ, season and habitat on Factor 2.

Full model: Factor 2 ~ plant_type + simple_organ + season + site + (1+season.code||species) + (1| plant.id)

Null Model: Factor 2 ~ 1 + (1+season.code||species) + (1|plant.id)

Test of the LMM for Factor 2

	Df	AIC	BIC	LogLik	deviance	Chisq	DF	Pr(>Chisq)
null.f2	5	502.660	519.740	-246.330	492.660	NA	NA	NA
full.f2	17	369.718	427.792	-167.859	335.718	156.942	12	0

Influence of each of the fixed effects on the LMM for Factor 2

	Df	AIC	LRT	Pr(Chi)
<none>	NA	369.718	NA	NA
plant_type	3	369.660	5.942	0.114
simple_organ	4	479.175	117.457	0.000
season	1	376.177	8.459	0.004
site	4	393.965	32.247	0.000

Coefficients of each of the states of each fixed effects in the model for Factor 2.

	Estimate	Std. Error	t value
(Intercept)	0.443	0.299	1.480
plant_typegrass	-0.336	0.324	-1.039
plant_typesedge	-0.875	0.346	-2.529
plant_typetree	-0.277	0.323	-0.856
simple_organfruit	-0.247	0.326	-0.756
simple_organleaf	-0.079	0.127	-0.623
simple_organroot	-0.310	0.180	-1.727
simple_organstem	-0.940	0.128	-7.364
Seasonwet	0.211	0.068	3.093
siteDOW	0.081	0.167	0.482
siteKH	1.600	0.281	5.698
sitePV	-0.262	0.275	-0.953
siteTR	0.214	0.205	1.045

Model 3: Effect of plant type, plant organ, season and habitat on Factor 3.

Full model: Factor 3 ~ plant_type + simple_organ + season + site + (1+season.code||species) + (1| plant.id)

Null Model: Factor 3 ~ 1 + (1+season.code||species) + (1|plant.id)

Test of the LMM for Factor 3

	Df	AIC	BIC	LogLik	deviance	Chisq	DF	Pr(>Chisq)
null.f2	5	436.335	453.415	-213.167	426.335	NA	NA	NA
full.f2	17	327.648	385.721	-146.824	293.648	132.687	12	0.000

Influence of each of the fixed effects on the LMM for Factor 3

	Df	AIC	LRT	Pr(Chi)
<none>	NA	327.648	NA	NA
plant_type	3	327.329	5.682	0.128
simple_organ	4	419.008	99.360	0.000
season	1	325.661	0.013	0.910
site	4	347.305	27.658	0.000

Coefficients of each of the states of each fixed effects in the model for Factor 3.

	Estimate	Std. Error	t value
(Intercept)	1.566	0.278	5.642
plant_typegrass	-0.393	0.306	-1.282
plant_typesedge	-0.816	0.332	-2.457
plant_typetree	-0.355	0.309	-1.152
simple_organfruit	0.080	0.295	0.272
simple_organleaf	-0.622	0.114	-5.470
simple_organroot	-1.431	0.160	-8.926
simple_organstem	-0.166	0.114	-1.464
Seasonwet	-0.009	0.080	-0.114
siteDOW	-0.687	0.146	-4.713
siteKH	-0.648	0.262	-2.473
sitePV	-1.124	0.249	-4.513
siteTR	-0.998	0.181	-5.514

Model 4: Effect of plant type, plant organ, season and habitat on the absolute residuals from Model 1

Full model : Absolute residuals from model 1 ~ plant_type + simple_organ + Season + site + (1+season.code||species) + (1|plant.id)

Influence of each of the fixed effects on the LMM for the absolute residuals from Model 1

	Df	AIC	LRT	Pr(Chi)
<none>	NA	-67.604	NA	NA
plant_type	3	-64.047	9.557	0.023
simple_organ	4	-72.387	3.217	0.522
season	1	-69.457	0.147	0.701
site	4	-72.502	3.101	0.541

Coefficients of each of the states of each fixed effects in the model for the absolute residuals from Model 1.

	Estimate	Std. Error	t value
(Intercept)	0.282	0.085	3.324
plant_typegrass	-0.021	0.068	-0.315
plant_typesedge	0.169	0.063	2.707
plant_typetree	0.043	0.072	0.593
simple_organfruit	-0.021	0.119	-0.176
simple_organleaf	-0.089	0.053	-1.692
simple_organroot	-0.091	0.071	-1.279
simple_organstem	-0.078	0.053	-1.474
Seasonwet	0.014	0.035	0.384
siteDOW	-0.001	0.042	-0.017
siteKH	0.064	0.058	1.102
sitePV	-0.036	0.067	-0.534
siteTR	-0.041	0.051	-0.792

Model 5: Effect of plant type, plant organ, season and habitat on the absolute residuals from Model 2

Full model : Absolute residuals from model 2 ~ plant_type + simple_organ + Season + site + (1+season.code||species) + (1|plant.id)

Influence of each of the fixed effects on the LMM for the absolute residuals from Model 2

	Df	AIC	LRT	Pr(Chi)
<none>	NA	15.611	NA	NA
plant_type	3	15.321	5.710	0.127
simple_organ	4	16.660	9.049	0.060
Season	1	14.288	0.677	0.410
site	4	10.467	2.857	0.582

Coefficients of each of the states of each fixed effects in the model for the absolute residuals from Model 2.

	Estimate	Std. Error	t value
(Intercept)	0.539	0.118	4.568
plant_typegrass	-0.250	0.104	-2.393
plant_typesedge	-0.075	0.109	-0.686
plant_typetree	-0.119	0.109	-1.087
simple_organfruit	-0.306	0.140	-2.182
simple_organleaf	-0.114	0.060	-1.905
simple_organroot	-0.152	0.082	-1.850
simple_organstem	-0.047	0.060	-0.786
Seasonwet	-0.038	0.044	-0.854
siteDOW	0.072	0.079	0.910
siteKH	0.114	0.101	1.126
sitePV	0.003	0.107	0.025
siteTR	-0.058	0.093	-0.632

Model 6: Effect of plant type, plant organ, season and habitat on the absolute residuals from Model 3

Full model : Absolute residuals from model 3 ~ plant_type + simple_organ + Season + site + (1+season.code||species) + (1|plant.id)

Influence of each of the fixed effects on the LMM for the absolute residuals from Model 3.

	Df	AIC	LRT	Pr(Chi)
<none>	NA	-38.726	NA	NA
plant_type	3	-42.643	2.083	0.555
simple_organ	4	-36.056	10.671	0.031
Season	1	-38.639	2.088	0.149
site	4	-37.741	8.986	0.061

Coefficients of each of the states of each fixed effects in the model for the absolute residuals from Model 3.

	Estimate	Std. Error	t value
(Intercept)	0.202	0.100	2.031
plant_typegrass	0.082	0.081	1.015
plant_typesedge	0.033	0.085	0.390
plant_typetree	0.123	0.086	1.430
simple_organfruit	-0.061	0.122	-0.500
simple_organleaf	-0.105	0.053	-1.982
simple_organroot	0.036	0.073	0.497
simple_organstem	-0.040	0.053	-0.751
Seasonwet	-0.067	0.046	-1.469
siteDOW	0.059	0.068	0.864
siteKH	0.231	0.078	2.968
sitePV	0.122	0.086	1.419
siteTR	0.105	0.077	1.356

Model 7: Effect of plant type, plant organ, season and habitat on percent tannin
Full model: log.percent_tannin ~ plant_type + simple_organ + season + site + (1+season.code||species) + (1|plant.id)

Null Model: log.percent_tannin ~ 1 + (1+season.code||species) + (1|plant.id)

Chi Square Test of the LMM for percent tannin

	Chi	Df	Pr(>Chisq)
Full model	157.961	12	0.000

Influence of each of the fixed effects on the LMM for percent tannin

	Df	AIC	LRT	Pr(Chi)
<none>	NA	259.033	NA	NA
plant_type	3	289.880	36.847	0.000
simple_organ	4	355.057	104.024	0.000
season	1	262.478	5.445	0.020
site	4	259.365	8.332	0.080

Coefficients of each of the states of each fixed effects in the model for percent tannin.

	Estimate	Std. Error	t value
(Intercept)	1.209	0.224	5.408
plant_typegrass	-1.194	0.242	-4.926
plant_typesedge	-0.061	0.260	-0.236
plant_typetree	0.710	0.242	2.937
simple_organfruit	-0.192	0.247	-0.779
simple_organleaf	0.066	0.096	0.692
simple_organroot	0.104	0.122	0.848
simple_organstem	-0.546	0.096	-5.671
Seasonwet	0.127	0.050	2.555
siteDOW	-0.149	0.123	-1.210
siteKH	-0.566	0.211	-2.681
sitePV	-0.195	0.198	-0.986
siteTR	-0.331	0.151	-2.189

Model 8: Effect of plant type, plant organ, season and habitat on percent phenol
Full model: log.percent_phenol ~ plant_type + simple_organ + season + site + (1+season.code||species) + (1|plant.id)

Null Model: log.percent_phenol ~ 1 + (1+season.code||species) + (1|plant.id)

Chi Square Test of the LMM for percent phenol

	Chi	Df	Pr(>Chisq)
Full model	144.054	12	0.000

Influence of each of the fixed effects on the LMM for percent phenol

	Df	AIC	LRT	Pr(Chi)
<none>	NA	229.007	NA	NA
plant_type	3	259.863	36.856	0.000
simple_organ	4	309.194	88.187	0.000
season	1	234.048	7.041	0.008
site	4	229.388	8.382	0.079

Coefficients of each of the states of each fixed effects in the model for percent phenol.

	Estimate	Std. Error	t value
(Intercept)	1.438	0.201	7.138
plant_typegrass	-1.075	0.210	-5.120
plant_typesedge	-0.210	0.224	-0.937
plant_typetree	0.561	0.214	2.627
simple_organfruit	-0.157	0.229	-0.684
simple_organleaf	0.113	0.090	1.256
simple_organroot	0.144	0.115	1.251
simple_organstem	-0.414	0.091	-4.572
Seasonwet	0.167	0.057	2.952
siteDOW	-0.171	0.113	-1.513
siteKH	-0.467	0.190	-2.450
sitePV	-0.331	0.179	-1.848
siteTR	-0.357	0.138	-2.581

Model 9: Effect of plant type, plant organ, season and habitat on percent lignin
Full model: sqrt.true_lignin ~ plant_type + simple_organ + season + site + (1+season.code||species) + (1|plant.id)

Null Model: sqrt.true_lignin~ 1 + (1+season.code||species) + (1|plant.id)

Chi Square Test of the LMM for percent lignin

	Chi	Df	Pr(>Chisq)
Full model	37.847	12	0.000

Influence of each of the fixed effects on the LMM for percent lignin

	Df	AIC	LRT	Pr(Chi)
<none>	NA	531.973	NA	NA
plant_type	3	551.408	25.435	0.000
simple_organ	4	529.503	5.530	0.237
season	1	532.231	2.258	0.133
site	4	537.651	13.678	0.008

Coefficients of each of the states of each fixed effects in the model for percent lignin.

	Estimate	Std. Error	t value
(Intercept)	3.822	0.328	11.670
plant_typegrass	-0.674	0.311	-2.166
plant_typesedge	-0.003	0.327	-0.010
plant_typetree	1.094	0.319	3.431
simple_organfruit	-0.097	0.397	-0.246
simple_organleaf	-0.183	0.163	-1.122
simple_organroot	0.006	0.207	0.029
simple_organstem	0.006	0.165	0.035
Seasonwet	-0.127	0.083	-1.536
siteDOW	-0.312	0.198	-1.577
siteKH	-1.106	0.297	-3.719
sitePV	-0.464	0.295	-1.572
siteTR	-0.386	0.238	-1.617

Model 10: Effect of plant type, plant organ, season and habitat on percent cellulose
Full model: true_NDF ~ plant_type + simple_organ + season + site + (1+season.code||species) + (1|plant.id)

Null Model: true_NDF ~ 1 + (1+season.code||species) + (1|plant.id)

Chi Square Test of the LMM for percent cellulose

	Chi	Df	Pr(>Chisq)
Full model	150.316	12	0.000

Influence of each of the fixed effects on the LMM for percent cellulose

	Df	AIC	LRT	Pr(Chi)
<none>	NA	1743.766	NA	NA
plant_type	3	1781.881	44.115	0.000
simple_organ	4	1830.526	94.759	0.000
season	1	1741.880	0.114	0.736
site	4	1748.388	12.622	0.013

Coefficients of each of the states of each fixed effects in the model for percent cellulose.

	Estimate	Std. Error	t value
(Intercept)	21.152	2.691	7.861
plant_typegrass	13.667	1.966	6.950
plant_typesedge	3.830	1.641	2.334
plant_typetree	-2.015	2.090	-0.964
simple_organfruit	-5.818	3.926	-1.482
simple_organleaf	1.863	1.792	1.040
simple_organroot	-6.255	2.203	-2.840
simple_organstem	9.477	1.825	5.193
Seasonwet	-0.355	0.996	-0.356
siteDOW	-0.412	1.372	-0.300
siteKH	-0.584	1.765	-0.331
sitePV	5.895	1.982	2.975
siteTR	-0.970	1.621	-0.598

Model 11: Effect of plant type, plant organ, season and habitat on percent hemicellulose
Full model: true_ADF ~ plant_type + simple_organ + season + site + (1+season.code||species) + (1|plant.id)

Null Model: true_ADF ~ 1 + (1+season.code||species) + (1|plant.id)

Chi Square Test of the LMM for percent hemicellulose

	Chi	Df	Pr(>Chisq)
Full model	72.379	12	0.000

Influence of each of the fixed effects on the LMM for percent hemicellulose

	Df	AIC	LRT	Pr(Chi)
<none>	NA	1662.648	NA	NA
plant_type	3	1691.719	35.071	0.000
simple_organ	4	1684.519	29.871	0.000
season	1	1662.198	1.550	0.213
site	4	1656.103	1.455	0.835

Coefficients of each of the states of each fixed effects in the model for percent hemicellulose.

	Estimate	Std. Error	t value
(Intercept)	16.233	2.856	5.684
plant_typegrass	12.318	2.629	4.685
plant_typesedge	8.470	2.742	3.089
plant_typetree	-4.720	2.801	-1.685
simple_organfruit	-9.013	3.523	-2.558
simple_organleaf	-1.543	1.396	-1.105
simple_organroot	-5.983	1.767	-3.386
simple_organstem	0.917	1.418	0.647
Seasonwet	1.470	1.138	1.292
siteDOW	0.860	1.759	0.489
siteKH	2.069	2.607	0.794
sitePV	2.531	2.581	0.981
siteTR	2.441	2.108	1.158

Model 12: Effect of plant type, plant organ, season and habitat on percent protein
Full model: true_protein ~ plant_type + simple_organ + season + site + (1+season.code||species) + (1|plant.id)

Null Model: true_protein ~ 1 + (1+season.code||species) + (1|plant.id)

Chi Square Test of the LMM for percent protein

	Chi	Df	Pr(>Chisq)
Full model	219.394	12	0.000

Influence of each of the fixed effects on the LMM for percent protein

	Df	AIC	LRT	Pr(Chi)
<none>	NA	1054.596	NA	NA
plant_type	3	1055.554	6.957	0.073
simple_organ	4	1222.509	175.912	0.000
season	1	1060.202	7.605	0.006
site	4	1085.829	39.233	0.000

Coefficients of each of the states of each fixed effects in the model for percent protein.

	Estimate	Std. Error	t value
(Intercept)	9.200	1.078	8.535
plant_typegrass	-3.048	1.228	-2.483
plant_typesedge	-2.630	1.328	-1.981
plant_typetree	-1.317	1.223	-1.076
simple_organfruit	-3.537	1.128	-3.137
simple_organleaf	-0.602	0.420	-1.435
simple_organroot	-0.596	0.538	-1.108
simple_organstem	-4.141	0.426	-9.723
Seasonwet	0.793	0.276	2.870
siteDOW	0.529	0.553	0.956
siteKH	6.731	1.032	6.519
sitePV	-1.040	0.960	-1.083
siteTR	0.163	0.691	0.235

Model 13: Effect of plant type, plant organ, season and habitat on percent lipid
Full model: log.percent_lipid_dry ~ plant_type + simple_organ + season + site + (1+season.code|| species) + (1|plant.id)

Null Model: log.percent_lipid_dry ~ 1 + (1+season.code||species) + (1|plant.id)

Chi Square Test of the LMM for percent lipid

	Chi	Df	Pr(>Chisq)
Full model	177.700	12	0.000

Influence of each of the fixed effects on the LMM for percent lipid

	Df	AIC	LRT	Pr(Chi)
<none>	NA	179.012	NA	NA
plant_type	3	196.059	23.047	0.000
simple_organ	4	303.443	132.431	0.000
season	1	178.681	1.669	0.196
site	4	176.385	5.374	0.251

Coefficients of each of the states of each fixed effects in the model for percent lipid.

	Estimate	Std. Error	t value
(Intercept)	1.208	0.155	7.805
plant_typegrass	0.082	0.133	0.613
plant_typesedge	-0.182	0.138	-1.322
plant_typetree	0.585	0.141	4.160
simple_organfruit	0.339	0.200	1.695
simple_organleaf	-0.028	0.087	-0.325
simple_organroot	-0.193	0.119	-1.627
simple_organstem	-0.640	0.088	-7.295
Seasonwet	0.069	0.050	1.368
siteDOW	-0.127	0.089	-1.425
siteKH	0.154	0.126	1.224
sitePV	-0.016	0.131	-0.124
siteTR	-0.027	0.108	-0.250

Model 14: Effect of plant type, plant organ, season and habitat on calculated gross energy
Full model: calc_gross_energy ~ plant_type + simple_organ + season + site + (1+season.code||species) + (1|plant.id)

Null Model: calc_gross_energy ~ 1 + (1+season.code||species) + (1|plant.id)

Chi Square Test of the LMM for calculated gross energy

	Chi	Df	Pr(>Chisq)
Full model	72.681	12	0.000

Influence of each of the fixed effects on the LMM for calculated gross energy

	Df	AIC	LRT	Pr(Chi)
<none>	NA	785.691	NA	NA
plant_type	3	783.658	3.967	0.265
simple_organ	4	842.568	64.877	0.000
season	1	783.855	0.164	0.686
site	4	780.124	2.433	0.657

Coefficients of each of the states of each fixed effects in the model for calculated gross energy

	Estimate	Std. Error	t value
(Intercept)	19.908	0.471	42.297
plant_typegrass	-0.614	0.373	-1.646
plant_typesedge	-0.380	0.360	-1.057
plant_typetree	-0.215	0.404	-0.534
simple_organfruit	0.328	0.648	0.506
simple_organleaf	-0.701	0.262	-2.670
simple_organroot	-2.344	0.328	-7.149
simple_organstem	-0.383	0.266	-1.437
Seasonwet	0.096	0.237	0.405
siteDOW	0.327	0.297	1.102
siteKH	0.133	0.368	0.361
sitePV	0.143	0.394	0.363
siteTR	-0.116	0.342	-0.338

Model 15: Effect of plant type, plant organ, season and habitat on percent ash
Full model: log.true_ash ~ plant_type + simple_organ + season + site + (1+season.code||species) + (1|plant.id)

Null Model: log.true_ash ~ 1 + (1+season.code||species) + (1|plant.id)

Chi Square Test of the LMM for percent ash

	Chi	Df	Pr(>Chisq)
Full model	167.034	12	0.000

Influence of each of the fixed effects on the LMM for percent ash

	Df	AIC	LRT	Pr(Chi)
<none>	NA	74.397	NA	NA
plant_type	3	69.964	1.567	0.667
simple_organ	4	202.579	136.181	0.000
season	1	75.511	3.113	0.078
site	4	96.645	30.248	0.000

Coefficients of each of the states of each fixed effects in the model for percent ash.

	Estimate	Std. Error	t value
(Intercept)	1.347	0.152	8.859
plant_typegrass	-0.147	0.165	-0.889
plant_typesedge	0.073	0.179	0.408
plant_typetree	-0.006	0.168	-0.035
simple_organfruit	-0.130	0.167	-0.779
simple_organleaf	0.355	0.063	5.628
simple_organroot	0.501	0.081	6.215
simple_organstem	-0.068	0.064	-1.055
Seasonwet	0.089	0.047	1.905
siteDOW	0.386	0.081	4.740
siteKH	0.608	0.148	4.120
sitePV	0.473	0.133	3.551
siteTR	0.612	0.101	6.055

Model 16: Effect of plant type, plant organ, season and habitat on calculated metabolizable energy

Full model: sqrt.calc_met_energy ~ plant_type + simple_organ + season + site + (1+season.code|| species) + (1|plant.id)

Null Model: sqrt.calc_met_energy ~ 1 + (1+season.code||species) + (1|plant.id)

Chi Square Test of the LMM for calculated metabolizable energy

	Chi	Df	Pr(>Chisq)
Full model	24.729	12	0.016

Influence of each of the fixed effects on the LMM for calculated metabolizable energy.

	Df	AIC	LRT	Pr(Chi)
<none>	NA	315.997	NA	NA
plant_type	3	313.709	3.712	0.294
simple_organ	4	327.585	19.587	0.001
season	1	314.002	0.004	0.947
site	4	308.545	0.547	0.969

Coefficients of each of the states of each fixed effects in the model for calculated metabolizable energy.

	Estimate	Std. Error	t value
(Intercept)	2.031	0.184	11.007
plant_typegrass	-0.219	0.142	-1.546
plant_typesedge	-0.153	0.135	-1.132
plant_typetree	0.055	0.152	0.363
simple_organfruit	0.646	0.254	2.542
simple_organleaf	-0.198	0.109	-1.822
simple_organroot	-0.361	0.135	-2.668
simple_organstem	-0.215	0.111	-1.944
Seasonwet	0.006	0.082	0.069
siteDOW	-0.075	0.113	-0.663
siteKH	-0.007	0.138	-0.047
sitePV	-0.051	0.150	-0.340
siteTR	-0.021	0.131	-0.158

Model 17: Effect of plant type, plant organ, season and habitat on percent phytolith
Full model: log.percent_phyt_dry ~ plant_type + simple_organ + season + site + (1+season.code|| species) + (1|plant.id)

Null Model: log.percent_phyt_dry ~ 1 + (1+season.code||species) + (1|plant.id)

Chi Square Test of the LMM for percent phytolith

	Chi	Df	Pr(>Chisq)
Full model	121.706	12	0.000

Influence of each of the fixed effects on the LMM for percent phytolith

	Df	AIC	LRT	Pr(Chi)
<none>	NA	646.677	NA	NA
plant_type	3	672.205	31.529	0.000
simple_organ	4	705.333	66.656	0.000
season	1	650.272	5.596	0.018
site	4	652.178	13.501	0.009

Coefficients of each of the states of each fixed effects in the model for percent phytolith.

	Estimate	Std. Error	t value
(Intercept)	-0.19	0.325	-0.584
plant_typegrass	1.56	0.283	5.517
plant_typesedge	1.348	0.311	4.336
plant_typetree	-0.041	0.284	-0.145
simple_organfruit	-1.411	0.296	-4.763
simple_organleaf	-0.129	0.192	-0.668
simple_organroot	0.024	0.215	0.112
simple_organstem	-0.983	0.197	-4.983
Seasonwet	-0.309	0.129	-2.388
siteDOW	-0.31	0.243	-1.279
siteKH	-0.476	0.337	-1.414
sitePV	0.45	0.332	1.356
siteTR	0.418	0.275	1.517