A Holistic Approach to Examining Ancient Agriculture

A Case Study from the Bronze and Iron Age Near East

Alexia Smith and Natalie D. Munro

Department of Anthropology, University of Connecticut, Beach Hall U-2176, 354 Mansfield Road, Storrs, Connecticut 06269, U.S.A. (alexia.smith@uconn.edu). 3 V 09

CA+ Online-Only Material: Supplement A and Data File

Agriculture provided the foundation for the development and sustenance of Bronze and Iron Age civilizations in the Near East, yet remarkably little is known about how its practice varied across the region at this time. Archaeobotany and zooarchaeology have been used independently to study ancient agriculture, but there is a dire need for a more comprehensive and holistic approach, one that integrates the two data sets and better represents the reality of food production. Correspondence analysis can be an effective tool for quantitatively integrating regional Bronze and Iron Age plant and animal data spanning Syria and Jordan. Distinct regional patterns of food production and wild resource use are evident. The main variable driving this trend is available moisture. Theoretically, the method outlined here can be applied to any region and time period.

During the Bronze and Iron Ages, agricultural production provided the foundation for the development and maintenance of state-level societies in the Near East. Agriculture is also often cited in theories of collapse (e.g., Dalferes, Kukla, and Weiss 1997), yet despite this relationship with enormous social change, very little is known about regional variation in agricultural practices at this time. What we do know comes from independent analyses of zooarchaeological and archaeobotanical data sets (e.g., van Zeist 2003b; Zeder 1994). With the exception of a handful of studies (e.g., Grigson 2000; Miller 1996; Zeder 1994), zooarchaeologists and archaeobotanists have worked independently, and their research represents separate, albeit complementary, pursuits. Much can be gained by integrating the two data sets, particularly when investigating ancient subsistence strategies. By its very nature, agriculture integrates plants and animals in ways that render it more than the sum of two practices, particularly during the Bronze Age, when agriculture was well established. Consequently, the study of plants and animals in isolation does not reflect the true nature of agricultural practice, and alternative approaches are necessary to expose the dynamic ways in which people structured the production, gathering, and consumption of both animal and plant foods.

Our main goals for this research were to investigate methods for integrating plant and animal data quantitatively and to examine regional patterns in the Near Eastern data set. First, we surveyed the literature and compiled a list of Bronze and Iron Age sites—spanning Syria, Lebanon, Israel, and Jordan—that yielded both archaeobotanical and zooarchaeological reports. Second, we evaluated the reports to determine the best way to combine the data and enhance between-site comparisons. Third, we used correspondence analysis to assess whether regional patterns could be detected in the combined plant and animal data set. Correspondence analysis, a multivariate statistical technique, has been used with great success to examine modern plant and animal species data as well as archaeobotanical data (e.g., Bogaard 2004; Jones et al. 2000; Riehl 1999; Smith 2005), but it has not been used to analyze archaeobotanical and zooarchaeological records together.

Integrating Regional Ecological Data: Methods

Data Selection and Integration

A researcher integrating data from multiple investigators must consider a number of factors. Naturally, the research focus of each excavation and investigator differs, which results in differential sampling, recovery, and treatment of certain economic or wild taxa. Researchers also use different approaches when quantifying remains, an issue that has received much attention in the literature (e.g., Grayson 1984; Hastorf and Popper 1988). Although problems exist with published data, they do not prohibit our understanding of regional agricultural economies and environmental adaptations.

Sites were selected for analysis only if they met four criteria. First, both zooarchaeological and archaeobotanical reports had to be available for the site. Only published reports were considered, so that all primary data would be accessible to any researcher. In instances where the same data were published multiple times, the most recent report of each sample was selected. Second, reports had to provide lists of plants or animal taxa with their associated seed count or number of identified specimens (NISP). Reports that provided only summary or percentage data were eliminated. Third, the plant and animal data had to originate from the same chronological period. In some instances, we broadened the age category of remains (e.g., from Early Bronze II to Early Bronze) to retain data. Fourth, specialized reports that focused on one or a few select taxa were eliminated. Our survey of the literature yielded 10 sites spanning seven chronological phases, representing 417 archaeobotanical samples, 52 faunal samples, 265 plant taxonomic groups, 91 faunal taxonomic groups, and 205,090 specimens (fig. 1; table 1).

We then determined the best way to integrate the faunal
Figure 1. Map of study area, showing distribution of sites included in statistical analysis.

and botanical data sets for quantitative analysis. Integrating archaeobotanical and zooarchaeological data raises issues of differential preservation and different approaches to data generation and presentation. Plant data tend to be presented on a sample-by-sample basis, with each sample collected from a particular context dating to a single time period. Typically, the volume of the site represented by a sample is small, on the order of 20 liters. In contrast, faunal samples are often grouped from either the entire excavated area dating to a particular chronological phase (e.g., Early Bronze IV) or, if distinct use areas are defined, a large portion of the site. Because of this discrepancy in sample number and site coverage, we decided to analyze data at the phase level rather than at the sample level so that the two data sets would be more directly comparable. Consequently, all of the archaeobotanical samples dating to the same phase from any given site were condensed into one sample, and total seed counts for each species were calculated. This decision led to reduced resolution and a loss of contextual information for the plant remains, but this is not a significant problem in a regional study; Colledge, Conolly, and Shennan (2004) collapsed plant data to the phase level to successfully examine Neolithic agriculture in Europe and the Near East. Faunal data are typically presented at the phase level, but in instances where separate samples within a phase were reported, the samples were combined and NISPs summed with the same method applied to the plant data. This process resulted in 20 matched plant and animal samples, with each sample representing a single phase at a site. The matched samples were then combined to create 20 integrated lists of plant and animal data with associated seed count or NISP information (table 1; raw data are presented in the online edition).

From these lists, it was clear that different researchers group species differently and sometimes use different names for the same thing. This is obviously a problem in a synthetic study such as this. By broadening some taxonomic categories (from *Triticum aestivum* and/or *Triticum durum* to “free-threshing wheats”), we were able to diminish the “researcher effect” and lessen noise within the data set, but at the cost of reducing resolution.
### Table 1. Number of archaeobotanical samples (and zooarchaeological samples) combined per phase per site

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>'Ain Dara</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>35 (4)</td>
<td>Crawford 1999; Frey and Marean 1999</td>
</tr>
<tr>
<td>Tell Aifes</td>
<td></td>
<td></td>
<td></td>
<td>8 (1)</td>
<td></td>
<td></td>
<td>3 (1)</td>
<td>Wachter-Saskady 1998; Wilkens 2000</td>
</tr>
<tr>
<td>Tell Qarqur</td>
<td></td>
<td></td>
<td></td>
<td>63 (1)</td>
<td></td>
<td></td>
<td>13 (2)</td>
<td>Arter 1999; Smith 2005</td>
</tr>
<tr>
<td>Bab edh-Dhrá</td>
<td>12 (1)*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>35 (1)</td>
<td>Finnegan 1979; Mccreery 1980</td>
</tr>
<tr>
<td>Numeira</td>
<td>21 (1)*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Finnegan 1979; Mccreery 1980</td>
</tr>
<tr>
<td>Umm el-Marra</td>
<td>1 (1)</td>
<td></td>
<td></td>
<td>9 (4)</td>
<td></td>
<td></td>
<td>11 (1)</td>
<td>Schwartz et al. 2000</td>
</tr>
<tr>
<td>Tell es-Sweyhat</td>
<td>10 (1)</td>
<td>18 (1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Buitenhuus 1983; Hide 1990; Miller 1997; van Zeist and Bakker-Heeres 1985</td>
</tr>
<tr>
<td>Brak</td>
<td>24 (1)</td>
<td>32 (1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Charles and Bogaard 2001; Colledge 2003; Dobney, Jaques, and Neer 2003</td>
</tr>
<tr>
<td>Tell Sheikh Hamad</td>
<td>17 (1)*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Becker 1991; van Zeist 1999/2000, 2003a</td>
</tr>
</tbody>
</table>

Note. EB = Early Bronze Age; MBA = Middle Bronze Age; LBA = Late Bronze Age. These chronological phase designators do not strictly apply to the Mesopotamian sites included here but are used as convenient shorthand.

*Early Bronze Age.

*EB II/III.
Little can be done to quantitatively account for differential preservation of plant and animal remains. Since both zooarchaeologists and archaeobotanists recognize that certain species, plant parts, or elements are routinely preserved better than others and take this into account when interpreting their respective data, we adopted a similar approach here. Problems quantifying either zooarchaeological or archaeobotanical remains also exist, and these may be compounded when the two data sets are combined, particularly if sample sizes differ dramatically. Frequency data, such as the NISP for animal bones and plant part counts for archaeobotanical remains, are the most comparable, but they are differentially biased. The NISP, for example, may be artificially increased because of fragmentation, a phenomenon that does not affect plant remains to the same degree. Furthermore, while the number of animal bones per skeleton is typically constant, the number of seeds per plant can vary widely both within and between species. The use of presence/absence data can offset some of these discrepancies but does not circumvent the problems entirely. Since these issues have not yet been explored with integrated plant and animal data sets, we analyzed both abundance and presence/absence data.

Data Cleaning and Analysis

Complex data sets, such as the one examined here, are best understood with an iterative approach that employs multivariate techniques to examine both the whole data set and subsets within it. We chose correspondence analysis (CA) to analyze the data because of its suitability for heterogeneous presence-absence and abundance data with many zero values, both of which are typical of zooarchaeological and archaeobotanical records (Gauch 1982; ter Braak, 1986, 1994). CA is an open-ended ordination approach that allows the main variables affecting the species composition of the data set to be evaluated. CA can suffer from the “arch effect,” whereby a spurious quadratic arch in the ordination scores of the first (horizontal) axis is distorted in the second (vertical) axis (Jongman, ter Braak, and van Tongeren 1987, 104–105). Arches are clearly visible in attribute plots and can lead observers to “interpret the first and second axes as being two separate gradients” when there may be only a single gradient affecting data structure (Gauch 1982, 150). Detrended correspondence analysis (DCA) helps alleviate this problem, so DCA was carried out when CA yielded a spurious arch. Detrending was carried out by means of second-order polynomials. Canonical correspondence analysis (CCA) allows specific variability in the data set caused by a known variable to be examined by extracting the variation caused by that particular variable or set of variables. Sample age was used to constrain data to assess the extent to which species composition changed through time, and the statistical significance of the results were evaluated with a Monte Carlo permutations test.

Since the inclusion of rare taxa in correspondence analysis can result in noise that masks real trends, taxa present in fewer than five of the 20 samples were eliminated. Functionally similar taxa, such as cereal chaff, were combined to minimize the extent of data loss. This resulted in retention of 17 faunal taxa, 90 botanical taxa, and 180,714 specimens (see the raw data in the online edition). Raw counts were transformed logarithmically to normalize the data as much as possible, by means of the following formula: transformed count = ln (raw count + 1.0). Presence-absence data were created for each sample, with 1 denoting presence and 0 denoting absence. Data analyses were conducted with CANOCO, version 4.53, and attribute plots of the results were drawn with CanoDraw, version 4.12 (ter Braak and Šmilauer 2002).

All biplots were constructed by means of biplot scaling focused on intersample distances. Rules for interpreting such biplots are outlined in detail by ter Braak (1986, 1171–1172) and ter Braak and Verdonschot (1995). In summary, two axes are visible in each plot, together with all of the species and samples included in the analysis. Ordination places the samples along axes on the basis of variation in species composition. The primary, or most distinct, variation in the data is represented along axis 1, while secondary variation is represented along axis 2 (figs. 2, 3). The variation depicted by each axis is assumed to be driven by a different variable. The relative positions of species and sample points with respect to each axis can be interpreted with various methods. The centroid principle was used here, whereby “species scores are near the points for samples in which they occur with the highest relative abundance and, similarly, the sample points are scattered near the positions of species that tend to occur in those samples” (Lepš and Šmilauer 2003, 159; emphasis added). In other words, samples with similar contents cluster together, as do species that often co-occur within samples. Rare species tend to lie farthest from the origin. For short gradient lengths (ca. 2 SD), the biplot rule is more appropriate than the centroid principle (Lepš and Šmilauer 2003, 149–167; for a diagram illustrating the biplot rule, see fig. A1, in CA + online supplement A).

Results

Correspondence analysis of the integrated data set using both abundance and presence/absence data yielded similar ordination diagrams with a distinct arch. Since the arch represents a distortion of the first axis within the second axis, evaluating the cause of differences in species composition between samples is extremely difficult. DCA successfully removed the quadratic arch and allowed patterning in the data to be more readily observed.

The results of the DCA analyses using abundance (fig. 2) and presence/absence (fig. 3) data show a similar clustering of sites along axes 1 and 2. Moving left to right along axis 1, the clusters are as follows: (1) all included phases from Tell es-Sweyhat and Umm el-Marra; (2) all phases from Tell Qarqu; (3) all phases from Brak, ‘Ain Dara, and EBII/III (Early
Figure 2. Detrended correspondence analysis biplot of integrated plant and animal abundance data displaying 36.9% of the variance within the species data. The eigenvalues of axes 1 and 2 are 0.321 and 0.245, respectively. The total inertia is 1.532. *Ficus carica*, with coordinates of (1.2561, 6.4308), is omitted from the diagram to enhance clarity of the other species. EB = Early Bronze Age; MBA = Middle Bronze Age; LBA = Late Bronze Age.
Figure 3. Detrended correspondence analysis biplot of integrated plant and animal presence/absence data displaying 36.4% of the variance within the species data. The eigenvalues of axes 1 and 2 are 0.296 and 0.197, respectively. The total inertia is 1.356. *Ficus carica* (0.3483, 4.3518) and *Equus asinus* (2.8531, 3.7761) are omitted. EB = Early Bronze Age; MBA = Middle Bronze Age; LBA = Late Bronze Age.
The results of the CCA test conducted to assess the extent to which plant and animal use changed through time are displayed in figure A2, in supplement A. The Monte Carlo permutations test conducted to test the significance of the results indicates that there is insufficient evidence to reject the null hypothesis that plant and animal remains do not change through time.

Discussion

All of the DCA results indicate that distinct zones of cultivation, herding, and wild-resource use existed within the Near East during the Bronze and Iron Ages. Samples dating to different phases from the same site cluster together, as do geographically proximate sites (figs. 2–5). Since samples cluster by site and locale rather than time period, we conclude that environment has a stronger influence on the range of plants and animals present at a site than chronology. Such clustering of sites, while not surprising, is encouraging, because it demonstrates the robustness of CA as a method for examining regional food production and resource use. While no temporal variation within the data set could be detected through the CCA analysis, this does not mean that change did not occur. Changes in agricultural production during the Bronze and Iron Ages are to be expected, given the drastic social changes evident at this time. Nevertheless, since local environmental factors are highly influential, studies of temporal change must occur on a local scale within each cluster. Such studies are currently not possible because of the paucity of data within each locale (table 1).

From figures 2 and 3, it is evident that the remains from Numeira and Bâb edh-Dhrâ are distinct from those in northern Syria. The Jordanian sites show a heavier reliance on grapes and olives. In some respects, however, they share similarities with sites in western Syria, hinting at an agricultural system that spans the Levant. Similarities with sites in western Syria include mixed cropping of wheat, barley, legumes, and horticulture as well as the presence of flax (Linum spp.), donkeys (Equus asinus), and onagers (E. hemionus) in limited frequencies.

Across northern Syria, clear trends are visible. The clustering of sites along the first axes in figures 4 and 5 strongly suggest that available moisture is the prime determinant of variation within the integrated plant and animal data set: the gradient runs from left to right, corresponding to sites in wetter and drier regions, respectively. The DCA results enable us to make a number of observations. First, sites that lie in relatively high rainfall zones, such as Tell Qarqur and Tell Brak, demonstrate a heavier emphasis on cropping than do drier areas. Crops include einkorn (Triticum monococcum), emmer (T. dicoccum), free-threshing wheats, barley, lentil (Lens culinaris), and vetches (Vicia spp.). Second, Tell Sheikh Hamad appears to be an anomaly. The site lies in an area that currently receives approximately 200 mm rainfall per annum, placing it below the cutoff point for rain-fed agriculture, yet it clusters with the higher-rainfall sites. A series of Assyrian irrigation canals has been located around Tell Sheikh Hamad (Kühne 1990), so it is likely that irrigation was used at the site, elevating the amount of moisture available. Third, animals are also distributed along the precipitation gradient. Of the domestic animals, pigs are most dependent on permanent water sources and plot in wetter areas. Sheep and goats are much better adapted for arid conditions and figure more prominently in drier areas, such as Umm el-Marra in the steppic Jabbul Plain and Tell es-Sweyhat. Steppes are often favored grazing locations for sheep and goats because these animals can easily tolerate the dry conditions and do not have to compete with agricultural fields for land. Plants
Figure 4. Detrended correspondence analysis biplot of integrated plant and animal abundance data displaying 45.3% of the variance within the species data. Samples from Tell Afis, Late Bronze Age (LBA) Tell Bderi, Bāb edh-Dhrā, and Numeira were omitted. The eigenvalues of axes 1 and 2 are 0.269 and 0.179, respectively. The total inertia is 0.989. EB = Early Bronze Age; MBA = Middle Bronze Age.
Figure 5. Detrended correspondence analysis biplot of integrated plant and animal presence/absence data displaying 44.4% of the variance within the species data. Samples from Tell Afis, Late Bronze Age (LBA) Tell Bderi, Bâb edh-Dhrâ, and Numeira were omitted. The eigenvalues of axes 1 and 2 are 0.198 and 0.162, respectively. The total inertia is 0.811. EB = Early Bronze Age; MBA = Middle Bronze Age.
typically found in pastures, such as grasses (Gramineae) and small legumes (Trifolium, Melilotus, and Trigonella), are more prominent in these drier regions, underscoring the importance of herding. Fourth, most of the wild animal taxa also fit this pattern: more arid-adapted species, such as gazelle, equids, and hare, are found in greater relative abundances at sites in dry areas, while forest-adapted species, such as red deer and fallow deer, are more common in wet areas. Exceptions include roe deer and cattle, which are more important in drier areas than expected. Small taxa, such as turtles and fish, are also most important in wet areas, in particular at the site of Tell Qarqur, which had access to the marshes of the Ghab Valley.

Conclusions

This study demonstrates that correspondence analysis is an effective and powerful tool for integrating plant and animal data at the regional level, assuming that care has been taken to enhance comparability of the data. When data cleaning is not rigorous enough, CA can help isolate problematic samples. The trends visible from presence/absence and abundance data in this data set are largely similar, but we highly recommend exploration of both data sets. By combining plants and animals at the analytical rather than the discussion phase of research, CA allows for simultaneous exploration of the data sets and moves us one step closer to a truly integrative approach. The approach adopted here can theoretically be applied to any location or time period. The technique allows regional agroeconomies and spheres of wild-resource use to be identified and provides the quantitative rigor required to establish the veracity of what may otherwise be an impressionistic trend.

Given that the vast majority of Bronze and Iron Age settlements in the Near East owe their existence to agriculture and that hundreds of Bronze and Iron Age sites have been excavated in the study area, it is remarkable that only 10 sites met the criteria for inclusion in this study. This shortcoming resulted in a spotty distribution of site samples, particularly in the Levant, where the scarcity of published botanical reports is to be lamented. Given that vast portions of the study region are severely underrepresented, it is difficult to present a truly regional synthesis at present. Nevertheless, we must start somewhere: from the data available, distinct regional patterns of agriculture and wild-resource use are evident. In particular, this study demonstrates similarity between closely situated sites located in similar environments, suggesting that Near Eastern agriculture does indeed have a regional flavor that is strongly influenced by environmental factors. Available moisture appears to be the strongest governing factor.

Faunal and botanical data independently augment our understanding of the past, but when integrated, the two data sets provide more comprehensive and enriched insights into ancient subsistence and landscape use. The next step is for archaeobotanists and zooarchaeologists to integrate their sampling strategies at the planning stage of an excavation and coordinate their approach to identifying, quantifying, and publishing the remains. Minimum reporting standards that outline methods and use standardized taxonomic grouping would also be useful. Since archaeology is inherently destructive and typically the published record is all that remains, published data must have some utility beyond the immediate scope of the report at hand and should be presented with this in mind. Large geographic gaps in the data set also must be filled. Given the link between agriculture and dramatic social change during the Bronze and Iron Ages, the paucity of paired archaeobotanical and zooarchaeological studies is astonishing. This must change. Only when these gaps are filled can the regional nature of agriculture be truly understood and knowledge of the rise, maintenance, and collapse of early state-level societies within the Near East be placed on a firmer footing. We hope that this effort will underscore the need for regional studies of agriculture and encourage the collection and publication of the data required to meet this goal.

Acknowledgments

We thank Melinda Zeder, for valuable suggestions and discussion of the data presented here, and Daniel S. Adler and five anonymous reviewers, for comments on earlier drafts of the paper. This research was supported in part by a National Science Foundation grant (0210651); a Grant-in-Aid of Research from the National Academy of Sciences, administered by Sigma Xi, the Scientific Research Society; and a Harris/Torch Grant (Committee on Archaeological Policy, American Schools of Oriental Research) awarded to Alexia Smith. We thank them all. We dedicate this paper to the memory of Stine Rossel.

References Cited


Charles, M., and A. Bogaard. 2001. Third millennium BC


