1
2 Organic geochemical approaches to identifying formation processes
3 for middens and charcoal-rich features
4
5

6 X. Gao^{a,*}, M. Norwood^b, C. Frederick^c, A. McKee^d, C.A. Masiello^{a,e}, P. Louchouart^{b,f}
7

8 ^a *Department of Earth Science, Rice University, 6100 Main St., Houston, TX 77005, USA*
9

10 ^b *Department of Marine Science, Texas A&M University at Galveston, Galveston, TX 77553,*
11 *USA*
12

13 ^c *Department of Geography and the Environment, University of Texas at Austin, Austin, TX*
14 *78712 and Consulting Geoarchaeologist, 2901 FM 1496, Dublin, TX 76446, USA*
15

16 ^d *Versar, Inc. Plano, TX 75074 and Geospatial Information Sciences Department, University*
17 *of Texas at Dallas, Richardson, TX 75080, USA*
18

19 ^e *Departments of Chemistry and BioSciences, Rice University, 6100 Main St., Houston, TX*
20 *77005, USA*
21

22 ^f *Department of Marine Science, Texas A&M University at Galveston, Galveston, TX 77553,*
23 *USA*
24
25
26
27
28
29
30

31 *Corresponding author. Tel: 713-348-4114.
32 *E mail address: xdgao@rice.edu(Xiaodong Gao).*
33
34
35

36 ABSTRACT

37 Charcoal is an important component of archaeological sites that can yield fundamental
38 information useful in understanding past human activities. We applied a suite of organic
39 geochemical tools to identify and quantify anthropogenic molecular charcoal fragments in an
40 archaeological site in Panola County, N Texas to determine the provenance of a buried
41 melanic horizon. The site is a small Middle-Late Caddo occupation inhabited principally
42 between ca. 1450 and 1600 AD. We tested three molecular techniques to determine their
43 ability to constrain site formation processes: solid state ^{13}C nuclear magnetic resonance
44 spectroscopy (NMR), lignin free phenols and lignin oxidation products (LOPs) and
45 anhydrosugar analysis. We demonstrated that, in conjunction, the three afforded the
46 identification and quantification of anthropogenic molecular charcoal fragments at this
47 archaeological site.

48 The approach allowed reconstruction of the fire history and vegetation type around the site
49 during the occupation period. The NMR and anhydrosugar analysis pointed to temperature
50 conditions associated with thermal refuse in different contexts. In particular, geochemical
51 signatures in a smudge pit indicated a higher combustion temperature than for the charcoals
52 in the other paleosol samples, suggesting multiple fire uses. The data also showed differential
53 persistence of fire signals in environmental samples, likely as a result of differential
54 decomposition rate and solubility, which may lead to heterogeneous fates of different forms
55 of pyrogenic C in the environment. The techniques expand the tool kit available to
56 archaeologists and soil scientists in detecting the human footprint in soil formation,
57 especially in warm, humid climates where decomposition is rapid.

58
59 **Keywords:**

60 Midden, charcoal, smudge pit, pyrogenic C, biomarker, NMR, lignin free phenols, LOPs,
61 anhydrosugar, Caddo

1. Introduction

A midden is loosely classified in archaeological literature as a localized area of dark earth and elevated artifact concentrations that resulted from the deliberate discard of refuse (e.g. food remains and domestic materials; Perttula, 2004; Hays and Weinstein, 2010; Jackson et al., 2010). Middens, especially the organic fraction, are important for reconstructing past human behavior, such as social organization, status, wealth, ritual, domestic and specialized activities (Boone 1987, Knight, 2004; Eberl et al., 2012). Given that they may have complex formation processes resulting from varied environmental settings, diverse human activities and post-depositional alteration, it may not be sufficient to simply identify such contexts as middens. Instead, it may be more informative to apply a range of tools to better understand the origins (Schillito and Matthews, 2013). Thin section micromorphology, molecular geochemistry and other micro-archaeological techniques have been applied to the examination of complex midden deposits at higher resolution than resolved by the unaided eye (Beck, 2007; Holliday and Gartner, 2007; Wilson et al., 2009; Misarti et al., 2011; Eberl et al., 2012; Shillito and Matthews, 2013).

Decomposition often makes it challenging to identify the sources of midden organic matter (OM) and the problem is especially acute in humid tropical and subtropical regions, where heat and moisture can drive rapid soil OM decay. However, one form of OM, charcoal, tends to be preserved at the molecular level even in warm, humid conditions, and is commonly found at many archeological sites where other forms of OM have been completely degraded. In ecosystems where wildfire is not a natural occurrence, it is reasonable to assume that charcoal in soil indicates human activity. Indeed, one study has shown that charcoal layers coincide with periods of known human activity from the mid-Holocene onwards in Spain (Cortizas et al., 2005).

Although to most archeologists charcoal is a singular substance, recent organic geochemistry research has demonstrated that it is a heterogeneous material exhibiting a wide range of physicochemical properties. For example, some low temperature charcoals may contain incompletely combusted lignocellulose fragments, whereas high temperature charcoal may form a disordered graphitic-like structure under pyrolysis conditions (Hedges et al., 2000; Masiello et al., 2004; Keiluweit et al., 2010; Wiedemeyer et al. 2015). Biomass burning is a complex process and often includes a sequence of stages, including ignition,

flaming plus glowing plus pyrolysis, glowing plus pyrolysis (smoldering), glowing, and extinction (Andreae and Merlet, 2001). The initial flaming stage can use up atmospheric O₂ quickly and result in O₂ depleted pyrolysis conditions in the latter stages, where high temperature charcoal is produced. Indeed, charcoal produced from wildfires is often a mixture of low temperature and high temperature (> 500 °C) pyrogenic-derived materials (Singh et al., 2012). Production conditions (e.g. fire temperature and fire exposure time) can result not only in differences in charcoal composition, but also strongly influence post-depositional alteration (Cheng et al., 2007; Keiluweit et al., 2010). Using fire biomarkers unique to production conditions and fuel material, it is possible to reconstruct site-specific fire history and associated human activity. Therefore, charcoal is an important component of archaeological sites that can yield fundamental information useful in understanding past human activity.

Identification and quantification of charcoal in soils is, however, not straightforward. Charcoal may be broken into molecular fragments that are unidentifiable under a microscope in humid, tropical or subtropical environments (Braadbaart et al., 2009; Spokas et al., 2014). Also, charcoal is often present heterogeneously in soils as a mixture with other forms of soil OM (SOM), rendering identification and quantification of ‘ancient’ charcoal even more difficult (Kaal et al., 2008b). Indeed, one study reported that charcoal was abundant in the extractable SOM fraction from an archeological site where macroscopic charcoal was absent, suggesting that the macroscopic charcoal record used in archaeological studies often provides an incomplete picture of the fire history (Kaal et al., 2008b). Furthermore, despite the stability and recalcitrance of charcoal, it does experience some environmental alteration and biodegradation in soil and may develop various surface functional groups over time.

A variety of techniques have been proposed in recent years for characterizing and quantifying the molecular signal of charcoals in environmental media, including ¹³C nuclear magnetic resonance spectroscopy (NMR; Schmid et al., 2001; Kleber et al., 2003; Kaal et al., 2007), pyrolysis-gas chromatography-mass spectrometry (Py-GC-MS (Kaal et al., 2007, 2008a,b), benzene polycarboxylic acids (BPCAs; Glaser et al., 1998; Kaal et al., 2008b; Schneider et al., 2011), hydrolysis and methylation (THM; Kaal et al., 2008b) and molecular fire markers (e.g. levoglucosan; Kuo et al., 2008a; Knicker et al., 2013). Each of these has advantages and limitations, providing some unique information but an incomplete picture of

the charcoal material (Preston and Schmidt, 2006). For example, many charcoal analysis methods (e.g. Py-GC-MS, BPCAs and THM) require pre-treatment (wet chemical extraction) to remove non-combustion (pyrogenic) derived OM by assuming reactivity differences between operationally defined charcoals and non-charcoal material.

The molecular fire biomarker methods provide important information on fire history but often have an analytical window for production temperature because the biomarkers are sensitive to fire temperature. For instance, low temperature pyrogenic biomarkers (e.g. levoglucosan and free lignin phenols) are lost at higher temperatures due to molecular transformation and volatilization, where other higher temperature pyrogenic biomarkers (e.g. PAHs with ordered condensed aromatic rings) begin to appear (Kuo et al., 2008a, 2011a; Harvey et al., 2012). Lignin oxidation products (LOPs) are also lost in higher temperature combustion (Kuo et al., 2008b), so their presence in pyrogenic material can serve as an indirect marker for lower temperature combustion.

Our objective was to apply a suite of complimentary analytical approaches to identify and quantify fire biomarkers deposited among Caddo settlement soils in East Texas. We reconstructed the fire history and vegetation type (angiosperm vs. gymnosperm) around the sites during the occupation period using charcoal recovered from midden deposits. We employed a suite of organic geochemical approaches, including ^{13}C CPMAS (cross polarization magic angle spinning) NMR in conjunction with a molecular mixing model (MMM), analysis of free lignin phenols and lignin oxidation products, and analysis of anhydrosugar content (levoglucosan, mannosan and galactosan). Each of these tools brings a slightly different perspective to human use of fire. The combination of ^{13}C CPMAS NMR with MMM can provide quantitative information on the total amount of charcoal in a sample, along with the distribution of other forms of organic carbon (e.g. carbohydrate, protein, lignin, lipid, charcoal, carbonyl) from the ^{13}C NMR spectral peak distribution (Baldock et al., 2004). Anhydrosugars, on the other hand, are produced exclusively from the thermal breakdown of cellulose and hemicellulose during low temperature combustion (150-350 °C) and can be used as a unique low temperature fire biomarker (Simoneit, 1999; Kuo et al., 2008a). Free lignin phenols can also be used as a fire biomarker (Kuo et al., 2011a). However, unlike charcoal and anhydrosugars, which are produced exclusively from pyrogenic processes, free lignin phenols may have two different sources: thermal breakdown

of lignin during combustion and/or biodegradation products of lignin during post-depositional alteration. Although free lignin phenols alone are not conclusive indicators of fire, their distribution can be informative when paired with other fire markers. Schmid et al. (2001) used similar complementary techniques to examine the composition of organic carbon at an archaeological site. Here we use these tools together to reconstruct the fire history and local vegetation type around the site during the occupation period.

2. Material and methods

2.1. Site description

The Murvaul Creek site (41PN175) is in Panola County, NE Texas (Fig. 1). The site is a small Middle-Late Caddo occupation that was inhabited principally between ca. 1450 and 1600 AD. It is within 60 m of the edge of an upland alluvial terrace, likely at least Eocene in age. Artifacts from the occupation are contained within sandy sediments in the upper 75 cm of the soil profile.

We identified four major stratigraphic units at the site (Fig. 2). At the base of the excavations was a reddish brown argillic (or Bt) horizon presumably formed in Eocene bedrock (Wilcox Group). This layer was overlain by ca. 1 m of loose unconsolidated sand, which could be divided into three horizons from bottom to top: (i) an Ehorizon, (ii) a paleosol (or Abhorizon) and (iii) a recent colluvial deposit (see Fig.2). The Caddo lived upon the paleosol, and this deposit in one portion of the site was notably darker colored and contained more artifacts and identified as a Caddo refuse dump or midden. Overlying the paleosol was a body of loose sand that pinched out upslope and thickened down slope and clearly represents a period of soil erosion. Single grain optically stimulated luminescence (OSL) dates obtained from this unit indicate that it was deposited in the last 300 yr (Bateman, 2012) and a thin, weak Ahorizon had formed at the top of the colluvium since deposition ceased.

Similar to other small Caddo occupational sites, the Murvaul Creek site contains structural and other habitational features (e.g. post molds, storage pits and cooking hearths) bordering the periphery of a midden (ca. 30 m in diameter), which contains a higher artifact concentration than habitational areas. The challenge with large midden features such as this for archaeologists lies in understanding the behavioral processes involved with creating such

a feature. Analysis of midden deposits in the archaeological literature is often limited in scope to the macro-remains, such as faunal and floral material and ceramic and lithic debris. However, in a highly leaching environment, soil acidity limits the potential for the recovery of macro-remains of faunal and floral material, restricting archaeologists' ability to interpret subsistence practices. Identifying soil biogeochemical methods for interpreting the principal constituents that may have been added to the soil should provide a useful and complementary tool for understanding past behavioral practices.

2.2. Field sampling

We collected five representative soil samples from the archeological site, including one modern natural surface soil, three paleosol samples (two within the midden, one outside it) and one smudge pit sample (Figs. 1 and 2):

Sample 1: Modern Ahorizon soil formed in Colluvium (lab ID, 1982_0-5cm)

This sample was collected from the surface (0-5 cm) of a very faintly melanized A horizon soil that formed in the top 10 cm of the colluvium. The primeval forest that covered the site in the Caddo period was cleared sometime prior to 1938 and the pine-sweetgum-hickory forest, with an understory of yaupon holly, regenerated during the middle to late 20th century. There is no known history of either deliberate or natural burning of the surface during the 20th century, but the sediment comprising the colluvium was derived from erosion of the underlying soil horizons upslope to the south.

Sample 2: Midden Paleosol - strongly melanized (lab ID, 1982_40 cm)

This sample was collected from the northern end of the site within the midden, where it appeared to be the most strongly melanized and artifacts were recovered with the highest density.

Sample 3: Midden Paleosol - weakly melanized (lab ID, TU-80)

This sample was collected from the eastern end of the excavations within the midden, where the dark color was less pronounced than for the previous sample (1982_40 cm).

Sample 4: Non-Midden Paleosol (lab ID, PN175_16)

This is a weakly melanized paleosol (presumed unaltered forest soil) inferred to be a vestige of the Late Caddo period forest topsoil where it experienced little anthropogenic alteration.

Sample 5: Smudge pit feature (lab ID, F-10)

Smudge pits are small depressions used to combust OM, with the primary goal of creating smoke (Binford, 1967). In functional terms they are recognized in the field as shallow pits with abundant charcoal, as opposed to hearths which, although they may contain charcoal, typically have a wide range of thermal decomposition products owing to a presumably higher burning temperature and a more freely oxidizing thermal environment.

2.3. Elemental analysis

The contents of organic C and N were measured for the bulk soil samples and the HF treated material for ^{13}C NMR analysis (discussed in detail in the next section) by catalytic combustion and subsequent chromatographic separation and detection of CO_2 and N_2 with a Costech ECS 4010 elemental analyzer. All samples were oven-dried at 60 °C overnight and ground prior to analysis. For soil samples, inorganic C was removed prior to measurement. Soil samples were treated with full strength (7 M) HCl fumigation in open-top Ag capsules (Costech Analytical), following the procedure of Harris et al. (2001).

2.4. Lignin free phenols, LOPs and anhydrosugar analysis

We extracted the free monomers (anhydrosugars and lignin free phenols) from the 5 soil samples using modifications of the methods of Kuo et al. (2008) and Louchouart et al. (2009). Briefly, homogenized soil samples (ca. 400 mg) were mixed with pre-combusted sand and extracted using accelerated solvent extraction (ASE 200 Dionex) with 9:1 dichloromethane (DCM):MeOH at 100 °C and 1500 psi (103.4 bar). Prior to extraction, samples were spiked with deuterated *trans*-cinnamic acid (*d7-trans*-cinnamic acid), used as a surrogate standard for lignin phenols, and *d7*-levoglucosan, a surrogate for levoglucosan analysis. Each extract was dried with Na_2SO_4 , then in a LabConcoTM solvent concentrator and re-suspended in pyridine (500 μl). Both free phenol and anhydrosugar extracts were

derivatized for 15 min using N,O-bis (trimethylsilyl)trifluoroacetamide containing 1% trimethylchlorosilane (99:1 BSTFA:TMCS) at 75 °C and immediately analyzed using a Varian triple quadrupole 480-300 GC-MS system with a fused silica column (J&W DB-5MS, 30m x 0.25mm i.d., 0.25µm film thickness; Agilent Technologies, Santa Clara, CA, USA). Each sample was injected splitless using He as a carrier gas. The GC injector and GC-MS interface were at 280 °C and 270 °C, respectively. The mass spectrometer was operated in electron ionization mode (EI, 70 eV) using full scan (FS) mode. Compound assignment was performed using GC retention times and by comparing mass spectra with those of commercially available standards. The average recovery of levoglucosan and lignin phenols was $98 \pm 4\%$ and $100 \pm 9\%$, respectively.

LOPs were analyzed using the method developed by Hedges and Ertel (1982) and Goñi and Hedges (1992) with modifications (Kuo et al., 2008b; Louchouart et al., 2010). The method is able to convert lignin polymers to lignin monomers in charcoal network and has been successfully tested on a number of charcoals produced at different temperatures (Kuo et al., 2008b). Briefly, each ground homogenized sample containing 2-5 mg OC was combined with CuO (ca. 330 mg), $\text{Fe}(\text{NH}_4)_2(\text{SO}_4)_2 \cdot 6\text{H}_2\text{O}$ (ca. 100 mg) in a stainless steel reaction vessel (3 ml; Prime Focus Inc.). the reaction vessel was filled with ca. 3 ml of 2 N Ar sparged NaOH solution, placed in a customized purging block (Prime Focus Inc.) and closed under an absence of O_2 . The vessel was heated at 155 °C for 3 h in a customized Hewlett-Packard 5890 gas chromatograph fitted with a rotating carousel. After cooling, d7-*trans*-cinnamic acid was added to the vessel as a surrogate standard, centrifuged free of solids and rinsed 2x with 3 ml of 1 N NaOH. The aqueous solution was acidified with 6 N HCl ($\text{pH} \leq 1$) and solvent-extracted 3x with EtOAc. Like the free phenol extraction method, the EtOAc extract was dried with Na_2SO_4 , then in a LabConco™ solvent concentrator and re-suspended in pyridine. The GC-MS method used to analyze LOPs was identical to that described above for free lignin monomers.

2.5. ^{13}C NMR spectroscopy

Due to the relatively low C content and high Fe content of the soil samples, it was necessary to demineralize them prior to ^{13}C NMR analysis to remove magnetic material and concentrate OC. For each sample, 5 g were added to a 50 ml polyethylene centrifuge

284 tube, to which 50 ml of 3% HF was added. The sample was then mixed on a vortex stirrer
285 and shaken for 2 h on an orbital shaker, before centrifuging at 2000 rpm for 10 min, after
286 which the supernatant was discarded. The sample was subjected to five 2 h digestions and
287 one 8 h digestion for a total of 18 h, after which it was rinsed 3X with deionized water,
288 decanted into a glass vial and dried in an oven at 60 °C. Through the process of rinsing,
289 decanting and drying, the samples became laminated, with a thin organic surface layer
290 covering a white mineral layer. We separated the surface organic layer from the mineral layer
291 using tweezers and used the organic layer for elemental analysis and ^{13}C NMR analysis.

292 We used solid state CPMAS NMR spectroscopy to examine the chemical structure
293 based on ^{13}C signal intensity and chemical shift. A CP with radio frequency ramp - variable
294 amplitude CP was used to restore a broad matching profile and enable more efficient
295 polarization transfer at higher spinning speed (Metz et al., 1994). The spectra were recorded
296 using a Bruker Avance 200 MHz NMR spectrometer (Bruker Corp., MA), equipped with 4
297 mm MAS and operated at a rotor spinning frequency of 7 kHz. CP spectra were acquired by
298 applying a 90 degree ^1H pulse, a 1.0 ms ^{13}C contact pulse, composite pulse proton decoupling
299 and a 5 s recycle delay. The CPMAS pulse sequence was chosen because it can improve the
300 signal to noise ratio (S/N) up to a factor of 4 VS. direct polarization (DP) MAS. In addition,
301 CP can significantly reduce total data acquisition time because of its reliance on proton
302 relaxation ($T_1\text{H}$) rather than carbon relaxation ($T_1\text{C}$) used by DP pulse sequences, where $T_1\text{H}$
303 is at least an order of magnitude faster than $T_1\text{C}$. However, CP pulse sequences are prone to
304 under-detection of ^{13}C nuclei remote from protons. Large aromatic clusters formed during
305 higher temperature combustion or pyrolysis-containing core C more than 3 to 4 bonds away
306 from the next proton were found to be underestimated with solid state CP MAS ^{13}C NMR
307 (Alemany et al., 1983; Freitas et al., 1999; 2001; Smernik et al., 2002). On the other hand, no
308 significant difference was observed between the spectra of CP and DP pulse sequences for
309 charcoal produced at temperatures < 600 °C (Freitas et al. 1999; 2001), suggesting that
310 charcoal formed in natural wildfire is unlikely to be under represented with CP MAS ^{13}C
311 NMR.

312 Spinning sidebands (SSBs) for some of the functionalities (e.g. aromatic, carbonyl)
313 with large chemical shift anisotropy (CSA) were observed under the experimental conditions.
314 With a 200 MHz NMR spectrometer (50 MHz for ^{13}C) with a MAS rate of 7 KHz, SSBs

appear at 140 ppm (7 KHz/50MHz) away from their corresponding central bands at both sides. Therefore, the signal loss from the central bands can be approximately corrected by adding twice the intensity of the low field SSBs.

After SSB correction, the relative allocation of signal was assigned to seven individual spectral regions corresponding to different functionalities: alkyl C (0-45 ppm), N-alkyl/methoxyl (45-60 ppm), O-alkyl (60-95 ppm), di-O-alkyl (95-110 ppm), aromatic C (110-145 ppm), phenolic (145-165 ppm) and amide/carboxyl (165-215 ppm). We used a MMM (Baldock et al., 2004) to estimate the fraction of charcoal and other biochemical components in each sample. The method assumes that SOM exists in the form of several distinct biomolecules (e.g. carbohydrate, protein, lignin, lipid, carbonyl and charcoal), with each type having a characteristic ^{13}C NMR spectrum and C/N ratio. The C/N ratio and signal distribution across the seven predefined ^{13}C NMR spectral regions for each biomolecule were mathematically mixed together to fit the measured spectrum of a sample (Baldock et al., 2004). Concentration determined with the model are expressed as the proportion of OM contributed by each molecular component. The Solver add-in of MS Excel was used to determine the linear combination of molecular components that best-fit the NMR spectrum of each sample. The model also reports the difference in the sum of squares between measured and predicted NMR signal distributions, which can be used to determine the quality of the model fit. We additionally report the ratio of alkyl to O-alkyl signal, which provides information on the degree of decomposition of a sample (Baldock et al., 1997).

The C observability (Table 3) obtained from spin counting (Smernik et al., 2000, 2002) is consistent with other studies using CP MAS to detect charcoal content (Kaal et al., 2008b; McBeath et al., 2014). Our data exhibit a close association between C observability and soil charcoal content (r^2 0.86), with the samples having the highest charcoal content having the lowest C observability. Results from a prior study by Kaal et al. (2008b) suggested that, even though the CP pulse sequence achieves much lower C observability than that of the DP pulse sequence for charcoal, the two pulse sequences reported similar charcoal content using the MMM (CP charcoal content on average was 2-12% lower than charcoal content reported from DP pulse sequence), suggesting that the NMR CP pulse sequence is a reliable, non-destructive technique for quantifying charcoal formed in wildfires with a relatively low temperature in complex matrices such as SOM.

3. Results and discussion

3.1. Elemental content and C/N data suggest high pyrogenic C content of paleosols

We expected that middens receiving both charcoal and household refuse would have higher %OC and %N relative to the modern surface soil, leading to lower C/N values. On the other hand, we anticipated that the smudge pit site would have higher %OC and lower %N, leading to higher C/N than for the modern surface soil. Instead, we found that all the five soil samples contained very low %OC content (Table 1). The weakly melanized midden (sample 3) had the lowest OC (0.25%) and the smudge pit feature (sample 5) had the highest (1.31%), suggesting that the sandy soils at the archeological site have a low capability for preserving OM. In addition, the smudge pit feature (sample 5) has the highest C/N value, consistent with its use to produce smoke, not as a receptacle for higher-N household waste. The modern natural surface soil (sample 1) had the lowest C/N and low %OC and %N, again indicative of soil and climatic conditions that interfere with OM storage.

In highly humified SOM, high C/N values and aromatic content are consistent with a high proportion of pyrogenic OM, which tends to be low in N. Thus, the C/N data are consistent with our initial speculation that the smudge pit feature may have the highest charcoal content, whereas the modern natural surface soil has the lowest charcoal content (if any). Surprisingly, the paleosol outside the midden (sample 4), which was expected to have limited anthropogenic influence, also exhibited higher C/N than the modern surface soil, indicating that the paleosol may also contain a high content of pyrogenic OC, either from natural wildfire or anthropogenic fire.

3.2. Solid state ^{13}C NMR and MMM results indicate high charcoal content of paleosols

The ^{13}C CP NMR spectra were obtained on the organic portion of the soils concentrated by HF treatment. The spectra of the 5 soil samples were distinctly different (Fig. 3), as indicated by the signal distribution among the seven chemical shift regions (Table 2). The spectrum of the modern natural surface soil (sample 1) exhibited typical features of natural SOM (Baldock et al., 2004), dominated by peaks in the alkyl C (0-45 ppm), oxygen-

substituted alkyl (O-alkyl) C (65-95 ppm) and the carboxylic/amide C (165-215 ppm) regions. The alkyl and O-alkyl peaks originate mainly from lipids and carbohydrates, respectively.

The spectra of the three paleosol samples (2, 3, and 4) all contained a large peak centered at ca. 130 ppm, with diminished alkyl and O-alkyl peaks compared with the natural surface soil. The peak at ca. 130 ppm is indicative of aromatic structures, which may originate from either lignin or charcoal. In addition to aromatics, the lignin NMR spectrum also contains a large peak in the region of 145-165 ppm, corresponding to phenolics, that are very low in the charcoal structure. The ^{13}C NMR peak intensity ratio of aromatics/phenolics is ca. 1.5 and 5.0 for lignin and charcoal, respectively (Baldock et al., 2004). Therefore, the higher aromatic/phenolic values for the 3 paleosol samples (Table 2) vs. that of lignin are consistent with the presence of charcoal. The intensity of the aromatic peak from the smudge pit (sample 5) was higher than that of the three paleosol samples and became the predominant peak in the spectrum, suggesting that the OC in the smudge pit was dominated by highly aromatic charcoal material.

The alkyl/O-alkyl ratio of a sample, an NMR-based index of SOM decomposition, can provide information on the degree of decomposition/decomposability of SOM samples, where higher values indicate a higher degree of decomposition and lower residual decomposability (Baldock et al., 1997). The alkyl/O-alkyl ratio of the natural surface soil was higher than that of the paleosol samples and the smudge pit (Table 3), suggesting that the OC in the paleosol samples and the smudge pit was less decomposed than the OC in the modern surface horizon soil. The observation that the old OC was less decomposed than the modern OC may reflect site-specific OC preservation mechanisms. We speculate that this could result from the sudden burial of the paleosol following the deposition of the colluvium.

We further processed the NMR spectra through the MMM to obtain modeled concentrations of the six major biochemical components (carbohydrate, lignin, protein, lipid, carbonyl and charcoal; Table 3). The modern natural surface soil also contained a significant amount of charcoal (ca. 8%), suggesting that surface soil may have experienced historical wildfire burning. However, it is important to note that the colluvial sediment making up the modern soil is derived from erosion of the paleosol upslope and this charcoal could be associated with the Caddoan occupation. Support for this lies in the radiocarbon age of

isolated charcoal fragments collected from the colluviums, which were found to be in reverse order, becoming older towards the modern ground surface (McKee et al., 2014).

Compared with the natural surface soil, the paleosol samples contained much higher charcoal content, ranging from ca. 33-39% of total OC. The MMM results unambiguously show the presence of molecular charcoal fragments in the paleosol samples due to past human activity and/or historical wildfire burning. As expected, the smudge pit feature had the highest estimated charcoal content (ca. 67% of OC) with very small contributions from other biochemical components. The low concentrations of carbohydrates, proteins, lignin and lipids in our HF-treated paleosol samples closely resemble the composition of black carbon in soils measured with other techniques (e.g. pyrolysis GC-MS) at other sites (Kaal et al., 2008a).

In addition to charcoal, the lignin content calculated from the MMM can be used as an indirect fire biomarker and provides additional information on the pyrogenic material. As shown in Table 4, the lignin content for the two paleosol samples within the midden (samples 2 and 3) and the smudge pit feature (sample 5) was remarkably low and was associated with a high content of carbonyl component. The carbonyl component in the MMM represents functional groups formed upon degradation. Our observation of a relationship between low lignin content and high carbonyl component is consistent with studies that examined black carbon content of an acidic colluvial soil in Spain (Kaal et al., 2008a, b). The lack of lignin in the charcoal samples can be explained by significant thermal alteration during biomass burning (Kuo et al., 2008b) and/or post-depositional biodegradation. As a consequence, the high content of the carbonyl component is the product of oxidation, in particular carboxylation (Kaal et al., 2008; Ascough et al., 2011). Indeed, Cohen-Ofri et al. (2006) reported that fossil charcoal contained abundant carboxyl groups, which were absent from modern charcoal.

Despite all three paleosols containing a similar level of charcoal, the paleosol outside the midden (sample 4) had significantly higher lignin concentration than the two paleosols within the midden, as indicated by the MMM data. We speculate that pyrogenic C (charcoal) within and outside the midden deposit may have different origins or have experienced different post-depositional alteration. The charcoal within the midden was likely created by habitation burning (e.g. heating, cooking), whereas the charcoal in the paleosol outside the

midden may have originated from natural wildfire or different cultural processes (e.g. open field burning). Different burning processes (e.g. fire temperature, controlled vs. open burning) may have resulted in different thermal alteration and composition of the pyrogenic material remaining behind within and outside the midden.

3.3. Anhydrosugar analysis suggests different uses of fire by Caddo

Environmental pyrogenic C is a continuum containing both low and high temperature-derived fire biomarkers (e.g. anhydrosugars, lignin free phenols and highly aromatic molecular constituents). Due to differences in their molecular structure, we expect these fire-derived biomarkers to behave differently in a post-depositional setting. For example, low-temperature pyrogenic OM produced <450 °C is enriched with depolymerized and functionalized lignocellulose constituents that are highly soluble and decompose on timescales of days to weeks when dissolved in water (Harvey et al., 2012; Norwood et al., 2013), whereas high temperature pyrogenic OM formed > 450 °C is comprised almost entirely of aromatic structures with differing degrees of condensation (Harvey et al., 2012; Wiedemeier et al., 2015), which are much less soluble and labile than the low temperature components (Harvey et al., 2012; Myers-Pigg, 2015).

Monosaccharide anhydrosugars (levoglucosan, mannosan and galactosan) are formed exclusively from the thermal breakdown of cellulose/hemicellulose during combustion (Simoneit, 1999). Therefore, anhydrosugars can be used as unique biomarkers for low temperature (150-350 °C) combustion of any biomass containing cellulose/hemicellulose (Kuo et al., 2008a; Myers-Pigg, 2015). All the soil samples yielded detectable concentrations of anhydrosugars, including the modern natural surface soil (Table 4). Despite a short half-life (3-7 days) in solution (Norwood et al. 2013), these fire-derived anhydrosugars were preserved in the soil, suggesting a role for long term physical protection in the soil. This is consistent with reports of anhydrosugar presence in sediments hundreds to thousands of years in age (Kuo et al., 2011b; Hopmans et al., 2013), which demonstrate the potential for the preservation of this labile pyrogenic constituent in ancient deposits.

The anhydrosugar content varied among the five soil samples, reflecting different pyrogenic processes and post-depositional degradation experienced by each of the samples (Table 4). The occurrence of anhydrosugars in the modern surface soil unambiguously

demonstrates the presence of carbonaceous residues from biomass burning. Likewise, all three paleosol samples contained some level of anhydrosugars. The smudge pit feature had the lowest anhydrosugar content, consistent with pyrogenic byproducts derived from higher temperature combustion than the charcoal in the other paleosol samples and the modern surface soil sample. Indeed, as revealed by the MMM data (Table 3), negligible lignin content was observed in the smudge pit sample, suggesting that the thermal decomposition of lignocellulose fragments was complete at such high fire temperatures. This contrasts with the initial assumption that the smudge pit was created in a relatively low temperature burning event.

In addition, the anhydrosugar content did not correlate with charcoal content, suggesting differential production rate and/or landscape fate for the two different fire markers. Charcoal and anhydrosugar contents may track different components of the pyrogenic C continuum based on the production fire temperature, and may experience heterogeneous fates in the environment.

3.4. Lignin analysis serves as an indirect fire biomarker and provides additional information on fire history

Lignin free phenols and LOPs typically occur in three monomer forms: cinnamyl (C), vanillyl (V) and syringyl (S) units. The $\Sigma 8$ value is the sum of the total concentration of individual lignin monomers from the whole soil (w/w), whereas $\lambda 8$ is the total concentration of the monomers normalized to OC (w/100 mg OC). Lignin free phenols and LOPs can provide a wide range of information on ecosystem and fire history (Kuo et al., 2008b; Myers-Pigg et al., 2015).

LOPs can provide information on regional vegetation type at the time of the paleosol formation based on the ratios of the individual monomers. For instance, angiosperm-derived lignin is enriched in syringyl (S) units and non-woody tissue (e.g. leaves and grasses) is enriched in cinnamyl (C) units. These relative enrichments have led to the use of C/V and S/V ratios to determine the vegetation source of natural OM in soil (Hedges and Mann, 1979; Hedges and Ertel, 1981). The LOPs from the modern natural soil, paleosols and smudge pit feature all contained V, S, and C monomers (Table 5). The S/V and C/V ratios suggest that the major biomass source of the OC was dominated by woody and non-woody angiosperm

tissue (Hedges and Mann, 1979), suggesting that the Caddo consumed (and burned) both local trees and leaves. The paleosols and the smudge pit feature had significantly lower S/V but higher C/V values than the modern natural soils, suggesting a higher input of non-woody (e.g. grass, leaves) derived OM. The burning of agricultural fields would yield a higher C/V value, as the soft tissue of edible grasses (e.g. corn leaves) would be expected to have a higher value. The Caddo have been reported to at least occasionally maintain their fields through burning (Newcomb, 1961; Swanton, 1996).

In addition, the ratio of LOPs containing carboxyl groups to those containing aldehyde groups - the acid/aldehyde ratio and abbreviated as $[Ad/Al]_{v-s}$ - can be used to determine a sample's history of decomposition (Hedges et al., 1988) and /or experience to fire (Kuo et al., 2008b). Both decomposition and fire tend to raise $[Ad/Al]_{v-s}$ ratio values. The significantly higher values (Table 5) in the paleosols and the smudge pit than in the modern soil are in line with reported increased values in low to mid-temperature charcoals, and support the interpretation of increased proportion of charcoal content in the paleosols and fire smudge pit feature (Kuo et al., 2008b).

The lignocellulose component of fire fuel material is often completely combusted in high temperature fires, but a small portion may remain in low temperature fires. Thus, LOPs can serve as an indirect fire biomarker, providing additional information on fire history, such as fire temperature. For instance, pyrogenic carbon derived from low temperature fires often contains higher lignin content due to incomplete thermal decomposition of lignocellulose components than pyrogenic C derived from high temperature fire (Kuo et al., 2008b). Our LOP data show much higher lignin content for the modern surface soil (sample 1) and significantly lower content for the paleosol samples (2-4) and the smudge pit feature (sample 5; Table 5). The higher lignin content of the modern surface soil indicates that a large portion of the OC in the modern surface soil was not fire-derived or only experienced minimal alteration from low temperature fire. The small amount of lignin in the paleosols indicates that the thermal decomposition of lignocellulose fragments in these samples was also incomplete, most likely due to the low to medium fire temperature (e.g. ca. 300-450 °C).

Free lignin phenol content ($\Sigma 8_{free,phenols}$) exhibited wide variability among the soil samples (Table 5), likely due to variation in fire conditions and in post-depositional alteration. The modern surface soil contained (1) contained the highest concentration of free lignin

phenols (93.6 ng/g, w/w whole soil), whereas the strongly melanized midden paleosol (2) contained the lowest concentration of free lignin phenols (2.1 ng/g). Unlike charcoal and anhydrosugars, which may be produced only via pyrogenic processes, free lignin phenols may have two different sources: thermal breakdown of lignin during combustion and biodegradation during post-depositional alteration. Thus, lignin free phenols alone cannot be used to understand the pyrogenic history of samples. Norwood et al. (2013) documented that these low temperature pyrogenic biomarkers are readily degraded when exposed to microorganisms (bacteria and fungi) in solution over a half-life of 3-7 days. In-situ degradation, along with variability in pyrogenic OM production temperature are potential mechanisms that could explain the wide variability in the free lignin phenols among the soil samples (cf. Kuo et al., 2011a; Norwood et al., 2013).

Although both the MMM method and LOP method ($\lambda 8$ values) provide data on soil lignin content, it is challenging to make direct comparison between the two methods. The lignin content from the MMM was much higher than the lignin content from the LOP method (except for the smudge pit feature; Tables 3 and 4). For instance, the MMM estimated that 12.3% of the total OC pool for the modern surface soil is derived from lignin (Table 3), which is twice that (5.5%) estimated from LOPs using the conversion of individual lignin oxidation phenols into units of OC (Louchouart et al., 1997; Table 5). There are several potential technical reasons that may contribute to the difference between the two approaches. First, the LOP method requires an initial CuO oxidation step followed by wet chemical extraction before GC quantification. Both steps may result in significant loss of the target compounds (as much as 75%) and thereby lead to underestimation of the total lignin content (Hedge and Ertel, 1982). Consistent with this, we generally observed more lignin in samples from CP NMR and MMM than from the LOPs method.

Nevertheless, both the LOPs analysis and the MMM showed higher lignin content for the modern surface soil (1) and significantly lower content for the paleosol samples (2 and 3). The MMM reported 0% lignin for the smudge pit sample, while LOPs reported a value in the same range as that observed for the other paleosols. The most unexpected result is the lignin content of the non-midden paleosol (4), where MMM and LOPs reported contrasting values. One possible explanation for the discrepancy is that lignin might be preferentially concentrated during the HF treatment for the NMR analysis (Hockaday et al., 2009). As

shown in Table 1, the OC concentration efficiency is significantly lower for the non-midden paleosol sample than for the other samples.

4. Conclusions

We used ^{13}C NMR spectroscopy to identify and quantify molecular charcoal fragments in a series of paleosol samples from a humid, subtropical environment where OM preservation was otherwise poor (Table 6). NMR is a non-destructive and relatively rapid technique that can provide coarse resolution information on the amounts of charcoal and other organic components in samples. We extended our analyses using three other techniques, measuring anhydrosugars, free lignin monomers and LOPs.

We used this multi-tracer integrated approach to test what human behavior or natural processes created the archaeological deposits observed. Each tracer provided slightly different information, which we coupled together to give a fuller picture of Caddo land use practices and the regional ecosystem. We showed that, even for soil with low OC, smudge pits can be identified from high C/N values and very high charcoal/OC fractions, as determined from NMR plus a molecular mixing model. The smudge pit contained 1.31% OM of which 67% was identified as charcoal. The midden samples were also very low in OC (0.25-0.51%) and were preferentially enriched in charcoal (33-39%), but did not show the high C/N values of the smudge pit, suggesting that these soils possibly received both charcoal and N-rich household waste. The presence of a high charcoal level in the paleosol sample not identified as a midden (sample 4) suggests that charcoal was mobile on this landscape, and may have been a ubiquitous feature of the Caddo people's environment.

LOP analysis helped identify the sources of the biomass used in fires, and pointed to burning of both woody and soft angiosperm tissues. Smudge pit and paleosol samples had a LOP signature richer in cinnamyl (C) units, suggesting an input of biomass from grassy angiosperm tissue (consistent with corn).

The NMR and anhydrosugar analysis provided information on the temperature conditions associated with thermal refuse production. In particular, both NMR and anhydrosugar results suggested that charcoal in the smudge pit was likely produced at higher temperature than the charcoals in the other paleosol samples, in contrast to our initial assumption that charcoal in the smudge pit was produced in a low temperature fire.

The results highlight the value of integrating organic geochemical methods of charcoal analysis with traditional archaeological approaches to identify midden constituents that cannot be macroscopically identified. Our initial assumption was that field melanization would be an indicator of charcoal content. This was shown not to be the case, whereby the non-midden sample (4) contained a relatively equal amount of charcoal to the midden samples (2 and 3). This leaves the question as to what drives the dark color of the midden if it is not charred OM.

Acknowledgements

This research was initially funded through a contract with Versar, Inc. and the Texas Department of Transportation (TxDOT) in partial fulfillment of Texas Antiquities Permit 5879. J. Abbott, W. Troell and S. Pletka of TxDOT and D. Peter of Versar all deserve thanks for their support of the project. C.A.M. and X.G. acknowledge the support of NSF awards EAR 0949337 and EAR 0911685 and all authors thank two anonymous reviewers for constructive comments.

Associate Editor – S. Derenne

Table titles

Table 1

Concentration of organic C and N in bulk soil samples and HF treated organic layer (BDL, below detection limit).

Table 2

^{13}C NMR signal distribution in the seven predefined chemical shift regions.

Table 3

The C observability (%) from CP NMR, biochemical composition (%) and alkyl/O-alkyl values of the OC in the soil samples.

Table 4

Concentration of anhydrosugars.

Table 5

Concentration of lignin free phenols and lignin oxidation products.

Table 6

Summary of results vs. expectation.

Figure captions

Fig. 1. Map of Murvaul Creek site.

Fig. 2. Generalized profile of stratigraphic locations of the samples.

Fig. 3. ^{13}C CP NMR spectra of soil samples from Murvaul Creek site.

References

- Aleman, L., Grant, D., Pugmire, R., Alger, T., Zilm, K., 1983. Cross polarization and magic angle sample spinning NMR-spectra of model-organic compounds .2. Molecules of low or remote protonation. *Journal of the American Chemical Society* 105, 2142-2147.
- Andreae, M.L., Merlet, P., 2001. Emission of trace gases and aerosols from biomass burning. *Global Biogeochemistry Cycles* 15, 955-966.
- Ascough, P.L., Bird, M.I., Francis, S.M., Thornton, B., Midwood, A.J., Scott, A.C., Apperley, D., 2011. Variability in oxidative degradation of charcoal: Influence of production conditions and environmental exposure. *Geochimica et Cosmochimica Acta* 75, 2361-2378.
- Baldock, J.A., Masiello, C.A., Gelinas, Y., Hedges, J.I., 2004. Cycling and composition of organic matter in terrestrial and marine ecosystems. *Marine Chemistry* 92, 39-64.
- Baldock, J.A., Oades, J.M., Nelson, P.N., Skene, T.M., Golchin, A., Clarke, P., 1997. Assessing the extent of decomposition of natural organic materials using solid-state ¹³C NMR spectroscopy. *Australian Journal of Soil Research* 35, 1061-1083.
- Bateman, M.D., 2012. Quartz optical dating report, 18 July 2012, 41PN175, East Texas. Sheffield Centre for International Drylands Research, Department of Geography, The University of Sheffield, UK.
- Beck, M.E., 2007. Midden formation and intrasite chemical patterning in Kalinga, Philippines. *Geoarchaeology* 22, 453-475.
- Binford, L.R., 1967. Smudge pits and hide smoking: The use of analogy in archaeological reasoning. *American Antiquity* 32, 1-12.
- Boone, J.L., 1987. Defining and measuring midden catchment. *American Antiquity* 52, 336-345.
- Braadbaart, F., Poole, I., van Brussel, A.A., 2009. Preservation potential of charcoal in alkaline environments: an experimental approach and implications for the archaeological record. *Journal of Archaeological Science* 36, 1672-1679.
- Cheng, C.H., Lehmann, J., Engelhard, M.H., 2008. Natural oxidation of black carbon in soils: Changes in molecular form and surface charge along a climosequence. *Geochimica et Cosmochimica Acta* 72, 1598-1610.
- Cohen-Ofri, I., Weiner, L., Boaretto, E., Mintz, G., Weiner, S., 2006. Modern and fossil charcoal: aspects of structure and diagenesis. *Journal of Archaeological Science* 33, 428-439.
- Cortizas, A.M., Mighall, T., Pomal, X.P., Munoz, J.C.N., Varela, E.P., Rebol, R.P., 2005. Linking changes in atmospheric dust deposition, vegetation change and human activities in northwest Spain during the last 5300 years. *Holocene* 15, 698-706.
- Eberl, M., Álvarez, M., Terry, R.E., 2012. Chemical signatures of middens at a late classic Maya residential complex, Guatemala. *Geoarchaeology* 27, 426-440.
- Freitas, J.C.C., Bonagamba, T.J., Emmerich, F.G., 1999. ¹³C high-resolution solid-state NMR study of peat carbonization. *Energy and Fuels* 13, 53-59.
- Freitas, J.C.C., Bonagamba, T.J., Emmerich, F.G., 2001. Investigation of biomass- and polymer-based carbon materials using ¹³C high-resolution solid-state NMR. *Carbon* 39, 535-545.
- Glaser, B., Haumaier, L., Guggenberger, G., Zech, W., 1998. Black carbon in soils: the use of benzene carboxylic acids as specific markers. *Organic Geochemistry* 29, 811-819.

- Goñi, M.A., Hedges, J.I., 1992. Lignin dimers; structures, distribution, and potential geochemical applications. *Geochimica et Cosmochimica Acta* 56, 4025-4043.
- Harris, D., Horwath, W.R., van Kessel, C., 2001. Acid fumigation of soils to remove carbonates prior to total organic carbon or carbon-13 isotopic analysis. *Soil Science Society of America Journal* 65, 1853-1856.
- Harvey, O.R., Herbert, B.E., Kuo, L.-J., Louchouart, P., 2012. Generalized two-dimensional perturbation correlation infrared spectroscopy reveals mechanisms for the development of surface charge and recalcitrance in plant-derived biochars. *Environmental Science & Technology* 46, 10641-10650.
- Hays, C.T., Weinstein, R.A., 2010. Tchefuncte? and early woodland. In: Rees, M.A. (Ed.), *Archaeology of Louisiana*. Louisiana State University Press, Baton Rouge, Louisiana, pp. 97-134.
- Hedges, J.I., Blanchette, R.A., Weliky, K., Devol, A.H., 1988. Effects of fungal degradation on the CuO oxidation products of lignin: A controlled laboratory study. *Geochimica et Cosmochimica Acta* 52, 2717-2726.
- Hedges, J.I., Eglinton, G., Hatcher, P.G., Kirchman, D.L., Arnosti, C., Derenne, S., Evershed, R.P., Kögel-Knabner, I., de Leeuw, J.W., Littke, R., Michaelis, W., Rullkotter, J., 2000. The molecularly-uncharacterized component of nonliving organic matter in natural environments. *Organic Geochemistry* 31, 945-958.
- Hedges, J.I., Ertel, J.R., 1982. Characterization of lignin by gas capillary chromatography of cupric oxide oxidation products. *Analytical Chemistry* 54, 174-178.
- Hedges, J.I., Mann, D., 1979. The characterization of plant tissues by their lignin oxidation products. *Geochimica et Cosmochimica Acta* 43, 1803-1807.
- Hockaday, W.C., Masiello, C.A., Randerson, J.T., Smernik, R.J., Baldock, J.A., Chadwick, O.A., Harden, J.W., 2009. Measurement of soil carbon oxidation state and oxidative ratio by ¹³C nuclear magnetic resonance. *Journal of Geophysical Research* 114, G02014.
- Holliday, V.T., Gartner, W.G., 2007. Methods of soil P analysis in archaeology. *Journal of Archaeological Science* 34, 301-333.
- Hopmans, E.C., Lopes dos Santos, R.A., Mets, A., Sinninghe Damsté, J.S., Schouten, A., 2013. A novel method for the rapid analysis of levoglucosan in soils and sediments. *Organic Geochemistry* 58, 86-88.
- Jackson, H.E., Scott, S.L., Schambach, F.F., 2010. At the house of the priest. In: Perttula, T.K., Walker, C.P. (Eds.), *The Archaeology of the Caddo*. University of Nebraska Press, Lincoln, Nebraska, pp. 47-85.
- Kaal, J., Baldock, J.A., Buurman, P., Nierop, K.G.J., Pontevedra-Pombal, X., Martínez-Cortizas, A., 2007. Evaluating pyrolysis-GC/MS and ¹³C CPMAS NMR in conjunction with a molecular mixing model of the Penido Vello peat deposit, NW Spain. *Organic Geochemistry* 38, 1097-1111.
- Kaal, J., Martínez-Cortizas, A., Nierop, K.G.J., Buurman, P., 2008a. A detailed pyrolysis-GC/MS analysis of a black carbon-rich acidic colluvial soil (Atlantic ranker) from NW Spain. *Applied Geochemistry* 23, 2395-2405.
- Kaal, J., Brodowski, S., Baldock, J.A., Nierop, K.G.J., Martínez-Cortizas, A., 2008b. Characterization of aged black carbon using pyrolysis-GC/MS, thermally assisted hydrolysis and methylation (THM), direct and cross-polarization ¹³C nuclear

749 magnetic resonance (DP/CP NMR) and the benzene polycarboxylic acid (BPCA)
 750 method. *Organic Geochemistry* 39, 1415-1426.
 751 Keiluweit, M., Nico, P.S., Johnson, M.G., Kleber, M., 2010. Dynamic molecular structure of
 752 plant biomass-derived black carbon (biochar). *Environmental Science & Technology*
 753 44, 1247-1253.
 754 Kleber, M., Rossner, J., Chenu, C., Glaser, B., Knicker, H., Jahn, R., 2003. Prehistoric
 755 alteration of soil properties in a Central German chernozemic soil: In search of
 756 pedologic indicators for prehistoric activity. *Soil Science* 168, 292-306.
 757 Knicker, H., Hilscher, A., de la Rosa, J.M., Gonzalez-Perez, J.A., Gonzalez-Vila, F., 2013.
 758 Modification of biomarkers in pyrogenic organic matter during the initial phase of
 759 charcoal biodegradation in soils. *Geoderma* 197, 43-50.
 760 Knight, V.J., Jr., 2004. Characterizing elite midden deposits at Moundville. *American*
 761 *Antiquity* 69, 304-321.
 762 Kuo, L.J., Herbert, B.E., Louchouart, P., 2008a. Can levoglucosan be used to characterize
 763 and quantify char/charcoal black carbon in environmental media? *Organic*
 764 *Geochemistry* 39, 1466-1478.
 765 Kuo, L.-J., Louchouart, P., Herbert, B.E., 2008b. Fate of CuO-derived lignin oxidation
 766 products during plant combustion: Application to the evaluation of char inputs to soil
 767 organic matter. *Organic Geochemistry* 39, 1522-1536.
 768 Kuo, L.-J., Louchouart, P., Herbert, B.E., 2011a. Influence of combustion conditions on
 769 yields of solvent-extractable anhydrosugars and lignin phenols in chars: Implications
 770 for characterizations of biomass combustion residues. *Chemosphere* 84, 797-805.
 771 Kuo, L.-J., Louchouart, P., Herbert, B.E., Brandenberger, J., Wade, T., Crecelius, E., 2011b.
 772 Combustion-derived substances in deep basins of the Puget Sound: Historical inputs
 773 from fossil fuel and biomass combustion. *Environmental Pollution* 159, 983-990.
 774 Louchouart, P., Amon, R., Duan, S., Pondell, C., Seward, S.M., White, N., 2010. Analysis of
 775 lignin-derived phenols in standard reference materials and dissolved organic matter
 776 by gas chromatography/tandem mass spectrometry. *Marine Chemistry* 118, 85-97.
 777 Louchouart, P., Kuo, L.J., Wade, T.L., Schantz, M., 2009. Determination of levoglucosan
 778 and its isomers in size fractions of aerosol standard reference materials. *Atmospheric*
 779 *Environment* 43, 5630-5636.
 780 Masiello, C.A., 2004. New directions in black carbon organic geochemistry. *Marine*
 781 *Chemistry* 92, 201-213.
 782 McBeatha, A.V., Smernika, R.J., Krull, E.S., Lehmann, J., 2014. The influence of feedstock
 783 and production temperature on biochar carbon chemistry: A solid-state ¹³C NMR
 784 study. *Biomass and Bioenergy* 60, 121-129.
 785 McKee, A., Frederick, C.D., Perttula, T.K., Selden, R.Z., Bush, L., Kemp, L., Gregory, B.,
 786 Yost, C., Ferguson, J.R., Glascock, M.D., Tomka, S., Cecil, L., Masiello, C., Gao, X.,
 787 Goodmaster, C., Beasley III, V., 2014. Data Recovery Investigations: Murvaul Creek
 788 Site (41PN175), Panola County, Texas. (Draft). Miscellaneous Reports of
 789 Investigations Number 598. Versar, Inc., Plano, Texas.
 790 Metz, G., Wu, X., Smith, S.O., 1994. Ramped-amplitude cross polarization in magic-angle-
 791 spinning NMR. *Journal of Magnetic Resonance, Series A* 110, 219-227.
 792 Misarti, N., Finney, B. P., Maschner, H., 2011. Reconstructing site organization in the
 793 eastern Aleutian Islands, Alaska, using multi-element chemical analysis of soils.
 794 *Journal of Archaeological Science* 38, 1441-1455.

- Myers-Pigg, A.N., Louchouart, P., Amon, R.M.W., Prokushkin, A., Pierce, K., Rubtsov, A.,
2015. Labile pyrogenic dissolved organic carbon in major Siberian Arctic rivers:
Implications for wildfire-stream metabolic linkages. *Geophysical Research Letters*
DOI: 10.1002/2014GL062762.
- Newcomb, W.W., Jr., 1961. *The Indians of Texas: From Prehistoric to Modern Times*.
University of Texas Press, Austin, TX.
- Norwood, M.J., Louchouart, P., Kuo, L.-J., Harvey, O.R., 2013. Characterization and
biodegradation of water-soluble biomarkers and organic carbon extracted from low
temperature chars. *Organic Geochemistry* 56, 111-119.
- Perttula, T.K., 2004. The prehistoric and Caddoan archeology of the northeastern Texas
pineywoods. In: Perttula, T.K. (Ed.), *The Prehistory of Texas*. Texas A&M
University Press, College Station, Texas, pp. 370-407.
- Preston, C.M., Schmidt, M.W.I., 2006. Black (pyrogenic) carbon: a synthesis of current
knowledge and uncertainties with special consideration of boreal
regions. *Biogeosciences* 3, 397-420.
- Schillito, L., Matthews, W., 2013. Geoarchaeological investigations of midden-formation
processes in the early to late ceramic Neolithic levels at Çatalhöyük, Turkey ca.
8550–8370 cal BP. *Geoarchaeology: An International Journal* 28, 25-49.
- Schmid, E.M., Knicker, H., Baumler, R., Kogel-Knabner, I., 2001. Chemical composition of
the organic matter in neolithic soil material as revealed by CPMAS ¹³C NMR
spectroscopy, polysaccharide analysis, and CuO oxidation. *Soil Science* 166, 569-584.
- Schneider, M.P.W., Smittenberg, R., Dittmar, T., Schmidt, M.W.I., 2011. Comparison of gas
with liquid chromatography for the determination of benzene polycarboxylic acids as
molecular tracers of black carbon. *Organic Geochemistry* 42, 275-282.
- Smernik, R.J., Oades, J., 2000. The use of spin counting for determining quantitation in solid
state ¹³C NMR spectra of natural organic matter: 2. HF-treated soil fractions.
Geoderma 96, 159-171.
- Smernik, R.J., Baldock, J.A., Oades, J.M., 2002. Impact of remote protonation on ¹³C
CPMAS NMR quantitation of charred and uncharred wood. *Solid State Nuclear
Magnetic Resonance* 22, 71-82.
- Simoneit, B.R.T., 1999. A review of biomarker compounds as source indicators and tracers
for air pollution. *Environmental Science and Pollution Research* 6, 159-169.
- Singh, N., Abiven, S., Torn, M.S., Schmidt, M.W.I., 2012. Fire-derived organic carbon in
soil turns over on a centennial scale. *Biogeosciences* 9, 2847-2857.
- Spokas, K.A., Novak, J.M., Masiello, C.A., Johnson, M.G., Colosky, E.C., Ippolito, J.A.,
Trigo, C., 2014. Physical disintegration of biochar: An overlooked process.
Environmental Science & Technology Letters 1, 326–332.
- Swanton, J.R., 1996. *Source Material on the History and Ethnology of the Caddo Indians*.
University of Oklahoma Press, Norman, OK.
- Wiedemeier, D.B., Abiven, S., Hockaday, W.C., Keiluweit, M., Kleber, M., Masiello, C.A.,
et al., 2015. Aromaticity and degree of aromatic condensation of char. *Organic
Geochemistry* 78, 135-143.
- Wilson, C.A., Davidson, D.A., Cresser, M.S., 2009. An evaluation of the site specificity of
soil elemental signatures for identifying and interpreting former functional
areas. *Journal of Archaeological Science* 36, 2327-2334.

840 Zimmerman, A.R., Gao, B., Ahn, M.-Y., 2011. Positive and negative carbon mineralization
841 priming effects among a variety of biochar-amended soils. *Soil Biology and*
842 *Biochemistry* 43, 1169-1179.
843

844

845 Figure 1

846

847

848
849 Figure 2

850
851
852
853
854 Figure 3
855

856
857
858
859
860
861
862

863 **Tables**864 **Table 1**

Samples	Bulk soil			HF treated		
	OC (%)	N (%)	C/N	OC (%)	N (%)	C/N
1. Modern surface soil	0.51	0.04	12.8	7.28 ± 0.64	0.47 ± 0.05	15.5
2. Midden paleosol	0.42	0.03	14.0	10.98 ± 0.47	0.57 ± 0.04	19.3
3. Midden paleosol	0.25	BDL *		8.97 ± 0.75	0.53 ± 0.01	16.9
4. Non-midden paleosol	0.38	0.02	19.0	3.50 ± 0.40	0.13 ± 0.02	28.0
5. Smudge pit feature	1.31	0.02	65.5	20.47 ± 0.25	0.31 ± 0.02	66.0

865

866

867 **Table 2**

Sample	Chemical shift region (ppm)							Aromatic: phenolic
	0-45	45-60	60-95	95-110	110-145	145-165	165-215	
	Alkyl	N-alkyl/methoxyl	O-alkyl	Di-O-alkyl	Aromatic	Phenolic	Amide/carboxyl	
1. Modern surface soil	30.1	7.7	27.4	6.6	13.4	4.3	10.6	3.1
2. Midden paleosol	13.1	5.1	18.0	5.8	35.6	7.8	14.3	4.6
3. Midden paleosol	17.1	6.3	18.3	5.3	31.8	8.5	12.6	3.7
4. Non-midden paleosol	12.4	4.8	15.3	7.3	37.1	10.9	12.2	3.4
5. Smudge pit feature	5.4	2.2	7.3	3.6	57.2	9.5	14.8	6.0

868

869

870

871

872

873

874

875

876

877 **Table 3**

Sample	C observability (%)	Carbohydrate (%)	Protein (%)	Lignin (%)	Lipid (%)	Carbonyl (%)	Char (%)	Alkyl/:O-alkyl
1. Modern surface soil	57.9	33.7	18.2	12.3	23.5	4.3	8.0	1.10
2. Midden paleosol	33.4	22.5	14.1	2.2	7.7	14.4	39.1	0.74
3. Midden paleosol	38.0	21.7	16.6	9.4	10.3	9.6	32.5	0.93
4. Non-midden paleosol	27.6	17.1	10.3	16.4	6.8	12.3	37.0	0.81
5. Smudge pit feature	21.9	6.4	4.2	0.0	3.3	19.5	66.6	0.97

878

879 **Table 4**

Sample	Galactosan (ng/g)	Mannosan (ng/g)	Levoglucosan (ng/g)	L/M	L/G	L/M+G	G/M
1. Modern surface soil	419.9	35.7	34.4	1.2	0.1	0.1	10.9
2. Midden paleosol	328.2	27.4	28.4	1.4	0.1	0.1	10.9
3. Midden paleosol	214.9	46.5	96.1	2.2	0.5	0.4	4.4
4. Non-midden paleosol	150.1	29.7	63.1	2.4	0.5	0.4	4.7
5. Smudge pit feature	141.8	11.3	13.2	1.9	0.2	0.2	10.1

880
881
882 **Table 5**

Sample	Free phenols	LOPs (CuO)										
	$\Sigma\delta_{\text{free.phenols}}$	$\lambda 8$	$\Sigma\delta_{\text{LOPs}}$	V	S	C	P	S/V	C/V	(Ad/Al) _v	(Ad/Al) _s	%Corg-phen ^a
	(ng/g)	(μg 100 mg/OC)	($\mu\text{g/g}$)	($\mu\text{g/g}$)	($\mu\text{g/g}$)	($\mu\text{g/g}$)	($\mu\text{g/g}$)					
1. Modern surface soil	93.6	3893	198.6	71.1	118.4	9.1	5.7	1.7	0.1	0.5	0.2	5.4
2. Midden paleosol	2.1	127.4	5.4	2.6	1.9	0.9	1.9	0.7	0.3	1.2	0.3	0.2
3. Midden paleosol	27.7	147.0	3.7	1.7	1.2	0.8	1.9	0.7	0.5	4.5	0.6	0.2
4. Non-midden paleosol	9.1	86.7	3.3	1.7	1.1	0.5	1.3	0.6	0.3	1.7	0.5	0.1
5. Smudge pit feature	26.0	118.9	15.6	7.1	6.0	2.5	4.3	0.8	0.4	1.0	0.5	0.1

883
884 ^a Estimated fraction of total OC pool represented by lignin according to Louchouart et al. (1997).

885 **Table 6**

Context	Expectation	¹³ C NMR	Anhydrosugar	Lignin free phenols and LOPs
Modern soil	Humified OM, little/no charcoal	Small amount of charcoal (8%); high lignin content	Significant amount of levoglucosan	Highest amount of lignin free phenols and LOPs, vegetation dominated by woody angiosperm
Non-midden paleosol	Humified OM together with small amount of pyrogenic OM/charcoal from anthropogenic and natural sources	High content of charcoal and lignin	Significant amount of levoglucosan	Small amount of lignin free phenols and LOPs, vegetation dominated by non-woody angiosperm
Midden	Mixture of thermal and humified OM, largely from anthropogenic sources	High content of charcoal, low content of lignin	Significant amount of levoglucosan	Small amount of lignin free phenols and LOPs, vegetation dominated by non-woody angiosperm
Smudge Pit	Dominantly thermally altered OM/charcoal	Highest content of charcoal, no lignin	Lowest amount of levoglucosan	Small amount of lignin free phenols and LOPs, vegetation dominated by non-woody angiosperm

886

887

888

889

890

891

892

893

894

895

896

897

898

899

900

901

902

903