Summary Overview

Microbial clocks for estimating the postmortem interval of human remains at three anthropological research facilities.

Jessica L. Metcalf[1,2], Rob Knight[2], David O. Carter[3], Franklin Damann[4]

1 – University of Colorado, Boulder, CO

2 - University of California San Diego, San Diego, CA

3 - Chaminade University, Honolulu, HI

4 – Defense POW/MIA Accounting Agency, Omaha, Nebraska

Key Collaborators: Sibyl Bucheli, Aaron Lynne, Melissa Conner, Dawnie Steadman, Giovanna Vidoli

Award number 2015-DN-BX-K016

GENERAL SUMMARY

Estimating the postmortem interval (PMI) is a critical step in many crime scene investigations. Current methods, such as those that rely on insects or chemical decay products, are limited because they may only work for a short duration, are specific to particular geographic regions, are not very accurate, or a combination of these factors. In our previous NIJ award (2011-DN-BX-K533), we demonstrated that microbes provide an accurate clock that starts at death and relies on ecological change in the microbial communities that normally inhabit a body and its surrounding environment (e.g., gravesoil). Here, we build on these findings in several ways that are beneficial to crime scene investigation with experiments using donated human bodies at forensic anthropology facilities and generating rapid, inexpensive amplicon sequencing data (16S rRNA and 18S rRNA genes). We determined that the microbial decomposition clock of corpse skin and gravesoil ticks at a similar rate in different geographic regions and to some extent in different seasons, if temperature is included in the model. Furthermore, by examining the microbes that invade bones after an extended decay period, we discovered that bone microbial community changes can be used to estimate PMI in bodies over a longer period of decay. This project brought together a multi-disciplinary team of scientists and the exploitation of advances in DNA sequencing technology that allow characterization of microbial communities at one-millionth of what the project would have cost a decade ago. A stakeholder meeting held at the American Academy of Forensic Sciences meeting in 2019 addressed how this new technology could be implemented into the justice system to aid practitioners in determining PMI at death scenes.

Purpose

Our overarching goals were to test the effects of regional location and seasonality on microbial-based estimates of the postmortem interval (PMI), and to further explore the

microbiome of decaying human bone to estimate PMI during the extended postmortem period. This project met NIJ's Applied Research Goal to improve our understanding of microbial evidence by conducting a highly collaborative three-phase project coordinated across three anthropological research facilities.

Goal 1: Determine whether changes in skin and gravesoil microbial communities during decomposition are similar at three anthropological research facilities that are located in different geographic regions.

Goal 2: Determine whether seasonal variations in microbial communities associated with decomposing human cadavers are consistent among the three geographic regions.

Goal 3: Determine whether microbial community change is useful for estimating PMI after Active Decay (extended PMI) by sampling human bone at a single anthropological research facility in Texas (SHSU).

Project design and methods

Skin and gravesoil sampling for Goals 1 & 2: Donated human remains were placed at three anthropological research facilities in environmentally distinct regions of the United States. The Colorado Mesa University Forensic Investigation Research Station (Grand Junction, CO; CMU) is located in a high-altitude arid environment, the Sam Houston State University Applied Anatomical Research Center (Huntsville, TX; SHSU) is located in a piney woods ecoregion with a humid subtropical climate and sparse forest covering, and the University of Tennessee Knoxville Forensic Anthropology Center (Knoxville, TN; UTK) is located in a temperate region naturally forested with oak, maple, and hickory trees.

During each season, three donated human remains were placed at each facility. In some cases, three non-autopsied, never frozen, fresh (not decomposed) bodies were not available and it

took two seasons/years to collect samples. All remains were obtained through willed-body donation programs at the facilities and placed outside and allowed to decompose. The remains were sampled daily for 21 days beginning on the day of placement. Samples were collected using sterile dual SWUBE swabs (BD; Franklin Lakes, NJ) taken from the skin of the face, skin of the hip, soil associated with the face, and soil associated with the hip, as well as soil from a control plot that was not associated with human remains. Swabs were frozen immediately after sampling and were transported to the University of Colorado (Boulder, CO) at the conclusion of the sampling period for DNA extraction. During the sampling period, additional metadata were measured and collected, including environmental data (temperature, humidity, soil temperature) and the condition of the remains (cause of death, initial body condition, Megyesi total body score, maggot presence, scavenging activity). In order to compare remains that were placed on a different day and in different locations, accumulated degree day (ADD) was used as a proxy for day to describe the postmortem interval (PMI). Accumulated degree day was calculated using a base of 0°C, where: Degree Day (DD) = ((max temp + min temp / 2) – base temperature) and ADD = $(DD_x) + (DD_{x+1})$, where x = 24h period.

Rib bone sampling for Goal 3:

Cadavers placed in the spring and summer at SHSU in Huntsville, TX were allowed to continue decomposing. Collection of rib bones, following Damann et al. (2015), began once decomposition progressed enough so that little dissection was needed. Right and left lower ribs were randomly selected and collected approximately every three weeks for a total of eight bones from each body (48 overall), with one exception, in which two ribs were mistakenly collected from the same cadaver, resulting in one subject with nine time-points and another subject with

only seven time-points. Each rib was individually bagged and immediately frozen at -10°C, then stored until shipping to Colorado State University for processing.

DNA Extraction for skin and gravesoils: DNA extraction was conducted at the University of Colorado (Boulder, CO) following Earth Microbiome Project (EMP) protocols, which can be accessed on the EMP website (www.earthmicrobiome.org/protocols-and-standards) (Thompson et al. 2017). DNA was extracted from one of the two swabs taken at each sample site using the PowerSoil DNA Extraction Kit (Qiagen; Hilden, Germany) following manufacturer instructions.

DNA Extraction for Bone samples: Rib bones were sampled for DNA extraction at Colorado State University. A fume hood was cleaned with 20% bleach solution before processing and between each bone sample. Each rib was mechanically abraded with a hand-held Dremel® Drill to remove any tissue and superficial layers of cortical bone. An approximate 40 mm x 15 mm section of bone was removed from the rib angle, as indicated in **Figure 1**. The samples were weighed and ultraviolet irradiated at 254 nm for 30 minutes on each side. Each sample was wiped down with 3% bleach, then abraded again with the Dremel Drill to ensure removal of the outer layer of bone. The sample was then divided into three equal segments, each of which were weighed and placed into a tube. Two segments were frozen for future use while the remaining sample was pulverized in a sterile Waring MC2 blender cup. The cup was washed and soaked in bleach for 3 minutes between each sample. Each of the bone powders were placed into a clean tube for extraction. DNA was extracted from 0.2-0.5 grams of pulverized bone. First, bone powder was demineralized and lysed using 30 µl of 10% sodium dodecyl sulfate (SDS), 20 µl proteinase K, and 500 µl 0.5 M ethylenediaminetetraacetic acid (EDTA) (Loreille et al. 2007). The samples were vortexed for 2 seconds and placed on a heating block at 55°C for 1 hour, with additional 2-second vortexes every 15 minutes. The lysed samples were centrifuged at 10,000 x g

for 1 minute at room temperature. The supernatant was removed, measured, and placed into a clean tube. The pellets were kept and frozen at -20°C. Fifteen extraction blanks were included to identify any potential contamination. DNA was purified using the PowerSoil DNA Isolation Kit from MoBio (Carlsbad, CA) with a modified protocol.

Figure 1: Rib bone sampling example.

Following extraction, samples were transported to the Knight Lab at San Diego State University (La Jolla, CA) for library preparation and sequencing. The total genomic DNA was amplified to target informative amplicons using PCR.

Microbial DNA sequencing: DNA was shipped to UC San Diego for next-generation sequencing. Bacterial, archaeal, and microbial eukaryote amplicons were generated following the EMP protocols (press.igsb.anl.gov/earthmicrobiome/protocols-and-standards). For both 16S and 18S rRNA gene amplicons, PCR products were quantified using Picogreen Quant-iT (Invitrogen, Life Technologies, Grand Island, NY) and pooled at equal concentrations. No-template controls were also included in the pools for each amplicon type. Each amplicon pool was purified using the UltraClean PCR Clean-up Kit (Qiagen; Hilden, Germany). Amplicon pools were sequenced using a 300 cycle kit the Illumina MiSeq sequencing platform (Illumina, San Diego, CA).

Data analysis

Machine Learning Model Construction

To determine whether the microbial community composition could predict PMI, informative models were constructed using the machine learning algorithm, Random Forest Regression. In the construction of these models, the microbial data was used as the predictor and ADD was used as the response. Data were normalized using total-sum scaling to avoid the loss of statistical

power by discarding reads and samples using the Calour library (Xu et al. 2019). Models were constructed with the Random Forest Regression algorithm with k-fold cross-validation. Cross-validation was conducted during model training by individual host (remains) so that data from the same host was included in either the training or testing set, but not both. The accuracy of these models was assessed during cross-validation and measured using the mean absolute error (MAE), calculated as the deviation of the predicted from observed values and represented in the same unit as the original data (ADD). The models with the lowest error after hyperparameter tuning were considered the most accurate. This method was applied to subsets of the data as was necessary to answer specific research questions.

Models were generated for samples taken at each facility (CMU, SHSU, UTK) and for all facilities combined (general) within each as well as for each season and all seasons combined (general). These models were constructed using microbiome data at several levels of taxonomic resolution (ASV, species, genus, family, order, class, phylum) to determine which resolution resulted in the most accurate model at each facility. Once models were generated within facility, the generalizability of these models was assessed by testing the model on the data from the other facilities or other seasons (i.e. cross-facility prediction). Once the best resolution was selected for each amplicon type, the data were merged to create a single dataset. Recursive feature elimination was used to determine which features contributed to the most accurate final model. All modeling was done using the python machine learning package scikit-learn v19.0 (Pedregosa et al. 2011). Results were visualized using R software, version 3.5.1, ggplot2, and matplotlib 2.0.0 (Wickham 2009; Droettboom et al. 2017).

Findings

Skin and gravesoil for predicting PMI across seasons and geography (Goals 1 & 2)

This study confirmed the findings of previous studies that demonstrated the predictable patterns of microbial community succession during PMI. The microbial communities shifted as PMI increased, though slightly differently at each facility. The microbial diversity associated with remains placed at SHSU and UTK proceeded similarly to each other, while those at CMU were more distinct. As a result, we approached model construction by first constructing a model to predict PMI using data from each facility (within-facility models). These models resulted in error rates that varied across seasons with spring season data resulting in the lowest errors after estimating chronological days from ADD (**Figure 2**). To determine the generalizability of data from one facility to other locations, we tested the model from a single facility on data from the other two facilities. These errors were much higher than the within-facility errors, which may be due to the differences in the overall temperature since the response variable was temperature based. For example, the highest error rate was associated with applying the CMU spring model, for which the highest ADD was 214, to SHSU spring data, which had a maximum ADD of 450, thus forcing the model to extrapolate far beyond the temperature range from which it was trained. Following this, we constructed a general model using data from all three facilities to train the model. This resulted in errors similar to, but slightly higher than, within facility models. Overall, these results suggest that models can be constructed to predict PMI across multiple geographic regions, and that a limiting factor in generating a robust model is sample size across a broad temperature range for the training model. We suspect that including more high-quality data sets in model training will result in more accurate predictions of test samples, at least within a season.

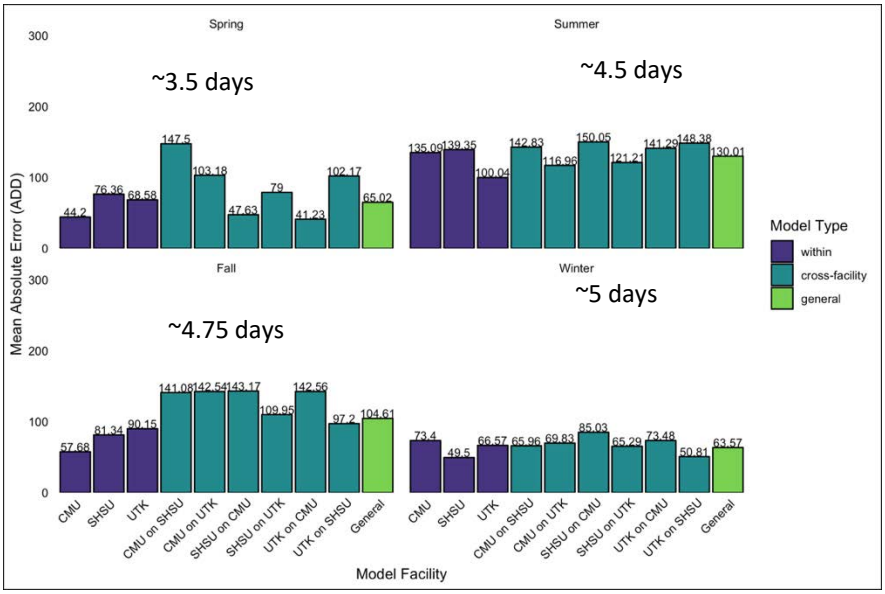

Figure 2. Mean Absolute Error (MAE) in units of Accumulated Degree Day (ADD) for PMI models by season and constructed in three different ways: a general model (included samples from bodies at all three locations in the training data and samples from a different set of bodies at all three facilities in the testing data), within facility models (included 2 bodies from a facility to train the model, and samples from 1 body to test the model), and cross-facility models (tests each within facility model with samples from a body from a different facility). We include the approximate number of chronological days represented by the lowest ADD for each season.

Next, we included data from all seasons and facilities into a single model, resulting in a general model of 110.85 ADD, or about 7 chronological days (**Figure 3**). All within facility models using data from all seasons generated MAEs higher than within season models shown in **Figure 2**. We are currently testing whether additional metadata such as humidity may help overcome the higher errors associated with PMI models generalized across seasons.

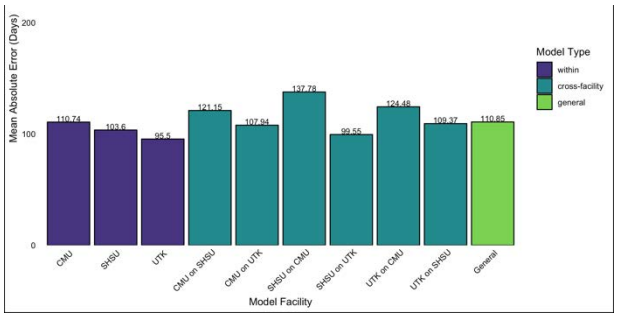

Figure 3. MAE in units of ADD for PMI models including data across four seasons.

Rib bone for predicting PMI during extended decomposition (Goal 3):

Rib bone alpha diversity analysis using Faith's Phylogenetic Diversity Index revealed that an increase in richness is significantly correlated with higher accumulated degree days for both the bacterial and archaeal communities as well as microbial eukaryotic communities ($p <$

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

0.05, **Figure 4**). Microbial communities become more diverse during decomposition, presumably as more environmental microbes are able to invade the bone over time.

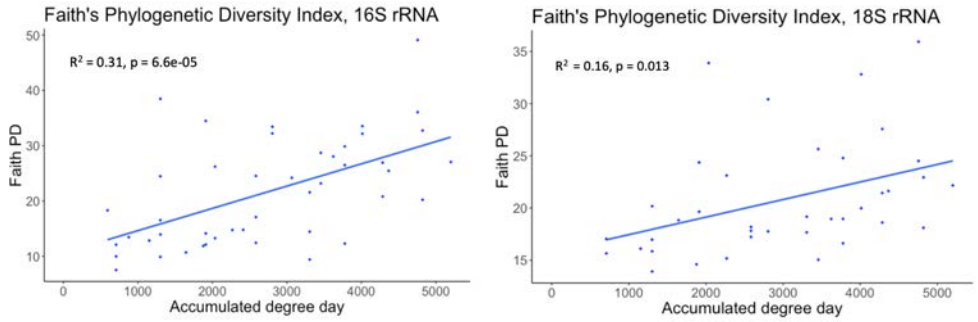

Figure 4. Figure 2: A measure of alpha diversity using Faith's Phylogenetic Diversity Index with increasing ADD for 16S rRNA and 18S rRNA datasets.

The model with the lowest error, an MAE of approximately 36 days, following random forest regression was the general model, which included 16S rRNA data and combined data from both summer and spring seasons (**Figure 5**). Modeling of the 18S rRNA data resulted in slightly higher errors than 16S rRNA data, with a general model error of approximately 43 days (**Figure 5**). Combined 16S rRNA and 18S rRNA models gave a general model error of approximately 39 days. These results are very encouraging given that sampling frequency was only every 21 days, making a low error unlikely. Expanding this study to include additional bodies and more frequent sampling will help us understand how accurate this tool could be for predicting extended postmortem intervals.

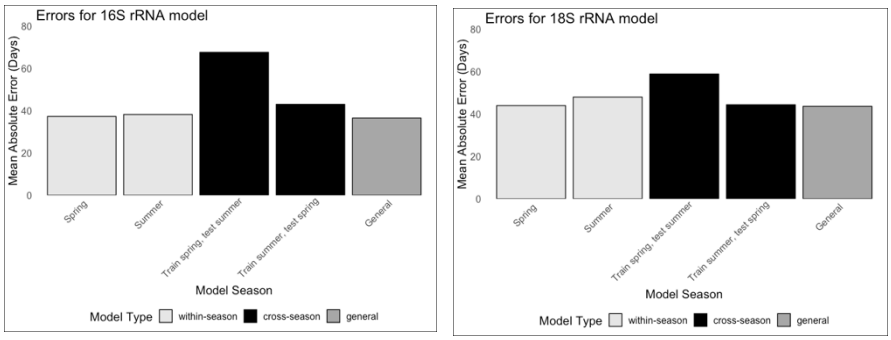

Figure 5. Mean absolute error (MAE) based on 16S rRNA (left) and 18S rRNA (right) Random Forests regression analysis to predict PMI of rib bones. Error for the best model, the 16S rRNA general model, was approximately 36 days. ADD was converted to chronological days.

Implications for criminal justice policy and practice

The projects funded by this award further demonstrate the potential of microbiome tools for estimating PMI in outdoor, terrestrial death scenes. These findings bring microbiome

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

technology closer to being used to solve crimes. On February 18, 2019, PI Metcalf led and organized a meeting entitled "Stakeholder meeting for estimating the postmortem interval using microbes" in collaboration with Co-PIs David O. Carter, Rob Knight, and the Forensic Technology Center of Excellence (https://forensiccoe.org/event/stakeholder-meeting-pmi-microbes-postmortem-interval-2-2019/). Stakeholders included researchers, forensic scientists, lawyers, pathology experts, and practitioners. The goals were to understand recent research and gaps in microbiome technology for PMI and to envision the steps for adopting this type of forensic technology into casework. Future steps will require coordination between these different stakeholder groups.

PRODUCTS

Additional publications are forthcoming.

Metcalf, JL. 2019 Estimating the Postmortem Interval Using Microbes: Knowledge Gaps and a Path to Technology Adoption. Forensic Science International. Genetics 38 (January): 211–18. https://www.fsigenetics.com/article/S1872-4973(18)30403-4/fulltext

Deel HL, Bucheli S, Belk AD, Ogden S, Lynne A, Carter DO, Knight R, Metcalf JL. (2019). Using microbiome tools for estimating the postmortem interval.
B. Budowle, S. Morse, Schutzer (Eds.), Microbial Forensics 3e, San Diego, Elsevier, 171-191.

Belk A, Xu ZZ, Carter DO, Lynne A, Bucheli S, Knight R, Metcalf JL. (2018) Microbiome Data Accurately Predicts the Postmortem Interval Using Random Forest Regression Models Genes, 9(2), 104 https://www.ncbi.nlm.nih.gov/pubmed/29462950

Video products: Tangled Bank (https://www.tangledbankstudios.org/)
https://www.youtube.com/watch?v=B_lHQSxz9GI

Damann, Franklin E., et al 2015. *Journal of Forensic Sciences* 60 (4): 844–50.
Droettboom, Michael, et al. 2017. "Matplotlib/matplotlib: v2.0.0 (Version v2.0.0)." *Zenodo*
Loreille, Odile M., et al. 2007. *Forensic Science International. Genetics* 1 (2): 191–95.
Pedregosa, Fabian, et al. 2011. *Journal of Machine Learning Research: JMLR* 12: 2825–30.
Thompson, L. R., et al. 2017. *Nature* 551: 457–63.
Wickham, Hadley. 2009. *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer-Verlag.
Xu, Zhenjiang Zech, et al. 2019. *mSystems* 4 (1). https://doi.org/10.1128/mSystems.00269-18.