Table of Contents

I. Purpose

The ultimate goal of trace forensic evidence in a criminal investigation is to identify the people, places and things involved in the commission of the crime. The reality is that, with most types of non-biological trace evidence, association (however strong) rather than positive identification is a more likely outcome. A commonly found trace material that does not appear to be widely analyzed in operational crime labs at this time is household dust. This is unfortunate since the potential for identification rather than merely association with this type of evidence is a realistic possibility. Dust bunnies (i.e. dust balls) appear to be a unique entangled conglomeration of fibers containing a variety of inorganic and organic particulates from the immediate environment that are formed over a period of time due to air flow and that accumulate inside rooms (inside homes or the workplace), vehicles (e.g. trunk) or even in some outdoor locations. They can be transferred onto, for example, the clothing of a body that has been dragged across the floor prior to the body being taken away and deposited elsewhere. Thus, in principle, if one or more dust bunnies are found associated with a crime it should be possible to positively identify the room from which it originated. The hypothesis of the current work was that, specifically by conjoining DNA and trace evidence methodologies, it might be feasible to

identify a particular room that might be the location of a crime scene. Uniquely DNA in this study was not used to develop a genetic profile of a perpetrator, but instead was used as an additional piece of evidence to individualize a particular room or location.

II. Project design

The objectives of this projects were: 1) microscopical characterization of organic and inorganic material from dust bunnies, 2) determination of the presence and amount of human DNA in dust bunnies, 3) development of suitable DNA recovery and typing methods for bio-particles isolated from dust bunnies, 4) analysis of dust within and between rooms within the same dwelling and between different dwellings to ascertain whether support can be obtained for the hypothesis that it is possible to definitely individualize a room and its occupants and 5) development of statistical models capable of assessing the probability of a specific source being responsible for a specific dust bunny.

Phase I involved an evaluation of numerous dust samples from various locations across the U.S. in order to determine if human DNA could be obtained. Whole dust bunny samples were extracted and the presence of human DNA of sufficient quantity and quality was confirmed. Allele recovery (at least 1 to up to 30 alleles) was observed in 98% of the dust samples tested (N=40), using increased cycle number for the STR amplifications. DNA profiles with random match probabilities of $>10^6$ were obtained in 10 of the dust samples tested.

With the confirmation of the presence of human DNA in the dust samples, enhanced genetic analysis was performed using micro-manipulation methods to collect and analyze individual bio-particles from the dust specimens. Twenty dust bunnies collected "locally" from private dwellings (e.g. areas within donors' homes) or high occupancy buildings were analyzed. From these twenty dust bunnies a total of 976 putative bio-particles were subjected to DNA analysis. DNA profiles with

21-30 alleles were detected in 22 samples (2%), 37 (4%) had 11-20 alleles, 212 (22%) had 1-10 alleles and 42 (4%) only had detection at Amelogenin (gender).Of the 20 dust bunnies tested, a probative (>20 alleles) profile was observed in at least 1 bio-particle from 10 of the dust bunnies.

After the successful demonstration of the ability to perform both genetic analysis on dust bunny samples in Phase I of the work, Phase II focused on the ability to combine the genetic and micro-chemical analysis. This phase involved a larger scale study to ascertain whether it is possible to not only distinguish between different dwellings, but also different rooms within the same dwelling. For this phase of the work, dust bunnies from two rooms (living room (LR) and bedroom (BR)) in 50 single- or multiple- occupant dwellings were collected. These two rooms in the dwelling were selected due to their expected differing occupancies and activity levels. The living room in many households has a high activity level or "traffic" with potentially numerous transient occupants in additional to the dwelling occupants. In contrast, a bedroom would be expected to have less "traffic" and be primarily occupied by the dwelling occupants. Within each room, two dust bunny specimens were collected so that each room was evaluated in duplicate. Reference buccal swabs were collected, if possible, from known adult dwelling occupants to obtain reference DNA profiles from the occupants. The remainder of the report will be focused on the results from the work from Phase II.

III. Methods

A. Genetic Analysis

To prepare the dust bunny sample for analysis using micro-manipulation, the dust bunny was "teased" apart over wax paper using sterile tweezers. The material was then picked on up a low-retention adhesive material WF Gel-Film® (Gel-Pak®) that had been adhered to a glass microscope slide support. We then stained the slide with Trypan blue (0.4%). This allowed

foreasier visualization of potential human bio-particles. The samples were viewed using a high-powered stereomicroscope (Leica M205C; high resolution). Before and after images were taken of all bio-particles collected during the study. For each dust sample, a set of twenty individual putative bio-particlesor small clump samples were collected. A '0' particle sample was collected as a negative control with each set. To collect the bio-particles for analysis, a water soluble-adhesive was collected onto the end of a tungsten needle. The adhesive was then used to collect targeted cells from the substrate (e.g. WF Gel-Film® (Gel-Pak®)). The cells were then transferred directly into a 0.2mL PCR tube containing a micro-volume STR amplification mix (3.5 µl, Identifiler Plus). A lysis mix (ForensicGEM, ZYGEM, 1.5 µl) was then added to all samples after the collection was completed. A combined lysis-amplification reaction was then performed (Identifiler Plus, 34 cycles).

B. Micro-chemical Analysis

Prior to mounting, the dust specimen was held over a glass microscope slide with a flat cover slip forceps, and taped several times. The particles which fell onto the microscope slide and paper were all collected and secured with a druggist fold. If the specimen appeared to be homogeneous, five representative sub-samples were mounted on a 75 mm x 75 mm glass microscope slide in either Cargille Melt Mount® with a dispersion (HD) RI oil 1.540 or Cargille Melt Mount® 1.539. Before mounting, the specimen was teased with two needles to loosen the fibers and debris composing the dust. A representative sample of a heterogeneous dust specimen could be mounted in the same manner. However, large particles that cannot be mounted were first sorted out for separate examination. After mounting, the specimen was examined with a polarized light microscope for characterization and identification. A dust data tabulation sheet was prepared for each dust specimen. The datasheet

contains approximately 600 physical and color attributes for these components.The collected data was transferred to an Excel data table and stored for statistical analysis.

C. Combined micro-chemical and genetic analysis

Statistical analysis of the information garnered from dust bunny contents consists of a two-phase approach. Phase I solely used the material content within the empirical Bayes "two-groups" approach of Efron (Efron 2013) to estimate same-source location probability when comparing dust bunnies. Phase II used Bayesian network methods to formally combine location findings with those obtained from human DNA found within the dust. Using this two pronged approach, "occupant" information provided by DNA profiles combines with the composite material information provided by the dust bunnies to add value to trace evidence analysis. This way, not only can a particular source be pinpointed for the location of a crime scene, for example, but also identify recent occupiers of said location who may or may not have any direct connection to the criminal matter at hand.

IV. Findings

A. Genetic Analysis

50 dust bunny sample sets were analyzed. Of the 50 sets, 48 were complete sets, which included two dust bunnies from living rooms (LR1, LR2) and bedrooms (BR1, BR2). The remaining two sets had either LR1/BR1 or LR1/BR1/BR2. Additional dust samples were not available from these donors. Overall 197 dust samples were collected and analyzed. With 20 bio-particles examined from each one, this comprised 3,940 bio-particles that were subjected to a initial DNA profiling analysis, with 1,970 from living room samples and 1,970 from bedroom

samples. For each dust specimen, we compared obtained DNA profiles to the reference profiles available for each set. We determined if the profile matched one of the known occupants of the dwelling. If the obtained profile did not match one of the reference profiles, the donor source was designated as 'unknown'. However, it should be noted that we did not have reference profiles for all dwelling occupants in some cases. Therefore, we can only make assessments based on the available known occupants. For many samples, there were too few alleles obtained to make an accurate determination of the potential source of the profile. Therefore, we designate these samples as 'inconclusive'.

To readily visualize the DNA profile recovery success and probative value from the various dust samples, we have constructed box plot diagrams to show the range of values of the random match probabilities (RMP) from the various samples (i.e. discrimination potential (DP) defined here as 1/RMP)) (Figure 1, DP order of magnitude shown for all BR1 samples). We have utilized an *ad hoc* threshold of 1 in a million (10^6, represented as '6' on the box plot axis) as signifying a probative profile. Additional sets of 20-bio-particle samples were collected and analyzed for any samples in which a probative profile was not obtained, and this work is currently still ongoing.

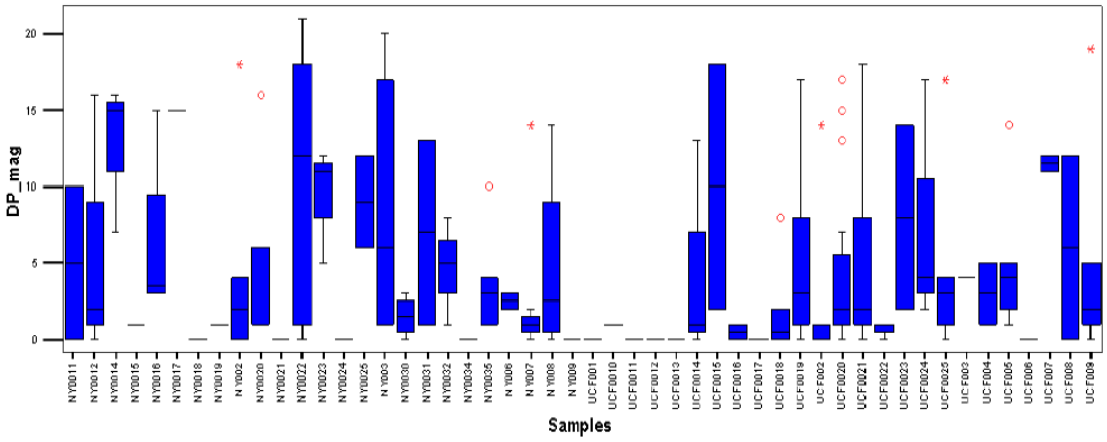


Figure 1. Discrimination Potential (DP) Box Plot for BR1 Samples from 50 Dust Bunny Sets. Y-axis: magnitude of the DP value; X-axis: Dust bunny sample.

Including the re-collections that have been performed, we have up until present analyzed 4,442 putative bio-particles from amongst the 50 dust samples. Human DNA was recovered in 640 them (~14%). Again, this is not unexpected since there is no prior confirmation that the bio-particles selected and collected actually contain human DNA or even comprise cellular material since in many instances nuclei are not visible. A number of the human cells actually present in these samples will also likely contain degraded DNA. Amongst the 640 bio-particles, 22% could be attributed to known habitual dwelling occupants. 14% were from an unknown individual who could be a habitual or transient occupant for whom we do not have a reference profile. A significant number (410, ~64%) were inconclusive which means too few alleles were present in order to make a determination of the origin. Of the 410 inconclusive samples, ~13% were samples in which only Amelogenin (sex-determination marker) was present. Mixtures were only observed in 11 samples (~0.25%), which confirms the effectiveness of the sampling method that targets only single cells or single clumps of cells.

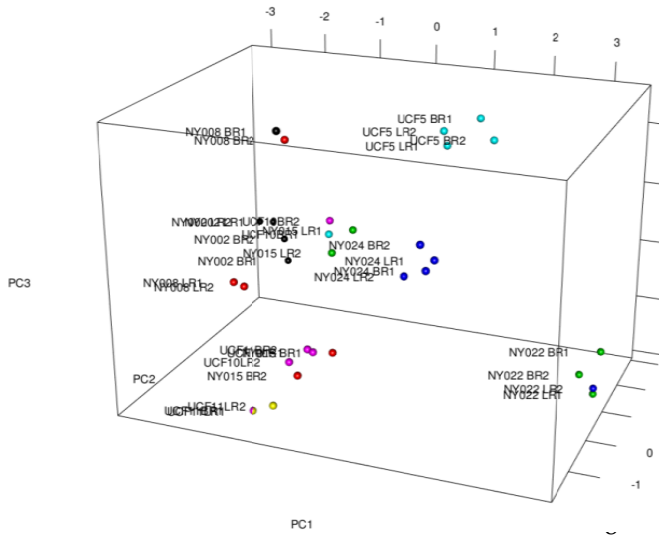
There have been 29 instances of living room – bedroom linkages amongst 14 of the 50 dust sets. A linkage is where a donor has been identified in both rooms (living room and bedroom). This is expected as habitual donors will occupy various rooms in the home. In some dust sets, the occupants providing reference samples were not always the known primary occupants of the bedrooms in which the BR1 and BR2 samples were collected. Therefore, it is not unexpected that some donors would not be found in both rooms. We also evaluated whether donor profiles were identified in the replicate samples collected from the same room (LR1/LR2 or BR1/BR2 linkages). Surprisingly, this occurs less frequently than the living room to bedroom linkages, with only 16 instances (13 BR1-BR2 and 3 LR1-LR2) amongst 11 of the 50 dust sets.

B. Combined micro-chemical and genetic analysis

Micro-chemical analysis was performed on the 50 dust bunny sets described above. A dust data tabulation sheet was prepared for each dust specimen. The collected data was transferred to Excel data table and stored for statistical analysis.

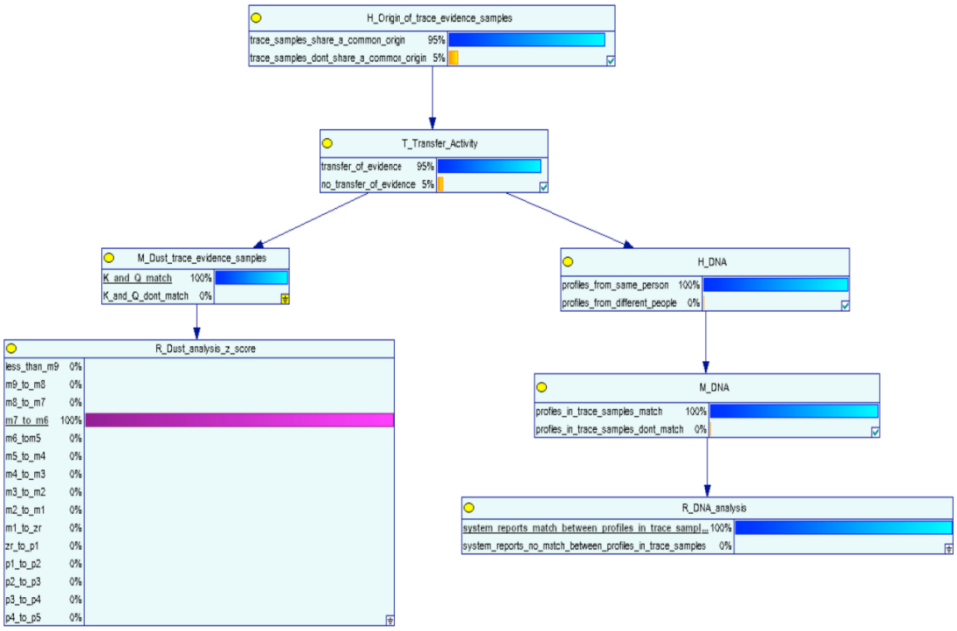
Phase I: Trace EvidenceBased Source Location Probability

A datasheet of items commonly found in dust components was used to record each dust samples composition. A "1" is recorded in an attribute cell if a dust component is observed with those given characteristics described by the cell. A "0" is recorded otherwise. Thus a dust bunny is represented as a binary feature vector which can be compared to other dust bunnies using a penalized Hamming score scheme (scale 0 to 1) first described by Daugman (Daugman 2007). Using a penalized Hamming score cut-off of 0.5 the false positive source location ID was estimated to be 0.1%. This is an average performance metric. What is of more case specific interest is the probability of a specific source being responsible for a specific dust bunny Pr(same source | Hamming score). This posterior probability was computed with the empirical Bayes methodology discussed above and yields "transformed" penalized Hamming score (scale of -9 to 5, z). Dust samples micro-chemically analyzed in this study can be projected into the space of their first three PCs, which retain 37% of the data's total variance. Figure 2 on the left shows a subset of the data. Dust bunnies from different locations cluster separately and those from the same dwelling house cluster together.



Phase II: Trace Evidence-DNA Data Fusion (Directed Acyclic Graphical Model)

The objective dust bunny trace evidence z-score provided by the non-DNA model is fused with DNA profile analysis results using Bayesian sub-network fragments first suggested by Taroni et al. (Taroni 2006). "Occupant" information provided by DNA profiles combines with the composite material information provided by the dust bunnies that speak more to location of origin. When comparing dust exemplars from two (or more) locations their z-score is entered into the model followed by DNA profile match results and corresponding DNA profile random match probability. The total information combines to change prior odds of trace evidence common origin to posterior odds from which a likelihood ratio for the evidence under trace evidence origin hypothesis can be easily extracted.



With a priori Bayesian network in place, we can now instantiate it by observing the evidence in a case. Say a body is found rolled up in a blue and red sleeping bag stuffed behind a dumpster. Underneath the body is a dust bunny composed of 16 different materials. A suspect is apprehended and their living accommodations searched. The search finds multiple dust bunnies

each with 19 different materials. When comparing the dust bunnies they produce a z-score in the range $-7 < z \leq -6$ under the dust model. This information is entered into the Bayesian network model and pushes prior odds of same source origin from 50/50 to 86/14 posterior odds in favor of the trace evidence at the two different locations sharing the same source. Later DNA results are returned indicating DNA from the body matches the profile found in skin cells recovered from the suspects living accommodations. The DNA findings are also entered into the network: pushing the posterior odds that the trace evidence shares a common origin further still to 95/5. Note that all of this is predicated on an *a priori* network that was parameterized with beliefs and data that accounted for a great deal of initial uncertainty. Less conservative values would have significantly increased the posterior odds ratio. The DAG model is under refinement to fully capture the probity of the DNA match evidence.

References

- Efron 2013. Efron B. "Large Scale Inference: Empirical Bayes Methods for Estimation Testing and Prediction", Cambridge University Press, Cambridge; (2013).
- Daugman 2007. Daugman J. "New Methods in Iris Recognition", IEEE Trans Sys, Man and Cybernetics-Part B: Cybernetics 37(5):1167-1175, 2007.
- Taroni 2006. Taroni F, Aitken C, Garbolino P, Biedermann A. "Bayesian Networks and Probabilistic Inference in Forensic Science", Wiley, Hoboken, NJ; (2006).

V. Implications for criminal justice policy and practice in the United States

In criminal investigations involving a dead body that has been moved from the original crime scene, it would be useful to law enforcement investigators (and triers of fact) to be able to obtain trace evidence from the body that would be capable of positively identifying the room or location from which the body came and who was the habitual occupier of that room. The current work, combining as it does highly discriminatory DNA and micro-chemical trace technologies that are available in most crime labs, permits the criminalist to be able to do just that. The physical evidence involved is simple household dust manifested in the form of dust bunnies.