Imagine a scenario where personal belongings such as pens, keys, phones, or handbags are found at an investigative site. It is often valuable to the investigative team that is trying to trace back the belongings to an individual to understand their personal habits, even when DNA evidence is also available. Here, we develop an approach to translate chemistries recovered from personal objects such as phones into a lifestyle sketch of the owner, using mass spectrometry and informatics approaches. Our results show that phones' chemistries reflect a personalized lifestyle profile. The collective repertoire of molecules found on these objects provides a sketch of the lifestyle of an individual by highlighting the type of hygiene/beauty products the person uses, diet, medical status, and even the location where this person may have been. These findings introduce an additional form of trace evidence from skin-associated chemistries from personal belongings. Such information could help a criminal investigator narrowing down the owner of an object found at a crime scene, such as a suspect or missing person.

Skin | phones | lifestyle chemistries | trace evidence

Trace evidence analysis plays a major role in forensic science because it reveals connections between people, objects, and places (1, 2). At a crime scene, traces collected from hairs, clothes or carpets fibers, soil, and pieces of glass (3–6) can shed light on events of the crime, and provide information about who was present before and after an unlawful act. Traditionally, both DNA and fingerprints at crime scene are used to identify a suspect with a high degree of certainty. However, both analyses have limitations. DNA analysis is limited because in many cases samples have degraded, contain insufficient amounts of DNA, are samples from which DNA cannot be obtained easily, or are contaminated (3), leading to difficulties in recovering the genomic profile of the suspect (7). Fingerprints collected at a crime scene have proven to be a very powerful approach but have limitations because they can be insufficient for matching to a specific individual if it is only a partial print (8) or overlap with other interfering fingerprints (9), and many materials, such as many fabric types and rough surfaces, are unsuitable for fingerprint analysis. Objects touched by subjects often provide only finger smears without visible fingerprints, rendering fingerprint matching to individuals useless. However, even if good fingerprints or high-quality DNA can be obtained, the person may simply not be in the database against which fingerprints and/or DNA is searched. Consequently, complementary approaches for matching samples either to individuals or to profiles that aid in identification would help an investigative team locate individuals from trace evidence commonly recovered at crime scenes.

In a scenario where personal belongings such as pens, keys, phones, handbags, or other personal objects are found at a crime scene, the chemical interpretation of skin traces recovered from these objects can help understand the individuals' personal routine. This can be accomplished by revealing information regarding habits, lifestyle, and disease status associated with certain types of medication, thus narrowing the pool of individuals to whom an object may have belonged.

Recent work demonstrates that the external environment influences the chemical composition of the outermost layer of the skin. Our daily routines leave chemicals on skin surface originating from our surroundings and the human habitats to which we are exposed (10–12). Skin-associated chemicals also arise from personal habits including diet, exercise, clothes, medications, and personal care products. Together, these sources represent the vast majority of identifiable chemical entities on the human skin surface (13). Although no approach has been developed yet to capture such lifestyle-derived skin-associated chemicals from objects, the detection of molecular traces from the skin has been used in the past and involves very sensitive techniques. Due to its superb sensitivity (14), mass spectrometry (MS) is a powerful tool widely used for forensic applications (15) by providing either molecular (16–19) or elemental analysis (20–22). Specific sets of known molecules are usually targeted in forensics applications, for example, to identify traces of illicit drugs or explosives (23–26).

However, individually targeting molecules for detection is not sufficient to describe the lifestyle of a given person, whereas a complete chemical signature obtained through the chemical analysis of a swab of the personal

Significance

This paper introduces the concept of skin-associated lifestyle chemistries found on personal belongings as a form of trace evidence. We propose a mass spectrometry-based approach to illuminate chemical traces recovered from personal objects. Using a chemical composite recovered from a swab of a phone, as a representative personal belonging, we can provide insights into personal lifestyle profile by predicting the kind of beauty product the individual uses, the food he/she eats, the medications he/she takes, or the places he/she has been. Therefore, the chemical interpretation of traces recovered from objects found on a crime scene can help a criminal investigator to learn about the lifestyle of the individual who used or touched these objects.
object might reveal personal habits and enable investigators to develop a composite sketch of a person’s lifestyle.

To test how much we can learn about personal lifestyles through the analysis of chemical traces recovered from personal belongings, we decided to use mobile phones. Phones offer a broad range of uses, in addition to staying in touch with loved ones, games, or connecting to the world through the Internet. They offer self-monitoring applications, including personal health tracking (27–32), environmental exposures (33, 34), and weather monitoring, and therefore we spend a lot of time on our phones. New applications increase the amount of time we spend on our phones (in part because this is an explicit goal of many application developers). Because of their utility, phones are handled frequently and provide ample opportunity to transfer skin molecules by contact.

In this proof-of-principle study, we aimed to (i) take an inventory of the chemistries recovered from phones of individuals and (ii) evaluate the relationships between chemistries that can be found on phones and skin of the owners. The goal of this work is not to identify the person based on the metabolomic signatures. Rather, it aims to determine how much we can learn about the personal lifestyle of the individuals based on skin-associated chemistries detected on their personal objects. We applied untargeted MS-based methods, combined with computational approaches, to take a broad look at the chemicals on hands and phones of 39 individuals. We show that the chemistries detected on phones and hands are distinct and reflect a personalized lifestyle profile that is even detected after 4 mo, and that phones’ molecular patterns match skin molecules detected on the hands of individuals. Lifestyle chemistries originating from hygiene products, beauty products, medications, and diet are detected on phones of individuals, and these traces provide information regarding personal habits and can construct a profile of individual lifestyle. Finally, spatial mapping of lifestyle chemicals on hands and phones correlate the abundance of chemicals detected on different parts of phones with those found on hands. Our results highlight the potential of molecular signatures from personal objects in assisting forensic analysis by providing insights into the personal habits of an individual.

### Chemical Patterns Among Phones and Hands of Individuals

Before we characterized specific molecules that illuminate the lifestyle of individuals, we first compared the chemical signatures of samples from each individual to test whether the chemical patterns recovered from phones match the molecules found on hands of the owners. We also evaluated how time affects such global comparisons. We collected eight samples from each right hand (three samples from the palm and one sample from each finger) and four samples from each phone (two from the screen and two from the back) (Fig. 1A). Additionally, we chemically monitored phones and hands of 10 individuals 4 months after the first sample collection to determine long-term reproducibility and signature stability. A total of 588 metabolomic samples were

---

**Fig. 1.** Representation of the three-step study workflow. The workflow includes (A) sample collection of four samples from their personal phones and eight samples from each individual’s right hand. (B) The samples are then analyzed using LC-MS/MS to characterize the chemical composition of hands and phones. (C) Data interpretation involves a statistical pipeline to compare and correlate phone and hand samples. Integrative computational methods were used for chemical characterization of hand and phone molecules, spatial mapping to display the localization of molecules, and hierarchical clustering to provide lifestyle information.
analyzed using a previously validated ultra-performance liquid chromatography/quadrupole time-of-flight (UPLC-Q-TOF) tandem-mass spectrometry (MS/MS) workflow (13), to detect metabolites present on hands and items (Fig. 1B).

We compared the molecular profiles collected from phones and hands of 39 individuals and 10 individuals 4 months later (Fig. 1C). We show that phone and hand samples, separately (Fig. 2A and B) or combined together (Fig. 2C), from different sites (phone, back and front; hand, palm and fingers) within the same individual are significantly different ($P < 0.001$) from the samples of the corresponding sites from other individuals, suggesting that phone chemistries are more similar to the hand of the owner than to hands of other individuals.

We also observed that phone chemicals collected from 39 individuals (Fig. S1A) and from 10 individuals 4 months later (Fig. S1B) are distinct [time 1: 33-fold better than random guessing (Fig. S1A); time 2: 9.4-fold of the accuracy of random guessing (Fig. S1B)]; each individual has the same number of samples; thus, the probability of guessing correctly is 1/39]. Similarly, hand metabolomic signatures discriminate each person with accuracies 38.5- and 10-fold of baseline for the first time point and the second time point, respectively (Fig. S1 C and D). These findings suggest that the chemical signatures from hands might reflect a personalized lifestyle that can be transferred to phones.

Therefore, we tested whether molecular patterns detected on phones match the chemistries of the corresponding hands. Although phones likely represent a longer-term accumulation of lifestyle chemistries than hands, chemicals recovered from phones (front and back) still match the hands of owners significantly (Fig. 2C and Fig. S1 E and F). Matching hands to samples from the back of the phone appears to be much more accurate [time 1: 69% (Fig. S1G); time 2: 85% (Fig. S1H)] than matching samples from the front of phones [time 1: 33% (Fig. S1J); time 2: 25% (Fig. S1L)], suggesting that molecules from hands transfer better to the back of phones. A significantly higher number of chemicals were detected on the back of phones than on the front ($P < 0.0005$) (Fig. S2 A–D).

The correlation between the number of molecules detected on hands and those recovered from phones shows that the transferability and recovery of hand chemistries on phones differs among individuals (Fig. S2E). Factors that could influence transferability include the contact and cleaning frequencies. The composition of the material of the surface in contact with the hands (front, glass, vs. back, plastic) may also affect the results. Mainly, as reported for bacterial communities while this manuscript was being prepared (35), chemicals recovered from the back of phones could come mostly from our hands, whereas those on the front of the phone from the face.

Although further studies are needed to determine the transferability of hands molecules on different surfaces such as plastic, wood, and glass, our results demonstrate that the chemical signatures detected on phones match the chemicals detected from the hand of the owner. Therefore, it may be possible to develop a lifestyle sketch of the person who has touched the objects based on the skin-associated molecules that are detected on phones. Specific chemistries may narrow a list of possibilities in an investigation by revealing the medications that an individual is taking, the food they eat, the clothes they wear, and personal care products they use.

**Molecular Networking to Characterize Matched Chemicals Between Phones and Hands of Individuals**

Because the results above show that the chemical signatures from phones match the chemistry found on the hands of individuals, we hypothesized that chemical traces of personal belongings can provide sufficient information to gain insights into the lifestyle of an individual such as diet, medical condition, and personal hygiene or personal care, such as the use of specific deodorants, lotions, or shampoo.

We recently demonstrated that molecular networking can be successfully used to organize large-scale MS dataset collected from the human skin surface (13). Molecular networking suggested that we could measure the medications, type of food, and personal care products each individual uses. Molecular networking is a MS visualization approach that matches MS/MS similarities and also uses the relationships between MS/MS spectra to dereplicate MS/MS signatures by matching them to reference MS/MS spectra of known chemicals (36–38). Briefly, molecular networking first uses MCluster algorithm (39) to merge all identical spectra, followed by spectral alignment to compare pairs of related MS/MS spectra via cosine similarity scoring to generate familial groupings, where 1 indicates a perfect match. We recently developed the infrastructure to perform these tasks via a web interface at [gnps.ucsd.edu/ProteoSAFe/static/gnps-splash.jsp](http://gnps.ucsd.edu/ProteoSAFe/static/gnps-splash.jsp) (40). The resulting network is then exported and visualized in Cytoscape software (41, 42) with nodes connected by edges where each node corresponds to a consensus MS/MS spectra.

Based on this potential, we used molecular networking to match specific skin-associated traces found on a phone to the hand of the corresponding individual (Fig. 1C). A molecular network was generated from hands and phones of 39 individuals and resulted in 10,383 nodes, where each node represents consensus of at least three identical MS/MS spectra (Fig. S3; [gnps.ucsd.edu/ProteoSAFe/status.jsp?task=eea01f9ec9ea44bc80d76de30da67be4](http://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=eea01f9ec9ea44bc80d76de30da67be4)).

![Fig. 2](https://example.com/f2.png) **Fig. 2.** Comparison of molecular profiles between phones and hands of individuals. Bray-Curtis distances over all metabolites were measured from collected LC-MS data to compare (A) phone samples, (B) hand samples separately, and (C) phone and hand samples combined, within the same individual (blue) and between different individuals (red). $P$ values were calculated using 1,000 random permutations.

Bouslimani et al.  PNAS Early Edition | 3 of 10
molecules on phones that are also found on hands of individuals (Dataset S1). On the basis of annotated hand-associated chemicals found on phones, we can gain insights into the lifestyle of the individual to whom the phone belonged by matching the annotated molecules to the type of hygiene/beauty products, pesticides/insecticides, plasticizer exposures, diet, and medications (Dataset S1). Some representative examples of unique lifestyle chemicals detected on hand and phone of a specific individual are shown in Fig. 4.

We were able to detect a skin inflammation medication (clolobetasol propionate) on both the hand and phone of volunteer 9, as well as a lipid, trihydroxy palmitic acid, used to produce soap and cosmetics (44) (Fig. 4 A and B, and Fig. S4 A and B). A molecular family of the trehalose fatty acids, usually used as pigment dispersant allowing pigments to be blended in various high-end cosmetics (α,α′-trehalose 6-palmitate [M+Na]⁺ m/z 603.334 and its analog [M+Na]⁺ m/z 631.366) (45), was found in samples from individual 38 (Fig. 4C and Fig. S4 C and D). A hair regrowth treatment, minoxidil, was used by person 2 (Fig. 4D and Fig. S4E); an active ingredient of eye drops, tetrahydrozoline, was used by individual 10 (Fig. S4F); and an antifungal medication was exclusively detected on person 16 (Fig. 4E and Fig. S4G). Furthermore, molecules likely derived from diet and/or hygiene products were also detected. Among them are sinensetin-type molecules, frequently found in citrus fruits (46) and detected only on the hand and phone of person 32, who is a large consumer of oranges and uses a lot of citrus-based cleaning products (Fig. 4F and Fig. S4H). Oral medications such as an antidepressant citalopram (consumed 20 mg/d) was also detected on the hand and phone of volunteer 21 (Fig. 4G and Fig. S5A). As highlighted by the specific molecular cluster in Fig. 3G, not only was the drug detected in volunteer 21’s samples but also its oxidized metabolite citalopram N-oxide ([M+H]⁺ m/z 341.169), displaying a mass shift of 15.994 Da and a similar MS/MS fragmentation pattern compared with the drug itself (Fig. 3G and Fig. S5B). Consistent with previous findings, consumed drugs that are detected on hands might originate from sweat secreted on the skin (47, 48) or can be transferred upon contact (25, 49).

Fig. 3. Representation of number of chemical features detected on phones and hands of 39 individuals. A–C were generated from LC-MS/MS clusters provided by the molecular network represented in Fig. S3, gnps.ucsd.edu/ProteoSAFe/static/gnps-splash.jsp using accurate parent mass and MS/MS fragmentation patterns (gnps.ucsd.edu/ProteoSAFe/result.jsp?task=0671dd76ec8334d528f5b4b5ae243e045&view=group_by_compound). Such annotations are in agreement with the level 2 of annotation defined by the 2007 metabolomics standards initiative (43). The annotated compounds lists were then correlated to the information provided by volunteers for this study, regarding the food they consume and the hygiene products and medications they were using.

Unique Chemicals Detected Only on Hands. Molecules from diet were often specific individual. For example, Piperine from black pepper (55) was only found on the hand of person 3 (Dataset S1), and capsaicin, the active ingredient in chili peppers (56), was detected only on the hand of person 28 (Dataset S1). Some
molecules such as antihistamine that were taken as solution (Fig. S8A and Dataset S1), ibuprofen (Fig. S8B and Dataset S1), and pet pesticides (Fig. S8C and Dataset S1), used occasionally by volunteers, were found on hands of individuals but not on their phones. As highlighted above, the detection of molecules on the hands can result from sweat secreted by the skin (47, 48) or can be transferred upon contact (25, 49). The detection of chemicals on hands but not on phones indicates that not all lifestyle signatures from the hands are recovered on the phones, reinforcing that the overall profile is important to an investigator rather than the detection of individual molecules.

The combined molecular networks of the two time points from the 10 individuals reveal that the percentage of stable chemicals varies from person to person (Figs. S7 and S9A). Ten percent to 28% of detectable MS/MS spectra that originate from the hand (Fig. S9B) and phone (Fig. S9C), respectively, are stable over time and are still detected on phones after 4 months. This result suggests that, although personal lifestyle might change over time, for the same person many lifestyle traces recovered from phone can still match the original phone’s and hand’s chemistries after considerable time has elapsed.

Predicting Lifestyle of Individuals Based on Phone Metabolomic Signatures

In addition to correctly and efficiently matching the molecular traces recovered from phones to the hands of the corresponding individuals, molecular networking helps gain insight into the individual’s lifestyle by identifying molecules detected on personal phones. For example, detecting sunscreen-derived molecules indicates that this person likely lives in a sunny area and spends time outside, caffeine indicates the person is a coffee consumer, nicotine indicates a smoker or at least someone who is exposed to smoke, medicines help establish the medical status of an individual, and detection of citrus-derived molecules observed on some individuals’ hands and phones indicate they may eat oranges or lemons. The combination of many such lifestyle routines defines a unique skin chemical signatures.

Because a combination of chemicals to predict lifestyle is more informative than a unique single molecule, we organized the annotated molecules (Dataset S1) using a hierarchically clustered heat map (Figs. 1C and 5, and Fig. S10) based on presence/absence data regarding detection of annotated MS/MS spectra on the phones and
hands of individuals. Therefore, hierarchical clustering reveals an overall view of an individual’s lifestyle that helps predict personal habits and routine (Fig. 5 and Fig. S10). Although some annotated molecules are found on the phones and hands of multiple individuals, the heat map reveals specific signatures that reflect a personalized lifestyle. The origin of each molecule highlighted in the heat map was represented using a lifestyle-associated legend, allowing an easier interpretation of the molecules detected and how they relate to lifestyle (Fig. 5 and Dataset S1). As shown in Fig. 5, a chemical signature reflects several lifestyle characteristic features of a person. The personal habit includes the kind of hygiene products a person uses, such as plant-based products, foaming products, or sunscreens; the kind of beauty products the individual uses, such as makeup and fragrance; and the type of food the person eats or drinks, such as spices or coffee. Skin-associated chemicals detected on phones also indicate where the person has been and their medical status. Below, we highlight representative examples that illustrate how an investigator might obtain insight into the lifestyle of the individual.

A group of chemicals detected on phone and matched to hand of individual 1 provides information regarding his/her lifestyle. Based on this signature (Fig. 5A), we can tell that this person uses plant-based products because rutin, a flavonoid glycoside found in many plants (57), was detected on his/her phone and hand (Dataset S1). This person likely lives in a sunny area and spends a lot of time outside (avobenzone and oxybenzone, two active ingredients of sunscreen). Volunteer 1 uses a skin moisturizer (dexpantenol) (Dataset S1) (58), and flavoring products due to the detection of a flavoring agent (m-anisic acid), a molecule commonly added to fragrances (Dataset S1) (59). Furthermore, DEET (Fig. S6E) was detected and is associated with mosquito lotions, suggesting that the person is likely a camper or backpacker. We confirmed that this person, a female, spends time outside and indeed went camping and used DEET.

Volunteer 16, in addition to using an antifungal treatment, uses moisturizer due to the detection of dexpantenol, often found in moisturizers (58) and personal care products that contain foaming agents (N-lauroyl sarcosine). This person eats herbs and spices, consistent with the detection of 7-hydroxyastrobailignan (Fig. 5B and Dataset S1).

Similarly, the medical status of volunteer 21 can be established upon the detection of antidepressant drug and its metabolite. Additionally, this person uses skin care products such as facial or body creams, as a skin-conditioning agent [glycerol tricaprylate (60)] was detected (Fig. 5C and Dataset S1). A flame retardant [tris(2-butoxyethyl) phosphate] usually used in plasticizers (61) and textiles was also detected on individual 21 (Fig. 5C and Dataset S1). In addition of the antifungal medication (clofibetasol...
propionate) used by person 9 (Fig. S4A), a lipid trihydroxy palmitic acid used to produce soap and cosmetics was also detected (Fig. 5D and Fig. S4B).

Volunteer 2, in addition to using a hair regrowth treatment (minoxidil), is a coffee or tea consumer (caffeine) (Fig. 5D) and is likely female based on the personal care product signatures. This individual uses a lot of cosmetic products, as we detect castor oil derivatives (sebacic and suberic acids) used as skin-conditioning and -moisturizing agents (62) and nonsynonyms (webdictionary.personalcarecouncil.org) often used in hair dyes and as wetting agents in cosmetics, including hair and skin care products (Fig. 5E and Dataset S1).

The detection of two pigment dispersants (trehalose fatty acids) (Fig. S4 C and D) (45) associated with individual 38 together with the detection of glycans (Dataset S1) that are usually incorporated in facial creams for their antiaging properties (63), likely indicate that this person is a female who uses makeup (Fig. 5F).

Each of these unique molecules mentioned above was confirmed by the volunteers and is in agreement with the lifestyle of each individual. According to the volunteers, none of the lifestyle molecules annotated was discordant with their own personal knowledge of their lifestyle. Currently, due to the absence of the correct databases, we do not have another option but to ask participants to study the effectiveness of the matches and how they correlate to their lifestyle. One can imagine, however, in the future to have a statistical approach and a confidence score for the type of lifestyle when appropriate databases become available that connect each molecule and associate such signatures with a lifestyle (e.g., beer drinker, jogger, hunter, diabetic, fruit eater, etc.).

Spatial Distribution of Lifestyle Chemistries on Phones and Hands of Individuals
To evaluate whether different parts of phones and hands have similar molecular features, spatial molecular maps (13) were generated (Fig. 1C), therefore allowing the visualization of the distribution and the abundance of specific chemicals detected on phones and hands. Fig. 6 illustrates the spatial distribution of specific molecules detected on phones and hands of four individuals. The hair loss treatment minoxidil is highly abundant on the index and middle fingers of volunteer 2, which correlates with its high localization on the top back of the phone (Fig. 6B), probably due to the frequent contact of these two fingers with that part of the phone.

For individual 32, the citrus-associated molecule annotated as sinensetin is mainly present on middle and ring fingers with correlation to a high abundance on the back of the phone (Fig. 6B). The drug citalopram, used by volunteer 21, is mostly present in the bottom palm of the hand and little finger and more abundant on hand rather than on the phone (Fig. 6C), whereas the oxidized metabolite of citalopram is mostly present on the phone (Fig. 6C). The distribution of α,α′-trehalose 6-palmitate ([M/z 603.336, [M+Na]+]) from the trehalose fatty acids molecular family displays a very similar spatial distribution to that of its analog ([M/z 631.336, [M+Na]+]) (Fig. 5D), revealing that these two molecules originate most likely from the same beauty product (Fig. 6D).

The use of spatial mapping thus allows visual comparison of molecular distributions on phones and hands, by displaying the localization and abundance of molecules of interest and correlating that with personal habits. Additionally, like molecular networking, spatial mapping can also be used to match chemical features detected on phones to hands of individuals (Fig. S11 and Dataset S2). Moreover, untargeted spatial mapping and molecular networking are complementary tools as they are generated based on different fractions of the same dataset: MS1 features (m/z + retention time) and MS/MS data, respectively. Thus, if molecules that might be found on hand and phone of a specific individual are not highlighted by molecular networking because they did not fragment during MS/MS analysis, they can be picked by spatial mapping. Additionally, the detection of molecules in corresponding spatial locations on hands and phones decreases the possibility of chance matches.

Fig. 6. Spatial mapping of lifestyle chemicals detected on hands and phones of individuals. Two-dimensional spatial maps generated from LC-MS1 data display the localization and abundance of molecules of interest on hands and phones of four individuals. Figure highlights phone and hand localization of (A) the hair loss treatment molecule minoxidil ([M/z 210.135; RT, 150.8 s]) on person 2, (B) sinensetin ([M/z 373.125; RT 230 s]) on person 32, (C) antidepressant citalopram ([M/z 325.171; RT, 228] and its metabolite citalopram N-oxide ([M/z 341.163; RT 230]) on person 21, and (D) trehalose fatty acids ([M/z 603.336; RT 361]) on person 38. Color scale is associated to the relative abundance of features, with red corresponding to the highest relative abundance; this is indicated as a plus (+), and the minus (−) is the lowest abundance.
Conclusion/Discussion

In this study, we established a methodology to characterize the chemistry present on mobile phones and to calculate the relationships between chemistries that can be found on objects and the skin. Our approach shows that many chemicals recovered from phones originate from the hands of the owners, highlighting therefore the transfer of our skin molecules on objects that we touch.

Furthermore, we show that molecular signatures recovered from phones can be used to gain insights into personal habits based on similarities of skin-to-object chemical signatures. Many molecular classes, revealed through molecular networking, were identified on phones. These include beauty and hygiene products, diet, pesticides/insecticides, plasticizers, and medications such as antifungal, skin inflammation, and oral antidepressant treatments. Furthermore, we show that even molecules that were applied more than 6 months before sample collection can still be detected on phones, such as DEET, active ingredient of mosquito lotions. Some molecular traces tested 4 months later persist on phones and match the hands of the same individuals sampled at the first time point.

In addition to highlighting unique lifestyle chemicals detected on phones of individuals, molecular networking correctly and efficiently links the molecular traces recovered from phone to the hand of the owner. Furthermore, we show that the combination of many such identifiable chemistries detected on phones defines chemical signatures that help to predict the lifestyle of an individual and to construct a composite sketch of personal lifestyle.

Our findings introduce an additional form of trace evidence based on molecular signatures of human hands. Such a MS-informatics–based approach has yet to be used to link traces of untargeted chemical data collected from an item, such as a phone, to the lifestyle of an individual. Because the computational challenges of such untargeted metabolomics analysis are intensive, only now can we begin to perform such calculations at the scale that is needed for forensic analysis, but only with the use of supercomputers. Therefore, the MS data collection and analysis, although only about 500 samples in this study, are still laborious tasks. The main challenge in implementing large-scale untargeted metabolomics is not the data collection, but the Big Data management and the development of the appropriate tools to integrate and effortlessly interpret the large datasets.

Although significant challenges remain in annotating such a large-scale dataset, the implementation of this methodology provides the key stepping-stone needed to use molecular traces detected on personal belongings to gain insights into personal habits and daily and hygiene routine of individuals, and also to show where they have been or even if they are on medication.

At present, the molecular lifestyle signatures analysis approach to track and find an individual can complement traditional physical evidence including DNA, fingerprints, and skin microbial fingerprints (64, 65) to provide additional information about a person. Furthermore, if the traditional physical evidence is not reliable, incomplete, or not available to accurately identify an individual, the molecular lifestyle signature might be used to provide information that is otherwise impossible to obtain. The molecular analysis would help a criminal investigator in narrowing down the owners of the object (e.g., a suspect of a crime scene or understanding the habits of a terrorist) by identifying specific lifestyle characteristics from objects they touch.

Perspective

Although it is unusual to write a dedicated section as a perspective in a research paper, we felt it is important to describe how this work might move beyond the proof-of-principle presented in this paper. We envision the development of confidence measures that can be associated with each lifestyle chemical signature identified on objects, similar to confidence that can be assigned to fingerprints matches. Despite the development of tools such as large-scale molecular networking and the increasing number of publicly accessible MS/MS reference datasets in the recent years, the current method described in this paper still requires manual interpretation of the data and needs further validation to be applied in forensics. However, we anticipate this proof-of-principle study to complement the traditional approaches in forensics science, through mass-spectral analysis of the objects someone touches frequently.

Ultimately, this work suffers from the chicken-and-egg conundrum where the approach is the egg and the database is the chicken: one of them had to come first. With all of the current MS/MS mass-spectral reference libraries METLIN (66), mzCloud (https://www.mzcloud.org), Massbank (67), the Human Metabolome Database (69), and GNPS (gnps.ucsd.edu/ProteoSAFe/static/gnps-splash.jsp) (40), the scientific community can only, on average, annotate 1.8% of metabolomics data (70) (2.3% in this study, gnps.ucsd.edu/ProteoSAFe/result.jsp?task=eea019ec9e44be3d07e30da67be4&view_all_annotations_DB), and yet we can already reveal specific lifestyle signatures of people with this level of annotations due to large number of molecules that are detected with a single metabolomics run. However, as the community makes more reference spectra available, the construction of a database of lifestyle/personal habit versus a reference spectra will lead to solid association of lifestyles to people. Currently, such a database does not exist and it will take significant community effort akin to the level of fingerprint databases.

Here, we show that the principle behind the approach works. We believe that an extensive database with appropriate metadata of where these molecules are found and their associated lifestyles can become available. Such a database would not only benefit forensics (tracking a subject) or a terrorist tracing (useful for the military) but also be useful for toxicology, as it would be a noninvasive way to measure environmental exposures, such as exposure to plasticizers and other pollutants (Fig. S10 B and C). Finally, we envision the use of this approach in medicine perhaps to understand patient compliance and patient-specific metabolism, or even monitor individualistic responses to medications, even without the need to use a needle, but again with the proper computational and reference infrastructure.

Materials and Methods

Subject Recruitment and Sample Collection. Thirty-nine healthy adults were recruited to participate in the study. All individuals signed a written informed consent and approved sample collection from their hands and personal objects, in accordance with the sampling procedure approved by the University of California, San Diego, Institutional Review Board (Approval 130537X). There was no skin preparation for sampling, and volunteers kept their same daily routine. The idea was to use chemistries that compose the traces on objects as a signature, including molecules from external environment and/or daily routine. Samples were collected from right hand and phone of each individual using swabs, following the previously validated protocol (13). Eight samples were collected from each right hand (three from the palm and one from each finger) and four samples from each phone (two from the screen and two from the back) (Fig. 1). Sampling was performed at 5 × 5-cm area for phones spots and 4 × 1-cm area for hands spots, with premoistened cotton swabs at each site in 50:50 ethanol/water. The absorbed material was then extracted in 200 μL of 50:50 ethanol/water. Time 2 samples were collected 4 mo later following the same protocol, from 10 individuals recruited in the first time of sampling. After collection, samples were stored at –80 °C until further analysis. The ethanol/water extracts were subjected to UPLC-Q-TOF for detection of smaller molecules, including MS/MS of molecules for molecular network analysis.

UPLC-Q-TOF MS Analysis. Hand and phone swab sample extracts (Dataset S1) were analyzed using previously described methods described in detail in ref. 13. We used a UPLite 3000 UPLC system (Thermo Scientific), controlled by Chromelon software (Thermo Scientific). Briefly, UPLC conditions of analysis were as follows: 1.7-μm C18 (50 × 2.1-mm) UHPLC Column (Phenomenex); column temperature, 40 °C; flow rate, 0.5 mL/min;
Supervised random forest classifier was used to classify the amplitude of metabolite features in featureXML format. The model performances of the time 1 and time 2 predictions were consistent. For each time point, the phone samples were collected from the residents with molecular signatures. The model performances of the time 1 and time 2 samplings were considered. One bucket table with a total number of extracted features of 32,487 for time 1 and 32,487 for time 2 were considered. The difference in 1,000 random permutation of the sample labels.

Bray-Curtis Dissimilarity Analysis. For the calculation of sample-sample dissimilarity, the Bray–Curtis metric was used. Bray–Curtis dissimilarity between a pair of samples is defined as:  
$$d_{ij} = \sum (M_i - M_j) / (M_i + M_j),$$
where M_i is the amplitude of metabolite i in sample i, and M_j denotes the sum of all metabolites in sample i. M_i = âˆ‡m_i.

For calculation of statistical significance, the difference in the mean of the within and the between group Bray–Curtis dissimilarity was compared with the difference in 1,000 random permutation of the sample labels.

Random Forest Models. Supervised random forest classifier was used to evaluate the discriminant power of molecular signatures collected from hands and phones of individuals. The classification model was tuned and validated with cross-validation with the bucket table generated from the samples collected from phones and hands. Then the model was applied to predict the phone owner, to determine whether they can trace back to their owners with molecular signatures. The model performances of the time 1 and time 2 are consistent. For each time point, phone samples collected from the front and the back parts were also predicted separately, to determine which phone surface (glass, front, and plastic, back) provides the highest recovery of hands molecular signatures.

Molecular Networking. LC-MS/MS data collected from hands and phones were converted to the mzXML open file format in Compass data analysis software (Bruker Daltonics) and were subjected to the molecular network workflow of GNPS (40) (gnps.ucsd.edu/ProteoSAFe/static/gnps-splash.jsp). MCLuster was used to group MS/MS spectra that originated from identical compounds together (71). These clusters were then represented with a consensus spectra used to construct the molecular network. Similarity between every pair of MS/MS spectra was determined using a spectral alignment algorithm calculating a cosine score. Cosine score ranges from 0 to 1, with 1 indicating a perfect match. A cosine threshold of 0.65 was used for the molecular network in this study. Molecular networking parameters included precursor m/z tolerance of 1 Da and a fragment mass tolerance of 0.5, only MS/MS spectral pairs with at least four matched fragments were included, and each MS/MS spectrum was only allowed to connect to its top 10 scoring matches, resulting in a maximum of 10 connections per node. The maximum size of connected components allowed in the network was 600 and the minimum number of spectra required in a cluster was 3. Molecular network parameters for MS/MS data collected from 39 hands and phones are accessible here: gnps.ucsd.edu/ProteoSAFe/status.jsp?task=eaa01f9ec93a4bc0d7763e30d5b6d4; molecular networks generated from MS/MS data collected from hands time 1 and phones time 2: gnps.ucsd.edu/ProteoSAFe/status.jsp?task=60dbae8123384d69802046650c400d5; phones time 1 and phones time 2: gnps.ucsd.edu/ProteoSAFe/status.jsp?task=072732e55c344fa9b9403c9e144df.

The resulting molecular network bucket table was generated using GNPS and then visualized in Cytoscape software (42); nodes represent consensus MS/MS spectra connected with edges. Network organization was performed using the FM3 layout plugin (72) (appsp.cytoscape.org/apps/fm3). To simplify the network, MS/MS spectra corresponding to the background (from swabs and solvents) were removed. MS/MS data were simultaneously searched against GNPS MS/MS spectral libraries, and putative identifications were imported into the network and displayed in Cytoscape as attributes.

Frequency Histogram and Heat Map. To build the frequency histogram, the GNPS precursor ion intensity table (available as “Create Cluster Buckets” in the workflow option on the GNPS web page) was used. The molecular networking output generated from hands and phones of 39 individuals (Dataset S1). The rows (samples) and columns (molecules) were ordered by hierarchical clustering using the Jaccard dissimilarity measure and Ward grouping method. The R package d3heatmap was used to draw and display the hierarchical clustering heat map in the shape of a square.

Two-Dimensional Spatial Molecular Mapping. Spatial mapping was used in this study to map the number of features detected in each sample and to display the localization and abundance of features on hands and phones of individuals. Spatial mapping for hands and phones was performed in the ili toolbox for 2D and 3D spatial mapping. This open-source toolbox was developed by us and freely available at https://github.com/ili-toolbox/ili. We used the workflow previously developed for molecular and microbial cartography of the human skin surface (13). Two-dimensional mapping involves assigning for each sampling spot a spot (x-y pixel coordinates of the center, radius) in the image or photo of hand and phone that spatially fits the best as judged by the visual examination. Visualization of such 2D maps was performed in ili by dragging and dropping the hand or phone model (Fig. S11) and a CSV file that combines coordinates information and the bucket table generated in section 3 (Dataset S2). A .CSV file includes hands/phones filenames, specific coordinates of each sampling spot, spot radii, and relative intensities of detected molecular features (Dataset S2). To display MS (UPLC-Q-TOF) data, spectral intensities are integrated over a given m/z and retention time regions, corresponding to the molecular abundance at each location on the hand and phone. These data were visualized using pseudocolormap and overlaid onto images of the hand and the phone, at specific coordinates based on sampling location. The so-called jet pseudocolormap was used to display the relative signal intensities at each spot location.

ACKNOWLEDGMENTS. We thank all volunteers who were recruited in this study for their participation. We also thank Tal Luzzato Knaan for her valuable input to the manuscript. This work was supported by National Institute of Justice Award 2015-DN-BX-K047. T.A., A.B., and P.C.D. acknowledge funding from the European Commission’s Horizon 2020 Programme (Grant 634462). R.R.D. was supported by the São Paulo Research Foundation (Award FAPESP-2015-03348-3). This work was partially supported by US National Institutes of Health (NIH) Grant 5R34GM120484-07; N.B. was also partially supported as an Alfred P. Sloan Fellow. We further acknowledge NIH Grant GM51089211 and Bruker for the shared infrastructure equipment that enabled this work.

Bouslimani et al.