J Forensic Sci, March 2008, Vol. 53, No. 2
doi: 10.1111/j.1556-4029.2007.00649.x
Available online at: www.blackwell-synergy.com

TECHNICAL NOTE

April Hemmila,[1] *B.A.; Jim McGill,*[2] *Ph.D.; and David Ritter,*[2] *Ph.D.*

Fourier Transform Infrared Reflectance Spectra of Latent Fingerprints: A Biometric Gauge for the Age of an Individual*

ABSTRACT: To determine if changes in fingerprint infrared spectra linear with age can be found, partial least squares (PLS1) regression of 155 fingerprint infrared spectra against the person's age was constructed. The regression produced a linear model of age as a function of spectrum with a root mean square error of calibration of less than 4 years, showing an inflection at about 25 years of age. The spectral ranges emphasized by the regression do not correspond to the highest concentration constituents of the fingerprints. Separate linear regression models for old and young people can be constructed with even more statistical rigor. The success of the regression demonstrates that a combination of constituents can be found that changes linearly with age, with a significant shift around puberty.

KEYWORDS: forensic science, fingerprints, chemical composition, fourier transform infrared spectroscopy, principal component analysis, least-squares analysis, age determination, biometrics

As part of an investigation of a child abduction incident, Oak Ridge scientists found that at moderate temperatures children's fingerprints may disappear in as little as 24 h, while the prints from adults last longer (1). Other studies report consistent results. For example, Blasdell found that after 7 days in a 64.5–76°C oven, 76% of the children's fingerprints had become inadequate for lifting, while only 8.5% of adult prints were no longer adequate for lifting with standard black powder (2). Up to 85% of the adult's fingerprint mass has been found to be lost over 2 weeks, children's even more quickly (3). The dominant process is evaporative loss of moisture and volatile unesterified free fatty acids. After the loss of moisture, lipids remain as the main components of fingerprints. Low boiling point lipids vaporize at high temperatures.

The Oak Ridge group used gas chromatography-mass spectrometry (GC-MS) to determine that the chemical compositions of fingerprints differ significantly between children under the age of 10 and adults. Fingerprint compositions from children had higher concentrations of volatile unesterified free fatty acids, while adults had higher concentrations of less volatile fatty acids esterified with long-chain alcohols. Most adults' latent prints are composed of perspiration and sebum (3). They conclude that sebaceous excretions that increase after puberty can account for the observed differences (4) (Michelle Buchanan, ORNL, and Art Bohanan, Knoxville Police Department, http://www.ornl.gov/info/ornlreview/rev29_3/text/war.htm). The main lipid components of sebum are free fatty acids (30%), glycerides (33%), wax esters (22%), squalene (10%), with the remaining 5% being cholesterols and hydrocarbons (5).

[1]Present Address: Southeast Missouri Hospital College of Nursing and Health Sciences, Cape Girardeau, MO.
[2]Department of Chemistry, Southeast Missouri State University, One University Plaza, Cape Girardeau, MO 63701.
*Presented at the 43rd Annual Meeting of the Missouri Academy of Science, Saint Joseph, MO, April 20, 2007.

Received 19 April 2007; and in revised form 11 Aug. 2007; accepted 11 Aug. 2007.

The dermatological literature also contains studies on differences in the chemical composition of human skin surface lipids at various ages. These analyses are typically performed by hydrolysis of the lipid materials, conversion to the corresponding fatty acid methyl esters, and GC-MS analysis (6–10). These studies report that the sebum of children is characterized by relatively high concentrations of long-chain free fatty acids and significant levels of cholesterol and cholesteryl esters. These compounds typically decrease in abundance with age, while the concentrations of wax esters tend to increase steadily through puberty, at which time they typically peak and begin to decline again with further aging. Within the wax ester class, branched and straight-chain fatty ester components tend to change systematically with age, the former decreasing and the latter increasing in concentration. While significant concentrations of squalene are seen throughout life, its concentration is reported to increase somewhat with age (1,4,11). Total sebum cholesterol generally decreases with age, while the proportion of cholesterol in esterified form tends to increase (1,4,9,10).

Experimental

These reports that fingerprint composition differs with age led us to examine the possibility that a change in the compositions linear with age could be found. Fingerprints were collected from 78 individuals ranging in age from 4 to 68 years old following a protocol approved by the Southeast Missouri State University College of Science and Mathematics Review Committee for Research Involving Human Subjects. No washing or other preparations were conducted prior to collection. Each person pressed one finger onto one tin plate. Tin was chosen because it is a representative, inexpensive, and relatively inert metal; some other metals show evidence of chemical reaction with print constituents (work in progress). Two infrared spectra were taken of each print; children 0–19 years of age within 24 h, and adults 20–70 years within 48 h on a Thermo NEXUS 670 FT-IR spectrometer (Fig. 1). A Smart Endurance diamond composite-crystal single-reflection ATR with 0.75 mm area

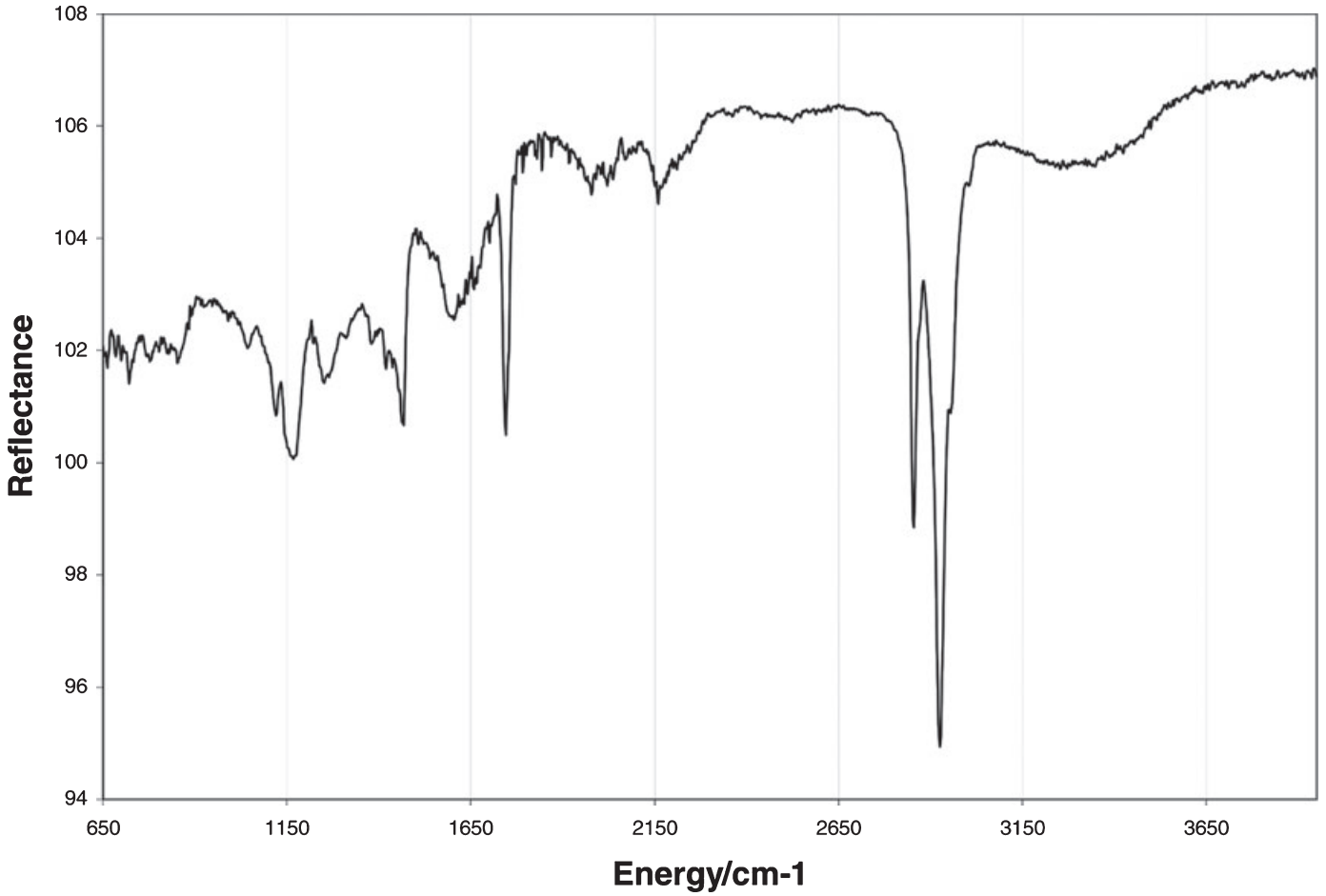

FIG. 1—*FTIR reflectance spectrum of the latent fingerprint of a 13-year-old male subject.*

was used over a range of 650–4000 cm^{-1} at a nominal resolution of 2 cm^{-1} to yield 1741 channels. Spectra were the average of 64 scans and were ratioed against air (background). Each spectrum was logged with age and sex information by an ID number that contained no age or sex coding. In accord with the approved experimental design, no demographic data other than age were considered.

Each spectrum was saved as a CSV text file, and then added to a Microsoft Excel spreadsheet. The entire set was arrayed into a single Excel matrix that was imported into The Unscrambler (12). A principal component analysis (PCA) of the entire set of all samples and all wavelengths initially yielded groups organized by day of taking spectrum. As a correction, each spectrum was modified by taking the Golay second derivative to eliminate baseline variations, and the CO_2 wave number range from 2378 to 2763 cm^{-1} was removed from the analysis to minimize the effect of day-to-day changes in background quality. These treatments removed the day-to-day grouping artifact, Fig. 2.

A partial least squares (PLS1) regression analysis of age versus IR spectra for the whole population showed a very strong correlation (92%) between actual and predicted age of the individual, and a root mean square error of calibration (RMSEC) of 3.6 years by using eight principal components (Fig. 3*a–c*). Full cross validation, in which one sample at a time is kept out of the calculation, was used in each regression. This was repeated until all samples had been kept out once. There are as many recalculation rounds as samples. The RMSEC was then computed to include all of the regressions. Thus, each sample is statistically treated as an unknown, and the RMSEC includes contributions from each. Inspection of the predicted versus measured plot indicated a discernible inflection in the plotted samples in the mid-twenties age range. To investigate, separate PLS1 analyses were done on the group younger and the group older than this age to see if they were statistically different (Figs. 4 and 5). Each group was found to produce a model that was better than the original population as a whole. The young group produced a correlation of 95% with RMSEC of 0.78 years with three PCs. The older group correlation was not quite as good: 95% correlation with RMSEC of 4.8 years with four PCs.

The inflection in the whole-group plot considered together with the improved fits for young and old suggest that the compounds that are important in the regression for young may be different than for older. To probe this, the regression coefficients from the three PLS1 analyses were compared a wavelength at a time (Figs. 3*c*, 4*c*, 5*c*). Areas of high significance were identified on each and then correlated between studies (Table 1). Each of the high-significance regions of the spectrum was re-analyzed separately by PLS1 three ways: on the entire population, on the young, and on the old. The results are summarized in column five of Table 1. The 1190–1260 cm^{-1} region separated the ages 18–22 from the rest of the population (Fig. 6*a*); 1440–1782 cm^{-1} separated the ages above 29 from those under (Fig. 6*b*). The 2828–2970 cm^{-1} region separated the samples into four populations: 4–5, 11–14, 18–26, and 29–70 years old (Fig. 6*c*).

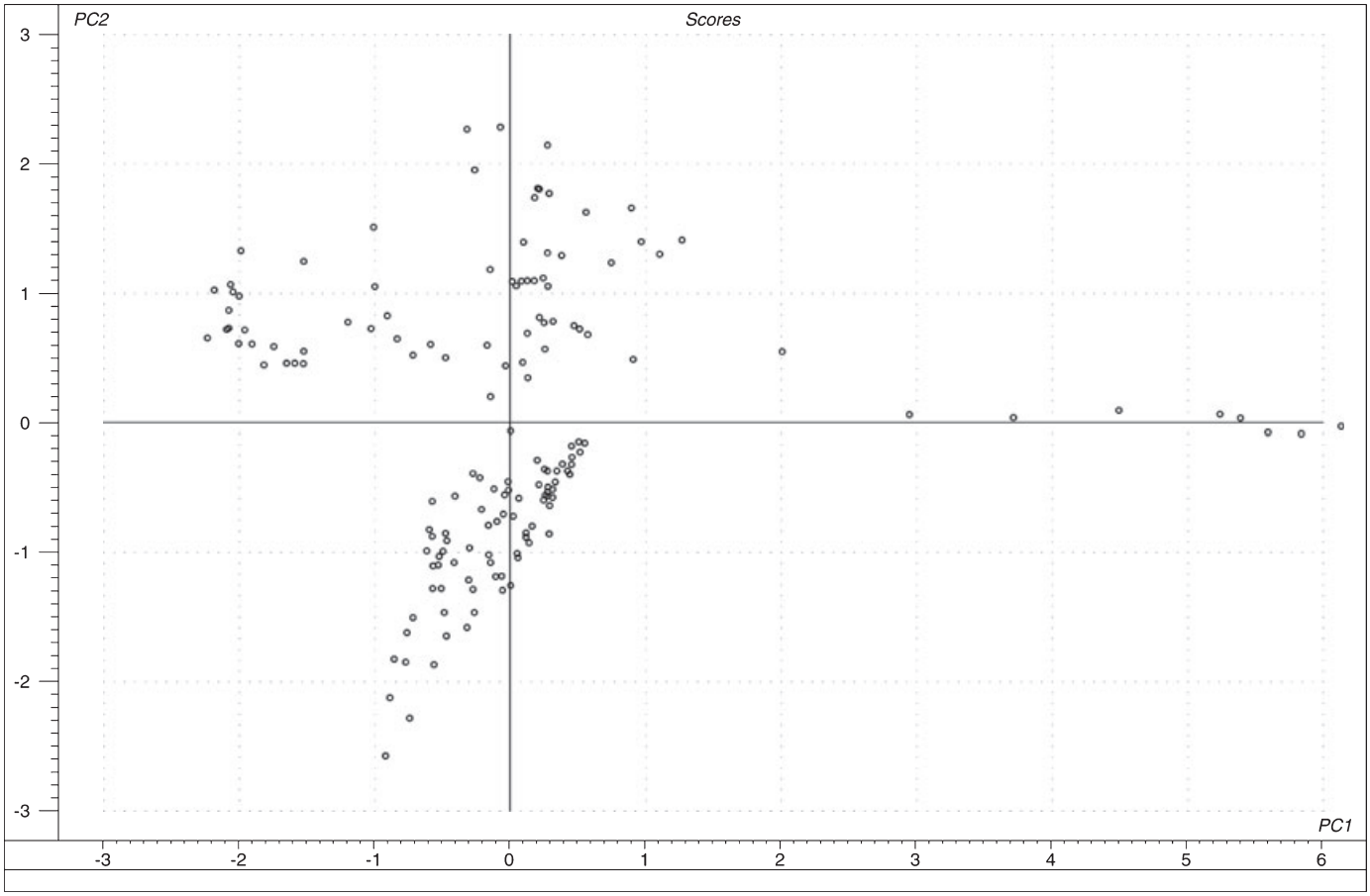

FIG. 2—*PCA2 versus PCA1 score plot for population. PCA1 explains 27% of variance in X spectra.*

Results and Discussion

The PCA method is unbiased (13). Its first component, PCA1, represents the main constituent variance of the fingerprint mixture. In contrast, the PLS1 algorithm actively uses the *Y*-data, age of person, in determining its first component, PCS1. The result is that PCA1 accounts for 27% of the *X* variance, while PCS1 accounts for a much smaller amount, 8%. Comparison of the score plots from the PCA to the PLS1 analyses (Figs. 2*a*, 3*a*) shows that PCA1 has a different orientation relative to the sample population than PCS1: PCA1 is (approximately) orthogonal to PCS1. Because the regression coefficients relate the variation in the spectrum to variation in the person's age, they must contain information about the composition. Because PCA1 represents the highest concentrations, the orthogonality of PCS1 indicates that the highest concentration constituent(s) of the fingerprint do(es) not change linearly with age. That both PCA and PLS1 account for a similar amount, approximately 70%, of the *X* variance with eight PCs suggests that a similar linear combination of minor constituents produces the observed results. Each analysis simply selects the constituents in its own order.

Regression coefficient plots (Figs. 3*a*, 4*b*, 5*b*) identify IR spectral ranges which in turn correspond to characteristic frequencies of functional groups, and thus to families of compounds (14). Regression coefficients and weighting factors consistently emphasized three regions (Table 1). The observed population age correspondence to the three high-significance regions of the spectrum (Table 1) indicates that the compounds that track linearly with age change several times over the age range 4–70 years. The 2828–2970 cm^{-1} region (C–H stretch, alkane groups) PLS1 score plot shows separation of the samples into four groups: 4–5, 11–14, 18–26, and 29–70 years old (Fig. 6*c*). However, the fact that it is the worst of the four spectral regions for predicting age, (Table 1 columns 3, 4) supports the hypothesis that a linear combination of compounds is necessary to understand age. The 1440–1782 cm^{-1} range corresponds to C–H and C=O. And finally 1190–1260 cm^{-1} range matches C–O oxygenated compounds. This breakdown of functional groups with age is consistent with the previous studies cited in the introduction: basically young people have primarily free fatty acids, and older people have longer chain fatty acids that have been esterified with alcohols. All of these observations are in accord with existing knowledge of the forensic and dermatology communities.

While our results do not allow us to identify specific chemical components in the fingerprint mixtures that account for the observed differences in infrared spectra with age, they do allow us to infer some information about relevant chemical classes and to draw comparisons with the existing literature. For example, it is not surprising that the 2800–3000 cm^{-1} region of the infrared spectrum is a significant wavelength range in discerning the four age groups. Absorptions in this region arise from C–H stretching vibration frequencies of aliphatic groups. The forensic and dermatological literature confirm that changes in the proportion of free long-chain fatty acids and their corresponding esters are generally predictable throughout life. Furthermore, well-documented changes in the proportion of branched versus straight-chain aliphatic long-chain

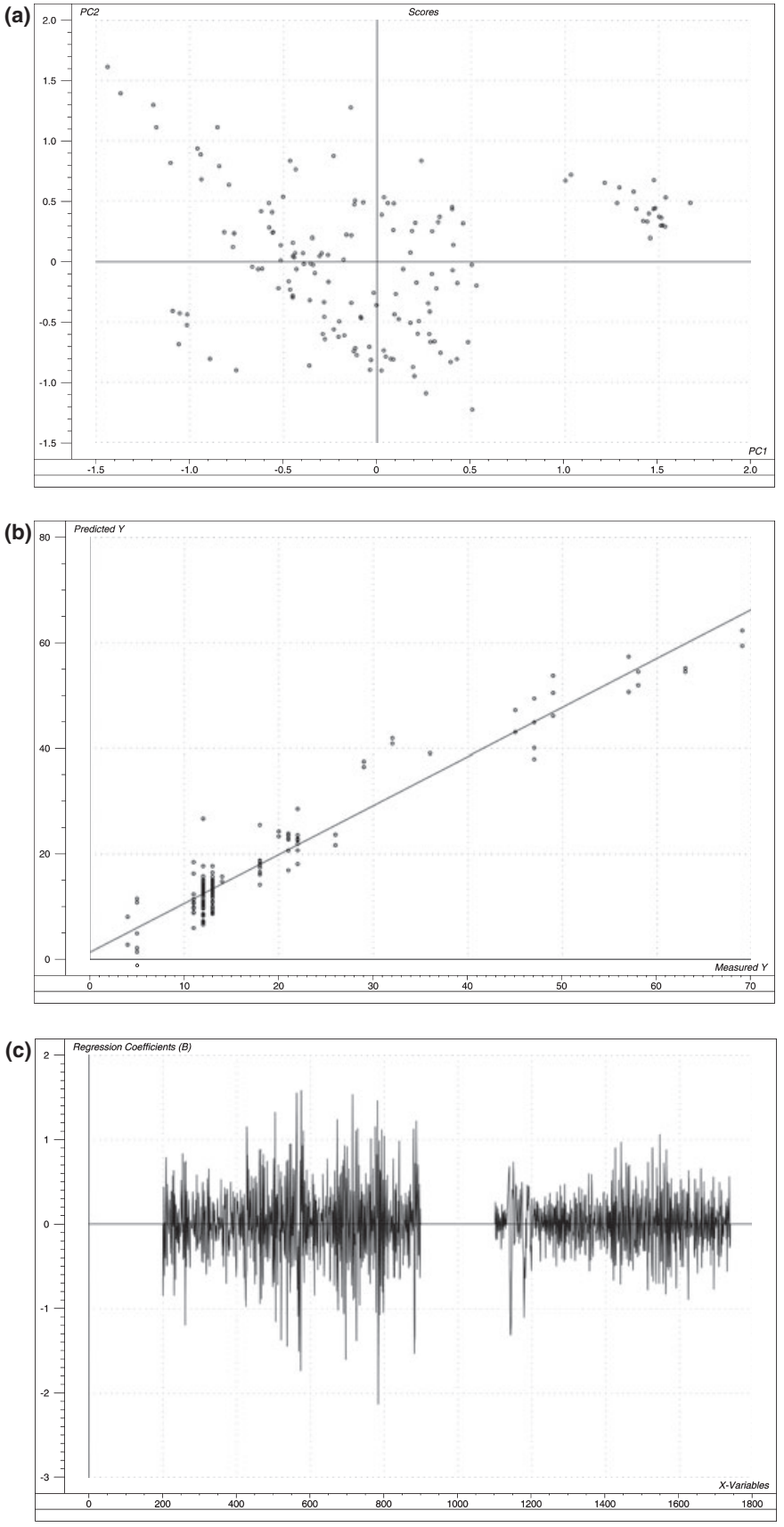

FIG. 3—*(a) PCS2 versus PCS1 score plot for population. PCS1 explains 8% of variance in X spectra. (b) Predicted versus measured for PLS1 regression of population using eight PCs. Slope 0.928; Offset 1.34 years; Correlation 0.963; RMSEC 3.68 years. (c) Regression coefficients for PCS1 of population regression.*

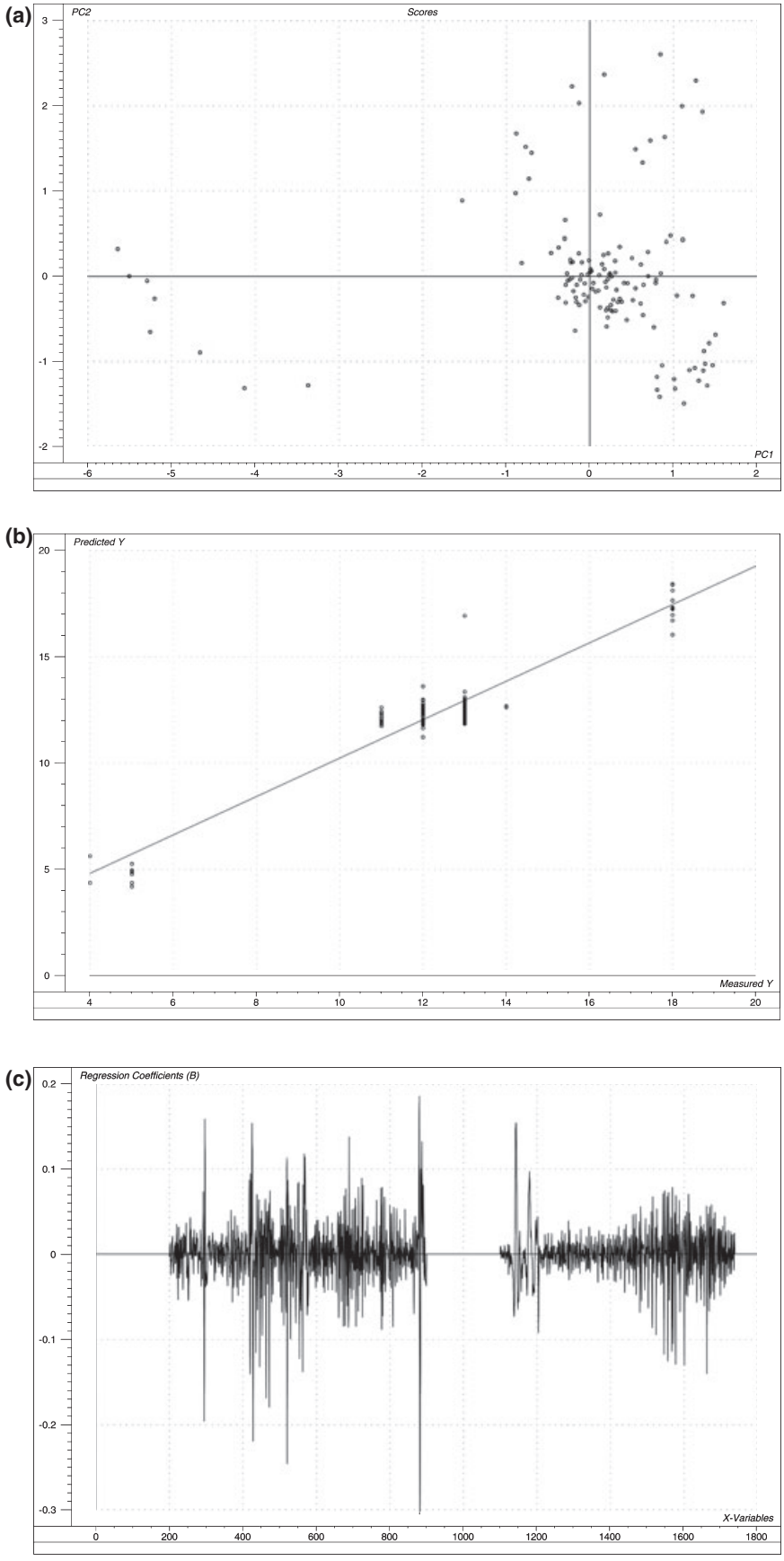

FIG. 4—*(a) PCS2 versus PCS1 score plot for young. PCS1 explains 28% of variance in X spectra. (b) Predicted versus measured for PLS1 regression of young using three PCs. Slope 0.902; Offset 1.21 years; Correlation 0.950; RMSEC 0.78 years. (c) Regression coefficients for PCS1 of young regression.*

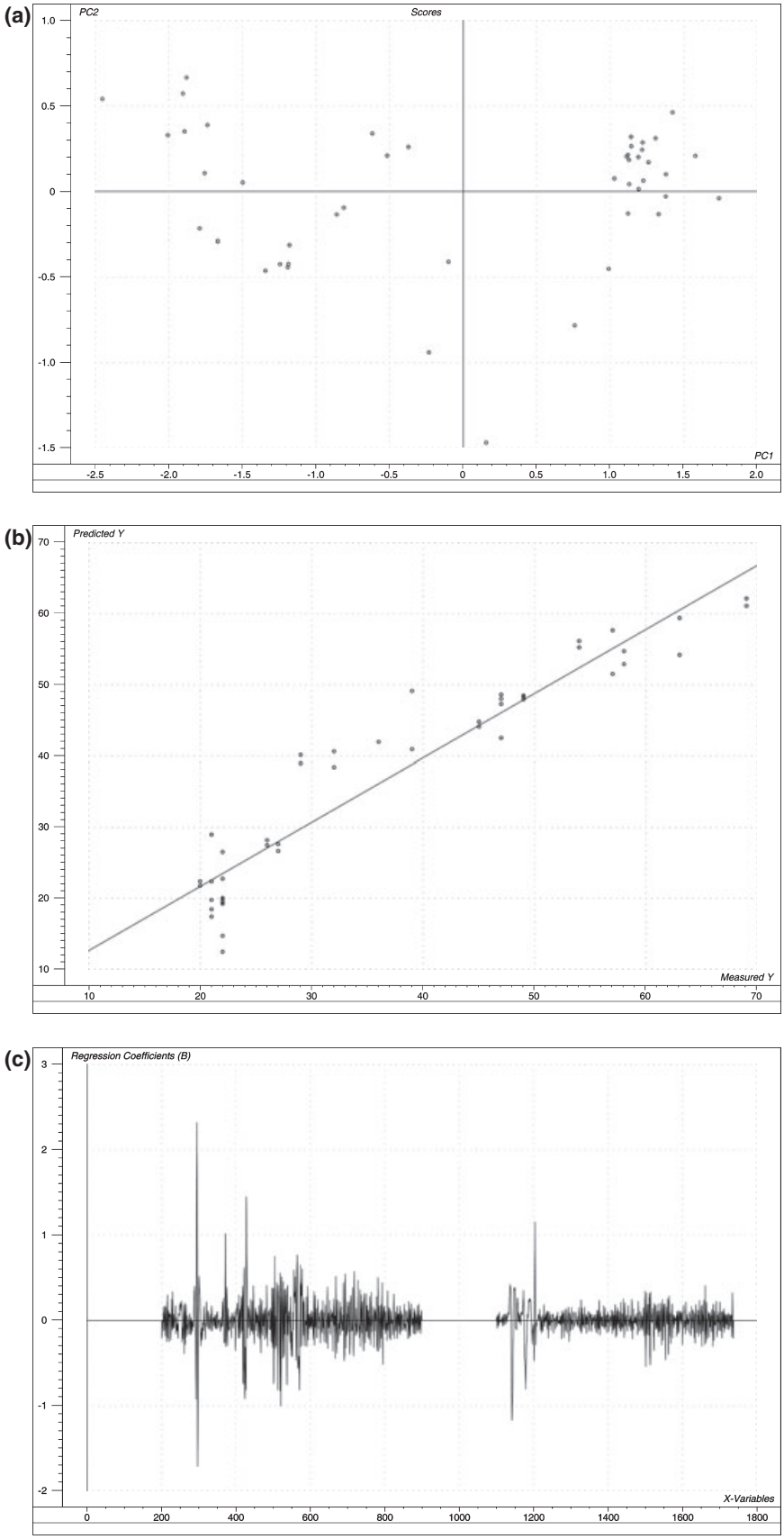

FIG. 5—*(a) PCS2 versus PCS1 score plot for older age group. PCS1 explains 36% of variance in X spectra. (b) Predicted versus measured for PC of older age group regression. Slope 0.903; Offset 3.59 years; Correlation 0.950; RMSEC 4.84 years. (c) Regression coefficients for PCS1 of older age group regression.*

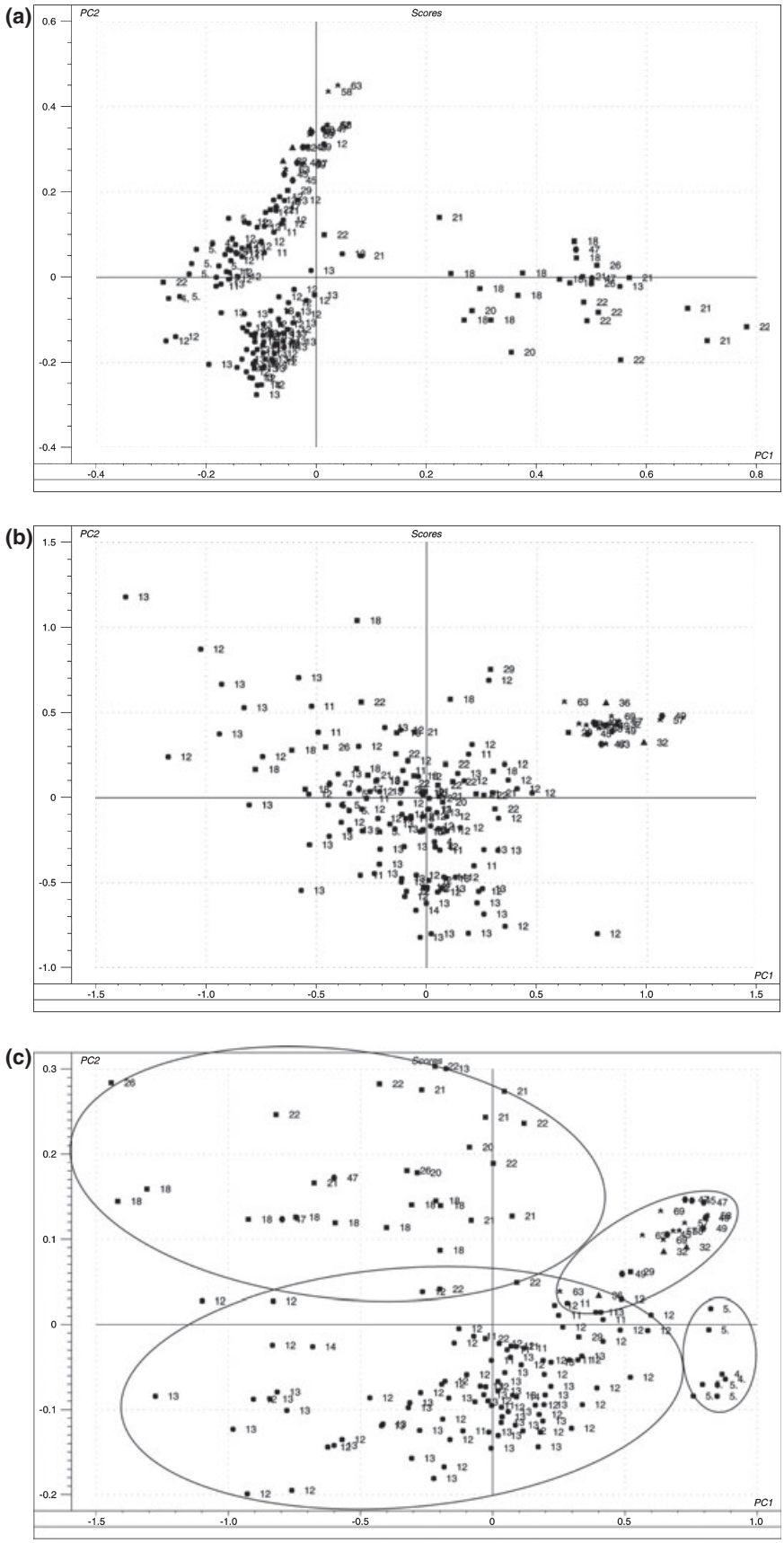

FIG. 6—*(a) PCS2 versus PCS1 1190–1260 cm^{-1} region score plot for population. PCS1 explains 79% of variance in X spectra. Ages shown. Ages 18–22 on right side. (b) PCS2 versus PCS1 1440–1782 cm^{-1} region score plot for population. PCS1 explains 12% of variance in X spectra. Ages shown. Older samples are in upper right. (c) PCS2 versus PCS1 2828–2970 region score plot for population. PCS1 explains 79% of variance in X spectra. Ages show four regions of populations: 4–5, 11–14, 18–26, and 29–70 years old.*

TABLE 1—*Summary of principal component analysis results for age determination.*

Channels	cm^{-1}	Younger	Older	Age Group
280–315	1190–1260	PC1 = 71, none	PC1 = 36, 13 = 90	18–22
415–580	1440–1782	PC1 = 32, 10 = 90	PC1 = 40, 7 = 91	>29
1130–1210	2828–2970	PC1 = 19, 8 = 90	PC1 = 38, 5 = 94	Four groups

Region-by-region analysis of portions of the spectrum for older and younger sample sets. Percent of variance explained by the first principal component PC1, and how many PCs required to explain 90% of the variance. Age group identifies that portion of the population that is best separated by analysis of the corresponding region.

esters should manifest in variations in the frequency and intensity of the C–H stretching vibrations.

Increases in long-chain fatty esters, and changes in the relative concentrations of cholesteryl esters, will show changes in the 1190–1260 cm^{-1} area corresponding to C–O single bond stretching of esters. This is observed and accounts in part for our ability to distinguish subjects in the later stages of puberty from the younger and older age groups. In concert with these changes come variations in the frequencies of C–H bending and C=O double bond stretching regions of the spectrum, from 1440 to around 1782 cm^{-1}. Changes in C–H bending frequencies could also be indicative of the branching of the fatty acid moieties of these long-chain esters. Variations in concentrations of unsaturated fatty acids and esters and squalene may also give rise to changes in the C=C double bond stretching region of the spectra, which also falls in the 1440–1782 cm^{-1} range.

An important application of this technique would be in a case in which a criminalist locates a latent fingerprint at a crime scene for which database searches do not provide a match. This technique could be used to provide an age for the individual even if no other information is known. This should prove very discerning. We recognize that an impediment to field application of this method is that in order to use infrared spectroscopy the latent print must first be located without damaging it. Standard Federal Bureau of Investigation print procedures call for "inherent fluorescence by laser or alternate light source" to be used to locate latent prints before any destructive processing is performed (15). The FBI "Handbook of Forensic Services" also specifically directs the examiner to "examine all evidence visually and with a laser or an alternate light source before using any other latent print development process" (16). While this cannot be done in every case, following the official protocol to nondestructively locate prints is the recommended procedure.

The chemical composition of latent fingerprints has been probed nondestructively to determine the age of the individual responsible for deposition of the prints. Spectral regions contributing to the discriminating power of the method have been identified. Elucidation of the individual compounds responsible for this discriminatory power as well as further investigations of this and other potential information hidden within latent fingerprints continue and may eventually prove valuable to criminal investigators.

Acknowledgment

This research was sponsored by the Southeast Missouri State University Grants and Research Funding Committee (2004).

References

1. Buchanan MV, Asano K, Bohanon A. Chemical characterization of fingerprints from adults and children. In: Hicks J, DeForest PR, Baylor VM, editors. Proc Int Soc for Opt Eng (SPIE) Photonics East Conference; 1996 Nov 20-21, Boston, MA. Bellingham, WA: The International Society for Optical Engineering, 1997;2941:89–95.
2. Blasdell R. The longevity of the latent fingerprints of children vs adults. Policing: Int J Police Strategies Manage 2001;24(3):363–8.
3. Ramotowski RS. Composition of latent print residue. In: Lee HC, Gaensslen RE, editors. Advances in fingerprint technology, 2nd ed. Boca Raton: CRC Press, 2001;63–104.
4. Noble D. Vanished into thin air: the search for children's fingerprints. Anal Chem 1995;67:435A–8A.
5. Victoria Police, Victoria Forensic Science Centre. Fingerprint branch training manual, module nine, latent fingerprint composition. Melbourne, Victoria, Australia, 2002 June 24. Available at http://www.nifs.com.au/F_S_A/Latent%20fingerprint%20composition.pdf (accessed April 1. 2007).
6. Sansone-Bazzano G, Cummings B, Seeler AK, Reisner RM. Differences in the lipid constituents of sebum from pre-pubertal and pubertal subjects. Br J Dermatol 1980;103(2):131–7.
7. Stewart ME, Downing DT. Proportions of various straight and branched fatty acid chain types in the sebaceous wax esters of young children. J Invest Dermatol 1985;84(6):501–3.
8. Yamamoto A, Serizawa S, Ito M, Sato Y. Effect of aging on sebaceous gland activity and on the fatty acid composition of wax esters. J Invest Dermatol 1987;89(5):507–12.
9. Stewart ME, Steele WA, Downing DT. Changes in the relative amounts of endogenous and exogenous fatty acids in sebaceous lipids during early adolescence. J Invest Dermatol 1989;92(3):371–8.
10. Stewart ME, Downing DT. Unusual cholesterol esters in the sebum of young children. J Invest Dermatol 1990;95(5):603–6.
11. Mong GM, Petersen CE, Clauss TRW. Advanced fingerprint analysis project final report: fingerprint constituents. Richland, WA: Pacific Northwest National Laboratory, 1999 Aug. Report No.: PNNL-13019.
12. CAMO Technologies, Inc. The Unscrambler V9.0 [computer program]. Trondheim, Norway: CAMO ASA, 2002.
13. Esbensen K. Multivariate data analysis in practice, 5th ed. Esbjerg, Denmark: CAMO Process AS, 2002.
14. Silverstein RM, Bassler GC, Morrill TC. Spectrometric identification of organic compounds. 5th ed. New York: John Wiley and Sons, 1991.
15. Federal Bureau of Investigation. Processing guide for developing latent prints. Washington, DC: Federal Bureau of Investigation, 2000;4–9. Available at http://www.fbi.gov/hq/lab/fsc/backissu/jan2001/lpu.pdf (accessed April 1, 2007).
16. Federal Bureau of Investigation. Handbook of forensic services. Washington, DC: Federal Bureau of Investigation, Progressive Management, 2003.

Additional information and reprint requests:
David Ritter, Ph.D.
Department of Chemistry
Southeast Missouri State University
One University Plaza
Cape Girardeau, MO 63701
E-mail: dritter@semo.edu