5-2017

Differentiation of Hair Using ATR FT-IR Spectroscopy: A Statistical Classification of Dyed and Non-dyed Hairs

Matthew S. Boll
University at Albany, State University of New York

Differentiation of Hair Using ATR FT-IR Spectroscopy: A Statistical Classification of Dyed and Non-dyed Hairs

An honors thesis presented to the
Department of Chemistry,
University at Albany, State University of New York
in partial fulfillment of the requirements
for graduation with Honors in Chemistry
and
graduation from The Honors College.

Mathew S. Boll

Research Mentor: Kyle C. Doty
Research Advisor: Igor K. Lednev
Collaborator: Ray Wickenheiser

May, 2017

Abstract

Although hair is one of the most common and abundant types of evidence found at a crime scene, the current forensic analyses employed underutilize its full potential evidentiary value. Microscopy is the fundamental technique used to analyze forensic hair evidence, but even this routine and well-accepted method has limitations. In this study, non-dyed and dyed hairs from individuals varying in race, biological sex, and age, were analyzed using attenuated total reflection Fourier transform-infrared (ATR FT-IR) spectroscopy. Through the incorporation of multivariate statistical analysis, spectra collected from dyed and non-dyed hairs were differentiated with high accuracy. After hair spectra were determined to be dyed or non-dyed, dyed hair spectra were successfully differentiated amongst themselves based on brand (or manufacturer) and dye color. The methodology developed here allowed for predicting whether an individual used a permanent hair dye, and then the brand and color of hair dye used, with at least 90% confidence. The high accuracy shown in this study illustrates the ease and robustness of coupling ATR FT-IR spectroscopy and multivariate statistics for forensic hair analysis, specifically for the analysis of dyed hairs. The use of spectroscopy for forensic hair analysis, as demonstrated by this proof of concept study, would advance the field of trace evidence as a whole, and can potentially be utilized to confirm conclusions drawn from methodologies employed currently, in turn leading towards increased individualization.

Acknowledgements

I would like to thank Jeremy Manheim for introducing me to the growing field of forensic hair research, and Kyle Doty for extending tremendous guidance and patience during my research endeavors. I would also like to extend a special thank you to Professor Igor Lednev, and Professor Paul Toscano for their meaningful advice regarding my research and course work. Last but not least, I want to express my gratitude to my family, whose encouragement throughout my academic career provided me with the motivation to excel in and out of the classroom and laboratory.

This project was supported by Award No. 2014-DN-BX-K016, awarded by the National Institute of Justice, Office of Justice Programs, U.S. Department of Justice. The opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect those of the Department of Justice.

Table of Contents

1. Introduction

Hair is found everywhere and therefore common to nearly every crime scene. On average a person loses up to 100 scalp hairs every day [1,2,3]. Since it can vary significantly between individuals, and is copious in nature, hair has the potential to be an important type of evidence used throughout an investigation. The utilization of such trace evidence can be instrumental to criminal investigations because of the many ways that hair can be classified. Yet, the forensic analysis techniques currently employed underutilize the amount and important types of information that hair evidence can provide. Hair can differ in type (e.g., curly, wavy, straight, etc.), length, density, color, thickness, area of body (somatic) origin, etc., and varies among populations [4]. Although the individualization of hair evidence is not currently feasible without obtaining a DNA profile, the ability to distinguish hair samples by several class characteristics is helpful when attempting to narrow down or identify suspects of a crime. Specifically, this information can aid in the exclusion or inclusion of a certain type of person. Therefore, the ability to acquire more information about forensic hair evidence quickly, nondestructively, and, most importantly, with associated statistical confidence levels, would significantly benefit forensic investigations on scene, in the laboratory, and in the court room. Investigative leads can be provided regarding suspects based on rapid microscopic hair comparison, and hairs can be non-destructively screened for DNA analysis suitability.

Traditionally, hair analysts have used transmitted light and polarized light microscopy to provide information about the physical characteristics of hair and other fiber evidence [3,5]. A trained forensic examiner can use light microscopy to classify hair and identify alterations to hair, such as if a hair was dyed, burned, or cut [5,6]. The reliability of microscopic hair classification and comparisons is restricted by several factors, which include, but are not limited to, analyst experience and training, the condition of the hair sample, and the capabilities of the instrumentation being used [3]. Obtaining a DNA profile from the genetic material in hair to identify the individual is a well-established technique, however there are several limitations to performing this analysis as well. The current methodology is expensive and it can take a significant amount of time to process each sample, not to mention the large number of backlogged cases experienced by most DNA laboratories [7]. Additionally, a hair follicle or root sheath, which is ideal for nuclear DNA extraction, is only present during the hair's major growth phase (anagen hairs) [3,7]. Hairs lost naturally or shed (catagen or telogen hairs) often lack the follicular material, which makes nuclear DNA analyses nearly impossible and is usually left to well-funded and resourceful labs, such as the Federal Bureau of Investigations (FBI) crime laboratory [7]. The shaft of the hair does not contain nuclear DNA, but a mitochondrial DNA profile can be obtained. However, the mitochondrial DNA profile is not a unique identifier, like nuclear DNA [7,8]. The extraction of mitochondrial DNA is typically only utilized under unusual circumstances because it is even more costly and obtaining a full comparable profile is not guaranteed [9]. DNA analysis is also destructive to hair, fully digesting all or a significant portion of the hair in question.

There have been several studies focused on the use of vibrational spectroscopy to analyze hair samples for forensic purposes [10-14]. Members of Queensland University of Technology (QUT) have evaluated the use of Fourier transform-infrared (FT-IR) spectroscopy to analyze hair and discriminate between individual characteristics such as biological sex and race, as well as between hair spectra based on hair treatment products (i.e., permanent and semi-permanent dyes). In their 2004 study [10], members of QUT preliminarily established the use of FT-IR spectroscopy for forensic hair analysis, and then in following studies focused on the advantages of using an attenuated total reflection (ATR) attachment and the classification capabilities of chemometrics [11,12]. More recently, Kurouski and coworkers demonstrated the successful use of surface-enhanced Raman spectroscopy (SERS) to detect and identify if hairs were dyed or non-dyed [13]. While having the challenge of a lengthy sample preparation procedure, Kurouski et al. took the analysis one step further and differentiated between the dyes themselves. The QUT studies had a shorter sample preparation procedure, however their 2004 study sample preparation included complete physical flattening (with a rolling device) of the hair samples prior to analysis, which significantly alters an evidentiary sample. An ATR attachment can be used in lieu of flattening entire hair fibers. Furthermore, in most of these studies details regarding the origin of hair samples came from is unclear, few classification models were used for analysis, and the models constructed lacked the use of a crucial external validation step. In contrast, a recent study by Manheim and coworkers demonstrated the use of ATR FT-IR spectroscopy for hair analysis to successfully and accurately classify and discriminate between hair from humans, cats, and dogs using chemometrics, which incorporated an external validation step [14].

There have been two major reports that scrutinized several methodologies of forensic science. The 2009 National Academy of Science (NAS) report emphasized the need for the incorporation of statistics in emerging methodologies to improve the reliability of analyzing evidence [15]. Although the report focused on the individualization (such as with DNA profiling) of hair and other types of evidence, the importance that new techniques utilize statistics in order to support the conclusions drawn from current hair analyses was also emphasized [14]. The more recent 2016 Presidential Council of Advisors on Science and Technology's (PCAST's) report to the president also cited perceived flaws associated with microscopic hair analyses [16]. The concepts outlined in each of these reports emphasize the need for new methodologies that are capable of acquiring information about forensic hair evidence quickly and nondestructively, with associated statistical confidence levels. Application of statistics improves objectivity and also provides for the comparison of data across bodies of information, whereas current microscopic hair comparisons can only be conducted on a one to one or side by side basis. Although the possibility for development of new methodologies is vast, they should only complement results obtained from microscopic evaluation of hair evidence, and by no means replace that science completely. Wickenheiser and Hepworth raised questions about the subjectivity and validity of microscopic hair analysis in 1990, however established the value of using features differentiating hairs [17]. Further, in 2004 Betty et al. showed that although hair evidence analysis is flawed, it is helpful as associative

evidence, and meets all requirements for admissibility in criminal and civil courts [18]. Current hair analyses are shown to be scientifically valid and accepted, but, by incorporating new techniques with the preexisting ones, more robust conclusions can be reached, thereby leading to increased acceptance and support of the findings and underlying science overall.

The study shown here proposes a technique that incorporates all of the aforementioned desired factors by utilizing ATR FT-IR spectroscopy for forensic hair analysis. Some of the previous studies, such as those carried out at QUT and the one by Manheim et al., have shown the strength and versatility of using ATR FT-IR spectroscopy for forensic hair analysis. This work serves as an expansion on the previous studies and emphasizes the capability of the technique as being an effective forensic tool for hair analysis. Specifically, differentiation of dyed and non-dyed hairs as well as distinguishing between dyed hairs themselves to identify color or brand of the dye was implemented.

2. Materials and Methods

2.1 Hair Samples and Dyes

All hair samples for this study were obtained from volunteers who were willing to donate a small amount of scalp hair. For all hair donors, only their age, biological sex, and race were recorded; donor names and all other personally identifiable information was not collected. Hair samples were acquired from eleven donors, varying in age, biological sex, and race. The donated hairs were all natural in color (non-dyed) and included blonde, brown, black, and red samples. Hair samples were collected from a diverse donor pool to ensure that the statistical models built for differentiation purposes in this study were not biased or affected by a person's biological sex or race. Table I lists specific information about all eleven hair donors, including their age, biological sex, race, natural hair color, hair dye(s) used, and whether they were included in the calibration or validation dataset for all classification models constructed.

Dyes purchased for this study were selected to ensure consistency between brands. Specifically, two different colors of dyes from three different brands were used to treat the hair samples. These included Revlon medium brown (RMB), Revlon black (RB), Clairol medium brown (CMB), Clairol black (CB), Just for Men® medium brown (JFMMB), and Just for Men® black (JFMB). All of the samples were dyed in individual petri dishes to avoid cross-contamination. The dyes were prepared and applied according to the in-box instructions, which usually included dye color-activator mixing/preparation, a timed dying period, and a conditioner wash. Subsequent to dye application, hairs were rinsed with tap water, allowed to dry and left to sit for a week on a laboratory bench under ambient conditions before spectra were collected.

Table I. Information about each hair donor including age, biological sex, race, and natural hair color. Also listed are the dye(s) that the donor's hair was treated with and whether the donor was included in the calibration (training) dataset or the external validation (test) dataset for the classification models built.

Donor #	Donor's age	Donor's biological sex	Donor's race	Donor's natural hair color	Dye(s) used	Calibration or Validation
1	20	Male	Caucasian	Red	RMB, CMB, JFMMB	Calibration
2	20	Female	Caucasian	Light brown	RB, CB, JFMB	Calibration
3	22	Male	African American	Brown	RB, CB, JFMB	Calibration
4	N/A	Female	African American	Black	RMB, CMB, JFMMB	Calibration
5	20	Male	Hispanic	Brown	RMB, CMB, JFMB	Calibration
6	20	Female	Hispanic	Brown	RB, CB, JFMMB	Calibration
7	24	Male	Asian	Black	RB, CB, JFMMB	Calibration
8	34	Female	Asian	Black	RMB, CMB, JFMB	Calibration
9	20	Male	Indian	Black	JFMB	Validation
10	20	Female	Caucasian	Blonde	CMB	Validation
11	16	Female	Asian	Black	RMB	Validation

2.2 Instrumentation and Spectral Collection

All of the hair samples were analyzed using a PerkinElmer Spectrum 100 FT-IR spectrometer with an ATR attachment. The spectral range used was from 650 cm^{-1} to 4000 cm^{-1}, at a resolution of 4 cm^{-1}. The ATR attachment utilized a single bounce reflection diamond/ZnSe crystal. Prior to any sample measurement, and in between samples, the ATR attachment, crystal, and sample platform were cleaned with ethanol then wiped dry before a background spectrum was collected and

subtracted to “blank” the instrument. To collect spectra, hair samples were simply placed over the crystal and compressed using the ATR attachment. Ten “trials” were performed on each hair sample, where each trial was at a different point along the length of the hair shaft. Ten spectra were collected at each of the ten points along the hair shaft and combined for each “trial”. For short hairs, 2-3 strands from the same donor were needed to accumulate the ten total trials for that sample. Ten trials of spectra were used for each of the hair samples prior to dying and after they were dyed for all eleven donors. A total of 110 non-dyed spectra and 270 dyed spectra were collected.

2.3 Hair Classification Models

All of the spectra were converted into data matrices and loaded into PLS Toolbox 7.0.3 (Eigenvector Research, Inc.) within the MATLAB (The MathWorks, Inc.) platform. Partial least squares discriminant analysis (PLSDA) was used for classification and differentiation of dyed and non-dyed hairs. Prior to PLSDA model construction/validation the spectra were first truncated to 650-1800 cm^{-1} to enhance discrimination performance [14]. All of the spectra were also preprocessed prior to model construction/validation with the following steps: conversion from transmittance to absorbance, second order derivative conversion (Savitzky-Golay algorithm), total area normalization, and mean centering.

Three different types of classification analyses were carried out. Classification models were constructed using PLSDA to: (a) determine if a hair was dyed or non-dyed, (b) distinguish between the different brands of dye (i.e. Revlon, Clairol, or Just For Men®), and (c) discriminate between the color of the hair dye (i.e. black or medium brown). For all three PLSDA models, eight of the eleven total hair donors (320 spectra) were arbitrarily chosen and used as the calibration (training) dataset. The three other donors (60 spectra) were used to externally validate the developed classification models. The success of each PLSDA model was assessed by its internal cross-validation and external prediction sensitivities and specificities.

3. Theory

IR spectroscopy can provide specific chemical information about forensic evidence, such as hair. Although detailed, the vast amount of information provided by the spectroscopic technique is often difficult to interpret because of the hundreds of variables that need to be considered. Chemometrics is the application of advanced statistics to chemical systems in order to enhance data interpretation [19]. Chemometrics augments the analysis of large datasets, like the 380 spectra analyzed in this study, which included over 438,000 data points (i.e., intensity values of each spectrum for each variable, or wavenumber). Specifically, distinctive relationships between certain types of samples can be extracted or determined.

PLSDA is a classification technique that is derived from partial least squares regression and utilizes latent variables (LVs), which are linear combinations of the original variables, to represent the variation in the dataset [20]. The LVs allow for characterization of specific and significant differences between spectra, instead of

analyzing every single difference in the variables [21]. Furthermore, the LVs "capture" the variance between classes of spectra and are used when making predictions. Through PLSDA, graphical presentations of data can be created, in the form of scores plots, which simplifies data interpretation [22]. More importantly, PLSDA is capable of differentiating and grouping together hundreds of spectra that contain thousands of variables and is therefore ideal for this study.

4. Results and Discussion

In this study, ATR FT-IR spectroscopy was used to collect spectra from natural human hairs prior to and after dying with over-the-counter commercial hair dyes. The dyes were from three different brand manufacturers (Revlon, Clairol, and Just For Men®) and of two different colors (black and medium brown). ATR FT-IR spectroscopy is a technique that almost exclusively probes the surface of hair samples due to the evanescent wave only penetrating a couple micrometers (at most) into the sample being analyzed. Whether or not the dye penetrates deep within a hair shaft, it would almost certainly be present on the surface of the hairs where it was deposited. Permanent hair dyes utilize precursor molecules to oxidize hair, which allows color molecules to diffuse into and bind to the hair [11]. Therefore, we would expect visual differences in the spectra from dyed and non-dyed hairs by using ATR FT-IR spectroscopy. For classification purposes, PLSDA was utilized to specifically target three types of hair differentiation: (a) dyed versus non-dyed, (b) Revlon versus Clairol versus Just For Men® (dye brand), and (c) black versus medium brown (dye color).

4.1. Differentiation between dyed and non-dyed hairs

The first step in this three-part study was to differentiate between dyed and non-dyed hairs. Since the chemistry of hair is altered subsequent to dying, spectral differences would be expected, which would result in strong differentiation ability. Figure 1a shows the raw mean spectra for the dyed (green trace) and non-dyed (red trace) hair classes used in the calibration dataset.

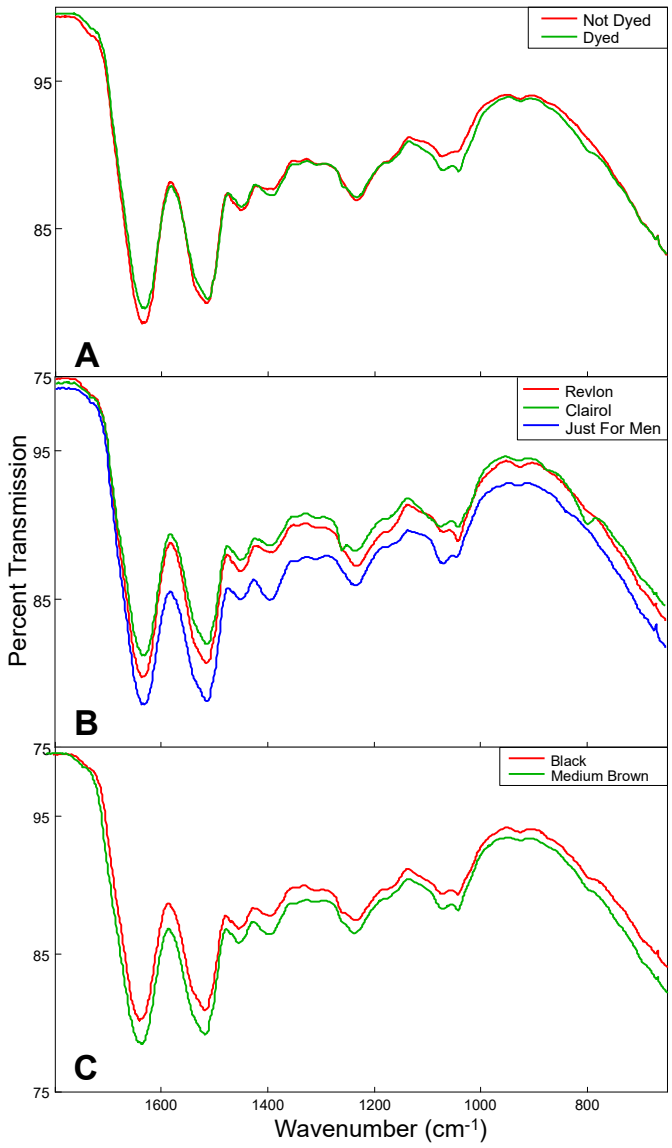

Figure 1. Overlay of the raw mean spectra used for PLSDA model training for differentiating between dyed and non-dyed hairs (A), dye brands (B), and dye colors (C).

There are few apparent differences between the two classes based on the raw mean spectra. However, the preprocessing steps allow for visualizing differences in the infrared signals' location and intensity. By visual inspection of the preprocessed mean spectra (not shown), the two classes of spectra differ extensively in the regions 662-673 cm^{-1} (sulfate stretching) and 1250-1270 cm^{-1} (ester stretching) [23, 24]. The two classes also differ in their intensity at 1040 cm^{-1}, a result of hair oxidation, specifically the oxidation of cysteine to cysteic acid [25]. These signals, among others were used to differentiate between dyed hair and non-dyed hair classes, and are highlighted in the latent variables (LVs) shown in Figure 2. The LVs emphasize particular differences between classes of spectra. They focus on a percentage of all of the differences between the classes and exclude possible

spectral noise to emphasize strong correlations. PLSDA was used to construct a model incorporating five LVs that captured 48.90% of the total variance in the 320 dyed and non-dyed hair spectra included in the calibration dataset. As shown in Figure 2a, LV1 of the first PLSDA model captures 16.54% of the total variance including a contribution from the sulfate stretching (662-673 cm^{-1}) and ester stretching (1250-1270 cm^{-1}) vibrational modes can be observed resulting in the differentiation of the dyed and non-dyed classes.

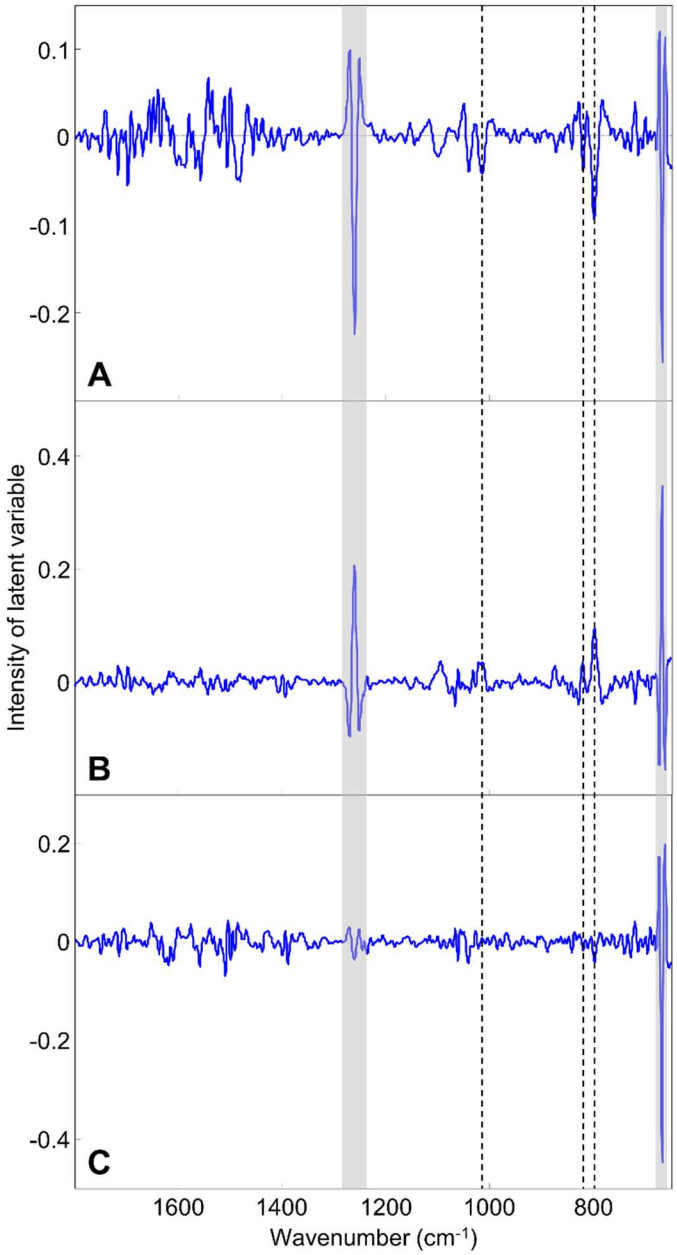

Figure 2. Latent variable one used to build PLSDA model for differentiating between dyed and non-dyed hairs (A), dye brands (B), and dye colors (C).

LV1, which captures the most variance compared to the other LVs, shows a large difference in the dyed and non-dyed classes for the sulfate (SO_4^{2-}) and ester (C-O) stretching modes. There are common esters included in the ingredients for all of the dyes used for this experiment, such as erythorbic acid, shown in Supplemental Table I, so intensity changes of ester signals are expected. Erythorbic acid is a stereoisomer of ascorbic acid and contributes to FT-IR signals at 794 and 819 cm^{-1}, which are also highlighted in LV1 of this model, and in the dye brand model that was constructed separately (Section 3.2). The sulfate signal can be explained by the chemistry of hair dyes. Most permanent dyes include ammonia, or other alkaline chemicals, to increase permeability in the cuticle for dye penetration. Once the cuticle has been compromised, an oxidizer, such as peroxide, can penetrate the cortex of the hair and break disulfide bonds so that new bonds with the dye molecules can be made [26]. The breaking of these bonds in hair results in the release of sulfur, instigating a change in the sulfate ATR FT-IR signal. Large chemical changes to hair, in addition to any latent changes, allow for strong classification capabilities. Shown in Figure 3 is an internally cross-validated (CV) prediction plot for hair spectra to be classified as belonging to the dyed hair class.

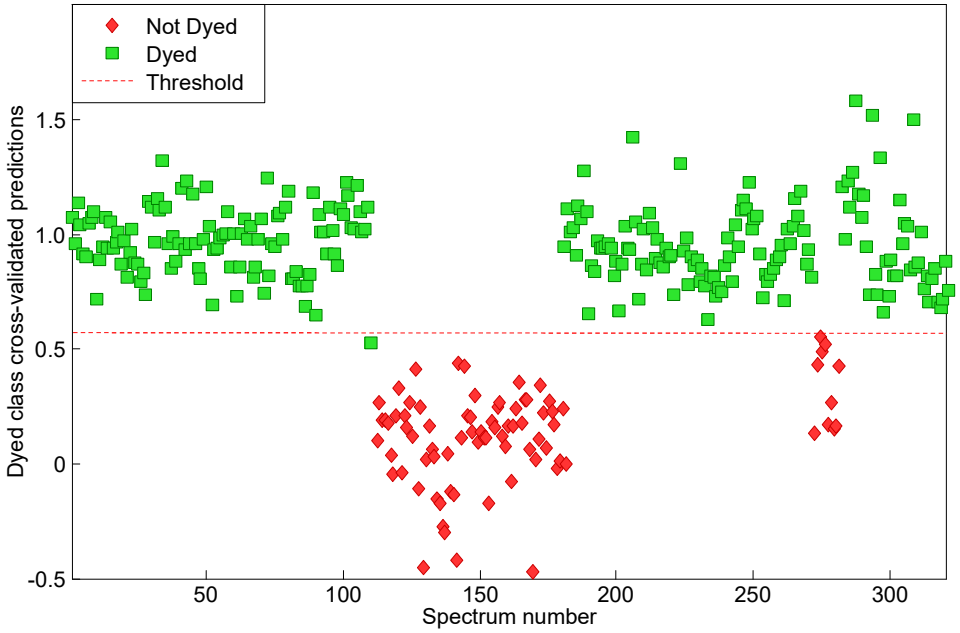

Figure 3. The cross-validated model predictions for dyed hair. The red dotted line separating dyed from non-dyed classes represents the default classification threshold.

A CV prediction plot shows the results for the most stringent type of calibration model testing. In this case the plot illustrates the predictive probability that a given spectrum will be classified as a dyed hair. The CV plot was constructed by performing the venetian blinds technique on the data matrix, where ten spectra were removed from the calibration dataset and loaded back in as “external” data for class predictions to be made. Each spectrum was assigned a class prediction probability of belonging to a particular class, with respect to the character of the LVs. In this

case, the lower the predicted probability, the more likely the spectrum was to be classified as non-dyed, whereas the higher probabilities corresponded to classification as the dyed class. The red dotted (threshold) line in Figure 3 represents the cut-off for whether a hair spectrum was classified as dyed or non-dyed. Only one of the dyed hair spectra fell below this threshold (green square; spectrum number 110). The spectrum was from a donor whose hair was dyed with the JFMMB hair dye. Spectrum 110 is only one out of the forty total spectra collected from that donor; all thirty-nine of the other spectra for that specific donor were correctly classified. Therefore, the prediction accuracy was 97.5% for that donor, but only decreased the cross-validation prediction accuracy to 99.6%. Additionally, the misclassified spectrum (number 110) from donor #1 was not predicted as belonging to the non-dyed class in other, less-stringent, calibration prediction plots such as strict class predictions or most probable class predictions (results not shown). Conversely, 100% of the non-dyed spectra were assigned probabilities below the threshold, and were not classified as dyed hair spectra. The successful separation of dyed and non-dyed spectra suggests that permanent hair dyes sufficiently change the chemistry of the hairs, at least on or near the surface (where the FT-IR beam penetrates), leading to effective differentiation.

Most of the dyed and non-dyed hair spectra had a near perfect probability of being predicted correctly by the first, binary PLSDA model. The CV sensitivity and specificity for the dyed class were 1.00 and 0.996, respectively, whereas the sensitivity and specificity for external predictions were 0.933 and 0.967, respectively. Sensitivity is the probability of the model predicting a spectrum as the correct class, and specificity is the probability of the model correctly not predicting a spectrum from one class as the other class. The high sensitivity and specificity of the model are exemplified by the exceptional classification results.

Strict class model predictions are determined based on which class a spectrum resembles most closely. Spectra are scored based on the resemblance to the character of LVs for each class. If a spectrum is predicted to have a probability above 0.50 for being assigned to one class, and only that class, then it is classified as such. All of the calibration spectra that were used to construct this model were correctly classified under strict class predictions. Sixty non-dyed and dyed hair spectra (thirty from each class) that were not used to construct the PLSDA model were used for external validation. Using completely unknown spectra for this external validation step is a method to rigorously test the model's prediction performance. It is similar to how the methodology would work if used for real criminal casework samples found at a crime scene. Of the sixty total validation spectra, only three were misclassified, resulting in an average prediction accuracy of 95.0% for the external validation spectra. The three misclassifications consisted of two spectra from undyed hair samples from donor #11 and one spectrum from a hair sample dyed with JFMB from donor #9. The preprocessed spectra from these misclassifications had lower than average intensities for characteristics of their respective classes and lacked the ester (C-O) stretching signal at 1250-1270 cm^{-1}.

4.2. Differentiation between dye brands

Once hair spectra were classified as either dyed or non-dyed, deeper analysis and further classification was performed on the dyed hair spectra only. The first approach for dyed hair classification was to distinguish between the brands of dyes used to color natural human hairs. Each brand contained signature ingredients that were consistent for the two colors of dyes used in this study, so it was expected that the brand classes would be differentiated according to these ingredients. In the Revlon hair dyes, oleic acid and sodium benzotriazolyl butylphenol sulfonate were common ingredients between the two colors. In the Clairol hair dyes, ammonium hydroxide and disuccinates were common. Petrolatum, a sealing agent, was a common ingredient between the two Just For Men® hair dyes used. The differences between ingredients from each of the three brands led to slight spectral differences. The mean truncated raw spectra of dyed hair samples grouped by dye brand, and the first LV from the PLSDA model used for their differentiation, are shown in Figure 1b and 2b, respectively.

Differences in the location and quantity of peaks are not observed between brands, but large intensity differences do exist between spectra for the three brand classes. Specifically, there are major intensity differences in the regions 662-673 cm^{-1} (sulfate stretching) and 1250-1270 cm^{-1} (ester stretching) as well as the peaks at 794, 819, (erythorbic acid) and 1040 cm^{-1} (cysteine oxidation). The intensity differences that led to spectral classification can be attributed to how the dyes mask or amplify particular signals by penetrating and reacting with the hair differently, most likely a result of the amount of oxidizing agents used by each brand. For example, differences in the 662-673 cm^{-1} sulfate signal or the 1040 cm^{-1} cysteine oxidation signal are a result of the oxidation of hair. For differentiation between dye brands, a second PLSDA model was constructed with six LVs, which captured 60.64% of the total variance in the calibration dataset. As shown in Figure 2b, LV1 exemplifies the differentiation between brand classes according to the observed signals in the mean preprocessed spectra, particularly for the sulfate stretching (662-673 cm^{-1}) and ester stretching (1250-1270 cm^{-1}) regions.

Despite the many spectral similarities between the hair samples dyed with different brands, the second PLSDA model was able to discriminate between them with high accuracy. In Figure 4 the CV predictions of each brand class can be observed as a three-dimensional scatter plot. The three axes depict how well the spectra of each brand of dye separate from each other as a result of the probability they will be predicted to their correct, respective, class. Although there is obvious separation between brand classes, and grouping within each dye class, some spectra do seem to overlap between brands. However, this figure only represents one way of visualizing the separation. It can be rotated in three-dimensional space to see the separation better from other angles.

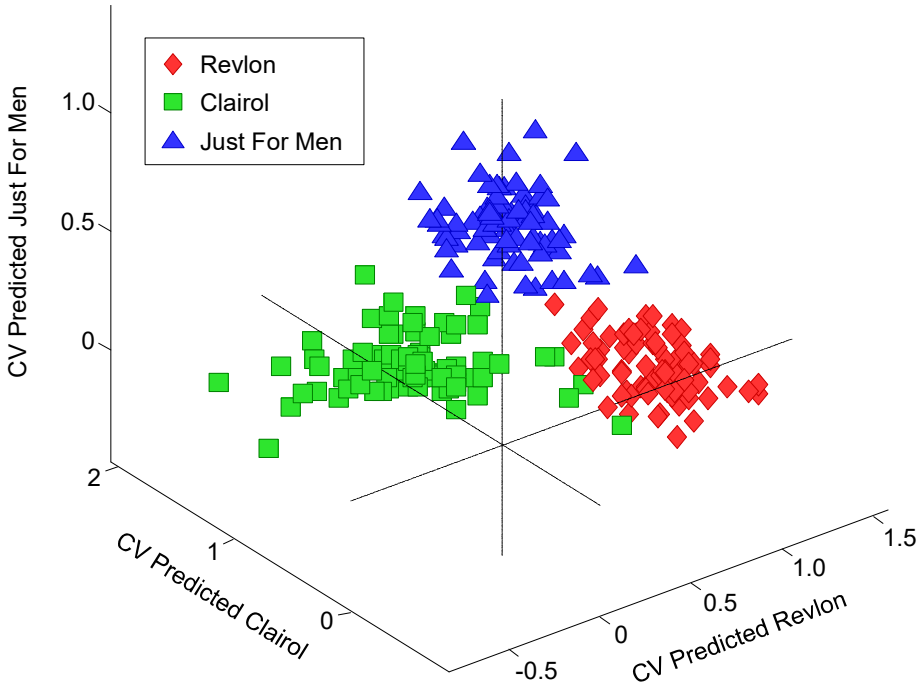

Figure 4. A three-dimensional scores plot showing the internally CV class predictions of the three brands of hair dye used to treat hairs.

A strict class predictions plot was also used to classify all of the spectra (calibration and validation) into the correct dye brand classes. During the strict class predictions, some of the spectra were predicted to the unclassified class. Spectra are predicted as unclassified if they have a probability of prediction of ≥ 0.50 for more than one class or < 0.50 for all classes, where 0.50 is the default prediction threshold for the model. The unclassified class prevents the model from producing unfavorable misclassifications, particularly for samples it has not been trained to take into account. The strict class predictions plot for the hair dye brand class model is shown in Figure 5.

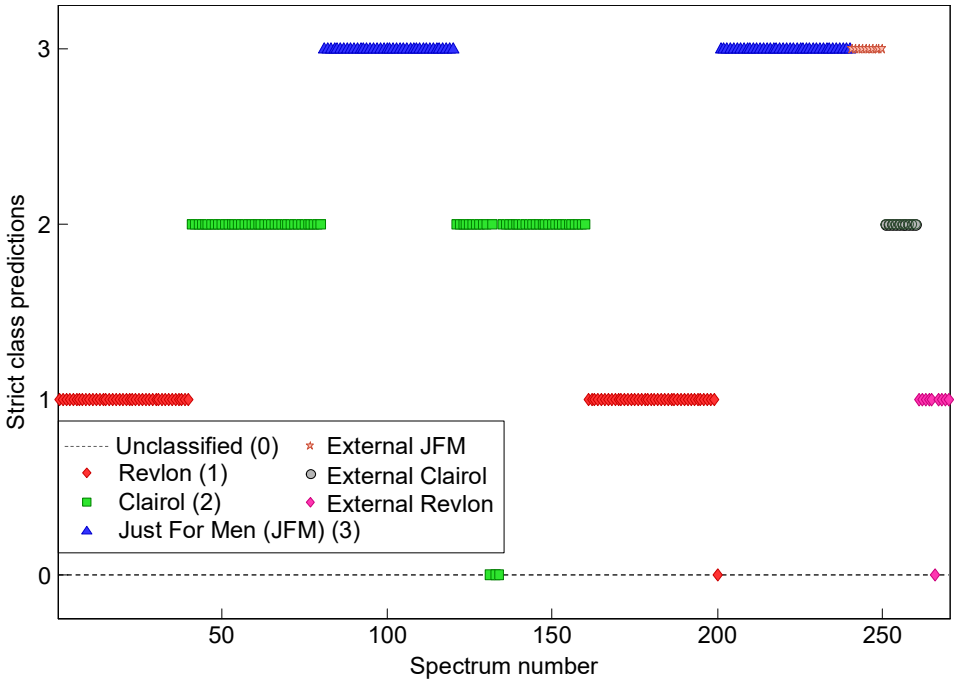

Figure 5. The strict class prediction plot for the hair dye brand classes, showing the classification of all 290 dyed spectra including external-validation (test) spectra.

Of the 240 calibration spectra, three were predicted as unclassified. In theory, this class prediction output is better than spectra being misclassified as one of the other dye brand classes, and is similar to reaching inconclusive results in other methodologies. Of the thirty validation spectra, twenty-nine were correctly classified as their respective brand of hair dye and one was predicted as unclassified, resulting in a 96.7% average prediction accuracy for the validation spectra. This PLSDA model had a CV sensitivity of 0.988, 0.963, and 1.00 for the Revlon, Clairol, and Just For Men® dye brand classes, respectively. The CV specificity for the same respective dye brands was 0.969, 1.00, and 0.988. The prediction sensitivity for each class was 1.00 and the prediction specificity were 1.00, 0.950, and 1.00 for Revlon, Clairol, and Just For Men® classes, respectively. Although all of the spectra were not correctly predicted as their appropriate brand classes, the model was still successful in avoiding the misclassification of spectra into the wrong brand class since all misclassified spectra were predicted as unclassified under strict class predictions.

4.3. Differentiation between dye colors

The third, and final, type of differentiation tested in this study was between hairs dyed with two different colors (from each of the three brands). For this approach, the spectra were labeled as “Black” or “Medium Brown”, irrespective of the dye brand. Both of these color classes contained dyed hair spectra from all three dye brands, and it was expected that spectral classification would not be as successful as the models previously discussed (3.1 and 3.2) as a result of a lack in class determining

ingredients. To classify the spectra, another PLSDA model was built, with five LVs, capturing 53.78% of the total variance in the calibration dataset.

Visually, there were very few differences between the mean spectra of these two dye color classes, especially in comparison to the number of spectral differences that exist between the two other forms of differentiation explained previously (Sections 3.1 and 3.2). Though the mean spectra, shown in Figure 1c, appeared to have little to no differences there was still an exaggerated difference between the mean spectra at 662-673 cm^{-1} (sulfate stretching), according to the first LV as shown in figure 2c. For the most part, the dyes did not vary greatly in the amount and types of ingredients between the two colors, as they did between brands. The lack of variation between the mean spectra of hair dyed with the two different colors was expected to lead to a markedly higher misclassification rate as compared to the other two PLSDA models built, however this was not observed.

The dye color PLSDA model was just as successful as the other two PLSDA models in being able to predict the spectra to their correct class. During strict class predictions two of the calibration spectra (from two different donors belonging to the medium brown color class) were misclassified as belonging to the incorrect dye color class. For the black dye color class, the model had a CV sensitivity and specificity of 0.983 and a prediction sensitivity and specificity of 1.00. Even with the two misclassifications of calibration spectra, the model correctly classified all of the thirty validation spectra, a 100% prediction accuracy.

4.4 Comparison of PLSDA models

Since this study incorporates three different types of classification analyses, it is important to demonstrate the overall effectiveness of this proof of concept methodology. To compare the performance of the three PLSDA models constructed, the prediction accuracies of the spectra used for internal cross-validation and external validation should be considered. These values are shown in Table II. With the lowest accuracy being 90.0%, on the spectral-level, this is very promising. This number increases on the donor-level, where the overall correct prediction accuracy of each donor (instead of all spectra for an entire class) was 100%. From a forensic standpoint this is much more important and practical because the requirement of such a method as the one developed here would be the ability to determine, for any single person, whether a hair is dyed or non-dyed, and, if dyed, the color and brand of the dye.

Table II. Comparison of prediction accuracies based on all spectra (spectral-level) used for internal cross-validation and external validation for each of the three PLSDA models constructed.

Model	Class	Spectral-level prediction accuracy	
		Internal Cross-Validation	External Validation
Dyed vs. non-dyed	Dyed	99.6%	96.7%
	Non-dyed	100%	93.3%
Dye brand	Revlon	98.8%	90.0%
	Clairol	96.3%	100%
	Just For Men®	100%	100%
Dye color	Black	98.3%	100%
	Medium Brown	97.5%	100%

The external validation portion of this study served as a real-world scenario test. The validation spectra were from donors that the model had never encountered before, ensuring that differentiation was not a result of hair donor characteristics, and that internal validation accuracy was not a result of overfitting the model(s). Hair spectra from donors 9, 10, and 11 were classified with high accuracy by each PLSDA model. It was demonstrated that by using the three models constructed for this study, the ability of classifying individuals, who have not yet been exposed to the model, as having hairs dyed or non-dyed with 90.0% confidence. If it was determined that the individual did have dyed hair, then we are capable of classifying the brand and color of the dye with 90.0 and 100% confidence, respectively.

5. Conclusions

The three PLSDA models constructed for this study resulted in an average spectral prediction accuracy of 98.1% ± 3.0% despite the race, age, or biological sex of the hair donor. Although some of the hair spectra were misclassified, the average prediction accuracy is representative of all 380 hair spectra used for the study, which took into account many variations between donors. Moreover, each misclassification only accounts for one of the ten spectra used for a single hair donor-dye combination, and in most of the misclassification occurrences the other nine spectra collected for the hair donor were classified correctly. Therefore, if all of the spectra

collected from a donor are taken into account, correct donor-wise classification was attained with at least 90.0% confidence in all three instances.

With the success of this study, in addition to the study conducted by Manheim et al., [14] a great deal of information about a hair found at a crime scene can be obtained using a simple and semi-nondestructive technique such as ATR FT-IR spectroscopy. Subsequent to determining that a hair or fiber is natural, and of human origin (based on Manheim et al.'s work), the hair can then be classified as dyed or non-dyed as shown here. Furthermore, if the hair is determined to be dyed, the classification can be more specific with identifying the brand, and color, of the dye. All of this information provided by applied spectroscopy could be helpful to criminal investigators, especially to complement the conclusions drawn by microscopic comparison techniques currently used by forensic hair analysts. It is particularly noteworthy that, unlike microscopic comparison techniques, a control sample is not needed for comparison.

Many of the studies previously performed were also successful in their analyses of hair using applied spectroscopy, but with longer, more complicated, procedures or sample preparation [10-13]. This study confirms that ATR FT-IR spectroscopy is a useful technique for analyzing human hairs, without additional sample preparations subsequent to the hair dying process per manufacturers' instructions. ATR FT-IR spectroscopy is an extremely favorable technique for forensic hair analysis because the method is quick and essentially non-destructive to the sample. Most importantly, the study demonstrates how ATR FT-IR spectroscopy can be coupled with chemometrics to provide statistical confidence to conclusions drawn from hair analysis, as was called for by the 2009 NAS report [15]. This study also provides helpful insights into how the utilization of spectroscopy to analyze hair samples can be successful despite the genetic origin (i.e. age, race, and biological sex) of the donor.

To enhance the practicality and further validate this method, additional studies should be performed, which feature other variabilities in hair donors and dye products. To ensure that the differentiation of the dye brands and colors in this study was not circumstantial, ATR FT-IR spectra of dyed hairs should be collected from more than one batch of a particular hair dye and classified with models similar to those produced here. Furthermore, the integrity of using methods suggested by this study, and several other studies, should be compared to current forensic hair analyses in order to associate reliability and accuracy, and to determine whether further studies would be useful.

ATR FT-IR spectroscopy is a worthwhile and reliable technique, which is currently underutilized in forensic trace evidence analysis. The applications discussed here, for hair analysis, demonstrate the sensitivity and selectivity of the technique. Additionally, by collecting spectra once, from the same hair source, a variety of useful and potentially identifiable characteristics can be obtained through multivariate statistical differentiation. In turn, more information about a hair sample can be provided, bringing the field one step closer towards individualization and helping to confirm findings provided by microscopic analysis.

References

[1] R. Dhurat, P. Saraogi, Hair evaluation methods: Merits and demerits, Int. J. of Trichology. 1 (2009) 108-119

[2] R. Saferstein, Forensic Science Handbook, Prentice Hall Regents, Englewood Cliffs, New Jersey, 1982.

[3] D.W. Deedrick, Part 1: Hair evidence, Forensic Sci. Commun. 2 (2000).

[4] G. Loussouarn, C.E. Rawadi, G. Genain. Diversity of hair growth profiles, Int. J. of Dermatol. 44 (2005) 6-9.

[5] Scientific Working Group for Material Analysis (SWGMAT), Forensic human hair examination guidelines, Forensic Sci. Commun. 7 (2005).

[6] Federal Bureau of Investigation, Quick quiz: Forensic hair examination, https://archives.fbi.gov/archives/news/stories/2005/march/hair031605 , 2005 (accessed 04.21.16).

[7] C. Hughes, Challenges in DNA testing and forensic analysis of hair samples, Forensic Mag. 2013.

[8] T. Melton, Mitochondrial DNA examinations of cold case crime scene hairs, Forensic Mag. (2009).

[9] R. Saferstein, What is DNA? (Part 2 of 2), https://expertpages.com/news/dna2.htm , 2017 (accessed 03.31.17).

[10] H. Panayiotou, Vibrational spectroscopy of keratin fibers: A forensic approach, Queensland University of Technology (2004).

[11] S. Brandes, Near-infrared spectroscopic studies of human scalp hair in a forensic context, Queensland University of Technology (2009).

[12] P.M.J. Barton, A forensic investigation of single human hair fibers using FTIR-ATR spectroscopy and chemometrics, Queensland University of Technology (2011).

[13] D. Kurouski, R.P. Van Duyne, In situ detection and identification of hairs using surface-enhanced Raman spectroscopy, Anal. Chem. 87 (2015) 2901-2906.

[14] J. Manheim, K.C. Doty, G. McLaughlin, I.K. Lednev, Forensic hair differentiation using attenuated total reflection fourier transform infrared (ATR FT-IR) spectroscopy, Appl. Spectrosc. 70 (2016) 1109-1117.

[15] National Research Council, Strengthening forensic science in the United States: A path forward (2009).

[16] President's Council of Advisors on Science and Technology (PCAST), Forensic science in criminal courts: Ensuring scientific validity of feature-comparison methods (2016).

[17] R.A. Wickenheiser, D.G. Hepworth, Further evaluation of probabilities in human scalp hair comparisons, J. of Forensic Sci. 35 (1990) 1323-1329.

[18] B. Betty, M. Houck, R. Bisbing, T. Watkins, The science of forensic hair comparison and the admissibility of hair comparison evidence: Frye and daubery considered, The McCrone Group (2004).

[19] R. Gautam, S. Vanga, F. Ariese, S. Umapathy, Review of multidimensional data processing approaches for Raman and infrared spectroscopy, EPJ Techniques and Instrumentation 2 (2015) .

[20] D. Ballabio, V. Cosonni, Classification tools in chemistry. Part 1: Linear models, PLS-DA, Anal. Methods 5 (2013) 3790-3798.

[21] M.L. Barker, W. Rayens, Partial least squares for discrimination, J. Chemometrics 17 (2003)166-173.

[22] R.G. Brereton, G.R. Lloyd, Partial least square discriminant analysis: Taking the magic away, J. of Chemometrics 28 (2014).

[23] D.A Skoog, S.R Crouch, F. J. Holler, Principles of Instrumental Analysis, sixth ed.,Thomson Brooks Cole, Belmont, 2006.

[24] Chemistry LibreTexts, Infrared spectroscopy absorption table, https://chem.libretexts.org/Reference/Reference_Tables/Spectroscopic_Parameters/Infrared_Spectroscopy_Absorption_Table , 2014 (accessed 11.04.16).

[25] G.J Weston, The infra-red spectrum of peracetic acid-treated wool, Biochem. Biophys. Acta. 17 (1955) 462-464.

[26] A.M. Helmenstine, Hair color chemistry https://www.thoughtco.com/salon-hair-color-chemistry-602183 , 2016 (accessed 11.04.16).

Supplemental Data

Supplemental Table I. List of ingredients for each of the six hair dyes used in the experiment.

RB	RMB	JFMB	JFMMB	CB	CMB
Color Blend:	Color Blend:	Color Blend:	Color Blend:	Color Blend:	Color Blend:
Water	Water	Water	Water	Water	Water
Oleic Acid	Oleic Acid	p-Phenylenediamine	Ethanolamine	Ethoxydiglycol	Ethoxydiglycol
Ethanolamine	Ethanolamine	Ethanolamine	Canola Oil Fatty Acids	Propylene Glycol	Propylene Glycol
Cetearyl Alcohol	Cetearyl Alcohol	Canola Oil Fatty Acids	Cetearyl Alcohol	Isopropyl Alcohol	Isopropyl Alcohol
Propylene Glycol	Propylene Glycol	Resorcinol	Resorcinol	Soytrimonium Chloride	Soytrimonium Chloride
Cocos Nucifera Oil	Cocos Nucifera Oil	m-Aminophenol	Steareth-21	Oleth-5	Ammonium Hydroxide
p-Phenylenediamine	Argania Spinosa Kernel Oil	Cetearyl Alcohol	p-Phenylenediamine	Trisodium Ethylenediamine Disuccinate	Oleth-5
Argania Spinosa Kernel Oil	Octyldodecanol	Steareth-21	Oleyl Alcohol	Ammonium Hyrdroxide	Trisodium Ethylenediamine
Octyldodecanol	Cera Alba	Oleyl Alcohol	Fragrance	Tall Oil Acid	Disuccinate
Cera Alba	Butyrospermum Parki Butter	2-Amino-4-Hydroxyethylamino ansisole Sulfate	Petrolatum	p-Phenylenediamine	Tall Oil Acid
Butyrospermum Parkii Butter	Mangifier Indica Seed Butter	Fragrance	Sodium Sulfite	Oleth-2	Oleth-2
Mangifera Indica Seed Butter	Helianthus Annuus Seed Oil	Petrolatum	p-Aminophenol	m-Aminophenol	Fragrance
Helianthus Annuus Seed Oil	Oleth-20	Sodium Sulfite	Erythorbic Acid	N,N-Bis(2-Hydroxyethyl)-p-Phenylenediame Sulfate	C11-15 Pareth-9

Oleth-20	Sodium Laureth Sulfate	Erythorbic Acid	m-Aminophenol	Fragrance	Cocamindopropyl Betaine
Sodium Laureth Sulfate	Dipropylene Glycol	Cetearyl Glucoside	N,N-Bis(2-Hydroxethyl)-p-Phenylenediamine Sulfate	C11-15 Pareth-9	C12-15 Pareth-3
Dipropylene Glycol	Polysorbate 60	Dicetyl Phosphate	Cetearyl Glucoside	Cocamidopropyl Betaine	p-Phenylenediamine
Polysorbate 60	Sodium Sulfite	Trisodium EDTA	Dicetyl Phosphate	Resorcinol	Erythorbic Acid
Sodium Sulfite	Steareth-21	Ceteth-10 Phosphate	Trisodium EDTA	C12-15 Pareth-3	Citric Acid
Steareth-21	Polyquaternium-10	Cetyl Alcohol	Ceteth-10 Phosphate	Erythorbic Acid	Resorcinol
Polyquaternium-10	Acrylates/C10-30 Alkyl Acrylate Crosspolymer		Cetyl Alcohol	Sodium Sulfite	Phenyl Methyl Pyrazolone
Acrylates/C10-30 Alkyl Acrylate Crosspolymer	C10-40 Isoalkylamidopropylethyldimonium Ethosfulate	Activator:	2,4-Diaminophenoxyethanol Sulfate	EDTA	p-Aminophenol
C10-40 Isoalkylamidoprpylethyldimonium Ethosulfate	Sodium Benzotriazolyl Butylphenol Sulfonate	Water			Sodium Sulfite
Sodium Benzotriazolyl Butylphenol Sulfonate	Buteth-3	Hydrogen Peroxide	Activator:	Activator:	1-Naphthol
Buteth-3	Tributyl Citrate	Cetearyl Alcohol	Water	Water	N,N-Bis(2-Hydroxyethyl)-p-Phenylenediame Sulfate
Tributyl Citrate	Erythorbic Acid	Laureth-23	Hydrogen Peroxide	Hydrogen Peroxide	m-Aminophenol
Erythorbic Acid	Tetrasodium EDTA	Ceteareth-20	Cetearyl Alcohol	Acrylates Copolymer	EDTA
Tetrasodium EDTA	p-Phenylenediamine	Acrylates/Steareth-20 Methylacrylate	Laureth-23	Streareth-21	

		Corsspolymer			
Resorcinol	Resorcinol	Steareth-2	Ceteareth-20	Oleth-2	Activator:
2,4-Diaminophenoxyethanol HCl	m-Aminophenol	Etidronic Acid	Acrylates/Steareth-20 Methylacrylate Corsspolymer	PEG-50 Hydrogenated Palmamide	Water
m-Aminophenol	2,4-Diaminophenoxyethanol HCl	Disodium EDTA	Steareth-2	Oleth-5	Hydrogen Peroxide
Parfum	Parfum	Sodium Stannate	Etidronic Acid	Acrylates/Steareth-20 Methacrylate Copolymer	Acrylates Copolymer
Linalool	Linalool		Disodium EDTA	Oleyl Alcohol	Streareth-21
Citronellal	Citronellal		Sodium Stannate	Etidronic Acid	Oleth-2
Geraniol	Geraniol			Disodium EDTA	PEG-50 Hydrogenated Palmamide
Courmarin	Courmarin			Simethicone	Oleth-5
Amytl Cinnamal	Amytl Cinnamal			Sorbitan Stearate	Acrylates/Steareth-20 Methacrylate Copolymer
Mica	Mica			Peg-40 Stearate	Oleyl Alcohol
Titanium Dioxide	Titanium Dioxide			Cellulose Gum	Etidronic Acid
					Disodium EDTA
Activator:	Activator:			Conditioner:	Simethicone
Water	Water			Water	Sorbitan Stearate
Hydrogen Peroxide	Hydrogen Peroxide			Bis-Hydroxy/Methoxy Amodimethicone	Peg-40 Stearate
Cetearyl Alcohol	Cetearyl Alcohol			Stearyl Alcohol	Cellulose Gum
Ceteareth-20	Ceteareth-20			Cetyl Alcohol	
Amodimethicone	Amodimethicone			Stearamidopropyl Dimethylamine	Conditioner:
Sodium Lauryl	Sodium Lauryl			Glutamic Acid	Water

Sulfate	Sulfate				
Steareth-10 Allyl Ether/Acrylates Copolymer	Steareth-10 Allyl Ether/Acrylates Copolymer			Fragrance	Bis-Hydroxy/Methoxy Amodimethicone
C11-15 Pareth-7	C11-15 Pareth-7			Benzyl Alcohol	Stearyl Alcohol
Sodium Stannate	Sodium Stannate			Panthenyl Ethyl Ether	Cetyl Alcohol
Laureth-9	Laureth-9			Citric Acid	Stearamidopropyl Dimethylamine
Phosphoric Acid	Phosphoric Acid			EDTA	Glutamic Acid
Glycerin	Glycerin			Sodium Chloride	Fragrance
Trideth-12	Trideth-12			Cocos Nucifera Oil	Benzyl Alcohol
Disodium Phosphate	Disodium Phosphate			Carthamus Tinctorius Seed oil	Panthenyl Ethyl Ether
EDTA	EDTA			Panthenol	Citric Acid
Methylparaben	Methylparaben			Trimethylsiloxysilicate	EDTA
				Aloe Barbadensis Leaf Juice	Sodium Chloride
Conditioner	Conditioner			Methylchlorosothiazolinone	Cocos Nucifera Oil
Water	Water			Methylisothiazolinone	Carthamus Tinctorius Seed Oil
Cetearyl Alcohol	Cetearyl Alcohol				Panthenol
Glycerin	Glycerin				Trimethylsiloxysilicate
Behentrimonium Chloride	Behentrimonium Chloride				Aloe Barbadensis Leaf Juice
Parfum	Parfum				Methylchlorosothiazolinone
Amodimethicone	Amodimethicone				Methylisothiazolinone
Argania Spinosa	Argania Spinosa				

Kernel Oil	Kernel Oil				
Cocos Nucifer Oil	Cocos Nucifer Oil				
Helianthus Annuus Seed Oil	Helianthus Annuus Seed Oil				
Mangifera Indica Seed Butter	Mangifera Indica Seed Butter				
Butyrospermum Parki Butter	Butyrospermum Parki Butter				
Hydroxylzed Silk	Hydroxylzed Silk				
Silk Amino Acids	Silk Amino Acids				
Panthenol	Panthenol				
Cera Alba	Cera Alba				
Octyldodecanol	Octyldodecanol				
Isopropyl Alcohol	Isopropyl Alcohol				
Trideceth-12	Trideceth-12				
Cetrimonium Chloride	Cetrimonium Chloride				
Sodium Benzotriazolyl Butylphenol Sulfonate	Sodium Benzotriazolyl Butylphenol Sulfonate				
Buteth-3	Buteth-3				
Citric Acid	Citric Acid				
Tributyl Citrate	Tributyl Citrate				
Benzyl Alcohol	Benzyl Alcohol				
Phenonxyethanol	Phenonxyethanol				
Yellow 5	Yellow 5				
Red 4	Red 4				
Red 33	Red 33				
Ext Violet 2	Ext Violet 2				