Forensic Chemistry 2 (2016) 15–21

Contents lists available at ScienceDirect

Forensic Chemistry

journal homepage: www.elsevier.com/locate/forc

Full length articles

Improving the confidence of "questioned versus known" fiber comparisons using microspectrophotometry and chemometrics

Georgina Sauzier [a,b], Eric Reichard [c], Wilhelm van Bronswijk [a], Simon W. Lewis [a,b], John V. Goodpaster [c,*]

[a] Department of Chemistry, Curtin University, GPO Box U1987, Perth, Western Australia 6845, Australia
[b] Nanochemistry Research Institute, Curtin University, GPO Box U1987, Perth, Western Australia 6845, Australia
[c] Department of Chemistry and Chemical Biology, Indiana University Purdue University Indianapolis (IUPUI), Indianapolis, IN 46202, United States

ARTICLE INFO

Article history:
Received 11 June 2016
Received in revised form 6 August 2016
Accepted 11 August 2016
Available online 24 August 2016

Keywords:
Forensic science
Textile fibers
Microspectrophotometry
Chemometrics

ABSTRACT

Microspectrophotometry followed by chemometric data analysis was conducted on pairs of visually similar blue acrylic fibers, simulating the "questioned versus known" scenarios often encountered in forensic casework. The relative similarity or dissimilarity of each pair was determined by employing principal component analysis, discriminant analysis and Fisher's exact test. Comparison of fibers from within each set resulted in a correct inclusion result in 10 out of 11 scenarios, with the one false exclusion attributed to a lack of reproducibility in the spectra. Comparison of fibers from different sets resulted in a correct exclusion result in 108 of 110 scenarios, with two sets that shared identical dye combinations being indistinguishable. Although the presented methods are not infallible, they may nonetheless provide a path forward for forensic fiber examiners that has a more scientifically rigorous basis on which to support their findings in a court of law.

1. Introduction

Textile fibers are a commonly encountered form of forensic trace evidence, and may be used to provide evidence of association due to their high tendency to shed and be transferred through physical contact [1–3]. Furthermore, certain classes of fibers can prove highly distinctive based on their morphology, composition and especially color. Over 7000 textile dyes and pigments are currently produced worldwide, with combinations of these often used to impart specific colors to textile products [2,4]. Additionally, textile dyeing processes are generally carried out in batches that may exhibit minor variations in dye form, shade or strength [5]. Although many colors can be distinguished visually, these assessments are subjective and may be affected by metamerism or the examiner's color vision [6,7]. More objective measurements can be obtained using instrumental methods such as microspectrophotometry (MSP), which is favored as a rapid and non-destructive method for characterizing the color of dyed fibers [8,9]. Several studies have demonstrated the capability of MSP to distinguish visually similar colored fibers based upon different chromophores in the molecular structure of their dyes [10–14].

Forensic fiber examinations frequently involve comparisons between a questioned (Q) sample recovered from a crime scene and a known (K) sample taken from a known source such as a suspect's home or belongings [15]. Such comparisons may result in one of three outcomes: inclusion (e.g., "Q and K may have originated from the same source"), exclusion (e.g., "Q and K did not originate from the same source"), or an inconclusive result. In the case of textile fibers, which are mass produced, a "source" can only be described in terms of a class of objects that share the same physical and chemical characteristics rather than any particular item.

The underlying logic of these "questioned versus known" (Q vs. K) comparisons can be expressed through the following if/then statement:

> "IF Q and K originated from the same source, THEN Q and K will share indistinguishable class characteristics."

It follows, then, that the contrapositive must also be true:

> "IF Q and K do not share indistinguishable class characteristics, THEN Q and K did not originate from the same source".

It is critical to realize, however, that it is a fallacy to affirm the consequent:

* Corresponding author.
E-mail addresses: Georgina.Sauzier@postgrad.curtin.edu.au (G. Sauzier), reicharde@yahoo.com (E. Reichard), W.vanBronswijk@curtin.edu.au (W. van Bronswijk), S.Lewis@curtin.edu.au (S.W. Lewis), jvgoodpa@iupui.edu (J.V. Goodpaster).

http://dx.doi.org/10.1016/j.forc.2016.08.001
2468-1709/

"IF Q and K share indistinguishable class characteristics THEN Q and K originated from the same source".

This means, therefore, that any forensic comparison that results in an association based upon indistinguishable characteristics may be a false inclusion. There is also an implicit assumption for any given Q vs. K comparison that the source(s) of Q and K are homogenous, i.e., that both the questioned and known are representative samples of their original source(s). If this assumption is invalid, it may lead to false exclusions (see contingency table below).

Ground truth	Result of comparison	
	Indistinguishable characteristics	Distinguishable characteristics
Same source (Q = K)	True inclusion	False exclusion
Different sources (Q ≠ K)	False inclusion	True exclusion

In general, the relative strength of an inclusion depends upon the rarity of the class to which Q and K are assigned, as the chance of a coincidental inclusion goes down as the source of Q and K becomes smaller and less common. A more quantitative approach to describing these outcomes can be expressed by stating two competing hypotheses:

Prosecutor's Hypothesis (H_p). The questioned fiber(s) originate from the individual/object which is the source of the known. This hypothesis represents a "true inclusion".

Defense Hypothesis (H_d). The questioned fiber(s) originate from another individual/object than the one suspected. This hypothesis represents a "false inclusion".

In turn, a likelihood ratio (LR), as derived from the Bayes Theorem can be defined as:

$$LR = P(E|H_p)/P(E|H_d)$$

where evidence (E) can be a quantitative score of similarity or dissimilarity between the questioned and known, $P(E|H_p)$ is the probability of observing the evidence given the prosecutor's hypothesis and $P(E|H_d)$ is the probability of observing the evidence given the defense hypothesis.

Ultimately, a fiber comparison that utilizes an analytical method such as MSP includes a determination of whether the spectra from a Q and K are truly "indistinguishable". Such a determination depends upon the variation between spectra of the questioned and known samples. Specifically, guidelines published by the Scientific Working Group for Materials Analysis (SWGMAT) dictate that a 'spectral inclusion' can be made if the questioned spectrum lies within the range of the known spectra in terms of the curve shape and absorbance values [16].

Traditionally, assessment of whether two or more fibers exhibit similar spectral characteristics has relied upon an examiner's visual interpretation of the data. The subjective nature of these comparisons has led to trepidations regarding potential human error or bias [17]. Substantial research in recent decades has therefore examined the utility of analytical techniques with chemometric analysis to provide more objective fiber examinations [18–22]. Liu for example employed Raman spectroscopy with chemometrics to distinguish cotton cellulose fibers based upon their color, crystalline fraction and strength [23]. Morgan et al. also described several inter-laboratory studies employing chemometrics with UV–vis and fluorescence MSP to a large database of dyed fibers, discriminating fibers according to both their dye composition and loadings with high levels of accuracy [24]. However, these studies have largely focused on the simultaneous discrimination of several fibers, rather than the Q vs. K comparisons more typical of forensic casework. Furthermore, there is presently a lack of quantitative measures for assessing sample similarity. The establishment of cut-off criteria for an 'inclusion' or 'exclusion' result would provide an additional statistical basis on which forensic practitioners could support their findings in a court of law.

This study investigated the potential use of MSP spectroscopy followed by chemometrics to assess the similarity or dissimilarity of several blue-dyed acrylic fiber sets. Chemometric data analysis was conducted on spectra acquired from various fiber pairs in order to simulate casework Q vs. K comparisons. Quantitative determination of the similarity was then made by comparing the resultant data using hypothesis testing.

2. Materials and methods

2.1. Samples

Fiber samples were provided by the University of South Carolina. The sample population consisted of eleven sets of bilobal blue acrylic fibers colored with varying combinations of cationic (basic) dyes, as shown in Table 1. Representative images of each fiber set are provided in the Electronic supplementary information (Fig. S1). Fibers from each set had varying diameters as indicated.

2.2. Microspectrophotometry

Individual fibers from each set were randomly removed and mounted on glass microscope slides using Permount mounting media (Fisher Scientific, U.S.A.) for analysis. MSP spectra were acquired between 400 and 800 nm using a CRAIC QDI 2000 microspectrophotometer in transmission mode, operated at 150× magnification. The spectrometer was calibrated using NIST traceable standards prior to use. An autoset optimization, dark scan and reference scan were also obtained prior to each sample analysis. Ten fibers were analyzed from each set, with five spectra recorded along the length of each to account for intra-fiber variation. Fifty averaged scans at a resolution of 5 nm were obtained for each spectrum.

2.3. Data analysis

Data pre-processing and chemometric analysis was conducted using XLSTAT (AddInSoft, Paris, France) and Unscrambler X 10.3 (Camo Software AS, Oslo, Norway). All spectra were first baseline offset to 0% absorbance and normalized to account for variations associated with the fiber diameter. In this case normalization to "unit vector length" was chosen as it was appropriate for UV–vis spectra [25]. Other normalizations were explored (i.e., normalization to unit area) but the performance of the model was not improved. Principal component analysis (PCA) was then conducted on the entire dataset of known sample spectra using Unscrambler X 10.3.

The Q vs. K approach was undertaken by conducting PCA on pairs of fiber sets using XLSTAT. In each comparison, the "known" sample was defined as a group of 45 spectra originating from the first nine fibers of the set, and the "questioned" sample was defined as the five spectra acquired from the last fiber analyzed in the same set, or the last fiber analyzed in a different set. Discriminant analysis (DA) was performed in XLSTAT on each pair based on their PCA scores against the first three PCs (accounting for >98% of total variance in each comparison), calculating prior membership probabilities from each training set. The number of PCs used to construct each model was selected according to the corresponding scree

Table 1
Dye compositions of eleven blue acrylic fiber sets utilized in this study.

Fiber set	Dye composition	Diameter (μm)
Fiber A	Blue 3, Red 18, Yellow 28	17.5
Fiber B	Blue 41, Red 46, Yellow 28, Yellow 29	15
Fiber C	Blue 41, Red 46, Yellow 28	15
Fiber D	Blue 41, Red 29, Yellow 21	21.25
Fiber E	Blue 147, Red 29, Yellow 28	23.75
Fiber F	Blue 3, Blue 147	23.75
Fiber G	Blue 147, Red 46, Yellow 28	18.75
Fiber H	Blue 3, Red 18, Yellow 28	22.5
Fiber I	Blue 41, Red 29, Yellow 28	22.5
Fiber J	Blue 41, Red 18, Yellow 28	25
Fiber K	Blue 3, Red 46, Yellow 28	25