NPA White Paper: DNA Barcoding for Botanical Authentication

Issue

New York State Attorney General Eric Schneiderman ordered four major retailers to stop selling store-brand herbal supplements. The four retailers received cease-and-desist letters demanding that they halt the sales of these supplements because a third-party testing lab used DNA barcoding to authenticate botanical ingredients listed on the label. The attorney general’s use of DNA barcoding as a type of “genetic fingerprint” purportedly demonstrated evidence, which has still not been released publicly or available for peer review to suggest that many of the ingredients claimed on the label were not in the product and that some contaminants may have been present. A total of 19 out of the 24 products tested supposedly contained DNA that was either unrecognizable or from a plant other than what was claimed on the label. The resulting media coverage about the attorney general’s decision to use DNA barcoding for authentication in botanical dietary supplements containing extracts and highly processed ingredients rather than raw botanical materials is concerning from a scientific and regulatory pursuit. Attorney General Schneiderman made an interesting choice in the academic lab to perform highly technical recoveries of DNA from highly processed products.

Members of the Natural Products Association (NPA)\(^1\) representing the natural products industry are concerned regarding the study methodology used by Attorney General Schneiderman’s office. DNA fingerprinting is a powerful tool for natural product authentication, including raw materials before extraction, however the information provided by the attorney general fails to mention the listing of botanical extracts as dietary ingredients in these herbal supplement products. Botanical extracts involve the use of alcohol or other solvent to extract the final dietary ingredient from the source botanical. The use of DNA barcoding methodology on extracts of raw ingredients is neither a good, better, or best standard of practice in the dietary supplement industry. Moreover, if the study actually was scientifically valid and fit for purpose, the science would appear to be inconsistent with the dietary supplement current Good Manufacturing Practices (GMPs).

For the past six weeks since the announcement, NPA has repeatedly asked the attorney general to release the results of his study out of government transparency, scientific validation and

---

\(^{1}\) NPA was founded in 1936 to promote and protect the unique values and shared interests of retailers and suppliers of natural nutritional foods and natural products. NPA is a non-profit 501(c)(6) association, whose mission is to advocate for the rights of consumers to have access to quality products that will maintain and improve their health, the rights of retailers and suppliers to sell these products, and to promote natural products for healthy lifestyles. NPA unites a diverse membership from the smallest health food store to the largest natural products supplier. We are the oldest and largest trade association in the natural products industry, representing over 2,000 members accounting for almost 10,000 retail, manufacturing, wholesale, and distribution locations of natural products, including foods, dietary supplements, and health/beauty aids. We champion consumers’ freedom of choice in our marketplace, strengthen and safeguard retailers and suppliers, and build strong markets to fuel industry growth. We act together with uncompromising integrity, and we encourage all to reach ever higher standards of quality. In short, we are the trade association advocating for the dietary supplement industry.
research integrity. NPA has also asked the Food and Drug Administration (FDA) to issue a judgment as to whether this DNA barcoding technology is appropriate, sensitive, specific and scientifically valid for routine use by the dietary supplement industry to definitely identify botanicals used as the starting source material to create botanical extracts. The concern is that DNA will be sufficiently degraded during the extraction and manufacturing process that intact DNA markers specific to a particular botanical will not be detected. Neither the attorney general nor the FDA has responded to our requests.

**Executive Summary**

DNA-based authentication of herbals and botanicals in finished dietary supplement products continues to be a work in progress, offering a powerful, novel tool for quality assurance quality control but if, and only if, key challenges/obstacles are overcome. DNA barcoding is a technology in its infancy with a potential for broad use in the future toward botanical authentication. The technique needs refinement for use in plants. It must undergo further testing and validation. There is no universal DNA barcode consisting of a single loci or loci combination with a high enough specificity rate to reliably conclude the presence or absence of a botanical in a product for all varietals at the species level. The method requires a large DNA library where sequences are inexorably linked to voucher specimens, and the natural products industry does not have those assurances at present.

Experts have not agreed on the universal DNA barcode(s) to be used in all botanical authentication studies. Once the ideal barcode has been universally accepted, verified and validated, one can use DNA barcoding on raw materials before extraction and on some appropriate finished products, this universal barcode for plants companies can use DNA barcoding to ensure they are producing quality products, in addition to providing consumer confidence.

NPA is not sure what can be made of the DNA results conducted by Dr. Schulte at Clarkson University until the study data is released. Contamination of DNA can readily occur in a facility that is not ISO-accredited and certified for performing such work. NPA has concerns over the chain of custody throughout the process during handling of both the test and reference materials. Furthermore, DNA barcoding involves amplification of DNA and therefore does not provide for any quantitative assessment. The presence of any contaminants found in a product from DNA barcoding does not convey a quantitative level of contaminants in the product. Contamination with foreign DNA can occur as a result of mishandling the test article in the laboratory. The dietary supplement cGMPs in 21 CFR part 111 also allows for setting limits on the levels of contaminants. DNA barcoding does not provide quantitative information to confirm whether a listed ingredient found in the product meets label claim.

NPA does not support the use of DNA based-approaches alone to authenticate botanicals in extracts or finished dietary supplement products. There are numerous DNA-based methods that could be appropriate for different products. There are many more products which simply
don't have DNA and therefore never appropriate for DNA-based tests whatsoever. The identification of botanical raw materials and finished products is multifaceted and many factors should be taken into account when ensuring conclusive positive authentication. DNA testing seldom is able to properly identify chemically complex herbal extracts as little or no DNA is extracted in many commercial extraction processes. There are numerous steps which take place during the botanical extraction process and manufacture of finished dietary supplement products which will degrade DNA. DNA-based approaches for identification such as barcoding should be partnered with chemical methods (chromatographic and spectroscopic) in an authentication toolkit to definitively conclude whether a dietary supplement product has met label claim for a botanical ingredient.2

Background

Anyone with a basic working knowledge of the natural products industry has spent time pondering the quality and authenticity of raw botanical starting materials. The legitimate dietary supplement industry must continuously combat news outlets fanning the flames of fly-by-night companies, who have no interest in maintaining reputation for quality, operating within the regulatory structure set forth by the FDA, or compliance within the FDA’s GMPs or long-term goals to remain as a fixture in this industry. Stories of adulteration for failed GMPs, contamination, economically motivated adulteration, poor quality, and misbranding plants doubt in minds of consumers regarding the botanical industry and seeds mistrust and apprehension to purchase quality products.

Americans have a right to have access to a diverse array of botanical ingredients to supplement their diet, and ones in which they can have confidence regarding quality and limits on contaminants. Since the majority of these botanical ingredients are not grown in the United States, manufacturers, distributors, re-packers, and retailers of natural products containing raw botanical materials relinquish control in the life cycle of the botanical material somewhere along the supply chain. However, title 21 of the U.S. Code of Federal Regulations (CFR) part 111 calls for 100 percent identity testing for all ingredients. While intentional and unintentional substitution of raw botanical ingredients and economically motivated adulteration with cheaper alternatives can occur, the FDA’s dietary supplement GMPs require that quality be tested in both the incoming raw material and the finished product released for consumption. The dietary supplement GMPs are in place to ensure the identity, strength, purity, and composition of finished products and that products have limits on the level of contaminants.

Canadian researchers from the University of Guelph started the debate over whether DNA technology should be used by the food and supplement industry as an acceptable method for identity testing in dietary supplement GMPs. Their 2013 study,3 suggested that out of the 44

---

herbal products they tested, many were determined to be adulterated or mislabeled. To arrive at their startling conclusions that little, if any, of the herb listed on the label is present in the final, highly processed, finished product, the researchers used a technique called DNA barcoding to identify a small segment of DNA, akin to scanning a product code or stock keeping unit (SKU) at a store, to authenticate botanical ingredients. Additionally, they suggested their test products were contaminated with other substances but did not provide a level for the contaminants. Researchers at the university discovered “considerable product substitution, contamination, and use of fillers,” including rice, wheat and soybeans. A month-long investigation and media coverage emerged as a result of this DNA study: results which are strikingly similar to the New York Attorney General’s action.

In the case of the Canadian study, however, both the National Institutes of Health’s (NIH) Office of Dietary Supplements chief and the FDA’s former director of the Division of Dietary Supplement Programs released similar statements indicating botanical extracts do not contain intact DNA, and the case was essentially closed. The flawed application of DNA barcoding to highly processed finished dietary supplement products does not mean the botanical industry can brush aside and ignore this emerging technology, but there should be a broader discussion as to when DNA authentication is fit for purpose for dietary supplements, which are highly processed, and when it is not. Many other experts in the field immediately questioned the Canadian study’s findings.

What is DNA Barcoding?

DNA barcoding is a relatively novel but as yet unproven and unaccepted method to extract an organism’s DNA and classify it to a particular taxa at the species level. DNA barcoding involves multiple steps including DNA extraction, removal of secondary metabolites which interfere with sequencing and the polymerase chain reaction (PCR), selection of the DNA barcode loci, gene amplification, DNA sequencing, and comparison with a DNA sequence library tied to voucher specimens (see figure 1). This novel taxonomic method of authentication was proposed in 2003 by Dr. Paul Hebert of the University of Guelph as a molecular tool for species identification. Researchers isolated the cytochrome c oxidase, subunit 1 (cox1 or CO1) gene from the mitochondrial genome of Lepidoptera, a family of insects belonging to moths and butterflies. Subsequent studies using CO1 as a genetic marker of comparison for mosquitoes, fish and birds were successful. The frequent genome rearrangements and transfer of genes between

---

4 http://m.prevention.com/mind-body/natural-remedies/safety-herbal-supplements
different genomes (plastid, nuclear, and mitochondrial) and across species of herbs and botanicals, in addition to limited base substitution rates, or wobble in the genome, makes CO1 a poor choice for barcoding in the plant kingdom.\textsuperscript{10,11}

DNA barcoding involves molecular methods to extract a short gene sequence (400 – 800 base pairs in length) from the starting material (i.e., raw botanical material). DNA barcoding can be used for species authentication, species delimitation, novel species discovery, identification of geographical distribution over time, and plant-animal interactions. The gene sequence used for authentication should be from a standardized region of the genome. In fact the inability of CO1 to work as a barcode in plants\textsuperscript{12} fueled efforts to find a better candidate.\textsuperscript{13}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{dna_barcode_diagram.png}
\caption{Steps in DNA barcoding.}
\end{figure}

DNA barcoding has not managed to avoid its detractors.\textsuperscript{14} DNA barcoding has often not provided reliable information above the species level (genus), and some have suggested it should not be applied to the species level but higher levels of taxonomic groupings.\textsuperscript{15} Experts in the natural products industry have criticized both the 2013 University of Guelph study and the recent New York attorney general findings, suggesting DNA would not be detected from

botanical extracts. Even staunch critics of quality in the natural products industry have criticized the findings from Attorney General Schneiderman. DNA barcoding naysayers include a vocal critic from Harvard Medical School and a consumer advocate from Consumer Labs. Laboratories in the natural products industry, including those that use DNA barcoding to authenticate botanicals using raw material before extraction, have criticized the New York attorney general’s choice of laboratory and method for authentication. NPA employs two former FDA chiefs from the Division of Dietary Supplement Programs, and they have questioned both the attorney general’s study’s findings and the unwillingness to release the data out of transparency, scientific validity and research integrity. NPA and its members support chemical fingerprinting using a chromatographic method as a gold standard. DNA barcoding without a chemical method for the phytochemicals may only tell you that DNA was unable to be extracted or amplified in the process, leading to false negative results.

**Intact DNA (i.e., Barcodes) Is Not Recoverable from Highly Processed Goods like Many Finished Dietary Supplements and Botanical Extracts**

The ideal DNA barcode should be readily recoverable, amplified, and sequenced from herbarium samples, extracts, and finished products; however, DNA is not typically recoverable from highly processed foods like finished dietary supplements and botanical extracts. While many DNA-based methods are appropriate for some finished products that are minimally processed, many finished products have been processed to the point that they do not have DNA. These products are inappropriate for DNA-based tests. There are multiple steps in the manufacture of botanical dietary supplements between farming and final processing, which degrade DNA (see figure 2). Post-harvesting treatments of botanicals in the final manufacturing of botanical dietary supplements cause single and double stranded breaks in the DNA. Even improper storage of raw botanical materials can cause DNA degradation. The extraction process for botanicals utilizes solvents like alcohol, which significantly degrades DNA. Unfortunately, the ideal DNA barcode for plants, accepted by the scientific community, has not been found, and therefore chemical methods must be used to verify the authenticity of botanical materials in finished supplements and extracts. DNA-based methods can be used on raw botanical materials before extraction as long as they are used with a confirmatory method like chemical fingerprinting for the phytochemicals present and expected in the raw material.
Challenges to DNA Barcoding for Botanicals

The Ideal Barcode Loci or Loci Combination for Extraction and Amplification from Plants Has Not Yet Been Found

A number of ideal characteristics should be possessed by any candidate plant barcode.\(^\text{16}\) First, it should be universally applicable across all taxa using standardized primers, simple enough to be amplified and short enough to sequence in one reaction with present-day sequencing capabilities (i.e., a limit for technology today to 750 bp). Second, the chosen barcode should be universally accepted by the scientific community to exhibit sufficient variability for authentication at the species level. Third, the barcode should always display high interspecies divergence but low intraspecies divergence for proper species delimitation. Finally, the barcode should not contain areas with DNA insertions or deletions. The barcode sequenced from a tested material should be able to be easily aligned with any other sequence in the DNA reference library, generated from voucher specimens. Unaligned segments complicate comparison of the sequenced barcode with DNA in the library, challenging interpretation of the final results.\(^\text{17}\)

\*\*\*\*

\textbf{Figure 2.} Steps in processing of raw botanical material to an extract.

The mitochondrial DNA CO1 barcode used for animals was ideal and therefore quickly universally accepted, but this has not been the case for plants. In 2004, the Consortium for the Barcode of Life (CBOL) was founded, and it quickly realized CO1, which works for the animal

\(^{16}\) Chase MW et al. (2007). A proposal for a standardized protocol to barcode all land plants. Taxon 56: 295-299.  
kingdom, was not suitable for plants. In contrast to animal groups, plants offer far lower rates of base substitution in their mitochondrial DNA, prompting researchers to find alternative barcoding regions. A variety of candidate DNA barcode loci \( (rbcL, \text{matK}, \text{trnH-psbA}, rpoB, rpoC1, ndhJ, accD, ITS, ycf5, \text{and TS2}) \) and various two-loci combinations have been tried for plant authentication. The plant working group (PWG) of CBOL settled on a two-locus combination involving \( rbcL \) and \( \text{matK} \). The combination was shown to discriminate at the species level with only a probability of 72 percent,\(^{18}\) an identification rate with such low specificity that is far from ideal as a candidate method that should be universally adopted by the natural products industry. The DNA sequences chosen for the method takes great care as it is not a one-size fits all approach. One 600 bp DNA sequence used for a particular genus to differentiate species may not be able to be used for another genus. We saw this with the failed adaptation to use CO1 from the animal to plant kingdom.

**Reliance of DNA Barcode Sequences on the Plant Taxonomy Studied and Setting Genetic Distance Limits**

DNA barcoding relies heavily on congruence between the barcodes and the species taxonomy of the group in question. Simply stated, it suggests that there is a perfect one-to-one match between barcodes and taxa, and there should be no overlap of absolute barcode sequence between species.\(^{19}\) While this is the case for animals, it is fulfilled to a much lower percentage in plant studies.\(^{20}\)

Genetic distance approaches through DNA barcoding has proven to be limited when it comes to defining species boundaries.\(^{21}\) Rates of evolution of mitochondrial DNA, which is used as a molecular clock,\(^{22}\) vary substantially in plants between and within species and between different groups of species, resulting in broad overlaps of intra- and interspecific distances.\(^{23,24,25}\) While a limit of 3 percent for genetic divergence is used to separate species that are the result of geographic isolation and mutation accumulation, the genetic system of plants may also undergo an accelerated speciation process, which happens in plants undergoing instantaneous speciation through hybridization or polyploid mechanisms. Plants have the ability to speciate instantaneously through allopolyploidy by changing the ploidy or

---

number of sets of their chromosomes.\textsuperscript{26,27} Thus, complete DNA sequence libraries for botanicals which are tied to voucher specimens are all the more critical during comparison.

The Importance of a Database Reference Library Tied to Voucher Specimens

For DNA barcoding on plant material to be reliable, you need a fairly complete, robust, and accurate database of already-identified DNA samples from appropriately vouchered specimens for comparison. A voucher is a preserved sample of plant material representing a single gathering or batch in the field as a permanent record as to the identity of that material. DNA barcoding performed on plant material requires a significant investment to obtain vouchered specimens, and a complete voucher specimen library is not simply created overnight. They are usually painstakingly collected and meticulously catalogued by large laboratories and Centers of Excellence dedicated to that venture. While there are voucher specimen libraries available, there are no complete databases of already-identified DNA samples based upon voucher samples for comparison.

Publicly available DNA sequence databases are not tied to vouchered specimens (i.e., herbarium botanical specimen) and their use by DNA barcoding researchers can lead to erroneous findings. NIH developed GenBank, which comprises the DNA DataBank of Japan, GenBank at NCBI and the European Molecular Biology Laboratory (EMBL). There is no system in place for any of those databases to verify a sequence came from a vouchered specimen. It is critical that the DNA reference library be tied to vouchered specimens as the DNA barcoding method is only as good as the fidelity of the vouchered botanical reference material in the library. One could have hundreds of variant sequences, uploaded and “identified” to the level of species, for corn, for example, to choose from in a public database of sequences. Since none of the sequences are based upon a voucher specimen, the DNA sequence chosen as the reference of comparison in the library is a guess.

NPA is eager to know what DNA database was used by Dr. Jim Schulte of Clarkson University in the New York attorney general study and how the sequences used in that database as a reference of comparison were tied to vouchered specimens. All we can conclude at this point is that he used publicly available DNA sequences that did not involve voucher specimens, leading to erroneous results. This is a challenge with scientific merit and warrants transparency on the part of the study investigators. Until the data is released by Attorney General Schneiderman, we can only speculate at this point. It is imperative that the attorney general or Clarkson University release the study report and accompanying data to either exonerate or debunk the study. All we can say at the moment is that the Clarkson University test results are preliminary and fail to provide a degree of certainty until they are validated with substantiation.

Use of DNA Barcoding for Authentication in the Absence of a Chemical Confirmatory Method for Phytochemicals

\textsuperscript{26} Templeton AR. (1980). Modes of speciation and inferences based on genetic distances. Evol 34: 719-729.

DNA based-approaches for identification such as barcoding should be partnered with other chemical (chromatographic and spectroscopic) methods in an authentication toolkit to definitively conclude whether a dietary supplement product has met label claim for a botanical ingredient.  

While some extracts could contain fragments of DNA, the fragments are of such low quality and short length to be used to generate an uninterrupted barcode consisting of 500–750 base pairs. The DNA is often degraded to a point that makes it impossible to perform proper authentication. Other post-harvest processes, including extensive heat treatment, submission of plant material to UV light, and extraction with alcohol impact the overall quality of the DNA. In the end, if the method is unable to detect DNA, all it can conclude is that it was unable to detect a DNA barcode for a particular botanical extract. It would be presumptuous in the absence of a confirmatory chemical method for phytochemicals to suggest the product failed to meet label claim. All we can conclude from the headlines generated by the New York Attorney General is that the results are preliminary and require substantiation and validation.

Concerns over the Laboratory Expertise and Facilities Selected by the New York Attorney General

The high degree of sensitivity of DNA-based tests necessitates that DNA barcoding tests be executed by qualified experts in botanical DNA authentication, performed in quality-controlled facilities, and DNA sequences compared to those developed from appropriate (i.e., voucher specimen) reference materials.

Dr. Schulte of Clarkson University is a self-proclaimed evolutionary biologist with a goal to “reconstruct the tree of life.” His expertise, according to his university website, is “lizards and snakes,” for which he uses DNA to trace their evolutionary branches. In other words, Dr. Schulte is a DNA researcher and academic with experience in terrestrial animals, not the plant kingdom. The plant kingdom has proved challenging for DNA barcoding in comparison to the animal kingdom. The choice of loci has proven difficult as a universal DNA barcode for plants has yet to be discovered. Accurate botanical species authentication and detection of adulterants is a major challenge when raw botanical materials are highly processed into dietary supplements or transformed into phytochemical extracts. There is an inverse relationship between the probability of recovering DNA from a sample and the degree to which that raw botanical material has been processed into an extract or finished product. NPA eagerly awaits release of Dr. Schulte’s data to determine which loci he chose for the various botanicals investigated and the size of the DNA fragments he recovered from the finished product extracts.

The facilities at Clarkson University were not ideal for a study of this scope and magnitude as directed by the New York Attorney General. The International Organization for Standardization

---

(ISO) accredits laboratories who perform DNA tests involving foodstuffs to ensure that DNA testing services rendered are safe, reliable and of good quality. The data generated by Clarkson University was not collected in an ISO-accredited and certified laboratory using validated methods. It is therefore quite possible that the Clarkson University laboratory was the source of contaminants reported by the New York Attorney General. There is a reason why DNA is carefully handled because of the ease with which cross-contamination can occur. DNA barcoding is a method with exquisite sensitivity but lower specificity and involves DNA amplification steps. Therefore, the method can generate false positives (i.e., presence of contaminants) due to its low rate of specificity. Any contamination by DNA introduced during the process will be amplified, leading to erroneous findings.

**What is the Gold Standard for Identifying Botanicals in Extracts and Finished Products**

The use of DNA barcoding or other DNA-based methods should not be used alone to authenticate botanicals. The public and press may believe that DNA-based methods are the gold standard for identification of botanicals in a finished product. The onslaught of crime scene television shows and movies applying the DNA-based methods in forensic pathology has created a public awareness as to the power and sensitivity of these methods to positively identify or exonerate individuals. Botanicals produce a variety of secondary metabolites (alkaloids, polysaccharides, terpenoids, tannins, glycosides, glycoproteins, natural phenols, polyphenols, phenazines, waxes, pigments and resins) at high levels for normal plant growth and development and self-defense. These secondary metabolites also include the beneficial phytochemicals sought for their flavoring and positive physiological properties on consumer health. If not properly removed, these secondary metabolites of plants may interfere with PCR, DNA extraction and cycle sequencing, making DNA unusable for further molecular analysis.

---

29 Wen XP, Deng XX. (2002). The extraction of genomic DNA from five species of Rosa Seed. 126: 18-21.
Chromatographic methods are the gold standard as stand-alone methods for the identification of botanicals in diverse matrices such as finished dietary supplement products. These tests produce a chemical fingerprint of peaks, corresponding to masses, which are characteristic of the phytochemical profile for a particular species. Chemical tests like high performance liquid chromatography with mass spectrometry provide a chemical fingerprint in a spectral array of peaks rather than a molecular fingerprint of DNA. DNA barcoding should always partner with a chemical fingerprint method for positive authentication because of the high probability of isolating no DNA and a negative result when used alone. It will be interesting to find out whether the investigators at Clarkson University used a chromatographic technique as a confirmatory method to rule out predictable explanations for finding no DNA. It is imperative that their study data be made publicly available to apply scientific scrutiny to the findings, provide validation to the results, or exonerate those companies undergoing current scrutiny. Scientific scrutiny of the test results is the only way to fairly adjudicate the allegations made and to prevent further actions by states to remove products, produced in accordance with federally opted GMP standards, from retailer shelves.