Application of Epigenetics to Identify Plant Organs in Herbal Supplements

Authors: Lorianny Rodriguez[1] and Heily Rosa[1]
Mentors: Dr. Damon Little[2]; Teacher: Allison Granberry[1]
[1]Hostos-Lincoln Academy, [2]New York Botanical Garden

Cold Spring Harbor Laboratory DNA Learning Center

Funded by the Thompson Family Foundation

Abstract

The purpose of this research was to design and test a protocol that can easily indicate the presence of roots and/or shoots in herbal supplements. Although herbal products in the United States are regulated by the FDA, testing standards are reduced in comparison to prescription or over-the-counter medication. Comparing methylation patterns within a DNA region can be used to determine which structure of the plant is being examined. A primer set flanking hypermethylated sites was designed from a well-known plant, *Arabidopsis thaliana.* 21 different PCR protocols were performed to test the primer set. Some DNA samples produced incorrect sized products, whereas other samples of the same species did not in the same PCR conditions. Protocols that did not produce amplicons can be explained by primers that either bound to each other and not to the template DNA or one or both primers hairpinned.

Introduction

Lorianny Rodriguez

Damon Little

Figure 1. *Arabidopsis thaliana* grown in lab.

As of 2016, over half of the American population takes some form of dietary supplements (Cohen et al., 2018). Herbal supplements are an expanding market projected to reach $86.7 billion by 2022 (Lamb, 2017). The regulatory powers of the Food and Drug Administration (FDA) extend to herbal supplements. Since the FDA's testing standards are reduced in comparison to prescription or over-the-counter medication they have to find more efficient and effective means of enforcement. The FDA only requires purity and identity testing for dietary supplements, but cannot mandate specific tests. The FDA relies on manufacturers' reports to determine whether a product is safe. After a product is on the market, the FDA does not test it unless there are sufficient complaints from consumers to demonstrate that it is not safe.

DNA analysis could provide an efficient, valid, and reliable method to identify the different components of an herbal supplement. DNA methylation and demethylation are epigenetic processes of adding or removing a CH_3 (methyl) group on cytosine. Throughout plants, methylation patterns vary. Comparing methylated patterns within a DNA region can be used to determine which structure of the plant is being examined. Shoots are more likely to be methylated than roots (Windman et al., 2011). Some supplements are labelled as containing leaves (shoots) or roots of a particular plant; methylation patterns can be used to test this by differentiating between methylated and unmethylated sites. Treatment with ammonium bisulfite changes methylated cytosines to uracil (detected as thymine). If this change occurs, it means that the site was methylated and therefore came from a shoot and if not methylated, a root.

Not methylated in roots, methylated in shoots	Genes
Extensin genes	AT5G35190, AT1G23720, AT2G24980, AT3G28550, AT3G54580, AT3G54590, AT4G08410, AT4G13390, AT5G06630, AT5G06640
Other genes	AT4G08380, AT4G08400

Table 1. Differentially methylated genes in *Arabidopsis thaliana* reported by Windman et al. (2011).

The purpose of this research was to design and test a protocol that can easily identify the presence of roots and shoots in herbal supplements. A primer-set flanking hypermethylated sites was designed from a well-known plant, *Arabidopsis thaliana,* (Fig. 1) in which shoot hypermethylated sites have been documented (Windman et al., 2011). The targeted gene regions are dominated by extensin genes (Table 1; Windman et al., 2011). The primer sets were electronically tested against eight fully sequenced plant genomes. Primer sets that did not electronically amplify in all eight genomes were discarded. *Arabidopsis thaliana*, *Zea mays* (corn), *Linum usitatissimum* (flax), *Salvia hispanica* (chia), *Foeniculum vulgare* (fennel seed), *Glycine max* (soybean), *Helianthus annuus* (sunflower), *Triticum aestivum* (wheatgrass), *Allium tuberosum* (garlic chives), *Beta vulgaris* (beet), *and Petroselinum crispum* (parsley) were grown and both roots and shoots were harvested. DNA extracted separately from both organs.

Materials and Methods

In order to find differentially methylated DNA regions, twelve genes that are reported to have differential methylation between roots and shoots were used (Windman et al., 2011). Sequences were downloaded from NCBI genbank (accession GSE52762). University of California at Santa Cruz's genome browser (http://genome.ucsc.edu) was used to view the gene sequence of *Arabidopsis thaliana* (TAIR 7) and the methylated sites. A hundred base pair region before and after each cluster of hypermethylated (methylated in 99% of shoot sequencing reads) sites was input to PRIMER 3 to design primer sets that: had a 50-200 bp product size; a primer size between 18-26; primer T_m between 57-64℃ (with monovalent cation concentration of 20 mM, divalent cation concentration of 2 mM, and dNTP concentration of 0.8 mM).

Assay Design

The primers designed by PRIMER 3 were electronically tested with rePCR against the genomes of *Arabidopsis thaliana*, *Brachypodium distachyon, Glycine max, Lotus japonicus, Medicago truncatula, Populus trichocarpa, Schrenkiella parvula,* and *Sorghum bicolor* (the only angiosperms with chromosomal-level genome assemblies). Primers that amplified more than once within each genome were eliminated from consideration. The remaining primers were compared across the species to identify universal primer sets. One primer set amplifies only once per genome and is present in all species (Table 2).

Arabidopsis thaliana gene	Forward primer (5' to 3')	Reverse primer (5' to 3')	Amplicon size
AT5G06640	CGCCTTGTAGACATA TGGGGGTGGTG	GTGTATAACTCTCC ACCTCCACCAT	80-243 bp

Table 2. Primer set for universal hypermethylated region.

Assay Validation

Seeds were grown in a basic potting soil until they produced a second set of true leaves. The shoots were immediately harvested, preserved in 91% isopropanol, and DNA was extracted. Plant samples included *Arabidopsis thaliana*, *Zea mays, Linum usitatissimum, Salvia hispanica, Foeniculum vulgare*, *Glycine max, Helianthus annuus, Triticum aestivum, Allium tuberosum, Beta vulgaris, and Petroselinum crispum*. To extract the DNA, a silica resin protocol was followed. PCR reactions were run: each 25μL reaction contained 2μL of sample, 1x PCR buffer, 0.2 mM dNTPs, 25μg/μL BSA, 0.5μM of each primer, and 1u *Taq* polymerase.

Results

Tests with *Arabidopsis thaliana* DNA using Protocols T and U (Annealing temperatures 65.9°C and 68.0°C; Table 3) produced only correct sized products. In addition, correct sized products were also solely produced at annealing temperatures ranging from 55.4°C-59.9°C (Protocols C-E; Table 3). Some protocols produced a mix of both correct and incorrect sized products, for example Protocols C-E (55.4°C, 57.9°C, 59.9°C) and Protocol S (63.7°C). Seven Protocols (L-R; Table 3) yielded only incorrect sized products. Some DNA samples produced incorrect sized products, whereas other samples of the same species did not in the same conditions, for example, Protocols F/I and G/J. PCR with DNA from species other than *Arabidopsis thaliana* were unsuccessful (Fig. 2).

Protocol	Annealing (°C)	Annealing (s)	Extension (°C)	Extension (s)	Correctly sized products	Incorrectly sized products
A	50.0	30	72.0	30	No	Yes/No
B	52.8	30	72.0	30	No	Yes/No
C	55.4	30	72.0	30	Yes	Yes/No
D	57.9	30	72.0	30	Yes	Yes/No
E	59.9	30	72.0	30	Yes	Yes/No
F	53.0	30	72.0	30	No	Yes/No
G	53.0	120	72.0	10	?	Yes/No
H	53.0	30	72.0	10	No	Yes/No
I	53.0	30	72.0	30	Yes	Yes/No
J	53.0	120	72.0	10	?	Yes/No
K	53.0	30	72.0	10	No	No
L	50.2	30	72.0	30	No	Yes
M	51.7	30	72.0	30	No	Yes
N	53.7	30	72.0	30	No	Yes
O	55.9	30	72.0	30	No	Yes
P	58.0	30	72.0	30	No	Yes
Q	60.1	30	72.0	30	No	Yes
R	61.6	30	72.0	30	No	Yes
S	63.7	30	72.0	30	Yes	Yes
T	65.9	30	72.0	30	Yes	No
U	68.0	30	72.0	30	Yes	No

Table 3. *Arabidopsis thaliana* PCR protocols and results. Each protocol began with a 95°C denaturation for 150 s; followed by 35 cycles of 95°C for 30 s, a variable annealing step, and a variable extension step; the reaction was completed with a 72°C extension for 600 s.

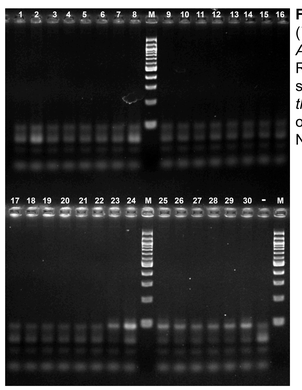

Figure 2. PCR results for Protocol U (Table 3). Reactions 23-30 are *Arabidopsis thaliana* samples. Reactions 1-22 used DNA from species other than *Arabidopsis thaliana*. Even numbers are roots and odd numbers are shoots. M indicates NEB 100 bp marker.

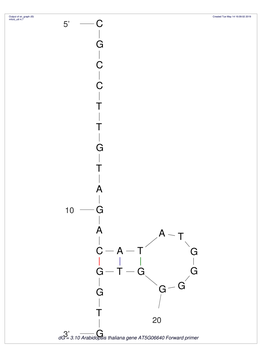

Figure 3. This structure has a 12-base hairpin at 61.6°C in 1x PCR buffer near the 3' end that likely would interfere with the ability of *Taq* polymerase to elongate the primer even if the primer managed to bind to the template.

Discussion

Protocols that did not produce amplicons can be explained by primers that either bound to each other and not to the template DNA or one or both primers hairpinned (folded upon themselves thus preventing extension; Fig. 3), therefore annealing did not occur and no extension was possible, resulting in no product. In Protocols A, B, and L-R (Table 3), there were no correctly sized products, but there were incorrectly sized products. This may have been a result of hairpinning: one primer hairpins while the other primer binds to its target site and another unintended site or the hairpinned primer is able to bind to an unintended site—the repetition within the extensin gene results in many sites that are similar to the target site. An extremely high annealing temperature is required for these primers to accurately anneal and extend and therefore produce correctly sized amplicons.

Protocols that had identical conditions except for annealing temperature can be used to test the hypothesis that hairpinning is responsible for lack of appropriately sized amplicons. Electronic folding of the primer sequences (MFold) using the same ionic conditions as the PCR reaction show that hairpinning is a controlling factor: the reverse primer folded identically at 61.6°C and 68.0°C (Protocols R and U; Table 3); the forward primer produced four structures at 61.6°C that were not found at 68.0°C. One of these structures has a 12-base hairpin near the 3' end (Fig. 3) that likely would interfere with the ability of *Taq* polymerase to elongate the primer even if the primer managed to bind to the template.

The success of PCR reactions were not entirely consistent: in some cases samples of *Arabidopsis thaliana* produced amplicon, while other samples did not. This may be due to different DNA concentrations and purity levels. Compounds co-purified with DNA may inhibit PCR. For example, Protocols G and J (Table 3) are identical, Protocol J is a replicate of Protocol G, conducted because G produced inconsistent results: some samples amplified while others did not; each sample performed identically in both replicates.

References

Cohen, P.A., Wen, A., & Gerona, R. (2018). Prohibited Stimulants in Dietary Supplements After Enforcement Action by the US Food and Drug Administration. *JAMA Intern Med*, 178(12).

Darst, R.P., Pardo, C.E., Ai, L., Brown, K.D., & Kladde, M.P. (2010). Bisulfite Sequencing of DNA. *Current Protocol of Molecular Biology, 7*(917), 1-20

Lamb, J. (2017, May 17). Herbal Supplements Market Size, Share, Report, Analysis. Trends & Forecast to 2022. https://www.reuters.com/brandfeatures/venture-capital/article?id=5660

Widman, N., Feng, S., Jacobsen, S.E., & Pellegrini, M. (2014) Epigenetic differences between shoots and roots in *Arabidopsis* reveals tissue-specific regulation, Epigenetics, 9(2), 236-242.

Acknowledgements

New York Botanical Garden
Christine Marizzi, Melissa Lee and Lydia Paradiso of Harlem DNA Lab
Cold Spring Harbor Laboratory