Increasing Bioavailability of Anthocyanin in *Daucus carota (V. Imperator & Deep Purple) Through CRISPR/Cas12-Mediated Editing*

Miya Zhao

Yale University

MCDB201L Molecular Biology Laboratory

Dr. Maria Moreno

May 7, 2024

Abstract

How does one make a carrot as nutritious as a blueberry? In order to create more bioavailable commercial carrots, scientists want to remove the acylation chemical modification characteristic on anthocyanin, a nutrient found in berry-colored fruits and vegetables. The attachment of this acyl group makes anthocyanin more stable and structured within plants; however, it also makes it difficult for humans to break down anthocyanin for its multiple health benefits like increased cardiac health and better sugar metabolism. Previous research suggests that the *DcSCPL1* gene plays a crucial role in acylating carrot anthocyanin. Thus, our goal is to induce early termination in a conserved region in *DcSCPL1* through a gRNA-mediated deletion.

The 2023 cohort used the SeqBuilder Platform to identify and place primers to amplicons of interest within exonic regions that contained conserved sequences. This year's cohort then determined the compatibility of the previously designed primers and potential guide RNAs within two strains of the *Daucus carota (v. Imperator or Deep Purple)* genome in order to design crRNAs. Experimentally, *DcSCPL1* was found within both the *Imperator* and *Deep Purple* strain, which may indicate its importance in acylation. Clones of the gRNAs within the *Imperator* and *Deep Purple* were successfully created using TOPO Blunt II ligation. Furthermore, crRNAs were created through Cas12 guide integration into the pMM260 vector through Golden Gate Cloning. Ultimately, it was found that gRNAs and crRNAs were successfully designed and integrated into their vector through sequenced data through a SeqBuilder analysis of the experimental clone colonies. In vitro, the two vectors can be used to confirm the functionality of the gRNAs. Finally, the crRNAs can be used to confirm the functionality of the *DcSCPL1* gene and induce in-vivo truncation through a single gRNA inducing early termination and thus loss-of-function of acylation within commercial carrots.

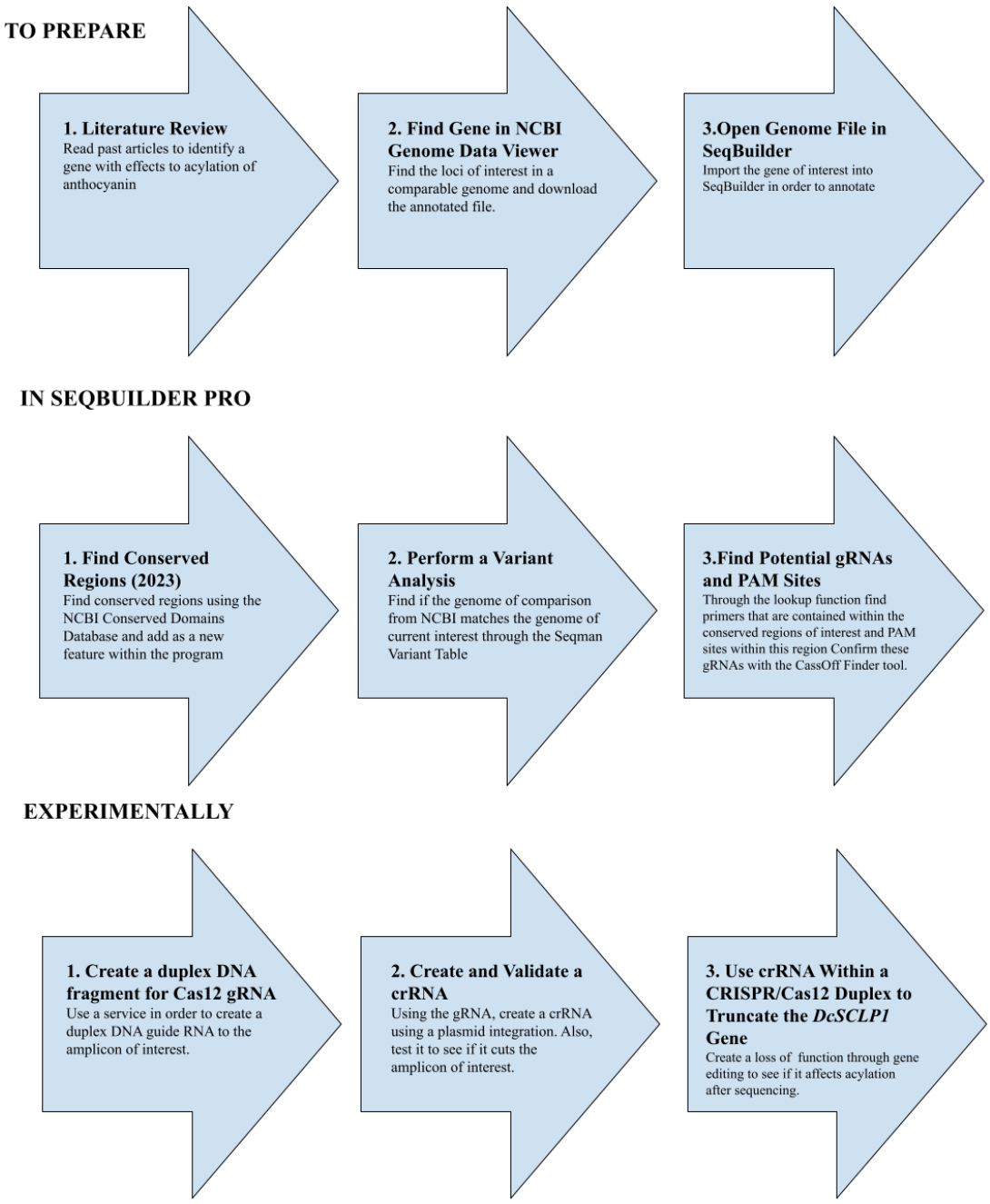


Figure 1. Graphical Abstract: Strategy to Identify and Functionally Analyze the DcSCPL1 Gene *Flowchart of the process used to annotate the gene and conserved domains from ideation to steps aiding in designing methodology to functionally analyze the gene of interest.*

Background and Significance

While commercial carrots of today are famous for their bright orange appearance, pre-domestication, many appeared purple. Anthocyanins are a phytochemical found in fruits and vegetables that also serve to protect the cells from light and create red, blue, and purple pigments. Thus, they are popularly used to replace artificial food colorings (Charron et al., 2009). Anthocyanins are beneficial to reduce chances of cardiovascular disease, cancer, and improve sugar metabolism (Wallace, 2011). Additionally, there are structural differences in the way these anthocyanins can be modified, primarily acylated or non-acylated. Importantly, acylated anthocyanins are more structurally stable and are more easily soluble in solvents. This allows for better use within the plant system, which is why the domesticated carrot is largely acetylated (Luo, 2022). Non-acylated anthocyanin are largely found in fruits like berries and grapes, are easily taken in by the consumer for antioxidant effects, yet still hold the same photoprotective effects (Khoo et al., 2017). Consequently, this lies the foundation that motivates the field to find ways to prevent acylation in commercial carrots for nutritional benefit.

In previous research, it was determined that non-acylated anthocyanins are more easily broken down for nutritious value by the body, but how much is unknown (Charron et al., 2009). The authors measured the amount of anthocyanins in blood plasma after feeding participants purple carrot juice. They found that throughout time, absorption of the non-acylated anthocyanin was higher by 11-14 fold, which was additionally corroborated by a study with steamed red cabbage (Charron et al., 2007).

After a series of gene mapping and gene expression analysis in *Daucus carota*, it was found that a cluster of three genes that code for SCPL-acetyltransferases correlated to a specific region known to control anthocyanin acylation (Curaba et al., 2019). One of these genes,

DcSCPL1, consistently was expressed alongside the purple phenotype and a gene known to promote carrot anthocyanin production (*DcMYB7*). The essentiality of this gene in controlling acylation of anthocyanins was confirmed after an exon splicing experiment that resulted in a low-acylation phenotype (Curaba et al., 2019). In a supporting study, overexpression of *DcMYB7* in orange carrots led to purple roots and silencing expression of solid purple (Deep Purple) carrots displayed yellow roots (Xu et al., 2019). This suggested that *DcSCPL1* is the primary gene controlling anthocyanin acylation and thus reducing its function with attention to reducing off-target modification, is the key to creating more nutritious carrots. Additionally, the expression of this gene can be measured visually by the color of the root of gene-edited carrots.

Similarly, this project seeks to create a truncation of the *Daucus carota DcSCPL1* gene within conserved regions in order to induce a loss of function for acylation. To do so, we will design gRNAs and crRNAs for CRISPR/Cas12 to cut out the amplicon of interest as found by previous cohorts of the project that studied *Scarlet Nantes* and the efforts of Dr. Fragoso efforts in the Dellaporta lab. Should the gRNAs be validated it will confirm the functionality of these regions for acylation as well as serve as a guide to create early termination within acylation-coding regions carrots in vivo. Ultimately, just one successful gRNA within an exon of interest should knock-out the function; however, off-target effects must be studied through experimentation and sequencing.

Materials and Methods

Isolation of Carrot gDNA from Leaf Tissue

gDNA of *Daucus carota v. Deep Purple* (MM75 s175) was isolated using the DNeasy Plant Handbook (2020) with the following modifications. The tissue was disrupted manually through the use of liquid nitrogen, a mortar, and pestle. Around 20 mg of the leaf from *D. carota* was ground into a fine powder and added into the preloaded AP1 Buffer and vigorously vortexed until the RNase was added. Furthermore, the DNeasy Mini spin column was spun and eluted a second time was added in order to create a second, more dilute extraction.

Nanodrop Spectrophotometer Analysis of gDNA Elutions

The clarity and yield of the extraction were then confirmed with the Denovix NanoValue Plus Spectrophotometer ds-11 following the DeNovix Protocol (2023) for both elutions.

Agarose Gel Electrophoresis Diagnostic of gDNA Elutions

An Agarose Gel Electrophoresis Diagnostic was completed as described by the Moreno Lab (2024) modified to use 20uL of the elution gDNA mixture in each well as opposed to 20uL of PCR reactions. In order to run a sample at around 100ng/ul, the final mixtures to be run on the gel included 14ul dH20, 2ul Elution 1, 4ul 6X loading dye then 10ul dH20, 6ul Elution 2, 4ul 6X loading dye respectively. Lane 1,7, and 13 were loaded with the 1kb ladder while Lane 2-6 was loaded with our first elutions and Lane 8-12 with the second elutions.

Identification and Analysis of the Gene of Interest

Following a presentation of information from Dr. Fragoso and research of primary literature on genes affecting acylation, the gene of interest, DcSCPL1, was designated to the group. Analysis of the candidate gene was done through studying the 2023 cohort's annotated

sequence on *Scarlet Nantes* and observation of the LOC108214129 within the NCBI Genome Data Viewer. The gene of interest was then downloaded and viewed in the SeqBuilder platform.

Identification and Annotation of Conserved Domains

The 2023 cohort also determined conserved domains through the NCBI Conserved Domains Database to identify potential editing sites to lose acylation function. Using the Lasergene-SeqBuilder, the file was annotated for these conserved sequences that was later used to design primer and guide RNAs for knockouts. The conserved regions within the *S. Nantes* sequences were confirmed to match in the *Imperator* and *Deep Purple* variety as well through a SeqMan Ultra Variant analysis. Through reading the annotated file on SeqBuilder, the conserved domains were noted within amplicons of interest using the add Features element of the program.

Strategy of Functional Analysis

For a functional analysis of the *DcSCPL1*, we strategized to create a truncation and therefore a loss-of-function analysis through designing a gRNA that can create an early termination in a susceptible and compatible coding sequence (within Exon 11). This included designing a Cas12 gRNA that could be turned into a crRNA compatible with the CRISPR/Cas12 complex to cut out the amplicon of interest in-vitro within the *Daucus carota* sequence. Should the knockout induce reduced functionality of acylation, it would confirm the functionality of DcSCPL1. In order to design the Cas12 gRNAs, qualifying PAM sites (5'TTTN3') were found through investigating the 2023 annotated file within SeqBuilder in regions that would encapsulate a conserved domain (near Exon 11). The 24 base pairs after this sequence then became gRNA guide candidates confirmed by the CASOFFinder tool to eliminate mismatches and bulges. The gRNA pMB24-02 was then made into a duplex DNA fragment by IDT.

Amplification of Interval of Interest

The sequence of interest containing Exon 11 was extracted using the 2023 cohort's designed primers and amplified using PCR.

Primer Sequences and Characteristics

The primers designed by the previous year were found in the annotated LOC108214129 of the DcSCPL1 gene in SeqBuilder and included the following information: Forward Primer pMB23-09 (5'AAGCTTCCAACTAAAACCCA3') at Tm 55.05°C and Reverse Primer pMB23-10 (5'GAGGGCCCTATTCAAAAGAA3') at Tm 54.99°C. This amplicon should be spaced 4512-5542 in the gDNA locus with a length of 1.031kb.

PCR Amplification of Amplicon (s99 and s146)

The PCR was performed as per the Moreno Lab (2024) with one set of four reactions receiving HF Phusion PCR Master Mix and another four using GC Phusion PCR Master Mix. Each set involved two controls and 2 experimentals with each set either receiving or not receiving DMSO. The ThermoFisher ™ Calculator optimized the primer characteristics for cycling with the final protocol involving initial denaturation of 98°C 30 seconds, 30 cycles of 98°C 15 seconds, 63°C 15 seconds, and 72°C 30 seconds, a final extension of 72°C of 5 minutes, then was held at 4°C for a week. This was repeated with the s99 Deep Purple gDNA sample as well as the s146 Imperator sample.

Agarose Gel Electrophoresis Diagnostic

An AGE diagnostic was performed to confirm the amplicon length as per the Moreno Lab (2024). This was repeated with the s99 *Deep Purple* gDNA sample as well as the s146 *Imperator* sample.

Cloning of gRNA Amplicons into PCR-Blunt II TOPO

The gRNA amplicon, pMB24-02, was cloned into a pCRBluntII-TOPO vector using the Zero Blunt TOPO (LifeTechnologies, 2014) protocol with the following modifications. The PCR reaction from the GC DSMO(-) conditions were used for the integrations. Additionally, two plates were created for both the *Imperator* and *Deep Purple* samples, one plate spread with 50uL of transformation and the other spread with 150uL. Furthermore, within the plasmid, it also has the KanR and CCdB genes integrated thus the agar also includes kanamycin (50mg/L).

Creation of crRNA: Cloning of Synthesized Cas12 gRNAs into pMM260

The ligation and cloning transformations were done by Dr. Moreno by the following protocols. A Cas12-gRNA Duplex was created through Golden Gate/BsaI Cloning and a ligation prescription in order to integrate the gRNA into pMM260. Then, this was transformed into NEB-10B cells that can be further tested to confirm integration of gRNA. The plasmid also contained the AmpR and SacB gene thus the agar included Ampicillin and Sucrose.

Confirmation of gRNA pMB24-02 in s99 TOPO Vector, s146 TOPO Vector, and p260

The confirmation of the gRNA amplicons in each of the vectors was confirmed by the following:

Screening Colonies by PCR

The colonies were screened through a PCR performed as per the Moreno Lab (2024) with the following modifications. All reactions were done in -DSMO and GC PCR Mix conditions. For the TOPO vectors, the forward primer P-264 (5'GCTATGACCATGATTACGCCAAG 3') and reverse primer P-265 (5' CTGGCCGTCGTTTTACAACGT3') were used. For the p260 vector,

the primers P-1422 and P-1384 were used. 7 experimental colonies were chosen and screened by PCR. Again, the final protocol involved initial denaturation of 98°C 30 seconds, 30 cycles of 98°C 15 seconds, 63°C 15 seconds, and 72°C 30 seconds, a final extension of 72°C of 5 minutes, then was held at 4°C for a week.

Agarose Gel Electrophoresis Diagnostic

An AGE diagnostic was performed to confirm the amplicon length as per the Moreno Lab (2024). This was repeated with the s99 *Deep Purple* TOPO sample, the s146 *Imperator* TOPO sample, and p260 sample.

Isolation of Plasmid DNA

Liquid cultures grown from the confirmed colonies (colony 1 for s99, colony 2 for s146, and colony 1 for p260) then DNA was isolated through the QIAprep Spin Kit Protocol with the following modifications. PB buffer was not used due to not needing the additional wash. Lastly, elutions were diluted to create 50ng/uL aliquots.

Nanodrop Spectrophotometer Analysis of gDNA

The clarity and yield of the extraction were then confirmed with the Denovix NanoValue Plus Spectrophotometer ds-11 following the DeNovix Protocol (2023) for the s99 *Deep Purple* TOPO sample, the s146 *Imperator* TOPO sample, and p260 sample.

SeqBuilder Sequence Analysis

After the DNA samples were sequenced by Plasmidsaurus using Oxford Nanopore technology, SeqMan Ultra was used to see if the target amplicon was cloned through comparison with the reference sequence (5’ ggatagtagagtgtatcatgcaaa 3’).

Results

Isolation of Carrot gDNA from Leaf Tissue

Carrot gDNA was successfully isolated from leaf tissue as confirmed by AGE and spectrometry. The results of the nanodrop for the first extraction of genomic DNA of the *Deep Purple* variant were all within the normal ranges as indicated by Figure 3. Additionally, the AGE in Figure 2 indicated the extracted DNA sequence is over 15kb for both the elutions.

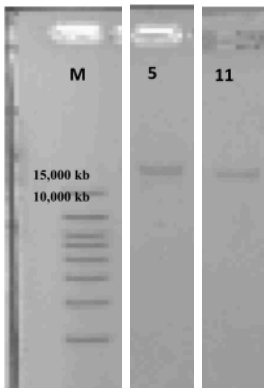


Figure 2. MM75 s146 Deep Purple gDNA Extraction Elution AGE analysis

All bands ran through the agarose gel electrophoresis with genomic DNA extracted displayed band sizes above 15,000 kb with the MM75 s146 Elution 1 ran in Lane 5 and MM75 s146 Elution 2 ran in Lane 11.

Figure 3. Spectrophotometric Data for MM75 s146 (Deep Purple) Elution 1 and Elution 2

Results from a Nanodrop Spectrophotometer displayed s146 Elution 1 and Elution 2

	MM75 s146 Elution 1	MM75 s146 Elution 2
A260	1.3902	0.3571
260/230	2.677	3.458
260/280	1.860	1.851
Yield (ng/ul)	69.511	17.856

Identification and Analysis of the Gene of Interest

The *DcSCPL1* gene was determined to be of interest through literature review and found within the larger *D. carota* genome in NCBI Genome Data Viewer as demonstrated in Figure 4. The 2023 cohort's annotations on conserved domains as displayed in Figure 5 in SeqBuilder, were analyzed and Exon 11 was assigned to me and located. Specifically, my section of the project involved the amplicon created by pMB23-09 and pMB23-10 as found in Figure 5. Furthermore, it was found that there was 100% conservation between the *Imperator* sequence and *Sativas* comparison variant of carrot for the *DcSCPL1* gene as demonstrated by the blank Variant table in Figure 6. The pathway to the identification of the gene of interest and the strategy of to see if it has an effect on acylation was created as outlined in Figure 1. Following the pathway, a Cas12 gRNA, pMB24-02, was created with details outlined in Figure 8 from the IDT company. The guide was chosen from the list outlined in Figure 7.

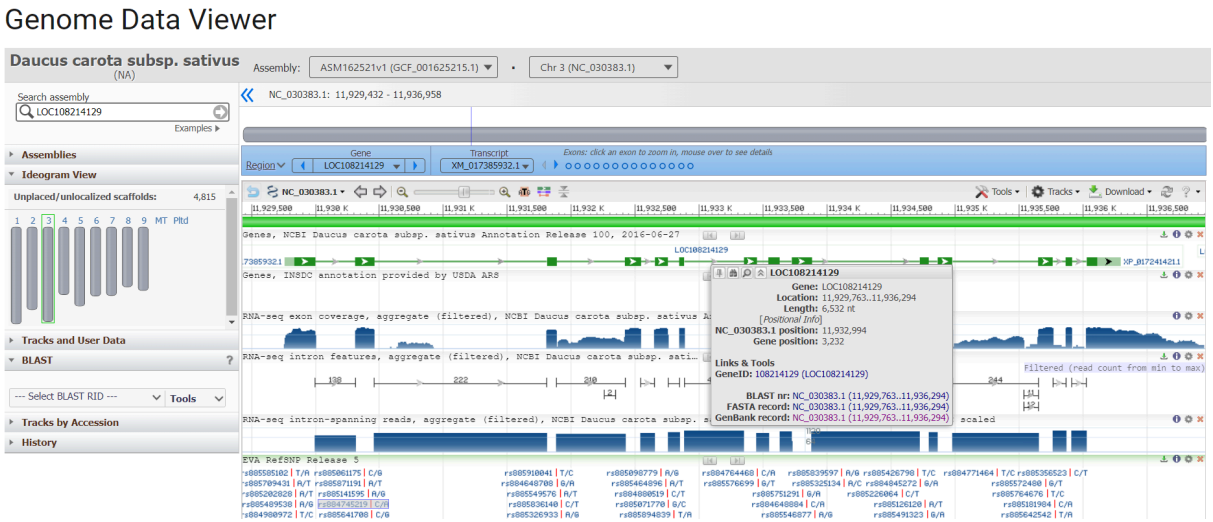


Figure 4. The location of the gene of interest DcSCPL1 (LOC108214129)

Through the NCBI Genome Data Viewer, the gene of interest DcSCPL1 (LOC108214129) is displayed in its reference to a general D. Caratus genome (subsp. sativus).

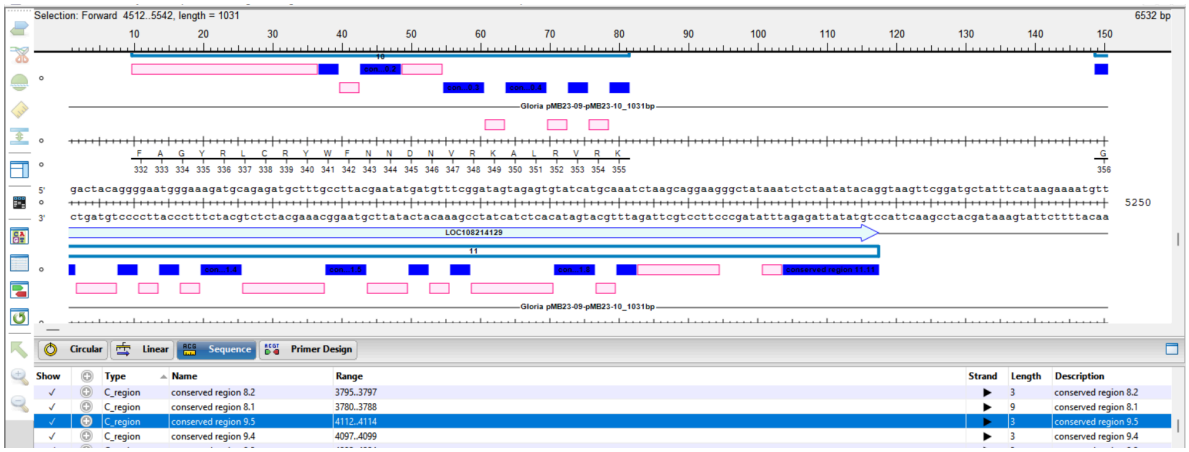


Figure 5. Annotated Conserved Domains within the Gene of Interest

Through the annotated files of the previous year, the conserved domains of the DcSCPL1 gene was found in various places highlighted in violet that concentrated in specific regions between their designed primers. The pink regions show the variant sequences and the purple regions show the conserved sequences in Imperator Vs. Sativas. This shows the conserved regions within the amplicon created by pMB23-09 and pMB23-10 between exons 10 and 11 of the gene.

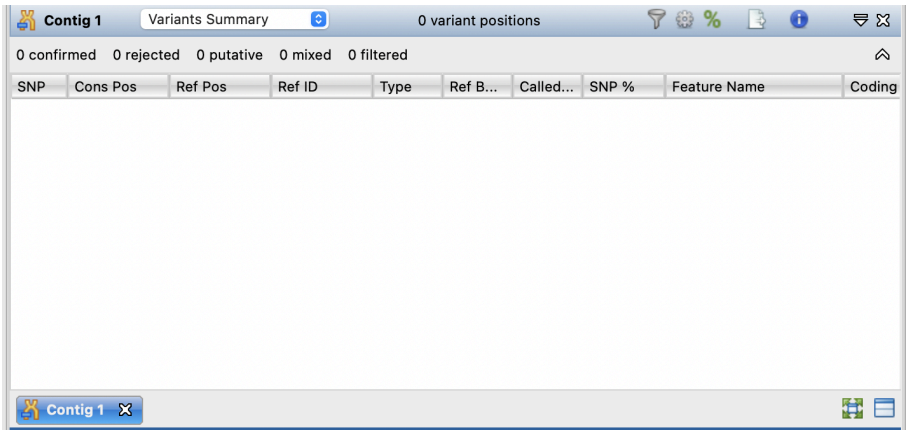


Figure 6. SeqMan Ultra Variant Table of Annotated *Imperator* Vs. *Sativas*

The annotated file showed a fully conserved genome with no variants to the genome sequenced online.

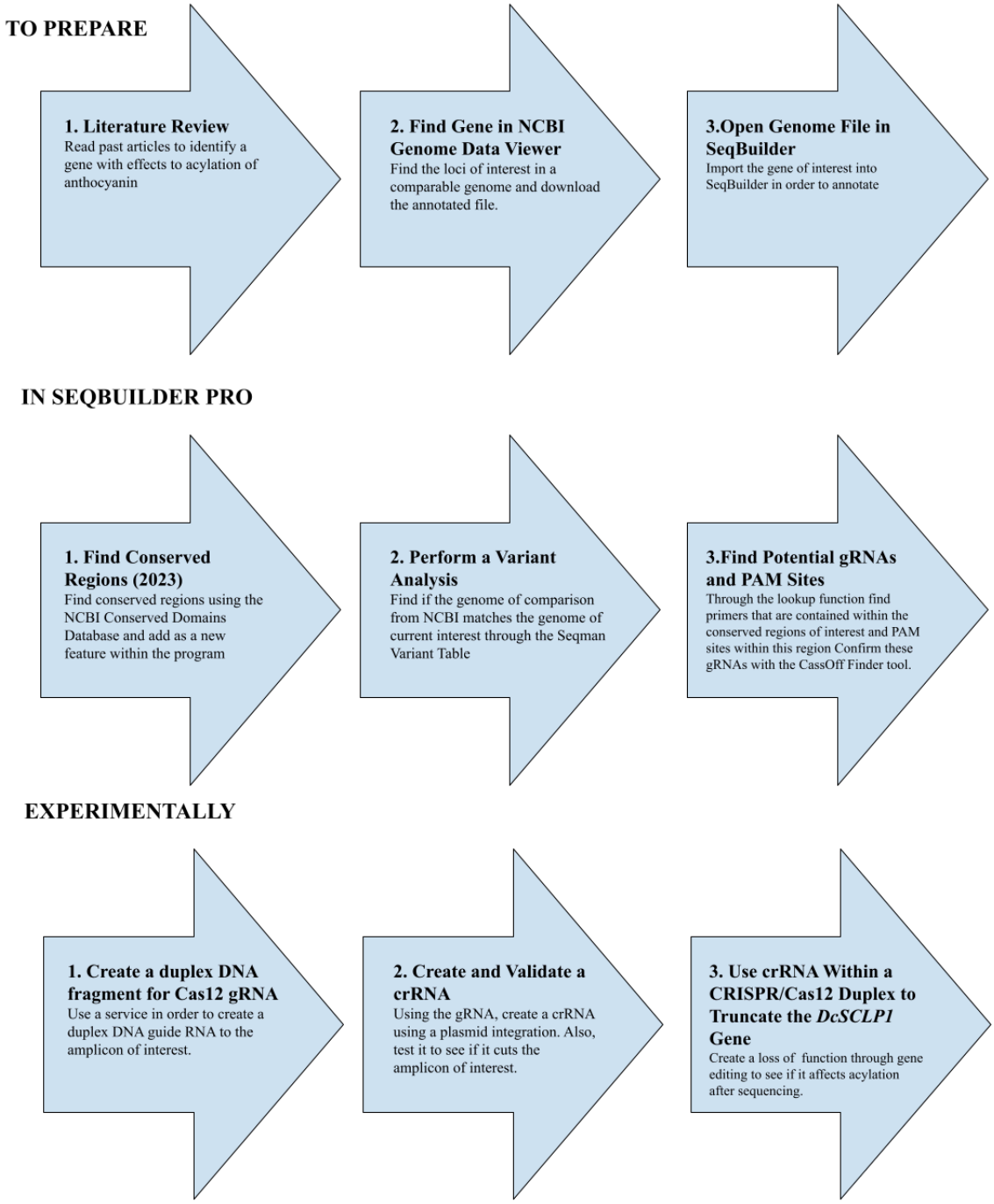


Figure 1. Strategy to Identify and Functionally Analyze the DcSCPL1 Gene

Flowchart of the process used to annotate the gene and conserved domains from ideation to steps aiding in designing methodology to functionally analyze the gene of interest.

Exon of Interest	PAM Sequence (5'-3')	gRNA Sequence 24 bp (5' - 3')	sgRNA Direction	Genome Location	Location Relative to Exon	Contributor	Mismatches	Bulge Size	No of targets
Exon 11	tttg	aaaaaagaatgtagaactaaacag	Forward	5061-5084	Left intron	Miya	0	0	1
Exon 11	tttg	ccttacgaatatgatgtttcggat	Forward	5138-5161	Within	Miya	0	0	1
Exon 11	tttc	ggatagtagagtgtatcatgcaaa	Forward	5158-5180	Within	Miya	2	1	1
Exon 11	tttc	ttatgaaatagcatccgaacttac	Reverse	5218-5241	Right Intron	Miya			
Exon 11	tttc	ataagaaaatgttgatttattaat	Forward	5238-5261	Right Intron	Miya			
Exon 11	ttta	aatataactacaatcatgctcaat	Forward	5425-5448	Right Intron	Miya	0	0	1
Exon 11	tttg	acaaaggagttgagggagggagga	Forward	5477-5500	Right Intron	Miya	0	0	1

Figure 7. Potential Cas12 gRNAs Designs

The table outlines the potential gRNAs found from the conserved regions. Highlighted in purple is the Cas12 gRNA chosen.

Gene ID	DcSCLP1	**Purification**	STD
Exon	11	**nmoles**	76.3
Coordinates	6654..6681	**Vol 1X TE (uL)**	763
Contributor	Miya	**pMM260 Cloning**	Yes
Guide ID	pMB24-02	**Colony PCR**	Yes
forward oligo	AGATggatagtagagtgtatcatgcaaa	**Fwd Primer gDNA**	pMB23-09

reverse oligo	GGCCtttgcatgatacactct actatcc	**Rev Primer gDNA**	pMB23-10
Scale	100nm	**s146_Deep Purple (kb)/ s99_Imperator (kb)**	1.27

Figure 8. Table of Information of Designed Cas12 gRNA pMB24-02

The table outlines the Oligos' information designed from the gRNA sequences found in the genome from identified PAM sites to create the guide pMB24-02.

Amplification of Interval of Interest

The interval of interest was found and amplified in both variants of *D. carota*. The primers for the amplicon of interest, pMB23-09 and pMB23-10, were found within the 2023 cohorts annotated sequence in SeqBuilder as demonstrated in Figure 9. This amplicon was successfully amplified in both the *Deep Purple* and *Imperator* sequence as demonstrated in the AGE analysis having a band at the 1.031kb mark within Figure 10A and 10B respectively.

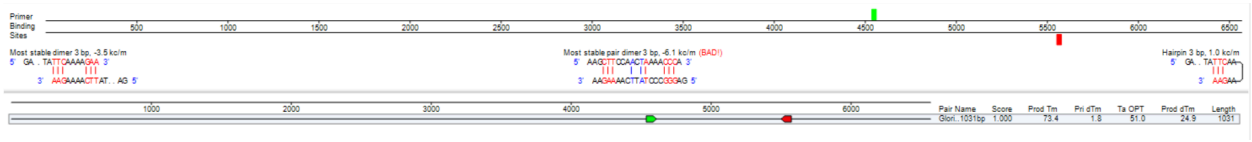
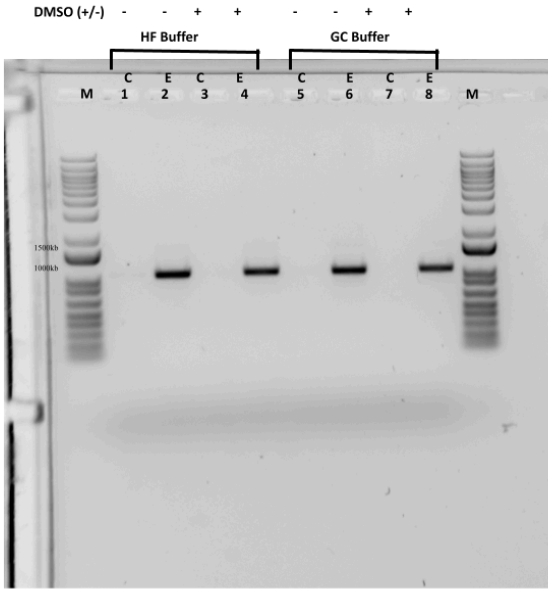


Figure 9. The Position Primers pMB23-09 and pMB23-10 in LOC108214129 of the *DcSCPL1* gene

The Forward Primer pMB23-09 is marked with the green flag at the 4512 position and the reverse primer pMB23-10 is marked with the red flag at position 5542 in context of the entire

LOC108214129 for an amplicon of length 1.031kb. The bottom view is a slightly smaller view of the information demonstrated at the top of the image.

A. 

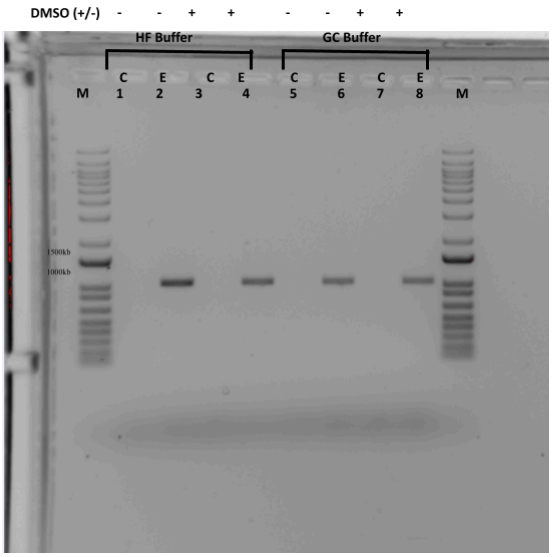
B. 

Figure 10. AGE Analysis of Amplicon from Primers pMB23-09 and pMB23-10 on *Deep Purple*(A.) and *Imperator* (B.)

In both the Deep Purple and Imperator sequences, the primers produced amplicons of lengths around 1000bps for all experimental trials in both the HF and GC fusion mix and with and without DMSO. C means control while E means experimental.

Cloning of gRNA Amplicons into PCR-Blunt II TOPO

The gRNA attached amplicon of interest was integrated into the TOPO vector for both the *Imperator* and *Deep Purple* samples as demonstrated by the slightly larger band of 1300 bps as illustrated in Figure 11. This was further confirmed by the presence of the gRNA sequence (5' ggatagtagagtgtatcatgcaaa 3') in the sequenced clone colonies (colony 1 for s99 and colony 2 for s146) as found by SeqBuilder in Figure 12 for both samples.

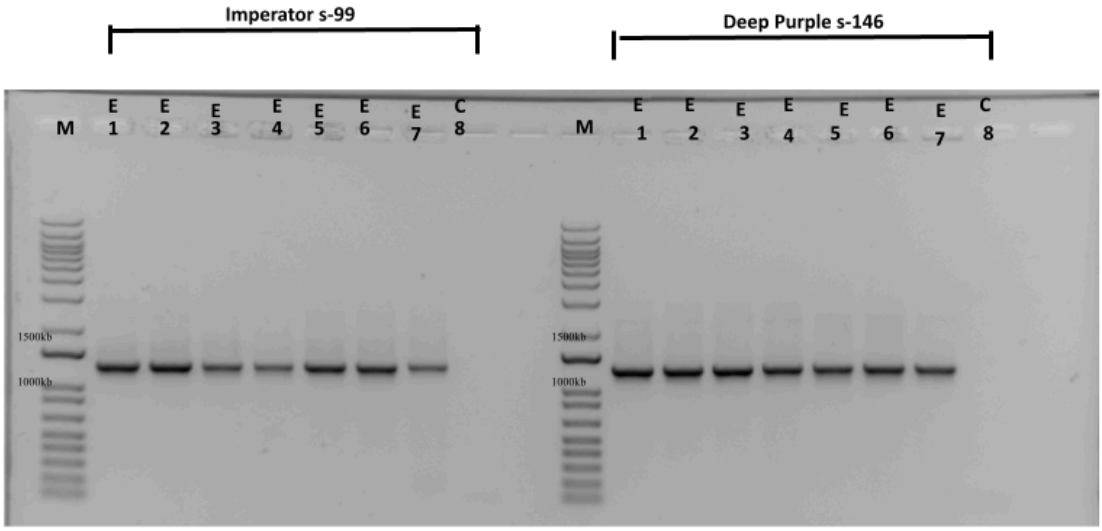
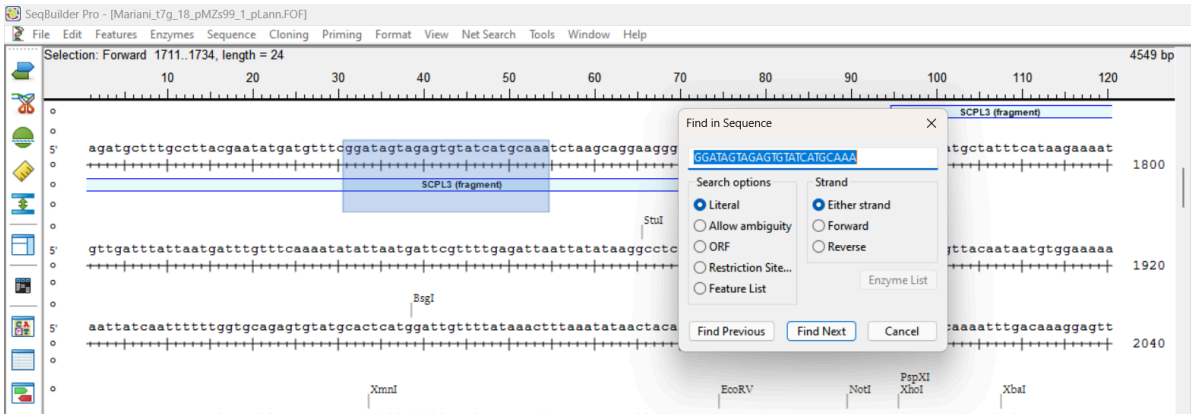


Figure 11. AGE Analysis of gDNA Amplicons in Blunt II TOPO PCR

In both the Deep Purple and Imperator sequences, the gDNA amplicons appear to be integrated into the clones as the TOPO added sequence of 240 bps should be integrated with the original 1030 bps of the amplicon of interest resulting in bands at around the 1270bps mark. All colonies tested indicated a band around the 1300 bp mark as illustrated. C means control while E means experimental.

A.



B.

Figure 12. SeqBuilder Confirmation of gRNA Presence in s99 *Imperator* (A.) and s146 *Deep Purple* (B.) Sequences

The sequenced genome from the two colonies was shown to contain the gRNA sequence of interest as highlighted (5' ggatagtagagtgtatcatgcaaa 3').

Creation of crRNA: Cloning of Synthesized Cas12 gRNAs into pMM260

The crRNA was successfully synthesized as demonstrated by the extracted amplicon appearing near the 200bp line in Figure 13 and the gRNA sequence (5' ggatagtagagtgtatcatgcaaa 3') transformed colony 1's sequenced DNA in SeqBuilder in Figure 14.

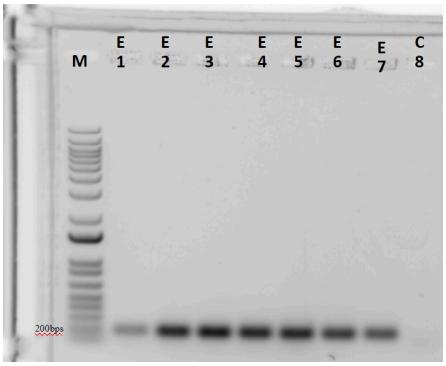


Figure 13. AGE Analysis of Cas12 gRNAs into pMM260

Illustrated is the AGE analysis of recombinant colonies with the vector pMM260 with the target insert fragment, the crRNA of ~165 bps in each experimental lane, which are all approximately 200 bps.

Figure 14. SeqBuilder Confirmation of gRNA Sample in pMM260

The sequenced genome from the pMM260 vector was shown to contain the crRNA sequence of interest as highlighted.

Discussion

Successful Elution of gDNA

The 260/280 ratio is above the standard 1.8 ratio at 1.860, indicating low contamination of aromatic amino acids and the 260/230 ratio is within the standard 2.0-2.2 range at 2.677, indicating low contamination of salts and polysaccharides. Additionally the concentration of DNA is decent at 69.511ng/uL. The second dilute elution is slightly contaminated and at a lower concentration at 17.856 ng/ul and the 260/230 ratio at 3.458 and the 260/280 ratio at 1.851 indicating polysaccharide, salt, and amino acid contamination. Moving forward, it will be sufficient to use the first elution to use as the gDNA for further experimentation on *Deep Purple*. Further research can use these aliquots for in vitro testing of primers and sequencing can be done to confirm target edits in knockouts and clone vectors.

Confirmation of DcSCPL1 Conservation within Dacus carota

Upon analysis in the SeqMan Ultra Variant table, it was confirmed that the DcSCPL1 gene is perfectly conserved within the *Daucus carota (*v. *Imperator* and v. *Deep Purple)* genome in comparison to the *Scarlet Nantes* genome suggesting the necessity of this sequence for acylation function and sparking interest in the amplicons chosen. Ultimately, upon analysis of the annotated sequence from 2023's cohort, this year's cohort successfully created gRNAs for CAS12 in the exons of interest within previously identified conserved regions and integrated them into vectors.

Confirmation of 2023 Primer Designs (pMB23-09 and pMB23-10)

The primers (pMB23-09 and pMB23-10) designed from the previous year successfully extracted the amplicon of interest within Exon 11 as confirmed by gel electrophoresis and PCR.

This demonstrates the presence of the gRNA sequence in *Imperator* and *Deep Purple*, insinuating compatibility to be edited. Furthermore, the AGE results demonstrated the darkest bands for the -DSMO and GC mix conditions, which should be used for further clone PCRs.

Confirmed Compatibility of gRNA pMB24-02 with Exon 11 Amplicon

Importantly, gene sequencing that confirmed the potential gRNA (pMB24-02) was successfully cloned into both the transformed *Imperator* and *Deep Purple* vector colonies which confirms the guide's presence within *Dacus carota*. This suggests the genome's susceptibility to be edited by the guide designs with CRISPR/Cas12. Interestingly, the primers worked for the *Deep Purple* sequence, further confirming homology between the variants.

Creation and Integration of a valid crRNA Sequence

Sequencing also confirmed that gRNA pMB24-02, was successfully integrated into the pMM260 vector. This indicates that the ligation techniques were successful and the vectors are compatible with the gRNAs of interest. Thus, the crRNA can be tested to see if it can cut out the amplicons of interest with the Cas12. However, the sequenced data did not line up completely within SeqMan Ultra Variant analysis. The guides are integrated within the vector but it is unclear which guides put them into place. This requires perhaps a replication of procedure or new guides to be synthesized.

In Continuation: In Vitro and In Vivo Functional Testing of Created crRNA

In future experimentation, the amplicon of interest can be silenced through CRISPR/CAS12 by using the designed crRNA to create truncations that cause early termination of acylation-coding sequences (within Exon 11). Ultimately, genetic sequencing can confirm if the complex can appropriately cut the sequence with minimal bulging, mismatches, or off-target sequences. Then, an in vivo study can be conducted to see if the loss of function displays any unexpected effects. Ideally, the carrots grown with the edited sequences can maintain their anthocyanin production without the acylation regions and conserve their color. In early experimentation, an albino carrot in the leaves would indicate successful knockout of the region due to an integration of the phytoene desaturase gene as a positive control (Curaba et al., 2019). Upon confirmation of phenotypic test in these knockouts, this leads to carrots with the conserved aesthetics and commercial viability of today with the anthocyanin bioavailability of interest.

References

Charron, C. S., Kurilich, A. C., Clevidence, B. A., Simon, P. W., Harrison, D. J., Britz, S. J., Baer, D. J., & Novotny, J. A. (2009). Bioavailability of anthocyanins from purple carrot juice: effects of acylation and plant matrix. *J Agric Food Chem*, 57(4), 1226-1230. https://doi.org/10.1021/jf802988s

Curaba, J., Bostan, H., Cavagnaro, P. F., Senalik, D., Mengist, M. F., Zhao, Y., Simon, P. W., & Iorizzo, M. (2019). Identification of an SCPL Gene Controlling Anthocyanin Acylation in Carrot (Daucus carota L.) Root. *Front Plant Sci*, 10, 1770. https://doi.org/10.3389/fpls.2019.01770

DeNovix. (2023, November 16). ds-11-series-user-guide. DeNovix.

Khoo, H. E., Azlan, A., Tang, S. T., & Lim, S. M. (2017). Anthocyanidins and anthocyanins: colored pigments as food, pharmaceutical ingredients, and the potential health benefits. *Food Nutr Res*, 61(1), 1361779. https://doi.org/10.1080/16546628.2017.1361779

Life technologies. (2014, February 24). Zero Blunt TOPO PCR Cloning Kit User Manual.

Luo, X., Wang, R., Wang, J., Li, Y., Luo, H., Chen, S., Zeng, X., & Han, Z. (2022). Acylation of Anthocyanins and Their Applications in the Food Industry: Mechanisms and Recent Research Advances. *Foods*(Bioactive Compounds, Antioxidants, and Health Benefits). https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9316909/

Wallace T. C. (2011). Anthocyanins in cardiovascular disease. Advances in nutrition (Bethesda, Md.), 2(1), 1–7. https://doi.org/10.3945/an.110.000042

Xu, Z. S., Yang, Q. Q., Feng, K., & Xiong, A. S. (2019). Changing Carrot Color: Insertions in DcMYB7 Alter the Regulation of Anthocyanin Biosynthesis and Modification. *Plant Physiol*, 181(1), 195-207. https://doi.org/10.1104/pp.19.00523

Acknowledgments

We thank Dr. Moreno for the mentorship, project overview, and resources in order to implement this project. We thank Dr. Fragoso for presenting the need and foundation for the project as well as the computational background/implementation of our data as well as the larger Dellaporta Lab. We thank April Pruitt and Pray Miao for facilitating each weekly experiment as well as their feedback and mentorship on maintaining a lab notebook and creating the final report. We thank the support of the laboratory teaching staff: Carol, Mariani, Geetha Suresh and greenhouse personnel Christopher Bolick. I thank Raj Letchuman for sharing supplies and collaborating on ideation as a lab benchmate. Finally, I thank my group members: Nico Paredes, Arden Parrish, Himani Pattisam and Manuel Perez for our collaboration on the overview of this project with everything from ideation of truncation strategy to exchanging elutions of gDNA to the weekly discussion of data and protocol. This work was supported by funding granted for the MCDB201L course from Yale University.