Kentucky Journal of Undergraduate Scholarship

Volume 1 | Issue 1 Article 10

May 2017

Rosa hybrid gene GAPC is mutated in the presence of the Rose Rosette Virus

William A. Beams
Brescia University, william.beams@brescia.edu

Eva L. Burelos
Brescia University, eva_burelos@yahoo.com

Charles C. Hundley
Brescia University, connorhundley@yahoo.com

Landon R. Dame
Brescia University, landon.dame@brescia.edu

Laura G. Sims
Brescia University, lzs0061@auburn.edu

See next page for additional authors

Rosa hybrid gene GAPC is mutated in the presence of the Rose Rosette Virus

Cover Page Footnote
Footnotes: 1 First two authors equally contributed to this work. 2 Address Correspondence to Jacob Adler, 717 Frederica Street, Owensboro, Kentucky, 42301; email: jacob.adler@brescia.edu; Phone: 270-686-4284 Acknowledgments: We would like the thank Dr. William Tyler and the Western Kentucky Botanical Garden for introducing us to the global issue of Rose Rosette Disease and Dakota Durrett for help with the final figure. We would also like to thank Brescia University for allowing us to research this project as undergraduate students and funding our work.

Authors
William A. Beams, Eva L. Burelos, Charles C. Hundley, Landon R. Dame, Laura G. Sims, and Jacob J. Adler

Rosa hybrid Gene *GAPC* is Mutated in the Presence of the Rose Rosette Virus[1]

William A. Beams
Brescia University

Eva L. Burelos
Brescia University

Charles C. Hundley
Brescia University

Landon R. Dame
Brescia University

Laura G. Sims
Brescia University

Jacob J. Adler, PhD
Brescia University

Abstract: *Rose Rosette Disease (RRD) harms the global rose supply by modification of the growth and development in rose cultivar. RRD spreads via a negative-sense RNA plant virus transmitted by eriophyid mites. Importantly, there is no pre-existing knowledge about the biochemistry by which this virus debilitates roses. Here we implicate glyceraldehyde-3-phosphate dehydrogenase (GAPDH), one of the major metabolic enzymes in plants, as a possible target of the virus. Genomic DNA of the cytosolic form of the protein encoded by GAPC was extracted from both virally-infected and non-infected samples of the Rosa hybrid cultivar Rosa Tropicana. The sequence results provided several distinct differences in the GAPC gene of the non-infected rose compared to the virally-infected rose. Importantly, these modified nucleotide bases resulted in a putative protein sequence containing four unique non-conserved amino acid substitutions in the GAPDH enzyme. This study provides the first evidence of a gene impacted in virally-infected rose plants.*

Keywords: GAPDH, Rose, Witches' Broom, sequencing, growth control, Rosa Tropicana

Rose Rosette Disease (RRD) is one of the most devastating diseases of roses (Conners, 1941). These diseased plants show symptoms of "witches' broom" with bright red misshapen leaves developing from the buds, distortion of the stem, and weak apical growth (Gergerich and Kim, 1983; Keiffer, 1975) (Figure 1). Recent PCR and electron microscopy studies have identified a viral origin to the disease (Windham, et al., 2014), characterized by a negative-sense RNA plant virus, termed the Emaravirus (Connors, 1941; Laney, et al., 2011).

1 The authors would like the thank Dr. William Tyler and the Western Kentucky Botanical Garden for introducing us to the global issue of Rose Rosette Disease and Dakota Durrett for help with the pathway figure. We would also like to thank Brescia University for allowing us to research this project as undergraduate students and funding our work.

The Emaravirus can be compared to the known Fig mosaic virus, which is currently infecting Fig trees in Japan (Laney, et al., 2011). The Emaravirus is transmitted to most rose cultivar by an eriophyid mite, called *Phyllocoptes fructiphilus*, which allows for rapid spread of the RRD (Amrine and Zhao, 1998). There are limited management techniques that exist to combat this viral disease (Windham, et al.,2014). Highlighting a majority of treatments is the use of miticides to combat the spread of the disease by eriophyid mites (Epstein, et al., 1997). Treatment techniques are complicated because the disease can be confused with herbicide damage due to its variation of symptoms (Windham, et al., 2014). Therefore, there is a need to understand how the virus manipulates the growth and metabolism of the rose. By understanding the genome of the infected roses, we can provide more specific management and treatment strategies. In pursuit of possible impacted pathways, it was hypothesized to include those involved in metabolism and cell growth, due to the visible changes associated with "witches' broom". The protein we examined in this study is the classical glycolytic protein, glyceraldehyde-3-phosphate dehydrogenase (GAPDH). GAPDH shows important functional diversity in plants and animals, demonstrating important roles in energy production and influences in membrane fusion, microtubule bundling, endocytosis, and DNA repair. Some intriguing influences of GAPDH include viral pathogenesis, regulation of apoptosis, and overall plant cell metabolism (Sirover, 1999). GAPDH genes are believed to have been formed by a series of gene duplication events occurring during the time when plants on land first emerged (Brinkmann, et al., 1989; Teich, et al., 2007). One of these genes includes the cytosolic form of GAPDH, the *GAPC* gene. *GAPC* encodes a GAPDH protein that catalyzes an enzymatic reaction for energy production in plants during glycolysis (Plaxton, 1996). It was this *GAPC* gene that we examined in this study. The goal of this study was to examine virally-infected *Rosa hybrid* plants to determine if they have an altered sequence in the *GAPC* gene.

Method

The rose plants used in this study were selected from a population at the Western Kentucky Botanical Garden. Plant leaves were taken fresh from the *Rosa hybrid* cultivar Rosa Tropicana from plants presenting the RRD and others not visibly presenting RRD (Figure 1). The diseased plants were verified by local botanists via observed symptoms of the disease. The cultivar Rosa Tropicana was used because of accessibility. Most local cultivars are removed from the population as soon as they display disease symptoms. Additionally, by focusing on one cultivar, multiple sample replicates insured reproducibility of the results. Leaves from two virally-infected and two non-infected independent plants were gathered fresh and taken to the laboratory, where the genomic DNA was extracted and purified. Then, the *GAPC* gene was amplified and isolated from the genomic DNA, using a two-step nested PCR technique, via manufacture instructions, which

uses *Taq* polymerase to specifically amplify DNA (Bio-Rad, 2008). Purified partial *GAPC* genes were sub-cloned into a pJET1.2 vector and sequenced using primers overlapping both the plasmid and the gene itself (Figure 2). Bioinformatic analysis was performed on the raw sequences using the iFinch for Educators software (Geospiza, Inc.). The sequences with a minimal quality chromatogram greater than 20 were chosen.

Results

Six sequence files were obtained for two independent non-infected plants and three sequence files were obtained for two independent virally-infected plants (Figure 2). While multiple sequence results were obtained for all four plants harvested, those not meeting the stringent minimum threshold for quality were not included for processing. The corresponding high-quality individual genomic sequences were merged into contig files for both the virally-infected plant's *GAPC* gene and non-infected plant's *GAPC* gene. The sequence assemblies for both contig sequences are presented (Figure S1 and Figure S2). Importantly, any sequences that came from only one sequencing reaction or one plant were not used in the finalized contig sequences for either condition. Both contig files were then verified to contain the *GAPC* gene using BLASTn against the reference genomic sequence database (NCBI), including crosschecks with verified *GAPC* sequences from *Arabidopsis thaliana* and other plant species. Finally, the two contig sequences were uploaded to GenBank (Figure 2). Several base differences were noted between the non-infected (KT806117) and the virally-infected (KT806118) contig DNA sequences upon examination.

Putative mRNA sequences were then obtained by annotating the gene using BLASTn against the GenBank mRNA sequence database of known *GAPC* mRNA (NCBI). These mRNA sequences identified 5 conserved coding regions in the non-infected plant's *GAPC* sequence (KT806117) and 3 conserved coding regions in the virally-infected plant's *GAPC* sequence (KT806118). Importantly, both putative mRNA sequences mapped back to other reported wild-type *Rosa hybrid* sequences by BLASTn analysis (NCBI). These coding regions were then used to obtain putative amino acid sequences via EMBOSS Transeq software (Li, et al., 2015) and verified using BLASTp against the protein database of known GAPDH proteins (NCBI). The amino acid sequence of the *Rosa hybrid* non-infected (KT806117) was compared to a previously reported sequence from another *Rosa hybrid* study (AEQ49677.1) and *Arabidopsis thaliana* (NP_187062.1) via PSI-BLAST analysis (NCBI) (Figure 3). Congruent with the previously reported amino acid sequence for *Rosa hybrid*, we report here 100 % identity with our non-infected *Rosa hybrid*. Therefore, we have independently obtained the likely partial amino acid sequence encompassing the catalytic domain of the GAPDH protein in *Rosa hybrid*. Interestingly, when the *Arabidopsis thaliana* amino acid sequence was compared to either *Rosa hybrid* sequence, it had 9 substitutions noted in the sequence. Of these 9 substitutions, 6 are

considered non-conserved amino acids. That many amino acid changes between these two species for a major glycolytic enzyme could provide evidence for phenotypical differences between these two plant species as it relates to their metabolism. Currently, no evidence exists that suggests that these substitutions modify the structure nor the function of GAPDH in roses compared to *Arabidopsis thaliana*.

Finally, the virally-infected and the non-infected putative GAPDH amino acid sequences were compared and indicated 10 substitutions in the virally-infected protein sequence (Figure 4). Of these 10 substitutions, 4 were non-conserved. Additionally, the sequence was also compared to another Rosa hybrid study (AEQ49677.1) and showed the same 10 modifications compared with the non-infected (KT806117) sequence (Figure S3). This level of amino acid substitutions is similar to that seen in the comparison of *Rosa hybrid* and *Arabidopsis thaliana*, two different species of plants. Thus, this degree of modification is similar to species differences between plants. It is hypothesized that this level of change could impact the structure and thus the functioning of GAPDH in virally-infected plants, however, future studies will be required to determine if such a change would impact protein structure.

Discussion

The results obtained in the present study suggest that the presence of RRD impacts the genomic sequence of *GAPC* in *Rosa hybrid*. Importantly, these genomic changes in virally-infected roses result in a putative amino acid sequence that differs from non-infected roses from two different studies. Thus, the putative amino acid sequence is likely modified in *Rosa hybrid* plants displaying RRD. With little known about the pathways of this virus and with minimal improvement in disease treatment strategies, the RRD is becoming a potential epidemic for rose cultivar. The sequenced results gathered here provide the primary evidence for a genomic modification seen in rose plants impacted by RRD. Further, the results serve as the baseline evidence for the *GAPC* gene as a potential gene editing target for the treatment of the RRD. Moving forward, further research will be needed to determine if there are implications on GAPDH signaling and other cellular pathways in rose cultivar.

References

Amrine, J. W., & S. Zhao. (1998). Research on Aerial Dispersal of *Phyllocoptes fructiphilus (Acari: Erio-phyidae)*, Vector of Rose Rosette Disease. *American Rose, 3*, 28-29.

Biotechnology Explorer Cloning and Sequencing Explorer Series [Curriculum manual, Catalog #166-5000EDU]. Hercules, CA: Bio-Rad.

Brinkmann, H., Cerff, R., Salomon, M., & Soll, J. (1989). Cloning and sequence analysis of cDNAs encoding the cytosolic precursors of subunits GAPA and GAPB of chloroplast glyceraldehyde-3-phosphate dehydrogenase from pea and spinach. *Plant Mol Biol. 13*, 81–94.

Conners, I.L. (1941). *Twentieth annual report of the Canadian Plant Disease Survey 1940*. Department of Agriculture: Ottawa, Canada, 98.

Epstein, A., Hill, J., & Nutter, F. (1997). Augmentation of rose rosette disease for biocontrol of multiflora rose (Rosa multiflora). *Weed Sci. 45*, 172-178.

Gergerich, R. C., & K. S. Kim. (1983). A description of the causal agent of rose rosette disease. *Arkansas Farm Research. 32*, 7.

Huang, X. & Madan, A. (1999). CAP3: A DNA sequence assembly program. *Genome Res. 9*, 868-877.

Keiffer, M. M. (1975). Eriophyoidea. In L. R. Jeppson, M. M. Keiffer, and E. W. Baker (Eds.), *Mites Injurious to Economic Plants* (pp.327-396). Berkeley, CA: University of California Press.

Laney A. G., Keller K. E, Martin R. R., & Tzanetakis I. E. (2011). A discovery 70 years in the making: characterization of the Rose rosette virus. *J Gen Virol. 92*, 1727-1732.

Li W., Cowley A., Uludag M., Gur T., McWilliam H., Squizzato S...Lopez R. (2015). The EMBL-EBI bioinformatics web and programmatic tools framework. *Nucleic Acids Research, 43*, 580-584. doi:10.1093/nar/gkv279

Plaxton, W.C. (1996). The organization and regulation of plant glycolysis. *Annual Review of Plant Physiology and Plant Molecular Biology, 47,* 185–214.

Sirover, M.A. (1999). New insights into an old protein: the functional diversity of mammalian glyceraldehyde-3-phosphate dehydrogenase. *Biochimica et Biophysica Acta, 1432*, 159–184.

Teich, R., Grauvogel, C., & Petersen, J. (2007). Intron distribution in Plantae: 500 million years of stasis during land plant evolution. *Gene, 394*, 96–104.

Windham, M, A. Windham, F. Hale & J. Amrine, Jr. (2014). Observations on rose rosette disease. *American Rose. 3,* 56-62.

Figure Legends

Figure 1.

The Rose Rosette Disease (RRD) presents itself in *Rosa hybrid* cultivar. This figure illustrates an unaffected *Rosa hybrid* cultivar, Rosa Tropicana (top) and another Rosa Tropicana plant infected with the Rose Rosette Virus (bottom). Distinct physical changes are present in roses that have been infected with the virus. These changes include bright red misshapen leaves, distortion, and unusual growth. Four rose plants were selected from the Western Kentucky Botanical Garden and young-growth leaves were taken from two plants not visibly presenting RRD (non-infected) and two plants presenting RRD (virally-infected). Each of the leaves were weighed out individually to 0.079 grams and broken down into 1-mm pieces and disrupted with lysis buffer with dithiothreitol. Cellular debris was removed by centrifugation and each DNA sample was purified in a silica-base column and concentrated with 70 % ethanol, all according to manufacture instructions (Bio-Rad, 2008). These images were taken by the authors of this study.

Figure 2.

GAPC genomic DNA sequences from *Rosa hybrid* samples. Contig sequences of *GAPC* genes presented here were obtained from both two infected and two non-infected plants. The non-infected plant and the infected plant sequences are contig sequence of 6 and 3 overlapping sequence files, respectively. These sequences were isolated by a two-step PCR amplification of genomic DNA from corresponding plants. Exonuclease 1 was added to the first round of PCR products before the nested PCR was done in order to remove any single-stranded DNA. The nested PCR, in this two-step PCR technique, corrects for non-specific products of the PCR reaction, specifically isolating the *GAPC* gene from other GAPDH family genes that could have been amplified via the first reaction. Following amplification of *GAPC*, size-exclusion chromatography was used to purify the PCR-amplified *GAPC* gene using PCR Kleen spin columns according to manufacture instructions (Bio-Rad, 2008). Purified *GAPC* genes were sub-cloned into a pJET1.2 vector utilizing a blunt-end ligation technique, previously described (Bio-Rad, 2008). Purified *GAPC* genes in pJET1.2 vectors were sequenced using primers overlapping both the plasmid and the gene itself: Forward pJET Primer: CGACTCACTATAGGGAGAGCGGC, Reverse pJET Primer: AAGAACATCGATTTTCCATGGCAG, Forward *GAPC* Primer: GGHATTGTTGAGGGTCTNATGAC, and Reverse *GAPC* Primer: CCAGTGGTGCTRGGAATGATGTT. Sequencing was performed by Eurofins MWG/Operon.

The *GAPC* genomic sequences shown are results after processing the raw sequences with iFinch for Educators assembly trim program (Geospiza, Inc.). These trimmed sequences for both the non-infected and the virally-infected plants were then assembled into contig sequences using the CAP3 Sequence Assembly Program (Huang and Madan, 1999) and submitted to GenBank (NCBI). Both contig sequences were analyzed against the reference genomic sequences with BLASTn (NCBI) to confirm identity as the *GAPC* gene.

Figure 3.

GAPDH amino acid sequence analysis of non-infected *Rosa hybrid*. Putative amino acid sequences were obtained by a two-step process. First, putative mRNA coding regions were determined by annotating the genomic sequences using BLASTn against the GenBank mRNA sequence database (NCBI). Then, the putative amino acid sequence was determined via EMBOSS Transeq (Li, et al., 2015) and verified using BLASTp against the protein database (NCBI). The resultant *Rosa hybrid* protein sequence (KT806117) was checked for sequence similarity with PSI-BLAST against another reported *Rosa hybrid* (AEQ49677.1) amino acid sequence and the sequence of *Arabidopsis thaliana* (NP_187062.1). * Represents a non-conserved amino acid substitution in a *Rosa hybrid* amino acid sequence compared to *Arabidopsis thaliana*. __ Represents an amino acid substitution in a *Rosa hybrid* amino acid sequence compared to *Arabidopsis thaliana*.

Figure 4.

GAPDH putative amino acid sequence analysis of virally-infected *Rosa hybrid*. Illustrative representations of two plant cells with GAPDH protein sequences compared for similarity with PSI-BLAST (NCBI). The putative amino acid sequence is modified at 10 locations, 4 of which are non-conserved changes in protein structure. * Represents non-conserved amino acid in the virally-infected *Rosa hybrid* amino acid sequence compared to the non-infected *Rosa hybrid*. __ Represents an amino acid substitution in a virally-infected *Rosa hybrid* amino acid sequence compared to the non-infected *Rosa hybrid*.

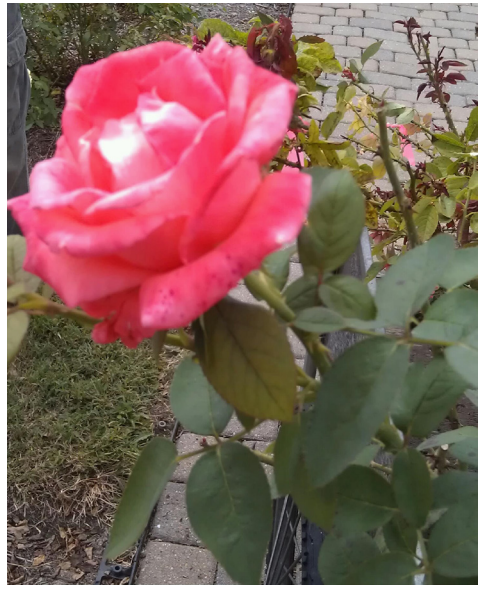

Figure 1

Non-infected *Rosa hybrid*

Plant 1
pJET Reverse Sequence
GAPC Reverse Sequence
Plant 2
pJET Forward Sequence
pJET Reverse Sequence
GAPC Forward Sequence
GAPC Reverse Sequence

Contig Sequence (KT806117)

AGACAAACAGAACATATTCAAACCAGTGCAACAACAATGTCATCACATTCACATTCAAAACAAATAAGGAAGC
AGAACAATTCGTTTACCTGTTGTCACCGATGAAGTCACGTTGACACAACATCGTCTTCGGTGTAACCCAAGATACCCT
TGAGCTTTCCCTCGGACTCCTCCCTCTCATAATGGAGTTAGAGGTTAACCAAAAAAATGCATGGTAATGTTGATGAC
TGTCCCATATTGTAATGGGATAAACGAGAAAACTTGATGCTACTTACTTGATAGCAGCCTTAATCTGGTCATATGTG
GCCTTCTTCTCAAGCCTGACAGTGAGGTCAACGACTGAAACATCAACAGTGGGAACACGGAAGGCCATTCCAGTCAA
TTTGCCATTGAGAGCTGGTAGAACTTTTCCAACAGCCTAAACAAAACACAAAACAGGCAGACAAATTGATCAAGCTA
CCAAAAAAAAGGCTCATTAAAATACAAAGAAAGACAATGATATTTAAAAGGGGCATTGTCACATAAGTTATGAAATA
CCTTGGCAGCTCCAGTCGTGCTAGGAATGATGTTGAATGAAGCAGCACGTCCGCCTCTCCAGTCCTTACTTGATGGG
CCATCAACAGTTTTCTGTGTGGCCGAATCAAAGACAGAAAAATTCTTTGCTTAGTTTCAAATTATAAAATCTCAACATT
CATACAACGAAAAAAGCCTGAAGCTAAAACTCACCTGTGATGGAGTGAACAGTGGTCATAAGGCCCTCAACAATTCC
GAATCTATCATTGATAACCTATAGACCAAGCAAACAATACGATACATAATAAGCAACAACCATTGCATCAGGGTTGCG
ATTCAAAACACAACACAAACAGTAGCACATTGACACACCTTGGCAAGGGGAGCAAGACAGTTGGTCGTGCAGCTGG
CATTAGACACAATGGTAAGGTCTGACTTGTATTCCTTCTCATTGACACCCACAACACAAACATGGGAGCATCCTTACTT
GGGGCAGAAATGACAACCTTTTTGGCACCACCCTGGTCGCACAAACAATACAAGTTAAATCCAGAATGACCATCAAA
GTAGATATAGAACTATCCAAATGA

Virally-infected *Rosa hybrid*

Plant 1
pJET Forward Sequence
pJET Reverse Sequence
Plant 2
pJET Forward Sequence

Contig Sequence (KT806118)

CAGTGAGACTTGAGAAGAAGGCTACTTATGATGAGATTAAAAATGCTATCAAGTAAGGATCAGGGCATTCACTATGA
TTTCTTTCTTTATATTGGAATTTAAAACAAACTAAACTACTAAACAGTTGATTGGTTTTATATTTTTGGTGCATCTAT
ATGCCTATTTTTCAGTTTGATGATTTAACTTACTTGTCTTTCAAGATCTGCTCACCTCTCCTTATCTGAGTTTCGTT
TTGGTTGTGAATTTCAGGGAAGAATCTGAGGGTAACCTAAAGGGAATCCTCGGATATATGGACGATGATTTAGTGTC
CACTTGACTTTGTGGATAACCACAGGTGAGAATGTTAATGATGATTGCTGACTCTTGCCTGTTGGCCTCTGTAAAACCT
GCAGCACTTGGCTATGTGTCTGCCCTATTCTTTACTTTGGTTAATTGATTTAGGCTCTGAGGGTGTGCATACATGTT
ATGCTGCAGGTCAAGCATCTTTGATGCCAAGGC

Figure 2

A. thaliana (NP_187062.1) STGVFTDKDKAAAHLKGGAKKVVISAPSKDAPMFVVGVNEHEYKSDLDIVSNASCTTNCLAPLAKVINDR
R. hybrid (AEQ49677.1) STGVFTDKDKAATHLKGGAKKVVISAPSKDAPMFVVGVNEKEYKSDLTIVSNASCTTNCLAPLAKVINDR
R. hybrid (KT806117) GGAKKVVISAPSKDAPMFVVGVNEKEYKSDLTIVSNASCTTNCLAPLAKVINDR

FGIVEGLMTTVHSITATQKTVDGPSMKDWRGGRAASFNIIPSSTGAAKAVGKVLPALNGKLTGMSFR
FGIVEGLMTTVHSITATQKTVDGPSSKDWRGGRAASFNIIPSSTGAAKAVGKVLPALNGKLTGMAFR
FGIVEGLMTTVHSITATQKTVDGPSSKDWRGGRAASFNIIPSSTGAAKAVGKVLPALNGKLTGMAAFR

VPTVDVSVVDLTVRLEKAATYDEIKKAIKEESEGKLKGILGYTEDDVVSTDFVGDNRSSIFDAKA
VPTVDVSVVDLTVRLEKKATYDQIKAAIKEESEGKLKGILGYTEDDVVSTDFIGDNRSSIFDAKA
VPTVDVSVVDLTVRLEKKATYDQIKAAIKEESEGKLKGILGYTEDDVVSTDFIGDNR

Figure 3

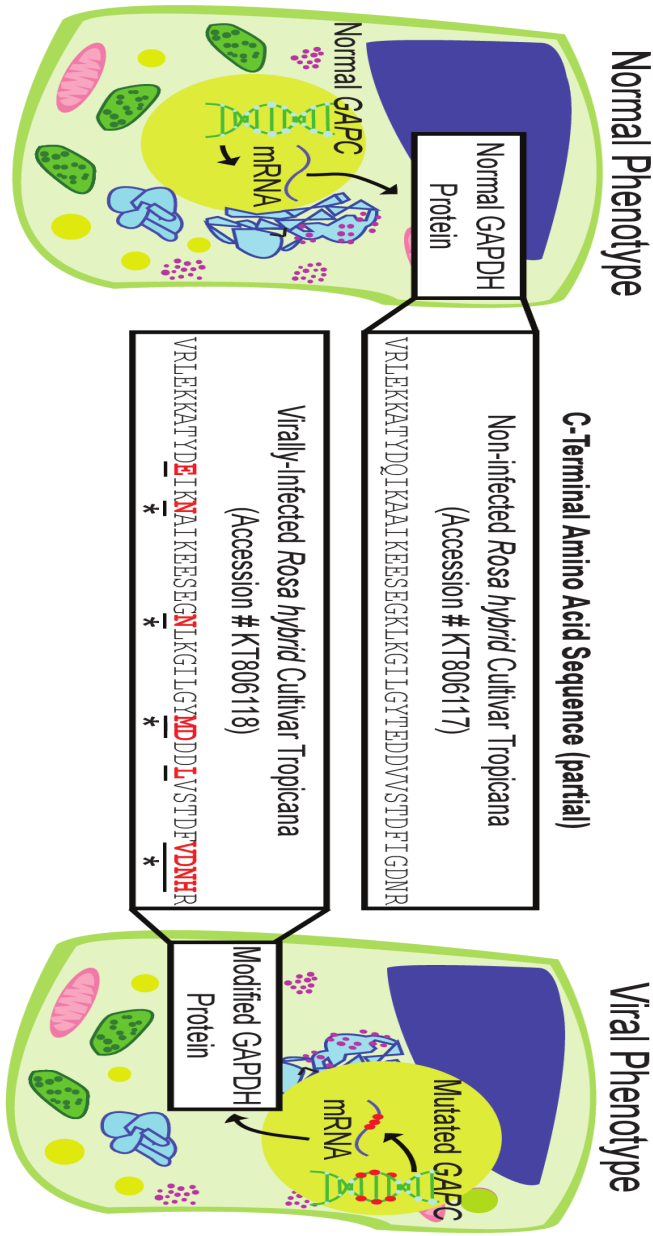
Normal Phenotype
Normal GAPC
mRNA
Normal GAPDH Protein
C-Terminal Amino Acid Sequence (partial)
Non-infected Rosa hybrid Cultivar Tropicana
(Accession # KT806117)
VRLEKKATYDQIKAALKEESEGKLKGILGYTEDDVVSTDFIGDNR
Virally-Infected Rosa hybrid Cultivar Tropicana
(Accession # KT806118)
VRLEKKATYDEIKNALKEESEGNLKGILGYMDDDLVSTDFVDNHR
Modified GAPDH Protein
mRNA
Mutated GAPC
Viral Phenotype

Figure 4