

Tools and computational resources for the design of CRISPR/Cas9 sgRNA for *NPR3* gene knockout in sour orange (*Citrus aurantium* L.)

Mao Yupanqui-Celestino^{1*}, Lizet Karina Azañero-Huaynacari¹, Carlos Roberto Pesantes-Rojas², Miguel Angel Aguilar-Luna-Victoria³, Desiderio Elías Cotos-Durán¹, Linder Ramírez-Viena⁴

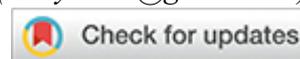
¹Escuela Profesional de Biología con mención en Biotecnología, Facultad de Ciencias, Universidad Nacional José Faustino Sánchez Carrión, Huacho, Perú.

²Escuela Profesional de Matemática Aplicada, Facultad de Ciencias, Universidad Nacional José Faustino Sánchez Carrión, Huacho, Perú.

³Escuela Profesional de Estadística e Informática, Facultad de Ciencias, Universidad Nacional José Faustino Sánchez Carrión, Huacho, Perú.

⁴Departamento de Ciencias Naturales, Facultad de Educación, Universidad Nacional José Faustino Sánchez Carrión, Huacho, Perú.

*Correspondence: Mao Yupanqui-Celestino (Cusy2000@gmail.com)



ABSTRACT

Citrus fruits are the most nutritious foods widely used in flavoring, beverages, and medicines due to their outstanding curative effects. Sour orange (*Citrus aurantium* L.) is the predominant rootstock in most citrus growing areas due to its good agronomic attributes, such as high quality, yield, and tolerance to various pathogens. However, the citrus tristeza virus (CTV) is the leading epidemic agent of sour and sweet orange. This study aimed to design in silico guide RNA (sgRNA) for CRISPR/Cas9-mediated inactivation of the Non-expression of Pathogenesis-Related genes 3 (*NPR3*) in sour orange (*CaNPR3*). The protein sequence of the *CaNPR3* gene is 584 amino acid residues long. The amino acid sequence of the *CaNPR3* gene was compared with the homologous sequences of other nearby vegetative species, showing a close similarity with *Citrus sinensis* and *Citrus Clementina* with 100% and 97.27%, respectively. CRISPR RGEN Tools provided 61 results for exon two of the *CaNPR3* gene, filtering to 19 sequences and selecting four sgRNA sequences for genetic editing, which were: sgRNA 1 (5'-CATCAGGAAAAGACTTGAGT-3'), sgRNA 2 (5'-AGAAC-CTCAGACAACACACCTT-3'), sgRNA 3 (5'-CATCAGATTTGACCCTGGAT-3') and sgRNA 4 (5'-

TTCTGGAGGGAGGGAGAGAAATGAGGAGG-3'). The predicted secondary structures of the four selected sgRNAs present efficient structures for gene editing of the target gene, allowing it to recognize, interact with Cas9 protein, and edit the target region.

Keywords: Gene editing, guide RNA, *CaNPR3*, in silico.

INTRODUCTION

Citrus is one of the most abundant fruit crops and the essential economic mainstay for many domestic and foreign farmers^{1,2}. In the previous cycle, sour orange (*Citrus aurantium* L.) was the predominant rootstock in

most citrus growing areas due to its good agronomic attributes such as high quality, yield, tolerance to various pathogens and resistance to abiotic stresses³⁻⁵. Sour oranges belong to the Rutaceae family, which are medium-sized cosmopolitan angiosperms. Recently, the fruit industry has faced unavoidable problems due to climate change and global warming, causing an increase in epidemics of new and invasive pathogens in agricultural fields⁶⁻⁹.

Citrus tristeza virus (CTV) is the main epidemic agent causing the worldwide loss of nearly 100 million trees of sweet orange (*Citrus sinensis* (L.) Osb.), grapefruit (*C. paradise* Macf.), mandarin (*C. reticulata* Blanco), and lime (*C. aurantifolia* (Christ.) Swing.) propagated on sour orange (*C. aurantium* L.)¹⁰. CTV causes collapse, obliteration, and necrosis of the crypt tubes and accompanying cells near the bud junction and induces the appearance and excessive development of non-functional phloem, which decreases water and mineral transport in the plant^{10,11}. It has been shown that silencing of the Non-expression of Pathogenesis-Related genes 3/4 gene, associated with these defense pathways, improved virus propagation and accumulation in sour orange plants compared to non-silenced controls¹².

The non-expression of pathogenesis-related genes (*NPR*) is a genuine transcription cofactor in the salicylic acid (SA) signal transduction pathway and plays critical regulatory roles in plant immunity¹³. Within the *NPR* family, the most representative gene is *NPR1*, containing a BTB/POZ domain and an anchor protein repeat domain. The genome of *Arabidopsis thaliana* contains six genes related to the *NPR* gene family¹⁴, *NPR3* and *NPR4* have been identified as homologous genes of *NPR1*¹⁵. The *NPR3* gene is involved in the transduction of the SA signal transduction pathway and also functions as a negative regulator of the basal defense response¹⁶. In sour grapes, it has been shown that gene silencing decreases CTV accumulation and increases growth concerning control.

The CRISPR/Cas (Clustered Regularly Interspaced Short Palindromic Repeats-CRISPR-associated) system has become one of the most widely used systems due to its low cost, easy adaptation and high efficiency during gene editing in plants^{17,18}. The CRISPR Cas9 nuclease is directed by a short single guide RNA (sgRNA) that recognizes the target DNA within the genome. The sgRNA combines CRISPR RNA (crRNA) and the crRNA transactivator (tracrRNA). The CRISPR/Cas system requires a well-defined adjacent short protospacer motif (PAM) located immediately downstream of the protospacer on the non-target DNA strand¹⁹⁻²³. In addition, Double-Strand Breaks (DSBs) of target DNA occur three nucleotides upstream of the PAM motif by incorporation of the HNH and RuvC domains belonging to the Cas9 endonuclease^{24,25}, DSB is repaired by homology-directed repair (HDR) pathway or non-homologous end joining (NHEJ) pathways in the genome²⁶. This repair pathway has been exploited as NHEJ causes indel mutation-producing Knockout (KO) genes. HDR causes point mutation accompanied by an exogenous gene of interest, producing a Knock-In (KI) gene.

Therefore, in recent years, *in silico* designs have been performed because they do not require time-consuming experimental procedures, which are based on practical and optimized molecular biology tools adapted to the issue of the organism that has been previously developed, with the ability to clone, predict protein structure, analyze gene expression, predict mutagenesis sites directed by CRISPR/Cas. The design of sgRNA is determinant in the specificity of targeting and the efficiency of excision or insertion because if off-target Cas9 activity is present (off-target), it can lead to unexpected effects²⁶⁻²⁸; cases of off-target occur because similar sgRNA sequences exist in other parts of the genome, this type of designs in computational machines is called *in silico*²⁹. As a result, current researchers are forced to develop new tolerant varieties and rootstocks with the support of genetic engineering, allowing industries to gain experience to face these and similar future problems³⁰.

The present study aims to design in silico sgRNA editing sgRNAs for *NPR3* gene knockout for positive regulation against CTV in sour orange (*Citrus aurantium* L.) mediated by CRISPR/Cas9, being a previous step in the genetic improvement of orange against these pathogens, using computational tools (bioinformatics).

MATERIALS AND METHODS

Identification of amino acid sequences of CaNPR3

The amino acid sequence of the *Citrus aurantium NPR3* gene (*CaNPR3*) was based on the reference sequence GCF_022201045.2. It was compared with other NPR amino acid sequences in plant species with a higher similarity percentage than 77%. The sequences were obtained from NCBI (<http://www.ncbi.nlm.nih.gov/>). The sequences were aligned using the Bioedit program, and a phylogenetic tree was formed using the Neighbor-Joining method with 1000 Bootstrap analysis repeats with the MEGA 11 program. The determination of the motifs and conserved regions of the aligned amino acid sequences was performed using the virtual platforms The MEME Suite (<https://meme-suite.org/meme/>) and WebLogo (<https://weblogo.berkeley.edu/logo.cgi>).

In silico design of sgRNA for CaNPR3 gene

The design of the sgRNAs for the target region of the *CaNPR3* gene was performed using the online CRISPR RGEN Tools (<http://www.rgenome.net/>) on the sour orange genome with the Cas-Designer function in order to perform frameshift recognition and mutation on the target sequence. In addition, the LOC102621158 gene was selected with the accession XM_006468378.4 mRNA, which is located NC_023047.1 (8114278.8117323) furthermore the gene is referenced by Gómez-Muñoz et al.¹². The mRNA has 4 exons, in this work we will work with exon 2 of *CsNPR3* from accession XM_006468378.4 exons. The RGEN-eligible target sequences (5' to 3') obtained were evaluated with the criteria of GC content between 40% to 60% and out-of-frame score greater than 70.

In silico prediction of sgRNA structure

The selected sgRNA structure was predicted using the online tool The Vienna RNA Web Services (<http://rna.tbi.univie.ac.at/#webservices>) in the RNAfold web server function. The prediction focused on the selected mRNA secondary structures complemented by an RNA scaffold set up for CRISPR/Cas9-mediated gene editing methodology. The selection of sgRNAs based on their secondary structures was established by forming two hairpins at the upper ends and a RAR stem-loop. Further microhomology prediction of the selected gRNAs with the Microhomology function (CRISPR RGEN Tools) allows prediction of mutation patterns caused by the NHEJ or MMEJ pathway and assessment of in- or out-of-frame target deletion.

RESULTS AND DISCUSSION

Identification of amino acid sequences of the CaNPR3 gene

The sequence of *CaNPR3* shows 100% similarity to *Citrus sinensis*, followed by *Citrus Clementina* with 97.27% (Table 1). Lengths of 584 to 590 amino acid residues were found in the compared sequences. All *CaNPR3*-like sequences contain an ankyrin repeat domain and a BTB/POZ dominium (Figure 1), indicating high levels of functional conservation in the *NPR*-like family^{31,32}. However, as the list of *NPR*-like proteins in different plant species increases, so does the complexity and variability of their function^{33,34}. Phylogenetic tree construction is a crucial method to analyze the functionality of *CaNPR3*-like genes; in this study,

based on the results of phylogenetic tree classification, *NPR3* protein sequences from various species were classified into two clades. Among these, clade II exhibited the most significant number of members, with 9, and clade I consisted of 5 members (Figure 1). From this, it can be observed that *Citrus sinensis* and *Citrus clementina* were assigned to clade I, of which their proteins are reported to function as positive and negative regulators of systemic acquired resistance (SAR) for each species, respectively³⁵. In addition, conserved motifs were identified in *CaNPR3*-like proteins (Figure 2). The protein sequences of the species analyzed share most of the motifs. However, all the sequences had ten motifs, except *Citrus Clementina*, which had motif 8 exclusively. The presence of shared conserved motifs within the *CaNPR3*-like gene family supports the classification of these genes as a multigene family³⁵. The conserved structural domain analysis reveals that all *CaNPR3*-like proteins (Figure 2) contain an N-terminal BTB/POZ structural domain and a central ankyrin repeat domain (Figure 3). In addition, genes on both the first and second branches of the evolutionary tree exhibit the C-terminal *CaNPR3*-like region, which is known to be a crucial component of the *NPR*-like gene family and assumes a key role^{36,37}.

Nº	Accession	Scientific name	Cover page of the query (%)	Percentage of identity (%)
1	XP_006468441.2	<i>Citrus sinensis</i>	100%	100.00
2	XP_024047266.1	<i>Citrus clementina</i>	100%	97.27
3	KAJ6294437.1	<i>Salix suchowensis</i>	100%	78.88
4	KAJ6761117.1	<i>Salix purpurea</i>	100%	78.53
5	KAJ6410587.1	<i>Salix udensis</i>	100%	78.53
6	XP_034887663.1	<i>Populus alba</i>	100%	79.90
7	XP_031280798.1	<i>Pistacia vera</i>	100%	82.00
8	KAJ0081498.1	<i>Pistacia atlantica</i>	100%	82.00
9	XP_015575545.1	<i>Ricinus communis</i>	98%	77.72
10	KAB1222144.1	<i>Morella rubra</i>	99%	77.00
11	NP_001238674.1	<i>Glycine max</i>	95%	77.30
12	KHN14362.1	<i>Glycine soja</i>	95%	77.13
13	XP_007024872.1	<i>Theobroma cacao</i>	96%	77.66
14	XP_044496684.1	<i>Mangifera indica</i>	100%	80.27

Table 1: Accessions of plant species with the highest sequence similarity to *CaNPR3*.

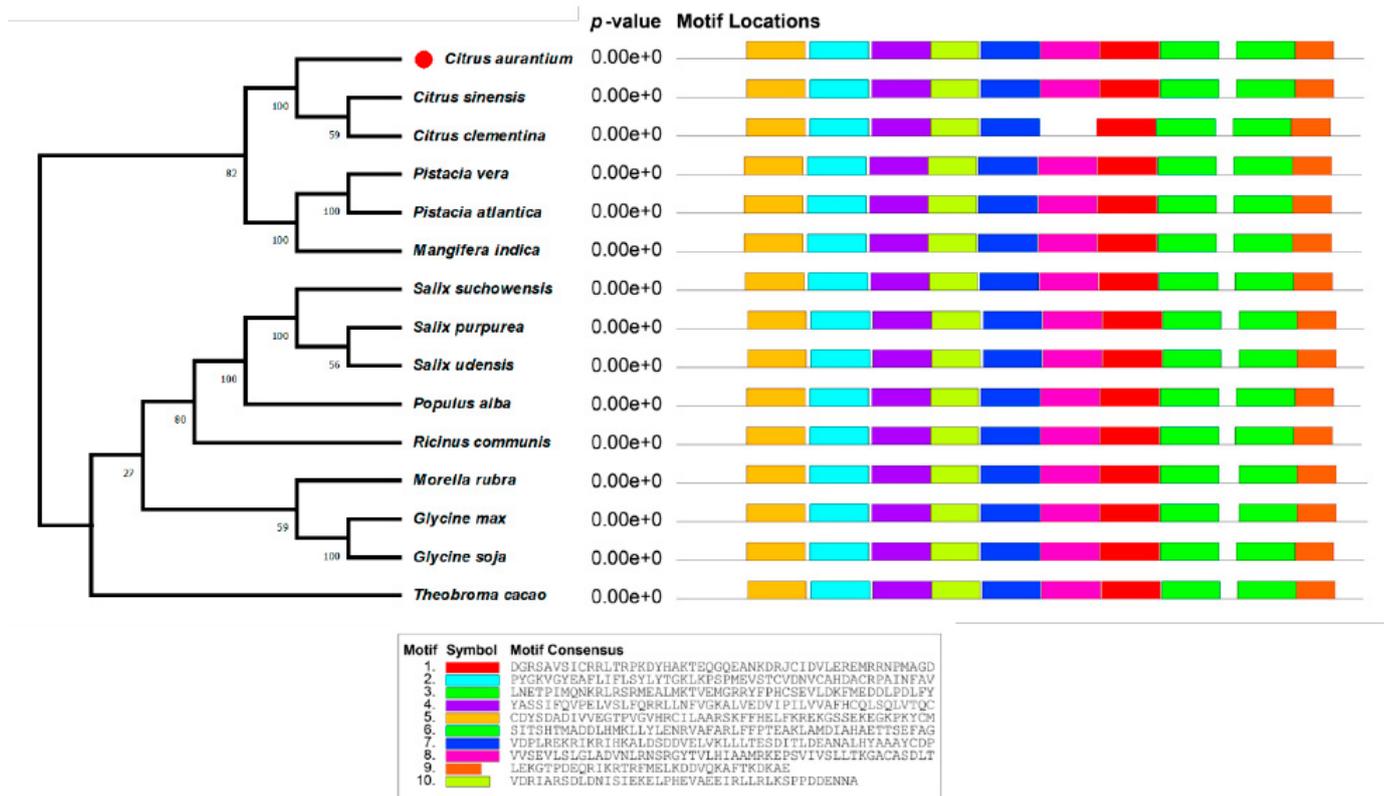


Figure 1: Phylogenetic tree analysis of neighbor-joining and conserved motifs of the *NPR3* protein from sour orange and nearby species.



Figure 2: Sequence comparison of conserved *NPR3* motif sequences of sour orange and nearby species.

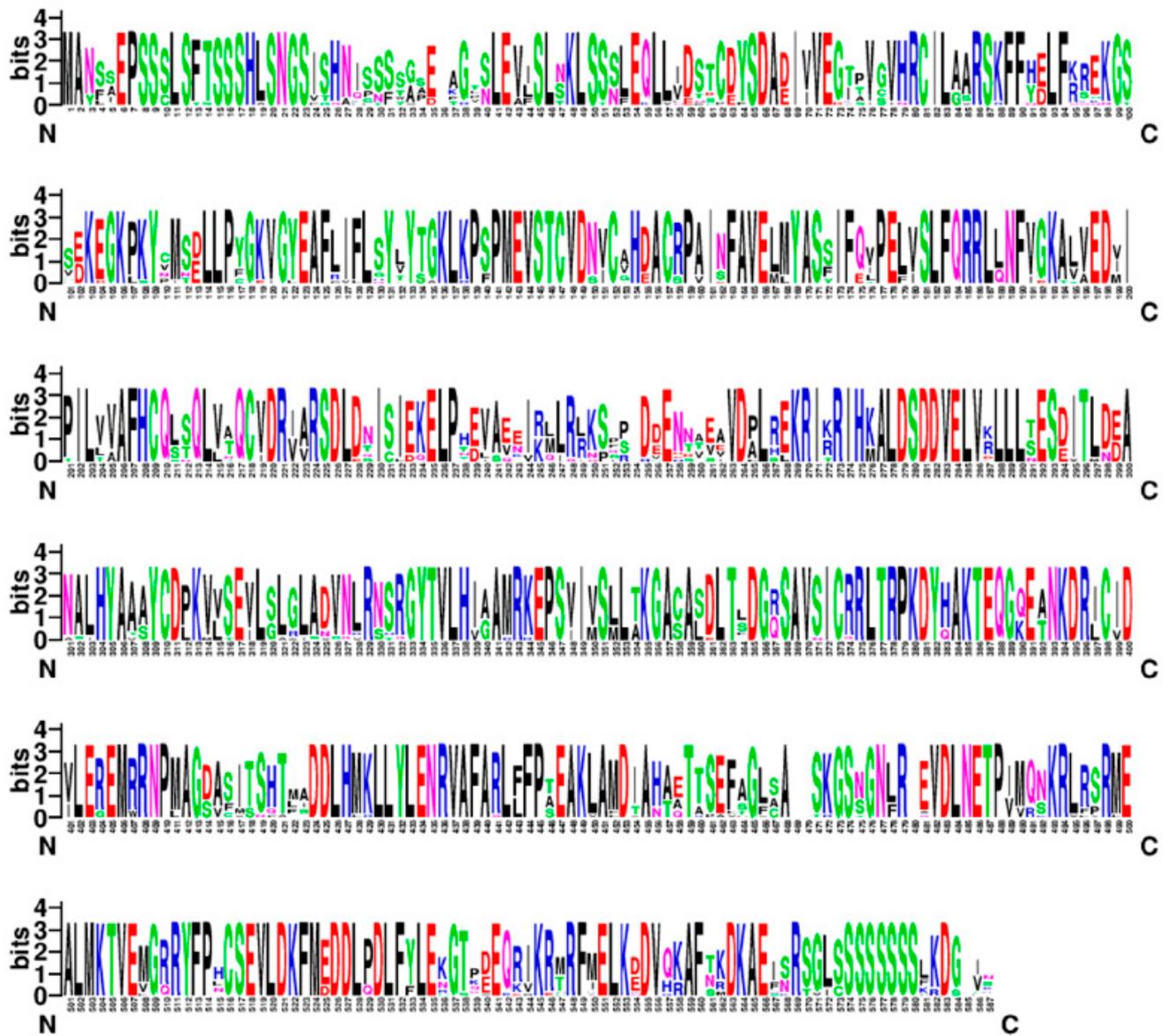


Figure 3: Conserved regions of the *NPR3* protein sequence of sour orange and related species.

In silico design of sgRNA for *CaNPR3* gene

In silico design, they identified 61 candidate sgRNA sequences for *Citrus aurantium* gene editing in the *CaNPR3* exon 2 region. By filtering, 19 candidate sgRNA sequences or RGENs are decreased (Table 2), finally reduced to 16 sgRNA sequences by employing additional evaluations on Mismatch "1" values in Mismatch 0, which confirms the target locus within the target region and proceeds by adding filter "0" in Mismatch 1 and 2 number of targets found outside or inside the target region.

The guanine (G) and cytosine (C) content of sgRNA was based on Liu et al.³⁸, who mentions that sgRNA efficacy is positively influenced by increasing GC content, but at the same time, increasing GC content decreases cleavage activity significantly. We established that the GC content of gRNA for this work should present a 40-60% range. It was estimated that the 16 sgRNA for the second exon of the *CaNPR3* gene has a high reliability because the GC values are between 30 to 80%, evidenced in the selection of highly efficient sgRNA for plant genome editing mediated by CRISPR/Cas9³⁹. However, if exon 2 *CaNPR3* is an inefficient region, sgRNAs are found to be between 41 and 55% tested in evaluating the efficacy of CRISPR/Cas9 sgRNAs targeting inefficient regions in *Arabidopsis thaliana*⁴⁰.

Out-of-frame scores are essential because choosing high-scoring sites increases the probability of obtaining permanent mutant clones. In addition, the out-of-frame score mentioned by Bae et al.⁴¹ correlates with the probability that sgRNA-induced mutations disrupt the open reading frame (ORF). All 16 sgRNAs for exon 2 *CaNPR3* possess an out-of-frame score ≥ 70.2 , and the gRNA scores are above the minimum score recommended (Out-of-frame ≥ 66) to create knockouts.

RGEN Target (5' to 3')	Position	Cleavage Position	Direction	GC Contents	Out-of-frame Score	Mismatches		
						0	1	2
TGGAAGGCAGCCAAAAGGATTGG	49	7.2	-	50	68.4	1	0	3
GACAATGGAAGGCAGCCAA-AAGG	54	7.8	-	50	71.2	1	0	0
CATCAGGAAAAGACTTGAGTCGG	189	25.8	-	40	83.9	1	0	0
TGCAGCGTACTGTGATCCCAAGG	363	50.3	+	55	80.5	1	0	0
TGATCCCAAGGTGTTGTCTGAGG	375	51.9	+	50	72.4	1	0	0
AGAACCTCAGACAACACCTTGGG	379	51	-	45	76.7	1	0	0
GAGAACCTCAGACAACACCTTGG	380	51.1	-	50	73.5	1	0	0
TTGTCTGAGGTTCTCAGCCTCGG	388	53.7	+	50	66.6	1	0	0
TATAACCCCGAGAACTTCTAAGG	426	57.2	-	40	71.5	1	0	0
ACATATTGGTGCAATGCGCAAGG	456	62.7	+	45	66.7	1	0	0
TTGTGCATCAGATTTGACCCTGG	516	70.7	+	45	74.3	1	0	0
GCATCAGATTTGACCCTGGATGG	520	71.2	+	50	74.9	1	0	0
CATCAGATTTGACCCTGGATGGG	521	71.3	+	45	77.6	1	0	0
GCATCTGTAGAAGATTGACAAGG	557	76.1	+	40	76.5	1	0	0
GACAAGGCCGAAAGATTATCAGG	573	78.2	+	45	72.6	1	0	0
ATGTTCTGGAGGGAGAAATGAGG	644	87.6	+	45	79.6	1	0	0
TTCTGGAGGGAGAAATGAGGAGG	647	88	+	50	79.5	1	0	0
ATGAGGAGGAATCCGATGGCTGG	661	89.9	+	55	81.5	1	0	0
ATAAAGGCATCACCAGCCATCGG	673	90	-	45	86.7	1	0	0

Table 2: Accessions of plant species with the highest sequence similarity to *CaNPR3*.

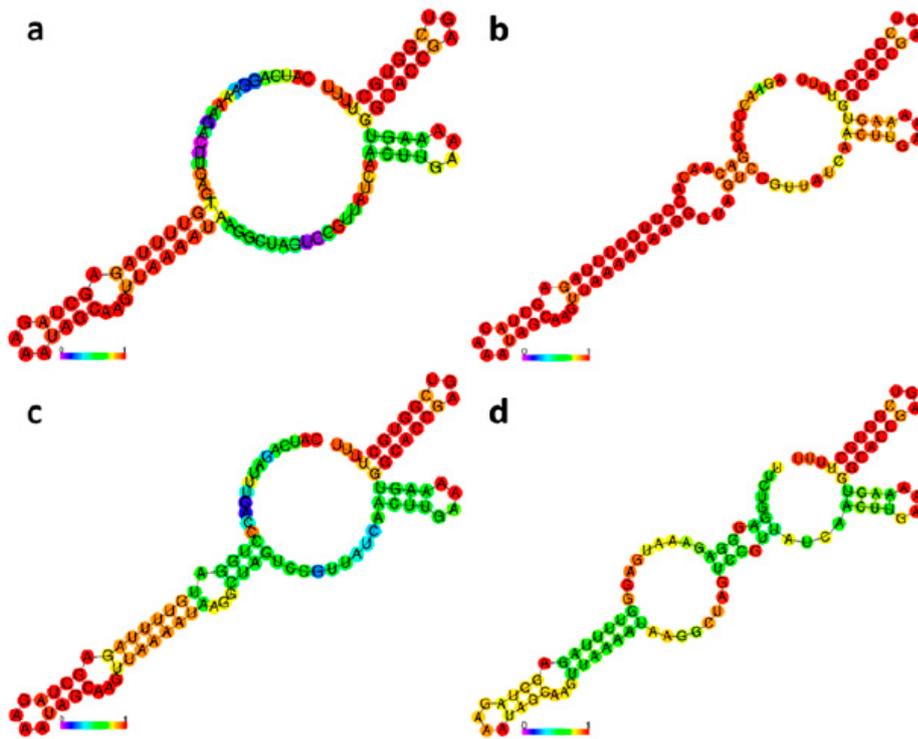


Figure 4: Secondary structures of sgRNAs for CaNPR3 gene editing: (a) sgRNA 1, (b) sgRNA 2, (c) sgRNA 3 and (d) sgRNA 4.

Objective sequence	Position	Genome location	Microhomology Score	Out-of-frame Score	Deletion Length
sgRNA 1(-): CATCAGGAAAAGACTT-GAGT	188	NC_023047.1	3908.5	83.86849175	16
Predicted Patterns				Microhomology	Pattern Score
WT: AGTTGCGGAGGAAATT-AGAATGCTCCGACTCAAGTCTTTTCCTGATGATGAGAACAACACTGC Mutation 1: AGTTGCGGAGGAAATTAGAATGCTCC-----TGATGATGA-GAACAACACTGC				TCC	224.5
sgRNA 2(-): AGAACCTCAGACAACA-CCTT	378	NC_023047.1	4009.2	76.71106455	11
Predicted Patterns				Microhomology	Pattern Score
WT: CATTATGCTGCAGCGTACTGTGATCCCAAGGTGTT-GTCTGAGGTTCTCAGCCTCGGATTA Mutation 2: CATTATGCTGCAGCGTACTGTG-----TTGTCTGAGGTTCTCAGCCTCGGATTA				GTG	288.5
sgRNA 3(+): CATCAGATTTGAC-CCTGGAT	520	NC_023047.1	4095.3	77.57429248	4
Predicted Patterns				Microhomology	Pattern Score
WT: AAAGGAGCTTGTGCATCAGATTTGACCCTGGATGGGCGAAGTGCCGTCAG-CATCTGTAGA Mutation 3: AAAGGAGCTTGTGCATCAGATTTGACCCTG----GGCGAAGTGCCGTCAGCATCTGTAGA				TG	245.7
sgRNA 4(+): TTCTGGAGGGAGAA-ATGAGG	646	NC_023047.1	5677.3	79.47439804	14
Predicted Patterns				Microhomology	Pattern Score
ATATGCATTGATGTTCTGGAGGGAGAAATGAGGAGGAA-TCCGATGGCTGGTGTATGCCTTT Mutation 4: ATATGCATTGATGTTCTGGAGG-----AATCCGATGGCTGGTGTATGCCTTT				GGAGG	447.3

Table 3: Calculation of microhomology-associated scores for 4 sgRNA for *CaNPR3*.

The 16 selected sgRNAs pose as a 1-0-0 filter in the Mismatches. This indicates that each sgRNA only has a single target site within the sour orange genome, in addition to not possessing off-target hits that differ by several nucleotides from the on-target sites, evidencing that the 16 designed sgRNAs are site-specific and probably efficiently in gene editing. Cho et al.²⁷ mention that RNA-guided endonucleases (RGENs) can distinguish on-target from off-target sites with mismatches starting from two bases but not those with a single-base mismatch.

In silico* secondary sgRNA structures for *CaNPR3

The secondary structures of the four selected sgRNAs stably present stem loops 2 and 3 at the upper end and stem-loop RAR stably possessing between 1 to 3 bubbles (Figure 4). From the structural evaluation, four sgRNA sequences were selected, the sequences sgRNA 1 (5'-CATCAGGAAAAGACTT-GAGT-3'), sgRNA 2 (5'-AGAACCTCAGACAACAACACCTT-3'), sgRNA 3 (5'-CATCAGATTTGACCCTGGAT-3') and sgRNA 4 (5'-TTCTCTGGAGGGAGGGAGGGAGAAATGAGGAGG-3'). Thus, even the evaluation of mutation patterns produced by the 4 sgRNAs identified that the four predicted patterns of mutations produced in target sites P1x, Px2, Px3 and Px4 were caused by the MMEJ repair pathway and located in

NC_023047.1 (Table 3). The predicted target site patterns were selected by their pattern score and microhomology score, which are related to silico-predicted deletion patterns resulting from microhomology-associated DNA repair and deletion length⁴¹.

Determining secondary structures is important because it allows the identification of sgRNA sequences that may present structural variations that impair the ability to recognize and edit the target region^{42,43}. It is known that a sgRNA functions by interacting its secondary structure with the Cas9 protein⁴⁴. Ma et al.⁴⁵ mentioned that the secondary structure of sgRNAs can hinder the efficiency of gene editing. The secondary structure of the four sgRNAs (Figure 4) presents the RAR stem-loop region, which triggers the processing of the pre-crRNA precursor by the RNase III enzyme and subsequently activates the cleavage of crRNA-directed DNA by Cas9, as the four sgRNAs present second and third stem-loops which promote the formation of stable complexes and improve in vivo activity^{44,46}. However, the structures of the 4 sgRNAs do not present stem loop 1, but Liang et al.³⁹ have shown that it is not related to editing efficiency in plants.

CONCLUSIONS

Four sgRNAs were designed in silico for CRISPR/Cas9-mediated inactivation of the *CaPRN3* gene susceptible to the citrus tristeza virus. The design was achieved by comparing the *CaPRN3* gene to the *CsPRN3* gene. The amino acid sequence of PRN3 from sweet orange was compared with sour orange and more plant species and showed close similarity with *Citrus sinensis* and *Citrus Clementina* with 100% and 97.27%, respectively. The sgRNAs found have excellent GC and out-of-frame content values and provide high efficiency in gene editing. The predicted secondary structures of the four selected sgRNAs show reliable structures in gene editing with CRISPR/Cas9. The sgRNAs were designed using silico computational tools, which are crucial in predicting the accuracy and exhibiting the success of gene editing on the target target.

Author Contributions: Conceptualization, MYC; methodology, MYC, LKAH; software, MYC and LKAH; validation, DCD, CRPR and LRV; formal analysis, DCD, CRPR, MAALV and LRV; research, MYC and LKAH; data curation, MYC; writing the original draft, MYC and LKAH; writing, revising and editing, DCD, CRPR, MAALV and LRV; supervision, LRV; All authors have read and accepted the published version of the manuscript.

Conflicts of interest: The authors declare no conflicts of interest.

REFERENCES

1. Dala-Paula, B.M.; Plotto, A.; Bai, J.; Manthey, J.A.; Baldwin, E. A.; Ferrarezi, R. S.; Gloria, M.B. Effect of Huanglongbing or Greening Disease on Orange Juice Quality, a Review. *Frontiers in Plant Science* **2019**.
2. Jia, X.; Jiang, X.; Li, Z.; Mu, J.; Wang, Y.; Niu, Y. Application of Deep Learning in Image Recognition of Citrus Pests. *Agriculture* **2023**, 13(5).
3. Acosta-Pérez, J.A.; Ortiz-García, C.F.; Zaldívar-Cruz, J.M.; Rodríguez-Cuevas, M.; Bautista-Muñoz, C.C.; Cruz Castillo-Aguilar, C. de la. Identificación del agente causal e importancia de la gomosis en la zona citrícola de Huimanguillo, Tabasco, México. *Universidad y ciencia* **2012**, 28(3), 245-258.

4. Ghorbel, R.; Domínguez, A.; Navarro, L.; Peña, L. High-efficiency genetic transformation of sour orange (*Citrus aurantium*) and production of transgenic trees containing the coat protein gene of citrus tristeza virus. *Tree Physiology* **2000**, 20(17), 1183-1189.
5. Hernández-Amasifuen, A.D.; Pineda-Lázaro, A.J.; Díaz-Pillasca, H. B. *In vitro* micropropagation of sour orange (*Citrus aurantium* L.) from nodal segments. *Bionatura* **2021**, 6(4), 2216-2221.
6. Teixeira, D.C.; Saillard, C.; Couture, C.; Martins, E.C.; Wulff, N.A.; Eveillard-Jagoueix, S.; Yamamoto, P.T.; Ayres, A.J.; Bové, J.M. Distribution and quantification of *Candidatus Liberibacter americanus*, agent of huanglongbing disease of citrus in São Paulo State, Brazil, in leaves of an affected sweet orange tree as determined by PCR. *Molecular and Cellular Probes* **2008**, 22(3), 139-150.
7. Bassanezi, R.B.; Montesino, L. H.; Stuchi, E.S. Effects of huanglongbing on fruit quality of sweet orange cultivars in Brazil. *European Journal of Plant Pathology* **2009**, 125(4), 565-572.
8. Endo, T.; Fujii, H.; Omura, M.; Shimada, T. Fast-track breeding system to introduce CTV resistance of trifoliolate orange into citrus germplasm, by integrating early flowering transgenic plants with marker-assisted selection. *BMC Plant Biology* **2020**, 20, 224.
9. Díaz-Pillasca, H.B.; Hernández-Amasifuen, A.D.; Machahua, M.; Pineda-Lázaro, A.J.; Argüelles-Curaca, A., & Lugo, B. (2021). Código de barras de ADN de tres especies de árboles frutales con potencial económico del valle de Huaura, Lima, Perú. *Bionatura* **2021**, 6(3), 1992-2000.
10. Moreno, P.; Ambrós, S.; Albiach-Martí, M.R.; Guerri, J.; Peña, L. Citrus tristeza virus: A pathogen that changed the course of the citrus industry. *Molecular Plant Pathology* **2008**, 9(2), 251-268.
11. Da-Silva, M.; Germano, S.; Duarte, A.; Pinto, P.; Marques, N.T. Callose synthase and xyloglucan endotransglucosylase gene expression over time in *Citrus × clementina* and *Citrus × sinensis* infected with citrus tristeza virus. *Phytoparasitica* **2023**.
12. Gómez-Muñoz, N.; Velázquez, K.; Vives, M.C.; Ruiz-Ruiz, S.; Pina, J.A.; Flores, R.; Moreno, P.; Guerri, J. The resistance of sour orange to the Citrus tristeza virus is mediated by the salicylic acid and RNA silencing defense pathways. *Molecular Plant Pathology* **2016**, 18(9), 1253-1266.
13. Wang, P.; Zhao, Z.; Zhang, Z.; Cai, Z.; Liao, J.; Tan, Q.; Xiang, M.; Chang, L.; Xu, D.; Tian, Q.; Wang, D. Genome-wide identification and analysis of NPR family genes in *Brassica juncea* var. *Tumida*. *Gene* **2021**, 769, 145210.
14. Fu, Z.Q.; Yan, S.; Saleh, A.; Wang, W.; Ruble, J.; Oka, N.; Mohan, R.; Spoel, S.H.; Tada, Y.; Zheng, N.; Dong, X. NPR3 and NPR4 are receptors for the immune signal salicylic acid in plants. *Nature* **2012**, 486(7402).

15. Canet, J.V.; Dobón, A.; Roig, A.; Tornero, P. Structure-function analysis of npr1 alleles in *Arabidopsis* reveals a role for its paralogs in the perception of salicylic acid. *Plant, Cell & Environment* **2010**, 33(11), 1911-1922.
16. Shi, Z.; Maximova, S.; Liu, Y.; Verica, J.; Gultinan, M.J. The Salicylic Acid Receptor NPR3 Is a Negative Regulator of the Transcriptional Defense Response during Early Flower Development in *Arabidopsis*. *Molecular Plant* **2013**, 6(3), 802-816.
17. Yin, K.; Gao, C.; Qiu, J.L. Progress and prospects in plant genome editing. *Nature Plants* **2017**, 3(8).
18. Jung, C.; Capistrano-Gossmann, G.; Braatz, J.; Sashidhar, N.; Melzer, S. Recent developments in genome editing and applications in plant breeding. *Plant Breeding* **2018**, 137(1), 1-9.
19. Deveau, H.; Barrangou, R.; Garneau, J.E.; Labonté, J.; Fremaux, C.; Boyaval, P.; Romero, D.A.; Horvath, P.; Moineau, S. Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. *Journal of Bacteriology* **2008**, 190(4), 1390-1400.
20. Chylinski, K.; Makarova, K.S.; Charpentier, E.; Koonin, E.V. Classification and evolution of type II CRISPR-Cas systems. *Nucleic Acids Research* **2014**, 42(10), 6091-6105.
21. Hsu, P.D., Lander, E.S., Zhang, F. Development and Applications of CRISPR-Cas9 for Genome Engineering. *Cell* **2014**, 157(6), 1262-1278.
22. Wu, Y.; Battalapalli, D.; Hakeem, M.J.; Selamneni, V.; Zhang, P.; Draz, M.S.; Ruan, Z. (2021). Engineered CRISPR-Cas systems for the detection and control of antibiotic-resistant infections. *Journal of Nanobiotechnology* **2021**, 19(1), 401.
23. Sharma, V.K.; Marla, S.; Zheng, W.; Mishra, D.; Huang, J.; Zhang, W.; Morris, G.P.; Cook, D.E. CRISPR guides induce gene silencing in plants in the absence of Cas. *Genome Biology* **2022**, 23(1), 6.
24. Zhang, X.H.; Tee, L.Y.; Wang, X.G.; Huang, Q.S.; Yang, S.H. Off-target Effects in CRISPR/Cas9-mediated Genome Engineering. *Molecular Therapy - Nucleic Acids* **2015**, 4, e264.
25. Miri, S.M.; Tafsiri, E.; Cho-Shing, W. C.; Ghaemi, A. CRISPR-Cas, a robust gene-editing technology in the era of modern cancer immunotherapy. *Cancer Cell International* **2022**, 20(1), 456.
26. He, X.; Tan, C.; Wang, F.; Wang, Y.; Zhou, R.; Cui, D.; You, W.; Zhao, H.; Ren, J.; Feng, B. Knock-in of significant reporter genes in human cells via CRISPR/Cas9-induced homology-dependent and independent DNA repair. *Nucleic Acids Research* **2016**, 44(9), e85.
27. Cho, S.W.; Kim, S.; Kim, Y.; Kweon, J.; Kim, H.S.; Bae, S.; Kim, J.S. Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. *Genome Research* **2014**, 24(1), 132-141.

28. Leenay, R. T.; Beisel, C. L. Deciphering, communicating, and engineering the CRISPR PAM. *Journal of Molecular Biology* **2017**, 429(2), 177-191.
29. Biswas, A.; Gagnon, J.N.; Brouns, S.J.J.; Fineran, P. C.; & Brown, C.M. (2013). CRISPRTarget: Bioinformatic prediction and analysis of crRNA targets. *RNA Biology* **2013**, 10(5), 817-827.
30. Hernández-Amasifuen, A.D.; Argüelles-Curaca, A.; Cortez-Lázaro, A.A.; Díaz-Pillasca, H.B. In vitro induction of callus from foliar explants in rocoto (*Capsicum pubescens* Ruiz & Pav.). *Granja* **2021**, 34(2), 127-135.
31. Li, J.; Mahajan A.; Tsai, M.D. Ankyrin repeat: a unique motif mediating protein-protein interactions. *Biochemistry* **2006**, 45 15168–15178.
32. Spoel, S.H., Mou, Z.; Tada, Y.; Spivey, N.W.; Genschik, P.; Dong, X. Proteasome-mediated turnover of the transcription coactivator NPR1 plays dual roles in regulating plant immunity. *Cell* **2009**, 137 860–872.
33. Backer, R.; Naidoo, S.; Van-Den, B.N. The NONEXPRESSOR OF PATHOGENESIS-RELATED GENES 1 (NPR1) and Related Family: Mechanistic Insights in Plant Disease Resistance. *Plant Sci* **2019**; 10:102.
34. Zhang, Y.; Cheng, Y.T.; Qu, N.; Zhao, Q.; Bi, D.; Li, X. Negative regulation of defense responses in Arabidopsis by two NPR1 paralogs. *The Plant Journal* **2006**. 48 647–656.
35. Jiang, D.; Yang, G.; Chen, K.; Yu, P.; Chen, J.; Luo, Y.; Li, N.; Huang, L.J. Identification and Functional Characterization of the Nonexpressor of Pathogenesis-Related Genes 1 (NPR1) Gene in the Tea Plant (*Camellia sinensis*). *Forests* **2023**; 14(8):1578.
36. Rochon, A.; Boyle, P.; Wignes, T.; Fobert, P.R.; Després, C. The Coactivator Function of Arabidopsis NPR1 Requires the Core of Its BTB/POZ Domain and the Oxidation of C-Terminal Cysteines. *Plant Cell* **2006**, 18, 3670–3685.
37. Boyle, P.; Le Su, E.; Rochon, A.; Shearer, H.L.; Murmu, J.; Chu, J.Y.; Fobert, P.R.; Després, C. The BTB/POZ Domain of the Arabidopsis Disease Resistance Protein NPR1 Interacts with the Repression Domain of TGA2 to Negate Its Function. *Plant Cell* **2009**, 21, 3700–3713.
38. Liu, X.; Homma, A.; Sayadi, J.; Yang, S.; Ohashi, J.; Takumi, T. Sequence features associated with the cleavage efficiency of CRISPR/Cas9 system. *Scientific Reports* **2016**, 6, 19675.
39. Liang, G.; Zhang, H.; Lou, D.; Yu, D. Selection of highly efficient sgRNAs for CRISPR/Cas9-based plant genome editing. *Scientific Reports* **2016**, 6, 21451.

40. Malik, A.; Gul, A.; Munir, F.; Amir, R.; Alipour, H.; Babar, M.M.; Bakhtiar, S.M.; Paracha, R.Z.; Khalid, Z.; Hayat, M.Q. (2021). Evaluating the cleavage efficacy of CRISPR-Cas9 sgRNAs targeting ineffective regions of *Arabidopsis thaliana* genome. *PeerJ* **2021**, 9, e11409.
41. Bae, S.; Kweon, J.; Kim, H.S.; & Kim, J.S. Microhomology-based choice of Cas9 nuclease target sites. *Nature Methods* **2014**, 11(7).
42. Robins, H.; Li, Y.; Padgett, R.W. Incorporating structure to predict microRNA targets. *Proc Natl Acad Sci USA* **2005**, 102(11), 4006-4009.
43. Konstantakos, V.; Nentidis, A.; Krithara, A.; Paliouras, G. CRISPR-Cas9 gRNA efficiency prediction: an overview of predictive tools and the role of deep learning. *Nucleic Acids Res* **2022**, 50(7), 3616-3637.
44. Nishimasu, H.; Ran, F.A.; Hsu, P.D.; Konermann, S.; Shehata, S.I.; Dohmae, N.; Ishitani, R.; Zhang, F.; Nureki, O. Crystal structure of Cas9 in complex with guide RNA and target DNA. *Cell* **2014**, 156(5), 935-949.
45. Ma, X.; Zhang, Q.; Zhu, Q.; Liu, W.; Chen, Y.; Qiu, R.; Wang, B.; Yang, Z.; Li, H.; Lin, Y.; Xie, Y.; Shen, R.; Chen, S.; Wang, Z.; Chen, Y.; Guo, J.; Chen, L.; Zhao, X.; Dong, Z.; Liu, Y.G. A Robust CRISPR/Cas9 System for Convenient, High-Efficiency Multiplex Genome Editing in Monocot and Dicot Plants. *Molecular Plant* **2015**, 8(8), 1274-1284.
46. Jinek, M.; Chylinski, K.; Fonfara, I.; Hauer, M.; Doudna, J.A.; Charpentier, E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **2012**, 337(6096), 816-821.

Received: October 9th 2023/ **Accepted:** January 15th 2024 / **Published:** 15 February 2024

Citation: Yupanqui-Celestino M., Azañero-Huaynacari L. K., Pesantes-Rojas C. R., Aguilar-Luna-Victoria M. A., Cotos-Durán D. E., Ramírez-Viena L. Tools and computational resources for the design of CRISPR/Cas9 sgRNA for *NPR3* gene knockout in sour orange (*Citrus aurantium* L.). *Bionatura Journal* 2024; 1 (1) **35**. <http://dx.doi.org/10.70099/BJ/2024.01.01.35>

Additional information Correspondence should be addressed to Cusy2000@gmail.com

ISSN.3020-7886

Peer review information. Bionatura Journal thanks anonymous reviewer(s) for their contribution to the peer review of this work using <https://reviewerlocator.webofscience.com/>

All articles published by Bionatura Journal are made freely and permanently accessible online immediately upon publication, without subscription charges or registration barriers.

Publisher's Note: Bionatura Journal stays neutral concerning jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2024 by the authors. They were submitted for possible open-access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Copyright: © 2024 by the authors. They were submitted for possible open-access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).