bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448838; this version posted June 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Characterization of endogenous Rubus yellow net virus in raspberries

Thien Ho[1,*], Janet C. Broome[1], Jason P. Buhler[1], Wendy O'Donovan[1], Tongyan Tian[2], Alfredo Diaz-Lara[3], Robert R. Martin[4], Ioannis E. Tzanetakis[5]

[1] Driscoll's Inc., Watsonville, CA 95076, USA

[2] California Department of Food and Agriculture, Sacramento, CA 95832, USA

[3] School of Engineering and Sciences, Tecnologico de Monterrey, Campus Queretaro, Queretaro 76130, Mexico

[4] Department of Botany and Plant Pathology, Oregon State University, and USDA-ARS Horticultural Crops Research Unit, Corvallis, OR 97330, USA

[5] Department of Entomology and Plant Pathology, Division of Agriculture, University of Arkansas System, Fayetteville, AR 72701, USA

[*] Corresponding author: thien.ho@driscolls.com

Abstract

Rubus yellow net virus (RYNV) belongs to genus *Badnavirus*. Badnavirids are found in plants as endogenous, inactive sequences, and/or in episomal (infectious and active) forms. To assess the state of RYNV infections, we sequenced the genomes of various *Rubus* cultivars and mined eight additional published whole genome sequencing datasets. Sequence analysis revealed the presence of a diverse array of endogenous RYNV (endoRYNV) sequences that differ significantly in their structure; some lineages have nearly complete, yet non-functional genomes whereas others have rudimentary, small sequence fragments. We developed assays to genotype the six main endoRYNV lineages as well as the only known episomal lineage in commercial *Rubus*. This study discloses the widespread presence of endoRYNVs in commercial raspberries, likely because breeding programs have been using a limited pool of germplasm that harbored endoRYNVs.

Keywords: *Rubus yellow net virus*, raspberry, *Rubus*, *Badnavirus*, virus diversity

1. Introduction

Raspberry is an economically important crop with global production in 2018 being over 820,000 tons grown in 125,000 hectares in all continents except Antarctica ("Production of raspberries in 2018". United Nations, Corporate Statistical Database (FAOSTAT) 2019, retrieved January 14, 2021). Commercial breeding for red raspberry (*Rubus idaeus*) began about 200 years ago, and most of the currently available cultivars share the same germplasm pedigree dating back to the late 1800s and early 1900s (Jennings, 2018).

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448838; this version posted June 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

More than 40 virus species are known to infect *Rubus*; yet Rubus yellow net virus (RYNV) is one of only two badnaviruses known to infect the genus (Diaz-Lara et al., 2015; Shahid et al., 2017). The genus *Badnavirus*, family *Caulimoviridae*, includes viruses that have an inactive, endogenous form and an infectious, episomal form. *Banana streak OL virus* (BSOLV), *Banana streak GF virus* (BSGFV), and *Banana streak IM virus* (BSIMV) integrants have been shown to reactivate from an inactive integrated counterpart (Chabannes et al., 2013; Gayral et al., 2008; Ndowora et al., 1999), whereas no other badnavirid is known to reactivate from integrated sequences (reviewed by Bhat et al., 2016).

RYNV is a component of raspberry mosaic, an important disease first described in the 1920s (Bennett, 1927; Stace-Smith, 1955). The virus infects all red raspberry and most blackberry and hybrid berry cultivars in North America and Europe (Stace-Smith and Jones, 1987a) and can reduce yield from 30-75% in the first year and up to 15% in subsequent years in mixed infections with black raspberry necrosis virus (Stace-Smith and Jones, 1987b). Partial RYNV sequences were first obtained at the turn of the century (GenBank accession number AF468454; Jones et al., 2002). The plant used in the Jones et al. (2002) study had virus-like symptoms and bacilliform particles were observed under the electron microscope. The first RYNV genome (RYNV-Ca, GenBank accession number KF241951), assembled from two PCR amplicons, was obtained by Kalischuk et al. (2008). Another genome (RYNV-BS, KM078034, Diaz-Lara et al., 2015) was sequenced from red raspberry 'Baumforth's Seedling A' using DNA from rolling circle amplification. Since then, several RYNV sequences were published using small RNA (Rajamäki et al., 2019) or whole genome sequencing (MN245240).

Diaz-Lara (2016) observed that red raspberry plants, supposedly free of RYNV based on aphid or graft transmission onto *R. occidentalis* 'Munger' indicator, yielded positive results when indexed by PCR-based assays. Moreover, those plants were tested positive for RYNV by PCR even after heat therapy and meristem-tip culture for virus elimination. It was demonstrated that RYNV integrates into the red raspberry genome (Diaz-Lara et al., 2020), but no further analysis was conducted for the reported endogenous RYNV (endoRYNV) sequences. In this study, multiple cultivars were assayed to determine the prevalence of endoRYNV and the lineages identified were validated and characterized in-depth.

2. Materials and methods

2.1. Plant material

Twenty-five raspberry cultivars maintained as tissue culture plantlets in Watsonville, California were used in the study (Table 1). For 'Baumforth's Seedling A', an additional mature plant was obtained from Corvallis, Oregon with the RYNV-BS (Diaz-Lara et al., 2015; Table 1) and used as a positive control for the episomal form, hereafter referred to as epiRYNV-BS. In addition, the genome of 75 proprietary red raspberry and 100 proprietary blackberry cultivars were sequenced and assayed for integration of RYNV-BS and the episomal form of the virus but their identity is not provided to protect intellectual property rights.

2.2. DNA purification, sequencing, and virus discovery

DNA was extracted using either the DNeasy(R) kit (Qiagen) or the method described by Poudel et al. (2013). All DNA libraries were constructed using a TruSeq DNA HT Sample Prep(R) kit and sequenced individually using paired-end (2 × 300 bp) Illumina HiSeq configuration by Novogene (Sacramento, CA).

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448838; this version posted June 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Raw Illumina reads were subject to de novo assembly using Spades (Bankevich et al., 2012). BLASTn search (Camacho et al., 2009) was performed on the output contigs with e-value=10 against published RYNV nucleotide sequences (nt) downloaded from GenBank nt database (January 16, 2021). After RYNV hits were filtered out, the remaining contigs were processed using BLASTx against a database containing all RYNV protein sequences downloaded from GenBank nr (January 16, 2021). All Illumina datasets were also submitted to VirFind (http://virfind.org, Ho and Tzanetakis, 2014) for virus detection and discovery. Bowtie2 (Langmead and Salzberg, 2012) was used for mapping raw reads to RYNV contigs for visual confirmation of the mapping assemblies with Tablet (Milne et al., 2013). BioEdit (Hall, 1999) was used to calculate sequence identity matrix, and ClustalW (Thompson et al., 1994) of the MEGA X software (Kumar et al., 2018) applied to align nucleotide and amino acid sequences. Expasy (https://web.expasy.org/translate/) was used to predict open reading frames (ORF). Conserved domain search was done using the NCBI homonymous tool (Lu et al., 2020). Breaking points of the RYNV lineages were identified by aligning raw Illumina reads with BLASTn against the assembled sequences and partially aligned reads were manually analyzed for sequence identities.

2.3. Electron microscopy

Tissues were homogenized in 100 mM potassium phosphate pH 7.0, 2% polyvinylpyrrolidone (MW:10,000) and 0.2% Na_2SO_3 at 1:20 (w:v). After a low-speed centrifugation at 10,000 g, the supernatant was used for immunosorbent electron microscopy (ISEM) according to Lockhart (1986). Briefly Formvar/carbon coated copper grids were floated on 10 µl of the capture antibodies of *s*ugarcane bacilliform virus (NanoDiagnostics, Arkansas) diluted 1:10 in 50 mM potassium phosphate buffer pH7.0. After 30 min incubation at 37°C, grids were rinsed with 50 mM potassium phosphate buffer pH7.0, and then floated on 30 µl of the sample preparations for 20-22 hrs at 4°C. The grids were then rinsed with 50 mM potassium phosphate buffer pH7.0 containing 100 µg/ml bacitracin and stained with 0.5% phosphotungstic acid pH7.0 containing 100 µg/ml bacitracin. The grids were examined using a Hitachi H-7500 transmission electron microscope (Hitachi High-Tech Corporation, Fukuoka, Japan) with an AMT Biosprint 12M-B CCD camera (Advanced Microscopy Techniques, Woburn, MA). Virus particles were measured using the camera software.

2.4. Data mining

Published datasets were mined for RYNV sequences (Table 1) including raw Illumina data of red raspberry cultivars 'Caroline', 'Cascade Bounty', 'Comox', 'Glen Cova', 'Meeker', and 'Willamette' from the Diaz-Lara et al. (2020) study, 'Glen Moy' from Hackett et al. (2018), and the assembled 'Joan J' genome, obtained using PacBio and Illumina sequencing (Wight et al. 2019). These eight datasets were processed using the procedures described above.

2.5. Development of RYNV lineage-specific primers and validation

For each endo/epi RYNV lineage, 20 PCR primer sets were designed by processing the corresponding sequences using PrimerQuest at default parameters for 'qPCR Intercalating Dyes' option (Integrated DNA Technologies, IDT). The outputs were aligned with all RYNV sequences and 5-10 oligo pairs were selected with each oligo, when possible, having at least 2nt mismatches to other RYNV lineages. SYBR Green quantitative PCR was performed for each set against cultivars 1-25 (Table 1) using 20 ng plant DNA, 5 µl Maxima SYBR Green qPCR Master Mix (2X) (Thermo Scientific, catalog number K0253), 1 µM

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448838; this version posted June 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

each of forward and reverse primers, and water to 10 µl. Amplification was performed on QuantStudio 6 Flex instrument (Applied Biosystems) with the amplification program consisting of 95°C for 10 min, followed by 40 cycles of 95°C, 53°C, and 72°C, for 20s each. The melting stage started at 53°C for 1m, increased by 0.05°C/s and stopped after reaching 95°C for 15s. To investigate further the possibility of integration of the epiRYNV-BS lineage, we used the most consistent assay against the epiRYNV-BS lineage developed here against a panel of 271 public and proprietary genotypes bulked into 294 DNA samples consisting of 876 plants (Supplemental table). Samples were considered positive for a lineage if there was amplification with the correct melting point. The previously published assay of Diaz-Lara et al. (2020) was also included in this validation for specificity comparison.

3. Results

3.1. A diverse array of endoRYNV is present in raspberry

The presence of RYNV DNA in commercial raspberries was investigated by whole genome sequencing and mining data of 25 and 8 cultivars, respectively (Table 1). Sequencing produced approximately 9 Gbp for each cultivar, representing ~30 X coverage of the predicted 300-Mbp raspberry genome (Wight et al., 2019). BLASTn and BLASTx steps using contigs assembled from the raspberry genomes found no evidence of RYNV DNA in cultivars ‘Korbfüller’, ‘Lloyd George’, ‘Malling Jewel’, ‘Octavia’, ‘Yellow Antwerp’, or the black raspberry cultivar ‘Munger’. The remaining 27 cultivars showed a diverse array of RYNV sequences, ranging from six partial genomic segments with duplicated and rearranged sequences, to rudimentary fragments of a few hundred base pairs. None of the newly discovered lineages possesses an intact genome. Bacilliform virus particles were consistently observed using electron microscope from ‘Baumforth’s Seedling A’ from Corvallis OR, the only sample in the study with a verified episomal form of the virus (Diaz-Lara et al., 2020) (Figure 1). The size of the virions is approximately 149 nm X 33 nm (n=102). No additional badnavirid sequences were detected in this or any other sample using VirFind (Ho and Tzanetakis, 2014) indicating that the observed particles belong to RYNV. Bacilliform particles were not detected from any of the other samples, including the ‘Baumforth’s Seedling A’ from Watsonville CA indicating that all other RYNV sequences are integrated in the raspberry genome.

3.2. Structure of main endoRYNV lineages

RYNV-derived sequences more than 4 Kbp in length were named based on the oldest cultivars they were identified in. RYNV has five conserved badnavirid domains including reverse transcriptase, ribonuclease H (RNaseH), pepsin-like aspartate protease, a zinc knuckle which is a zinc binding motif from retroviral gag proteins (nucleocapsid), and a ribosomal L25/TL5/CTC N-terminal 5S rRNA binding domain (Diaz-Lara et al., 2015). The reverse transcriptase and RNaseH (RT_RNaseH) domains were concatenated and used for sequence comparison against those of the three available genomes (GenBank accession numbers KF241951, KM078034, and MN245240). The new lineages shared >78.5% in nucleotide identities to each other and the epiRYNV-BS (Table 2).

<u>3.2.1. endoRYNV-CU1</u>

‘Cuthbert’ (1865 release, New York), has three endoRYNVs (namely endoRYNV-CU1, -CU2, and -CU3) and is the oldest cultivar in the study having an endoRYNV. endoRYNV-CU1 was discovered in 12 cultivars (Table 1). The 7268-nt segment has intact 5’ intergenic region (IG), ORF1, ORF2, but when compared to epiRYNV-BS, ORF3 is missing 549 bp after nt1691, corresponding to 183 amino acids (aa),

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448838; this version posted June 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

at the site for the ribosomal L25/TL5/CTC N-terminal 5S rRNA binding domain, essential to the badnavirus movement (Figure 2). The four other domains are similar to the epiRYNV-BS. ORFs 4 and 6 are embedded intact within ORF3.

endoRYNV-CU1 RT_RNaseH region shares 100% nt identity with that of RYNV-Ca (Kalischuk et al., 2013). However, there is 12.9% nt diversity between the two lineages. endoRYNV-CU1 3' IG is intact but very different from RYNV-Ca 3' IG, sharing only 29% nt identity (Figure 3). The RYNV-Ca has two inverted repeats, at nt4325-4693 and nt7564-7932. The endoRYNV-CU1 integrant is highly fragmented and has breaking points at genomic nt positions 12, 15, 767, 815, 1321, 1426, 2473, and 6527 (Table 3). One plant-virus junction was detected at nt7268.

3.2.2. endoRYNV-CU2

The second 'Cuthbert' endoRYNV (CU-2) is present in 11 cultivars. The 7252-nt sequence starts with the complete 5' IG, ORF1 and ORF2. When aligned against epiRYNV-BS, ORF3 is lacking a 141-nt stretch after nt2455, corresponding to 47 aa. Similar to endoRYNV-CU1, it has four conserved domains similar to the epiRYNV-BS but missing the ribosomal L25/TL5/CTC N-terminal 5S rRNA binding domain as well as about 428 nt of the 3' IG. Alike endoRYNV-CU1, its ORF4 and ORF6 are embedded intact within ORF3. The endoRYNV-CU2 RT_RNaseH region shares 79% nt identity with that of epiRYNV-BS and the integrant is fragmented at genomic nt positions 2217 and 5319, with plant-virus junctions detected at nt1, 188, and 5863.

3.2.3. endoRYNV-CU3

The third 'Cuthbert' endoRYNV (CU-3) is present in five cultivars and is heavily truncated. It is lacking 5' IG, ORF1 and ORF2. Its sequence starts with a truncated ORF3 and together with the 3' IG accounting for a 4550-bp stretch. Its RT_RNaseH region shares 99.6% to RYNV-Cu from Chile (MN245240). The integrant is fragmented at nt4423 and 4550, and has plant-virus junctions at nt1 and 1492.

3.2.4. endoRYNV-BS

The endoRYNV-BS was first detected in 'Baumforth's Seedling A' (1880 release, UK) and is present in 17 cultivars. The lineage is 7602 bp, and has intact 5' IG, ORF1, and ORF2. When aligned against epiRYNV-BS, ORF3 is missing a 132-nt stretch after nt2445 corresponding to 44 aa, and the 3' IG lacks 83 bp. This lineage has all five conserved badnavirid domains. The lineage is present in the assembled genome of 'Joan J' as a single copy on chromosome 4. The integrant is 12,143 bp and composed of two fragments. The first is in the forward orientation and contains complete 5' IG that forms a junction with the plant DNA at its 5' end, followed by complete ORF1, ORF2, and part of ORF3 that is truncated at nt6317. The second follows immediately after in the reverse orientation, with a truncated 3' IG at nt7602, then continues with ORF3 but truncated at nt1777 fusing to the plant genome. No full-length ORF3 is present in either of the fragments.

3.2.5. endoRYNV-PH1

First detected in Phoenix (1896 release, UK), endoRYNV-PH1 is present in nine cultivars. The 6631nt sequence starts with the 5' IG, followed by the ORF1 of 177 nt and missing 380 nt after nt561 when aligned against the epiRYNV-BS, before an intact ORF2. ORF3 is missing 675 nt after nt1307 as well as

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448838; this version posted June 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

the ribosomal L25/TL5/CTC N-terminal 5S rRNA binding domain. It has the four other conserved domains similar to the epiRYNV-BS. The integrant is fragmented at genomic nt positions 335, 1373, 1814, and 5859, with nt1 and nt6631 connected to the plant DNA.

3.2.6. endoRYNV-PH2

The last substantial integrated RYNV sequence, endoRYNV-PH2, was only found in 'Phoenix'. The 7091-nt fragment's 5' end has the intact 5' IG, ORF1 and ORF2. ORF3 misses 141 nt after nt2461 corresponding to 47 aa, and the fragment terminates at nt7091. This sequence has all conserved domains found in the epiRYNV-BS. The integrant is fragmented at nt2201 and 2432, and has two plant-virus junctions at nt3094 and 7091.

3.3. Validation

SYBR Green qPCR assays were able to differentiate between the epiRYNV-BS and each of the main endoRYNVs in all cultivars used in this study, either in single or multiple integration events (Figure 4, Table 1). Except the positive control 'Baumforth's Seedling A' OR, the epiRYNV-BS lineage was not detected in any of the genetics used in this study (Figure 5). The epiBS-2604F/2715R assay did not have off-target melting points and amplifications, compared to the Diaz-Lara et al. (2020) assay, presently considered the better assay for RYNV detection, with off-target in 43 cases (Supplemental figure).

4. Discussion

We analyzed the genome sequence data of commercial cultivars from around the globe released as early as 1802 and as recent as 2006. Integrated RYNV sequences were present in 27/33 cultivars (82%). The endoRYNV population could be categorized into six main lineages and other short endogenous fragments. The diversity of endoRYNV is complex with sometimes sequences having inversions, duplications, or deletions.

Rubus domestication has resulted in a reduction of genetic diversity (Haskell, 1960; Jennings, 1988), and modern cultivars are genetically similar to each other (Dale et al., 1993; Graham and McNicol, 1995). This can be seen in the case of the cultivars analyzed in this study. All breeding programs share the same endoRYNV lineages, which in turn were discovered in three cultivars commercialized in the 19th century: 'Cuthbert' (1865), 'Baumforth's Seedling A' (1880), and 'Phoenix' (1896). These endogenous sequences presumably became widespread as the three aforementioned cultivars were used as parents, or are in the lineages of most raspberry breeding programs worldwide.

endoRYNV-CU1 lineage is the closest isolate to the published RYNV-Ca sequence (KF241951) (Kalischuk et al., 2013). Since RYNV-Ca has two inverted repeats, misses the true 3' IG, and hence likely is an endogenous sequence, we consider epiRYNV-BS as the sole episomal RYNV lineage known to infect *Rubus*. It is important to note that when aligned against the epiRYNV-BS sequence, all endoRYNV lineages are truncated and missing genomic DNA stretches. From this data, we hypothesize that the endoRYNVs are unable to reactivate and become episomal due to their incomplete genomes. In addition to the raspberry cultivars of this study, we sequenced the whole genome of an additional 75 proprietary red raspberry cultivars, and epiRYNV-BS was absent in all (data not shown), indicating that this lineage may be unable to integrate in the raspberry genome. We also sequenced the genomes of 100

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448838; this version posted June 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

proprietary blackberry cultivars (data not shown) but did not find any evidence of endoRYNV, suggesting that endoRYNV sequences may be limited to red raspberry.

Diagnostic tests for infectious agents are necessary so that phytosanitary agencies can protect a country's natural resources and agriculture. However, the Rubus industry could be significantly impacted if a diagnostic test was positive for RYNV but inadvertently a no-risk endoRYNV was detected. Published PCR primers were designed to target either RYNV-Ca or epiRYNV-BS as they had been the only known RYNV lineages (Diaz-Lara et al., 2020; Jones et al., 2002; Kalischuk et al., 2008). Diaz-Lara et al. (2020) showed that primers currently used for RYNV detection could produce positive results in cultivars only harboring endoRYNV DNA, indicating the urgency to have a good diagnostic test that can clearly differentiate the two forms, similar to epiBS-2604F/2715R. This test should be developed and validated for accuracy and sensitivity against a wide range of episomal isolates.

Theoretically, endoRYNV can be removed from the red raspberry by traditional breeding. However, the effort required to remove endoRYNV DNA after multiple generations of backcrossing would be considerable, especially when desired traits must be retained. CRISPR-Cas9 could be used to remove the endoRYNVs, but for cultivars with multiple endoRYNV fragments, multiple gene-editing events will need to be done. We believe that such actions are not necessary as endoRYNV fragments could not reconstruct a full, infectious, genome.

5. Conflict of interest

T.H., J.C.B., J.P.B., W.O. are employees of Driscoll's Inc.

6. Acknowledgement

This work could not have been accomplished without the help of various colleagues. We thank the Driscoll's team including the raspberry breeding, molecular genetics, and greenhouse departments; and Melanie Kalischuk and Hanu Pappu for discussion of the RYNV-Ca work.

7. Data availability

The assembled sequences of endoRYNV-CU1, endoRYNV-CU2, endoRYNV-CU3, endoRYNV-BS, endoRYNV-PH1, and endoRYNV-PH2 were deposited on the NCBI's GenBank under accessions XXX-XXX, and the raw Illumina reads of all cultivars sequenced in this study to the Sequence Read Archive (SRA) under accessions XXX-XXX.

8. References

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., Pyshkin, A.V., Sirotkin, A.V., Vyahhi, N., Tesler, G., Alekseyev, M.A., Pevzner, P.A., 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19, 455–477. https://doi.org/10.1089/cmb.2012.0021

Bennett, C.W., 1927. Virus diseases of raspberries. Mich. Agr. Expt. Sta. Tech. Bull.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448838; this version posted June 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., Madden, T.L., 2009. BLAST+: architecture and applications. BMC Bioinformatics 10, 421–421. https://doi.org/10.1186/1471-2105-10-421

Chabannes, M., Baurens, F.-C., Duroy, P.-O., Bocs, S., Vernerey, M.S., Rodier-Goud, M., Barbe, V., Gayral, P., Iskra-Caruana, M.-L., 2013. Three infectious viral species lying in wait in the banana genome. Journal of Virology 87, 8624–8637.

Dale, A., Moore, P.P., McNicol, R.J., Sjulin, T.M., Burmistrov, L.A., 1993. Genetic diversity of red raspberry varieties throughout the world. Journal of the American Society for Horticultural Science jashs 118, 119–129. https://doi.org/10.21273/JASHS.118.1.119

Diaz-Lara, A., 2016. Identification of Endogenous and Exogenous Pararetroviruses in Red Raspberry (Rubus idaeus L.) and Blueberry (Vaccinium corymbosum L.). PhD thesis. Oregon State University.

Diaz-Lara, A., Mosier, N.J., Keller, K.E., Martin, R.R., 2015. A variant of Rubus yellow net virus with altered genomic organization. Virus Genes 50, 104–110. https://doi.org/10.1007/s11262-014-1149-6

Diaz-Lara, A., Mosier, N.J., Stevens, K., Keller, K.E., Martin, R.R., 2020. Evidence of Rubus Yellow Net Virus Integration into the Red Raspberry Genome. Cytogenetic and Genome Research 160, 329–334. https://doi.org/10.1159/000509845

Gayral, P., Noa-Carrazana, J.-C., Lescot, M., Lheureux, F., Lockhart, B.E.L., Matsumoto, T., Piffanelli, P., Iskra-Caruana, M.-L. 2008. A single Banana streak virus integration event in the banana genome as the origin of infectious endogenous pararetrovirus. Journal of Virology 82, 6697–6710.

Graham, J., McNicol, R.J., 1995. An examination of the ability of RAPD markers to determine the relationships within and between Rubus species. Theoretical and Applied Genetics 90, 1128–1132. https://doi.org/10.1007/BF00222932

Hackett, C.A., Milne, L., Smith, K., Hedley, P., Morris, J., Simpson, C.G., Preedy, K., Graham, J., 2018. Enhancement of Glen Moy x Latham raspberry linkage map using GbS to further understand control of developmental processes leading to fruit ripening. BMC Genetics 19, 59. https://doi.org/10.1186/s12863-018-0666-z

Hall, T., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symposium Series 41, 95–98.

Haskell, G., 1960. The raspberry wild in Britain. Watsonia 4, 238–255.

Ho, T., Tzanetakis, I.E., 2014. Development of a virus detection and discovery pipeline using next generation sequencing. Virology 471–473, 54–60. https://doi.org/10.1016/j.virol.2014.09.019

Jennings, D.L., 1988. Raspberries and blackberries: their breeding, diseases and growth. Academic, London.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448838; this version posted June 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Jennings, S.N., 2018. Advances in Rubus Breeding, in: Raspberry: Breeding, Challenges and Advances (Eds Graham J., Brennan R.) Pp 17–28. (Springer International Publishing, Cham, 2018).

Jones, A.T., McGavin, W.J., Geering, A.D.W., Lockhart, B., 2002. Identification of Rubus yellow net virus as a distinct badnavirus and its detection by PCR in Rubus species and in aphids. Annals of Applied Biology 141, 1–10. https://doi.org/10.1111/j.1744-7348.2002.tb00189.x

Jones, A.T., Roberts, I.M., 1976. Ultrastructural changes and small bacilliform particles associated with infection by rubus yellow net virus. Annals of Applied Biology 84, 305–310. https://doi.org/10.1111/j.1744-7348.1976.tb01773.x

Kalischuk, M.L., Fusaro, A.F., Waterhouse, P.M., Pappu, H.R., Kawchuk, L.M., 2013. Complete genomic sequence of a Rubus yellow net virus isolate and detection of genome-wide pararetrovirus-derived small RNAs. Virus Research 178, 306–313. https://doi.org/10.1016/j.virusres.2013.09.026

Kalischuk, M.L., Kawchuk, L.M., Leggett, F., 2008. First Report of Rubus yellow net virus on Rubus idaeus in Alberta, Canada. Plant Disease 92, 974–974. https://doi.org/10.1094/PDIS-92-6-0974A

Kumar, S., Stecher, G., Li, M., Knyaz, C., Tamura, K., 2018. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. Molecular Biology and Evolution 35, 1547–1549. https://doi.org/10.1093/molbev/msy096

Langmead, B., Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. Nature Methods 9, 357–359. https://doi.org/10.1038/nmeth.1923

Lockhart, B.E.L., 1986. Purification and serology of a bacilliform virus associated with banana streak disease. Phytopathology 76, 995–999.

Lu, S., Wang, J., Chitsaz, F., Derbyshire, M.K., Geer, R.C., Gonzales, N.R., Gwadz, M., Hurwitz, D.I., Marchler, G.H., Song, J.S., Thanki, N., Yamashita, R.A., Yang, M., Zhang, D., Zheng, C., Lanczycki, C.J., Marchler-Bauer, A., 2020. CDD/SPARCLE: the conserved domain database in 2020. Nucleic Acids Research 48, D265–D268. https://doi.org/10.1093/nar/gkz991

Milne, I., Stephen, G., Bayer, M., Cock, P.J.A., Pritchard, L., Cardle, L., Shaw, P.D., Marshall, D., 2013. Using Tablet for visual exploration of second-generation sequencing data. Briefings in Bioinformatics 14, 193–202. https://doi.org/10.1093/bib/bbs012

Ndowora, T., Dahal, G., LaFleur, D., Harper, G., Hull, R., Olszewski, N.E., Lockhart, B., 1999. Evidence That Badnavirus Infection inMusaCan Originate from Integrated Pararetroviral Sequences. Virology 255, 214–220. https://doi.org/10.1006/viro.1998.9582

Poudel, B., Wintermantel, W.M., Cortez, A.A., Ho, T., Khadgi, A., Tzanetakis, I.E., 2013. Epidemiology of Blackberry yellow vein associated virus. Plant Disease 97, 1352–1357. https://doi.org/10.1094/PDIS-01-13-0018-RE

Rajamäki, M.-L., Lemmetty, A., Laamanen, J., Roininen, E., Vishwakarma, A., Streng, J., Latvala, S., Valkonen, J.P.T., 2019. Small-RNA analysis of pre-basic mother plants and conserved accessions

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448838; this version posted June 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

of plant genetic resources for the presence of viruses. PLOS ONE 14, e0220621. https://doi.org/10.1371/journal.pone.0220621

Shahid, M.S., Aboughanem-Sabanadzovic, N., Sabanadzovic, S., Tzanetakis, I.E., 2016. Genomic Characterization and Population Structure of a Badnavirus Infecting Blackberry. Plant Disease 101, 110–115. https://doi.org/10.1094/PDIS-04-16-0527-RE

Stace-Smith, R., 1955. Studies on rubus virus diseases in British Columbia: Rubus yellow-net. Canadian Journal of Botany 269–274.

Stace-Smith, R., Jones, A.T., 1987a. Rubus yellow net virus, in: Virus Diseases of Small Fruits, USDA Agriculture Handbook No. 631. Converse, R H, pp. 175–178.

Stace-Smith, R., Jones, A.T., 1987b. USDA Agriculture Handbook No. 631 (USDA, Washington) 175.

Tamura, K., Nei, M., 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Molecular Biology and Evolution 10, 512–526. https://doi.org/10.1093/oxfordjournals.molbev.a040023

Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Research 22, 4673–4680. https://doi.org/10.1093/nar/22.22.4673

Wight, H., Zhou, J., Li, M., Hannenhalli, S., Mount, S.M., Liu, Z., 2019. Draft Genome Assembly and Annotation of Red Raspberry Rubus Idaeus. bioRxiv 546135. https://doi.org/10.1101/546135

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448838; this version posted June 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

9. Tables

Table 1. Description of the origin of the raspberry cultivars used as plants for Illumina sequencing and method development in this study, or mined from published literatures, sorted by year of release, and subsequent RYNV detection result* using analysis of whole genome sequencing data and validated by SYBR Green PCR detection.**

	Cultivar names	Year Release	Breeding Program	Endo RYNV -CU1	Endo RYNV -CU2	Endo RYNV -CU3	Endo RYNV -BS	Endo RYNV -PH1	Endo RYNV -PH2	Fragment	Epi RYNV -BS
1	Yellow Antwerp	1802	Belgium	No	No	No	No	No	No	No	No
2	Cuthbert	1865	New York	Yes	Yes	Yes	No	No	No	Yes	No
3	Baumforth's Seedling A	1880	UK	Yes	No	No	Yes	No	No	No	Yes/No***
4	Phoenix	1896	UK	No	Yes	No	Yes	Yes	Yes	Yes	No
5	Latham	1920	Minnesota	No	No	No	Yes	Yes	No	No	No
6	St. Regis	1920	New Jersey	No	Yes	No	Yes	Yes	No	Yes	No
7	Lloyd George	1923	UK	No	No	No	No	No	No	No	No
8	Malling Landmark	1943	UK	Yes	No	No	No	No	No	No	No
9	Korbfüller	1945	Germany	No	No	No	No	No	No	No	No
10	September	1947	New York	No	Yes	No	Yes	Yes	No	Yes	No
11	Mandarin	1955	North Carolina	No	No	No	Yes	No	No	No	No
12	Chilcotin	1965	British Columbia	No	No	No	Yes	No	No	No	No
13	Southland	1968	North Carolina	No	Yes	Yes	Yes	Yes	No	No	No
14	Heritage	1969	New York	No	Yes	Yes	Yes	Yes	No	No	No
15	Malling Jewel	1980	UK	No	No	No	No	No	No	No	No
16	Titan	1982	New York	No	No	No	Yes	Yes	No	No	No
17	Autumn Bliss	1984	UK	No	Yes	No	Yes	Yes	No	Yes	No
18	Summit	1989	Oregon	No	No	No	Yes	Yes	No	No	No
19	Tulameen	1991	British Columbia	No	No	No	Yes	No	No	Yes	No
20	Qualicum	1995	British Columbia	Yes	No	No	No	No	No	Yes	No
21	Prelude	1998	New York	Yes	Yes	No	Yes	No	No	No	No
22	Polka	2001	Poland	No	No	Yes	No	No	No	No	No
23	Octavia	2002	UK	No	No	No	No	No	No	No	No
24	Chemainus	2006	British Columbia	No	No	No	Yes	No	No	Yes	No
25	Munger	1890	Ohio	No	No	No	No	No	No	No	No
26	Willamette	1943	Oregon	Yes	No	No	No	No	No	Yes	No
27	Meeker	1967	Washington State	Yes	No	No	No	No	No	No	No
28	Glen Cova	1969	UK	Yes	No	No	No	No	No	Yes	No
29	Comox	1978	British Columbia	Yes	No	No	No	No	No	Yes	No
30	Glen Moy	1986	UK	Yes	No	No	No	No	No	Yes	No
31	Caroline	1998	Maryland	Yes	Yes	Yes	No	No	No	No	No
32	Cascade Bounty	2005	Washington State	Yes	Yes	No	Yes	No	No	Yes	No
33	Joan J	2005	UK	No	Yes	No	Yes	No	No	No	No

**: Yes = positive; No = negative.*

***: Samples 1-25 were plant samples in this study and validated using SYBR Green PCR detection, while 26-33 were data mined from the published literature and not subject to SYBR Green PCR. Further descriptions are presented in sections 2.1 and 2.4.*

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448838; this version posted June 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

****: Yes for the plant housed in Corvallis OR that has the episomal RYNV-BS; No for the plant housed in Watsonville CA.*

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448838; this version posted June 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Table 2. Percent identity matrix of the reverse transcriptase and RNaseH region of the six Rubus yellow net virus (RYNV) lineages discovered in this study against the three published RYNV sequences, *Gooseberry vein banding virus* (GBVBV), *Grapevine vein clearing virus* (GVCV), and *Birch leafroll associated virus* (BLRaV).

	endoRYNV-BS	endoRYNV-CU1	endoRYNV-CU2	endoRYNV-CU3	endoRYNV-PH1	endoRYNV-PH2	epiRYNV-BS KM078034	RYNV-Cu MN245240	RYNV-Ca KF241951	GBVBV HQ852249.1	GVCV KT907478.1
endoRYNV-CU1	80.5										
endoRYNV-CU2	80	82.5									
endoRYNV-CU3	79.8	90.3	81.8								
endoRYNV-PH1	80.3	90.3	81.3	88.5							
endoRYNV-PH2	81.1	87.1	89.3	85.2	85.4						
epiRYNV-BS KM078034	78.5	81.2	79	82.1	84.4	81.3					
RYNV-Chile MN245240	79.7	90.1	81.7	99.6	88.3	85	82				
RYNV-Ca KF241951	80.5	100	82.5	90.3	90.3	87.1	81.2	90.1			
GBVBV HQ852249.1	71.1	71	71.5	71.1	72.1	71.8	71	70.9	71		
GVCV KT907478.1	67.7	68.1	66.7	67.5	68.4	66.9	68.5	67.4	68.1	68.2	
BLRaV MG686419.1	66.9	66.4	65.4	66.4	67.3	65.7	65.7	66.3	66.4	67.3	69.8

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448838; this version posted June 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Table 3. Break points in the genomic sequence, and plant-virus junctions of the endogenous RYNV lineages discovered in this study. Nucleotide positions relative to the corresponding RYNV genomic sequence are shown within parentheses. Vertical bars indicate break/junction points.

<table>
<tr><th>Lineage</th><th>Break point</th><th>Nucleotide sequence</th></tr>
<tr><td>endoRYNV
-CU1</td><td>6527|12</td><td>(6390)GGCTTCCCAAGGAACTAGCGGAGCTCACGGCCGAACTGGTCAAAGGAAGGGACGAAGCCCTGGTGAACAA
GGAGGTACAGAGGAACATCTCATGTTTTCTCGAGACTGCCCTCCTCCAAGCGGAGAAATCCGTGACTA(6527) |
TAC(12)GCTCTGATACCA(1)(7268)GAAAAATGGCCCGAAAACATGGATCAAAACTTCTTAAGTGCGCTCTTGAACA
CTTAGAAGAACATAAAATTCAGAGACAATGCGGAAGATAGAACTGAATAAACTCTTTATTTCTCAAGAAAGTTAATA
CAATACAAGAGAAGGTT(7123)</td></tr>
<tr><td></td><td>15|815</td><td>(132)TTTGACACAGACACTCACTCCTCTTCCCTCCGAAAAAGAGAGCAAACTCAAGGTTCTGGTCTGAAACTTGGGT
TTCTCGACAAGAAGGAAAGTGTTTAAGCTACCATGGTGAGAGCTA(15)TG |
(815)AAGGCCACACAGTCAGCCCTGAAGCTAGGCTTCGCACAACTGCAGGAGGCAGTTCAGCTGATCATCACAAGG
GAAAACGATCCCAAACCAATCGAAGCAGCTACTGCACAAGTAGCCGAACAGCTGAGGAAGCAGCTTATTGAGGTC
AAGTCCGTCCTCGAGGAGACCAAGAAGATCGC(993)</td></tr>
<tr><td></td><td>767|1426</td><td>(685)CTCCAGACTCACCCAGTACCTGACAACAAAGGTTGGTTCACTACCAACAATCCCGGAGGATTCACCCCTCCTG
GACCAAGCCA(767) |
(1426)CAATACAATGAGCAGAAGCCAAACAAGGCTCCAGGCTCCGCCCGCAACTGAAAGGGCAACAAGCAGCTCG
GAATCAGGCACCCCCACCTTGGAGGACCAGATCCGAGGATACAGGCGCTCCGCAAGGTTACGACACCAGGC(1566)</td></tr>
<tr><td></td><td>2473|1321</td><td>(2367)ACGAGGCAGAGGACGAGGAGGATCAGGAAGTACAGGTGATAGGAGCTACCACCATTGAGGAGCCAGAGA
TGGAGTACCCAACTAGGCTCGAAGAAGTTATGGGCAAG(2473) |
(1321)CTCACTAAGGAGTTCGGAAAGGTCAACTTAGGGAAAGGAAAGGGGATAGAAGGAGCAGTCTCATCCAGAG
ACAAGAACTTCTACGTCTGGAAGAACCCCTTCAATCAATACAATGAGCAGAAGCCAAACAAGGCTCCAGGCTCCGC
CCGCAACTGAAAGGGCAACAAGCAGCTCGGAATCAGGCACCCCCACC(1513)</td></tr>
<tr><td></td><td>7268|plant</td><td>(7103)ATCGAGTAAGCTCAATAGCCAACCTTCTCTTGTATTGTATTAACTTTCTTGAGAAATAAAGAGTTTATTCAGTT
CTATCTTCCGCATTGTCTCTGAATTTTATGTTCTTCTAAGTGTTCAAGAGCGCACTTAAGAAGTTTTGATCCATGTTTT
CGGGCCATTTTTC(7268) |
(plant)TGGTAACATCGACTAAATATTGTTGAAATAATCTTCCACCTTTTAACAAAGTATCTGTACTTTTATTTCTGTCA
TGGATTTGATATGCAACAAAACTTCGCATTGACATTTTTTTTCTTTTGATTCCTTCATTT(plant)</td></tr>
<tr><td>endoRYNV
-CU2</td><td>2217|5319</td><td>(2087)TGGGCAGGTGGAGACAGCAATATGATCATCACCAGATCTCTGGTGGGACGATTGACTAACACCAGCATGAC
CAACTTTGAGTACCGGATAGAACAAGTCACAGACTACTTGGCGAGCAATGGAGTCGCATG(2217) |
(5319)ACTACTCAGAAGAAGGAGTAAGCACGCTGAGAAACCACAAGCTGCTACAAGAACTGAAAGAGCAGGGATA
CATCGGAGAAGAGCCCATGAAGCACTGGGCAAAGAATGGGA(5429)</td></tr>
<tr><td></td><td>plant|1</td><td>(plant)GAATTTAATTAGAGAAACAATTTAAACCATTCAAATTTCATTACAATGCGTGTATATGTGAAAATAGGAAAA
CTAGCAATAATTGGCACAAGGGGAGTTCCATAATTA(plant) |
(1)TGGTAGATTGTACTTTAGCTCTATTCAACATGTTAGCTTAAACTTTTCATTTCTTGTCGAGAAACCCAAGTTTCAGA
CTAGAACCTTGAGTTCCTCTCCTTTTTCAGGGAGGAGAGAGTAGTGAGCATTGCAAAACTTCTGCAACTTTCAAACC
CCCATGAAAGCTTTCC(279)</td></tr>
<tr><td></td><td>plant|188</td><td>(plant)TAACATCCCATCCTTGTTAAAAATACTTGTTATTCCTTGCCTGATATTTTCTGTATTTGAACTGTACAATCTGCT
GTCCTGTTATTTGCCTAGCCAC(plant) |
(188)TTCTGATTTTCACTAGTTCAAACCTACTGCTTAAACTGCAGGCTTAGGCGTCGAAGCGAAGTACCCTTGTAGCC
GTTAGCTGGAGTGCGTTAGGCGTTGATTGGGGAAAACGGACGTAAAGAAGCAGCAGCAACTAGGCAAGAAAACT(
335)</td></tr>
<tr><td></td><td>5863|plant</td><td>(5738)GGCAACGCCAAAATCTTCAGCAAGTTCGACTTGAAGTCTGGTTTTCACCAGGTACTAATGGATGAAGAGTCC
ATCCCATGGACGGCGTTTGTCACACCTGTAGGCTTCTACGAGTGGAAGGTGATG(5863) |
(plant)TCAGAAAATTGGCCGAACATAGGATCAAAAACTTTCTAAGTGCGCTCTTGAACACTCAGAAACTCATAAATT
CAGAGAAAATTGCGGAAGATAAGATGAACAAAAGCTCTTTATTTCTCAAGATGTTAATACAATTCTAGAGAAGGTT
GGCTATTGAGCTTACTCGATGCCTA(plant)</td></tr>
</table>

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448838; this version posted June 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

<table>
<tr><td>endoRYNV

-CU3</td><td>4550|4423</td><td>(4434)AGATTGTTCTTCAGTTTTCCGCAATATCTCTCAGTTTTTATGAGTTCTTAAGTGTTCAAGAGCGCATTTAAGAA
TTTTTGATATTGCGGAAAACTGAAGAACAATCTTTATTTCTCT(4550) |
(4423)TGAACGTTAATACAATACAGGAGAAGGTTGGCTATTGAGATTGCTCGATGCCTAAGCCTTCAGACTACACTC
CCTTATATAGGAAGTAGGCTAAGCAAAACGACAGACAACTACTAGCTCGGGTGCGCTTTCTGGGCCCCATCGACAG
CTAATAAGTTCTGCTTTACAGCTTTGCTAAAGCT(4242)</td></tr>
<tr><td></td><td>plant|1</td><td>(plant)TGCTAGCGTGCCAATATGATGAACTGAAGCTATGAACCATTTGCTAGTTTGTGACTGAATTGTGGTACTTGA
TTTGGTTGTAGGTTTGTATTAGGATGGGTATGCCAGGTAGGGGAAACTCTTGCAGTGAATAGCTACAGCAAGAGGA
AGTTGCAGCAAAATGCAAGACTCTTGCAGTGAATAGCTACAGCAAGAGGAAGTTGCAGCAAAATGCAAGAGGACC
TCTTGT(plant) | (1)ACAACGGCTGTATGCTAGTCCTACCGCAGGACTTCACCTTAGTCCCTGACGTGATCAACA(60)</td></tr>
<tr><td></td><td>plant|1492</td><td>(plant)AAACAAAAACAAAAATGTATGGTCCATTAGCCGAAATCAATTCTGCTGTAGGCATCATTTTGCACTGTGGGA
TCTGACTGATCTTGTCAAGGTCCTCTTGTGATAGCTCCCAGTCAAACAC(plant) |
(1492)TCCGATTTGGAACAACGATTCCCACCGATCACATCCAAAACCTTGAAAATGTAGCAAGGATGATCGAACAAT
GGAAGGAAACCCCCAGGGTAATCATCAAAGAAACAGCTGAGAGCAGCAGCAACACCATTGGAGCCCTCTTAGCAG
AAGAAGGAATAGAGGAGCTAGCCGCAGCTGTA(1670)</td></tr>
<tr><td>endoRYNV

-BS</td><td>7602|6317</td><td>(7445)CCAAGCGCACCCGGGCTAATTTTCTCTTGTCTTTTAGCATAAGCCCCCTTCCTATATAAGGAAGCTAAGTTAG
AAGGCTTAGGCATCGAGCAACCTCAATAGCCAACCTTCTCTTGTAATATCAGTATTCAAGAAATTAGTATTCAAGAA
ATTAGTTG(7602) |
(6317)CTATCCCCATGTTCACTTGTCTTGCTATACAGAGGCCCAAGCAAGGTTCCGCATTTGGGAATGTAGTTTCTGG
CGTAGTTGAGGACTCCTAGCCAGCTCCTTAATCCTTTGAGGGTC(6199)</td></tr>
<tr><td></td><td>plant|1</td><td>(plant)TCAAATTTTGTTGAAAATTTGTAGAAGTGATCTACTCATGAATATTTTAATACTGAACAGTTGATTTGTCGAG
ATGTAATCGAAAAATAGATATCACAACTATTAATCAAAATAATACTCATTTAAGTGAGATTTTTTC(plant) |
(1)TGGTATCAGAGCTTTAGCTCTCTTCATTATGTCAGCTTAAACACCTTTTTTCGTGTCGAGAAACCCAAGTTTCGGA
TCTGAACCTTGAGTTCCTCTCTTTTTCAGGAGGGTAGTGAGTAAGCCCAGCTTTGCAACTTTCAAACCCCCAGGAAA
ACTTTCCC(161)</td></tr>
<tr><td></td><td>plant|1777</td><td>(plant)AATTAAGTGATCGATTATCAGACCGAGAATTAATCTATTCATGGTTCAATAATCAGTCCGATAATATATATAT
ATCAATATATGCACTGATCAAACAATTAA(plant) |
(1777)TTCAGGTACATTGAAGAATCCTCTTACCAAAGGCTAGCACGAGAAGGCATGCAGTTTATACATGTAGGCATG
GCCATGGTAAGAATACAGATGCTGCACAGGACTGACGCA(1887)</td></tr>
<tr><td>endoRYNV

-PH1</td><td>335|5859</td><td>(452)AGTCGCCCGTTCAGATCTTTCCCAGGAGCTACTCTCTTCTTCGAAGGCAGAGATTTCGTGTGGTTGCATCAAAA
CCACCAAGATCTAGCCGCTTACTAGCTTTCCGGCCGGTGATCTG(335) |
(5859)TCACCCTGGTCAAATGGCTTCCAGAAGAACTGAAGGATCTCGCGGCCGAGCTAGCCAAGAAAGAAGGCAA
GACCTCCCTGAAGGGAGAGGTGCAGGAGGAGATCTCCTGTTTTCTCAGAACTGCCC(5984)</td></tr>
<tr><td></td><td>1373|1814</td><td>(1215)GGGACTTCCAACGGCAACTAGATCCGGATGCCGAGCTCTCTCTCAGCAGAAGGGGAAGAGCAAACCTGGTA
CCAGCAGAGGTACTACACACACTGAAAGAAGAAGGACCATCCGAGCCAGAAGGAAGGCCAGAAGGAGAGGACGA
GAGCACACATTATG(1373) |
(1814)ATCCCCGAAGAAGAAGAAGAAGTCTACCTCAAATATGAGGCAGAAGACGAAGAGGAGGATCAGGAGCTTC
AAGTGATTGGAGCCACCACCATAGAAGAGCCAGAAATGGAATACCCAACAAGGCTCGAGGAAGTTATGGGC(1954
)</td></tr>
<tr><td></td><td>plant|1</td><td>(plant)AAGGACGAGCTGATATTCAAGTGAAATAAAGATTGTTCATCATTTTCTTCCGCATTTTCTCTGAGTTTCATGA
GTTCTTAAGTGTTCCAGAGCGCACTTTCGAATATTAGATCCATGTTTTTCGGACCCCATTC(plant) |
(1)TGGTATCAGAGCTTTAGCTCTCACCATGGTAGCTTAAACACCTTTTTCTTGTCGAGAAACCCAAGTTTCAGATCTG
AACCTTGAGCTTTCTCTCT(95)</td></tr>
<tr><td></td><td>6631|plant</td><td>(6542)GCTTAGCCTACTTCCTATATAAGGGAGTATAGTCTGAAGGCTTAGGCACAGAGCAATCTCTTTAGCCAACCTT
CTCTTGAGTTGTATTAA(6631) |
(plant)TATTCAACCTTCCCTAAACCCTAAACCCTAAACCCTAAACCCTAAGCACACCCATGGTGAGAGCTAAAGCTCT
GATACCAGAATGGGGTCCGAAAAACATGGATCTAATATTCGAAAGTGCGCTCTGGAACACTTAAGAACTCATGAAA
CTCAGAGAAAATGCGGAAGAAAATGATGAACAATCTTTATTTTACCAAGTGAATCTGTTT(plant)</td></tr>
<tr><td>endoRYNV

-PH2</td><td>2201|2432</td><td>(1990)ACAATTGGTATACATAATACCAAACGCAATGATGTCAATACACGATTTTTACAATTGTATACAGGTCAGCGTG
CAAACCCGAGGCTACGGGACAGGATGGGCTGGAGGAGACAGCAGTATGATCATCACTAGATCACTGGTGGGGCGT
CTCACCAACACCAGCATGACAAACTTCGAGTACCGGATAGATCAAGTCACAGACTATCTAGCAA(2201) |</td></tr>
</table>

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448838; this version posted June 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

<table>
<tr><td></td><td></td><td>(2432)AGCACGCACTACGTGCTCGTCTTCAGACATAACTCCTATGAGACCAACCTCAGAGGAGAAAGGAGGCCAAG
GCAGAACGAACTTTC(2517)</td></tr>
<tr><td></td><td>7091|plant</td><td>(6970)GCAGAGATGGACGAGTCTAGAAGACTGGCCAAACAACGCCGAGACAAAGTCTTCGACGACGCAGGGCAAA
ACATCTGCGACACGGTCTACATCACCGGTGTCGACCTCGCCGCCGCCAAGGC(7091) |
(plant)ACGCGGCCACATTTGCAGCCTTGGTGGTGATAGTTTTGACTGTTGCTTGCAATGCTCCACTTGCTAAGTTGCA
ACATAAGTTTCAGAGTAAGCTTGTGGTGTCACAAGATGAAAAAGGCTCAAGGCTAGTTCTGAAGCTCTTGTAAACA
TGAAAGTATTAAAATTATATTCATGGGAA(plant)</td></tr>
<tr><td></td><td>3094|plant</td><td>(2924)AGTGGCATGGATAGCGAGTCCAGCATCACAGGTGGAGGATTCATACCACCAAGCCCAGTACCAGGAGCGC
AGGGATACCCACCTGCCACAGGAGCGTCTGCTTCAACCATAGGGCCAGCAGATCTGCAAGGATGGGGAGGACGAT
TGCCAAGGAGCAGATCGCCGATAGGA(3094) |
(plant)AAATGGCCCAACATATAGATAGGAAAATGGCCCAACATATAGATCGAAAAATGTCTAAGTGCGCGCTTGAA
CACTTAGAAACTCATAAAAACTCAGAGAATATGCGGAAGATAAGATGAATAAACTCTT(plant)</td></tr>
</table>

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448838; this version posted June 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

377 Table 4. Primer pairs used for Rubus yellow net virus lineage-specific SYBR Green PCR validation.

Lineage	Primer name and sequence	Amplicon size (bp)	Melting temperature (ºC)
endoRYNV-CU1	endoCU1-6376F: CACTGTAGTAAGATGGCTTCCC endoCU1-6475R: GTTCCTCTGTACCTCCTTGTTC	100	80.99
endoRYNV-CU2	endoCU2-2415F: CGGAAGAAGAAAGCACACATTAC endoCU2-2527R: CTCTGGCGTGAACTCTGAAA	113	79.94
endoRYNV-CU3	endoCU3-2450F: CACAAACTCAGGGACGTCTATC endoCU3-2575R: ATCCCTGGCAAGGTGTATTG	126	79.15
endoRYNV-BS	endoBS-4668F: CCTCCAGAACATCGAGAATGTC endoBS-4767R: GCTAGGCATTTCGTCACCTATTA	100	77.57
endoRYNV-PH1	endoPH1-2633F: GAGCTGTCGGAGAAGCTATTT endoPH1-2725R: TATCGTCAACCCTGGGTATCT	93	77.83
endoRYNV-PH2	endoPH2-3892F: GTGCTACCTTTGCGGAATAGA endoPH2-3989R: CTAAGCCCATGCCATTGAAATAC	98	79.02
epiRYNV-BS	epiBS-2604F: TCCTACGAGGTAAGCCTAAGAG epiBS-2715R: CCGAGTTCGAGAGTTGGTTTAG	112	79.81

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448838; this version posted June 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

10. Figures

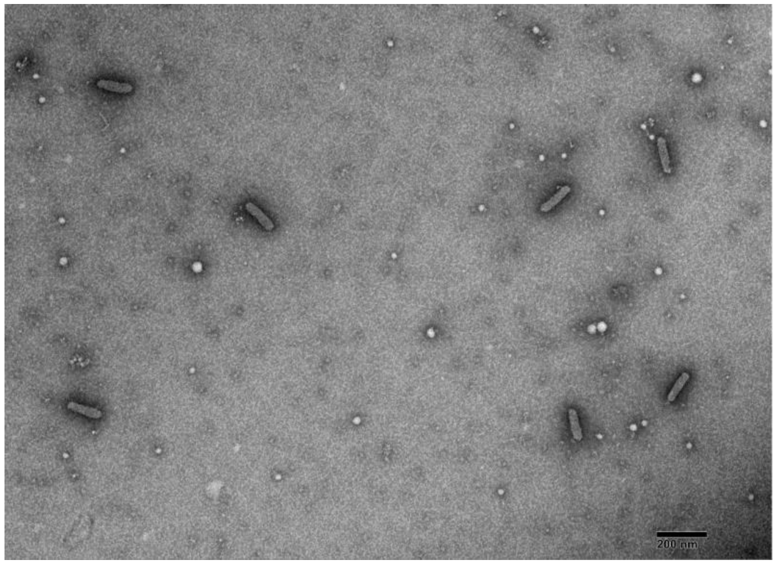


Figure 1. Bacilliform particles of rubus yellow net virus from 'Baumforth's Seedling A' (epiRYNV-BS) captured using antibody against sugarcane bacilliform virus and stained in phosphotungstic acid.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448838; this version posted June 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

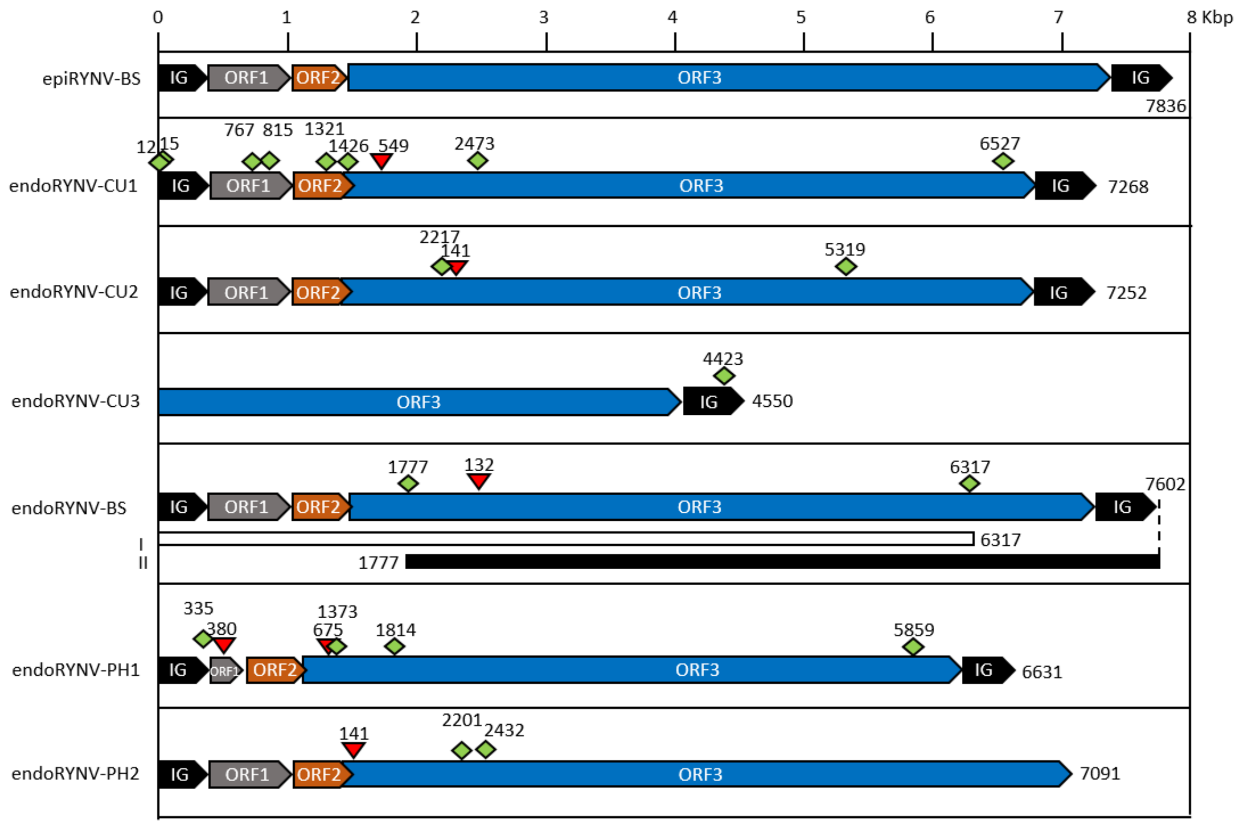


Figure 2. Overview of endogenous rubus yellow net virus (endoRYNV) lineage sequence structures in red raspberry. The sequences are represented with intergenic region (IG) in black; grey, orange, and blue boxes indicating ORF1, ORF2 and ORF3 of the virus, respectively. For simplicity, ORF4 and ORF6 embedded within ORF3 are not illustrated. Break points and deletions in the genomic sequence are represented in green diamond and red triangle shapes, respectively, with nucleotide location on top. The two fragments of endoRYNV-BS that were found in high-quality 'Joan J' genome are shown in white and black bars.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448838; this version posted June 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

```
endoRYNV_CU1        ACGTCTAGTGAAGTGACGCAATGAATGACTTCACAATTGCCAATGTCGTCACTGCTTACG
RYNVCaKF241951      ACGTCTAGTGAAGTGACGCAAGGAATGACTTCACAATTGCCAATGTCGTCACTGCTTACG
                    ********************* **************************************

endoRYNV_CU1        ACTTGGAACTTATCGTTTTGTGTCGGCAGCATCTCTTAGCTGTCATTTGTGTGTAAGTGC
RYNVCaKF241951      ACTTGGAACTTATCCTTTAGTGTCGGCAGCATCTCTTAGCTGTCATAAGTGTGTAAGTGC
                    ************** *** ***************************  ************

endoRYNV_CU1        GCCGGTAGTGCGCTGTGTCAGGATAAGGAATCTTATCTCCTTATCTTTTTGC--TTTGTT
RYNVCaKF241951      GCCAGTAGTGCGCTGTGTCAAGATAAGGAATCTTATCTCCTTATCTTCTTTCCCTTTGTT
                    *** **************** ************************** ** *  ******

endoRYNV_CU1        -AAAGCTA--GCTGTAAAGCAGATTCTCTTAGCTGTCGATGGGGCCCAGAAAGCGCACCC
RYNVCaKF241951      TAAAGGTAAAGCTGTAAAGCAGGACTAATTAGCTGCAGGTCA--TCAGGTTTGCGGTTGT
                     **** **  ************      *******  * *     *  *   ***

endoRYNV_CU1        GAG---CTAGTGGTCATCTGTCTTTTTGCTTAGCCTCTCCCCTATATAAGGGAGCTCAGT
RYNVCaKF241951      GGAACTCCTGCAGCTGACTGGTGAGCTCTTCGACTTTTCTAGTGAGGAAAGCGTTGTGCT
                    *     *  *  *    ***      *  *   * * **   *    ** *        *

endoRYNV_CU1        TAGAAGGCTTAGGCATCGAGTAAGCTCAATAGCCAACCTTCTCTTGTATTGTAT---TAA
RYNVCaKF241951      CATAGAGATGTCGTATAAGCTCATCTTTGTTGTGGTATTGCCACTTAGCTGCGTCCTTGG
                     * *  * *   * **    * * **   * *      * *   *    **  *   *

endoRYNV_CU1        CTTTCTTGAGAAATAAAGAGTTTATTCAGT--TCTATCT-TCCGCATTGTCT-CTGAATT
RYNVCaKF241951      TGTCCTTTGCGGCTATAAGCTTGATCCCATGATCCATGTATGCGCAAAGGGGACAGAGGT
                      * ***      ** *   ** ** *  *  ** ** * * ****  *    * **  *

endoRYNV_CU1        TTATGTTCTTCTAAGTGTTCAAGAG----CGCACTTAAGAAGTTTTGAT-CCATGTTTTC
RYNVCaKF241951      TGAGGTTGCAGGTGGTGCAGGTGACTCTTCGCCCATGAGGCGTTTCGTCGCTGCATATGC
                    * * ***       ***      **      *** * * **  **** *   *    * * *

endoRYNV_CU1        GGGCCATTTTTC------------------------------------------------
RYNVCaKF241951      TGCATATCCTCCCTTCGATCGGCACTTCTTGTGTATCACTCCAGGCGTGAGTGCATTCTT
                     *   **  * *

endoRYNV_CU1        ----------------------------------------------
RYNVCaKF241951      GCTGTGCCTTTGGTATCTCCTTCCTTCTTCTCCAGGAAGGTTTTTC
```

Figure 3. ClustalW alignment of the 3' intergenic regions of the endogenous rubus yellow net virus (endoRYNV) lineages Cuthbert and Canada (endoRYNV-CU1 and RYNV-Ca respectively. Although the two lineages share 100% nucleotide (nt) identity at the reverse transcriptase and RNaseH domain, their 3' IGs share 29% nt identity.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448838; this version posted June 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

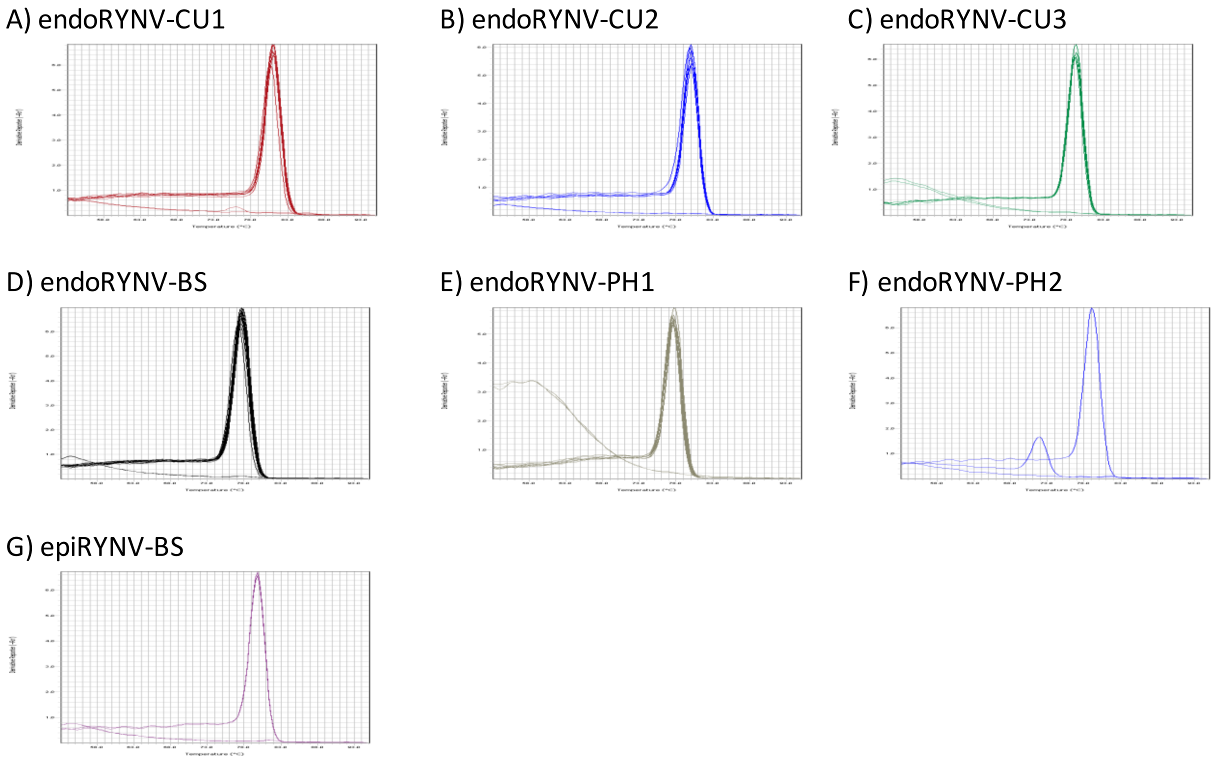
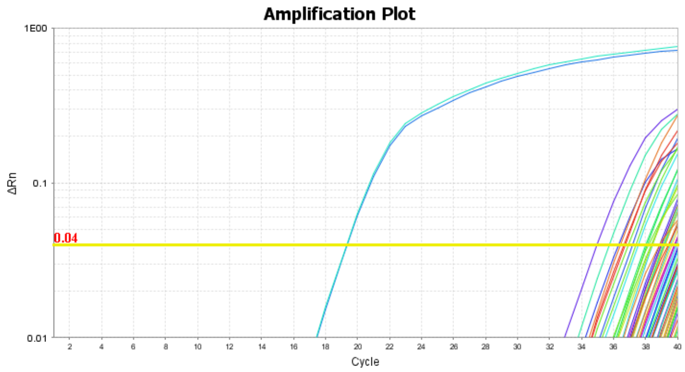


Figure 4. Melting curves for the lineage-specific SYBR Green PCR validation assays detecting the endogenous and episomal rubus yellow net virus (endoRYNV and epiRYNV respectively) lineages from the raspberry cultivars with corresponding endo/epiRYNVs. A) endoRYNV-CU1; B) endoRYNV-CU2; C) endoRYNV-CU3; D) endoRYNV-BS; E) endoRYNV-PH1; F) endoRYNV-PH2; G) epiRYNV-BS.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448838; this version posted June 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

(A)



(B)

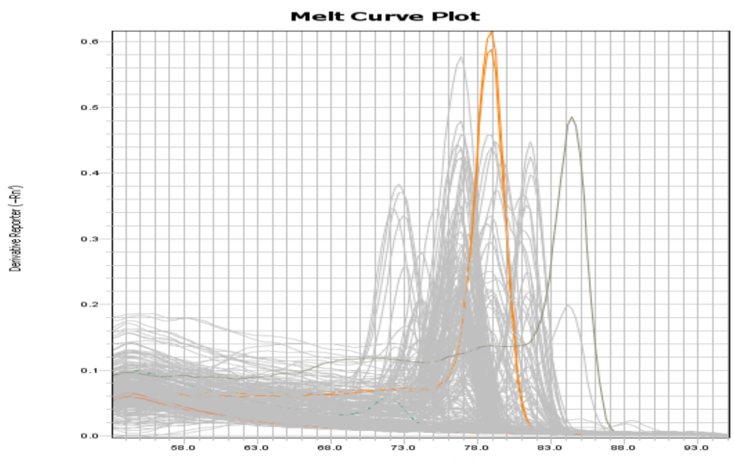


Figure 5. Detection of the rubus yellow net virus from ‘Baumforth’s Seedling A’ (epiRYNV-BS) lineage using the epiBS-2604F/2715R assay against a panel of 271 public and proprietary genetics. The plants were as described in 2.5. A) Amplification plot (the epiRYNV-BS positive control from Corvallis, Oregon, with two replicates were labeled in blue and light blue); B) Melting plot (the two positive control replicates were labeled in orange with unique melting point compared to all other non-specific amplifications, labeled in grey). epiRYNV-BS was detected at Ct=19, whereas non-specific amplifications initiated after Ct=35 but their melting points were different from that of the positive control.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448838; this version posted June 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

11. Supplemental data

>endoRYNV-CU1

TGGTATCAGAGCTTTAGCTCTCACCATGGTAGCTTAAACACTTTCCTTCTTGTCGAGAAACCCAAGTTTCAGACCAGAACCTTGA
GTTTGCTCTCTTTTTCGGAGGGAAGAGGAGTGAGTGTCTGTGTCAAAACCTTGAAAGATCAAACCCCCATGAAAACTTTCCTCA
CGGTACCATAAGTTTTCTATCCTTCACTAGTTTGAACCTACTGCTCAAACTGCAGGCTTAGGCGTCGAAGCGAAGTACCCTTGTA
GCCGTTAGCAGGAGGCGTTAGGCGTTGATTGGGGAAAACTGACGTAAAGAAGCAGCAGCAACTAGGCAAGAAACCTGACGG
GTAGATCACCGGCCGGAAAGCCAGTAAGCGGCTAGATCTGGGCAGTTTTGATGCAACCTCACGAAATCTCAGCCTTCGAAGAA
GAAAGCAGCTCTTGGGAAAGGTCTGAACGGGCGTATCGACAAGACTTTTTATTCAGAAATCTCAGAACGTATCCACGTTGGGA
GGCAAATCAGAAAACACCCTCTCTAGACTTTCCTTGCTACCACTTCAACACAACAACCGGACCACCAGTCCACCGCACTCTCTGC
AGACAAGAGAACAGTAAGGATTTACCATTTCTGGTAAACACCCTGTTCGATCTCAACATCACCGAGATCCACAATCAGGCGATT
CTGGACGATAAGATCTCCAGACTCACCCAGTACCTGACAACAAAGGTTGGTTCACTACCAACAATCCCGGAGGATTCACCCCTC
CTGGACCAAGCCACAATATCCTTAGATCTTCAAGCCCTCAAGGCAGATCTGAAGGAAATCAAGGCCACACAGTCAGCCCTGAA
GCTAGGCTTCGCACAACTGCAGGAGGCAGTTCAGCTGATCATCACAAGGGAAAACGATCCCAAACCAATCGAAGCAGCTACT
GCACAAGTAGCCGAACAGCTGAGGAAGCAGCTTATTGAGGTCAAGTCCGTCCTCGAGGAGACCAAGAAGATCGCGAGATCTC
TGTCCCCCGACGGATGAACCCTAGGTGGCAGGATACTGCAACCAAGGAAACCTACCTCAAAGCCATACAAGCTACCTCATCTCT
CACCTCCAACAACACAGGTCTAGGCTTCATCGAGCCACATACCTACACCGGAGGACAGCTATCTACCAACCTAGCAAAACAGA
ACAACACGCTCATCCAGCTGTTAGTTCAGGTGCTAGAAAAGAACCTCGACCTCGAGCAGGCAATTGTCAACCTCACAGCTCAG
GTCACAAGGCTAGAAAAGACCGTCTCGGAGAAAGACACAGTCAAACTCCCAGAAAGTGTCCTCAACGACCTCACTAAGGAGTT
CGGAAAGGTCAACTTAGGGAAAGGAAAGGGGATAGAAGGAGCAGTCTCATCCAGAGACAAGAACTTCTACGTCTGGAAGAA
CCCCTTCAATCAATACAATGAGCAGAAGCCAAACAAGGCTCCAGGCTCCGCCCGCAACTGAAAGGGCAACAAGCAGCTCGGA
ATCAGGCACCCCCACCTTGGAGGACCAGATCCGAGGATACAGGCGCTCCGCAAGGTTACGACACCAGGCGCAGCGAGCAATG
AGAAGGACCTTCAGTAGGGACTTCAGAAACACCATAGAACGGCAACTAGACCCAGATGCCGAGCTTTCCCTCAGCAGAAGAA
GGAGAGTAAACCGAGTACCAGCAGAGGTATACAGATCCGGAGAATGGGATTTACAACCCAGCAGGATCGTGGCACCACTAGC
AGTCCCAACAGAAGCAAGGCTTAGCCAAAACAGGAATGGCAATATAAGCCTCAGATTCACCGACTTCCGAGATCAGAGGATC
GTGGAGGAAGGAGAACCATCTGAGCCAGAAGGAAGGCCAGAAGGAGAAGATGATAGCACGCACTATGTGCTCATGTTCAAC
CACTCAAGGTGGGACACCTTAGGGCAACCAAGCGGGAAATATGATTACATGGTGAGGTATGATGCACCAGAACCTACCGCAT
GGCCAACATCCAACATCGGATGGGATGATGATAAGCCACCCAAACCGCCAAGCCCTACAAAAGGATCTTTTGAGGTAAACCTC
AAAGGAGAGAAGAAACTAAAAGAGAAGGAACTCGCGGAGTTCACACCGGAGACAGATCTGGTGAGCCAGTGGTTAAGTCAG
CTGTCAACATCCGCACATAATAGCGGAGCCTCAAGTTCAGACGAAGAACCAAAGTTCGACGAGGCAGAGGACGAAGACGATG
TGTACAACCAGCAAACCTGGCAAAAGGAAGACAAGGAGAAAAGAGACCTGGAACTACAGGGGTGGAAACCCACCGGGAGAC
CAGGAATCTACGAGATGATCCCCGAAGAAGAAGAAGAAATCTACCTCAGGTACGAGGCAGAGGACGAGGAGGATCAGGAAG
TACAGGTGATAGGAGCTACCACCATTGAGGAGCCAGAGATGGAGTACCCAACTAGGCTCGAAGAAGTTATGGGCAAGCTCAA
AAACGTGAGCATGGAAAAACTGTTCCCAGTAAGCGGAATGGACAGCGAATCCAGCATCACAGGTGGAGGATTCATCCCACCA
AGCCCAGTGCCAGGAGCACAAGGGTACCCACCAGCAACTGGAGCATCCGCGTCCACCATTGGACCAGCAGACATGCAAGGAT
GGGGAGGACGGCTACCTCGGAGCAGGTCGCCTATAGGCTATGGCAGACCCCAACAACCGTGGTCACTGCCCTCAGCACAGTC
TGATAACGGCTGCATGCTAGTCCTTCCACAGGACTTCACCCTAGTCCCCGACGTAATCAACAGATGGGAATCCATCACAGTCAA
CCTCATCAACAAGATGATGTTTGATTCCCTACAGGACAAGGCGGACTACGTAGAAAACCTCCTTGGAGAAAGAGAAAAGGAG
ACATGGATGACATGGAGAATGCAGTACGAGGAAGAGTACAGGCAACTCCTCACCATGAGCGGAGACGTAAGGAACCTTACTG
CCGCAGTCAAAAGGGTCTTTGGAGTACACGACCCGCACACAGGATCAGTACACATCCAGAATCAAGCGTACGCAGAGCTGGA
ACGCCTCTACTGCAAAAGAACGGATGATGTGATCCCCTTCCTCTACGACTACTACCAGTTAGCAGCCAAGTCAGGAAGGATGT
GGCTCGGACCTGAGCTATCTGAGAAGCTGTTCAGAAAGCTTCCACCGGAGATAGGCCCAACAATAGAGCAGGCCTATAAAGA
CAGGTATCCAGGCCTCACGATTGGAGTTTTGGCAAGGGCCAATTTCATCCTGGAATATCTACAAAACGTCTGCAAGCAAGCAG
CGTTGCAAAGGTCCCTAAAAAGCCTGAGCTTCTGCAGAAACATGCCAGTACCAGGGTACTACGAGAAGAAGCAATACGGCAT
CAGAAAGGCTAAAACCTATAAAGGAAAGCCTCACCCTACCCACGTGAAAGTCATCAAAAACAAGTACAAGCACACATCTGGGA
AGAAGTGCAAATGCTACTTATGTGGGATAGAAGGCCATTACGCCAGGGAATGCCCAAAGAAAGTGGTGAAGCCACAAAGAG
CGGCATACTTCAATGGCATGGGACTAGACGACAACTGGGATGTCGTGTCCGTCGAGCCCGGAGAATCAGATGACGATGAAAT
CTGTAGCATCTCCGAGGGAGAAAACGCTGGAGGAATGCATGAGCTTATGGCATTCAAGACTCAACTCCCATACCCAGTGGAGT
ACGAAGCCAGCACACCACAGTTCCTGATGCCATGGACACAAGTAACAGTGGAAAGAAGCGAAAAACCTTCCTGGAGAAGAAG
GAAGGAAATCCCGAAGGCACAACAGGATTGTACTCACACCTGGAGTGACACACAAGAAGTGCCTATCGAGGGAAGGATATGC
AGCATATGCAGTGATGAAACCCCTCATGGGCGAAGGATCACCTGCACCACCTGCAGCCTTAACCTCTGTCCGCTTTGCGCTTAC
ATGGATCATGGGATCAAGCTTATAGCCGCAAAGGACACCAAGGACGCAGCTAAGTGGCAATACCACAACAAAGATGAGCTTA

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448838; this version posted June 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

TACGACATCTCTATGAGCACAACGCTTTCCTCACTAGAAAAGTCGAAGAGCTCACCAGTCAGCTGCAGGAGTTCCACAACCGCA
AACCTGATGACCTGATCAGCTTAGCGGATGACTTGGAGGACGTGTCCATTCTGGACAACGCCTCAAAAAGGGGGAAGGAGAA
GGAATCTTTCCAATTCGGAACAACGATTCCCATCGACCACATCCAAAACTTGGAAAACGTGGCAAGGATCATCGAGCAATGGA
AGGATACCCCCAAGGTAATCATCAAAGAAACAGCTGAAAGCAGCAACAACACCATCGGAGCCCTCTTAGCAGAAGAAGGAAT
AGAGGAGCTAGCCGCAGCTGTAGACACGGCATACACAGAAATGCCAAAAGGAGGATTGAACAAGCTCTACAACACCATTGTT
GAGTTTGTAATACCCCAGGAAAAGGGGGCACCCACCAGGTTCAGGGTAAGAGCTGTAATAGACACAGGATGCACCTGTACAT
GTATCAACAGCAAGAAAGTCCCCAAAGAAGCCCTGGAGGAAGCGAAGTACCAGATGAACTTCGCAGGAGTAAATTCCACTGG
AGAAACGAAGCTTAAAATGAAGAACGGTAAGATGATCGTGTCTGGAAGCGATTTCTATACACCGTACATTGCAGCCTTCCCAA
TGGAACTACCAGACGTAGACATGCTCATCGGCTGCAACTTCTTGCGAGCCATGAAGGGAGGAGTCAGGCTTGAAGGTACTGA
AGTGACGATCTACAAGAAAGTCACCACAATCCAAACAACCCTGGAGCCCCAAAAGATATCTCTGCTCCGCGCAGAAGCAGAAG
TCGGAGAAGAGATCGAGCGTATGTACTACGCAAATGACTACTCTGAAGAAGGAGTCAGTCGCCTGAGAAACCACAAACTGCT
GCAGGAACTAAAAGAACAAGGCTACATAGGCGAAGAGCCAATGAAGCACTGGGCGAAAAACGGGATCAAGTGTAAGCTTGA
CATCAAGAACCCAGACATAGTAATCAGCAGTAAACCCCCGGATGCTGTCTCAAAGGAGACGAAGGCACAATACCAGCGGCAC
ATTGACGCTCTCCTGAAGATCAAAGTAATCCAGCCAAGCAAGAGCAAGCACAGAACCGCAGCCTTCATCACAAACTCGGGCAC
AACCGTTGACCCGATCACAAAGAAAGAAATCCGAGGAAAAGAAAGGATGGTGTTCGACTACAGAAGTCTGAACGACAACACC
CACAAAGACCAGTATACTTTGCCTGGGATCAACACCATCATATCGGCAATCGGCAATGCGAAGATCTTCAGCAAATTTGATCTG
AAGTCTGGATTCCACCAAGTATTGATGGACGAAGAATCCATCCCGTGGACCGCATTTGTCACACCAGTAGGGTTCTACGAGTG
GAAGGTAATGCCTTTCGGACTCGCAAACGCTCCGGCCGTCTTCCAGAGAAAGATGGACCAGTGTTTTGCAGGAACCTCAGAGT
TCATAGCCGTCTACATCGACGATATCCTGGTCTTCAGCAAGACCTTGAAGGAGCACGAAAAGCACCTGAGCATCATGCTTGGG
ATATGTCGAGACAACGGCCTGGTTTTGTCACCAAGCAAGATGAAGTTAGCAGCAACCGAGATCGACTTCTTGGGAGCCACCAT
TGGTGACGGAAAGATTAAACTCCAGCCTCACATAATCAAGAAGATAGCTGAGGTGGACGATGAATCTCTAAAAACCCTCAAAG
GGTTGAGAAGTTGGTTGGGAGTTCTCAACTATGCCAGGAACTACATCCCGAAGTGCGGAACACTCCTAGGCCCACTATACAGC
AAGACCAGTGAGCATGGAGACAGAAGGTGGCATGCTTCGGATTGGGCCTTAGTAAAGAAGATCAAGAGCCTGGTCCAAAATC
TCCCAGACCTCAAACTGCCCAGTGAGGAGGCCTATATGATCATCGAGACAGATGGTTGTATGGAAGGATGGGGCGGAGTCTG
TAAGTGGAAGCCCAACAAAGCAGACTCAGCTGGCAAGGAAGAAATCTGCGCTTACGCAAGCGGTAAGTTCCCAACGGTGAAA
TCTACCATTGACGCAGAAATCTTCGCGGTAATGGAGTCCTTAGAAAAATTCAAAATTTTCTACATGAATAAGGACGAGATCACC
ATCAGAACCGATTGCCACGCCATCATCACCTTTTACGAAAAGTTAAACGCCAAGAAACCTTCTCGGGTAAGGTGGTTAGCTTTT
TGTGATTATATAACAAACTCAGGGGTGAAGATGAAGTTCGAACACATCAAAGGCAAAGATAATCAGCTCGCTGACAATCTTAG
TCGCCTTACCCAACTCATCACTGTAGTAAGATGGCTTCCCAAGGAACTAGCGGAGCTCACGGCCGAACTGGTCAAAGGAAGGG
ACGAAGCCCTGGTGAACAAGGAGGTACAGAGGAACATCTCATGTTTTCTCGAGACTGCCCTCCTCCAAGCGGAGAAATCCGTG
ACTACTCGCCCATCAGAGCCGCACCATGTACTATGGCGGAGATGGACGAATCCCGAAGAGTGGCCATGCAGCGAAGAATCAA
GGTCTTCGACGATCTTGCACAAAACATCAGCGACGCCGTATACATCACAGGCATCGACCTCGCCGCCGCCAAGGCACGGGCAA
CCAGAGACAACTGGTACAATGACGTCACCCCAGCATTGGAAGAACGAGCAGCTGCAGCATGGAGACTCATGGCGGCCTACTC
AGACTTCGCCACGTGGAAGGACGTAAACGTCTAGTGAAGTGACGCAATGAATGACTTCACAATTGCCAATGTCGTCACTGCTT
ACGACTTGGAACTTATCGTTTTGTGTCGGCAGCATCTCTTAGCTGTCATTTGTGTGTAAGTGCGCCGGTAGTGCGCTGTGTCAG
GATAAGGAATCTTATCTCCTTATCTTTTTGCTTTGTTAAAGCTAGCTGTAAAGCAGATTCTCTTAGCTGTCGATGGGGCCCAGAA
AGCGCACCCGAGCTAGTGGTCATCTGTCTTTTTGCTTAGCCTCTCCCCTATATAAGGGAGCTCAGTTAGAAGGCTTAGGCATCG
AGTAAGCTCAATAGCCAACCTTCTCTTGTATTGTATTAACTTTCTTGAGAAATAAAGAGTTTATTCAGTTCTATCTTCCGCATTGT
CTCTGAATTTTATGTTCTTCTAAGTGTTCAAGAGCGCACTTAAGAAGTTTTGATCCATGTTTTCGGGCCATTTTTC

>endoRYNV-CU2

TGGTAGATTGTACTTTAGCTCTATTCAACATGTTAGCTTAAACTTTTCATTTCTTGTCGAGAAACCCAAGTTTCAGACTAGAACCT
TGAGTTCCTCTCCTTTTTCAGGGAGGAGAGAGTAGTGAGCATTGCAAAACTTCTGCAACTTTCAAACCCCCATGAAAGCTTTCC
TCACGGTACTATAAGCTTTCTGATTTTCACTAGTTCAAACCTACTGCTTAAACTGCAGGCTTAGGCGTCGAAGCGAAGTACCCTT
GTAGCCGTTAGCTGGAGTGCGTTAGGCGTTGATTGGGGAAAACGGACGTAAAGAAGCAGCAGCAACTAGGCAAGAAAACTG
CAAATCAGATCACCGGCCGGAAAGTCAGTAAGCGGCTAGATCGGGGCAGAAGCGGATGCAACCTCACGAAATCTCATCCTTT
GAAGAAGAGAGCAACTCTTGGGAAAGGTCTGAACGGGCGTATCGACAAGACTTCCTATTCAGAAATCTTAGATCCTATCCGAG
GTACGAAGCAAACCAGAAATCACCTTCCTGTGATTTCCCTTGTTATCACTTCAACACAACCACTGGACCTCCAGTCCACCGCACG
ATCTGCAAGCAGAAGAACAGTGAGGATTTACCCTACCTGGTAAATACACTGTTTGATCTCAACATCACTGAGATCCACAACCAG
GCAGTCCTAGACGACAAAGTCTCGAGACTCACACAGTACCTGGTCAAGAAAGTCGGAACGCTACCGACAATCCCGGAGGACT
CACCCCTCCTGGACCAAGGTTCCATAAACCTAGATCTGCAGGCCCTAAAAGCAGATCTGAAAGAAGTCAAGGCAACCCAGTCC
GCACTGAGGTTAGGCTTTGAACAGCTAAGAGAAGCTGTCCAGCTAATCATTGCCCGCGAAAACGATCCGAAGCCCATCGAAGC

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448838; this version posted June 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

TTCTACGGCACTCGTAGCGGAGCAGCTGAGGAAACAACTGATAGAGGTGAAATCTGTCCTCGAGGAGACAAAGAAGATCGCC
AGATCTCTCTCCCCTGACGGATGAACCCAAAGTGGCAGGAGACCGCCGCAAAAGAAACCTATCAGAAGGCTATCCAGGCAAC
CTCCTCGCTCACATCAAACAACACCGGTCTAGGGTTCATAGAACCTCATACTTACACCGGAGGACAACTGTCCACAAACCTAGC
AAAGCAGAACAACACGCTCATACAGCTGTTGGTCCAAGTGCTAGAAAAGAATCTGGACCTTGAACAGGCGATAGCAAACCTAT
CAGCACAGGTCACCAGACTCGAGAAGACCGTCGCAGAGAAGGACACGGTCAAACTCCCGGAAAGCGTCCTGAACGACCTCAC
CAAAGAATTCGGGAAGGTCAATCTCGGGAAAGGAAAGGGTTTAGAAGGTATTGTCTCTTCCAAAGACAAAAACTTCTACGTTT
GGAAGAACCCTTTCAACCAGTACAATGAGCAGAAGCCAAACAAGGCTCCAGGCTCCGCCCGCAACTGAAAGGGCAACAAGCA
GCTCGGACTCAGGAACCCCCACCCTGGAGGACCAGATCCGAGGATATAGACGCTCTGCAAGGTTAAGACACCAAGCACAGCG
AGCAGTTCGAAGGACCTTCAGCAGGGACTTCAGGAATACTATAGAAAGGCAACTGGACCCAGATGCCGAACTCTCCCTTAGTA
GGAGAAGAAGAGCGAATCTAGTACCAGCAGAGGTACTCTATGCTCACAACGGTTCGGAACCAGCGAATCGTGTGTACGAGCA
TTACAGTGAGCTCGGGGCCCACATAGTCGACAGGCAACAAGACTTCCGGTATATAGAGGAAGCCTCCTACCAAAGGCTAGTCA
GGGAAGGCATGCAGTTCATTCATGTCGGGATGGCCATGGTCAGGATCCAGATGTTGCATAGAACGGATGCGGGGATATCCGC
ATTAGTGGTGTTCAGAGACACCAGGTGGAGTGATGACAGGCAAGTCATCGGGAGCATGTCCGTGGACATGACAAGGGGAGC
ACAGTTGGTGTATATAATACCCAACGCAATGATGTCAGTACACGATTTCTACAATCGTATACAAGTAAGTGTGCAGACCCGAG
GCTACGGGACAGGATGGGCAGGTGGAGACAGCAATATGATCATCACCAGATCTCTGGTGGGACGATTGACTAACACCAGCAT
GACCAACTTTGAGTACCGGATAGAACAAGTCACAGACTACTTGGCGAGCAATGGAGTCGCATGCATACCCGGACAGAAGTGG
GACGTAGCCAATCGATCCGGAGAATGGGAACTTCAACCCAGCAGGATCACAGCGCCAGTCATGGCACCTACTGAGGCAAGGT
TGAGCCAGAACAGAAATGGCAGCATAAGCCTCAGATTCTCAGATTTCCGAGATCAGAGAATCACAGAAGAAAGGCCAGCTGA
AGAAGAGGGCAGGCCGGAAGAAGAAAGCACACATTACGTGCTCATGTTCCGTCACAGCTCCTACGAGACCAATCTCAGAGGA
GAAAGGAGGCCAAGGCAGAGTGAGCTTTCAGAGTTCACGCCAGAGACAGATCTGGTAAGCCAGTGGCTGAGCCAACTATCA
GCCTCAGCACACAACAGTGGAGCATCTAGTTCAGAAGACGAACCGCCCAGGTTTGATGACTCTGAAGAAGATAGTGACAACA
CCTATAATGAGAAGACCTGGCAAAGAGAAGACCAGGAAAGGCGAGATCTGGAGCTGCAAGGATGGAAGAAGACCAGCAGA
CCAGGAATCTATGAGCTGATCCCAGAGGAAGAGGAAGAAATCTACCTCAGGTATGAAGATGAAGAGGATCAGACAACAACAC
AGGTAATCGGGGCAACAACCATGGAGGAACCAGAGATGGAGTACCCTACCAGGTTGGAGGAAGTGATGGGAAGACTCAAAA
ACGTTAGCATGGAGAAGCTGTTCCCAGTTAGCGGCATGGATAGCGAGTCAAGCATCACAGGGGGAGGCTTTATCCCACCAAG
CCCAGTACCAGGAGCGCAGGGTTACCCACCTGCAACAGGAGCATCCTTTGGCTCGACAATTGGGCCAGCAGATCTACAAGGAT
GGGGAGGTAGGCTGCCACGCAGCAGATCACCACTTGGATATGGCCGCCCACAGCAGCCATGGTCATTACCATCCGCCCAATCT
GAGAACGGATGTATGCTAGTCCTCCCACAGGACTTCACTCTAGTCCCAGATGTAATAAACAGATGGGAATCCGTAACCGTCAA
CCTTATCAACAAGATGATGTTCGACTCCTTACAAGATAAGGCGGATTACGTGGAAAATCTCCTAGGAGAACGGGAGAAGGAG
ACATGGATGACCTGGAGGATGCAATATGAAGAAGAGTACAAGCAGCTCCTCACCATGGCAGGAGACGTGAGAAATCTCACGG
CAGCAGTCAAGAGGGTCTTTGGAGTGCATGACCCCCATACAGGATCAGTCCACATACAAAACCAGGCATATGCAGAACTGGA
GAGGCTCTACTGCAAGAGAACGGACGACGTGATCCCTTTCCTATATGACTACTACCAACTGGCGGCGAAATCTGGAAGAATGT
GGCTAGGGCCAGAATTATCAGAAAAGCTGTTCAGAAAGCTGCCACCAGAGATAGGCCCAACGATTGAGCAGGCCTATAAAGA
CAGGTATCCAGGGCTGACAATTGGAGTCTTGGCGCGAGCCAACTTCATACTGGAGTACCTACAGAATGTCTGCAAGCAGGCA
GCGCTACAGAGGTCTTTGAAGAGCCTCAGCTTTTGTAGAACCATGCCGGTGCCAGGGTATTATGAGAAGAAGCAATACGGCAT
CAGGAAGGCAAAGACCTACAAAGGGAAGCCACACCCCACACATGTAAAGGTAATCAAGAACAAGTATAAACACTCCGCAGGG
AGGAAATGCAAATGTTACCTTTGCGGGATAGAGGGTCACTACGCCAGGGAATGCCCAAAGCAAGTGGTTAAGCCACAGAGAG
CAGCATTCTTTAATGGCATGGGCCTTGATGATAACTGGGACGTCGTGTCCGTGGAGCCAGGAGAGACAGATGATGATGAGAT
CTGCAGCATCTCAGAAGGCGAGGGCACCGGGGGGATGAACGAGCTACTTGCATTCAAGACACAGCTCCCCTACCCAGTAGAA
TATGAGGCGAGCTCATCTCAGCAGTTCCTGCCATGGATTCAGGTAACTGTACAAAAGAGTGAGAAACCCTCATGGAGAAGAA
GGAAAGAGATCCAACCGGCACAACAAGAATGCGCTCACCTCTGGAGCGATACACAAGAAGTGCCTATAGAAGGAAGAATTTG
CAGCATCTGCAGTGACGAAACCCCTCTGGGAAGAAGGATCACCTGCACTGCCTGCAACCTAAATCTGTGTCCTCTTTGCGCATA
CATGGACCATGGTATCCAACTGATAGCCGCAAAGGACACCAGAGATGCAGCAAAATGGCAGTACCACAACAAGGATGAGCTA
ATCCGGCATCTCTATGAGCATAATGCTTTCTTAACCAGAAAGGTGGAGGAGCTGACAAGCCAGCTGCAACAATGGCAAAACCG
CAAACCAGAAGACCTTATCAGTCTTGCGGATGAGCCAGAGGACGTGTCCATCCTGGACAACGCCTCAAAAAACCGGGGGAAG
GAGAAGGATTCTTTCCAATTTGGAACAACAATCCCAGTCGATCATATACAAAACATCGAGAATGTTGGCAGAATGATCGAACA
ATGGAAGAACACCCCCAATGTCACGGTCAAAGAAGTGGCAGAATCAAGCAGCAACACCATAGGAGCACTCCTTGCAGAAGAA
GGAATTGAAGAGCTCGCGGCTGCAGTAGATACGGCATACACTGAAATGCCGAAAGGAGGGTTGAACAAACTCTACAACACCA
TCGTAGAATTTGTGATACCACAGGAAAAGGGGGCACCCTCTAGATTCCGGGTCAGAGCTGTCATTGACACAGGATGCACTTGT
ACCTGCATCAACAGCAGGAAGGTCCCGAAGGAAGCTCTCGAGGAAGCGAAGTTTCAGATGAATTTCGCCGGGGTAAACTCCA
CGGGAGAAACAAAGCTGAAAATGAAGAATGGCAAAATGATAGTATCAGGGAGTGATTTCTATACGCCATACATCGCGGCTTT
CCCCATGGAACTACCAGATGTGGACATGCTCATTGGCTGTAACTTCTTACGAGCCATGAAAGGAGGAGTCAGACTTGAAGGAA
CAGAAGTAACCATCTACAAGAAGGTTACCACAATCCAGACGACTCTAGAGCCACAGAAGATCTCTCTACTCCGTGCAGAGGCA

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448838; this version posted June 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

GAAGCAGGGGAAGAAATTGAGAGACTCTACTATGCCAATGACTACTCAGAAGAAGGAGTAAGCAGGCTGAGAAACCACAAG
CTGCTACAAGAACTGAGAGAGCAGGGATACATCGGAGAAGAGCCCATGAAGCACTGGGCAAAGAATGGGATCAAGTGCAAA
CTTGATATCAAAAATCCCGATATAGTTATCAACAGCAAGCCTCCGGATGCAGTGTCTAAAGAGACGAAGGCGCAGTACCAAAG
GCACATCGATGCTTTACTGAAGATCAAGGTGATTCAACCCAGTAAAAGTAAGCACCGAACAGCAGCCTTCATCACCAACTCGG
GTACCAGCATAGACCCGATAACAAAGAAGGAGATCCGAGGAAAGGAGAGGATGGTCTTCGACTACAGGAGTCTGAATACCCA
CAAGGATCAGTACACCTTGCCTGGGATAAATACCATCATCTCTGCAATTGGCAACGCCAAAATCTTCAGCAAGTTCGACTTGAA
GTCTGGTTTTCACCAGGTACTAATGGATGAAGAGTCCATCCCATGGACGACGTTTGTCACACCTGTAGGCTTCTACGAGTGGAA
GGTGATGCCTTTCGGACTCGCGAACGCTCCAGCAGTATTCCAAAGGAAGATGGATCAGTGTTTCGCAGGAACCTTTGAGTTCA
TAGCCGTATACATAGACGACATCCTAGTGTTCAGTAAGACACTCAAGGAACATGAGAAGCATCTCAGTATAATGCTGGGAATA
TGTCGTGACAACGGCCTAGTACTGTCGCCAAGCAAGATGAAACTGGCAGCAACAGAAATTGACTTTCTGGGAGCAACCATTGG
CGATGGGAAGATTAAGCTCCAACCTCATATCATCAAGAAGATAGCTGAGGTGGACGACGAGTCTCTTAAGACTCTGAAAGGAC
TAAGAAGCTGGTTGGGAGTGCTGAACTATGCCAGGAATTACATCCCAAAATGTGGAACCCTCCTCGGCCCACTCTACAGTAAG
ACGAGTGAGCACGGTGATAGAAGATGGCACGCTCAGGATTGGGCCTTAGTCAAAAGGATTAAAAGCCTGGTCCAGAACCTTC
CGGATCTGAAGCTACCAACTGAAGAGGCCTATATAATCATTGAGACCGATGGTTGTATGGAAGGATGGGGCGGAGTCTGCAA
ATGGAAGCCCAACAAGGCCGACTCAGCAGGAAAAGAAGAGATTTGCGCGTACGCAAGTGGGAAATTCCCAACGGTCAAATCA
ACAATAGACGCGGAAATATTTGCGGTCATGGAATCCTTGGAAAAATTCAAGATATTTTACATGAACAAGGAGGAAATCACCAT
CAGGACCGACTGCCACGCCATCATAACTTTTTACGAAAAGTTGAACGCGAAGAAGCCTTCCCGCGTAAGGTGGCTTGCTTTCTG
CGACTATATAACCAACTCCGGGGTCAGAATGAAGTTCCAACATATCAAAGGCAAGGATAATCAGTTAGCTGATAATCTCAGTC
GCCTAACCCAGCTGATCACCGCAGTGAGATGGCTACCAGAGGAAATGGCAGGAATCGCGGCAGAGCTCACCAAAGAGAGGG
GAATGAGCTCCGCTCTGAGCACAGTTCAGGAGAGCCTCTCAGGCTTTCTCAAAGCTGCCCTCCTCCAAGTCGAGAAGTCCTCGA
CTACACACCTGTCCGAGGAGCGCCCTGCTCTTTGGCAGAGATGGAAGAGTCTAGAAGACTGGCCAAACGGCTCAGAGACCGA
GTCTTCAACGAAGTCAGCCGGAGCATCAGCGACACTGTATTCATCACCGGCAGGGACCTCGCAACGGCCAAAGCATACGCCAC
CAGGGACAACTGGTATGGCGACCTCGTCCCACTCTTGGAGAAGAGAGCAGCAGCTGCATGGAAGCTCCTCGCTGCACACGCA
GAATTCTCCACATGGAAGGACGTAGACGTCTGAGAGGTGTGACGAAAAAGATGACCTCACAATTGCCAAGATC

>endoRYNV-CU3

ACAACGGCTGTATGCTAGTCCTACCGCAGGACTTCACCTTAGTCCCTGACGTGATCAACAGATGGGAGTCCATTACAGTCAACC
TCATCAACAAGATGATGTTTGACTCCCTGCAGGACAAGGCGGATTATGTGGAAAACCTCCTTGGAGAAAGGGAGAAAGAAAC
ATGGATGACATGGAGAATGCAGTACGAAGAAGAATACAAGCAACTCCTCACCATGAGCGGGGACGTGCGAAACCTAACGGCC
GCCGTCAAAAGGGTCTTCGGAGTACACGACCCACACACAGGATCTGTACACATCCAAAATCAGGCGTATGCGGAGCTAGAAC
GCCTCTACTGCAAAAGGACGGATGATGTGATCCCCTTCCTCTACGACTATTATCAGCTAGCAGCTAAGTCCGGAAGGATGTGG
CTCGGACCCGAGCTATCTGAGAAGCTGTTCAGGAAGCTTCCACCGGAGATAGGTCCAACCATAGAGCAGGCCTACAAAGATA
GATACCCAGGACTCACAATCGGAGTCTTAGCAAGGGCTAACTTCATCTTGGAGTATCTGCAGAATGTGTGCAAACAAGCTGCA
CTACAGAGGTCCCTCAAGAGCCTGAGCTTCTGCAGAAACATGCCAGTACCAGGATACTACGAGAAGAAGCAGTACGGCATCA
GAAAGGCCAAAACCTACAAAGGTAAGCCTCACCCAACCCACGTGAAAGTCATCAAGAACAAGTACAAGCACACCTCAGGGAA
GAAGTGCAAGTGCTACTTGTGTGGGATAGAAGGCCATTACGCAAGAGAATGCCCAAAGAAGGTGGTAA AACCACAGCGAGC
GGCATACTTCAATGGCATGGGATTGGACGACAATTGGGACGTCGTATCAGTCGAACCCGGAGAATCAGACGATGACGAGATC
TGTAGTATCTCCGAGGGAGAAAATGCTGGAGGAATGCACGAGCTGATGGCATTCAAGACTCAACTCCCCTACCCAGTGGAGT
ATGAAGCCAGCACACCACACCTTCTTATGCCTTGGACTCAGGTGACAATAGAAAAGAGTGAGAAACCATCCTGGAGGCGAAG
GAAGGAGATCCCAAAGGCACAGCATGACTGTACTCACACCTGGAGTGATACACAGGAAGTGCCAATTGAAGGAAGGATATGC
AGTATATGCAGTGACGAGACACCCCACGGGAGGAGAGTCACCTGCACCACCTGCAGCCTCAACCTCTGTCCTCTTTGCGCATAT
ATGGATCACGGGATCAAGCTGATAGCTGCAAAGGACACGAAAGATGCAGCTAAATGGCAGTATCATAACAAAGATGAGCTCA
TACGACACCTCTATGAGCACAACGCTTTTCTCACAAGAAAAGTAGAGGAGTTAACCAGCCAGCTGCAGGAGTTCCACAACCGC
AGGCAGGATGATCTGATCAGCCTGGCAGACGACCTGGAGGACGTGTCTATCCTAGACAACGCCTCAAAAAGAGGGAAGGAG
AAGGAATCATTCCGATTTGGAACAACGATTCCCACCGATCACATCCAAAACCTTGAAAATGTAGCAAGGATGATCGAACAATG
GAAGGAAACCCCCAGGGTAATCATCAAAGAAACAGCTGAGAGCAGCAGCAACACCATTGGAGCCCTCTTAGCAGAAGAAGG
AATAGAGGAGCTAGCCGCAGCTGTAGACACGGCATACACTGAAATGCCAAAGGGAGGACTAAACAAGCTCTACAACACCATT
GTTGAGTTTGTGATACCTCAGGAGAAAGGGGCACCCACAAGATTTAGGGTAAGGGCTGTAATCGACACAGGATGCACATGCA
TTAACAGCAAGAAAGTCCCCAAGGAAGCACTGGAGGAGGCGAAGTACCAGATGAACTTCGCAGGAGTAAACTCCACGGGAG
AAACAAAACTCAAGATGAAGAACGGTAAGATGATCGTATCTGGAAGCGACTTCTATACACCGTACATTGCAGCTTTCCCGATG
GAACTGCCAGACGTAGACATGCTCATAGGATGCAACTTCCTAAGGGCCATGAAAGGAGGAGTCAGACTTGAAGGGACAGAA
GTGACAATCTACAAAAAAGTCACCACCATCCAGACTACCTTGGAGCCACAGAAAATATCTCTCCTACGCGCAGAAGCAGAAGT

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448838; this version posted June 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

AGAGGAGGAGATTGAGCGCATGTACTACGCAACAGACTACTCTGAAGAGGGAGCTAGTCGTCTGAGAAATCACAAACTGCTG
CAGGAGCTGAAGGAGCAAGGCTACATAGGCGAGGAGCCAATGAAGCACTGGGCAAAAAATGGAATCAAGTGCAGACTTGAC
ATCAAGAACCCAGACATAGTAATCAGCAGCAAACCCCCGAATGCAGTCTCAAAGGAGACAAAAGCACAGTACCAGCGGCACA
TAGACGCCTTGCTAAAGATCGGAGTGATCCAACCAAGCAAGAGCAAGCACAGAACCGCTGCCTTCATCACAAACTCAGGGAC
GTCTATCGACCCTGTCACTAAGAAGGAAATCAGGGGAAAAGAAAGGATGGTGTTCGACTACAGGAGTCTGAACGACAACACC
CACAAGGACCAATACACCTTGCCAGGGATCAACACCATCATATCTGCGATCGGGAATGCGAAGATCTTCAGCAAGTTTGATCT
GAAGTCTGGATTCCACCAAGTGCTGATGGACGAAGAATCCATCCCGTGGACCGCATTCGTGACACCAGTAGGGTTCTACGAGT
GGAAGGTAATGCCCTTCGGACTCGCAAACGCTCCGTCCGTCTTTCAAAGGAAGATGGACCAGTGTTTTGCAGGAACCTCAGAA
TTCATAGCCGTCTACATCGACGACATCCTGGTCTTCAGCAAGACCTTAAAGGAACATGAAAAACATCTCAGCATCATGCTTGGG
ATATGTCGAGACAACGGCCTGGTATTGTCACCAACTAAAATGAAGATAGCTGCAACCGAGATAGATTTCTTGGGAGCCACTAT
AGGTGACGGAAGAATCAAACTCCAGCCTCACATAATCAAGAAAATAGCAGAGGTGGACGATGAGTCACTCAAGACCCTCAAG
GGATTGAGGAGTTGGTTGGGAGTCCTCAACTACGCGAGGAATTACATTCCAAAGTGCGGAACTCTTCTCGGCCCACTATACAG
CAAGACTAGTGAGCACGGAGACAGGAGATGGCACGCATCAGATTGGGCCTTGGTCAAGAAGATCAAAAGCCTGGTCCAGAAT
CTCCCAGACCTAAAACTGCCCAGCGAAGAAGCCTACATGATCATAGAGACTGATGGATGCATGGAAGGCTGGGGCGGTGTTT
GCAAATGGAAGCCCAACAAAGCAGACTCAGCAGGCAAAGAGGAAATCTGTGCCTACGCAAGTGGGAAGTTCCCCACGGTGA
AATCAACCATTGACGCAGAAATCTTCGCGGTAATGGAGTCCTTAGAAAAATTTAAAATTTTCTACATGAACAAGGACGAAGTCA
CCATCCGAACTGACTGCCACGCCATCATTACCTTCTACGAGAAGTTAAACGCCAAGAAACCTTCTAGGGTAAGGTGGTTAGCTT
TTTGCGACTATATAACGAACTCAGGGGTGAAGATGAAGTTCGAACACATCAAGGGTAAGGATAATCAGCTCGCTGACAACCTC
AGTCGCCTTACTCAACTAATCACAGTGGTCAGATGGCTTCCGGAAGAGCTGAAGGATCTCGCGGCCGAACTAACCAAAGGAG
AAGGCAAGACGCCAACGAAGAAGGAGACACAGGAAGAAATCTCCTGTTTTCTCAAAGCTGCCCTCCTCCGAGCAGAGAAATC
CGCGACTACTCACCCATCAGAGCCGCACCATGCACTATGGCAGAGATGGATGAATCCCGAAGACTGGCTCTGCAACGAAGAA
ACAAGGTCTTCGACAGCATCGCACAAGGCATCAGCGACGCAGTGTACGTCACAGGTGTCGACCTTGCCGCCGCAAAGGCACG
AGCCACCAGGGATAACTGGTACAACGACGTCACCCCGGCGCTGGAACAAAGAGCAGCTGCAGCATGGAGACTCATGGCAGCC
TACTCAGACTTCGCCACGTGGAAAGACGTGAACGTCTAGTGAAGTGACGCAAGGGATGACTTCACAATTGCCAATGTCGTCAT
TGCTTACGACTTGGAACTTATCAGTTAGTGTCGGCAGCATCTTCCAGCTGTCAGATATTTGTGTGTAAGTGCGCCATTAGTGCG
CTGTGTCAGGATAAGGAATCTTATCTCCTTATCTCTTCTTTAGCTTTAGCAAAGCTGTAAAGCAGAACTTATTAGCTGTCGATGG
GGCCCAGAAAGCGCACCCGAGCTAGTAGTTGTCTGTCTTTTTGCTTAGCCTACTTCCTATATAAGGGAGTGTAGTCTGAAGGCT
TAGGCATCGAGCAATCTCAATAGCCAACCTTCTCCTGTATTGTATTAACGTTCAAGAGAAATAAAGATTGTTCTTCAGTTTTCCG
CAATATCTCTCAGTTTTTATGAGTTCTTAAGTGTTCAAGAGCGCATTTAAGAATTTTTGATATTGCGGAAAACTGAAGAACAATC
TTTATTTCTCT

>endoRYNV-BS

TGGTATCAGAGCTTTAGCTCTCTTCATTATGTCAGCTTAAACACCTTTTTTCGTGTCGAGAAACCCAAGTTTCGGATCTGAACCT
TGAGTTCCTCTCTTTTTCAGGAGGGTAGTGAGTAAGCCCAGCTTTGCAACTTTCAAACCCCCAGGAAAACTTTCCTCACGGTATT
ATAAGTTTTCTGACCCTCACTAGTTCAACCCTACTGCTTGAACTGCAGACTTAGGCGTCGAAGCGAAGTACCCTTGTAGCCGTT
AGCCAGGTGGGCGTTAGGCGTTGATTGGGGAAAATGGACGTAAAGGAGCAGAGGCAACTAGGCGAGAAAACTGAGGACCA
GATCACCGGTCGGAAAGCCAGTAAGCGGCTAGATTGGGGCATAGCTGATGCAACCTCACGAGATCCTTTCCTTCGAAGAAGA
GAGCAACTCTTGGGAACGATCTGAACGGGCGAATCGACACGACTTCCTATTCAGAAATCTCAGATCTTATCCAAGATACGAGG
CTAACCAGAAATCACCTTCCTGTGATTTCCCTTGTTACCATTTCAACACCACCACCGGACCACCAGTCCACCGCACAATCTGCAA
ACAAAAGAACAGTGAAGACTTACCTTACCTGGTAAACACTCTGTTCGATTTATCTATAACAGGTATTCACAACCAGGCTGTCCT
CGACGACAAGGTCTCGAGACTCACACAGTACCTGGTAAAGAAAGTCGGCACACTCCCGACTATCCCGGAGGATTCACCCCTCC
TGGACCAGAACTCCCTGAGCTTAGATCTGCAGGCCCTAAGGGCAGATCTCAAGGAGGTCAAAGCAACTCAGTCCGCACTAAAA
CTCGGTTTTGAACAACTCCGAGAAGCAGTTCAATTGATCATCTCCAGGGAGAACGATCCAAAGCCAATCGAAGCCACGACTGC
ACTTGTGGCAGAACAGCTGAGGAAACAGTTGATTGAAGTCAGATCTGTGCTTGAGGAAACCAAGAAGGTCGTCAGATCGCTG
TCCCCTAACGGATGAACCCTAAGTGGCAGGACACCGCCACAAAGGAAACCTATCAGAAGGCACTTCAAGCAACAGCATCCCTC
ACCTCAAACAACACAGGCCTAGGGTTCATTGAACCACACACCTTCACGGGAGGACAGTTGTCCACTAACCTAGCCAAACAGAA
CAACACGCTCATCCAGCTGTTGGTTCAGGCGCTAGAGAAGGTCCTCGACCTTGAACAAGCCGTCGCAAACCTGACAGCTCAGG
TCACAAGGCTTGAAAAGACAGTCGCAGAGAAGGACACGGTAAAGCTACCGGAAAGCGTCCTCAACGACCTCACCAAGGAATT
TGGGAAGGTCAACATCGGAAAAGGAAAAGGGATCGAAGGCACAAACACTTCCAAAGACAAGAACTTCTACGTCTGGAAGAAT
CCCTTCACCCAGTACAATGAGCAGAAGCCAAACAAGGCTCCAGGCTCCGCCCGCAACTGAAAGGGCCACATCTAGCTCGGACT
CAGGTACCCCTACCTTGGAGGACCAGATCCGCGGCTACAGACGCTCGGCAAGGTTACGACACCAAGCCCAGCGAGCAGTTCG
GAGAACCTTCAGCAGGGATTTCCGAAACACAATAGAACGGCAACTTGACCCAGATGCCGAACTCTCCCTCAGCAGAAGGAGG

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448838; this version posted June 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

AGAGCAAACCTAGTACCAGCAGAGGTACTATACGCACACAACGGTTCAGAACCAGTAAACAGGGTGTATGAACACTACAGTGAGATCGGAGCTCACATAGTAGACAGGCAGCAGGACTTCAGGTACATTGAAGAATCCTCTTACCAAAGGCTAGCACGAGAAGGCATGCAGTTTATACATGTAGGCATGGCCATGGTAAGAATACAGATGCTGCACAGGACTGACGCAGGAATTTCTGCACTAGTAGTCTTTCGGGACACTAGATGGAGTGATGACAGGCAAGTCATTGGAAGCATGTCGGTAGACATGACAAGAGGTGCACAGCTGGTATATATCATACCAAATGCGATGATGTCAGTACACGACTTCTATAACCGCATACAAGTCAGCGTGCAAACCAGAGGATACGGTACAGGCTGGGCGGGAGGAGACAGCAACATGATCATCACCAGATCACTGGTGGGAAGGCTGACCAACACCAGCATGACAAATTTTGAGTACCGAATAGACCAGGTCACAGACTACCTTGCAAGCAATGGAGTCGCATGCATCCCAGGACAGAAATGGGACGTGGCCAACAGATCAGGAGAGTGGGAACTTCAGCCAAGCAGGATAGTGGCTCCACTCATGACACCTACAGAGGCAAGGCTGACCCAGAATAGAAGCGGCAGCATCAGCCTTAGGTTCACAGACTTCCGTGATCAGAGGATCGTGGAAGAAAGGCCGGCAGAAGAAGAAGGCAGGCCAGAGGAAGAAAGCACACACTATGTGCTCATGATCAGCCACAACTCAAGCTCCTATAGGACAAACCTCAAGGGAGAAAGAAGGCTGAGACAGGAAGAACTCGCTGAGTTCATGCCAGAATCAGATCTGGTAAGTCAATGGCTGAGTCAGTTATCAGCTTCAGCACACAACAGTGGAGCCTCAGACTCAGAGGACGAACCACCAAAGTTCGATGAGACTGACGAAGAAGCAGATGATGACACGTATAATCAGCGAACCTGGCAGAAAGAGGACCAGGAGAGAAGACAACTGGAACTGCAGGGGTGGAAGAAAACCAGCAGGCCAGGAATCTATGAAATGATCCCTCAGGAAGAAGAAGAAGTCTACCTCCGCTATGAGGCAGAATCAGAGGAAGAGGATCAACCCACACAGATCTTCGGAGCTACAAAGGTGGATGAACCCGAAATGGAGTACCCCACAAGGCTCGAAGAAATGATGAACAAGCTCAAAGGAGTGAGCATGGATAAGCTTTTCCCAGTGACCATGGAATCTGAGTCCAGCATAACTGGAGGAGGCTTCATACCACCAAGCCCAGTGCCCGGAGCACAAGGATACCCACCTGCAACAGGGCCAGCCTTTGGATCAACTATAGGCCCGGCAGACTTACAGGGATGGGGAGGAAGACTGCCCAGGAGCAGGTCTCCACTAGGATACGGCAGGCCGCAACAGCCATGGTCCTTGCCATCAGCCCAATCAGAAAACGGGTGTATGCTAGTGCTTCCACAAGACTTCACACTTGTTCCCGACGTGATAAACCGATGGGAATCAGTAACCGTCAACCTGATAAACAAAATGATGTTCGACTCATTACAGGATAAGGCGGACTACGTGGAGAATCTCCTTGGAGAAAGGGAGAAAGAAACATGGATGACCTGGAGAATGCAGTACGAAGAAGAATACAAGCAACTCCTCACCATGAGCGGCGATGTGAGAAACCTCACGGCTGCAGTTAAGAGGGTCTTTGGAGTACATGACCCCCACACTGGATCTGTGCATATCCAGAATCAAGCTTATGCAGAGCTAGAACGGCTGTACTGCAAAAGGACGGATGATGTGATACCTTTCCTGTACGACTACTACCAATTGGCAGCAAAGTCGGGAAGGATGTGGCTTGGCCCAGAACTGTCTGAGAAGCTATTCAGGAAATTACCTCCTGAAATCGGCCCTACAATTGAACAAGCCTACAAAGATAGGTATCCAGGCCTTACCATTGGAGTCTTGGCACGCGCCAACTTCATCCTTGAATACTTGCAGAACGTCTGCAAACAAGCAGCACTGCAGAGGTCCCTGAAGAGCCTCAGTTTCTGCAGAAACATGCCTGTGCCAGGATACTACGAAAAGGAGCAGTATGGCATTAGGAAGGCCAAGACGTACAAAGGGAAGCCCCATTCAACACACGTCAAGGTCATCAAAAACAAGTATAAGCACTCAGCAGGAAAGAAGTGCAAGTGCTATCTCTGTGGCATTGAAGGCCACTACGCTAGGGAGTGCCCAAAACAGGTGGTGAAACCACAAAGAGCCGCCTTTTTCAACGGCATGGGACTCGATGATAACTGGGATGTTGTATCTGTAGACCCAGGAGAGTCAGATGATGACGAGATCTGCAGCATATCCGAAGGAGAAGGAGCAGGAGGAATGAACGAGCTCCTAGCCTTTAAAACACAACTGCCATACCATGTGGACTACGAGGCCACAACATCATCACAGCAGTTAATGCCATGGATACAGGTCACTGTATCAAGAAGTGAAAAACCATCCTGGAGACGACGGAAGGAAGTCCGAAAGGAACAGCAGGATTGTTCTCACCAGTGGAGTGACACGAAGGAAGTGCCACTAGAGGCGAGAGTCTGCAGCATTTGCAGTGATGAAACTCCATTAGGCAGGAGGATCACCTGCGAAACCTGCAATCTGAATCTGTGTCCCATCTGTGCATATATGGATTACGGAATCCAATTAGTAGCAGCTAAAGACACAAGGGACGCTGCAAAATGGCAGTACAACAACAAAGACGAGCTCATCCGCCAACTATATGACCATAACGCTTTTCTAAGCAGGAAAGTCGATGAACTCACCAGCCAGCTGCAACAATACAGAGATCGCAAACCTGAAGATCTCATCAGCCTCGCAGATGACCCAGAGGACATGTCCATACTGGACGATGCCTCAAAAAGACGGGGGAAGGAGCATGAAATCTTCCAGTTCGGAACGACATTGCCCACGGATCACCTCCAGAACATCGAGAATGTCGGGCGAATGATCGAACAGTGGAAAAACAACACCCCCAGAGTCACCATCAAAGAAATAATAGGTGACGAAATGCCTAGCAGCAGCAGCACAATAGGAGCACTCCTCGCAGAAGAAGGAATCGAGGAACTCGCTGTAGCAGTAGACACGGCGTATACTGAAATGCCCAAGGGAGGACTCAACAAGCTCTACAACACGGTTGTGGAATTTGTCATCCCCCAAGACAAGGGAGCGCCAACAAGATTTCGGGTAAGAGCAGTCATTGATACAGGATGCACCTGCACCTGCATAAATAGCAACAAAGTCCCGAAAGAAGCCATGGAAGAAGCAAAGTACCAGATGAACTTTGCAGGAGTAAATTCAGCGGGTAGTACAAGGATGAAGATGAAGAGCGGAAAAAATGATTGTATCTGGGAGTGACTTCTACACCCCGTACATAGCCGCCTTTCCCATGAACCTGCCTGAAGTAGACATGCTCATAGGCTGCAACTTCCTACGAGCCATGAAAGGAGGGGTCAGATTGGAAGGAACAGAGGTGACAATCTACAAGAAAGTCACGACTATTCAGACGACCCTCGAACCACAGAAAATCTCCCTGATGAAAGCAGAGATGGAAGCTGAAGAGGAGATTGAAAGGATATACTACGCCAGTGATCATTCGGAAGAAGGAATGAACAGGCTTAGAAACCACAAGCTGCTCCAGGAGCTTAAGGAGCAAGGATACATAGGGGAAGAGCCCATGAAGCATTGGGCCAAGAATGGAATCAAGTGCAAACTTGATATCAAAAACCCCGACATCGTCATCAGCAGCAAGCCGCCTGATGCTGTATCCAAAGAAACAAAGGCCCAGTACCAGCGACACATTGATGCTCTGTTAAAGATAGGGGTCATCCGGCCCAGCAAAAGCAAGCATCGCACCGCGGCCTTTATCACACATTCGGGCACAAGCATTGACCCGATAACCAAGAAGGAGATACGGGGAAAGGAAAGAATGGTCTTTGACTACCGAAGCCTGAACGACAACACCCACAAGGATCAGTACACCTTGCCAGGCATAAACACAATCATCTCCGCCATTGGAAACGCGAAGGTCTTCAGCAAGTTCGACCTGAAATCAGGATTCCACCAAGTCCTCATGGACGAGGAATCAGTACCATGGACCGCGTTTGTCACACCCGTAGGGTTCTACGAGTGGCTCGTCATGCCATTCGGCCTCGCAAATGCACCTGCAGTATTCCAAAGAAAGATGGACCAGTGCTTTGCAGGTACTTCAGATTTTATTGCTGTGTACATTGACGACATACTCGTTTTCAGCAAGACAATCAAAGAACATGAGC

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448838; this version posted June 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

GTCACCTGAGCATCATGTTGAGCATCTGTCGAGACAACGGCCTGATCCTCTCGCCTAGCAAAATGAAGATCGCCGCCACGGAA
ATTGATTTCCTCGGGGCAACCATAGGTGACGGAAAGATCAAGCTGCAACCGCATATCATTAAGAAGATTGCGGAAGTCGATG
ATGAATCCTTAAAGACCCTCAAAGGATTAAGGAGCTGGCTAGGAGTCCTCAACTACGCCAGAAACTACATTCCCAAATGCGGA
ACCTTGCTTGGGCCTCTGTATAGCAAGACAAGTGAACATGGGGATAGAAGGTGGCACGCACAGGATTGGGCTTTGGTTAGAA
AAATAAAGGCCTTGGTCCAAAATCTGCCAAACCTGCAATTACCCACTGAGGAAGCCTACATCATAATTGAGACTGATGGATGT
ATGGAAGGCTGGGGAGGCGTATGCAAGTGGAAGCCCAACAAGGCTGATTCACCAAGCAAAGAAGAAATATGCGCCTATGCA
AGCGGTAAATTCCCTACCGTAAAATCTACCATTGATGCGGAAATCTTTGCTGTCATGGAATCATTGGAGAAATTCATAATCTTCT
ACATGAACAAAGATGAGATCATCATCCGCACAGACTGCCACGCCATCATCACGTTCTACGAAAAGTTGAACGCAAAGAAACCT
TCTCGGGTAAGGTGGCTAGCCTTTTGTGACTATATAACTAACTCTGGTGTCAAGATGATATTCCAACATATCAAAGGGAAGGAC
AACCAGCTCGCGGATAATCTAAGCCGCCTCACACAGCTCATCACTGCAGTCAGATGGCTCCCAACAGAAATGGCAGGAATCGC
CAGCGAACTCACCAAGGAGAGAGCTCCGAGTCCAGCAATGGACGAGGTACAGAAGAACCTGTCCGGCTTTCTAGAAGCTGCC
CTCCACCAAGTCGAGAAATCCTCGACTATGAACCATTCCGCAGAGCACCATGCTCTATGGCGGACATGGAGGAATCTCGAAGA
GCTGCCAAACAGCTCCGGGACAGAGTTTTCAACGAGGCAGGAAGACAGATCAGCGACATCGTCTTCATCACAGGAAAGGACC
TCGCAGCAGCCAAGGCATACGCTACCAGGGACAATTGGTACGGCGACCTCGTCGGCCTCCTGGAGAAACGCGCCGAAGCGGC
GTGGAAGCTGTTAGTCGCATACTCCGAGTTCTCTACGTGGAAAGTGGATGTCTAGAAGACATGACGTAGGAGCTCACCTCAGT
AATTGCCGGGACGTCACTGCGTGCGTCACGGAACTTATCTTTTAGTGTCGGCAGCACGTATTAGCTGTCACGCATTTTGTAAGT
GCGCCATTAGTGCGCGTGAGTCAGGATAAGGAATCTTATCCTTTGCTTTAGTTAGTGGCAAGCTGTAAAGCAAGCATTATTAGC
TGTCGATGGGGCCCCAAGCGCACCCGGGCTAATTTTCTCTTGTCTTTTAGCATAAGCCCCCTTCCTATATAAGGAAGCTAAGTTA
GAAGGCTTAGGCATCGAGCAACCTCAATAGCCAACCTTCTCTTGTAATATCAGTATTCAAGAAATTAGTATTCAAGAAATTAGT
TG

>endoRYNV-PH1

TGGTATCAGAGCTTTAGCTCTCACCATGGTAGCTTAAACACCTTTTTCTTGTCGAGAAACCCAAGTTTCAGATCTGAACCTTGAG
CTTTCTCTCTTTTTCTAGGGCGAAAGGAGTGTGACCCTGTTCAAAACCTTACCTGATCAAACCCCCATGAAAACTTTCCTTACGG
TACTATAAGTTTTCTGTCTATCACTAGTCCAAACCTACTGCTCAAACTGCAGGTCTAGGCGTCGAAGCGAAGTACCCTTGTAGCC
GTTAGCAGGAGGCGTTAGGCATTGATTGGGGAAAACTGACGTAAGAAAGCAGCAGCAACTAGGCAAGGAACTTGACAGACA
GATCACCGGCCGGAAAGCTAGTAAGCGGCTAGATCTTGGTGGTTTTGATGCAACCACACGAAATCTCTGCCTTCGAAGAAGAG
AGTAGCTCCTGGGAAAGATCTGAACGGGCGACTCGGCAAGACTCTTTATTCAAGAATCTCAAAAACCTATCCACGTTGGGAGG
CGAATCAAAAGACACCCTCTCTCGATTTCCCGTGCTACCACTTCAACACAACATCTGAAGAAACAGCTCATCGAGGTCAAGACA
GTCCTCGAAGAGACCAAGAAGATCGCTAGGTCTTTATCCCCCGACGGATGAACCCTAGGTGGCAGGATACTGCAGCCAAGGA
GACCTACATCAAAGCGATCCAAGCTACCTCATCGCTTACCTCCAGCAACACTGGTCAGGGTTTCATCGAACCCCACATCTACAC
AGGAGGACAGCTATCCACTAACCTAGCAAAACAGAACAACACTATCATCCAGTTGTTGGTTCAAGTGCTAGAGAAGAACCTCG
ACCTTGAGCAGGCAGTCGCCAACCTCACAGCTCAGGTCACGAGACTAGAAAAGGCCGTCGCTGACAAAGACACGGTCAAGCT
CCCTGAGAGTGTCCTGAACGACCTCACAAAAGAGTTCGGGAAGGTTAACTTAGGGAAGGGCAAGGGGATAGAAGGAACCGT
CTCTTCTAGAGACAAGAACTTCTACGTCTGGAAGAACCCCTTCAACCAGTACAATGAACAGAAGCCACACAAGGCTCCAGGCT
CCGCCCGCAACTGAAAGGGCCACGAGCAGCTCGGACGCAGGAACCCCCACTCTGGAGGACCAGATCCGAGGCTACAGACGCT
CTGCACGGTTACGACACCAGGCGCAGCGCACCATGAGAAGGACCTTCAGCAGGGACTTCCAACGGCAACTAGATCCGGATGC
CGAGCTCTCTCTCAGCAGAAGGGGAAGAGCAAACCTGGTACCAGCAGAGGTACTACACACACTGAAAGAAGAAGGACCATCC
GAGCCAGAAGGAAGGCCAGAAGGAGAGGACGAGAGCACACATTATGTGCTAATGTTCAACCATTCCAGCCCCAGATGGGATA
CGCTCGGACAGCCAAGCGGAAAATATGACTACATGGTACGGTATGATGCACCGGAACCAACCGCATGGCCGACAACCAATAT
AGGATGGGATGACGACCCACCAAAGCCACCAAGCCCTAAAGGATCTTTTGAGATCAACCTAAGAGGCGAAAAACGACTAAAA
GAGAAGGAACTCTCAGAGTTCACTCCTGAAACTGACCTAGTCAGTCAGTGGTTGAGTCAGCTCTCCAACTCTGCACACAACAGT
GGAACTTCGAGCTCTGAAGAAGAGCCACAGTTCGACGAGGCAGACGACGAGAACGACGAGTACAACCAGCAAACCTGGCAA
CGAGAGGACCAAGAAAAGAGAGACCTGGAACTACAAGGGTGGAAACCTACTGGTAGACCAGGAATTTACGAAATGATCCCC
GAAGAAGAAGAAGAAGTCTACCTCAAATATGAGGCAGAAGACGAAGAGGAGGATCAGGAGCTTCAAGTGATTGGAGCCACC
ACCATAGAAGAGCCAGAAATGGAATACCCAACAAGGCTCGAGGAAGTTATGGGCAAGCTCAAGAACGTGAGCATGGAGAAG
CTGTTCCCAGTCAGTGGGATGGACAGCGAATCCAGCATAACAGGTGGAGGATTTATCCCACCTAGCCCGGTGCCAGGCGCAC
AAGGATATCCCCCAGCAACAAGCGCATCAGCGTCCACAATAGGACCAGCAGACATGCAGGGATGGGGAGGAAGAATGCCTA
GAAGCAGATCACCTCTGGGCTATGGCAGACCTCAACAACCTTGGTCATTGCCATCTGCACAGTCAGACAATGGCTGTATGCTA
GTCCTGCCACAGGACTTCACACTAGTCCCGGACGTCATCAACAGATGGGAGTCTATCACAGTCAATCTCATCAATAAGATGATG
TTTGACTCTCTCCAAGACAAGGCGGACTACGTCGAAAATCTCCTAGGCGAAAGAGAGAAGGAGACATGGATGACATGGAGGA
TGCAGTACGAGGAAGAGTACAAGCAGCTCCTAACCATGAGCGGGGACGTGAGAAATCTCACCGCCGCAGTAAAACGGGTCTT

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448838; this version posted June 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

CGGAGTACATGACCCACACACAGGATCAGTGCACATCCAGAACCAAGCATATGCGGAGCTAGAGCGGCTCTACTGCAAACGA
ATGGATGATGTGATCCCCTTCCTCTATGACTACTACCAGCTAGCAGCCAAATCAGGAAGGATGTGGCTCGGACCCGAGCTGTC
GGAGAAGCTATTTAGAAAGCTCCCCCCGGAAATAGGCCCAACAATAGAACAGGCCTATAAAGACAGATACCCAGGGTTGACG
ATAGGAGTCCTAGCAAGGGCCAACTTCATCCTGGAGTATTTACAAAACGTATGCAAGCAGGCAGCCTTGCAACGATCTCTCAA
GAGCCTGAGCTTCTGCAGGAACATGCCAGTACCAGGGTACTATGAGAAGAAGCAATATGGCATAAGGAAGGCTAAAACCTAC
AAAGGTAAGCCACATCCAACCCACGTGAAAGTGATCAAGAACAAGTACAAGCACACCCAAGGGAAGAAGTGCAAGTGCTACC
TATGTGGGATAGAAGGCCACTACGCCAGAGAATGCCCGAAGAAGGTGGTCAAACCTCAAAGAGCGGCATACTTCAACGGCAT
GGGCCTAGATGATAACTGGGATGTCGTATCTGTCGAACCAGGAGAGTCCGACGACGACGAAATCTGCAGTATCTCAGAAGGA
GAGAACGCTGGTGGAATGCACGAGCTGATGGCATTCAAGACCCAACTCCCGTACCCAGTGGAGTATGAAGCCAGCACACCAC
AGTTCCTGATGCCTTGGACACAGGTAACAGTGGAGAGAAGTGACAAACCTTCCTGGAGAAGAAGGAAGGAAATCCCAAAGGT
ACAGCAGGACTGCACGCACACCTGGAGCGACACGCAGGAGGTACCTATAGAAGGCAGGATATGCAGCATATGCAGCGATGA
GACACCTCACGGACGGAGAGTGACCTGCACAACATGCAGCATCAACCTGTGCCCCCTTTGCGCGTATATGGATCACGGGATCA
AGCTAATAGCCGCAAAGGACACCAAAGACGCAGCCAAATGGCAATATCACAACAAGGACGAGCTGATTCGTCACCTGTACGA
GCACAATGCTTTCCTCACCAGGAAGGTAGAAGAACTCACCAGCCAGCTGCAAGAATTCCACAGCCGCAGACCTGAAGACCTGA
TCAGTCTGGCGGACGACTTGGAGGACGTGTCCATTCTGGACAACGCCTCAAAAAGGGGGAAGGAGAAGGAATTGTTCCAATT
TGGAACGACAATTCCCATTGACCACATCCAAAACCTGGAAAACGTGGCCAAAATCATCGAACAATGGAAAGACACCCCCAGGG
TTGTGATCAAAGAGACACCTGAAAGCAGCAACAACACCATTGGAGCCCTCCTAGCCGAAGAAGGGATAGAAGAACTAGCTGC
AGCAGTGGACACAGCCTACACGGAGATGCCAAAAGGAGGCCTCAACAAACTTTACAACACCATTGTAGAGTTTGTAATACCAC
AAGAAAAAGGGGCACCCACCAGATTCAGGGTACGAGCTGTAATAGACACAGGATGCACCTGCACATGCATCAACAGCAAGAA
AGTCCCCAAAGAGGCACTAGAAGAGGCAAAGTACCAGATGAACTTTGCAGGAGTAAACTCAACGGGAGAAACAAAGCTCAA
GATGAAAAATGGTAAGATGATCGTCTCTGGGAGCGATTTCTACACACCATACATAGCAGCTTTCCCGATGGAGCTACCAGACG
TAGACATGCTCATAGGATGCAACTTCCTAAGGGCTATGAAAGGAGGCGTTCGTCTTGAAGGAACGGAGGTGACGATCTATAA
AAAGGTTACCACCATCCAGACCACCTTAGAGCCTCAGAAGATATCCCTCCTCCGCGCAGAAGCTGAAGCAGAAGAGGAGATG
GAGCGCGTGTACTACGCGAATGACTATTCCGAAGAAGGAATCAGTCGCCTGAAGAACCATAGGCTGCTACAGGAGCTGAAGG
AACAAGGCTACATTGGCGAGGAACCAATGAAGCACTGGGCCAAAAATGGTATCAAGTGCAAGCTTGATATTAAAAACCCAGA
TATAGTGATCAGCAGCAAGCCCCCTGACTCAGTGTCAAAAGAGACAAAGGCCCAATACCAGCGGCACATTGACGCCCTACTCA
AGATAGGAGTCATACAGCCTAGTAAGAGCAAGCACAGGACCGCGGCCTTCATCACACATTCGGGTACGTCCATTGACCCGATC
ACAAAGAAAGAAGTCAGAGGAAAAGAAAGGATGGTGTTCGACTACAGAAGTCTAAACGACAACACCCACAAAGACCAATACA
CTTTGCCTGGGATCAACACCATCATCTCAGCGATTGGCAATGCAAAGATCTTCAGTAAGTTTGATCTGAAGTCTGGATTCCACC
AGGTGCTGATGGACGAAGAATCCATCCCATGGACTGCATTTGTCACGCCAGTAGGGTTCTACGAGTGGAAGGTCATGCCCTTC
GGACTCGCAAACGCTCCGGCCGTCTTCCAGAGAAAGATGGACCAGTGTTTTGCAGGAACCTCAGAGTTCATCGCCGTCTACAT
CGACGACATCCTGGTCTTTAGCAAGACCTTGAAGGAGCACGAAAAGCACCTTAGCATCATGCTAGGGATCTGTCGAGACAACG
GCCTGGTATTGTCACCCAGCAAAATGAAGCTTGCAGCAACCGAGATCGATTTCTTGGGAGCTACCATAGGTGACGGAAGGATC
AAGCTCCAGCCTCACATAATCAAGAAGATTGCAGAGGTGGACGACGAGTCCCTAAAAACCCTCAAAGGGTTAAGAAATTGGTT
GGGAGTTCTCAACTATGCCCGAAACTACATCCCGAAGTGTGGAACACTCCTCGGCCCACTATACAGCAAAACTAGTGAGCATG
GAGACCGTAGGTGGCATGCCTCAGATTGGGCCTTAGTAAAGAAGATCAAAGGCCTGGTCCAAAATCTCCCAGACCTGAAACT
GCCCACAGAAGAAGCCTACATGATCATCGAAACCGATGGATGCATGGAAGGTTGGGGAGGAGTCTGCAAGTGGAAGCCCAA
CAAAGCAGACTCAGCTGGAAAAGAAGAAATCTGCGCTTACGCAAGTGGGAAATTCCCCACAGTAAAATCCACGATTGACGCA
GAAATCTTCGCGGTAATGGAGTCCTTGGAGAAGTTCAAAATCTTCTACATGAACAAGGACGAGGTCACAATCAGGACTGACTG
TCAGGCCATCATCACCTTCTATGAGAAGCTGAACGCTAAGAAACCTTCACGGGTAAGGTGGTTAGCTTTTTGCGACTATATAAC
AAACTCAGGGGTGAAGATGAAGTTCGAACACATCAAAGGCAAAGACAATCAGCTTGCCGATAACCTCAGCCGATTGACCCAA
CTCATCACCCTGGTCAAATGGCTTCCAGAAGAACTGAAGGATCTCGCGGCCGAGCTAGCCAAGAAAGAAGGCAAGACCTCCCT
GAAGGGAGAGGTGCAGGAGGAGATCTCCTGTTTTCTCAGAACTGCCCTCCGCCGAGCAGAGGAATCCGCGACTACTCGCCCA
TCAGAGCCGCGCCATGTACTATGGCAGAGATGGATGAATCCCGAAGACTGGCCTTACAACGAAGAAACAAAGTCTTCGATGA
CATCGCACAGAACATCAGCGATGCGGTCTACATCACCGGCATCGACCTCGCAGCGGCAAAAGCAAGAGCTACCAGGGACAAC
TGGTACAATGACGTCACACCCGCACTGGAGCAAAGAGCAGCTGCAGCATGGAGACTCATGGCAGCCTACTCAGACTTTGCCAC
GTGGAAGGACGTAAACGTCTAGTGAAGTGACGCAAGGGATGACTTCACAATTGCCAATGTCGTCATTGCTTACGACTTGGAAC
TTATCAGTTTGTGTCGGCAGCATCTATTAGCTGTCATACTTTATGTAAGTGCGCCAGTAGTGCGCTGAGTCATAAGGATAAGGA
ATCTTAGCTCCTTATCGTCCTTTAGCTTTAGCAAAGCTGTAAAGATGAATTTATTAGCTGTCGATGGGGCCCAGAAAGCGCACC
TGAGCTGATATTTTCTATCTTTTTGCTTAGCCTACTTCCTATATAAGGGAGTATAGTCTGAAGGCTTAGGCACAGAGCAATCTCT
TTAGCCAACCTTCTCTTGAGTTGTATTAA

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448838; this version posted June 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

>endoRYNV-PH2

TGATACCAGAAACAAAGAATTACCATGGTCGCTTAAATTTTCTTTTCTTGTCGAGAAACCCAAGTTTCAGACCAGAACCTTGAG
TTGTCTCTCCTTTTTTACAGGGGAGGAGGGAGTACTGAGTGTTGCTAAAATTCTGCAAATTTCAAACCCCCATGAAAACTTTCCT
CACGGTATTATAAGTTTTCTGACCCTCACTAGTTCAAACCTACTGCTGAAACTGTAGACTTAGGCGTCGAAGCGAAGTACCCTT
GTAGCCGTTAGCAGAGTGGCGTTAGGCGTTGATTGGGGAAAACTGACGTAAAGGAACAGCTGCAACTAGGCAAGAAAACTG
AAGGTCAGATCACCGGCCGGAGAGCTAGTACGCGGCTAGATCTGGGCAGATCTCTTGATGCAACCTCACGAAATCTCATCTTT
TGAAGAAGAGAGCAACTCTTGGGAAAGGTCTGAACGGGCGTATCGACAAGACTTCTTATTCAGGAATCTAAGATCCTACCCTC
GTTACGAAGCAAACCAGAAGTCACCTTCCTGTGATTTCCCTTGCTATCACTTCAACACAACAACCGGACCACCAGTCCACCGCA
CTCTCTGCAGACAGAAGAACAGCGAAGATTTACCCTTCCTGGTAAATACTCTGTTCGATCTTAGTATCACCAAGATTCACAACC
AAGCGGTCTTAGACGATAAGATCTCAAGACTCACCCAGTACTTGGTGGCAAAAGTCGGCACTCTACCGACAATCCCGGAGGAA
TCACCCCTCCTGGACCAGACAGCCATCTCTCTAGATCTTCAAGCCCTCAAGGCAGATCTAAAAGAGATCAAGGCTACCCAGTCA
TTCTTGAAGCTAGGCTTTGCACAGCTCCAAGAAGCGGTACAGCTGATCATCACAAGGGAGAACGATCCCAAACCGATTGAGGC
AGCTACAGCTCAGGTGGCTGAACAACTGAGGAAGCAGCTTATCGAGGTTAAAGCCGTCCTAGAGGAAACAAAGAAGATCGCG
AGATCTCTGTCCCCCGACGGATGAACCCCAGGTGGCAAGACACCGCCGCAAAAGAAACCTACATAAAAGCTATCCAGGCTACC
TCATCTCTTACATCCAACAACACCGGGCTAGGGTTCATTGAACCTCACACCTATACCGGAGGACAACTGTCCACAAGCCTAGCA
AAACAAAACAACACGCTCATACAGCTGTTGGTTCAGGTGCTAGAGAAGAATCTCGACCTTGAACAGGCCATCGCGAACCTGTC
CGCACAGGTCACGAGACTTGAGAAGACCGTCGCAGAGAAAGACACGGTCAAACTCCCAGAAAGTGTCCTCAACGACCTCACT
AAGGAGTTCGGAAAGGTCAACCTTGGGAAAGGAAAGGGGCTAGAAGGAGCAGTCTCATCCAGAGACAAAAACTTCTACGTTT
GAAAGAATCCCTTCAATCAATACAATGAGCAGAAGCCAAACAAGGCTCCAGGCTCCGCCCGCAACTGAAAGGGCAACAAGCA
GCTCGGACTCAGGCACCCCCACTCTGGAGGACCAGATCCGAGGATACAGGCGCTCTGCAAGGTTACGACACCAAGCACAGAG
AGCGGTACGAAGGACCTTCAGCAGGGACTTCAGGAACACCATAGAAAGGCAGCTAGATCCGGACGCCGAGTTATCACTCAGC
AGAAGGAGAAGAGCAAACCTAGTACCAGCAGAGGTACTCTACGCTCACAACGGTTCAGAACCTGTGAATCGTGTGTATGAGC
ACTATAGTGAGCTCGGGGCTCATATAGTAGACAGGCAACAAGACTTCCGATATATCGAGGAAGCATCCTACCAGAGACTAGTC
AGAGAAGGCATGCAATTCATACATGTCGGCATGGCAATGGTTAGGATCCAAATGTTGCACAGGACAGATGCAGGTATATCTG
CGTTGGTGGTGTTCAGAGACACTAGATGGAGCGATGACAGGCAGGTCATCGGCAGCATGTCCGTAGACATGACCAGGGGAG
CACAATTGGTATACATAATACCAAACGCAATGATGTCAATACACGATTTTTACAATCGTATACAGGTCAGCGTGCAAACCCGAG
GCTACGGGACAGGATGGGCTGGAGGAGACAGCAATATGATCATCACTAGATCACTGGTGGGGCGTCTCACCAACACCAGCAT
GACAAACTTCGAGTACCGGATAGATCAAGTCACAGACTATCTAGCAAGCAACGGAGTCGCTTGTATACCCGGACAGAAGTGG
GATGTGGCCAACAGGTCCGGCGAATGGGAGCTACAGCCCAGCAGGATTGTTGCACCACTCATGACACCAACAGAGGCGAGAC
TCAGCCAGAACAGGAGTGGCAGTATAAGTCTCAGATTCACCGACTTCCGAGATCAGAGGATCGTGGAAGAAAGGCCAGTCGA
GGATGAGGGCAGGCCAGAAGGTGAAGAGGAGAGCACGCACTACGTGCTCGTCTTCAGACATAACTCCTATGAGACCAACCTC
AGAGGAGAAAGGAGGCCAAGGCAGAACGAACTTTCAGAGTTCACACCAGAGACAGATTTGGTGAGCCAATGGCTAAGCCAG
CTATCCGCATCAGCACACAACAGTGGAGTGTCAAGCTCAGAAGAAGAACCGCCTAGGTTCGATGAAACTGATGAAGACAGTG
ATGGCACATACAATGAGAAAACCTGGCAGAAGGAAGACCAGGAAAGGAGAAATCTGGAGCTGCAAGGATGGAAGAAGACCT
GCAGACCAGGCATATACGAACTGATCCCAGAGGAAGAAGAGGAAATCTACCTCAGATACGAGGCAGAAGACGAGGATCAGG
AGGTACAGGTACTAGGGGCAACAACCATGGAGGAACCAGAAATGGAGTACCCCACCAGACTGGAAGAGGTTATGGGCAAGC
TAAAGAATGTCAGCATGGAGAAACTTTTTCCAGTAAGTGGCATGGATAGCGAGTCCAGCATCACAGGTGGAGGATTCATACCA
CCAAGCCCAGTACCAGGAGCGCAGGGATACCCACCTGCCACAGGAGCGTCTGCTTCAACCATAGGGCCAGCAGATCTGCAAG
GATGGGGAGGACGATTGCCAAGGAGCAGATCGCCGATAGGATATGGCCGCCCACAGCAGCCATGGTCCTTACCATCAGCCCA
GTCTGACAACGGCTGTATGCTAGTCCTACCTCAGGATTTCACCTTAGTTCCCGATGTGATCAACCGATGGGAGTCTATTACTGT
CAACCTCATCAACAAAATGATGTTTGATTCCCTACAGGATAAGGCGGACTATGTAGAAAATCTCTTGGGAGAAAGAGAAAAGG
AGACATGGATGACTTGGAGGATGCAGTACGAAGAAGAATATAAGCAACTCCTCACCATGAGCGGAGACGTGAGGAATCTCAC
TGCCGCAGTCAAAAGGGTCTTTGGCGTACATGACCCTCATACTGGATCAGTCCACATACAGAACCAGGCGTATGCAGAACTGG
AGAGACTGTACTGTAAGCGGACAGACGACGTGATCCCCTTCCTCTACGACTACTACCAGCTAGCAGCTAAGTCTGGAAGGATG
TGGCTCGGACCAGAGCTATCAGAAAAACTTTTCAGAAAGCTTCCACCTGAGATAGGGCCAACTATTGAACAGGCCTACAAAGA
CAGGTACCCAGGTCTCACCATTGGAGTCTTGGCAAGGGCCAACTTCATCCTGGAATACCTACAAAACGTCTGCAAGCAAGCAG
CGTTACAGAGGTCCCTGAAGAGCCTCAGTTTTTGCAGAAACATGCCGGTGCCTGGATACTATGAGAAGAAACAGTATGGCATC
AGGAAGGCAAAAACTTACAAGGGAAAGCCTCACCCAACGCACGTAAAGGTGATCAAAAACAAGTACAAGCACACCTCCGGGA
AGAAGTGCAAGTGCTACCTTTGCGGAATAGAAGGCCATTACGCCAGGGAATGTCCAAAGAAGGTGGTAAAACCACAAAGGGC
GGCGTATTTCAATGGCATGGGCTTAGATGACAACTGGGATGTTGTATCCGTAGAGCCAGGAGAGTCAGATGACGACGAAATC
TGCAGCATCTCCGAAGGAGAGAACGCTGGGGGAATGCATGAACTCATGGCATTCAAAACTCAACTCCCGTACCCAGTCGAGTA
CGAAGCCAGCACATCACAACAGTTCCTGCCATGGATACAGGTAACTGTCGAGAAAAGTGATAAGCCCTCTTGGAGGAGAAGA

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448838; this version posted June 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

AGAGATATCCCACAGGCACAAAAGGACTGCGCTCACACTTGGAGCGACACACAGGAAGTGCCGATAGAAGGGAGGATATGC
AGCATTTGCAGTGATGAAACGCCTCATGGAAGGAGAGTCACCTGCACGACCTGCAGCCTAAACCTGTGTCCAATTTGCGCATA
TATGGATCATGGAATTAAGCTGATCGCCGCAAGGGACACCAGAGACGCAGCAAAGTGGCAGTACCACAACAAAGATGAACTC
ATCAGGCACCTCTATGAACACAATGCCTTCCTAACCAGGAAAGTGGAGGAACTGACGACCCAGCTGCAAGAATTCCAAAACCG
CAAGCCTGAAGACCTAATCAGCTTGGCGGATGATATGGAGGACGTGTCCATTCTGGACAACGCCTCAAAAAGGGGGAAGGAG
AAGGAATCTTTCCAATTCGGAACGACGATTCCCATCGATCACATCCAGAATCTTGAAAACGTAGCAAGAATGATCGAACAATG
GAAGGAGACCCCCAGAGTCACTATCAAGGAAACAGCAGAAAGCAGCAACACAATAGGAGCCCTCCTAGCAGAAGAAGGGAT
TGAAGAGCTAGCTGCAGCTGTTGATACGGCATACACAGAAATGCCAAAAGGAGGTCTGAACAAACTATACAACACCATTGTTG
AGTTTGTAATACCTCAGGAAAAAGGGGCACCCACAAGATTCAGGGTTAGAGCAGTCATAGACACAGGATGCACCTGCACATG
CATCAACAGCAAGAAGGTCCCTAAGGAAGCCCTGGAAGAAGCAAAGTATCAGATGAACTTTGCAGGAGTAAATTCCACTGGA
GAAACCAAGCTGAAAATGAAGAATGGCAAGATGATAGTATCAGGAAGTGATTTCTACACGCCATATATTGCAGCTTTCCCAAT
GGAATTACCAGATGTGGACATGCTCATTGGCTGCAACTTCTTGCGAGCCATGAAAGGAGGAGTCAGGCTGGAAGGAACGGAG
GTAACCATCTACAAGAAGGTTACAACTATCCAGACGACTCTAGAACCACAGAAGATCTCTCTACTCCGTGCAGAAGCTGAAGC
AGGAGAAGAAATCGAGAGACTTTACTACGCCAACGACTACTCTGAAGAAGGAGTAAACAGGCTGAGAAACCATAAGCTGCTA
CAAGAGCTCAAGGAGCAGGGTTACATTGGGGAAGAACCAATGAAGCACTGGGCGAAGAACGGAATCAAGTGCAAGCTTGAT
ATCAAGAACCCAGATATAGTCATCAGCAGCAAACCTCCGGATGCAGTCTCAAAAGAGACGAAGGCACAGTACCAGAGACACA
TCGATGCTCTATTGAAGATCAAGGTAATCCAACCCAGCAAGAGCAAGCATCGAACTGCCACCTTCATCACAAACTCGGGTACA
ACCATAGACCCGATAACAAAGAAGGAGATCCGAGGAAAAGAGAGGATGGTCTTTGATTACAGAAGCCTCAACGATAACACCC
ACAAGGATCAGTACACCTTGCCTGGGATTAATACCATCATCTCCGCAATTGGCAACGCAAAAATCTTCAGCAAATTTGACTTGA
AGTCTGGCTTCCACCAGGTACTTATGGATGAGGAGTCCATCCCTTGGACCGCGTTTGTCACACCAGTAGGTTTCTACGAATGGA
AGGTAATGCCTTTCGGACTCGCGAACGCTCCTGCAGTCTTCCAGAGGAAGATGGACCAGTGCTTCGCAGGAACCTCTGAGTTC
ATAGCCGTCTACATCGACGACATTCTGGTCTTCAGTAAGACGCTGAAGGAACATGAGAAGCATCTCAGCATCATGCTGGGAAT
ATGTCGAGATAACGGCCTGGTACTGTCACCCAGCAAAATGAAGCTCGCCGCAACAGAGATAGATTTCTTGGGAGCAACCATTG
GTGACGGAAAAATCAAGCTCCAGCCTCACATCATCAAGAAGATTGCCGAGGTGGACGATGAATCCCTAAAGACCCTCAAAGG
GCTGAGAAGCTGGTTGGGAGTTCTGAACTATGCCAGGAACTACATCCCCAAGTGTGGGACACTCCTTGGCCCACTCTACAGCA
AGACGAGTGAGCACGGGGACAGAAGATGGCATGCATCGGATTGGGCCTTAGTCAAGAAAATCAAAAGCCTGGTCCAAAACCT
CCCAGACCTCAAACTGCCCACTGAAGAGGCCTATATGATCATCGAGACAGATGGTTGTATGGAAGGATGGGGCGGAGTCTGT
AAATGGAAGCCCAACAAGGCAGACTCAGCAGGAAAGGAAGAAATCTGCGCATATGCAAGTGGGAAATTCCCAACGGTCAAAT
CAACTATTGACGCAGAAATCTTCGCGGTCATGGAATCCTTGGAAAAATTCAAGATATTTTACATGAACAAGGATGAGATCACCA
TCAGGACCGACTGCCACGCCATCATAACCTTTTACGAGAAGTTAAACGCCAAGAAGCCTTCACGCGTAAGGTGGTTAGCTTTTT
GCGACTATATAACAAACTCCGGGGTGAGGATGAAGTTCGAACATATCAAAGGCAAAGATAATCAATTAGCTGATAATCTCAGT
CGCCTAACCCAACTCATCACCGCAGTAAGATGGCTTCCAGAGGAAATGGCAGGAATCGCGGCAGAACTAGTCAAAGACAGGG
AAGAGTCCCCCGCGATGCAGAAGGTACAGGAGAGCCTCTCAGGCTTTCTCAAAGCTGCCCTCCTCCAAGTCGAGAAATCCTCG
ACTACACGCCCATCAGAGCCGCACCATGCTCTCTGGCAGAGATGGACGAGTCTAGAAGACTGGCCAAACAACGCCGAGACAA
AGTCTTCGACGACGCAGGGCAAAACATCTGCGACACGGTCTACATCACCGGTGTCGACCTCGCCGCCGCCAAGGC

>epiRYNV-BS

TGGTATCAGAGCTTTAGCTTTCACCATGGTAGCTTAAACCCCCCCATCTTGTCGAGAAACCCAAGTTTCAGATCCGAACCTTGA
GTTTGCTCTCTTTTCGAAGAGGAGAGGGAGTGTGACCTTACTCAAAACTTTTGCGGATCAAACCCCCACGAAAACTTCCTTCAC
GGTATCATAAGTTTTCTGTCCTTCACTAGTCCAAACCTACTGCTCAAATCGCAGGTCTAGGCGTCGAAGCGAAGTACCCTTGTA
GCCGTTAGCAGAGTGGCGTTAGGCGTTGATTGGGGAAAACTGACGTAAGAAAGCGACAGCAACTAGGCGAGGAAACTGACG
GACAAATCACCGGCCGGGAAGCTAGTAAGCGGCTAGATCTGTGTGGTTTTGATGCAACCACACGAAATCTCCGAATTCGAAGC
AGAGAGCAGCTCTTGGGAAAGATCTGAACGGGCGTATCGACAAGACTTCTCATTCAGGAATCTCAGAACCTATCCACGTTGGG
AATCAAACCAGAGGACACCCTCTCTTGAATTTCCGTGCTACCATTACAACACAACAACTGGACCTCCAGTCCACCGTACTCTCCT
CAGACAAGGTGACGGCAAGGATTTACCATACTTAGTAAACACCTTGTTCGATCTCAATATCACCGAGATCCACAATCAGGCGAT
CCTTGACGATAAGATCTCAAGGCTCACCCAGTACCTGACAACAAAAGTCGGTTCCCTACCGACGATCCCGGAGGATTCGCCCCT
CCTGGACCAAATAGCCCTTTCCCTAGATCTACAGGCCCTCAAGGCAGATCTGAAGGAAATCAAGGCCACACAGTCAGCACAAA
AGCTAGCCTTCTCACAACTGCAGGAGGCAGTCCAGCTTTTTATCGCTCGGGAGAACGATCCCAAGCCGATCGAGGCAGCTACT
GCACAAGTAGCCGAACAGCTGAGGAAGCAACTCATCGAGGTCAAGACAGTCCTCGAAGAGACCAAGAAGATCGCCAGGTCCT
TATCCCCCGACGGATGAACCCTAGGTGGCAGGAAACTGCAGCCAAGGAAACCTATATCAAAGCGATCCAAGCAACCGCTTCCC
TCACCTCCAACGGCACCGGCCAAGGCTTCATTGAGCCCCACACCTACACCGGAGGACAGTTATCTACCAATCTAGCCAAGCAAA
ACAACACCATAATCGAATTATTGGTGCAAGTGCTAGAAAAGAATCTCGACCTTGAACGGGCCGTAGCCAACCTCACAGGTCAA

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448838; this version posted June 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

GTCACTAGGCTCGAAAAAGCCGTAGCAGACAAAGAAGCAGTCAAGCTTCCGGAAAAAGTCCTAGAAGACCTAACCAAGGAAT
TTGGAAAGGTTAACTTAGGCAAAGGAAAGGGGAAAAAAAGGAAAGTCTCAAGCCGAGACAGAACTTCTACGTGTGGAAAAC
CCCTACACTCAATACAATGAGCAGAAGCCACACAAGGCTCCAAGCGCCCCCGCAACTGAAAGGGCCACAAGCAGCTCGGATT
CAGGGACACCAACCCTGGAAGACCAGATCCGAGGATACAGGAGATCTGCACGCATGAGGCATCAAGCGCAGCAGAGACTAC
GGAGGACCTTCGGAAGGGACTTCAGAAACACGATCGAGAGACAACTCGATCCTGATGCAGAGCTCTCCCTCAGTCGAAGGAG
AAGAGCCAACTTGGTACCGGCGGAAGTGCTCTACGCACACAACGGACAAGAACCGGTCAACCGAGTATACGAGCACTACAGT
GAGCTCAGCGCTCATGTGGTAGACAGGCAGCAGGACTTCCGATTTATAGAGGAAGCGTCCTATCAAAGGCTGACCAGAGAAG
GAATGCAGTTCATCCACGTAGGTATGGCGATGGTCAGGATACAAATGCTGCACAGGACAGATGCGGGTATATCCGCACTAGT
GGTGTTCCGAGACACCCGATGGAGCGATGACAGGCAGGTCATCGGTAGTATGTCAGTCGATATGACACGAGGAGCGCAGTTG
GTATACATCATACCCAACGCCATGATGTCGATTCATGACTTCTACAACCGGATTCAGGTCAGCATACAGACCAGAGGATACGG
CACAGGCTGGGAAGGAGGCGACAGCAACATGATTATCACAAGATCATTAGTCGGCCGACTCACAAACACCAGCATCACAAGC
TTCGAGTACAGGATAGACAACGTGACAGACTACCTAGCCAGCAACGGCGTAGCCTGCATTCCCGGACAGAAGTGGTCCGTGG
CAAACAGGTCTGGAGAATGGGAACTCCAGCCAAGCAGGATAGCAGCACCACTTGCAGTCCCCACAGATGCCAGGCTAAGACA
GAACCCAAACGGCAACATCAGCCTGAGGTTCACGGATTTCCGCGACCAAAGGATCGTGGAAGAAGGAGAGACATCCGAGCCA
GAAGGAAGACCGGAGACGAAGGAGGATGAGAGCACGCACTACGTGCTTATGTTCAAACACTCCAGCCCTAGGTGGGATACG
CTCGGACAGCCCAGCGGTAAATACGACTACATGGTCCGATATGACGCACCAGAACCGACCACATGGCCAACAACAAACAGAG
GATGGGACGATGACCCTCCTAAGCCACCAAGCCCAAAAGGATCCTACGAGGTAAGCCTAAGAGGCGAAAAGAAGCTGAAAG
AAAAGGAACTCGCAGAGTTCACCCCAGAGACAGATCTGGTCAGCCAGTGGCTAAACCAACTCTCGAACTCGGCACACAACAGT
GGAGCTTCAAGCTCAGACGATGAGCCAAAGTTCGATGAGGCAGACGACGAGGACGACGTCTACAATCAGAAACCTGGGAG
AAGGAAGACCAGGAAAAGAGGGAGCTGGAACTCCAAGGGTGGAAACCCACAGGGAGGCCAGGCCTCTATGAAATGATCCCT
GAGCAAGAAGAAGAAGTCTACCTCAGGTACGAGGCAGAGGACGAAGAGGAGGATCAGGAGTTGCAGGTCATAGGAGCCGC
AACAATGGATGAGCCAGAGATGGAATACCCAACCAGGCTCGAAAAAGTAATGGGCAAACTCAAAAACGTAAGCATGGAGAA
GCTGTTCCCAGTGAGCGGGATGGATAGCGAATCCAGCATAACAGGAGGTGGTGGAGGTTTTATACCACCAAGCCCAGTACCG
GGAGCACAAGGATACCCCCCAGCAACAACATCCACTATGTCCACCATTGGACCAGCAGACATGCAGGGATGGGGAGGGAGA
GTGCCCAGAAGTAGGTCACCATTAGGGTATGGCAGACCACAACAGCCATGGTCACTACCATCAGCACAGTCAGACAACGGCT
GCATGCTAGTCCTTCCACAGGACTTCACCCTAATACCGGATGTCATCAACCGATGGGAATCCATAACAGTCAACCTCATCAACA
AAATGATGTTTGACTCCCTGCAGGACAAGGCCGACTACGTCGAGAACCTTCTTGGAGAACGAGAAAAGGAGACGTGGATGAC
ATGGCGGATGCAATATGAAGAGGAGTATAAACAACTCCTTACGATGAGCGGAGATGTGAGGAATATCACAGCCGCAGTCAAG
CGGGTCTTCGGAGTACACGATCCGCATACAGGATCAGTCCACATCCAGAATCAAGCATATGCAGAGCTCGAAAGGCTCTACTG
CAAACGGACGGACGACGTGATCCCCTTCCTATACGACTACTATCAGCTAGCAGCAAAATCGGGAAGGATGTGGCTCGGACCC
GAGCTATCAGAGAAGCTGTTCAGAAAGCTTCCACCTGAGATAGGACCTACCATCGAGCAGGCCTATAAGGATCGATATCCAGG
GCTGACCATTGGAGTTTTGGCCAGAGCAAACTTCATCCTGGAATATCTACAGAACGTGTGCAAGCAGGCAGCACTGCAGCGTT
CGCTAAAAAGCCTGAGCTTCTGCAGGAATATGCCGGTCCCCGGATACTATGAGAAGAAGCAATATGGTATCAGAAAGGCTAA
AACCTACAAAGGTAAGCCTCATCCGACCCACGTTAAGGTGATAAAAAACAAGTACAAGCATACGCAGGGGAAGAAATGCAAG
TGCTACTTGTGTGGGATCGAAGGTCACTATGCCCGAGAGTGCCCAAAGAAGGTGGTCAAACCACAACGAGCGGCCTATTTCAA
TGGCATGGGACTAGACGATAACTGGGATGTAGTCTCGGTCGAACCAGGAGAAGAAGACGACGACGAGATCTGCAGCATCTCA
GAAGGAGAAAACACTGGCGGAATGCACGAACTTATGGCATTCAAAACACAACTCCCTTATCCAGTGGAGTATGAAGCCAGCA
CACCACAGTTCCTGATGCCATGGACACAGGTACCTGTGGAAAAGAGCGACAAACCTTCCTGGAGAAGACGGAAGGATATCTC
ACAAGTCCAGAAGGACTGCACACACACCTGGAGTGACACCCAGGAAGTGCCTATCAGCGACAGGGTTTGCAGCATCTGTAGT
GACGAAACACCTCACGGTAGAAGAGTCACCTGCACTACATGCAACATCAACCTCTGTCCGATATGTGCAAGGATGGACTATGG
GATCATGCTGATAGCAGCAAAAGACACCAAGAGCGCAGCACATTGGCAGTACCAAAACAAGGATGAGCTCATACAGCACCTG
TACGAACACAACGCTTTTCTCACCAGGAAAGTAGCAGAGCTCACTAGCCAGCTGCAGGAATTCCACAACCGCAGGCCTGAAGA
CCTGATCAGCCTAGCGGATGACCTGGAGGACGTGTCCATCCTGGACAACGCCTCAAACAGGGGGAAGGAGGAGAAGGAATT
GTTCCAATTTGGAACTACAATTCCCATCGACCACATACAAAACCTTGAAAATGTGGCCAAGATCATAGAAAAGTGGAAAGATA
CCCCCAGGGTCGTAATCAAGGAAACACCAGAAAGCAGTACCAGTAACACCATCGGAGCACTTCTAGCTGAGGAAGGAATCGA
AGAACTGGCTGCAGCGGTTGACACCGCCTATACCGAGATGCCAAAAGGAGGCCTCAACAAACTTTACAACACAATTGTGGAGT
TCGTCATACCTCAGGAAAAGGGAGCACCCACGAAGTTCAGAGTTCGTGCTGTGATAGATACTGGATGCACCTGCACGTGCATT
AACAGTAAAAAGGTTCCCAAGGAGGCACTCGAAGAAGCGAAGTACCAGATGAACTTCGCAGGGGTTAACTCAACAGGGGAG
ACCAAGCTGAAAATGAAGAACGGGAAGATGATCGTGTCAGGCAGCGACTTCTACACCCCGTATATAGCAGCTTTCCCAATGGA
GTTGCCAGACGTAGACATGCTGATTGGCTGCAACTTCTTGCGAGCCATGAAGGGAGGCGTAAGACTCGAAGGAACGGAAGTG
ACTATCTACAAGAAAGTCACCACAATCCAAACAACCCTAGAGCCACAGAAGATATCCCTCCTCCGAGCAGAGGCTGAAGTCGG
AGAAGAACTAGAGCGCATGTACTACGCCAATGACTATTCCGAGGAAGGAATAAGTCGGCTGAAGAACCACAGGCTGCTGCAG
GAACTCAGAGAACAAGGGTACATTGGTGAAGAGCCAATGAGACACTGGGCAAAGAACGGCATCAAATGCAAGCTGGATATC

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448838; this version posted June 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

AAGAATCCAGACATAGTCATCAGCAGCAAACCGCCTGACTCTGTATCAAAAGAGACGAAAGCCCAATACCAAAGGCATATAGA
TGCCCTGCTCAAAATCGGAGTAATCCAGCCCAGCAAGAGCAAACACAGGACGGCGGCTTTTATCACACACTCGGGTACGTCAA
TTGACCCGATTACCAAGAAAGAAGTCAGAGGGAAAGAACGGATGGTATTCGACTACCGAAGTCTCAACGACAATACCCACAA
AGATCAGTACACACTGCCGGGTATCAATACCATCATATCCGCGATTGGCAACGCTAAAATATTTAGCAAGTTCGATCTAAAGTC
CGGATTCCATCAGGTGCTCATGGACGAAGAATCCATACCATGGACGGCTTTCGTAACGCCAGTCGGATTCTATGAATGGAAGG
TCATGCCCTTTGGCCTTGCCAATGCTCCAGCTGTCTTCCAAAGGAAGATGGACCAATGCTTCGCTGGAACTTCGGAATTCATCG
CAGTCTACATCGATGACATCCTGGTGTTCAGCAAAACCCTAAAGGAGCACGAGAAACACCTTAGCATCATGCTAGGGATATGC
CGTGATAACGGTTTGGTTTTATCGCCCAGCAAAATGAAGTTGGCCGCCACAGAGATAGACTTCCTTGGCGCCACCATAGGCGA
TGGAAGGATCAAGCTCCAGCCTCACATCATAAAGAAGATAGCCGAGGTGGATGACGAATCCCTGAAAACCCTCAAAGGGTTA
CGAAGCTGGTTGGGAGTGCTCAACTACGCGCGCAACTACATCCCAAAGTGTGGCACACTGTTAGGCCCACTATACAGCAAGAC
CAGCGAGCATGGTGATCGTAGGTGGCACGCGTCTGATTGGGCCTTAGTCAAAAGAATTAAGGGCCTGGTCCAAAACCTCCCA
GACCTAAAACTCCCCACGGAAGAGGCATACATGATCATTGAGACTGATGGATGCATGGAGGGCTGGGGAGGAGTCTGCAAAT
GGAAGCCCATGAAGGCAGACTCAGCAAGCAAGGAAGAAATCTGCGCTTACGCCAGTGGTAAATTCCCCACGGTAAAATCAAC
AATAGACGCAGAAATCTTCGCAGTTATGGAGTCCTTGGAAAAATTCAAGATTTTTTACATGAACAAGGACGAGGTCACTATCA
GGACTGATTGTCAAGCAATAATCACCTTCTACGAGAAGCTGAATGCAAAGAAACCTTCGAGGGTAAGGTGGCTTGCCTTTTGC
GATTATATAACGAACTCCGGGGTAAGAATGAAGTTCGAACATATAAAAGGTAAAGACAACCAGCTCGCAGATAACCTCAGCC
GTCTCACACAACTGATTACATTTGTGAAATGGCTTCCAACCGAGCTCAAGGACCTCGCGGCAGAACTAACCAGGAAAGACGAC
GGGACGCCCGCGAAGAAGGAAGTGCAGGAGGAAATCTCCTGTTTTCTCGAAGCTGCCCTCCGCCGAGCCAAGAGATCCGTGA
CTACTCACCAATCCGAGCCACGCCATGTACTATGGCAGAAATGGCAGAATCCAGAAGGCTGGCTCTACTGCGACGAGAGGAG
ATCTTCAACAGCCTTGCCCAACACATCAGCGACACGGTCTTCATCACCGGAGTCGACCTTGCGGCAGCAAAGGCCAGAGCAAC
CAGGGACAACTGGTATGCTGACATCACACCAACACTGGAACGACGAGCCACCGCAGCATGGAAGCTCATGGCCGCTTACGAG
GAATTCGCCACGTGTAAGGATGTGAACGTTTAGTGAAGCGACGTCAGCAATGACTTCACAATCGCCCAAGTGCGTCACTGCTT
ACGCTTGGGAACTTATCTTTTAGTGTCGGTAGCATCTTCTAGCTGCCATACTTTATTGTAAGTGCGCCGATAGTGCGCTGAGTC
ATAGTGATAAGGAATCTTATCACCTTATCGTCCTTTCTTAGCTTTAGTAGCTGTAAAGACGAACTTATTAGCTGTCGATGGGGCC
CAGAAAGCGCACCCGAGCTGATATTTTCTCTCTTTCTGCTAAGCCCTCCCCCTATATAAGGGAGAAGAGTTTGAAGGCTTAGGC
ACAGAGCAATCTCTCTAGCCAACCTTTCTCTTGAGTTGTATTAAAACATTCAAGTGAAATAAAGACTTGTTCATCTTTTTCCGCAT
ATCTCTGAGTTTTTATGAGTTCTTAAGTGTTCGAAAGCGCACTTTTGAAATTAGATCCATGTTTTTCGGACCCCATTC

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448838; this version posted June 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Supplemental table. 271 public and proprietary genetics were bulked into 294 samples and used in validation of epiRYNV-BS detection assay.

Bulked sample number	Genetic/cultivar
1	Selection 1, Selection 2
2	Selection 2, Selection 3
3	Selection 4, Selection 5
4	Selection 5, Selection 6
5	Selection 7
6	Selection 7
7	Selection 8, Selection 9
8	Selection 9
9	Selection 10, Selection 9
10	Selection 10, Selection 11
11	Selection 12, Selection 13
12	Selection 13, Selection 14
13	Selection 15, Selection 16
14	Selection 16, Selection 17
15	Selection 18, Selection 19
16	Selection 19, Selection 20
17	Selection 20
18	Selection 20, Selection 21
19	Selection 22, Selection 23
20	Selection 23, Selection 24
21	Selection 25, Selection 26
22	Selection 26
23	Selection 27
24	Selection 27, Selection 28
25	Selection 28, Selection 29
26	Selection 29
27	Selection 30
28	Selection 30, Selection 31
29	Selection 31, Selection 32
30	Selection 32
31	Selection 33
32	Selection 33
33	Selection 33, Selection 34
34	Selection 35, Selection 36
35	Selection 36, Selection 37

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448838; this version posted June 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

36	Selection 38, Selection 39
37	Selection 39, Selection 40
38	Selection 41, Selection 42
39	Selection 42
40	Selection 42, Selection 43
41	Selection 43, Selection 44
42	Selection 45, Selection 46
43	Selection 46, Selection 47
44	Selection 47, Selection 48
45	Selection 48
46	Selection 49
47	Selection 49, Selection 50
48	Selection 51, Selection 52
49	Selection 53
50	Selection 53, Selection 54
51	Selection 54, Selection 55
52	Selection 55
53	Selection 56
54	Selection 56, Selection 57
55	Selection 57, Selection 58
56	Selection 58
57	Selection 59
58	Selection 59, Selection 60
59	Selection 60, Selection 61
60	Selection 61
61	Selection 62
62	Selection 62, Selection 63
63	Selection 63, Selection 64
64	Selection 64
65	Selection 65, Selection 66
66	Selection 66
67	Selection 67, Selection 68
68	Selection 68, Selection 69
69	Selection 70, Selection 71
70	Selection 71
71	Selection 72
72	Selection 72, Selection 73
73	Selection 74, Selection 75

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448838; this version posted June 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

74	Selection 76
75	Selection 76, Selection 77
76	Selection 77, Selection 78
77	Selection 78, Selection 79
78	Selection 80
79	Selection 80, Selection 81
80	Selection 81, Selection 82
81	Selection 82
82	Selection 83
83	Selection 83, Selection 84
84	Selection 84, Selection 85
85	Selection 85
86	Selection 86
87	Selection 86, Selection 87
88	Selection 87, Selection 88
89	Selection 88
90	Selection 89
91	Selection 89, Selection 90
92	Selection 90, Selection 91
93	Selection 91
94	Selection 92, Selection 93
95	Selection 93, Selection 94
96	Selection 103, Selection 94
97	Selection 103, Selection 104
98	Selection 105
99	Selection 105, Selection 106
100	Selection 106, Selection 107
101	Selection 107
102	Selection 108, Selection 109
103	Selection 109, Selection 110
104	Selection 111, Selection 94
105	Selection 94, Selection 95
106	Selection 96, Selection 97
107	Selection 97, Selection 98
108	Selection 99
109	Selection 100, Selection 99
110	Selection 101, Selection 102
111	Selection 102

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448838; this version posted June 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

112	Selection 103, Selection 111
113	Selection 112
114	Selection 112, Selection 113
115	Selection 113, Selection 114
116	Selection 114
117	Selection 115, Selection 116
118	Selection 117, Selection 118
119	Selection 118
120	Selection 119
121	Selection 119, Selection 120
122	Selection 120, Selection 121
123	Selection 122, Selection 123
124	Selection 123, Selection 124
125	Selection 124, Selection 125
126	Selection 125
127	Selection 125, Selection 126
128	Selection 126, Selection 127
129	Selection 127, Selection 128
130	Selection 128
131	Selection 129, Selection 130
132	Selection 130
133	Selection 131
134	Selection 131, Selection 132
135	Selection 132, Selection 133
136	Selection 133, Selection 134
137	Selection 135, Selection 136
138	Selection 137
139	Selection 137, Selection 138
140	Selection 138, Selection 139
141	Selection 139
142	Selection 140
143	Selection 140, Selection 141
144	Selection 141, Selection 142
145	Selection 142
146	Selection 143
147	Selection 143
148	Selection 144
149	Selection 144, Selection 145

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448838; this version posted June 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

150	Selection 145, Selection 146
151	Selection 146
152	Selection 147, Selection 148
153	Selection 148, Selection 149
154	Selection 150
155	Selection 150, Selection 151
156	Selection 151, Selection 152
157	Selection 152
158	Selection 153
159	Selection 153, Selection 154
160	Selection 154, Selection 155
161	Selection 156
162	Selection 156, Selection 157
163	Selection 157, Selection 158
164	Selection 158
165	Selection 159
166	Selection 159, Selection 160
167	Selection 160, Selection 161
168	Selection 161
169	Selection 162
170	Selection 162, Selection 163
171	Selection 163, Selection 164
172	Selection 164
173	Selection 165
174	Selection 165, Selection 166
175	Selection 166, Selection 167
176	Selection 168
177	Selection 168
178	Selection 169
179	Selection 169, Selection 170
180	Selection 171
181	Selection 171, Selection 172
182	Selection 172, Selection 173
183	Selection 173
184	Selection 174
185	Selection 174, Selection 175
186	Selection 176
187	Selection 176, Selection 177

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448838; this version posted June 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

188	Selection 177, Selection 178
189	Selection 178
190	Selection 179
191	Selection 179, Selection 180
192	Selection 180, Selection 181
193	Selection 181
194	Selection 182
195	Selection 182, Selection 183
196	Selection 183, Selection 184
197	Selection 184
198	Selection 185
199	Selection 185, Selection 186
200	Selection 186, Selection 187
201	Selection 187
202	POLKA, QUALICUM
203	QUALICUM, YELLOW ANTWERP
204	CHILCOTIN, LATHAM
205	CHILCOTIN, KORBFUELLER
206	SOUTHLAND, SUMMIT
207	MALLING JEWEL, SUMMIT
208	MANDARIN, Selection 188
209	PHOENIX, Selection 188
210	BAUMFORTH SD.A, HERITAGE
211	SEPTEMBER, BAUMFORTH SD.A
212	Selection 189, ST. REGIS
213	CUTHBERT, ST. REGIS
214	MALLING LANDMARK, Selection 190
215	MALLING LANDMARK, TITAN
216	AUTUMN BLISS, CHEMAINUS
217	Selection 248, CHEMAINUS
218	Selection 191, Selection 192
219	Selection 192, TULAMEEN
220	MEEKER, OCTAVIA
221	OCTAVIA, Selection 193
222	Selection 194
223	Selection 194, Selection 195
224	Selection 195, Selection 196
225	Selection 196, Selection 197

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448838; this version posted June 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

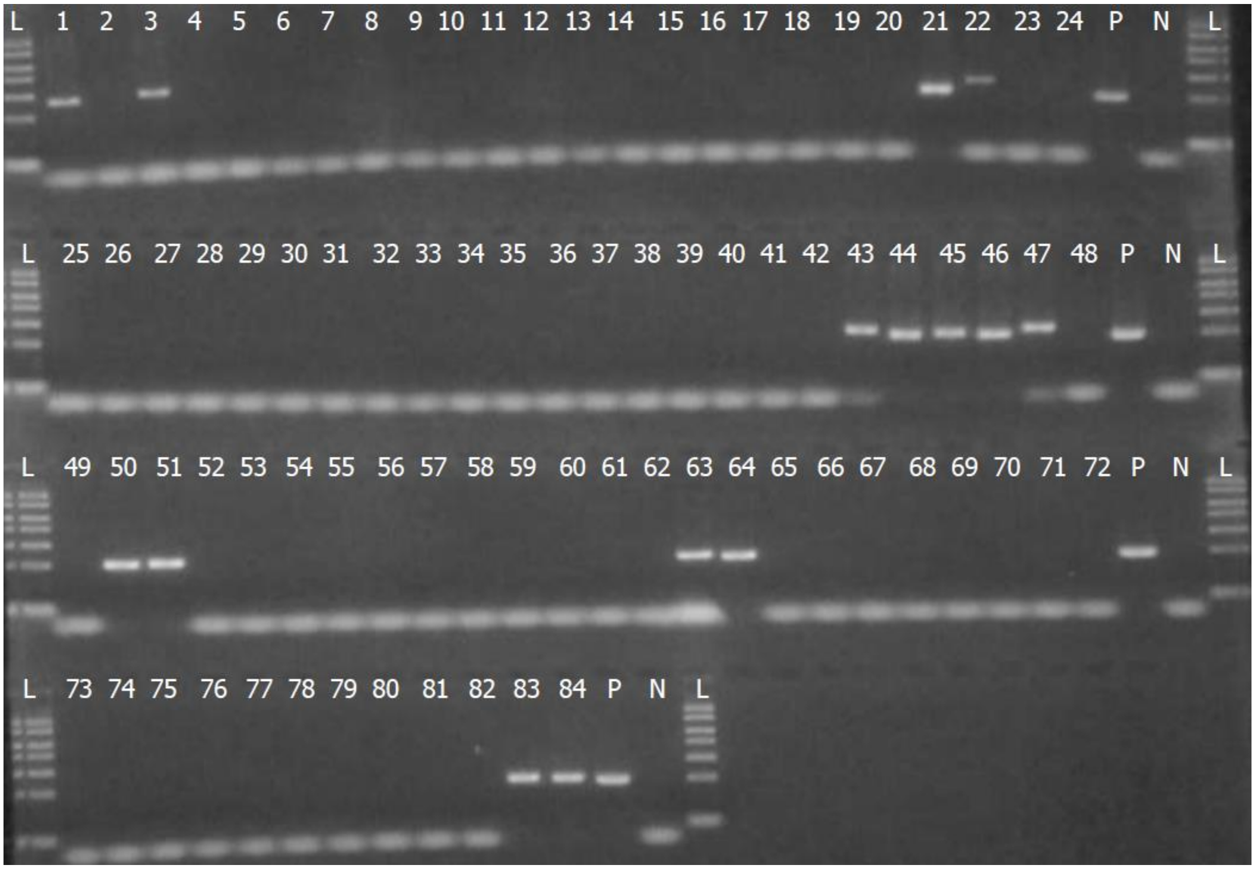
226	Selection 197, Selection 198
227	Selection 198
228	Selection 199, Selection 200
229	Selection 200
230	Selection 201
231	Selection 202
232	Selection 202, Selection 203
233	Selection 203, Selection 204
234	Selection 204
235	Selection 205
236	Selection 205, Selection 206
237	Selection 206, Selection 207
238	Selection 207
239	Selection 208, Selection 209
240	Selection 209, Selection 210
241	Selection 210, Selection 211
242	Selection 211
243	Selection 212
244	Selection 212, Selection 213
245	Selection 213, Selection 214
246	Selection 214, Selection 215
247	Selection 215, Selection 216
248	Selection 216, Selection 217
249	Selection 217, Selection 218
250	Selection 218, Selection 219
251	Selection 219
252	Selection 220
253	Selection 220, Selection 221
254	Selection 221, Selection 222
255	Selection 222, Selection 223
256	Selection 223, Selection 224
257	Selection 224
258	Selection 225
259	Selection 225, Selection 226
260	Selection 227
261	Selection 228
262	Selection 228, Selection 229
263	Selection 230, Selection 231

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448838; this version posted June 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

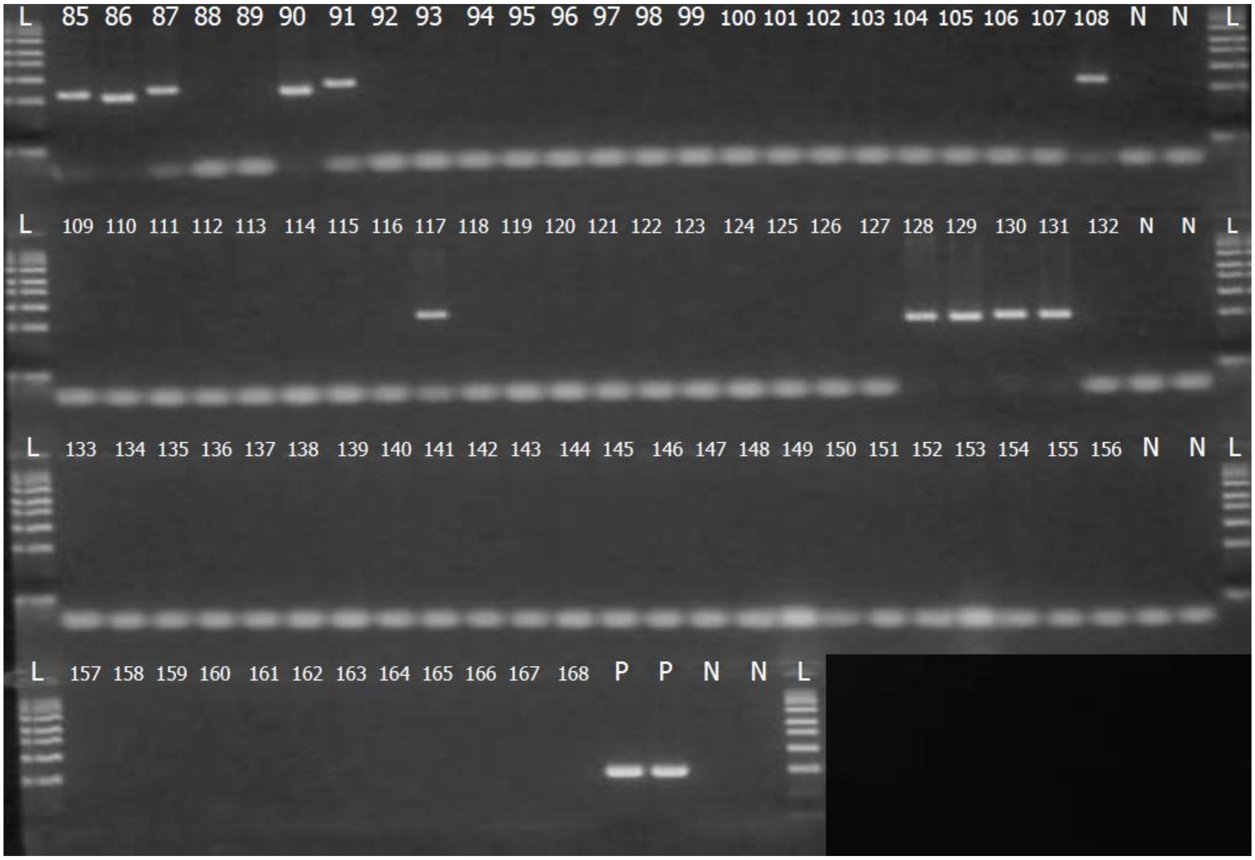
264	Selection 231
265	Selection 232
266	Selection 232, Selection 233
267	Selection 233, Selection 234
268	Selection 234
269	Selection 235
270	Selection 235, Selection 236
271	Selection 236
272	Selection 236, Selection 237
273	Selection 238, Selection 239
274	Selection 239, Selection 240
275	Selection 241, Selection 242
276	Selection 242, Selection 243
277	Selection 244, Selection 245
278	Selection 245, Selection 246
279	Selection 246
280	Selection 246
281	Selection 246, Selection 247
282	Selection 1, Selection 247
283	Selection 3, Selection 77
284	Selection 120, Selection 129
285	Selection 147, Selection 175
286	Selection 175, Selection 196
287	Selection 199, Selection 208
288	Selection 208, Selection 209
289	Selection 214, Selection 216
290	Selection 216, Selection 218
291	Selection 222
292	Selection 222, Selection 226
293	Selection 227, Selection 229
294	Selection 230

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448838; this version posted June 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

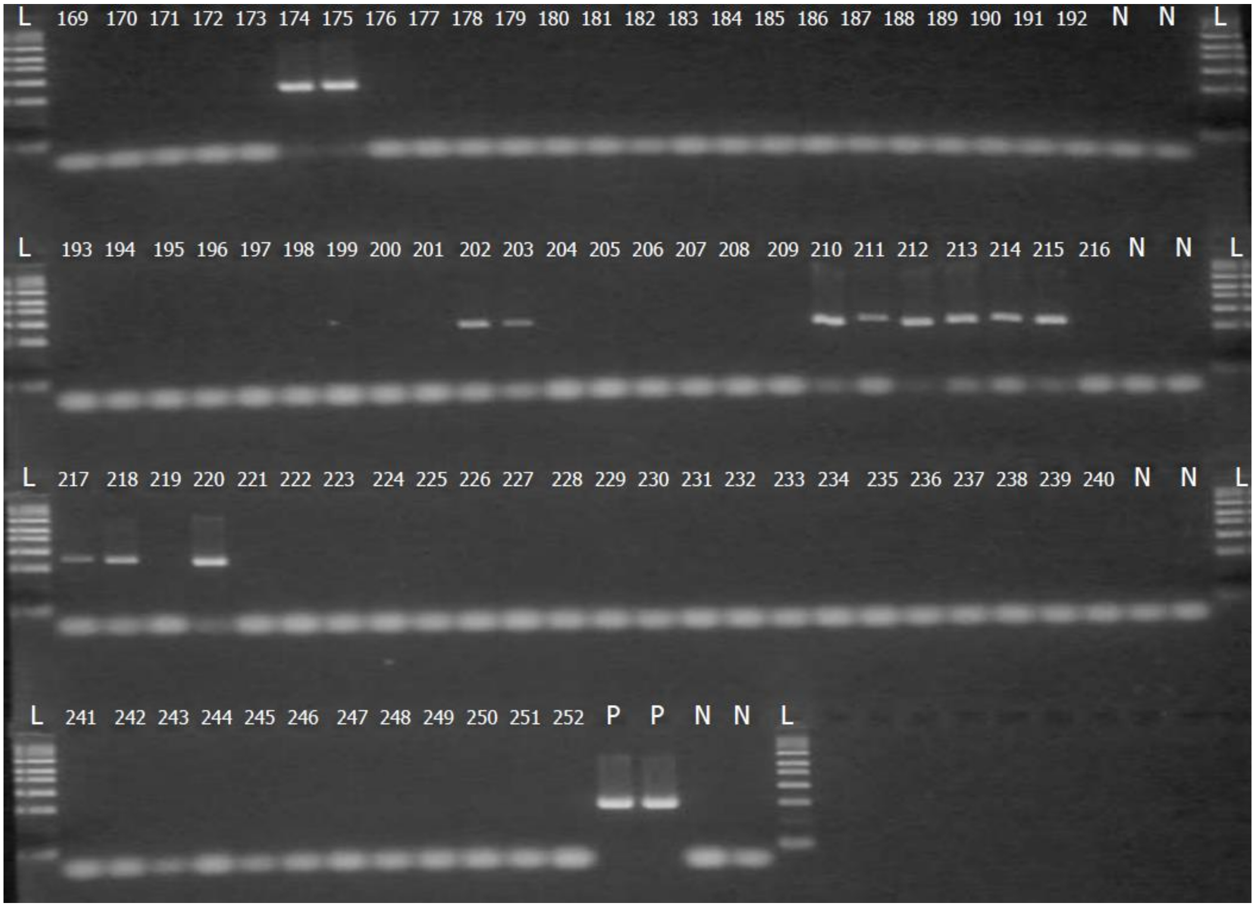
Supplemental figure. epiRYNV-BS detection using RYNV6-F/R PCR assay. 271 public and proprietary genetics were bulked into 294 samples. Off-target was observed in 43 cases.

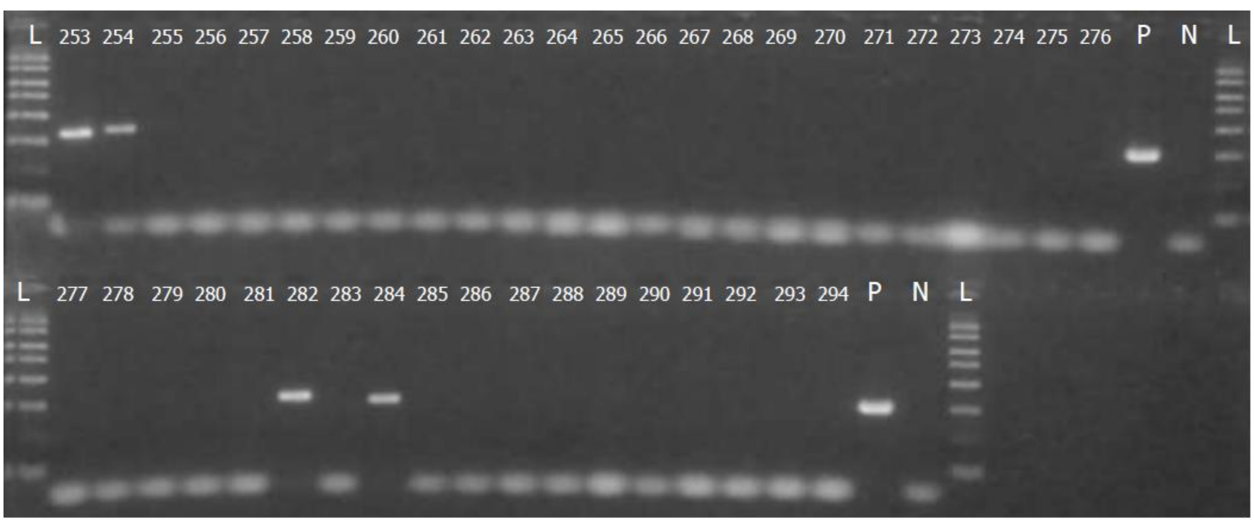


bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448838; this version posted June 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.



bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448838; this version posted June 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.





bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448838; this version posted June 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.