



# NGS of Virus-Derived Small RNAs as a Diagnostic Method Used to Determine Viromes of Hungarian Vineyards

Nikoletta Czotter<sup>1</sup>, Janos Molnar<sup>2,3</sup>, Emese Szabó<sup>1</sup>, Emese Demian<sup>1</sup>, Levente Kontra<sup>1</sup>, Ivett Baksa<sup>1</sup>, Gyorgy Szittya<sup>1</sup>, Laszlo Kocsis<sup>4</sup>, Tamas Deak<sup>5</sup>, Gyorgy Bisztray<sup>5</sup>, Gabor E. Tusnady<sup>2</sup>, Jozsef Burgyan<sup>1</sup> and Eva Varallyay<sup>1\*</sup>

<sup>1</sup> National Agricultural Research and Innovation Center, Agricultural Biotechnology Institute, Gödöllo, Hungary, <sup>2</sup> Research Center of Natural Sciences, Institute of Enzymology, HAS, Budapest, Hungary, <sup>3</sup> Department of Biotechnology, Nanophage-therapy Center, Enviroinvest Corporation, Pécs, Hungary, <sup>4</sup> Department of Horticulture, Georgikon Faculty, University of Pannonia, Keszthely, Hungary, <sup>5</sup> Department of Viticulture, Institute of Viticulture and Oenology, Szent-Istvan University, Budapest, Hungary

## OPEN ACCESS

### Edited by:

Gabor Jakab,  
University of Pécs, Hungary

### Reviewed by:

Massimiliano Morelli,  
Istituto per la Protezione Sostenibile  
delle Piante (CNR), Italy  
Stefanie Christine Becker,  
University of Veterinary Medicine  
Hannover, Germany

### \*Correspondence:

Eva Varallyay  
varallyay.eva@abc.naik.hu

### Specialty section:

This article was submitted to  
Plant Microbe Interactions,  
a section of the journal  
Frontiers in Microbiology

**Received:** 30 October 2017

**Accepted:** 18 January 2018

**Published:** 06 February 2018

### Citation:

Czotter N, Molnar J, Szabó E,  
Demian E, Kontra L, Baksa I,  
Szittya G, Kocsis L, Deak T,  
Bisztray G, Tusnady GE, Burgyan J  
and Varallyay E (2018) NGS of  
Virus-Derived Small RNAs as a  
Diagnostic Method Used to Determine  
Viromes of Hungarian Vineyards.  
*Front. Microbiol.* 9:122.  
doi: 10.3389/fmicb.2018.00122

As virus diseases cannot be controlled by traditional plant protection methods, the risk of their spread have to be minimized on vegetatively propagated plants, such as grapevine. Metagenomic approaches used for virus diagnostics offer a unique opportunity to reveal the presence of all viral pathogens in the investigated plant, which is why their application can reduce the risk of using infected material for a new plantation. Here we used a special branch, deep sequencing of virus-derived small RNAs, of this high-throughput method for virus diagnostics, and determined viromes of vineyards in Hungary. With NGS of virus-derived small RNAs we could detect not only the viruses tested routinely, but also new ones, which had never been described in Hungary before. Virus presence did not correlate with the age of the plantation, moreover phylogenetic analysis of the identified virus isolates suggests that infections are mostly caused by the use of infected propagating material. Our results, validated by other molecular methods, raised further questions to be answered before this method can be introduced as a routine, reliable test for grapevine virus diagnostics.

**Keywords:** grapevine, vineyard, virus, virome, diagnostics, metagenomics, small RNA, NGS

## INTRODUCTION

Grapevine can host more than 60 viruses and viroids (Al Rwahnih et al., 2009; Martelli, 2014), often as multiplied infection. Vegetative propagation and long lifetime of the plantation increase the risk of virus infection which cannot be controlled by traditional plant protection methods. The use of highly adaptive cultivars globalizes not only the presence of the particular cultivar but also the spread of new pathogens and their vector organisms. Safety regulations deal only with a limited number of known viruses and ignore new invading pathogens, which can lead to the use of infected propagating material and produce a new level of persistent infection risk. Traditional diagnostic methods can only answer the question whether or not the investigated virus is present in our sample, and need preliminary information about the pathogen (antigen for ELISA and sequence information of the particular variant for PCR-based methods). In striking contrast, deep sequencing

offers a unique opportunity to reveal any virus or viroid present in the sample, expected or not (Boonham et al., 2014). Indeed, different platforms were used for the description of new grapevine viruses (Martelli, 2014) and also to create the virome of a vineyard (Coetzee et al., 2010). During virus infection small interfering (si) RNAs having the same sequence as the infecting viruses are formed (Baulcombe, 2004; Molnar et al., 2005; Donaïre et al., 2009; Kreuzer et al., 2009; Szittyá et al., 2010) by the RNA interference (RNAi) based defense reaction of the plant. Deep sequencing of the small RNA (sRNA) population extracted directly from field plants offers a unique opportunity in virus diagnosis to identify several variants of grapevine infecting viroids (Navarro et al., 2009) or viruses (Pantaleo et al., 2010) even if they are alien on the plant or have never been described before (Zhang et al., 2011; Giampetruzzi et al., 2012; Wu et al., 2012). In our work we used this cutting-edge technique, the deep sequencing of virus-derived siRNAs to reveal the sanitary status of vineyards in our country. Analysis of the sRNA sequence dataset obtained using our bioinformatics pipeline enabled us to describe viruses never before reported from our country (Grapevine Syrah virus 1 and Grapevine Pinot Gris virus, Grapevine Satellite virus). Beside these new descriptions we analyzed our samples for the most widespread viruses with RT-PCR (using published diagnostic primers) and sRNA NGS in parallel. In most cases our results could be validated, but we also found contradictions in several cases, which is also discussed.

## MATERIALS AND METHODS

### Plant Material, Sample Preparation

Samples were collected from 14 vineyards representing 18 different varieties and nine wine-growing-regions of Hungary. Shoots of 1–10 randomly chosen individual plants per plantation were collected on the field or from sprouted canes. RNA was extracted from various organs: shoot tips, young leaves (until the 3rd internode from the shoot apex), older leaves (lower than the 3rd internode), tendrils, and inflorescences by CTAB method (Gambino and Gribaudo, 2008). RNA pools representing each plant were generated mixing equal amounts of RNA originating from the different organs. These individual plant RNA pools (library 14–18) or a plantation pool generated by the same strategy, representing all of the sampled plants from the same plantation (library 1–13) was used for small RNA library preparation (18 libraries in total) and sequenced using single index on a HiScanSQ by UD Genomed (Debrecen, Hungary) 50 bp, single end (8 samples/1 sequencing lane). Fastq files of the sequenced libraries are deposited to the GEO and can be accessed through series accession number GSE106240.

### Pipeline for Data Evaluation of NGS Results (Bioinformatics)

The resulting reads were sorted according to their indexes. Adapters of the sequenced reads were removed by the Trimmomatic program (Bolger et al., 2014), their quality was checked by the FastQC program (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc>) and deduplicated by the Picard MarkDuplicates tool (<http://broadinstitute.github.io/picard>). For

virus detection we used two different pipelines in parallel: (A) Short reads were mapped to viral reference genomes (Refseq viral database of NCBI from only plant and invertebrate hosts were used) by the BWA-aln short read aligner (Li and Durbin, 2009) with default options. Mapped reads were counted both with and without deduplication using samtools idxstats (Li, 2011). Redundant reads of the resulted hits were equalized to read/million read. Consensus viral sequences from the aligned deduplicated reads were generated using the samtools/bcftools (Li and Durbin, 2009) pipeline. Coverage of the appropriate genome was counted as % of the genome covered by nucleotide information from the mapped small RNA reads. (B) De novo assembling of the deduplicated reads was performed using Velvet with k-mer: 13, 15, 17 (Zerbino and Birney, 2008). The generated contigs were annotated by BLAST megablast (Morgulis et al., 2008) to the RefSeq of NCBI.

### Sequence Comparison

To compare consensus sequences of virus variants of the different libraries or sequenced PCR products we used the CLUSTAL Omega program (Sievers et al., 2011) and neighbor-joining algorithm implemented in MEGA v.6 (Tamura et al., 2013). Bootstrap values >70% (1000 bootstrap replicates) were used.

### Validation of Predicted Virus Diagnostics by RT-PCR

cDNA was synthesized from pooled RNA extracts representing each plantation using random primer and the RevertAid First Strand cDNA Synthesis Kit (Thermo Fisher Scientific, USA) (according to the manufacturer's instructions). The cDNA generated was used as templates for PCR reactions using Phire Hot Start II DNA Polymerase (Thermo Fisher Scientific, USA) and published diagnostic primers or new ones (see Supplementary Table 5) designed according to the consensus sequence generated by mapping our small RNA reads to the reference genomes. To detect GRVFFV, we used cDNA generated with a GRVFFV-specific GRVFFV-R/6391 primer. PCR products were analyzed by agarose gel electrophoresis. For Sanger sequencing cDNA was synthesized from pooled RNA extracts of individual plants and virus-specific PCR was done using Phusion Hot Start II High-Fidelity DNA Polymerase (Thermo Fisher Scientific, USA) or Q5 Hot Start High-Fidelity DNA Polymerase (New England Biolabs, UK) DNA polymerase. The purified products were cloned into pJET 1.2 vector (Thermo Fisher Scientific, USA) and sequenced. Sequences were deposited into GenBank (for GenBank Accession Numbers see Supplementary Figure 6).

### Validation by Northern Blot

For Northern blot analyses 4–5 µg of total RNA was separated on formaldehyde-1.2% agarose gel and blotted to Amersham Hybond-NX membrane (GE Healthcare, UK), by capillary method using 20xSSC. Hybridization was carried out at 65°C in Church buffer (0.5 M Phosphate buffer, pH 7.2 containing 1% BSA, 1 mM EDTA, 7% SDS) overnight with the appropriate radioactively labeled probe, washed for 5 min in 2 × SSC, 0.1%

SDS and for 15 min in  $0.5 \times$  SSC, 0.1% SDS at the temperature of the hybridization and exposed to an X-ray film.

Virus-specific,  $P^{32}$ -labeled, DNA probes were prepared by using the DecaLabel DNA Labelling Kit (Thermo Fischer Scientific, USA). As a template we used the PCR-amplified and purified product of cloned region of viral genome. The virus piece cloned was a 1663-bp part amplified with GPGV5557F and GPGV7220R for GPGV, a 1324-bp product amplified by RBDV\_RNA1F\_4082 and RBDV\_RNA1R\_5406 for RBDV RNA1, and a 927-bp product amplified by GSVsatF72 GSVsatR999 for GSV.

## RESULTS AND DISCUSSION

### Sample Collection and Sequencing

As a survey to detect virus infections in Hungarian vineyards samples were collected directly from the field, in a random fashion, independently of any apparent symptom, in May 2014 or from sprouts of single bud cuttings. 14 vineyards differing in the variety grown and the age of the plantation from 9 wine-growing regions of the country were sampled (Supplementary Figure 1 and Supplementary Table 1, 2). Small RNA libraries were prepared from pooled samples representing either the plantation (libraries 1–13), or different varieties at the same plantation (library 14–18) and sequenced.

### Initial Statistics

As a result of sequencing 8–14 million raw reads/library were generated (Supplementary Table 3). After trimming of the adapters, duplicates were removed and non-redundant reads (560000–1.6 million/library), without removing grapevine-specific sRNAs, were used for virus diagnostics. In different libraries 3.3–13% of the total non-redundant reads and 2.3–11.7% of the total redundant reads were mapped to viral reference genomes, representing 2–13 different viruses and viroids.

### Size Distribution of Sequenced Srnas

Size distribution of redundant sRNA sequences showed that the majority of the reads was between 21 and 24 nt, indicating that the library preparation was successful (Figure 1A). Most of the reads were 21 nt long and contained miRNA sequences in accordance with our previous report (Pantaleo et al., 2010). As for non-redundant reads the 24 nt-long size class was overrepresented (Figure 1B), likely responsible for transcriptional gene silencing (TGS) (Borges and Martienssen, 2015). sRNAs are products of different plant DICERs (DCL1, –2, –3, and 4) and each DCL enzyme activity produces a specific size class. The products of DCL1 and DCL4 are 21 nt-long, whereas DCL2 generates 22 nt-long, and DCL3, 24 nt-long sRNAs. Whereas DCL1 has a key role in miRNA biogenesis, DCL3 generates siRNAs for TGS against parasitic nucleic acids (e.g., transposons) (Parent et al., 2012). The *Vitis vinifera* genome encodes four DCLs homologous to DCLs characterized in *Arabidopsis thaliana* (Zhao et al., 2015). Although according to that work VvDCL1 contains only one RNaseIII domain, VvDCL2 and VvDCL3 lacks a dsRB domain, moreover VvDCL4 lacks a PAZ domain, we think they must be

fully functional since all characteristic sRNA size classes (21, 22, and 24 nt) were present in our samples. Figures 1A,B show the size distribution of all host-derived sRNA reads.

During antiviral silencing DCL4 and DCL2 process virus-derived dsRNAs into sRNAs. Virus- and viroid-specific sRNAs in our samples were almost exclusively 21–22 nt long, supporting the idea that in grapevine DCL2 and DCL4 are the key enzymes in virus-derived sRNA biogenesis (Figure 1C).

### Origin of Viral SRNAS

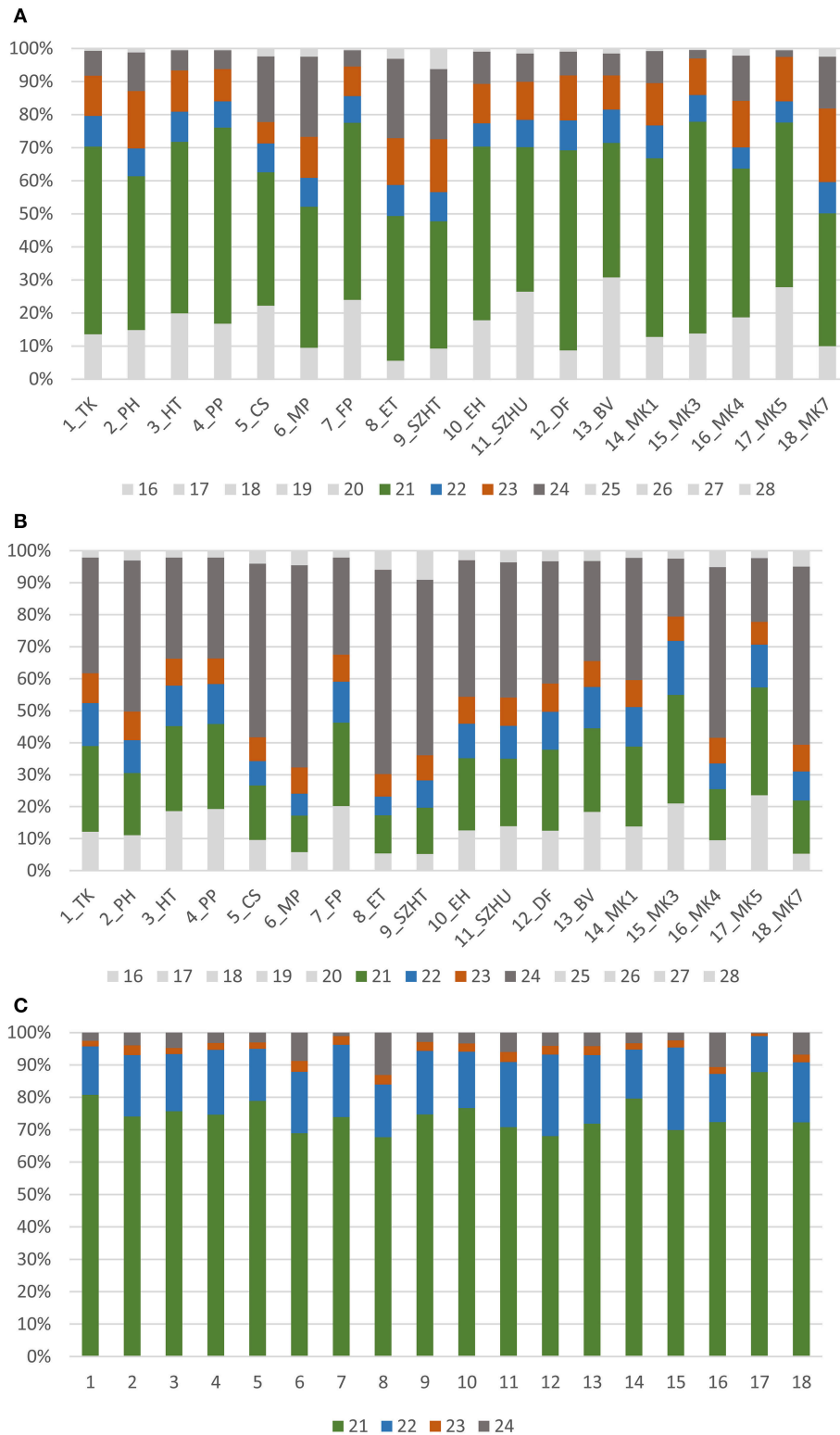
In order to identify viruses present in our plantations, sRNA reads were aligned and mapped to reference genomes of all known viruses of plant or insect host origin. Coverage (in %) of the whole viral reference genome was also calculated. During this analysis virus-specific contigs were assembled with different k-mers (kmer13, 15, 17), and the resulting contigs were also aligned to this set of reference genomes. In Supplementary Table 4 the results of the bioinformatics analysis for quarantine viruses: Grapevine fanleaf virus (GFLV), Arabis mosaic virus (ArMV), Grapevine leafroll-associated virus 1-3 (GLRaV1-3), Grapevine virus A (GVA), Grapevine virus B (GVB), Grapevine fleck virus (GFkV), together with Grapevine chrome mosaic virus (GCMV), Grapevine red globe virus (GRGV), Grapevine asteroid mosaic-associated virus (GAMaV), Grapevine vein feathering virus (GRVfV), Grapevine Syrah virus 1 (GSyV1), Grapevine rupestris stem pitting-associated virus (GRSPaV), Grapevine Pinot gris virus (GPGV), Raspberry bushy dwarf virus (RBDV), Grapevine satellite virus (GSV) and viroids: Hop stunt viroid (HSVd) and Grapevine yellow speckled viroid 1-2 (GYSVd-1 and 2) are summarized. A virus or viroid was diagnosed as present if any virus/viroid specific contigs (if any with any kmer) was present, and coverage of the viral genome by small RNA reads was higher than 40% (in case of viruses) or 80% (in case of viroids). According to these results the tested plantations are free from GFLV, ArMV, and GLRaV2, but we usually found simultaneous presence of up to 13 of different viruses and viroids in the same plantation.

### Validation of Deep Sequencing Results

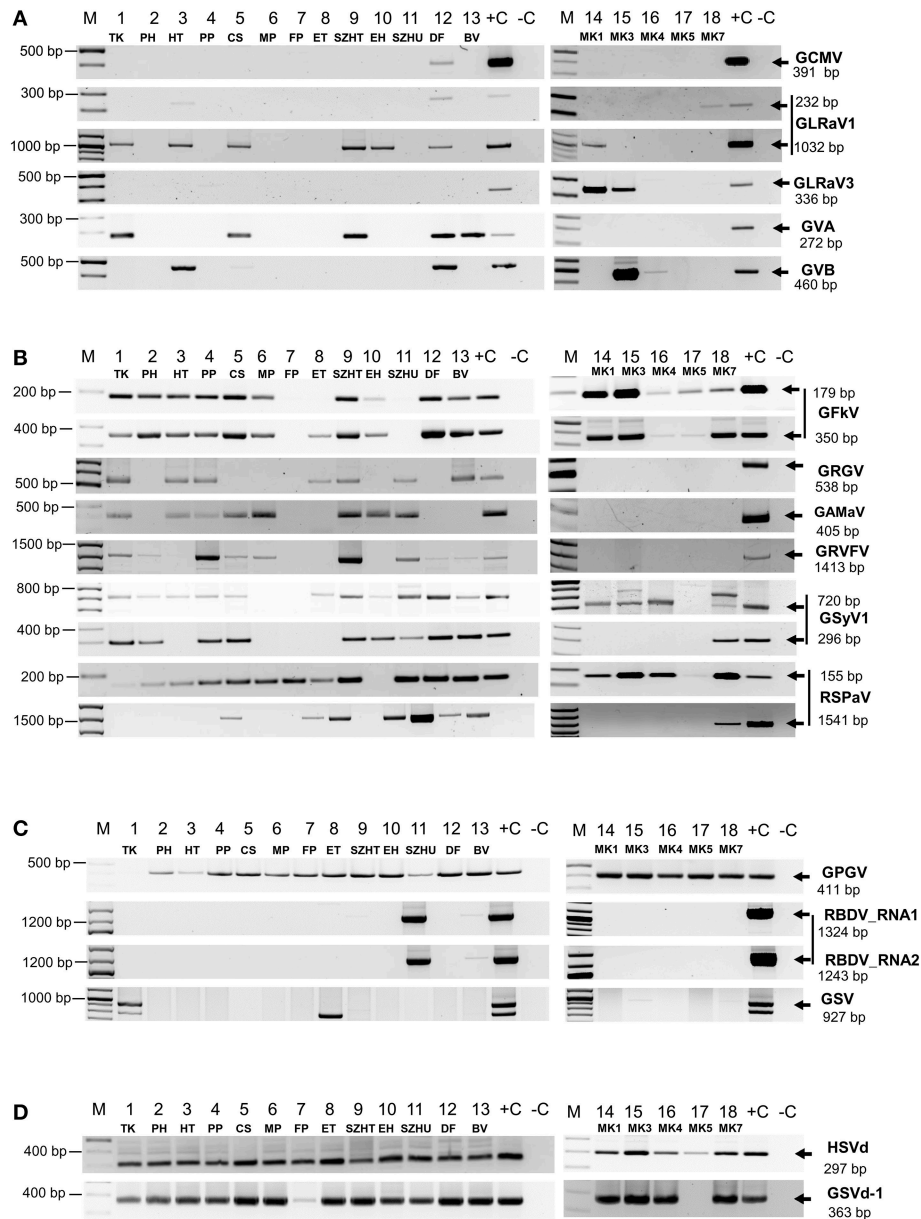
In order to validate our deep sequencing results we synthesized cDNA from RNAs representing plantation pools and set up PCR reactions by published diagnostic primers or with primers designed according to the sequenced sRNA reads (Supplementary Table 5). Positive controls (cDNA from virus-containing samples) and negative controls were always included. PCR products were analyzed by separation on 1.2 % agarose gels (Figure 2) and traditional Sanger sequencing. Results of the sRNA NGS virus diagnostics and its comparison with RT-PCR are summarized in Table 1. Sequences were deposited to GenBank and used for phylogenetic comparison (see Supplementary Table 6 for summary). Our results showed that the reliability of sRNA NGS as a diagnostic tool varied from virus to virus, and we discuss it for each of the identified viruses.

### Nepoviruses

The only *Nepovirus* detected was GCMV present in 12\_DF where partial CP could be amplified and sequenced (Figure 2A)



**FIGURE 1** | Size distribution of (A) trimmed, (B) non-redundant, (C) viral, redundant sequenced reads of the sequenced libraries.



**FIGURE 2 |** RT-PCR validation of sRNA NGS for **(A)** *Nepo*-, *Leafroll*-, and *Vitiviruses*, **(B)** *Tymoviruses*, **(C)** viruses which presence is not routinely tested and **(D)** *viroids*. cDNA was synthesized from pooled RNA extracts representing each vineyard using random primer and used as templates for PCR reactions with published diagnostic primers or new ones designed according to the consensus sequence generated by mapping our small RNA reads to the reference genomes. To detect GRVfV we used cDNA generated with a GRVfV-specific primer. PCR products were analyzed by agarose gel electrophoresis. (M), GenRuler 100 bp+; (+C), cDNA containing the tested virus was used as positive, or (-C), water as negative control.

(accession number MF100927). Interestingly enough, all of the GCMV RNA2 sequences in GenBank originated from Hungarian accessions (Elbeaino et al., 2014), and are more closely related to each other (more than 96% identity in this CP region) than to the isolate what we found (88–89% identity compared to other GCMV RNA2 sequences in this CP region) (Supplementary Figure 2A). The geographical origins of these Hungarian GCMV accessions are unknown. According to a recombination analysis of full GCMV RNA2 sequences, GCMV suggested to be a

putative interspecies recombinant of GARSV and TBRV (Digiario et al., 2015), which question could be further investigated incorporating HUDF isolate, however for its recombination analysis full RNA2 must be sequenced.

## Grapevine Leafroll Associated Viruses

GLRaV1 was detected in nine of our samples, but validation with published diagnostic primers failed in six samples (**Figure 2A**, **Table 1**). Testing the assumption that diagnostic primers were



**TABLE 1** | Summary of the small RNA NGS virus diagnostics and its RT-PCR validation for each library.

Library code	Age of the plantation in years	type of analysis	Nepoviruses				Grapevine leafroll viruses			Vitiviruses			Tymoviruses				Viruses not routinely tested						Viroids		total number of viral pathogens						
			GFLV	AMV	GCMV	GLRaV1	GLRaV2	GLRaV3	GVA	GVB	GFRV	GRGV	GMAaV	GRVFV new	GRVFV	GSV	GRSPaV	GPGV	RBDVs	GSV	HSVd	GYSVd-1									
1_TK	20	contig blast read bwa PCR	0	0	0	1	0	0	1	1	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	11	
2_PH	7	contig blast read bwa PCR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	11
3_HT	14	contig blast read bwa PCR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8
4_PP	16	contig blast read bwa PCR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	11
5_CS	?	contig blast read bwa PCR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10
6_MP	6	contig blast read bwa PCR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8
7_FP	4	contig blast read bwa PCR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7
8_ET	13	contig blast read bwa PCR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7
9_SZHT	?	contig blast read bwa PCR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9
10_EH	100<	contig blast read bwa PCR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	11
11_SZHU	1	contig blast read bwa PCR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10
12_DF	100<	contig blast read bwa PCR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7
13_BV	?	contig blast read bwa PCR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9
14_MK1	stock collection regularly virus tested	contig blast read bwa PCR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	12
15_MK3		contig blast read bwa PCR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9
16_MK4		contig blast read bwa PCR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4
17_MK5		contig blast read bwa PCR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5
18_MK7	number of plantation where the virus	contig blast read bwa PCR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
			0	0	0	1	8	0	1	5	3	12	9	8	13	14	13	2	2	14	13	2	2	2	14	13				7	

Result of the bioinformatics analysis (contig blast and read bwa) was marked by 0 or 1. 0 was used if there was no hit for the appropriate virus, whereas 1 was used if any contigs could be built, or the coverage of the virus genome was higher than 40% (in case of viruses) or 80% (in case of viroids). Result of RT-PCR validation was marked as 1, if PCR product was present or by S if not only the product was present but it was Sanger sequenced. Any hit present was highlighted with gray color. Summary of the total number of viral pathogens in the library, and number of the plantations where the virus is present, was highlighted with red and green color respectively.

designed to variable regions, new primers were designed to the HSP70 coding region based on the sequenced sRNAs reads. With this new set of primers we could successfully validate the presence of GLRaV1 in five additional libraries (**Figure 2A**). In 13\_BV we could only detect the virus in two of the individuals (Supplementary Figure 3A), which might explain the lack of the PCR product in the pool. Phylogenetic analysis of our isolates showed that they clustered into two distinct groups, the same E and A as was suggested by Kominek et al. (2005) (Supplementary Figure 2B) and are only 82–92% identical to the reference genome, supporting the operation of high-level variability, due to which GLRaV1 can be easily overlooked by traditional diagnostic methods (Esteves et al., 2013). The only Hungarian GLRaV1 isolate (CSE\_6.4.1.H) in GenBank (Cseh et al., 2013) (clustered to group E) was collected at the same region of the country from where HUTK and HUHT (clustered to group A) originated, suggesting that the source of the infection is more likely the propagation material. These results show that detection of GLRaV1 by sRNA NGS seems reliable, but its validation by RT-PCR may be problematic due to the high variability of the virus.

For GLRaV3 the result of sRNA NGS and its validation by RT-PCR using published diagnostic primer pairs correlated well (**Table 1**), detecting its presence in 14\_MK1 and 15\_MK3. Phylogenetic analysis of the sequenced part of the CP showed that these isolates (HUMK1 and HUMK3) are very closely related (98% identity on nt level) but share only 91% identity with the NY1 Reference genome from the USA (Supplementary Figure 2C). They clustered together with isolates from Brazil, Israel, and South Africa, but share only 90–91% identity with European strains. Their phylogenetical relationship based on the HSP70 region showed that they cluster into two distinct group together with isolates from different regions of the country (Cseh et al., 2013) (Supplementary Figure 2D). Samples 14\_MK1 and 15\_MK3 are different varieties in different rows of the same plantation. As these geographically linked variants tended to be divergent, it seems possible that the infection originated from an infected propagation material and is not the result of an onsite infection.

## Vitiviruses

GVA was detected at five, whereas GVB at two plantations (**Table 1, Figure 2A**). In 12\_DF we could clone an RT-PCR product from both of these viruses although we failed to detect GVB by sRNA NGS. Testing individuals for their presence we have found that in this sample only one plant was infected (Supplementary Figures 3B,C), which could decrease the concentration of the virus below the detection limit by sRNA NGS, in the pool. GVA isolates were only 85–90% identical to the Italian Reference genome and grouped into two distinct clades within Group I, together with other European strains (Goszczyński, 2014) (Supplementary Figure 2E). GVB isolates showed higher variation: they were only 85% identical to each other and grouped into different clades (Fonseca et al., 2016) (Supplementary Figure 2F). We can conclude that sRNA NGS based virus diagnostics worked well for vitiviruses, but our results

showed that using plantation pools containing extracts of non-infected plants can lower the virus concentration, and without further investigation slight and uneven infections can easily be overlooked.

## Tymovirales

### Grapevine Fleck Virus (GFKV)

GFKV was one of the most widespread viruses present in 14 samples. Virus diagnostics by sRNA NGS and RT-PCR validation for GFKV presence usually correlated well (**Table 1, Figure 2B**). Although using diagnostic primers we failed to validate its presence in 8\_ET, validation was successful with a new set of primers, designed according to the sequenced sRNA reads (**Figure 2B**). Sequence comparison of the isolates showed high variability: they were 85–95% identical to the Italian reference (Supplementary Figure 2G). Sequencing of the cloned PCR product in 5\_CS showed that this is a product of GRVfV (MF461275) and not of GFKV, highlighting the high-level identity of these two viruses. However, validation of the presence of GFKV failed, it is possible that sequencing more clones would have yielded a GFKV-specific product.

Tymoviruses, involved in fleck complex, are closely related; they can coexist in the same plant and GFKV is often presents in co-infection together with GRGV, GAMaV, GRVfV (Sabanadzovic et al., 2000; Cretazzo et al., 2017) and sometimes with GSyV1, therefore it is not surprising that we could also detect these viruses in most of our samples.

### Grapevine Red Globe Virus (GRGV)

GRGV was detected in many of our samples but its presence could only be validated in seven libraries (**Table 1, Figure 2B**). GRGV has been identified in different parts of Europe (Sabanadzovic et al., 2000; Beuve et al., 2015; Cretazzo et al., 2017; Voncina et al., 2017) and also in California (El Beaino et al., 2001) and in Brazil (Fajardo et al., 2017). We detected GRGV in different regions of the country; the sequenced strains have 87–95% identity with the Reference strain (NC\_030693) and clustered separately (Supplementary Figure 2H). Variability of different isolates and the fact that GRGV was not found in Czech and Slovak accessions—not even with NGS (Eichmeier et al., 2016)—further supports the idea that it is not originally present in Central-Europe, and possibly originates from infected propagation material of a geographically different origin.

### Grapevine Asteroid Mosaic Associated Virus (GAMaV)

We could detect GAMaV in eight samples (**Table 1, Figure 2B**). Although described and well-known since 1994 (Boscia et al., 1994), there is still only limited sequence information about GAMaV. The Reference genome was only uploaded to GenBank in 2016 (NC\_031692), but since that time—thanks to NGS surveys - it has been reported from Canada (Xiao and Meng, 2016) and from France (Candresse et al., 2017a). Our isolates are 94–96% identical to the Reference and 93–96% identical to each other, showing less variability than GRGV. As Hungarian GAMaV isolates clustered with isolates from different geographical origins (Supplementary Figure 2I), the use of

virus-infected propagation material is the most straightforward explanation for their presence.

### Grapevine Rupestris Vein Feathering (GRVfV)

GRVfV was detected in 11 and 13 libraries using a full genome (AY706994) and a Reference Genome (NC\_034205) (Reynard et al., 2017) for its detection respectively (Table 1 and Supplementary Table 4). The new Reference is only 77% identical to the first, Californian full genome, which shows the very high variability of this virus and the reason why the validation of its widespread presence has usually been failed (Pantaleo et al., 2010; Reynard et al., 2017). Our RT-PCR validation using cDNA produced by random probe also failed in all cases. We tried to increase the concentration of virus specific cDNA using virus-specific primer for cDNA synthesis and could amplify GRVfV in 9 libraries, but still failed in four samples (Figure 2B). Our isolates share 79–96 and 79–87% identity with AY706994 and NC034205, respectively, further confirming the diversity of this virus. Moreover, sequencing different individuals from the same plantation (HUcS, HUPP, HUTK, HUdF) revealed the presence of distinct variants (78–92% identical to each other) at the same plantation, clustering separately (Supplementary Figure 2J). This case shows that sRNA NGS has difficulties in accurate diagnostics of viruses with high variable genomes, however the presence of a Tymovirus could be accurately detected.

### Grapevine Syrah Virus 1 (GSyV1)

As we have reported previously we have found GSyV1 in Hungarian vineyards (Czotter et al., 2015 present in 15 libraries (Table 1). With DetF-DetR (Al Rwahnih et al., 2009#3) amplifying putative MP, we could validate its presence in 10 while with primers amplifying part of the CP (Sabanadzovic et al., 2009) in further six samples (Figure 2B). High prevalence of GSyV1 have also been found in Czech and Slovak grapevines which also showed high variability at the 5' putative MP coding region (Glasa et al., 2015), resulting false negative result if the above primers were used for diagnostics. Our isolates grouped distantly both into the two major and the diverged third clades suggested by Glasa (Glasa et al., 2015) (Supplementary Figure 2K). Moreover sequences of GSyV1 from different individuals of the same plantation (HU11TK2 and HU11TK9) showed high variability what supports the idea that GSyV1 population in Central Europe is more diverse than the North American ones.

### Grapevine Rupestris Stem Pitting-Associated Virus (GRSPaV)

GRSPaV is known to be the most widespread virus infecting grapevine, but to our surprise we could detect its presence by both contig blast and read bwa in only three of our samples (Supplementary Table 4). In a striking contrast, with RT-PCR using diagnostic primers amplifying a very short part of the replicase we obtained a product in 16 of our samples (Figure 2B). GRSPaV's viral RdRp has a very low proofreading activity and frequent recombination events, because the coexistence of different variants in the same plant led to the evolution of diverse variants (Morelli et al., 2011; Glasa et al., 2017). To test that if the number of virus-specific reads or

coverage will increase if we use a different reference during bioinformatics analysis, we made the direct sRNA BWA using 5 distinct full GRSPaV genomes (AF057136=NC\_001948\_Ref\_1\_USA, AY881627\_BS\_Canada, KR054734\_JF\_China, AY881626\_SG1\_USA, AY368590\_SY\_USA) (Supplementary Table 7). According to this analysis we obtained higher than 40% coverage to different GRSPaV strains only in two additional libraries. GRSPaV is mostly spread by vegetative propagation; as a result, it is particularly difficult to eliminate by sanitation techniques and coexists with grapevine for a long time (Meng et al., 2006). Their coevolution led to gene expression changes of the host with mutual advantages, resulting in slight down-regulation of stress genes in the presence of the virus (Gambino et al., 2012). Because of these advantages it is possible that during this coevolution an acceptable balance of the virus and the host defense reaction was achieved. The presence of a GRSPaV-coded silencing suppressor with an activity to block virus-derived sRNA biogenesis can also explain these results, but must be further investigated.

To be able to analyse phylogenetical relationship of GRSPaV isolates, we amplified and sequenced a longer 3' part of their genome (Figure 2B). Hungarian variants clustered with the GRSPaV-1 and Tannat variants from the USA and Uruguay, but it is possible that sequencing more clones or full genomes would alter this phylogenetic picture (Supplementary Figure 2L).

### Grapevine Pinot Gris Virus (GPGV)

According to our survey, GPGV, never described before, seems widespread in our country: it is present in 17 libraries (Table 1). In striking contrast to the predominance of 21 nt GPGV-derived reads with both sense and antisense orientation, a 24 nt-long antisense excess was found in 1\_TK and 11\_SZHU (Supplementary Figure 4A). In these samples the number of virus-specific reads was very low, but we have found 2 GPGV-annotated contigs in each of them. Aligning these contigs to GPGV revealed that they were generated from the 5' part of the genome (Supplementary Table 8). In this region (155–235) the Italian Reference (NC\_015782) differs from all of the other sequenced genomes and contains a stretch with exact match to *V. vinifera* shotgun sequences. Similarly to GPGV's original description (Giampetruzzi et al., 2012), our bioinformatics pipeline does not contain removal of host/grapevine-specific reads before virus diagnostics. This suggests that GPGV-identified sRNA reads in these libraries could be false positives of host origin, which what could explain their size distribution and the contradiction between contigs and the results based on direct sRNA alignment. To investigate the question why we could get a GPGV-specific RT-PCR product in 11\_SZHU (Figure 2C), if the virus-derived sRNAs are false positives, extracts of individuals and different tissues were investigated by RT-PCR for the presence of the virus. The analysis showed that at this very young plantation only one plant was infected, and GPGV was present only in its young leaves and shoot tips (Supplementary Figure 3D). We proved the presence of GPGV by Northern blot in 11 vineyards (using pooled RNA for 14–18 libraries, originating from different varieties of the



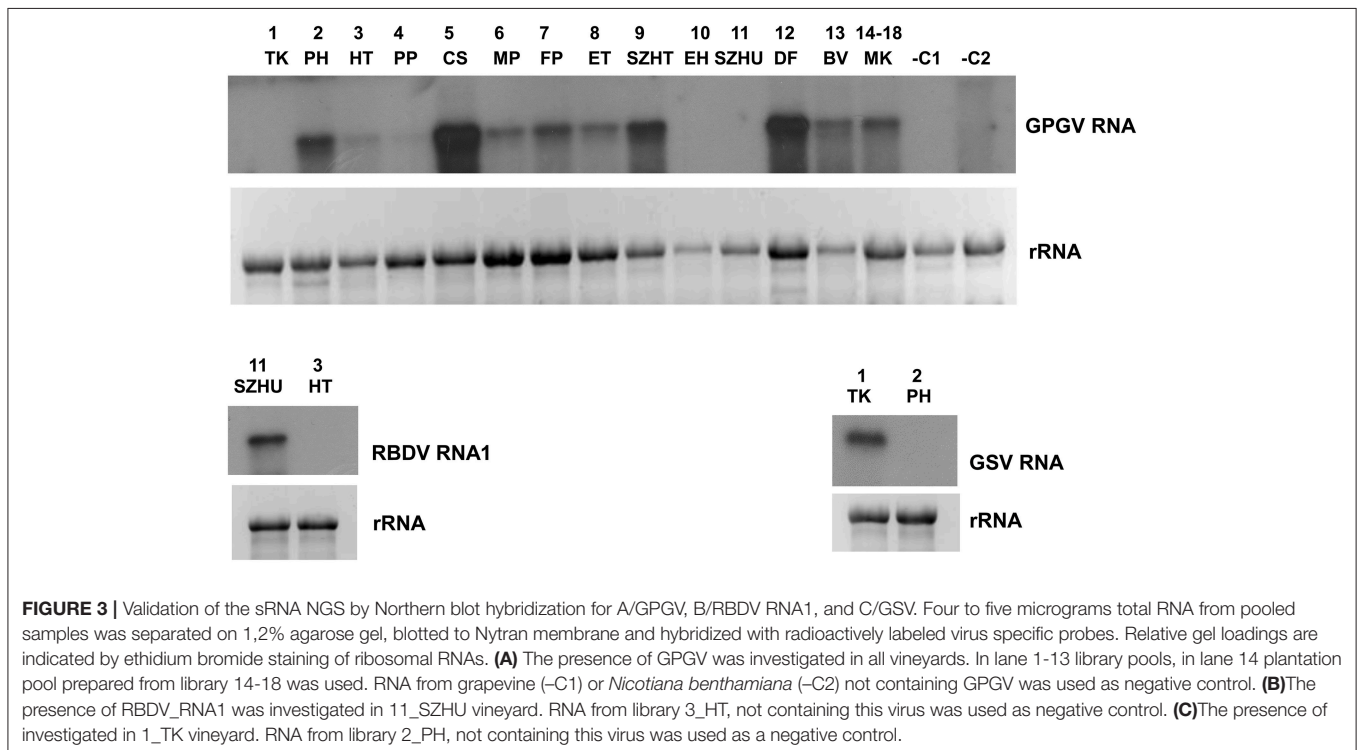
same plantation) (**Figure 3A**). In 10\_EH the low amount of loaded RNA, whereas in 11\_SZHU low amount of GPGV could be the reason why we couldn't get a signal. Since its first description (Giampetruzzi et al., 2012) GPGV has been reported from all over the world, including Slovakia (Glasa et al., 2014), Slovenia (Plesko et al., 2014), Croatia (Voncina et al., 2017), Serbia, Romania, and Ukraine (Bertazzon et al., 2016), i.e., almost all of Hungary's neighboring countries. According to their CP sequences our isolates showed slight variation, but grouped distantly and together with isolates of different geographical origins (Supplementary Figure 2M), which supports the possibility that GPGV spread from Eastern Europe to Italy (Bertazzon et al., 2016; Malagnini et al., 2016), and from Europe to other parts of the world (Wu and Habili, 2017). Its spread by a putative slow-moving eriophyid mite vector (Bertazzon et al., 2017) from surrounding infected plants could explain its unequal presence in our youngest (1-year-old) plantation (11\_SZHU). In spite of its high prevalence, symptoms caused by GPGV are rare, and the varieties we sampled did not show symptoms connected to GPGV. Requirements for being latent or virulent strain are still elusive (Saldarelli et al., 2015); however, according to the polymorphism at the end of the MP, all isolates from our country belong to the latent group, having MPs shorter by six amino acids.

### Raspberry Bushy Dwarf Virus (RBDV)

Since its first description in Slovenia (Mavric et al., 2003), RBDV has been reported rarely, but also from Hungary, to infect grapevine (Plesko et al., 2012). We found the presence of

RBDV in 11\_SZHU, where presence of contigs, high number of normalized virus-specific and high coverage of both RNA1 and RNA2 by small RNA reads were present (Supplementary Table 4). In this very young, 1-year-old Furmint plantation we successfully validated the presence of both RNAs by RT-PCR (**Figure 2C**), and in case of RNA1 by Northern blot (**Figure 3B**). Phylogenetical analysis of the cloned part of RNA1 showed that it differs from the RBDV RNA1 sequences, which was only available from *Rubus* host (Supplementary Figure 2N upper panel). The same analysis of the MP part of RNA2 (Supplementary Figure 2N lower panel) showed that RBDV at Tokaj, in the north-eastern part of the country, is closer to the isolate originating from the Slovenian *Vitis* host than to Hungarian isolates from the same host, from the southwestern part of the country (JQ928628 and JQ928629), which strongly suggests its origin from the plantation material. Although we couldn't find RBDV-specific RT-PCR product from the 13\_BV plantation pool, we tested individuals at this plantation, because RBDV-specific sRNAs were present and coverage of both RNA1 and RNA2 was about 40%. An RT-PCR product with primers amplifying RNA1 at the expected size was found in BV2 plant (Supplementary Figure 3E) and proved to be RBDV1-specific by Sanger sequencing, having 98.6% identity to the RBDV RNA1 sequence from 11\_SZHU (Supplementary Figure 5). BV2 is Balafánt, an ancient Hungarian variety, which has been present in the collection for a long time, raising the question how it could be infected by this virus in the north-eastern part of the country.

In addition to viruses, we also identified a viral satellite and viroids in our vineyards.



## Grapevine Satellite Virus (GSV)

Grapevine satellite virus was first described in California by deep sequencing of dsRNAs (Al Rwahnih et al., 2013) and later from Iran by the same technique (Candresse et al., 2017b). During our survey we found its presence in two of our libraries (Supplementary Table 4). In 1\_TK its presence was validated by RT-PCR (Figure 2C) and also by Northern blot (Figure 3C). In 13\_BV we obtained a virus-specific product by RT-PCR only when we tested individuals of this plantation (Supplementary Figure 3F). Although a helper virus of GSV has not been identified yet, its original source was infected by Vitiviruses and various GLRaVs (Al Rwahnih et al., 2013). 1\_TK and 13\_BV contained GVA and GLRaV1 in parallel, suggesting their helper virus function for GSV. Sequence comparison of the 927 bp cloned from GSV (HUTK) showed that is closer to the Californian Reference (NC\_021480) than to the one from Iran (Supplementary Figure 2O), and it is slightly different from our other isolate (Supplementary Figure 6), making its origin more elusive.

## Viroids

### Hop Stunt Viroid (HSVd)

According to our survey, almost all of our libraries contained massive amounts of HSVd-derived sRNAs. Their size distribution and polarity showed the same pattern: simultaneous presence of 21-, 22-, and 24nt-long reads originating from both strands, as it was expected on the basis of the detailed analysis of Navarro et al. (2009) (Supplementary Figure 4B). We could validate its presence in our samples by RT-PCR (Figure 2D). As HSVd sequences from different strains share high identity, we only sequenced 1 cloned PCR product originating from Tokaj. HSVd variants from the sequenced ENTAV115 Pinot noir originating from Eger (the northern part of the country) contained the Riesling variant reported from Germany (Navarro et al., 2009). Phylogenetic analysis of our isolate HUMK4 showed that although it is 95% identical to the reference genome, it is not very closely related to this Riesling strain and to another Hungarian HSVd isolate reported from Hungary (Supplementary Figure 2P).

### Grapevine Yellow Speckled Viroid (GYSVd1, 2)

GYSVd was found in 16 samples (Table 1). Although they are 81% identical our analysis generated results for GYSVd1 and GYSVd2, as they both have reference genomes. According to our bioinformatics analysis GYSVd1 is present in 16 of our libraries. The same analysis for GYSVd2 was not that straightforward (Supplementary Table 4). Size distribution of the GYSVd1 derived sRNAs showed again the presence of 21, 22, 24 nt-long reads, but in our samples the presence of 22 nt-long reads was not as high as it was for those originating from the HSVd or GYSVd1 published from the Pinot noir ENTAV (Navarro et al., 2009) (Supplementary Figure 4C). RT-PCR validation by GYSVd1-specific primers was successful for all of our samples (Figure 2D), but we never obtained any product if we used GYSVd2-specific primers, which would suggest that variants at Hungarian plantations are closer to GYSVd1. Sequences of three isolates, originating from different parts of the country,

showed divergent variation (Supplementary Figure 2Q). Detailed analysis by multiple alignment showed hot spot regions of the genome, where its structure allows mutations, but we did not find variation at the point where variants from PN had been reported (Navarro et al., 2009) (Supplementary Figure 7); however, it might be possible to identify variants at that point by sequencing more clones.

## Sanitary Status of Hungarian Plantations

In most cases we could validate the result of sRNA NGS by RT-PCR, but interestingly the number of validated viruses and viroids was sometimes higher than expected according to contig blast or read bwa (Table 1). The investigated vineyards were free of GFLV, ArMV and GLRaV2. GCMV and GLRaV3 was present at one, GVB at three, whereas GVA at five places. As expected, GLRaV1 and GFkV were present frequently (8 and 12 plantations, respectively), and we also found other Tymoviruses: GRGV, GAMaV, and GRVfV together with GFkV at several locations (9, 8, and 13, respectively). The most widespread infection was found for viroids (present at 14 and 13 plantations, respectively), GRSPaV and, unexpectedly, for GSYV1 and GPGV at 14, 13, and 13 locations, respectively. To our surprise, we found RBDV at two plantations, and we first described GSV in our country.

Most of the vineyards were infected with different viral pathogens in parallel, and we couldn't find any correlation between the age of the plantation and the number of viruses and viroids present. However, the most infected plantation 12\_DF is about 100 years old (infected by 12 viral pathogens), the youngest, 1-year-old plantation 11\_SZHU contained 9 viral pathogens, which suggests that the infection is usually brought about by the use of infected propagation material. Moreover, the sanitary status of the MK plantation at Tokaj (14-18\_MK), containing different varieties, differed from one variety to another. The least infected variety was the rootstock variety Teleki-Kober 125 (infected only with GPGV and slightly with 1 viroid), whereas the most infected varieties were the frequently used Furmint and Hárslevelu clones (infected with eight viruses). This significant difference in viral status between the rows of the same plantation reduces the possibility of an on-site infection from the neighborhood by viral vectors.

## CONCLUSIONS

Combination of next generation sequencing, bioinformatics, and molecular biology techniques provide us a powerful new high-throughput diagnostic tool to monitor grapevine plantations and get a deep insight into the virus infection status of Hungarian vineyards.

During our comprehensive survey, beside routinely tested viruses, presence of viruses never described in Hungary before, was revealed. Phylogenetic analysis of partial sequence of the virus isolates suggests that infection usually happened through the infected propagation material, highlighting the importance of virus diagnostic surveys by more sensitive methods than the routinely used ones.

Traditional diagnostic methods can only answer the question if a particular pathogen is present in the sample or not. As a contrast sRNA NGS, as a metagenomics technique, seems to be a powerful new high-throughput diagnostic tool. Using this approach, presence of grapevine-infecting regulated viruses can be reliably determined. Randomly collected, pooled samples, instead of samples from selected symptomatic plants, offer an alternative, unbiased way to reveal the presence of viral pathogens in the vineyards, but uneven presence of viruses in the sampled plants could lower the concentration of the virus in the pooled library, which could lead to false negative result. However our results showed that false negative results originating from the presence of SNPs in the genome of the isolates, at the place of the genome where diagnostic primers were designed, can be avoided.

Surveys by high-throughput methods, like sRNA NGS, will continuously provide us information about the presence and importance of different viral pathogens and can support the development of new sensitive tests for their routine diagnostics. Moreover, NGS has shown to be superior, both in reliability and speed, to biological indexing in grapevine (Al Rwahnih et al., 2015). NGS could lead to frequent discovery of new viruses (Massart et al., 2017), but beside that it can become a diagnostic method itself, in the future. Although we show that sRNA NGS-based metagenomics is one of the most reliable virus diagnostic methods, our work emphasizes that before becoming a diagnostic tool, to be adopted in the certification protocols, not only drop in the sequencing cost, but standardization and improvement of the bioinformatics pipeline is highly needed. Regular check of stock collections of varieties and rootstock plantations with this new sensitive method would prevent false negative diagnostics of the propagating material and could help maintaining the health of the vineyards in the future.

## REFERENCES

- Al Rwahnih, M., Daubert, S., Golino, D., Islas, C., and Rowhani, A. (2015). Comparison of next-generation sequencing versus biological indexing for the optimal detection of viral pathogens in grapevine. *Phytopathology* 105, 758–763. doi: 10.1094/PHYTO-06-14-0165-R
- Al Rwahnih, M., Daubert, S., Golino, D., and Rowhani, A. (2009). Deep sequencing analysis of RNAs from a grapevine showing Syrah decline symptoms reveals a multiple virus infection that includes a novel virus. *Virology* 387, 395–401. doi: 10.1016/j.virol.2009.02.028
- Al Rwahnih, M., Daubert, S., Sudarshana, M. R., and Rowhani, A. (2013). Gene from a novel plant virus satellite from grapevine identifies a viral satellite lineage. *Virus Genes* 47, 114–118. doi: 10.1007/s11262-013-0921-3
- Baulcombe, D. (2004). RNA silencing in plants. *Nature* 431, 356–363. doi: 10.1038/nature02874
- Bertazzon, N., Filippin, L., Forte, V., and Angelini, E. (2016). Grapevine Pinot gris virus seems to have recently been introduced to vineyards in Veneto, Italy. *Arch. Virol.* 161, 711–714. doi: 10.1007/s00705-015-2718-2
- Bertazzon, N., Forte, V., Filippin, L., Causin, R., Maixner, M., and Angelini, E. (2017). Association between genetic variability and titre of Grapevine Pinot gris virus with disease symptoms. *Plant Pathol.* 66, 949–959. doi: 10.1111/ppa.12639
- Beuve, M., Candresse, T., Tannières, M., and Lemaire, O. (2015). First report of Grapevine redglobe virus (GRGV) in Grapevine in France. *Plant Dis.* 99, 422–422. doi: 10.1094/PDIS-10-14-1009-PDN

## AUTHOR CONTRIBUTIONS

NC: sRNA library preparation, molecular biology work, evaluation of data; JM, ES, and LeK: Bioinformatics analysis; ED: Molecular biology work; IB: sRNA library preparation; GS, LaK, GB, and GT: Evaluation of data, manuscript preparation; TD and JB: Experimental design, evaluation of data, manuscript preparation; EV: Experimental design, evaluation of data, writing of the manuscript.

## ACKNOWLEDGMENTS

We thank Éva Pájtli and Tibor Csorba for the supply of grapevine samples from Tokaj. Our work was supported by the Hungarian Ministry of Agriculture, the KTIA\_AIK\_12–1-2013-0001 project and by a Hungarian Scientific Research Fund (OTKA; K119783 and K119701) grant from the National Research, Development and Innovation Office. NC participated in the Program for Reinforcement of Scientists of the Hungarian Ministry of Agriculture and is a PhD student of the Fesztetics Doctoral School at the University of Pannonia. ED is a PhD student of the Doctoral School of Biological Sciences at Szent István University. The author(s) would like to acknowledge the contribution of the COST Action FA 1407–DIVAS (Deep Investigation of Virus Associated Sequences).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2018.00122/full#supplementary-material>

- Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170
- Boonham, N., Kreuze, J., Winter, S., van der Vlugt, R., Bergervoet, J., Tomlinson, J., et al. (2014). Methods in virus diagnostics: from ELISA to next generation sequencing. *Virus Res.* 163, 262–268. doi: 10.1016/j.virusres.2013.12.007
- Borges, F., and Martienssen, R. A. (2015). The expanding world of small RNAs in plants. *Nat. Rev. Mol. Cell Biol.* 16, 727–741. doi: 10.1038/nrm4085
- Boscica, D., Sabanadzovic, S., Savino, V., Kyriakopoulou, P. E., Martelli, G. P., and Laforteza, R. (1994). A non-mechanically transmissible isometric virus-associated with asteroid mosaic of the Grapevine. *Vitis* 33, 101–102.
- Candresse, T., Faure, C., Theil, S., Beuve, M., Lemaire, O., Spilmont, A. S., et al. (2017a). First report of Grapevine asteroid mosaic-associated virus Infecting Grapevine (*Vitis vinifera*) in France. *Plant Dis.* 101, 1061–1061. doi: 10.1094/PDIS-01-17-0012-PDN
- Candresse, T., Marais, A., Theil, S., Faure, C., Lacombe, T., and Boursiquot, J. M. (2017b). Complete nucleotide sequence of an isolate of Grapevine satellite virus and evidence for the presence of multimeric forms in an infected Grapevine. *Genome Announc* 5:e01703-16. doi: 10.1128/genomeA.01703-16
- Coetzee, B., Freeborough, M. J., Maree, H. J., Celton, J. M., Rees, D. J., and Burger, J. T. (2010). Deep sequencing analysis of viruses infecting grapevines: virome of a vineyard. *Virology* 400, 157–163. doi: 10.1016/j.virol.2010.01.023
- Cretazzo, E., Padilla, C. V., and Velasco, L. (2017). First report of Grapevine Red Globe virus in grapevine in Spain. *Plant Dis.* 101, 264–265. doi: 10.1094/PDIS-06-16-0932-PDN

- Cseh, E., Takacs, A. P., Gaborjanyi, R., Palkovics, L., and Kocsis, L. (2013). Rt-PCR analysis and evolutionary relationship of some Hungarian Grapevine leafroll associated virus 1 and 3 Isolates. *Am. J. Plant Sci.* 4:5. doi: 10.4236/ajps.2013.410250
- Czotter, N., Szabó, E., Molnar, J., Kocsis, L., Deák, T., Bisztray, Gy., et al. (2015). First description of Grapevine Syrah virus 1 in vineyards of Hungary. *J. Plant Pathol.* 97, S74. doi: 10.4454/JPP.V97I4SUP.022
- Digiario, M., Yahyaoui, E., Martelli, G. P., and Elbeaino, T. (2015). The sequencing of the complete genome of a Tomato black ring virus (TBRV) and of the RNA2 of three Grapevine chrome mosaic virus (GCMV) isolates from grapevine reveals the possible recombinant origin of GCMV. *Virus Genes* 50, 165–171. doi: 10.1007/s11262-014-1094-4
- Donaire, L., Wang, Y., Gonzalez-Ibeas, D., Mayer, K. F., Aranda, M. A., and Llave, C. (2009). Deep-sequencing of plant viral small RNAs reveals effective and widespread targeting of viral genomes. *Virology* 392, 203–214. doi: 10.1016/j.virol.2009.07.005
- Eichmeier, A., Kominkova, M., Kominek, P., and Baranek, M. (2016). Comprehensive virus detection using next generation sequencing in grapevine vascular tissues of plants obtained from the wine regions of Bohemia and Moravia (Czech Republic). *PLoS ONE* 11:e0167966. doi: 10.1371/journal.pone.0167966
- Elbeaino, T., Kiyi, H., Boutarfa, R., Minafra, A., Martelli, G. P., and Digiario, M. (2014). Phylogenetic and recombination analysis of the homing protein domain of grapevine fanleaf virus (GFLV) isolates associated with 'yellow mosaic' and 'infectious malformation' syndromes in grapevine. *Arch. Virol.* 159, 2757–2764. doi: 10.1007/s00705-014-2138-8
- El Beaino, T., Sabanadzovic, S., Digiario, M., Abou Ghanem-Sabanadzovic, N., Rowhani, A., Kyriakopoulou, P. E., et al. (2001). Molecular detection of Grapevine fleck virus-like viruses. *Vitis* 40, 65–68.
- Esteves, F., Santos, M. T., Eiras-Dias, J. E., and Fonseca, F. (2013). Molecular data mining to improve antibody-based detection of Grapevine leafroll-associated virus 1 (GLRaV-1). *J. Virol. Methods* 194, 258–270. doi: 10.1016/j.jviromet.2013.09.004
- Fajardo, T. V. M., Silva, F. N., Eiras, M., and Nickel, O. (2017). High-throughput sequencing applied for the identification of viruses infecting grapevines in Brazil and genetic variability analysis. *Trop. Plant Pathol.* 42, 250–260. doi: 10.1007/s40858-017-0142-8
- Fonseca, F., Duarte, V., Santos, M. T., Brazao, J., and Eiras-Dias, E. (2016). First molecular characterization of grapevine virus B (GVB) in Portuguese grapevine cultivars and improvement of the RT-PCR detection assay. *Arch. Virol.* 161, 3535–3540. doi: 10.1007/s00705-016-3024-3
- Gambino, G., Cuzzo, D., Fasoli, M., Pagliarani, C., Vitali, M., Boccacci, P., et al. (2012). Co-evolution between Grapevine rupestris stem pitting-associated virus and *Vitis vinifera* L. leads to decreased defence responses and increased transcription of genes related to photosynthesis. *J. Exp. Bot.* 63, 5919–5933. doi: 10.1093/jxb/ers244
- Gambino, G. P., and Gribaudo, I. (2008). A Rapid and effective method for RNA extraction from different tissues of grapevine and other woody plants. *Phytochem. Anal.* 19, 520–525. doi: 10.1002/pca.1078
- Giampetruzzi, A., Roumi, V., Roberto, R., Malossini, U., Yoshikawa, N., La Notte, P., et al. (2012). A new grapevine virus discovered by deep sequencing of virus- and viroid-derived small RNAs in Cv Pinot gris. *Virus Res.* 163, 262–268. doi: 10.1016/j.virusres.2011.10.010
- Glasa, M., Predajna, L., Kominek, P., Nagyova, A., Candresse, T., and Olmos, A. (2014). Molecular characterization of divergent grapevine Pinot gris virus isolates and their detection in Slovak and Czech grapevines. *Arch. Virol.* 159, 2103–2107. doi: 10.1007/s00705-014-2031-5
- Glasa, M., Predajna, L., Soltys, K., Sabanadzovic, S., and Olmos, A. (2015). Detection and molecular characterisation of Grapevine Syrah virus-1 isolates from Central Europe. *Virus Genes* 51, 112–121. doi: 10.1007/s11262-015-1201-1
- Glasa, M., Predajna, L., Soltys, K., Sihelska, N., Nagyova, A., Wetzels, T., et al. (2017). Analysis of Grapevine rupestris stem pitting-associated virus in Slovakia reveals differences in intra-host population diversity and naturally occurring recombination events. *Plant Pathol. J.* 33, 34–42. doi: 10.5423/PPJ.OA.07.2016.0158
- Goszczynski, D. E. (2014). Complete genome sequence of a natural mutant of grapevine virus A (GVA). *Arch. Virol.* 159, 2523–2528. doi: 10.1007/s00705-014-2085-4
- Kominek, P., Glasa, M., and Bryxiova, M. (2005). Analysis of the molecular variability of Grapevine leafroll-associated virus 1 reveals the presence of two distinct virus groups and their mixed occurrence in grapevines. *Virus Genes* 31, 247–255. doi: 10.1007/s11262-005-3236-1
- Kreuzer, J. F., Perez, A., Untiveros, M., Quispe, D., Fuentes, S., Barker, I., et al. (2009). Complete viral genome sequence and discovery of novel viruses by deep sequencing of small RNAs: a generic method for diagnosis, discovery and sequencing of viruses. *Virology* 388, 1–7. doi: 10.1016/j.virol.2009.03.024
- Li, H. (2011). A statistical framework for SNP calling, mutation discovery, association mapping and population genetic parameter estimation from sequencing data. *Bioinformatics* 27, 2987–2993. doi: 10.1093/bioinformatics/btr509
- Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25, 1754–1760. doi: 10.1093/bioinformatics/btp324
- Malagnini, V., de Lillo, E., Saldarelli, P., Beber, R., Duso, C., Raiola, A., et al. (2016). Transmission of grapevine Pinot gris virus by Colomerus vitis (Acari: Eriophyidae) to grapevine. *Arch. Virol.* 161, 2595–2599. doi: 10.1007/s00705-016-2935-3
- Martelli, G. P. (2014). Directory of virus and virus-like diseases of the grapevine and their agents. *J. Plant Pathol.* 96, 1–136. doi: 10.4454/JPP.V96I1SUP
- Massart, S., Candresse, T., Gil, J., Lacomme, C., Predajna, L., Ravnkar, M., et al. (2017). A framework for the evaluation of biosecurity, commercial, regulatory, and scientific impacts of plant viruses and viroids identified by NGS technologies. *Front. Microbiol.* 8:45. doi: 10.3389/fmicb.2017.00045
- Mavric, I., Virscek, M., Marn, M. V., and Koron, D. (2003). First report of raspberry bushy dwarf virus on red raspberry and grapevine in Slovenia. *Plant Dis.* 87, 1148–1148. doi: 10.1094/PDIS.2003.87.9.1148D
- Meng, B., Rebelo, A. R., and Fisher, H. (2006). Genetic diversity analyses of grapevine Rupestris stem pitting-associated virus reveal distinct population structures in scion versus rootstock varieties. *J. Gen. Virol.* 87(Pt 6), 1725–1733. doi: 10.1099/vir.0.81533-0
- Molnar, A., Csorba, T., Lakatos, L., Varallyay, E., Lacomme, C., and Burgyan, J. (2005). Plant virus-derived small interfering RNAs originate predominantly from highly structured single-stranded viral RNAs. *J. Virol.* 79, 7812–7818. doi: 10.1128/JVI.79.12.7812-7818.2005
- Morelli, M., Minafra, A., Boscia, D., and Martelli, G. P. (2011). Complete nucleotide sequence of a new variant of grapevine rupestris stem pitting-associated virus from southern Italy. *Arch. Virol.* 156, 543–546. doi: 10.1007/s00705-011-0936-9
- Morgulis, A., Coulouris, G., Raytselis, Y., Madden, T. L., Agarwala, R., and Schaffer, A. A. (2008). Database indexing for production MegaBLAST searches. *Bioinformatics* 24, 1757–1764. doi: 10.1093/bioinformatics/btn322
- Navarro, B., Pantaleo, V., Gisel, A., Moxon, S., Dalmay, T., Bisztray, G., et al. (2009). Deep Sequencing of viroid-derived small RNAs from Grapevine provides new insights on the role of RNA silencing in plant-viroid interaction. *PLoS ONE* 4:e7686. doi: 10.1371/journal.pone.0007686
- Pantaleo, V., Saldarelli, P., Miozzi, L., Giampetruzzi, A., Gisel, A., Moxon, S., et al. (2010). Deep sequencing analysis of viral short RNAs from an infected Pinot Noir grapevine. *Virology* 408, 49–56. doi: 10.1016/j.virol.2010.09.001
- Parent, J.-S., Martinez de Alba, A. E., and Vaucheret, H. (2012). The origin and effect of small RNA signaling in plants. *Front. Plant Sci.* 3:179. doi: 10.3389/fpls.2012.00179
- Plesko, I. M., Marn, M. V., Nyerges, K., and Lazar, J. (2012). First report of raspberry bushy dwarf virus infecting grapevine in Hungary. *Plant Dis.* 96, 1582–1582. doi: 10.1094/PDIS-04-12-0383-PDN
- Plesko, I. M., Marn, M. V., and Zezlina, I. (2014). First report of grapevine pinot gris virus infecting grapevine in Slovenia. *Plant Dis.* 98, 1014–1015. doi: 10.1094/PDIS-11-13-1137-PDN
- Reynard, J. S., Brodard, J., Dubuis, N., Yobregat, O., Kominek, P., Schumpp, O., et al. (2017). First report of Grapevine rupestris vein feathering virus in Swiss Grapevines. *Plant Dis.* 101, 1062. doi: 10.1094/PDIS-01-17-0140-PDN
- Sabanadzovic, S., Abou-Ghanem, N., Castellano, M. A., Digiario, M., and Martelli, G. P. (2000). Grapevine fleck virus-like viruses in Vitis. *Arch. Virol.* 145, 553–565. doi: 10.1007/s007050050046
- Sabanadzovic, S., Ghanem-Sabanadzovic, N. A., and Goralbanya, A. E. (2009). Permutation of the active site of putative RNA-dependent RNA polymerase



- in a newly identified species of plant alpha-like virus. *Virology* 394, 1–7. doi: 10.1016/j.virol.2009.08.006
- Saldarelli, P., Giampetruzzi, A., Morelli, M., Malossini, U., Pirolò, C., Bianchedi, P., et al. (2015). Genetic Variability of Grapevine Pinot gris virus and its association with grapevine leaf mottling and deformation. *Phytopathology* 105, 555–563. doi: 10.1094/PHYTO-09-14-0241-R
- Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., et al. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* 7:539. doi: 10.1038/msb.2011.75
- Szittyá, G., Moxon, S., Pantaleo, V., Toth, G., Pilcher, R. L. R., Moulton, V., et al. (2010). Structural and functional analysis of viral siRNAs. *PLoS Pathog.* 6:e1000838. doi: 10.1371/journal.ppat.100838
- Tamura, K., Stecher, G., Peterson, D., FilipSKI, A., and Kumar, S. (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30, 2725–2729. doi: 10.1093/molbev/mst197
- Voncina, D., Al Rwahnih, M., Rowhani, A., Gouran, M., and Almeida, R. P. P. (2017). Viral diversity in autochthonous croatian grapevine cultivars. *Plant Dis.* 101, 1230–1235. doi: 10.1094/PDIS-10-16-1543-RE
- Wu, Q. F., Wang, Y., Cao, M. J., Pantaleo, V., Burgyan, J., Li, W. X., et al. (2012). Homology-independent discovery of replicating pathogenic circular RNAs by deep sequencing and a new computational algorithm. *Proc. Natl. Acad. Sci. U.S.A.* 109, 3938–3943. doi: 10.1073/pnas.1117815109
- Wu, Q., and Habili, N. (2017). The recent importation of Grapevine Pinot gris virus into Australia. *Virus Genes* 53, 935–938. doi: 10.1007/s11262-017-1475-6
- Xiao, H., and Meng, B. (2016). First report of Grapevine asteroid mosaic-associated virus and Grapevine rupestris vein feathering virus in Grapevines in Canada. *Plant Dis.* 100, 2175–2175. doi: 10.1094/PDIS-03-16-0413-PDN
- Zerbino, D. R., and Birney, E. (2008). Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res.* 18, 821–829. doi: 10.1101/gr.074492.107
- Zhang, Y., Singh, K., Kaur, R., and Qiu, W. (2011). Association of a novel DNA virus with the grapevine vein-clearing and vine decline syndrome. *Phytopathology* 101, 1081–1090. doi: 10.1094/PHYTO-02-11-0034
- Zhao, H., Zhao, K., Wang, J., Chen, X., Chen, Z., Cai, R., et al. (2015). Comprehensive analysis of dicer-like, argonaute, and RNA-dependent RNA polymerase gene families in Grapevine (*Vitis Vinifera*). *J. Plant Growth Regul.* 34, 108–121. doi: 10.1007/s00344-014-9448-7

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Czotter, Molnar, Szabó, Demian, Kontra, Baksa, Szittyá, Kocsis, Deak, Bisztray, Tusnady, Burgyan and Varallyay. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.