

Article

Not peer-reviewed version

---

# Integrating High-Throughput Sequencing Data from Herbarium and Contemporary Samples Reveal a Novel Carlavirus Long-Established in European Beech

---

[Pier P.M. de Koning](#)\*, Anne K.J. Giesbers, [Susanne von Bargaen](#), Stephanie T.G. Rensen, [Carmen Büttner](#), Marcel Westenberg, Marleen Botermans, [Artemis Rumbou](#)

Posted Date: 26 March 2026

doi: 10.20944/preprints202603.2172.v1

Keywords: beech; *Fagus sylvatica*; carlavirus; HTS; herbarium; data sharing; virome; baseline studies



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

# Integrating High-Throughput Sequencing Data from Herbarium and Contemporary Samples Reveal a Novel Carlavirus Long-Established in European Beech

Pier P.M. de Koning <sup>1</sup>, Anne K.J. Giesbers <sup>1</sup>, Susanne von Bargaen <sup>2</sup>, Stephanie T.G. Rensen <sup>1</sup>, Carmen Büttner <sup>2</sup>, Marcel Westenberg <sup>1</sup>, Marleen Botermans <sup>1</sup> and Artemis Rumbou <sup>2</sup>

<sup>1</sup> Netherlands Institute for Vectors, Invasive plants and Plant health (NIVIP), National Plant Protection Organization (NPPO), Netherlands Food and Consumer Product Safety Authority (NVWA), Geertjesweg 15, 6706 EA, Wageningen, the Netherlands

<sup>2</sup> Humboldt-University of Berlin, Faculty of Life Sciences, Albrecht Daniel Thaer-Institute of Agricultural and Horticultural Sciences, Division of Plant Pathology and Plant Protection, Lentzeallee 55/57, 14195, Berlin, Germany

\* Correspondence: p.p.m.dekoning@nvwa.nl

## Abstract

High-throughput sequencing (HTS) was applied to investigate the virome of European beech (*Fagus sylvatica*) from asymptomatic leaves and leaves exhibiting chlorosis, line patterns and malformation. Total RNA extracted from six samples, including herbarium material collected in 1967 and 1968 and contemporary samples from France, Germany, and the Netherlands, was subjected to Illumina sequencing followed by *de novo* assembly, sequence similarity searches and phylogenetic analyses. In each sample, contigs belonging to a previously undescribed virus within the genus *Carlavirus* were obtained. The virus was tentatively named beech carlavirus. The detection of the virus over more than five decades and in three European countries indicates its long-term and a probable wider occurrence and circulation. Moreover, its prolonged unnoticed presence suggests that it does not induce noticeable and acute disease outbreaks. These findings underscore the value of integrating historical and recent field samples through collaborative data sharing to improve insight into virus diversity and ecology in forest trees.

**Keywords:** beech; *Fagus sylvatica*; carlavirus; HTS; herbarium; data sharing; virome; baseline studies

## 1. Introduction

Historically, forest trees have received little attention in plant virus research, possibly because they were difficult to investigate using classical diagnostic methods. European beech (*Fagus sylvatica*, also known as common beech) is an example of a tree species whose virome remains largely unexplored. Early reports of viral infections in beech in Europe emerged in the 1960s [1]. This initiated systematic investigations into virus–host interactions in beech. Nienhaus and Castello [2] summarized observations of virus-like symptoms in broadleaf forest trees. In degenerating beech trees at different forest sites of West Germany, several virus species were detected, among which: bean yellow mosaic virus (BYMV; *Potyvirus phaseoliteum*), brome mosaic virus (BMV; *Bromovirus BMV*) and cherry leaf roll virus (CLRV; *Nepovirus avii*) [3,4] as the most commonly reported [5–7]. These viruses were initially identified by electron microscopy and a serological test and several were further characterized. However, the causal agent of the reported leaf symptoms and degeneration was not proven. Collectively, these foundational studies represent the first viral infections in

European beech and laid the groundwork for contemporary research. No other viruses have been described from European beech in the decades following these initial discoveries [8,9].

Over the past two decades, high-throughput sequencing (HTS) has become widely adopted in plant virology as a powerful method to detect and/or identify plant viruses, including novel viruses, without requiring prior knowledge. HTS has, for instance been applied to sequence historical virus collections (e.g., herbarium material) enabling the recovery and characterization of viral genomes that were previously undetectable using classical diagnostic methods and for which no sequence data were available [10,11]. Moreover, HTS is widely used in virus baseline studies to characterize the virome in both symptomatic and asymptomatic plants, including studies focused specifically on forest tree species [12].

Combining sequence data from historical material (e.g., herbarium specimens) with contemporary field samples from baseline studies provides insights into the geographical distribution of viruses, documents their long-term presence, refines potential associations between infection and symptom expression, and strengthens pest risk assessments as well as evidence-based regulatory decision-making [13]. Recent publications have highlighted the need for baseline studies and recommend systematic investigations of virus diversity and distribution across countries, hosts and ecosystems [14,15]. In response, several initiatives have been launched to perform such baseline studies [16–19].

Although research on viruses in forest and urban trees intensified over subsequent decades, knowledge of virus occurrence remains particularly limited compared with fruit trees [8,9]. To address this gap, the Euphresco project VIRNOTree was launched in 2023 with the aim of characterizing the virome of forest trees. The present study, conducted within VIRNOTree, reports the discovery of a putative novel carlavirus (family *Betaflexiviridae*, order *Tymovirales*) infecting European beech. The virus appears to be geographically widespread, having been detected in France, Germany and the Netherlands, and its presence can be traced back to 1967.

## 2. Materials and Methods

### 2.1. Sampling Information

Six European beech leaf samples from France, Germany and the Netherlands were included in this study (Table 1). Three field samples (11647, 11648, 40757) were collected based on the presence of virus-like symptoms from Germany and France. Two samples were obtained from herbarium specimens preserved in the virus symptom herbarium collection at NIVIP, the Netherlands. One additional sample (6165949) did not exhibit virus-like symptoms and was collected within the framework of the Euphresco VIRNOTree project from the Netherlands.

**Table 1.** Overview of the samples included in this study.

No.	Sample name	Year	Pool/single	Sample type	Country	Location
1	11647	2014	Single	Leaves	Germany	Siebengebirge
2	11648	2014	Single	Leaves	Germany	Siebengebirge
3	40757	2018	Pool (3 trees)	One leaf per tree	France	Corsica
4	6165949	2024	Pool (5 trees)	One leaf per tree	The Netherlands	Wageningen
5	WAG0452949	1967	Single	Herbarium, leaves	The Netherlands	Groningen (city)
6	WAG0453175	1968	Single	Herbarium,	The Netherlands	Paterswolde

leaves

Netherlands

## 2.2. RNA Extraction, Library Preparation and Sequencing

### 2.2.1. Samples 1 to 3

Samples were sequenced using Illumina technology as described in Rumbou, *et al.* [20]. Total RNA was isolated from 0,5-1 g of leaf tissue using the InviTrap® Spin Plant RNA Mini Kit (STRATEC Molecular, Germany), followed by removal of remaining DNA with rDNase according to the supplier protocol (Macherey-Nagel, Germany) and RNA purification using NucleoSpin® RNA Clean-up (Macherey-Nagel, Germany). Ribosomal RNA depletion was performed using the RiboMinus Plant Kit for RNA-Seq (Invitrogen). One to two micrograms of RiboMinus RNA of each sample were used for cDNA synthesis with the Maxima H Minus double-stranded cDNA synthesis Kit (Thermo Scientific) primed with random hexamers. Between 1 and 2 µg of purified double-stranded cDNA from each sample was submitted to BaseClear (Leiden, the Netherlands) for RNA-seq analysis. Libraries were sequenced on an HiSeq 2500 platform (Illumina) generating 100 nt paired-end reads, yielding approximately 50–100 Mb of data per sample.

### 2.2.2. Samples 4 to 6

Samples were sequenced using Illumina technology as described in Eppo [21]. Total RNA was extracted from approximately 1g of leaf tissue using the RNeasy Plant Mini Kit (Qiagen, Venlo, the Netherlands). Library preparation and sequencing were performed at GenomeScan (Leiden, the Netherlands). Ribosomal RNA was depleted with the QIAseq FastSelect rRNA Plant Kit (Qiagen), after which sequencing libraries were prepared using the NEBNext Ultra II Directional RNA library Prep Kit (New England Biolabs, Ipswich, MA, USA). Libraries were sequenced on a NovaSeq 6000 platform (Illumina), generating 150 nt paired-end reads and at least 2 Gb data per sample.

## 2.3. HTS Pipeline and Sequence Analysis

Raw sequencing reads were imported into CLC Genomics workbench (v 25.0.1, Qiagen) and processed using the standard workflow [21]. Reads were quality trimmed and *de novo* assembled. Contigs were retained for downstream analyses when they exceeded 100 nt with minimum read depth of 10. Consensus sequences were screened against locally installed NCBI databases using MegaBLAST (nt) and DIAMOND [22] (nr). Taxonomic visualization of BLAST outputs was carried out with Krona using a bitscore threshold of 25 [23]. Putative carlavirus contigs were queried on BLASTn against the NCBI core\_nt database on 2026-02-04 to confirm their novelty. Open reading frames (ORFs) were then predicted in Geneious Prime (v 2025.1.2) and translated protein sequences were assessed by BLASTp against the NCBI nr\_cluster\_seq database.

## 2.4. Phylogenetic Analysis

A reference dataset of carlavirus sequences was compiled from NCBI RefSeq, the ICTV Virus Metadata Resource (VMR\_MSL40.v2.20251013) and closely related taxa identified from the BLASTn output (see chapter 3.3). Redundant entries, (e.g. duplicates present in both VMR and RefSeq) were removed. The complete list of reference sequences is provided in Table S1. Reference sequences and the putative novel beech carlavirus sequences were aligned using MAFFT v 7.490 [24] with default parameters implemented in Geneious Prime v 2025.1.2. The coat protein (CP) and replicase polyprotein ORFs were translated into amino-acid sequences and aligned separately using MAFFT with default settings. Maximum-likelihood phylogenetic trees were inferred from (I) (near-)complete genomic nucleotide sequences and (II) the amino-acid alignments of the CP and replicase polyprotein amino acid sequences with IQ-Tree v2.3.6 [25], using the substitution models, as selected by ModelFinder [26], GTR+F+I+G4 for the genomic nucleotide dataset and the Q.pfam+F+I+G4 substitution model for both protein datasets. The RefSeq sequence of cherry virus B (*Foveavirus betavii*,

NC\_076603, family: *Betaflexiviridae*) was used as outgroup. Branch support was determined using the ultrafast bootstrap (UFboot) with 10,000 replicates [27].

### 2.5. Datamining using Serratus and NCBI SRA

To assess whether the putative novel carlavirus was present in publicly available sequencing data, we screened datasets in Serratus and the NCBI Sequence Read Archive (SRA). For Serratus, the “palmID: Viral-RdRP Analysis” (<https://serratus.io/palmid>) tool was queried using the replicase polyprotein amino-acid sequence of isolate WAG0452949 as the input, applying default parameters. In parallel, the NCBI SRA (<https://www.ncbi.nlm.nih.gov/sra>) was queried on 2026-01-29 using the following search string to retrieve RNA-seq runs from European beech published from 2021 onward: (Fagus sylvatica[Organism]) AND (RNA-Seq[Strategy]) AND (“2021/01/01”[Publication Date] : “3000”[Publication Date]). This time restriction was applied because Serratus indexing partially overlaps with SRA content and to focus on more recent datasets. The nucleotide sequence of the putative novel carlavirus isolate WAG0452949 was subsequently used as a query for BLAST searches against the retrieved SRA dataset.

### 2.6. Development of an RT-PCR Test

To establish a diagnostic test, an RT-PCR was developed using two primer pairs (Table 2). Primer pair S14-7254F/Carla-S14-7751R amplifies a 498 bp fragment spanning the TGB3-CP region (2 min at 94°C, 35 cycles of 30 s at 94°C, 30 s at 58°C, 30 s 72°C, 5 min final extension), whereas primer pair Carla-S14-7618F/Carla-S14-7751R amplifies a 130 bp fragment within the CP gene (2 min at 94°C, 35 cycles of 20 s at 94°C, 20 s at 60°C, 20 s 72°C, 5 min final extension). The RT-PCR using Maxima H Minus RT reverse Transcriptase and DreamTaq polymerase (Thermo Fisher Scientific, Schwerte, Germany) was performed on samples 1, 2 and 3 according to manufacturer’s instructions. Amplicons from both primer pairs were subjected to Sanger sequencing (Macrogen Europe, Amsterdam, the Netherlands) to confirm target specificity.

**Table 2.** RT-PCR primers used for detection of the novel beech carlavirus.

Set	Name	Gene	Sequence
1	Carla-S14-7254F	TGB3-CP	GCTGGGTATGCTAGTGGACT
	Carla-S14-7751R		GTTTGAGATCATCGACAGGCT
2	Carla-S14-7618F	CP	CAGTGATGAAATGTTGGAGCAG
	Carla-S14-7751R		GTTTGAGATCATCGACAGGCT

### 2.7. Testing of European Beech Trees with RT-PCR

Leaf samples (E58078, E58081) from two European beech trees growing in Göttingen, Germany and displaying virus-like symptoms were tested by RT-PCR. Total RNA was extracted according to Boom, *et al.* [28] and subjected to the RT-PCR with carlavirus-specific primer sets as described in the previous chapter (Table 2).

## 3. Results

Five symptomatic and one asymptomatic leaf samples from European beech were analyzed by HTS (Table 3, Figure 1). The HTS pipeline detected sequences corresponding to a carlavirus. In five samples, a single viral contig was reconstructed, whereas in sample 2, three distinct contigs were reconstructed, representing different genotypes. In total, eight putative novel carlavirus sequences were obtained, ranging from 8,724 to 8,872 nucleotides in length. No additional viral contigs were detected. All sequences have been deposited in GenBank under accession numbers PX705376-PX705383, with corresponding SRA BioSamples SAMN53882179-SAMN53882184 and BioProject PRJNA1379522. An overview of sequencing and assembly data is shown in Table 4.

**Table 3.** Overview of symptoms observed on the leaves of European beech trees in France, Germany and the Netherlands.

No.	Sample name	Leaf symptom description
1	11647	Curling and mottling
2	11648	Curling and interveinal chlorosis
3	40757	Chlorotic leaf tips and chlorosis along veins
4	6165949	No symptoms observed
5	WAG0452949	V-shaped yellow and orange patterns and chlorotic spots
6	WAG0453175	V-shaped yellow and orange patterns and chlorotic spots

A



B



C



D



**Figure 1.** Virus-like symptoms observed on European beech leaves analyzed in this study. (A) Sample 11647 (Siebengebirge, Germany; 2014) displaying leaf curling and mottling. (B) Sample 11648 (Siebengebirge, Germany; 2014) showing interveinal chlorosis and leaf curling. (C) Herbarium specimen WAG0452949 (Groningen, the Netherlands; 1967) and (D) herbarium specimen WAG0453175 (Paterswolde, the Netherlands; 1968) both showing V-shaped yellow and orange patterns and chlorotic spots. The images were taken in the 1960s prior to preparing the herbarium specimen.

**Table 4.** Overview and HTS metrics of eight putative novel carlavirus contigs identified in European beech.

No.	SRA BioSample	Total reads sample	GenBank accession	Total mapped viral reads	Mean read depth (x)	Sequenc e length (nt)	Carlavirus specific RT- PCR
1	SAMN538 82179	787.486	PX705376	110.642	1547	8774	Positive
2	SAMN538 82180	713.488	PX705377	4.487	63	8800	Positive
			PX705378	17.283	245	8784	
			PX705379	2.567	37	8724	
3	SAMN538 82181	6.930.926	PX705380	9.579	155	8871	Positive
4	SAMN538 82182	16.088.034	PX705381	17.862	298	8872	Not tested
5	SAMN538 82183	15.352.090	PX705382	66.216	1010	8871	Not tested
6	SAMN538 82184	14.169.406	PX705383	55.277	920	8872	Not tested

The putative novel carlavirus nucleotide sequences were queried using BLASTn. The closest match was birch carlavirus (NC\_076638.1), sharing 67-69% pairwise nucleotide identity and 69-75% query coverage (Table 5). Genome annotation identified six ORFs typical for members of the genus *Carlavirus* encoding a replicase polyprotein, the triple gene block proteins (TGB1-TGB3), the CP and a ribonucleoprotein (RBP) (Fig. 2). BLASTp analyses of the replicase polyprotein and CP yielded the highest similarity to birch carlavirus proteins (YP\_010799375.1 and YP\_010799379.1) with only 66-67% and 61-63% amino-acid identity, respectively (Table 5). According to the species demarcation criteria for carlaviruses (<80% amino-acid identity in either the replicase polyprotein or the CP) these results indicate that the recovered sequences represent a putative novel carlavirus species [29].

**Table 5.** Overview of the NCBI BLASTn/BLASTp results for eight genomic sequences and their corresponding replicase polyprotein and coat protein amino-acid queries; birch carlavirus was the top hit in all cases (E-value of 0.0; NC\_076638.1, YP\_010799375.1, YP\_010799379.1).

No.	GenBank accession	Genomic sequence	or protein	Pairwise identity (%)	Query coverage (%)
1	PX705376	complete genome	nt	68.06	80
		replicase polyprotein	aa	67.39	100
		coat protein	aa	62.16	96

		complete genome	nt	68.75	76
	PX705377	replicase polyprotein	aa	66.52	100
		coat protein	aa	61.56	96
2	PX705378	complete genome	nt	69.28	74
		replicase polyprotein	aa	67.10	100
		coat protein	aa	62.16	96
	PX705379	complete genome	nt	69.33	69
		replicase polyprotein	aa	66.68	100
		coat protein	aa	62.46	96
3	PX705380	complete genome	nt	68.45	82
		replicase polyprotein	aa	66.73	100
		coat protein	aa	63.30	96
4	PX705381	complete genome	nt	67.33	82
		replicase polyprotein	aa	66.42	100
		coat protein	aa	62.46	96
5	PX705382	complete genome	nt	68.61	75
		replicase polyprotein	aa	66.68	100
		coat protein	aa	61.86	96
6	PX705383	complete genome	nt	68.71	75
		replicase polyprotein	aa	66.60	100
		coat protein	aa	61.56	96

nt: nucleotide, aa: amino acid.



**Figure 2.** Genome organization of the putative novel carlavirus (isolate WAG0452949). The genome contains six ORFs encoding a replicase polyprotein, the triple gene block proteins (TGB1–TGB3), a coat protein (CP), and a ribonucleoprotein (RBP). Numbers above the genome indicate the start and end nucleotide positions of each ORF.

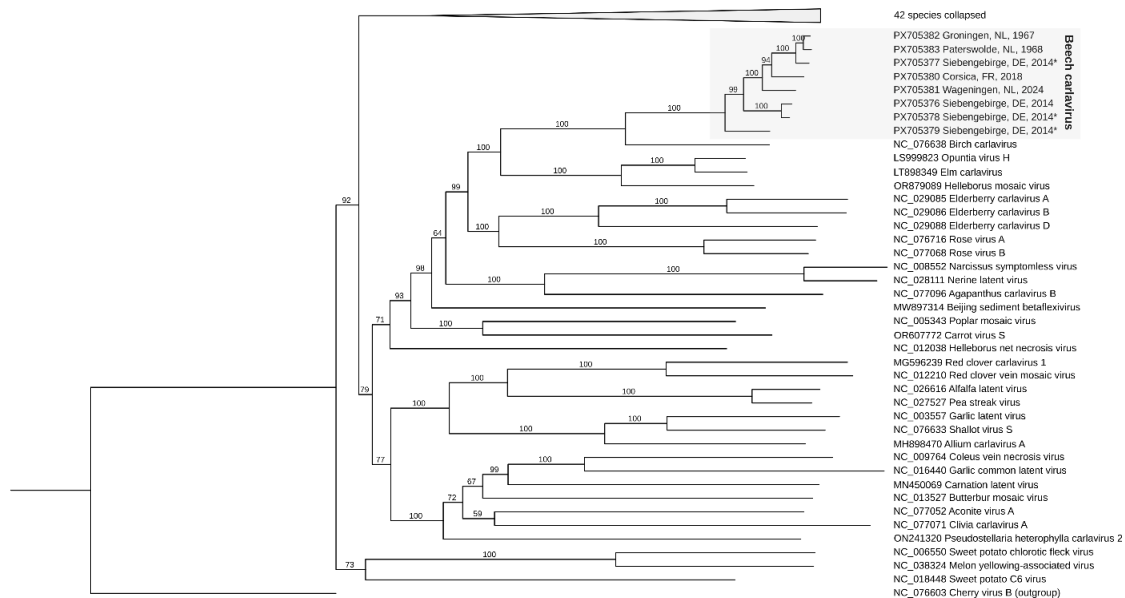
To determine whether the eight putative novel carlavirus sequences represent a single species, pairwise comparisons were performed using the genomic nucleotide sequences as well as the deduced amino acid sequences of the replicase polyprotein and CP. The nucleotide sequences shared 78.8% to 95.3% pairwise identity. While the amino acid sequences of the replicase polyprotein and CP shared 88.8 to 98.1% and 94.9 to 99.4% pairwise identity, respectively. Based on the established species demarcation criteria, these values indicate that all eight sequences belong to the same virus species.

In addition, three phylogenetic analyses were conducted using the (near-)complete genomic nucleotide sequences and the deduced amino-acid sequences of the replicase polyprotein and the CP of the putative novel carlavirus (Fig. 3). In all three analyses, the sequences obtained in this study clustered together as a distinct, monophyletic group, within the genus *Carlavirus*. Consistent with the pairwise sequence comparisons and species-demarcation criteria, these findings support the assignment of the virus to a new carlavirus species, tentatively designated beech carlavirus.

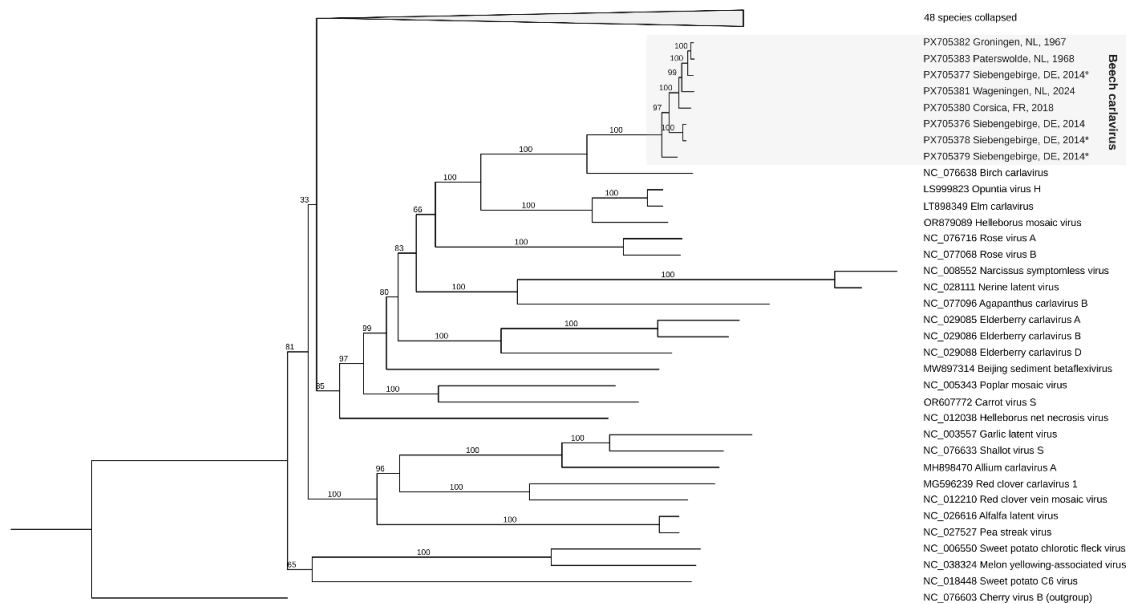
Beech carlavirus was also detected in samples 1, 2 and 3 by RT-PCR followed by Sanger sequencing (Table 4). Although the RT-PCR was not performed for samples 4 to 6, *in silico* primer-template mapping indicated that these sequences should also be amplified by the same primer pair.

Furthermore, beech carlavirus was detected in two leaf samples from European beech (E58078, E58081) displaying tip chlorosis and mottling (data not shown). Finally, data mining in the NCBI SRA (320 runs across seven BioProjects) did not result in additional detections of beech carlavirus.

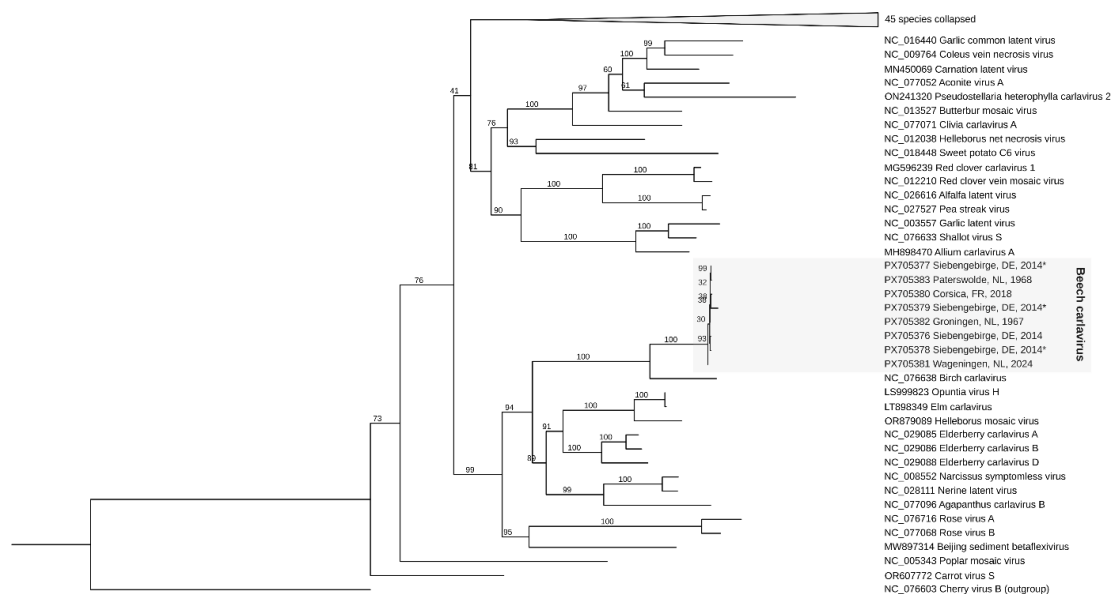
A



B



C



**Figure 3.** Maximum-likelihood phylogenetic trees including reference carlavirus species and the putative novel beech carlavirus (shaded in gray). Asterisks (\*) denote sequences originating from the same sample. Branch support values indicate UFBoot percentages based on 10,000 bootstraps replicates. (A) Tree inferred from (near-)complete genomic nucleotide sequences using the GTR+F+I+G4 substitution model. (B) Tree inferred from replicase polyprotein and (C) from coat protein amino-acid sequences using the Q.pfam+F+I+G4 substitution model.

#### 4. Discussion

Pre-publication data sharing between two institutes suggested the presence of the same novel carlavirus from beech in several samples, prompting further analysis. Eight putative novel carlavirus genomes were recovered from six European beech leaf samples originating from herbarium and contemporary material. Across three phylogenetic reconstructions, based on (near-)complete genomic nucleotide sequences and the deduced replicase polyprotein and coat protein amino-acid sequences, these isolates consistently grouped into a single, distinct clade within the genus *Carlavirus*. In line with established species demarcation criteria for carlaviruses, and supported by the closest sequence similarity to birch carlavirus, the virus is regarded as a distinct species, tentatively designated as beech carlavirus with the proposed species name *Carlavirus fagi*.

Notably, the eight beech carlavirus genomic sequences displayed substantial sequence diversity, with pairwise nucleotide identities ranging from 78.8% to 95.3%. Such divergence is consistent with patterns reported for other members of the family *Betaflexiviridae*, in which high levels of intraspecies variability are frequently observed [30]. Although revised species demarcation criteria have been proposed for *Betaflexiviridae* [31], these thresholds have not yet been formally adopted by the ICTV which underscores the continued need to interpret classification within a broader phylogenetic and genomic context.

The presence of beech carlavirus in herbarium specimens collected in the late 1960s alongside contemporary field samples indicate that this virus has been circulating in beech populations for at least five decades and is possibly widely distributed in Europe. Yet, it has not been formally described prior to this study. Early studies using electron microscopy, serology and PCR documented several viruses associated with beech; however, none reported carlaviruses [2,3,5–7]. Notably, Winter and Nienhaus [3] detected a potyvirus by electron microscopy, from which the particles are morphologically similar to a carlavirus, but confirmed their finding using a serological test specific for a potyvirus. Therefore, it can be concluded that Winter and Nienhaus [3] described another virus. In the last decade, research on forest tree viromes has increased, but no additional viruses have been formally reported in beech [9]. Vainio, *et al.* [32] included an image of symptomatic beech leaves,

resembling those observed in the herbarium samples in this study, suggesting a potential viral aetiology and possible infection with beech carlavirus, but did not provide virome data or diagnostic confirmation. To our knowledge, this is the first report of a complete genome from a virus infecting European beech.

Consistent with the absence from literature, beech carlavirus was not detected in publicly available RNA-sequencing datasets screened via NCBI SRA and Serratus. This indicates that beech carlavirus sequences have not been previously deposited in public RNA-seq datasets and that the sequences presented here may represent the first publicly deposited sequences of this virus. The absence of detections in NCBI SRA likely reflects the comparatively limited number of European beech RNA-seq runs currently available and the fact that these datasets were generated largely for purposes other than virus discovery; consequently, they may not include tissues, seasons, or symptom states in which high virus titres are expected. Moreover, the lack of detection in Serratus – where SRA datasets from hosts beyond beech were also screened – could be consistent with a restricted host range, but it may equally result from under-sampling of potential hosts. This interpretation aligns with the still limited baseline knowledge of forest tree viromes [32,33]. Overall, these results highlight a surveillance gap and underscores the need for targeted surveys, including systematic HTS-based virome profiling (symptomatic vs. asymptomatic trees, multiple tissues and timepoints) and/or assay-based screening (e.g., RT-PCR), to estimate the prevalence and distribution of beech carlavirus.

The prolonged, previously unrecognized presence of beech carlavirus suggests that it is unlikely to cause acute disease outbreaks and may frequently occur as a low-impact and/or latent infection, with visible symptoms potentially manifesting only occasionally, for instance when the tree is under stress. This hypothesis is supported by its detection in at least one asymptomatic sample. Nevertheless, beech carlavirus was identified as a sole viral agent detected in five trees exhibiting virus-like symptoms. At present, however, a causal relationship between infection and symptom expression cannot be inferred, especially given the heterogeneity of symptom phenotypes observed across samples.

The symptoms described in this study, particularly those reported from the herbarium samples from the Netherlands showing discolored V-shaped patterns across the secondary veins, can appear superficially similar but should not be confused with beech leaf disease (BLD). This disease is associated with the nematode *Litylenchus crenatae* subspecies *mccanii*, which affects European and American beech (*F. grandifolia*) in North America [34,35]. BLD is typically characterized by banding between the lateral leaf veins, leaf thickening, distortion and leathery texture which markedly differs from the symptoms recorded in this study.

Experience from other forest tree species indicates that variability in symptom expression may reflect complex viromes and mixed infections rather than a single causal agent. For example, individual birch trees have been reported to harbour multiple virus species and different genotypes of the same virus can co-occur within a single host [33]. While this study focusses on a virus found in single infection in beech, we cannot exclude the possibility that additional viruses may occur in symptomatic beech trees and that mixed infections could influence symptom expression. Furthermore, the potential impact of environmental factors, such as changing climatic conditions on virus prevalence and symptom development in beech is currently unclear. Taken together, these findings highlight the importance of virus surveillance and the integration of historical and contemporary samples to better understand virus diversity, ecology, and host interactions in forest trees.

The epidemiology of beech carlavirus remains to be determined. At present, there is no scientific evidence for its transmission pathways or dispersal mechanisms (e.g., insect vectors, seed, or pollen). Other carlaviruses are often transmitted vegetatively or via aphids, while seed transmission is uncommon [36]. As transmission via seed for carlaviruses is uncommon, vector mediated transmission via aphids may be a candidate pathway to investigate further.

## 5. Conclusions

This study demonstrates how HTS, applied to both historical herbarium and contemporary plant material, can substantially advance our understanding of virus–host associations in forest trees. Through pre-publication data sharing, we integrated results that were generated independently at two institutes, providing robust evidence for a novel carlavirus infecting European beech, tentatively named beech carlavirus. The virus appears was detected in samples originating from France, Germany and the Netherlands and has probably a wider distribution in Europe. Its detection over multiple decades, including in asymptomatic trees, is consistent with long-term presence and the potential for latent infections. Virus-like symptoms, including line patterns, were observed in infected beech, but their association with beech carlavirus remains to be formally established, for example through larger association studies. Based on the current knowledge, we do not have indications that beech carlavirus appears a concern for plant health. Still, its epidemiological impact in natural and managed beech populations as well as its presence on other continents, remains to be determined.

**Supplementary Materials:** The following supporting information can be downloaded at: Preprints.org, Table S1: reference sequences for phylogenetic analysis.

**Author Contributions:** Conceptualization, A.R., C.B., M.B. and P.K.; methodology, A.R., M.W. and S.B.; software, P.K.; formal analysis, A.R. and P.K.; investigation, A.G., A.R., C.B., S.B., S.R., M.B.; resources, A.G., S.R., M.B.; data curation, A.R.; writing—original draft preparation, P.K.; writing—review and editing, A.G., A.R., C.B., S.B., M.W. and P.K.; visualization, P.K.; supervision, M.B.; funding acquisition, A.R. and M.B.. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was supported by the COST DIVAS Action (FA1407, Application of next generation sequencing for the study and diagnosis of plant viral diseases in agriculture).

**Data Availability Statement:** The nucleotide sequences from this study can be found in NCBI GenBank (<https://www.ncbi.nlm.nih.gov/genbank/>) under accessions PX705376-PX705383 and the SRA data in NCBI SRA (<https://www.ncbi.nlm.nih.gov/sra>) under BioSamples SAMN53882179-SAMN53882184 and BioProject PRJNA1379522.

**Acknowledgments:** The sampling and analyses at NIVIP were conducted within the framework of two Euphresco projects: 2023-A-446, *Surveillance for viruses and other non-culturable pathogens of forestry tree species (VIRNOTree)* (Sample 6165949) and 2019-E-312 *Using High Throughput Sequencing to gain insights from virus collections and strengthening the infrastructure of Plant Virus Collections (VirusCurate)* (Sample WAG0452949 and WAG0453175). The authors thank Anne Sophie van Bruggen from the NPPO-NL for her input on Beech Leaf Disease and Dirk Jan van der Gaag and Martijn Schenk for proofreading and contributions to the manuscript.

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

## References

1. Schmelzer, K.; Schmidt, H.; Schmidt, H. Virus diseases and phenomena suggesting a virus in forest trees. **1966**.
2. Nienhaus, F.; Castello, J. VIRUSES IN FOREST TREES. *Annual Review of Phytopathology: 1989* **1989**, *27*, 165-186.
3. Winter, S.; Nienhaus, F. Identification of viruses from European beech (*Fagus sylvatica* L.) of declining forests in Northrhine-Westfalia (FRG). **1989**.
4. Nienhaus, F.; Ebrahim-Nesbat, F.; Fricke, M.; Büttner, C.; Welter, K. Investigations on viruses from declining beech trees (*Fagus sylvatica* L.) in Rhineland and Westfalia, Federal Republic of Germany. *European journal of forest pathology* **1985**, *15*, 402-411.
5. Werner, R.; Mühlbach, H.P.; Büttner, C. Detection of cherry leaf roll nepovirus (CLRV) in birch, beech and petunia by immuno-capture RT-PCR using a conserved primer pair. *European Journal of Forest Pathology* **1997**, *27*, 309-318.

6. Jones, A.; Koenig, R.; Lesemann, D.E.; Hamacher, J.; Nienhaus, F.; Winter, S. Serological comparison of isolates of Cherry leaf roll virus from diseased beech and birch trees in a forest decline area in Germany with other isolates of the virus. *Journal of Phytopathology* **1990**, *129*, 339-344.
7. Hamacher, J.; Quadt, A. Isolation of cherry leaf roll and brome mosaic viruses from European beech and transmission to beech seedlings. **1994**.
8. Büttner, C.; Bargaen, S.v.; Bandte, M.; Mühlbach, H.-P. Forest diseases caused by viruses. In *Infectious forest diseases*; CABI Wallingford UK: 2013; pp. 50-75.
9. Büttner, C.; Landgraf, M.; Colino, H.L.F.; von Bargaen, S.; Bandte, M. Virus diseases of forest and urban trees. In *Forest microbiology*; Elsevier: 2023; pp. 61-97.
10. Hammond, J.; Adams, I.P.; Fowkes, A.R.; McGreig, S.; Botermans, M.; van Oorspronk, J.J.; Westenberg, M.; Verbeek, M.; Dulleman, A.M.; Stijger, C.C. Sequence analysis of 43-year old samples of *Plantago lanceolata* show that Plantain virus X is synonymous with Actinidia virus X and is widely distributed. *Plant Pathology* **2021**, *70*, 249-258.
11. Adams, I.P.; Boonham, N.; Jones, R.A. Full-genome sequencing of a virus from a 33-year-old sample demonstrates that Arracacha mottle virus is synonymous with Arracacha virus Y. *Microbiology Resource Announcements* **2018**, *7*, 10.1128/mra.01393-01318.
12. Rumbou, A.; Vainio, E.J.; Büttner, C. Towards the forest virome: high-throughput sequencing drastically expands our understanding on virosphere in temperate forest ecosystems. *Microorganisms* **2021**, *9*, 1730.
13. Jones, R.A.; Boonham, N.; Adams, I.P.; Fox, A. Historical virus isolate collections: An invaluable resource connecting plant virology's pre-sequencing and post-sequencing eras. *Plant Pathology* **2021**, *70*, 235-248.
14. MacDiarmid, R.; Rodoni, B.; Melcher, U.; Ochoa-Corona, F.; Roossinck, M. Biosecurity implications of new technology and discovery in plant virus research. *PLoS pathogens* **2013**, *9*, e1003337.
15. Fox, A.; Botermans, M.; Ziebell, H.; Fowkes, A.R.; Fontdevila Pareta, N.; Massart, S.; Rodoni, B.; Chooi, K.M.; Kreuze, J.; Kumar, P. Implications of high throughput sequencing of plant viruses in biosecurity—a decade of progress? *Peer Community Journal* **2025**, *5*.
16. Schönegger, D.; Marais, A.; Babalola, B.M.; Faure, C.; Lefebvre, M.; Svanella-Dumas, L.; Brázdová, S.; Candresse, T. Carrot populations in France and Spain host a complex virome rich in previously uncharacterized viruses. *PloS one* **2023**, *18*, e0290108.
17. Fox, A.; Gibbs, A.J.; Fowkes, A.R.; Pufal, H.; McGreig, S.; Jones, R.A.; Boonham, N.; Adams, I.P. Enhanced apiaceous potyvirus phylogeny, novel viruses, and new country and host records from sequencing Apiaceae samples. *Plants* **2022**, *11*, 1951.
18. Schnabel, E.; Xavier, C.A.D.; Whitfield, A.E.; Dubrow, Z.; Pham, G.; Cieniewicz, E. Exploring the virome of blackberry and wild *Rubus* spp. in South Carolina. *Phytobiomes Journal* **2025**, *9*, 80-94.
19. Maclot, F.; Debue, V.; Malmstrom, C.M.; Filloux, D.; Roumagnac, P.; Eck, M.; Tamisier, L.; Blouin, A.G.; Candresse, T.; Massart, S. Long-term anthropogenic management and associated loss of plant diversity deeply impact virome richness and composition of Poaceae communities. *Microbiology Spectrum* **2023**, *11*, e04850-04822.
20. Rumbou, A.; Candresse, T.; Marais, A.; Theil, S.; Langer, J.; Jalkanen, R.; Büttner, C. A novel badnavirus discovered from *Betula* sp. affected by birch leaf-roll disease. *PLoS One* **2018**, *13*, e0193888.
21. EPPO. Addendum – New supporting information for PM 7/151 Considerations for the use of high throughput sequencing in plant health diagnostics. *EPPO Bulletin* **2024**, *54*, 253.
22. Buchfink, B.; Xie, C.; Huson, D.H. Fast and sensitive protein alignment using DIAMOND. *Nature methods* **2015**, *12*, 59-60.
23. Ondov, B.D.; Bergman, N.H.; Phillippy, A.M. Interactive metagenomic visualization in a Web browser. *BMC bioinformatics* **2011**, *12*, 385.
24. Katoh, K.; Standley, D.M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular biology and evolution* **2013**, *30*, 772-780.
25. Minh, B.Q.; Schmidt, H.A.; Chernomor, O.; Schrempf, D.; Woodhams, M.D.; Von Haeseler, A.; Lanfear, R. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Molecular biology and evolution* **2020**, *37*, 1530-1534.

26. Kalyaanamoorthy, S.; Minh, B.Q.; Wong, T.K.; Von Haeseler, A.; Jermini, L.S. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature methods* **2017**, *14*, 587-589.
27. Minh, B.Q.; Nguyen, M.A.T.; Von Haeseler, A. Ultrafast approximation for phylogenetic bootstrap. *Molecular biology and evolution* **2013**, *30*, 1188-1195.
28. Boom, R.; Sol, C.; Salimans, M.; Jansen, C.; Wertheim-van Dillen, P.; Van der Noordaa, J. Rapid and simple method for purification of nucleic acids. *Journal of clinical microbiology* **1990**, *28*, 495-503.
29. King, A.M.; Lefkowitz, E.; Adams, M.J.; Carstens, E.B. *Virus taxonomy: ninth report of the International Committee on Taxonomy of Viruses*; Elsevier: 2011; Volume 9.
30. Tian, Z.; Hu, T.; Holmes, E.C.; Ji, J.; Shi, W. Analysis of the genetic diversity in RNA-directed RNA polymerase sequences: implications for an automated RNA virus classification system. *Virus Evolution* **2024**, *10*, veae059.
31. Silva, J.M.F.; Melo, F.L.; Elena, S.F.; Candresse, T.; Sabanadzovic, S.; Tzanetakis, I.E.; Blouin, A.G.; Villamor, D.E.V.; Mollov, D.; Constable, F. Virus classification based on in-depth sequence analyses and development of demarcation criteria using the Betaflexiviridae as a case study. *Journal of General Virology* **2022**, *103*, 001806.
32. Vainio, E.J.; Rumbou, A.; Diez, J.J.; Büttner, C. Forest Tree Virome as a Source of Tree Diseases and Biological Control Agents. *Current Forestry Reports* **2024**, *10*, 153-174. <https://doi.org/10.1007/s40725-024-00214-8>.
33. Rumbou, A.; Candresse, T.; Marais, A.; Svanella-Dumas, L.; Landgraf, M.; von Bargen, S.; Büttner, C. Unravelling the virome in birch: RNA-Seq reveals a complex of known and novel viruses. *PLoS One* **2020**, *15*, e0221834.
34. Carta, L.K.; Handoo, Z.A.; Li, S.; Kantor, M.; Bauchan, G.; McCann, D.; Gabriel, C.K.; Yu, Q.; Reed, S.; Koch, J. Beech leaf disease symptoms caused by newly recognized nematode subspecies *Litylenchus crenatae mccannii* (Anguinata) described from *Fagus grandifolia* in North America. *Forest Pathology* **2020**, *50*, e12580.
35. EPPO. *Litylenchus crenatae mccannii*. Available online: <https://gd.eppo.int/taxon/LITYMC/> (accessed on
36. Peters, D.; Matsumura, E.E.; van Vredendaal, P.; van der Vlugt, R.A. The plant virus transmissions database. *Journal of General Virology* **2024**, *105*, 001957.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.