

**UNIVERSIDAD POLITÉCNICA DE MADRID**

ESCUELA TÉCNICA SUPERIOR DE INGENIERÍA AGRONÓMICA, ALIMENTARIA Y DE  
BIOSISTEMAS



**Evolution and ecology of virus host range and  
host utilisation in a heterogeneous  
ecosystem**

**DOCTORAL THESIS**

Submitted for the degree of Doctor by:

**Bisola Mercy Babalola**

M.Sc. in Plant health  
Bachelor of Agriculture

Madrid, 2024



UNIVERSIDAD POLITÉCNICA DE MADRID

Escuela Técnica Superior de Ingeniería Agronómica,  
Alimentaria y de Biosistemas

**Doctoral Degree in Biotechnology and Genetic Resources of Plants  
and Associated Microorganisms**

# **Evolution and ecology of virus host range and host utilisation in a heterogeneous ecosystem**

## **DOCTORAL THESIS**

Submitted for the degree of Doctor by:

**Bisola Mercy Babalola**

M.Sc. in Plant health  
Bachelor of Agriculture

Under the supervision of:

Dr. Fernando García-Arenal Rodríguez

Dr. Michael John McLeish

Madrid, 2024

Title: Evolution and ecology of virus host range and host utilisation in a heterogeneous ecosystem

Author: Bisola Mercy Babalola

Doctoral Programme: Biotechnology and Genetic Resources of Plants and Associated Microorganisms

Thesis Supervision:

Dr. Fernando García-Arenal Rodríguez, (Dr. Ingeniero Agrónomo, Universidad Politécnica de Madrid)

Dr. Michael John McLeish, (Dr. in Molecular Ecology and Evolution, Universidad Politécnica de Madrid)

External Reviewers:

Thesis Defense Committee:

Thesis Defense Date:

“This thesis has been funded by the European Union’s Horizon 2020 Research and Innovation Programme under Grant Agreement No. 813542. INEXTVIR -Marie Skłodowska-Curie Innovative Training Network (MSCA-ITN)”.

## ACKNOWLEDGEMENT.

*“Not that we are sufficiently qualified in ourselves to claim anything as coming from us, but our sufficiency and qualifications come from God” - 2 Corinthians 3:5 - AMP.*

The above quote is to first and foremost acknowledge my creator, to whom I am grateful for life and sustenance, without which this thesis would not exist.

I would like to express my sincere appreciation to the European Union for its Horizon 2020 Research and Innovation program, under which I was selected for the Marie Skłodowska-Curie doctoral fellowship within the INEXTVIR (Innovative Network for Next Generation Training and Sequencing of Virome) project, grant agreement no. 813542. I am immensely grateful for the opportunity to be an early-stage researcher (ESR) at my host institution, the Centre for Plant Biotechnology and Genomics (CBGP), Madrid, constituted by Universidad Politécnica de Madrid (UPM) and Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA), which afforded me the privilege to conduct my doctoral research. I would like to extend my appreciation to Prof. Maja Ravnikar and Dr Ion Gutierrez of National Institute of Biology, Slovenia for their excellent leadership and coordination of the INEXTVIR project and other members of the project consortium.

My appreciation goes to my directors Prof. Fernando Garcia-Arenal Rodriguez and Dr Michael John McLeish for their continuous guidance, professional inputs and support during the entire course of my doctoral program. I also appreciate Prof. Aurora Fraile for assisting during field sampling. Special thanks to Dr. Thierry Candresse and Dr Armelle Marais who hosted me at INRAE, Bordeaux during my three-month secondment. Indeed, I have been privileged to stand on the shoulders of these giants whose groundbreaking work and relentless passion have paved the way, inspiring me to push the boundaries of knowledge and contribute to the ever-growing tapestry of our field.

I am thankful for all the other members of our research group who were part of my research journey: Miguel Ángel Mora and Antolín López Quirós for their technical assistance during the

course of my laboratory work. Also, Dr. Cristian Carrasco, Dr. Julio L. Rodríguez, Dr. María Ángeles Ayllón, Dr. Adrián Peláez, Dr. Ana Ruiz, Dr. Zamfir Adrian, Rafael de Andrés, Miriam Gil-Valle (Nena) who helped in various ways during my entire stay in Lab227. I would like to appreciate all the INEXTVIR ESRs for the interactive ESR meetings and teamwork that allowed peer to peer learning and motivation. Specifically my thanks goes to Dr. Maryam Khalili, Dr Ayoub Maachi, Dr Deborah Schönegger, Dr Mark Paul Selda Rivarez, Grigorii Sukhorukov, Dr Johan Rollin, whom I was opportuned to work closely with and also exchange good ideas.

I truly appreciate my close friends: Christy Ugba, Dr Tolulope Ayeyemi and Ebunoluwa Ajobiewe who despite their own struggles, were always available to chat and encourage me to keep up the good work.

I am deeply thankful for my dear parents, Arc. and Mrs J.S Babalola, who backed me up in various ways beyond explanation. When the going was tough, they kept encouraging and praying for me. To all my siblings too, I grateful for your constant prayer support and understanding throughout these years.

A special acknowledgement goes to my husband: Dr Nasamu Musa, who initiated the PhD idea and has been my backbone from the very beginning to end. To our beloved son, Emmo-Vosine Musa, who joined the journey as a baby and has grown to see the end of it. You both have been my greatest joy and motivation.

Finally, I appreciate myself for staying through the process and hoping against hope even when the end was uncertain.

## ABSTRACT.

Understanding the evolution and ecology of virus host range and utilisation is pivotal for addressing viral emergence, especially in ecosystems where diverse plant communities, such as wild and cultivated species, coexist. Viral emergence is influenced by intrinsic genetic factors and ecological conditions that determine a virus's ability to infect new hosts. High-throughput sequencing (HTS) offers an advanced method for studying entire viral communities, allowing detailed insights into viral diversity, host interactions, and co-infections across different environments. Yet, the role of ecological factors in shaping virus host range and emergence remains underexplored. By comparing the viral communities of different subspecies of carrot and its wild relative and assessing the host range of eight viruses with different transmission modes in four different habitats, this thesis aims to investigate the factors influencing the host range and utilisation of plant viruses. This was achieved by analysing incidence, infection patterns and genetic diversity of virus populations in relation to host plants and habitat characteristics.

HTS identified 45 virus OTUs across wild and cultivated carrot populations. Wild carrots exhibited greater virome richness, harbouring a higher proportion of novel viruses, while cultivated carrots showed a higher incidence of known aphid-transmitted viruses. This indicates that wild carrots act as reservoirs of viral diversity with potential for novel virus emergence. The multi-host study found that ecological factors, including host abundance, richness and habitat heterogeneity, play a crucial role in shaping virus host range and use. The realized host range of viruses varied significantly between wild and anthropic habitats, with generalist viruses -CMV, PZSV, WMV utilising hosts in both habitats and specialist viruses -PPV, BChV being more restricted to hosts in anthropic habitats. Non-random interactions between viruses were observed, including both facilitative (positive associations) and competitive (negative associations) interactions. Co-infection was common among viruses sharing similar transmission modes, such as those transmitted by aphids, which suggests that transmission strategies and vector behaviours significantly influence virus community structures. The genetic diversity of CMV populations was influenced by habitat type rather than host specificity, indicating that ecological dynamics, such as habitat disturbance and connectivity, affect virus genetic structure. Within Edge habitat, CMV,

PZSV and TAV showed minimal genetic differentiation across different hosts, highlighting their low levels of host-specific genetic divergence. This study demonstrates that the interaction between ecological and genetic factors is crucial in determining the host range evolution and emergence potential of plant viruses. Wild plant populations, such as wild carrots, play a key role as reservoirs of viral diversity, potentially introducing new viruses into cultivated systems. The findings underscore the importance of habitat heterogeneity and host abundance and richness in shaping viral communities, emphasising the need for targeted management strategies that consider ecological dynamics to mitigate virus emergence in agroecosystems. This research provides a framework for understanding virus evolution in complex ecological contexts, offering insights into managing plant virus outbreaks.

## RESUMEN.

Comprender la evolución y la ecología de la gama de hospedadores de virus y su utilización es fundamental para abordar la emergencia viral, especialmente en ecosistemas donde coexisten diversas comunidades vegetales, como especies silvestres y cultivadas. La aparición de virus depende de factores genéticos intrínsecos y de condiciones ecológicas que determinan la capacidad de un virus para infectar nuevos huéspedes. La secuenciación de alto rendimiento (HTS) ofrece un método avanzado para estudiar comunidades virales enteras, lo que permite conocer en detalle la diversidad viral, las interacciones con los hospedadores y las coinfecciones en distintos entornos. Sin embargo, el papel de los factores ecológicos en la configuración de la gama de hospedadores y la aparición de virus sigue sin explorarse lo suficiente. Mediante la comparación de las comunidades virales de diferentes subespecies de zanahoria y su pariente silvestre y la evaluación del rango de hospedadores de ocho virus con diferentes modos de transmisión en cuatro hábitats diferentes, esta tesis pretende investigar los factores que influyen en el rango de hospedadores y la utilización de los virus de las plantas. Para ello, se analizaron la incidencia, los patrones de infección y la diversidad genética de las poblaciones de virus en relación con las plantas hospedadoras y las características del hábitat.

El HTS identificó 45 OTU de virus en poblaciones de zanahorias silvestres y cultivadas. Las zanahorias silvestres mostraban una mayor riqueza vírica, albergando una mayor proporción de nuevos virus, mientras que las zanahorias cultivadas mostraban una mayor incidencia de virus conocidos transmitidos por áfidos. Esto indica que las zanahorias silvestres actúan como reservorios de diversidad vírica con potencial para la aparición de nuevos virus. El estudio multihospedador reveló que los factores ecológicos, entre ellos la abundancia y riqueza de hospedadores y la heterogeneidad del hábitat, desempeñan un papel crucial en la configuración de la gama de hospedadores y el uso de virus. Los virus generalistas (CMV, PZSV, WMV) utilizan hospedadores de ambos hábitats, mientras que los virus especializados (PPV, BChV) se limitan más a los hospedadores de hábitats antropogénicos. Se observaron interacciones no aleatorias entre los virus, tanto facilitadoras (asociaciones positivas) como competitivas (asociaciones negativas). La coinfección fue común entre virus que comparten modos de transmisión similares,

como los transmitidos por pulgones, lo que sugiere que las estrategias de transmisión y los comportamientos de los vectores influyen significativamente en las estructuras de las comunidades de virus. La diversidad genética de las poblaciones de CMV se vio influida por el tipo de hábitat más que por la especificidad del hospedador, lo que indica que la dinámica ecológica, como la perturbación del hábitat y la conectividad, afecta a la estructura genética del virus. Dentro del hábitat Edge, el CMV, el PZSV y el TAV mostraron una diferenciación genética mínima entre distintos hospedadores, lo que pone de relieve sus bajos niveles de divergencia genética específica de hospedador. Este estudio demuestra que la interacción entre factores ecológicos y genéticos es crucial para determinar la evolución del rango de hospedadores y el potencial de emergencia de los virus vegetales. Las poblaciones de plantas silvestres, como las zanahorias silvestres, desempeñan un papel clave como reservorios de diversidad viral, introduciendo potencialmente nuevos virus en los sistemas cultivados. Los resultados subrayan la importancia de la heterogeneidad del hábitat y la abundancia y riqueza de hospedadores en la formación de comunidades virales, lo que pone de relieve la necesidad de estrategias de gestión específicas que tengan en cuenta la dinámica ecológica para mitigar la aparición de virus en los agroecosistemas. Esta investigación proporciona un marco para entender la evolución de los virus en contextos ecológicos complejos y ofrece ideas para gestionar los brotes de virus en las plantas.

## Table of Contents.

|                                                                               |           |
|-------------------------------------------------------------------------------|-----------|
| Abstract .....                                                                | 2         |
| Resumen.....                                                                  | 6         |
| <b>1. INTRODUCTION .....</b>                                                  | <b>13</b> |
| 1.1 Plant virus emergence.....                                                | 13        |
| 1.1.1 Ecological factors driving virus emergence. ....                        | 16        |
| 1.1.2. Host range evolution and virus emergence. ....                         | 18        |
| 1.1.3. Co-infection.....                                                      | 20        |
| 1.2 Genetic determinants of host range evolution. ....                        | 21        |
| 1.2.1 Across host trade-offs. ....                                            | 23        |
| 1.3 Virus communities and heterogeneous ecosystems.....                       | 24        |
| 1.3.1 Plant-virus interactions in wild communities.....                       | 26        |
| 1.4 Habitat heterogeneity of an ecosystem in Spain. ....                      | 27        |
| 1.4.1 Virus detection and surveillance: development of HTS technologies. .... | 30        |
| 1.4.2 Application of metagenomics to virus ecology and evolution. ....        | 31        |
| 1.5 The virome of a crop and its wild relative.....                           | 32        |
| 1.5.1 Description of the system: hosts strategies and host use.....           | 32        |
| 1.5 A multi-host – multi-virus study system.....                              | 34        |
| 1.5.1 Description of virus genera analysed.....                               | 36        |
| <b>2. OBJECTIVES.....</b>                                                     | <b>39</b> |
| <b>3. MATERIALS AND METHODS .....</b>                                         | <b>40</b> |
| 3.1 Virome study in single host: carrot. ....                                 | 40        |
| 3.1.1 Carrot sampling location and study area. ....                           | 40        |
| 3.1.1 Double-stranded RNA purification and Illumina sequencing.....           | 43        |
| 3.1.2 Metagenomic analysis of Carrot samples .....                            | 44        |
| 3.2 Virome study in multiple hosts. ....                                      | 45        |
| 3.2.1 Study area and sampling.....                                            | 45        |
| 3.2.2 Sample processing and extraction methods. ....                          | 48        |
| 3.2.3 Library preparation and High throughput sequencing.....                 | 48        |
| 3.2.4 Metagenomic detection of virus OTUs .....                               | 49        |
| 3.2.5 RT-PCR detection of viruses. ....                                       | 50        |
| 3.3 Data analyses .....                                                       | 53        |

|                                                                                                                                              |            |
|----------------------------------------------------------------------------------------------------------------------------------------------|------------|
| 3.3.1 Host range analyses.....                                                                                                               | 53         |
| 3.3.2 Co-occurrence analyses .....                                                                                                           | 53         |
| 3.3.3 Incidence .....                                                                                                                        | 54         |
| 3.3.4 Statistical analyses. ....                                                                                                             | 55         |
| 3.3.5 Genetic diversity and Phylogenetic analyses. ....                                                                                      | 57         |
| <b>4. RESULTS .....</b>                                                                                                                      | <b>60</b>  |
| 4.1 The virome of a crop and its wild relative.....                                                                                          | 60         |
| 4.1.1 Comparison of the virome of wild and cultivated carrot populations. ....                                                               | 60         |
| 4.1.2 Virome structure in cultivated and wild carrots.....                                                                                   | 66         |
| 4.1.3 Identification of two novel putative satellite RNAs with hammerhead structures in the virome of French and Spanish carrot samples..... | 69         |
| 4.2 A multi-host-multi-virus study. ....                                                                                                     | 77         |
| 4.2.1 Viral OTU detection by HTS and RT-PCR.....                                                                                             | 77         |
| 4.2.2 Host range by HTS2C detection. ....                                                                                                    | 85         |
| 4.2.3 Comparison of the host range of the different viruses. ....                                                                            | 97         |
| 4.3 Virus co-occurrence.....                                                                                                                 | 101        |
| 4.4 Patterns of virus incidence. ....                                                                                                        | 105        |
| 4.4.1 Ecological and genetic determinants of incidence.....                                                                                  | 114        |
| 4.5 Genetic diversity Analyses .....                                                                                                         | 129        |
| 4.5.1 Phylogenetic analyses.....                                                                                                             | 136        |
| <b>5. DISCUSSION .....</b>                                                                                                                   | <b>144</b> |
| 5.1 The virome of a crop and its wild relative.....                                                                                          | 146        |
| 5.2 A multi-host-multi-virus study. ....                                                                                                     | 148        |
| 5.2.1 Virus-Host detection by HTS and RT-PCR.....                                                                                            | 148        |
| 5.2.2 Host range evolution in heterogenous ecosystems. ....                                                                                  | 150        |
| 5.2.3 Co-occurrence and co-infection patterns.....                                                                                           | 153        |
| 5.2.4 Ecological and genetic determinants of incidence.....                                                                                  | 155        |
| 5.2.5 Population genetic diversity and structure. ....                                                                                       | 157        |
| <b>6. CONCLUSIONS .....</b>                                                                                                                  | <b>160</b> |
| <b>7. REFERENCES .....</b>                                                                                                                   | <b>162</b> |
| <b>ANNEXES .....</b>                                                                                                                         | <b>198</b> |

## List of acronyms.

|          |                                                 |
|----------|-------------------------------------------------|
| AI       | Association Index                               |
| AIC      | Akaike's Information Criterion                  |
| bam      | binary alignment map                            |
| BaTS     | Bayesian Tip-association Significance tests     |
| bp       | Base pair                                       |
| BLAST    | Basic Local Assignment Search Tool              |
| cDNA     | Complementary DNA                               |
| CI       | Cylindrical inclusion Protein                   |
| CP       | Coat Protein                                    |
| CTAB     | Cetyltrimethylammonium bromide                  |
| DNA      | Deoxyribonucleic acid                           |
| dsRNA    | double-stranded RNA                             |
| dsDNA    | double- and single-stranded                     |
| $D_{AE}$ | Asymptotic estimator of Shannon diversity       |
| EH       | Environmental heterogeneity                     |
| $\pi$    | nucleotide diversity                            |
| GFG      | gene-for-gene                                   |
| HhSats   | Hammerhead satellites                           |
| HV       | Helper viruses                                  |
| HTS      | high throughput sequencing                      |
| mRNAs    | messenger RNA                                   |
| NCBI     | National center for Biotechnology information   |
| PVP      | Polyvinyl pyrrolidone                           |
| RT-PCR   | reverse transcription polymerase chain reaction |
| RNA      | Ribonucleic acid                                |
| sgRNA    | subgenomic RNA                                  |

|         |                                  |
|---------|----------------------------------|
| ssRNA   | single-stranded RNA              |
| ssDNA   | single-stranded DNA              |
| SatRNA  | Satellite RNA                    |
| tlaRNAs | tombusvirus-like associated RNAs |
| WGCS    | whole genome consensus sequences |

## List of viral species.

|           |                                            |
|-----------|--------------------------------------------|
| BChV      | <i>Beet chlorosis virus</i>                |
| BYDV      | <i>Barley yellow dwarf virus</i>           |
| CABYV     | <i>Cucurbit aphid-borne yellows virus</i>  |
| CMV       | <i>Cucumber mosaic virus</i>               |
| CMD       | <i>Carrot motley dwarf disease complex</i> |
| CMoV      | <i>Carrot mottle virus</i>                 |
| CtRLV     | <i>Carrot red leaf virus</i>               |
| CtRLVaRNA | <i>Carrot red leaf associated RNA</i>      |
| CYDV      | <i>Cereal yellow dwarf virus</i>           |
| PepMV     | <i>Pepino mosaic virus</i>                 |
| PMMoV     | <i>Pepper mild mottle virus</i>            |
| PepGMV    | <i>Pepper golden mosaic virus</i>          |
| PHYVV     | <i>Pepper huasteco yellow vein virus</i>   |
| PZSV      | <i>Pelargonium zonate spot virus</i>       |
| PPV       | <i>Plum pox virus</i>                      |
| PVY       | <i>Potato virus Y</i>                      |
| PYFV      | <i>Parsnip yellow fleck virus</i>          |
| TAV       | <i>Tomato aspermy virus</i>                |
| TuMV      | <i>Turnip mosaic Virus</i>                 |
| TYCNDV    | <i>Tomato leaf curl New Delhi virus</i>    |
| WMV       | <i>Watermelon mosaic virus</i>             |
| WSMV      | <i>Wheat streak mosaic virus</i>           |
| YoMV      | <i>Youcai mosaic virus</i>                 |

## 1. INTRODUCTION.

### 1.1 Plant virus emergence.

Plant viruses pose a significant threat to global food security and agriculture sustainability, leading to substantial annual crop yield reductions of up to 7 % of the potential yield (Oerke and Dehne, 2004) and billions of dollars in economic losses every year (Nicaise, 2014; Sastry and Zitter, 2014). More so, crop losses translate to economic losses for farmers, impacting their livelihoods and economic flows within society (Jones and Naidu, 2019, He and Krainer, 2020). Farmers may resort to increased pesticide use to control virus vectors, leading to environmental and health concerns. Countries may impose import restrictions on agricultural products from regions experiencing virus outbreaks (Elena et al. 2014, Jones, 2021). These losses involve endemic virus diseases of major crops, but the impact of viral diseases on plant production is often much higher, up to 80-100%, for emerging viral diseases (Jones et al. 2008). Therefore, understanding disease emergence and the extrinsic or intrinsic factors involved is crucial for developing effective control strategies and mitigating the negative consequences of plant viral diseases mentioned above.

Emergent diseases are defined as infectious disease whose incidence is increasing following its first introduction into a new host population or whose incidence is increasing in an existing host population as a result of long-term changes in its underlying epidemiology (Woolhouse and Dye 2001). Virus emergence is a multifaceted process involving the appearance of new viruses or increased infectivity and/or virulence of existing ones in human, animal, and plant populations (Domingo, 2010). Emergent viruses often are multi-host pathogens that spillover into a new host population (target host) from the original host population in which the virus is permanently maintained, a reservoir host as defined by Haydon et al. (2002), often a wild host if emergence occurs in a crop (Roossinck and García-Arenal, 2015). For emergence to occur, the virus must move from its reservoir host to infect the target host, and later evolve to establish productive infections and effective transmission mechanisms in the new host (Woolhouse et al. 2005; Elena et al. 2014).

The process of emergence is divided in three phases: in the first phase, the virus from a reservoir host comes in contact with another host species present in the same community, or with a new host species (McLeish et al. 2021a; Moury et al. 2017). Such initial contact that results in an infection in the new host is favoured by changes in the distribution and abundance of natal or new host, vectors, and/or virus (Elena et al. 2014). In the second phase, continuous spillover of the virus from its reservoir to its new host (wild or crop) results in viral adaptation to the new host or environment so that infections are efficient enough to allow transmission between individuals of the new host (Woolhouse et al. 2005). For the third phase, the virus needs to optimize the between-host transmission in the new host population, which often requires adaptation to new vector species or new modes of transmission (Elena et al. 2014). While the first two phases may involve genetic changes in the virus (e.g., mutation/recombination, selection, adaptive trade-offs), the third step may also require changes on the part of the vector or host populations, such as through increased contact rates, host distribution changes, and other ecological or environmental shifts (Elena et al. 2011; Dennehy, 2017). The most understood aspect of plant virus emergence relates to the evolutionary mechanisms that constrain or favor host adaptation (Elena et al. 2014), and knowledge of ecological factors that favour new encounters and host shifts is still limited and derives from rather few experimental systems. Thus, the ecological factors involved in plant virus emergence will be further elaborated to identify possible management strategies for consideration.

A typical example of an emergent plant virus is Pepino mosaic virus (PepMV, genus *Potexvirus*), an important pathogen of tomato crops worldwide (Gómez et al. 2012; Hanssen and Thomma, 2010). PepMV was first isolated in Peru in 1974 from pepino (*Solanum muricatum*) plants showing yellow mosaic symptoms (Jones et al. 1980). It was not reported as a pathogen of tomato until 1999, when detected in greenhouses in the Netherlands (van del Vlugt et al. 2000). Soon after, its epidemic expansion in tomato crops was reported first in Europe and then in North America (French et al. 2001; Cotillon et al. 2002; Pagán et al. 2006), thus it became a major pathogen of tomato worldwide. PepMV is transmitted both by plant-to-plant contact and through the seeds, which may have contributed to its long-distance spread (Córdoba-Sellés et al. 2007; Hanssen et al. 2010). Emergence involved adaptation to the new host, the domesticated tomato, as was

shown by comparing isolates from wild tomato populations in Peru and tomato crops in Europe (Moreno-Pérez et al 2014). Although emerging viruses often have wild reservoirs, emergence can also occur from another crop, or from crops to native vegetation (Jones, 2009). The term 're-emerging' is often used to describe the resurgence of previously known viruses that had caused significant disease problems but had declined in importance before experiencing a sudden increase in incidence within crops (Jones, 2009).

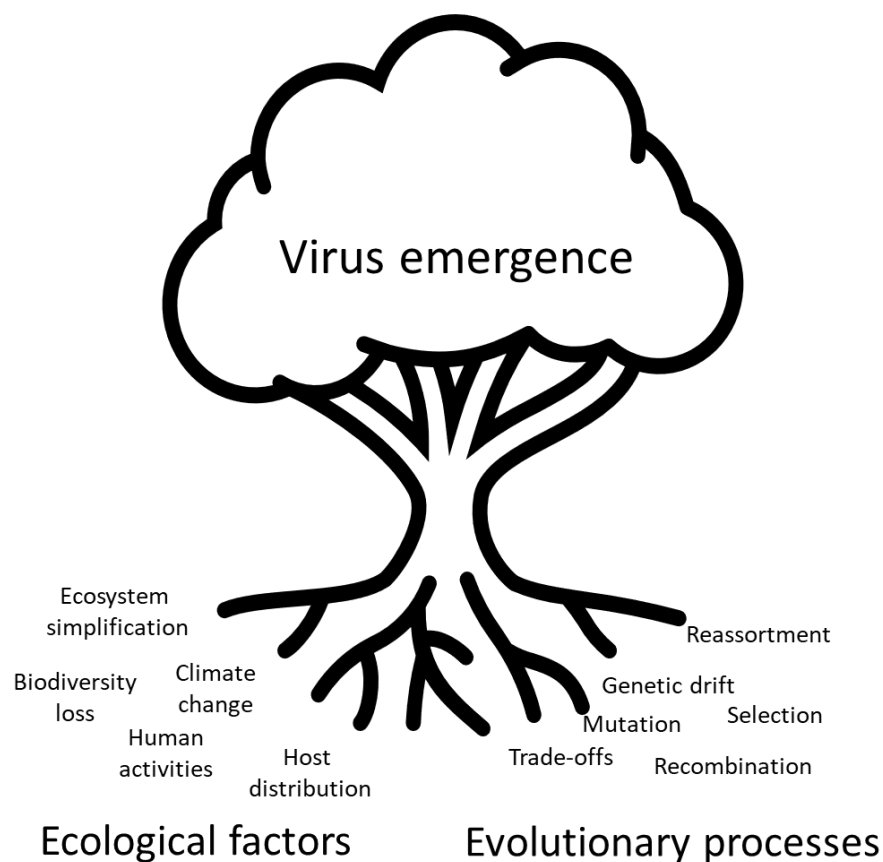


**Figure 1.1: Conceptual representation of ecological and evolutionary factors at the root of virus emergence.** Interactions among either or both ecological and evolutionary factors are present during the three phases of emergence.

### 1.1.1 Ecological factors driving virus emergence.

In the first phase of emergence, ecological factors must favor contact between the reservoir and the target host populations. The introduction of viruses to new hosts or environments is facilitated through various human activities such as international trade, agriculture and the associated loss of wild ecosystems (Anderson et al. 2004; Jones, 2009; McDonald and Stukenbrock, 2016). International trade of seeds and crops facilitates the global movement of infected plants and vectors and has been identified as a major factor contributing to plant virus long-distance dispersal (Jones 2018). Agricultural practices associated with deforestation, crop introduction and intensification create opportunities for virus spread and novel host encounters (Jones, 2009; McDonald and Stukenbrock, 2016). For example, expansion of agricultural land into natural habitats disrupts established virus-host dynamics (Jones 2009). Monocultures create vast swathes of susceptible hosts, amplifying virus spread, while the use of insecticides can disturb natural predator-prey dynamics, leading to an increase in virus vectors (Seal et al. 2006a,b).

Ecological changes such as biodiversity loss and habitat fragmentation are among factors that bring previously isolated viruses and susceptible hosts into closer contact (phase I of emergence), which facilitates host range expansion (phase II). Though empirical and experimental evidence is scarce, virus ecologists agree that ecosystem simplification associated with biodiversity loss will favour both new virus–host encounters, and new epidemiological dynamics that may facilitate emergence (Elena et al. 2014; Roossinck and García-Arenal 2015). In a study of the relationship between biodiversity and disease risk, wild pepper or chiltepin populations growing under three different levels of human management (cultivated, let-standing and wild populations) were monitored for virus infection (Pagán et al. 2012; Fraile et al. 2017). The prevalence of two begomoviruses, Pepper golden mosaic virus (PepGMV) and Pepper huasteco yellow vein virus (PHYVV), was found to be higher in cultivated than in let-standing and wild populations, which correlated with a lower diversity of the plant communities where the cultivated peppers grew. That is, loss of biodiversity was associated with an increase of infection and disease risks. Thus, expansion in human activity resulting in biodiversity loss can lead to changes in virus distributions

and prevalence that provide conditions that favour emergence of viruses and development of damaging epidemics at the interface between managed and natural vegetation.

Climatic changes have also been identified as a significant ecological factor influencing the emergence and spread of vector-borne diseases, (Jones, 2014; Alcaide, et al. 2020; Amari et al. 2021). The alteration of meteorological conditions such as temperature or precipitation affects the population, range and distribution of vectors and hosts, and seasonality of vectors and hosts, consequently impacting virus survival and transmission dynamics of vector-borne diseases (Canto, et al. 2009, Trebicki, 2020). In addition, climate change may affect the geographical distribution of host species and facilitate changes in host range by making possible new interactions between virus, plant species and vectors, as species distributions are among ecological factors that have been highlighted as central to virus emergence (McLeish et al. 2019). Differences in biotic and abiotic environmental conditions and that of hosts, and the availability of resources in each location, determine viral community composition (McLeish et al. 2019). Community composition refers to the identity of all taxa that a community consists of, and their relative abundances (Hanson et al. 2012). The interactions among species in a community, or in a given environment, determine what organisms are present in it, or the community structure (Lavorel and Garnier 2002; Van der Heijden et al. 1998). Species diversity (i.e., richness and evenness), biotic and abiotic interactions are some of the methods commonly used to measure community structure. Interactions among species and with the abiotic environment are largely the result of relationships among traits, with stochastic interactions also being a factor. Morphological, physiological, and behavioral traits that suit a species to a particular habitat, or resource (i.e., ecological specialisation), and facilitate or limit the ability to use environments, are an important cause of community structure (Via and Hawthorne, 2002). For example, a mycovirus conferred heat tolerance in a fungal endophyte associated with a grass and allowed the fungus and grass to grow at high soil temperatures (Márquez et al. 2007). The role a species has as a host or pathogen, may depend on its ability to compete with other taxa already present in the community. For example, plant pathogens may use physical barriers or toxins to exclude competitors from resource-rich niches within the host plant (Abdullah et al. 2017). The range or

variety of environmental conditions a species can exist or grow in may be captured in what is termed a niche (Sexton et al. 2017).

Ecological specialisation is similar in many respects to niche theory, which establishes the relationship between a species' resource exploitation strategy and the breadth of resources available to it (Hutchinson, 1957). Two hypothetical communities may have the same host species diversity but may differ according to the traits represented by, or resources available to, each community of hosts (i.e., functional diversity). Resource use, especially of the host species used by a virus, is usually generalised as the number of hosts used to fulfil their replication (McLeish et al. 2018). The breadth of environmental conditions a species can exist or grow in may be captured in what is termed a niche (Sexton et al. 2017). Hosts constitute a major environment of parasites, and classification of viruses by the number of hosts they are able to use gives two nominal groups: viruses that can only infect, multiply, and be transmitted to hosts of a single or few related species (specialist viruses), and viruses that can do so on hosts belonging to a wider group of species, often quite unrelated (generalist or multi-host viruses) (García-Arenal and Fraile, 2013). Plant viruses are largely generalists (McLeish et al. 2017) as is true for many pathogens of humans and other animals (Woolhouse et al. 2001), and the identification of ecological strategies of each virus species in various plant communities may therefore hold the key to predicting emergence events.

### **1.1.2. Host range evolution and virus emergence.**

Changes in host range are a critical factor in the emergence of viruses, and potentially influence transmission and survival (McLeish et al. 2019; García-Arenal and Fraile, 2013). Host range, defined as the number of species used or exploited by a pathogen, conditions the epidemiology of viruses, and is a key factor in the evolution of host relationships that has consequences for emergence. (Fraile and García-Arenal, 2018). To determine a pathogen's host range in absolute terms is generally impossible, and equally difficult is to identify "non-hosts". Thus, host range

should not be treated as a fixed value in natural systems since ecological factors such as distribution and abundance shape interactions between species (Arnold and Lutzoni, 2007).

Host range is a plastic trait that can vary as a result of changes to the ecology of the virus, host(s), or vector(s) (de Vienne et al. 2013). For example, a pathogen can acquire a new host in exchange for a previous host, with the number of hosts remaining unchanged; a phenomenon known as a host shift. The host-range mutations that result in host shifts can result in the virus being able to also infect species closely related to the new host (Bedhomme et al. 2015; Longdon et al. 2018). Such genetic changes may include mutations in key viral proteins or changes in host receptor binding that enhances adaptation to similar hosts species during host shifts and contribute to successful emergence in new species. In addition to changing hosts, a pathogen can also expand its host range (host breadth expansion) or contract it (host breadth contraction). Host breadth contraction occurs when a pathogen loses the ability to infect certain hosts, often due to specialization on a particular host species. For example, viruses that adapt strongly to a new host may lose compatibility with their previous hosts, leading to reduced host range (Poulin, 2010; Elena et al. 2014). Host range expansion refers to the process by which a pathogen increases the number of host species it can infect. The addition of new hosts through host range expansion creates an opportunity for a virus to spread and establish itself in new populations (Bera et al. 2018). An example of host range expansion is resistance breaking, which occurs when a virus acquires the capacity to infect and multiply in host plant genotypes that are immune or resistant to other virus genotypes. In crops, resistance breaking threatens the effectiveness of genetic resistance bred into cultivars, which is the only direct and most efficient way of controlling viral diseases (García-Arenal and McDonald 2003; Maule et al. 2007; Gómez et al. 2009). Host range expansions may also occur via ecological fitting where the virus, without adaptation, exploits traits already present in the new host (Janz et al. 2001; Engering et al. 2013; Peláez et al. 2020). Host shifts and host expansions are common in plant-virus relationships, often leading to devastating disease outbreaks (Zhong et al. 2016).

While generalist viruses with large host ranges should facilitate opportunities for transmission and survival, host range evolution is considered to largely favor the development of specialist strategies. This is because, differential host-associated selection would limit the fitness of

generalists in any one host, which would result in them being outcompeted by specialists (Woolhouse *et al.* 2001; Elena and Sanjuán, 2003). There is a large evidence of across host fitness trade-offs in plant viruses, but trade-offs can be modified by different environmental conditions (García-Arenal and Fraile 2013).

The distribution of viruses over their host range, that is, the frequency of infection in the different host species by a virus across plant communities, may vary largely. Habitat heterogeneity (defined as the differences in community composition among habitats) can determine the virus intrinsic traits of transmission mode and structure of the virus population, including aspects relevant for the role of hosts as reservoirs and inoculum sources (Haydon *et al.* 2002; Woolhouse *et al.* 2001; Woolhouse *et al.* 2005). Also, the distribution of viruses over their hosts will affect the frequency of co-infection of an individual host by different viruses (Malpica *et al.* 2006). Co-infecting viruses may interact in ways that alter their transmission modes within and between host species, influencing their overall host range. Thus, co-infection can influence host range expansions or contractions and promote virus emergence, making it a key factor in virus evolution and epidemiology.

### **1.1.3. Co-infection.**

The occurrence of co-infections has potential impact on the pathogenicity, evolution, epidemiology and control of viruses (Woolhouse *et al.* 2002). Co-infection, defined as infections of different virus strains or species within an individual host (Alcaide *et al.* 2020) also has important consequences for host range evolution (Read and Taylor 2001; Woolhouse *et al.* 2002). Co-infection affects host use through mechanisms such as resource competition, and may impact on the replication, virulence and transmission of viruses (Gómez *et al.* 2009; Marchetto and Power 2018). Competition for host resources will affect the fitness of each of the co-infecting viruses, hence, co-infection will be a factor in pathogen evolution in general (Sacristán, *et al.* 2004). Co-infection is necessary for genetic exchange via recombination or reassortment of segmented genomes among different virus genotypes or species (Pagán and García-Arenal 2018),

which can result in new genotypes that provide an opportunity for adaptation to new hosts (Fraile et al. 2008; Bedhomme et al. 2012) or to spread to new environments (Fiallo-Olivé et al. 2019).

Plant viruses co-infecting the same host may interact in either a synergistically or antagonistically (Syller, 2012; 2014; Syller and Grupa, 2016). In an antagonistic interaction, the presence and activity of one of the viruses reduces replication of the other, which results in less efficient transmission and lower fitness. Antagonism in transmission can also be due to cross-protection (Lacroix et al. 2014; Syller 2012). In a synergistic interaction both, or at least one of the co-infecting viruses replicates to higher levels than when in single infection, which would result in higher transmission rates (Syller, 2012). Increases of transmission rates of viruses in synergistic co-infections may also be independent of replication levels. For example, the helper component from one Potyvirus may assist aphid transmission of another with a less efficient or defective helper component (Murant, 1993; Syller, 2006), or the presence of a polerovirus may be required for the transmission on an umbravirus (Ryabov et al. 2001). However, synergism between viruses may lead to increased host mortality and result in lower transmission rates (Lacroix et al. 2014; Syller 2012). Co-infection may also result in synergisms through effects on the host. For example, infection by one virus may result in an increased host susceptibility to a second one, a common phenomenon among different types of pathogens named predisposition by plant pathologists (Woolhouse et al. 2002). Frequent co-infections may also contribute to the persistence in host populations of alleles conferring susceptibility to disease (McClelland et al. 2003).

## **1.2 Genetic determinants of host range evolution.**

A key trait that determines plant–virus interactions is genetic specificity that describes the subset of hosts that a virus can infect and replicate in, and how well a virus performs on different hosts (fitness across-hosts) (McLeish et al. 2019). Two major models have been presented to explain specificity of infection: 1) the gene-for-gene (GFG) model (Flor, 1956); and 2) the matching-alleles (MA) model (Grosberg and Hart, 2000). In the GFG model, a hierarchy of alleles that determine resistance in hosts and infectivity in viruses results in some alleles allowing the virus to infect a

larger range of host genotypes than other infectivity alleles. A key feature of the GFG model is that alleles are predicted to occur that allow infection of all host genotypes (universal infectivity). Such infectivity alleles will not become fixed in the virus population if increased infectivity is costly for the virus in the absence of resistance in the host, that is a cost of infectivity is required to maintain polymorphism of infectivity in the virus. Similarly, a cost of resistance is required to maintain polymorphisms for resistance in the host population. In the MA model, infection requires a specific match between the host and virus alleles. That is, a virus genotype must exactly match a host's genotype to elude the host's immune response and successfully infect, and universal infectivity is not possible under this model, with polymorphism at resistance and infectivity loci maintained by negative frequency-dependent selection (Agrawal and Lively 2002). An intermediate situation between these two models exists when viruses can partially infect a fraction of the full set of susceptible host genotypes under the GFG model (Parker, 1994).

The models can be extended to explain the evolution of generalism and specialism. The GFG model may be extended to explain host range expansions, as it allows a virus to adapt to an increasingly broader range of hosts. Over time, polymorphisms are maintained in both host and virus populations, allowing the virus to spread across species with varied resistance (Thrall et al. 2016). On the other hand, the MA model promotes host shifts, and host-virus coevolution continuously changes the genetic makeup of both populations, making it difficult for a single virus genotype to infect different host genotypes. This system restricts host range expansions but enhances adaptability to specific hosts. The assumptions in both models allow variation in fitness across hosts, but this results in different outcomes to host use; e.g., host expansions and host shifts. As an example, studies suggest that defence mechanisms are triggered upon recognition of a pathogen by non-specific immune response (And and Tsuda, 2010). Such defences are sufficient to prevent infection by generalist viruses and permit host shifts, however, specialist viruses harbour effectors that potentially switch off these defences, facilitating host expansion (Dodds and Rathjen, 2010). Thus, both GFG and MA models highlight the role of fitness costs associated with host-virus interactions and the evolution of host range (Thrall et al. 2016).

### 1.2.1 Across host trade-offs.

Across-host fitness trade-offs are considered to have a vital role in shaping host range evolution in plant viruses. Due to host-dependent fitness components, a pathogen cannot maximize its fitness across all potential hosts simultaneously. As a result, viruses often evolve to optimize their fitness in one or several closely related hosts, leading to fitness costs in others. These fitness trade-offs promote viral specialisation, which limits the expansion of host range (Elena and García-Arenal, 2023). Evidence for across-host trade-offs derive from the study of field isolates and from host passage experiments. For example, a twenty year-long study of the genetic composition of the tobamovirus population infecting pepper crops in southeastern Spain, in relation to the deployment of cultivars with different alleles at the resistance locus *L*, suggested costs of infectivity. Such costs were demonstrated by experiments in which the accumulation of the various virus genotypes was quantified in single infection or in competition in co-infections in susceptible pepper genotypes (Fraile et al. 2011). An example of across-host fitness trade-offs demonstrated by experimental evolution involves *Plum pox virus* (PPV), where serial passage in peas increased its infectivity and multiplication in the new host but decreased transmission efficiency in the original peach host (Wallis et al. 2007). Another example involves the rice stripe virus, which replicates in both plant and insect hosts, with different efficiency in each, which condition its host range evolution (Zhao et al. 2018).

There are different mechanisms underlying across-host fitness trade-offs. The simplest mechanism is antagonistic pleiotropy, in which mutations that increase fitness in one host are deleterious in another host (Whitlock, 1996). Evidence indicates that antagonistic pleiotropy is common in RNA viruses because of their small genomes, which are compacted with genetic information and encode few, multi-functional proteins (Bedhomme et al. 2015). A study was conducted by Moreno-Perez et al. (2016) to test whether resistance-breaking mutations in the coat protein of *Pepper mild mottle virus* (PMMoV) resulted in antagonistic pleiotropy. It was found that resistance-breaking mutations had pleiotropic effects on virus fitness, which were

antagonistic or positive depending on multiple factors, and underscored the complexity of host range evolution and the difficulty of predicting the overcoming of resistance factors in crops. Another mechanism causing across-host trade-offs is epistatic interactions between mutations that favour adaptation to a new host (Bedhomme et al. 2015). Epistasis is defined as the functional interaction between genetic mutations that influences a phenotype (Phillips, 2008). Epistasis is common in viruses and is a key determinant in adaptive processes because it affects the evolutionary pathways of genes (Bedhomme et al. 2015). Full genome sequences of experimentally adapted virus lineages showed that they frequently differ from their ancestors by several mutations opening the possibility of epistasis among mutations of ancestors (Hillung et al. 2015). Epistasis among many mutations show existing interactions rather than a multiplicative effect, and a degree of epistasis can be identified from experimental fitness measures of single and double mutants (Kouyos et al. 2007).

Both antagonistic pleiotropy and epistatic interactions among host-range mutations create an adaptive landscape that constrains virus host range expansions (Elena et al 2014). However, there is a large amount of evidence that many plant viruses have wide host ranges (Power and Flecker, 2003), and that the same genotype of one virus can infect many different hosts (Pelaéz et al. 2021; Zamfir et al. 2023). Both theory and experiments show that frequent transmission between hosts will prevent the development of across-host fitness trade-offs (Turner & Elena, 2000; Turner et al. 2010), which would explain this apparent discrepancy among data.

### 1.3 Virus communities and heterogeneous ecosystems.

Virus-host interactions occur across different levels of organisation from within cells to individual plants, plant populations and communities, and ecosystems (Lefeuvre et al. 2019; McLeish et al. 2020). Virus emergence is therefore a consequence of high-dimensional interactions of potentially many species. In multispecies systems, factors that explain the distribution of infections are largely unknown, because it is unclear how variation in the biotic and abiotic environment affects the susceptibility of hosts and their competence to transmit pathogens (Streicker et al. 2013). Environmental heterogeneity (EH) is an umbrella term that ecologists have

used to describe spatial heterogeneity (Ettema and Wardle 2002; Tamme et al. 2010; Stein et al. 2014) of key components of the abiotic and biotic environment, such as vegetation cover, climate, topography, soil (August 1983; Sparrow, 1999; Hortal et al. 2009; Barajas-Barbosa et al. 2020), and structural complexity (Poggio et al. 2010; Yang et al. 2015). Environmental heterogeneity has implications for transmission (Gurarie and Seto 2008; Yang et al. 2015) and the movement ecology of vector-borne pathogens (Ostfeld and Keesing 2000). For instance, a study of multi-host – multi-pathogen interactions showed that infection, transmission, and co-infection of fungal pathogens depended on functional groups of herbs, grasses, and legumes in communities of up to 60 species (Rottstock et al. 2014). The quantification of environmental heterogeneity has been very diverse because of the high variability in study systems, spatial scales and study taxa. The most frequently used measures for quantification of environmental heterogeneity include diversity of vegetation cover types and plant species richness (Barajas-Barbosa et al. 2020; Stein and Kreft 2015). The compositional characteristics of vegetation covers determine whether species interact with one another or not (Connor and McCoy 2001). In the context of this thesis, environmental heterogeneity refers to differences between plant communities as constituted by multiple habitat types, or habitat heterogeneity. Habitat heterogeneity modulates the ecology and evolution (Figure 1.2) of virus populations (Archie et al. 2009; Meentemeyer et al. 2012) and shapes virus emergence. A core focus for this thesis is to study the impact of dominant and distinct habitat types of an agricultural ecosystem on virus occurrence and genetic structure through analyses of host range, incidence and genetic diversity.

# Towards understanding the ecology and evolution of plant viruses at different tiers

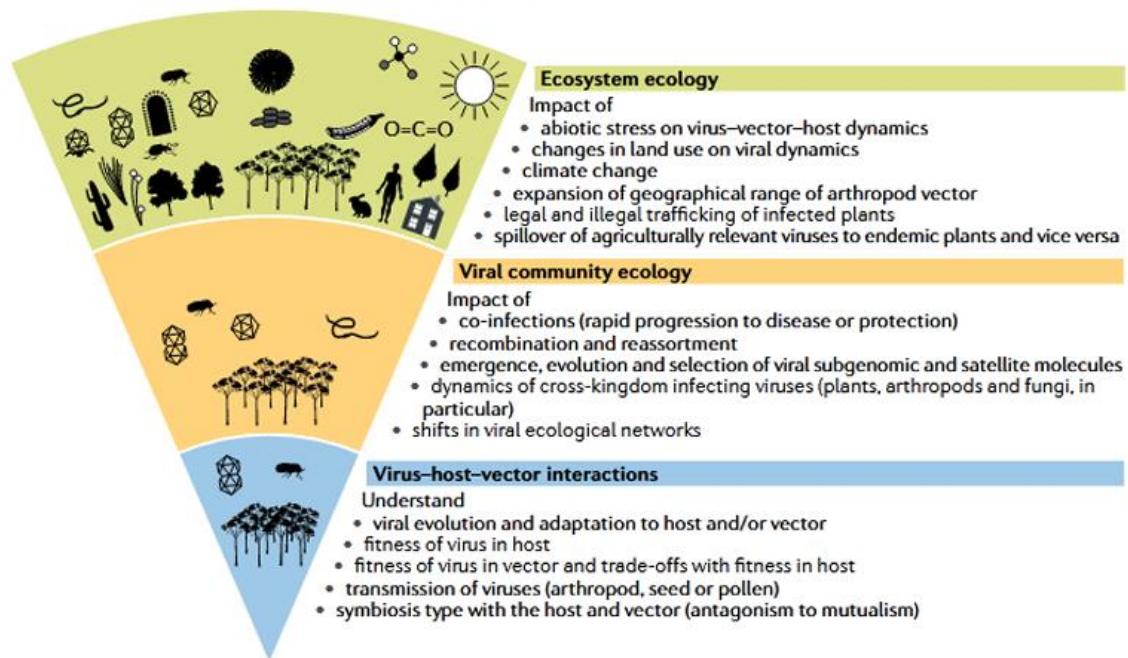


**Figure 1.2: Ecology and evolution of plant viruses at different levels.** A combination of HTS-based metagenomics and conventional virology methods can be used for virus discovery and study interactions up to the ecosystem scale (from Lefeuvre et al. 2019).

## 1.3.1 Plant-virus interactions in wild communities.

For a long time, crop plants have been the focus of virus studies, including studies on emergence, and knowledge on host ranges derives mostly from viruses infecting crops and plant communities of crop fields (Elena et al. 2009; Roossinck and García-Arenal, 2015). Thus, there is a shortage of knowledge about plant-virus interactions in wild plant communities as well as on transmission of viruses from wild plants to crops and vice versa. The need to consider wild plants in understanding virus emergence (Jeger et al. 2023) has been emphasized because of the role wild plants may have as reservoirs or inoculum sources in the epidemiology of viruses in crops, although this concept derives mostly from the study of wild plants growing within crop fields (i.e., weeds) or adjacent to them (Tugume et al. 2010; Moreno-Perez et al. 2016). Growing interest in the consequences of ecological processes on disease risk requires the study of disease systems in wild

plant communities, and not only in crop and adjacent communities. This is because plant diversity is temporally and spatially heterogeneous and should allow for a diversity of interactions with a high potential for host shifts (McLeish et al. 2019).

#### 1.4 Habitat heterogeneity of an ecosystem in Spain.

In agricultural ecosystems, anthropogenic disturbance has modified biodiversity and increased fragmentation (Tylianakis et al. 2007; Roossinck and García-Arenal 2015). Agricultural ecosystems in central Spain are characterized by variation among habitats that typically comprise diverse crop species, wild plant communities, and successional assemblies (McLeish et al. 2021a). Here, habitat refers to a given area that is characterised by plant community composition and the resources represented in each area. The key habitat categories in the ecosystem (Figure 1.3) include evergreen oak forest communities, (Oakwood), unused successional communities or “eriales” (Wasteland), communities on the edges (Edge) between crop fields, and the communities in crop fields (Crop). Variation among habitats is expected to affect virus host range evolution, as has been developed earlier in this Introduction. Differences in plant communities among habitats together with virus occurrence patterns can be used to understand how viruses might use specific hosts and this specificity might underlie the development of a reservoir (McLeish et al. 2017). Land-use practices in the Vega del Tajo-Tajuña agricultural region of the Tagus river basin, result in changes to host species distributions and abundances (McLeish et al. 2021b).



**Crop fields**



**Edges**



**Wastelands**



**Oakwoods**

**Figure 1.3: Habitats of the Vega del Tajo-Tajuña agricultural región.**

Crops are mostly seasonal annual monocultures, which are rotated or left fallow between seasons (Barbero-Sierra et al. 2016). In Crop, weed species are intermixed with the crop itself. In the alluvial soils of the river terraces (“vegas”) crops are irrigated, except for some winter crops (wheat, barley). Fields are irrigated and fertilized during cultivation, and are tilled between crops; therefore, weed populations cannot establish permanently but are colonizers that benefit from the nutrient-rich and well-prepared soil due to horticultural practices (Sacristán et al. 2004). The germination and survival of weed species within crop fields are influenced by both the cultivated crop species and the soil seed-bank inherent to the field (Wilson and Aebischer, 1995). However, changes in agricultural practices over the last decades, e.g., increased use of herbicides and fertilisers, have altered the composition of weed communities (Baessler and Klotz, 2006; Hyvönen, 2007; Travlos et al. 2018). In this habitat, the focus is on two irrigated summer crops,

*Cucumis melo* L. (melon) and *Zea mays* L. (maize), and two winter crops, *Brassica oleracea* L. (cabbage and cauliflower) and *Hordeum vulgare* L. (barley), since they represent a large fraction of the agricultural area of the region. The narrow borders (“lindes”) between crops constitute the Edge habitat, which is composed of relatively permanent plant assemblages of annual and perennial species (McLeish et al. 2021b; Sacristán et al. 2004). Plants that are present in Edge also benefit from the irrigation and fertilisation practices applied to the adjacent crops. Edge is generally left intact between seasonal cropping, with occasional low-level disturbance due to partial ploughing, fire or herbicide treatments (McLeish et al. 2017).

The native Iberian Mediterranean forests are mainly dominated by species of *Quercus* L. that cover large areas of central Spain under a broad range of altitudes, climates, and substrates (Gavilán et al. 2018). The Oakwood habitat primarily consists of several oak species, with *Quercus ilex* L. and *Quercus coccifera* L. being the common species in this area (Gavilán et al. 2018). Oakwood forests are considered natural communities of this region in Spain, which traditionally have been conserved as coppice stands and also as “dehesas” (Gavilán et al. 2018). Stands have been traditionally managed to obtain wood but have also been used for recreational activities such as hunting (Fabbio et al. 2003). In addition to species of oak, sclerophyllous shrubs, perennial herbs and annual species of grasses, legumes, and Asteraceae constitute a herbaceous component.

The Wasteland habitat supports sclerophyllous scrubland and grasslands (Ramírez-Sanz et al. 2000) that are occasionally grazed and presently not used by humans. Grassland can be in early- or late-successional stages that depend on community assembly processes. Both grassland stages can be related to quality indicators; e.g., medium disturbance regime, absence of nitrophilous species, and absence of encroachment of chamaephytes or shrub species (Molina et al, 2023). Wasteland varies as a consequence of episodes of human intervention that occurred during the second half of the 20<sup>th</sup> century; e.g., cultivation, intensive grazing, timber, charcoal exploitation, and land abandonment (Gallego Fernández et al. 2004). Once human interventions ceased, Wasteland came to support greater species diversity, which led to an improvement in soil structural stability, water-holding capacity, and input of organic matter to the soil (Valverde-Asenjo et al. 2020).

### 1.4.1 Virus detection and surveillance: development of HTS technologies.

The early detection and identification of existing and new viruses potentially causing disease is important for appropriate management procedures (Rubio et al. 2020) and crucial for timely intervention in virus emergence (Woolhouse et al. 2005). Virus surveillance was difficult until recently when 'omic technologies and computational tools for the analyses of big data greatly enhanced the ability to monitor entire communities.

High throughput sequencing (HTS) has been readily applied to the analyses of virome (collection of viruses found in a biological community) and incidence of viruses associated with wild plants, allowing descriptions of virus-host associations in plant communities of agricultural and non-agricultural ecosystems (Bernardo et al. 2018; Melcher and Grover, 2011; Muthukumar et al. 2009; Roossinck et al. 2010; Bačnik et al, 2020; Rivarez et al. 2021). The development of HTS technologies has revolutionised virus identification, detection, discovery, and community ecology (Roossinck, 2017; Villamor et al. 2019; Moubset et al. 2022; McLeish et al. 2024). Conventional methods for detection and identification such as serological (i.e. ELISA, Tissue blot-ELISA) or nucleic acid (PCR or probe-based) approaches require previously purified virus particles used for raising antibodies or nucleic acid sequences for the design of specific primers or target probes (Ward et al. 2004). Key advantages of HTS over these methods are: (i) nucleic acids (viral or non-viral) of the entire community can be sequenced in bulk; (ii) read sequence data are produced at relatively low cost; and (iii) the high volume of (big) data generated by HTS is analysed rapidly through high performance computing infrastructure and bioinformatic pipelines. These benefits allow entire plant, virus, and vector communities to be monitored across time and space. The ecosystem-scale perspective recognises that information about the influence of environmental heterogeneity on the movement ecology of interacting species is necessary for predicting disease risk. This thesis aims to leverage HTS to study the evolution and ecology of viruses in multiple plant communities across habitats of a heterogeneous ecosystem.

### 1.4.2 Application of metagenomics to virus ecology and evolution.

In the past decade, the field of metagenomics has been responsible for substantial advances in understanding microbial ecology and evolution. Metagenomics (Handelsman et al. 1998) is the study of genetic material recovered directly from environmental samples without the necessity of obtaining pure cultures in the laboratory (Thomas et al. 2012; Roossinck et al. 2015). The application of HTS techniques in metagenomics has indicated that plant virus diversity has been severely underestimated (Ma et al. 2019; McLeish et al. 2022 and 2024).

The analysis of viral nucleic acids with metagenomics sequencing (viromics) can be used to identify viruses that infect an individual host or a community of hosts with no bias towards those that cause diseases (Roossinck et al. 2010; Wu et al. 2015; Massart et al. 2017; Claverie et al. 2018; Ma et al. 2019) or visible symptoms (Wren et al. 2006, Gutiérrez-Aguirre et al. 2018; Pecman et al. 2022). HTS approaches have enhanced the study of interactions of multiple viruses within single hosts. Among the recently published studies we can single out several ones with different objectives. For instance, HTS was used to investigate the diversity of viral populations of tomato (*Solanum lycopersicum* L.) and wild *Solanum nigrum* L. populations, and revealed that *S. nigrum* may act as a reservoir host for viruses infecting solanaceous crops (Ma et al. 2020). In a study by Maachi et al. (2022) by means of HTS, virus inoculum fluxes between two hosts, a pluriannual cucurbit species, *Ecballium elaterium*, and melon plants were analysed. Population genetics showed no evidence of migration between the two hosts for *Cucurbit aphid-borne yellows virus* CABYV, whereas unidirectional migration of *Tomato leaf curl New Delhi virus* TYCNDV from melon to *E. elaterium* was observed. A community-based study (Thapa et al. 2015) of 400 plants from six wild plant species collected from 20 sites in the Tall-grass Prairie Preserve in Northeastern Oklahoma (USA), showed that host identity had a significantly stronger effect in virome composition than other factors such as location or sampling year (Thapa et al. 2015). An investigation of 1,725 georeferenced plant samples collected over two years in two biodiversity hot spots (Western Cape region of South Africa and Rhone River delta region of France), suggested that agriculture influenced plant virus distributions at the landscape scale (Bernardo et al. 2018). In an analysis of irrigation and surface water sources (Maksimović 2024), 20 families from known plant viruses were detected and ten viruses from economically important genera

*Tobamovirus*, *Aureusvirus*, *Betanecrovirus* and *Sobemovirus* were discovered. The study demonstrated the importance of environmental water samples as potential inoculum sources, and as indicators of the distribution and diversity of plant viruses for which only limited information was available from plants.

The ability to generate large data sets by HTS technology enables the study of systems of different complexity, from single pathogen – single host interactions to multi-host – multi-pathogen systems (Johnson et al. 2013; Woolhouse and Gowtage-Sequeria, 2005). Both approaches are followed in this Thesis, which includes the study of a system formed by different subspecies of a crop and its wild relative, and of multiple hosts in different plant communities.

## 1.5 The virome of a crop and its wild relative.

### 1.5.1 Description of the system: hosts strategies and host use.

Carrot represents a particularly interesting model plant as the cultivated crop (*Daucus carota* subspecies *sativus* L.) and its wild carrot relative, wild carrot (*Daucus carota* subspecies *carota* L.), belong to the same species and therefore represent a host pair with the lowest expected host genetic barrier to virus gene flow. Also, the cultivated and wild carrots often occur in nearby habitats. The carrot system provides a natural setting to observe how viruses adapt to use similar hosts in different environments (wild vs. cultivated), which can shed light on strategies for host use and range expansion. There is evidence that domesticated carrots likely developed from wild carrot in Central Asia, around 1,100 years ago (Iorizzo et al. 2013, Rong et al. 2014). Carrot is one of the important root vegetables of the family Apiaceae and is cultivated worldwide (Que et al. 2019). The family Apiaceae has about 3,700 species, some of them horticultural crops used as vegetables and condiments (Petruzzello, 2019). Cultivated carrot is a biennial plant. However, wild carrot shows variation in their life history strategies. It can be a winter annual, germinating in fall, surviving winter as rosettes and flowering the next summer, or occur as a perennial, surviving several years as rosettes with postponed flowering (Simon et al, 2008). These different strategies

may be highly dependent on genetic and environmental factors with, for example, annual plants being more common in resource-rich environments and perennials dominating in nutrient poor environments (Wijnheijmer, 1989; Mandel et al. 2019). The different nutrient and water availability, different plant density and different life cycles in cultivated and wild carrots may influence their susceptibility to viral infections and particularly to a range of viruses that affect their commercial production (Latham et al. 2004).

More than 30 viruses are known to infect carrot, of which the diseases caused by parsnip yellow fleck virus (PYFV), carrot red leaf virus (CtRLV), carrot mottle virus (CMoV), carrot red leaf associated RNA (CtRLVaRNA), and the disease of complex carrot motley dwarf (CMD) are the principal viral diseases that affect commercial carrot production (Adams et al. 2014). The economic importance of PYFV and CtRLV has been recognised due to the foliar symptoms (CtRLV) and viral die-back of seedlings (PYFV) (van Dijk and Bos 1985) they cause. These viruses affect carrot crops only sporadically but when they do occur they can be devastating.

In addition to viruses, subviral agents such as associated RNA (aRNAs) and satellite RNAs (satRNAs), often with unassigned taxa, have been reported infecting carrots. Examples are beet western yellows virus associated RNA (BWYVaRNA), carrot red leaf virus associated RNA (CtRLVaRNA) and carrot mottle virus satellite RNA (CMoVsatRNA) (Adams et al. 2014; Menzel et al. 2009). Satellite RNAs (satRNAs) are subviral agents that are dependent on a helper virus (HV) for critical functions such as replication, encapsidation, or transmission (Roossinck et al. 1992). Unlike defective RNAs, satRNAs have no sequence similarity to their HV. Some of them are capable of modifying, positively or negatively, the disease symptoms induced in host plants by the HV, while others do not seem to substantially affect their HV replication or the symptoms it causes (Hu 2009; Gnanasekaran and Chakraborty 2018). On the other hand, associated RNAs establish a mutualistic rather than parasitic interaction with the HV, to which they provide some lacking function often related to movement or transmission.

Reports of viruses, aRNAs or satRNA found affecting carrot and other Apiaceae species are mostly from cultivated crops. Less is known about viruses infecting wild carrots, which may represent an important reservoir or inoculum sources of carrot infecting viruses.

## 1.6 A multi-host – multi-virus study system.

This study was conducted in the Central Spain ecosystem described in section 1.3.3 of this Introduction. A metagenomic study of the 122 plant species for which at least five individuals were collected allowed the identification of 148 virus operational taxonomic units (OTUs), about half of which showed habitat specificity (McLeish et al. 2024). Habitat specificity is the richness and distinctiveness of a species' presence across different habitats (Halvorsen & Edvardsen, 2009). Out of the OTUs that were identified in more than one habitat, we selected eight with references in five different virus genera *Anulavirus*, *Cucumovirus*, *Potyvirus*, *Tobamovirus* and *Polerovirus*. These eight viruses allow a comparative analysis of virus transmission mode. Subsets of these viruses share a common evolutionary history, as represented at the level of genus, which can be used to determine within- and between-taxon responses to habitat heterogeneity. The five genera have messenger sense, single stranded RNA (ssRNA) genomes, as is the case for 75% of all reported plant viruses (ICTV 2017). In the next section information is given on the traits of the five virus genera, whose genome structure is shown in Figure 1.4.

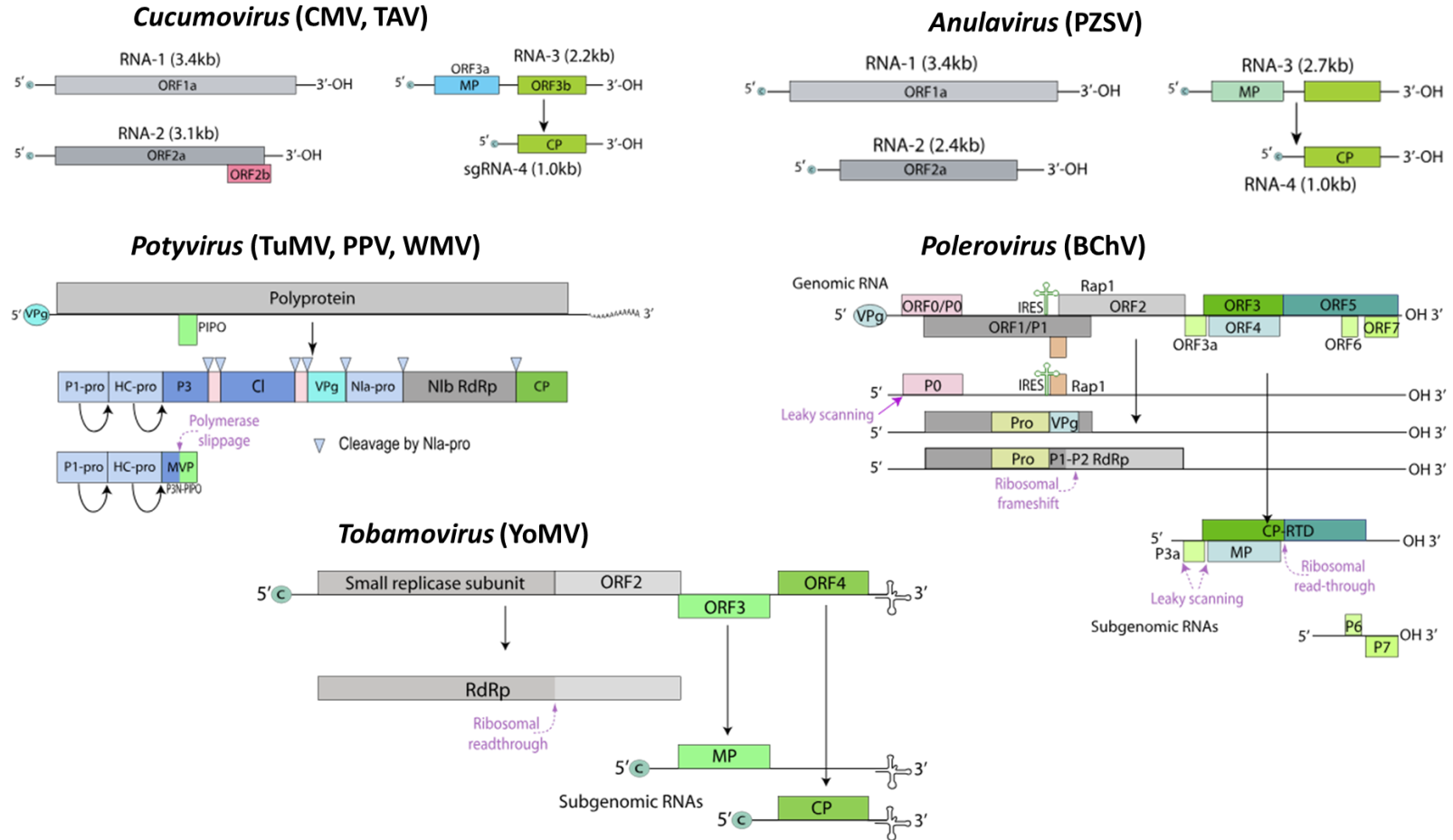


**Figure 1.4. Genome organization of virus genera analysed in this thesis.** Acronyms of the names of the analysed virus species of each genus are in brackets: cucumber mosaic virus CMV and tomato aspermy virus TAV, pelargonium zonate spot virus PZSV, turnip mosaic virus TuMV, plum pox virus PPV, watermelon mosaic virus WMV, beet chlorosis virus BChV, youcai mosaic virus YoMV. Source: ViralZone (<https://viralzone.expasy.org/>).

### 1.6.1 Description of virus genera analysed.

The genera *Anulavirus* and *Cucumovirus* belong to the virus family *Bromoviridae*. Viruses of this family cause major diseases in fruit, vegetable and fodder crops such as tomato, cucurbits, bananas, and alfalfa. *Bromoviridae* are characterised by three-partite genomes of about 8 kb, named RNA1, RNA2, and RNA3 according to decreasing size (Bujarski et al. 2019). Genomic RNAs are encapsidated in separate particles that may also contain subgenomic RNAs (sgRNAs), defective RNAs or satRNAs. RNA1 encodes the ~110 kDa protein 1a, with methyltransferase and helicase motifs, RNA2 encodes the 79–85 kDa protein 2a, which is the RNA-dependent RNA polymerase, RNA3 that encodes the 34 kDa movement protein (MP), and the 22–24 kDa coat protein (CP) that is translated from subgenomic RNA4. The 5' end of the genomic segments has a methylated cap structure, and the 3' end has a complex structure conserved among the three genomic RNAs. In the cucumoviruses this structure mimics a transfer RNA (tRNA) and can be amino acylated. A major difference between both genera is that Cucumoviruses, but not Anulaviruses, encode a second protein in RNA2, the ~19 kDa protein 2b, which is involved in suppression of post-transcriptional gene silencing and symptom induction (Palukaitis and García-Arenal, 2003). A unique feature of Anulaviruses is that RNA3 is slightly larger than RNA2, unlike other members of the family and its 2a protein is the smallest (78.9 kDa) reported for viruses in the family (King et al. 2012).

Anulaviruses such as *pelargonium zonate spot virus* (PZSV) are known to be vertically transmitted through seeds, and horizontally through pollen, facilitated by pollen-carrying flower thrips that mechanically injure leaves or flowers (Gallitelli, 1982; Lapidot, et al. 2010). Isolates of PZSV infect tomato and artichoke and are common in weeds (e.g. *Diplotaxis erucoides* where it is also seed-borne; Giolitti et al. 2013).

Cucumoviruses are transmitted by specific aphid species in a non-persistent manner, (Ferreles, et al. 2019). The particles of *cucumber mosaic virus* (CMV), the type member of the genus, are transmitted in a stylet-borne, non-persistent manner by more than 80 species of aphid in 33 genera (Palukaitis et al. 1992). Seed transmission of CMV also occurs in many hosts, with

frequencies ranging from <1% to 50% (Palukaitis and García-Arenal, 2003). Genetic diversity among CMV strains is high, and isolates have been classified into three subgroups (named IA, IB and II) according to the degree of similarity of their genomic RNAs (Palukaitis and García-Arenal, 2003). CMV has a very broad host range and is endemic in most temperate regions of the world (Palukaitis and García-Arenal 2019). The other cucumovirus in this study, *tomato aspermy virus* (TAV) predominantly infects composites and solanaceous plants (Palukaitis and García-Arenal, 2003).

The genus *Potyvirus* belongs to the *Potyviridae* family, the largest taxon of plant RNA viruses infecting a broad range of monocots and dicots in most climatic regions (Wylie et al. 2017; Moury and Desbiez 2020). Potyviruses have monopartite genomes of 10 kb encapsidated in filamentous, flexuous particles (Lecoq and Desbiez, 2012) and horizontal transmission is by aphids in a non-persistent, non-circulative manner (Wylie et al. 2017, Martínez-Turiño and García, 2020). The presence of the viral protein covalently linked (VPg) at the 5' terminus and a poly(A) tail at the 3' terminus are features of potyvirus genomes. The genome is organized into a single large open reading frame (ORF) that encodes a polyprotein, which is subsequently cleaved by viral proteases into functional proteins (P1, helper component protein (HC-Pro), P3, 6K1, cylindrical inclusions (CI), nuclear inclusions (NI a and b, CP) involved in genome replication, systemic movement, aphid transmission and countering of host defences. A short ORF embedded in the P3's in a different reading frame encodes PIPO protein, essential for cell-to-cell movement (ICTV 2021). In this study we focus on three potyviruses *plum pox virus* (PPV), *turnip mosaic virus* (TuMV) and *watermelon mosaic virus* (WMV), which are important pathogens of stone fruit, brassicas and cucurbit crops, respectively.

Poleroviruses (genus *Polerovirus*, family *Solemoviridae*) such as *beet chlorosis virus* (BChV), have monopartite genomes of 5.3-5.7 kb in size, which are encapsidated in spherical particles (Smirnova et al. 2015). The genome has a VPg at the 5' end, but no poly(A) tail at the 3' end. Poleroviruses are phloem-restricted and are horizontally transmitted by specific aphid species in a persistent, circulative, non-propagative manner (Harrison, 1999; Reinbold et al. 2001; Sömera et al. 2021). The genome organization of Poleroviruses includes seven overlapping ORFs. ORFs 1-2 are translated from the genomic RNA, ORF1 by leaky scanning of the ORF0 start codon, and

ORF2 by frame switching during the translation of ORF1, producing a fusion protein. These ORFs generate proteins with protease activity or involved in genome replication, plus the VPg. ORF3a, 3, 4 and 5 are translated from sgRNA1, and ORFs 6 and 7 from sgRNA2. These proteins are expressed by leaky scanning of the preceding start codons, except ORF5 that is expressed as a fusion protein with the product of ORF 3 by read-through of its stop codon. ORF 3 encodes the CP, and the ORF5 product RTD extension is required for aphid transmission. ORF 4 encodes the movement protein. The function of the proteins encoded by the other ORFs is unclear (ICTV, 2021). RNA recombination is frequent in Poleroviruses and contributes to the emergence of new species or strains (Chen et al. 2016), which renders taxonomy difficult. This is the case for beet-infecting Poleroviruses. The Polerovirus studied here, BChV is an important pathogen of sugar beet that causes, as for other Poleroviruses, a yellowing disease and other dicotyledon crops (Balke, 2023; Hauser et al. 2000).

Species of genus *Tobamovirus* (family *Virgaviridae*) are important pathogens of several horticultural crops world-wide and have had a central role in the development of different areas of virology, including virus evolution (Fraile and García-Arenal 2018). Tobamoviruses have 6.3-6.5 kb monopartite genomes encapsidated in highly stable rod-shaped particles that allow their long-term survival in the environment, particularly in the soil, where virions are liberated from the roots. In nature they are also transmitted through the seed and, horizontally, by plant-to-plant contact. Recently, passive (mechanical) horizontal transmission by pollinators and other insects has been reported (Dombrovsky and Smith 2017; Gooding 1986; Salem et al. 2022; Darzi et al. 2018; Levitzky et al. 2019; Okada et al. 2000). The genome is organized into four ORFs. ORFs 1 and 2 are translated from the genomic RNA giving a 126 kDa protein with activity methyltransferase and helicase, and a 183 kDa protein, the fusion with the product of ORF 2 that is translated by read through of the ORF1 stop, which is the RNA-dependent RNA polymerase. ORF3 is translated from a first sgRNA and encodes the MP, and ORF 4, translated from a second sgRNA, encodes the CP. Tobamoviruses are considered to have natural narrow host ranges associated with six plant families - *Solanaceae*, *Cucurbitaceae*, *Malvaceae*, *Fabaceae*, *Brassicaceae* and *Lamiaceae* (Stobbe et al. 2012), and it has been proposed that they have evolved closely with their eudicotyledonous hosts (Gibbs et al. 2015). Here we study *youcai*

*mosaic virus* (YoMV), initially described infecting cucurbits (Gibbs et al. 2015), which our metagenomic study identified as having a moderate host range.

## 2. OBJECTIVES.

The general objective of this PhD thesis is to improve current understanding of virus host range and host use in heterogeneous ecosystems, as it affects transmission dynamics and the process of virus emergence. This goal is approached by a combination of metagenomics and molecular detection methods, and bioinformatic analyses, to study virus distributions between plant communities. Virus resource use will be quantified through data of incidence and infection patterns at the individual plant level, in host populations, or in diverse plant communities. This objective was approached primarily through the study of viruses transmitted by aphids non-persistently, the major group of plant viruses according to mechanisms of horizontal transmission, considering viruses transmitted by other means for comparison

This general objective has been addressed by asking the following two questions:

(1) how do anthropic and wild plant communities' impact on virus-host interactions and transmission dynamics, and

(2) are the ecological strategies for host use similar for viruses that use the same mode of horizontal transmission,

which have been answered through the following specific objectives:

1. To compare and describe the virome of wild and cultivated carrots in France and Spain.
2. To determine host range and incidence of aphid and non-aphid transmitted viruses in four diverse habitats of an agricultural ecosystem in central Spain
3. To determine the role of virus intrinsic and extrinsic factors on host range and use.
4. To analyse co-infections according to virus intrinsic factors.
5. To analyse genetic diversity of virus populations according to host plants and habitat.

### 3. MATERIALS AND METHODS.

#### 3.1 Virome study in single host: carrot.

##### 3.1.1 Carrot sampling location and study area.

The study area was located in Segovia-Valladolid, an area in Central Spain. In early summer 2021, the sampling team visited six fields (1A, 1B, 2, 3, 4, 5) in total with field 4 belonging to wild carrots while the other five fields corresponded to cultivated carrots. Table 3.1 shows the names and locations of all six fields recorded with the geographic information system (GIS) provided by google maps.

**Table 3.1: Carrot sampling locations in Segovia.**

| Carrot Fields   | Area name               | GPS Co-ordinates |                |
|-----------------|-------------------------|------------------|----------------|
| 1A (Cultivated) | Gomezserracin           | 41° 17' 53.8" N  | 4° 20' 14.3" W |
| 1B (Cultivated) | Gomezserracin           | 41° 17' 53.8" N  | 4° 20' 14.3" W |
| 2 (Cultivated)  | Remondo                 | 41° 20' 43.1" N  | 4° 27' 46.9" W |
| 3 (Cultivated)  | Iscar                   | 41° 21' 01.3" N  | 4° 31' 28.8" W |
| 4 (Wild)        | Fresnedillas de Cuellar | 41° 19' 42.6" N  | 4° 26' 35.0" W |
| 5 (Cultivated)  | Fresnedillas de Cuellar | 41° 19' 42.6" N  | 4° 26' 35.0" W |

In each field, fifty individual plants were systematically sampled by collecting leaves from asymptomatic and symptomatic (chlorosis or redness) carrot plants (Figure 3.1). Common weeds such as *Datura sp*, *Solanum nigrum*, *Galium sp*, *Xanthium spinosum* were observed in the cultivated carrot field. Wild carrot plants growing in ditches by the motorway of Fresnedillas de Cuellar (SG-V-3341, Km 1) were very tall and branched. Plant samples were placed in labelled cooling boxes and transported back to the laboratory for processing. Samples were sorted by eliminating debris, weeds or insects before drying in an air-tight container with anhydrous calcium

chloride ( $\text{CaCl}_2$ ) for one week. For each field, a pool corresponding to the fifty sampled plants was assembled (0.1 g leaf per plant) and used for double-stranded RNA extraction.



**Figure 3.1: Carrot field samples from Segovia, Spain.** (A) Cultivated carrot samples from Field 1B showing chlorosis and redness. (B) Wild carrot samples with red leaves and white inflorescence. (C) Overview of cultivated carrot field showing patches of chlorotic leaves.

In addition to the six populations of wild and cultivated carrot sampled in Spain, a similar but larger scale sampling was performed in summer 2019 by French collaborators, Dr. Thierry Candresse and his PhD student Deborah Schöneegger. Sixteen different populations of cultivated and wild carrots were sampled locally in the Nouvelle-Aquitaine region of France, including four cultivated populations, four off-type populations growing within the fields (their early bolting phenology and root morphology suggest that they originated from pollen or seed contaminations of the planted commercial varieties and they are therefore referred to here as off-type) and eight wild carrot populations. In summer 2020, a similar sampling was conducted in multiple areas of France, covering a north–south country gradient. In total, 29 carrot populations were collected, comprising 11 cultivated carrots populations, three off-type populations and 15 wild populations. Therefore, a total of 45 carrot populations were sampled in France over 2 years. The locations of the sampled carrot populations used in the present study is shown in Figure 3.2.



**Figure 3.2: Location of carrot fields sampled in France (A) and Spain (B).** Sites of wild carrot populations are indicated with green pins, sites of cultivated populations are indicated by red pins, and sites of off-type populations are indicated by orange pins (Schönegger et al. 2023).

### 3.1.2 Double-stranded RNA purification and Illumina sequencing.

The same dsRNA extraction method was applied to both French and Spanish samples. Double-stranded RNA (dsRNA) was purified from each pool of carrot leaves. Samples were ground to a fine powder using liquid nitrogen in a mortar and pestle and dsRNAs were purified by two rounds of CF11 cellulose chromatography as described by Marais et al. (2018). RNA quality and concentration of purified dsRNAs were measured in nanodrop using 1  $\mu$ L volume (Table 3.2).

**Table 3.2: Nanodrop measure of purified dsRNA from Spanish carrot samples.** Total volume of elute is 20  $\mu$ L.

| Sample   | Concentration (ng/ $\mu$ L) | Absorbance 260/280 | Absorbance 260/230 |
|----------|-----------------------------|--------------------|--------------------|
| 1A       | 43.2                        | 2.02               | 1.30               |
| 1B       | 47.1                        | 1.79               | 1.29               |
| 2        | 27.6                        | 1.96               | 1.21               |
| 3        | 20.3                        | 1.83               | 0.89               |
| 4 (wild) | 47.0                        | 2.03               | 0.83               |
| 5        | 25.8                        | 1.84               | 1.00               |

Purified dsRNAs were converted to cDNA and randomly amplified while simultaneously adding MID tags by RT-PCR (François et al. 2018; Marais et al. 2018). Random amplification PCR products were purified using the MinElute PCR purification kit (Qiagen, Courtaboeuf, France) and their concentration determined spectrophotometrically. Equal amounts of amplification products from each sample/pool were sent for Illumina sequencing in a multiplexed format (2×150 bp) on the NovaSeq 6000 system at the GetPlaGe platform (GenoToul INRAE Toulouse, France).

### 3.1.3 Metagenomic analysis of Carrot samples

Read sequences were demultiplexed and quality trimmed on the CLC Genomic Workbench, version 21.0.3 (CLC-GW, Qiagen) using default settings and a minimum read length of 60 nucleotides (nt). Contigs were assembled de novo using a minimal contig length of 250 nt. Viral contigs were annotated by a comparison against the non-redundant sequence (nr) database in GenBank using BlastX (Altschul et al. 1990). Identified contigs were manually scaffolded if needed using the most closely related reference sequence from NCBI GenBank. Scaffolds were verified

and extended by repeated rounds of mapping of residual reads using CLC-GW. When a given virus was identified in multiple samples, the longest contig with the deepest coverage was finally selected to be used as a reference sequence for that virus and was deposited in GenBank. Novel virus species were identified and putative taxonomic assignments proposed based on the calculation of percent pairwise identities using Mega7 (Kumar et al. 2018) and multiple nucleotide or amino acid alignments with reference sequences selected according to the International Committee on Taxonomy of Viruses (ICTV) species demarcation criteria for each virus genus or family.

The identity of putative novel OTUs was further assessed by completing 5' and 3' genome ends by Rapid Amplification cDNA Ends (RACE) experiments using specific primers and the SMARTer RACE Kit (Takara Bio Europe, Saint-Germain-en-Laye, France), as recommended by the company. Specific two-step RT-PCR assays (Marais et al. 2011) were developed for selected novel viruses to identify infected individual plants from positive pools. Primers were designed manually in MEGA 7 using multiple alignments generated from contigs of the corresponding viruses and of related viruses and GenBank reference sequences, with the aim to design primers that would amplify all isolates of a given virus but be specific enough not to amplify isolates of closely related viruses. PCR products were visualized on 1.2 % agarose gels and selected amplicons were directly Sanger sequenced (Eurofins Genomics, France) to confirm the specificity of the amplification products.

## 3.2 Virome study in multiple hosts.

### 3.2.1 Study area and sampling.

The field samples for the virome study in multiple hosts were collected between July 2015 and June 2017, before I began my thesis. The strategic sampling design (McLeish et al. 2021) was conducted in a 2303 km<sup>2</sup> area of Vega del Tajo-Tajuña, an agricultural region of south-central Spain (Figure 3.3). Repeat sampling of multiple sites selected a priori, was conducted in each of four habitat categories: Crop, Edge, Oakwood, and Wasteland. The first two of the four habitats constitute what we termed the 'anthropic habitats.' Crops (Crop) are annual monocultures and

may include assemblages of wild plant species. Crops are left fallow between seasons or rotated. Narrow boundaries that separate crops (Edge) are relatively permanent communities of wild species that partially benefit from nutrient and water supplementation in adjacent crops. The second two habitat categories, which we termed 'non-anthropic', are firstly evergreen oak forests (Oakwood) that are the region's primary habitat for wild species. Oakwood is dominated by *Quercus* species with an understory of sclerophyllous shrubs and grasses. Structurally distinct from Oakwood are abandoned, undisturbed and patchy areas of successional scrubland (Wasteland) interspersed among the other habitats (McLeish et al. 2021).

Regardless of whether plants showed symptoms of virus infection or not, samples were taken systematically from the four habitats to minimise biases in the representation of plant species richness and relative abundance of these communities (Choset, 2000). A boustrophedon transect method (a line taken alternately from right to left and vice-versa, and so on) was used in all instances except for Edge that have narrow, linear configurations. In Oakwood and Wasteland, samples were collected from a 25 x 25 m area, and in Edge and Crop, samples were collected from a 25 m x 2 m area, at each site at each resampling. In total, 23 different sites were sampled: four sites each for Oakwood, Wasteland and Edge, while eleven sites belonged to Crop, with four sites corresponding to melon fields, two sites to barley fields, three sites to cauliflower fields, and two to maize fields. Collections at the non-anthropic habitats (Oakwood and Wasteland) were done in two different seasons: spring, and autumn when the plant diversity was maximal and minimal, respectively. Anthropic habitats (Crops and adjacent Edge sites) were sampled according to the crop season: melon and maize Crop collections were done in summer; barley Crop in spring and cauliflower Crop in winter.

From 23 sampling sites, a total of 78 collections yielded 6,291 individual samples, and formed an herbarium for subsequent taxonomic identification by Professor Rosario G. Gavilán (Universidad Complutense de Madrid, Madrid, Spain). She identified 329 plant species distributed over all four habitats. Rarefaction analyses (Hurlbert, 1971) performed by McLeish et al. (2021) was used to assess the sample representation and sampling effort in each habitat. Given the number of individual samples of each species at each collection ( $n = 78$ ) rarefaction analyses allowed the calculation of the number of species (species richness) expected in smaller subsamples of  $n$

individuals. By making predictions for successively smaller subsamples of individuals in a collection, each community can be represented by “rarefaction curves” that illustrate the number of species richness as a function of the number of individual samples in each collection (McLeish et al 2021a). The rarefaction analyses indicated near asymptotic relationships between the number of samples and the expected number of species.

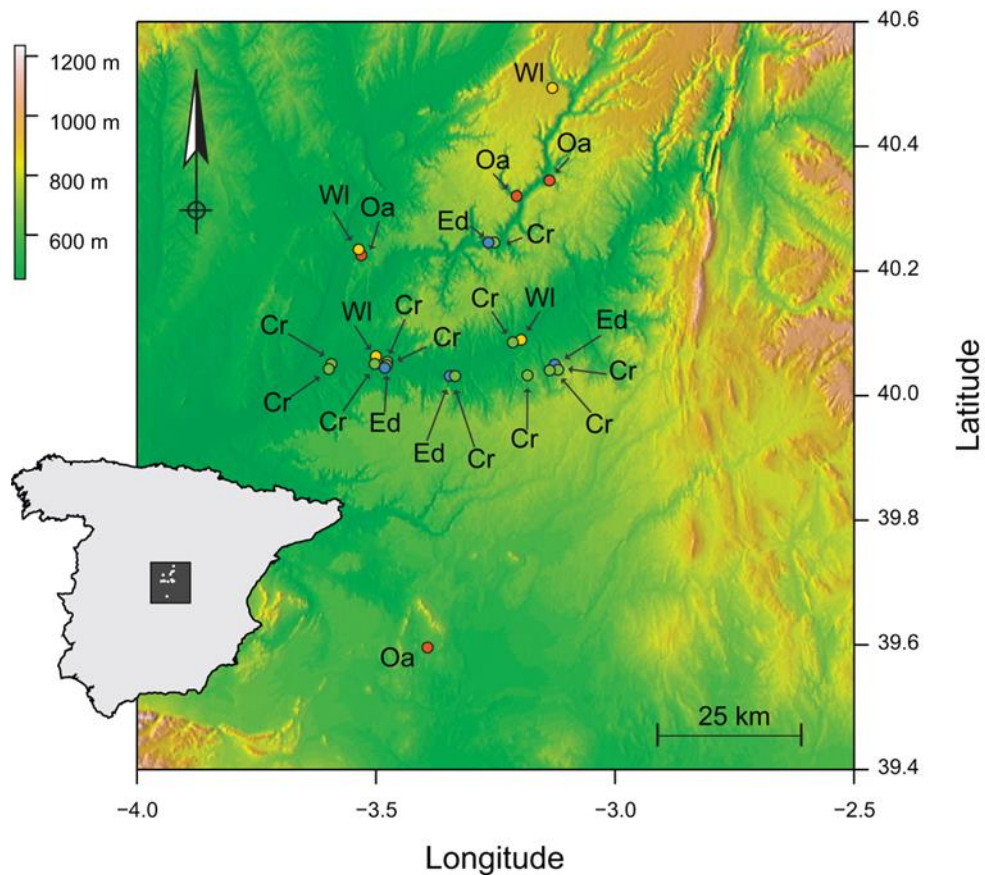


Figure 3.3: **Location of sampled sites in central Spain.** The study extent is shown, with the replicate sites of Crop (Cr, green), Edge (Ed, blue), Oakwood (Oa, red), and Wasteland (WI, yellow) sites in central Spain (inset). Elevation indicated at top left.

### 3.2.2 Sample processing and extraction methods.

In the field, all samples were placed into clear plastic bags and kept at 4°C for transport to the laboratory. The samples were processed in the laboratory by removing soil and non-plant material, bagged in sealable plastic bags, frozen in liquid nitrogen and stored at -80°C. Depending on the nature of plant species and leaf source, optimised procedures of tissue homogenisation were conducted to obtain high quality RNA preparations necessary for nucleotide sequencing and virus detection. Therefore, 0.3 g tissue of each individual plant sample was homogenized either by (I) mortar with a pestle and glass powder, (II) Eppendorf tube without glass powder, and (III) Eppendorf tube with glass powder. All tissue homogenisation methods were carried out using liquid nitrogen to prevent RNA degradation. Total RNA from individual plants of the species retained for HTS was extracted with the Agilent Plant RNA Isolation Mini Kit (Catalogue No.: 5185-5998), Spectrum™ Plant Total RNA Kit (Catalogue No.: STRN50), and the CTAB method (Chang et al. 1993) depending on the plant species. To minimise contamination across samples, extractions were done in groups of samples from the same species and collection (i.e. same time, same site). Plastic bags from field collections were opened in the laboratory and plant tissues were ground in liquid nitrogen with mortar and pestle (one for each individual sample) and sterile glass powder before adding the extraction kit reactants. Mortars and pestles were washed in 0.2 M NaOH and autoclaved before being used. Working surfaces were also washed with 0.2M NaOH between manipulation of groups of samples from the same plant species and collection. RNA extracts from same species and collection were pooled to obtain a single HTS library.

### 3.2.3 Library preparation and High throughput sequencing.

Rare species (defined here as those for which less than five individuals were collected in at least three of the four habitats) were removed from the sample of plant taxa for sequencing to minimise the potential for virus-negative libraries. Elimination of rare species is a common practice in studies of ecological interactions, as their inclusion may increase variances due to small sample sizes (Vázquez et al. 2005; Blüthgen et al. 2008). For abundant species (i.e., those with

more than five individuals collected in at least three of the four habitats), one or more individual RNA extracts of the same plant species and collection (numbers depending on the species and collection) were pooled to obtain a single HTS library. In total, 332 libraries of 2037 extracts were sent for sequencing at Centre de Regulacio Genòmica (Barcelona, Spain). Paired-end sequencing was performed on either HiSeq 2500, 3000 or 4000 Illumina platforms according to the provider's changes in infrastructure. Both forward and reverse read orientations of each pool were run in the same lane while each pool was run in a separate lane. All libraries included a step for rRNA depletion using the RiboZero kit. Paired end reads of 125 or 150 nt (2x125; 2x150 base pairs) were obtained with Phred quality scores greater than Q30. Trimming of adaptor contamination was conducted using cutadapt v1.8.3 (Martin 2011).

### 3.2.4 Metagenomic detection of virus OTUs

Viral sequences were detected in the HTS read libraries using a Basic Local Assignment Search Tool (BLAST) (Altschul et al. 1990). The queries were conducted against a database of plant virus genomic references of positive- and negative-sense single-stranded RNA (ssRNA), double-stranded RNA (dsRNA), and double- and single-stranded DNA (dsDNA and ssDNA) retrieved from the NCBI Viral Genome Browser (<https://www.ncbi.nlm.nih.gov/genomes> accessed in December 2018). Blast+ version 2.2.29 (Camacho et al. 2009) was used to identify OTUs derived from the HTS libraries. The results of the BLAST queries of each library were merged with taxonomic and study site details. The detection of an operational taxonomic unit (OTU) was considered credible if a read met the following *in-silico* selection criteria: (1) having a BLAST query coverage of 100%; (2) a query length of 125 nucleotides (together referred to as HTS2C detections). These two criteria were aimed at optimizing sequence similarity between read and reference, while allowing for mismatches and gaps expected from the presence of viruses that are divergent from the available genome references. Four additional criteria were also implemented in *the* procedure such that reads were retained only if: 3) they matched a single reference genome only; 4) they matched with a single region of that reference; 5) the difference between the maximum and minimum of query start positions relative to a reference genome had a span of more than 1% of

the length of the reference and; 6) the matches were with paired-reads only (HTS6C detections). The four additional criteria further optimised read–reference homology by taking advantage of read-pairing and set a minimum threshold of horizontal coverage of the reference genome.

The above metagenomic procedure focused on the detection of known (annotated) virus species from OTUs that fulfilled all six selection criteria. This resulted in the detection of 158 OTUs with high credibility of detection in the HTS libraries: In order to explore the role of habitat heterogeneity and transmission mode on infection patterns, viruses that occurred in more than one habitat and having different transmission modes were selected. Of the 158 OTUs of high credibility, thirteen OTUs were detected in plant species present in all four habitats, and thirteen more occurred in three habitats. Thus, I focused on eight OTUs that occurred in at least three of four habitats, all having ssRNA genomes but with different transmission modes, to analyse their infection pattern across the habitats: the eight OTUs are cucumber mosaic virus CMV, tomato aspermy virus TAV, pelargonium zonate spot virus PZSV, turnip mosaic virus TuMV, plum pox virus PPV, watermelon mosaic virus WMV, beet chlorosis virus BChV and youcai mosaic virus YoMV).

### 3.2.5 RT-PCR detection of viruses.

Although the *in-silico* criteria (HTS2C and HTS6C) were aimed to reduce false positive detections, the possibility of false positive results with the metagenomic detection protocol caused by sample crosstalk during handling and sequencing, and false negative RT-PCR results associated with the virus titre in the pool of total RNA extracts cannot be ignored (McLeish et al. 2022). Thus, the selected eight OTUs listed in section 3.2.4 underwent RT-PCR validation to confirm the detections by HTS. RT-PCR is known to be a reliable, sensitive and cost-effective option for virus detection in multiple samples and has been employed on a wide variety of RNA viruses in plants (Wang et al. 2022). Thus, the analyses of individual plants by RT-PCR allowed for estimation of virus incidence across multiple plant species and was used to identify co-infections of virus communities present in individual plants.

For OTUs such as CMV, PZSV, TAV, and WMV that were detected in a large number of hosts and in at least three habitats, a subset of the OTU-positive libraries was randomly sampled with the condition that about 50% of the hosts of each habitat were represented. Specific primers that amplified about 70-160 nucleotides from Coat Protein (CP) gene region references of each OTU were designed with the NCBI Primer-BLAST tool (Ye et al. 2012). In all cases, primer sequences were a consensus of the sequences for that taxon in the database (Table 3.3), and virus detection was performed from RNA extracts of each of the individual plant samples that were pooled for HTS in each OTU-positive library.

For RT-PCR, the complementary DNA (cDNA) was synthesised with the Maxima First Strand cDNA Synthesis Kit (ThermoFisher Scientific, Lithuania) on 0.5µg of the individual RNA extracts. The PCR was performed with the Supreme NZYtaq II polymerase kit (NZYtech, Lisboa, Portugal) with the following cycle conditions: initial denaturation at 95°C for 3minutes, forty cycles of denaturation at 94°C for 30seconds, annealing temperature based on each primer pair for 30 seconds, and extension at 72°C for 30seconds. Lastly, a final extension of 72°C for 7 minutes was carried out. The PCR products were visualised by gel electrophoresis in 2 % agarose, 1x TAE buffer (40mM Tris, 20mM acetic acid, 1 mM EDTA) (Sambrook et al. 1989) run for 40 minutes at 100V after ethidium bromide staining.

Specificity of RT-PCR primers was assessed with five distinct libraries from five different plant species that were unlikely hosts of each OTU, that is, libraries absent from the list of HTS2C detections, with no amplification expected in any case. The specificity of the RT-PCR detection was further verified by Sanger sequencing of 3 to 5 amplicons from RT-PCR-positive libraries of different plant species. For this, RT-PCR amplicons were purified with the Macherey-Nagel™ NucleoSpin™ Gel and PCR Clean-up Kit according to manufacturer's directions. 5-10µl of eluted DNA plus 5µl of forward primer were sent to Macrogen sequencing company and the chromatograms obtained from each sequence were analysed with MEGA X (Kumar et al. 2018). NCBI's BLASTn tool was used to check sequence identity and the BLAST output were compiled in a table. Both HTS and RT-PCR detections were used to estimate virus host range, defined as the number of host species used by a pathogen.

**Table 3.3. Specific primers used for RT-PCR detection of eight viruses.**

| <b>Virus</b> | <b>Primer name</b> | <b>Primer sequence</b> | <b>NCBI<br/>Accession</b> | <b>Position in<br/>Accession</b> | <b>Amplicon<br/>size (nt)</b> |
|--------------|--------------------|------------------------|---------------------------|----------------------------------|-------------------------------|
| CMV          | Fny-CMV_F          | CGTTGCCGCTATCTCTGCTAT  | NC_001440.1               | 1661-1681                        | 70                            |
|              | Fny-CMV_R          | GGATGCTGCATACTGACAAACC | NC_001440.1               | 1709-1730                        | 70                            |
| TAV          | TAV_83_F           | GCCATCCCTTCAACATCCGA   | NC_003836.1               | 1377-1396                        | 83                            |
|              | TAV_83_R           | TCGGTCGAACATCCAACGAA   | NC_003836.1               | 1440-1459                        | 83                            |
| PZSV         | PZSV_70_F          | GGGCTCTCTATGTCTGTTGA   | NC_003651.1               | 1721-1740                        | 70                            |
|              | PZSV_70_R          | GCCTACTTTCAGATTCCGTG   | NC_003651.1               | 1771-1790                        | 70                            |
| TuMV         | TuMV_85_F          | ACCTCCCCGAACATAAACGG   | NC_002509.2               | 9191-9210                        | 85                            |
|              | TuMV_85_R          | CGTGGTCAATGAGCGGTTTG   | NC_002509.2               | 9256-9275                        | 85                            |
| PPV          | PPV_94cp_F         | GACGAGGAGGAAGTTGATGC   | NC_001445.1               | 8592-8611                        | 94                            |
|              | PPV_94cp_R         | CTGGCGTGAAAATGGGGTTG   | NC_001445.1               | 8666-8685                        | 94                            |
| BChV         | BChV_83cp_F        | TGGAATGCTCAAGGCCTACC   | NC_002766.1               | 3894-3913                        | 83                            |
|              | BChV_83cp_R        | CGGAACTTTGGGAAGAGGCT   | NC_002766.1               | 3957-3976                        | 83                            |
| WMV          | WMV_93cp_F         | TCGCGGGAATTAACAGCAGG   | NC_006262.1               | 9653-9672                        | 93                            |
|              | WMV_93cp_R         | GATTCACGTCCCTTGCACTG   | NC_006262.1               | 9726-9745                        | 93                            |
| YoMV         | ORMV_160_F         | CCTTCTGAGTGCGATTGTGA   | NC_004422.1               | 5746-5765                        | 160                           |
|              | ORMV_160_R         | GATGGACGCGACTCTTCTTC   | NC_004422.1               | 5886-5905                        | 160                           |

### 3.3 Data analyses.

#### 3.3.1 Host range analyses.

Detections based on both HTS2C and RT-PCR were used to estimate host range. First, the HTS2C data set was used to show similarities or dissimilarities in host range breadth among OTUs by performing hierarchical clustering analyses. The clustering approach combines information from both the number of observations per host and number of observations for each OTU across host species to inform the clustering algorithm. Hierarchical clustering of host ranges per OTU was analysed by host species used and according to number of species per plant family. Host range analyses at both species and family level are necessary to provide evidence for the status of each viruses as specialists or generalists. Moreso, host range evolution of viruses to previously unreported host species and families could be an indicator of host expansion. Thus, the number of observations for each OTU across host species and families was used as input matrix and scaled using the R function 'scale', which first subtracts the column means from the columns (center step) and then divides the centered columns by the column standard deviations (scale step). The Euclidean distance method was selected with the Ward's ("ward.D2") option, implemented in the R package "*hclust*" (R Core Team, 2022), which uses agglomerative hierarchical clustering appropriate for objects of uneven variances (Murtagh and Legendre, 2014). The R package *heatmaply* was used to visualise the clusters. Bar graphs were used show habitat-related differences in the host ranges of each OTU by counting the number of species in which they occurred both by RT-PCR and HTS2C.

#### 3.3.2 Co-occurrence analyses.

Co-occurrence analyses were used to determine distribution patterns of OTUs and their potential interactions within communities. Co-infection of a host is common and viruses are assumed to colonise host species randomly with respect to one another. Therefore, a co-occurrence analyses

was performed both at the library level and individual plant level to provide evidence of non-random or random patterns of co-infection and identify virus-virus interactions that drives the risk of disease emergence. The identification of non-random patterns of infection is an important step for predicting transmission pathways in plant communities. Thus, the HTS2C data was converted into a binary matrix indicating virus presence (1) or absence (0) across hosts to calculate non-random virus-virus co-occurrences. Co-infection was analysed with individual plant samples. With the R package 'cooccur' (Griffith et al. 2016), a probabilistic and distribution free model based on interactions between different virus species, was used to calculate non-random virus associations. The algorithm compares the probabilities of observed frequencies of co-occurrence being significantly higher, lower or equal to the expected frequencies and then, in each case, deduces positive, negative or random associations between virus pairs (Veech, 2013).

### 3.3.3 Incidence.

The number of RT-PCR positive individual plant samples allowed to estimate the incidence of each virus, defined as the percentage of infected plant individuals over the total number of plant individuals analysed (Campbell and Neher 1994). Incidence data provides details of infection patterns in different hosts and resource utilization across habitats. For the multi-host study, incidence was measured in observed hosts ( $ih$ ) and in the entire plant community ( $iL$ ), constituting two variables. The first variable, incidence in observed hosts ( $ih$ ) is defined as the ratio of RT-PCR positive individuals to total number of analysed individuals for each host species, shown as:

$$ih = \sum_j i_j = \sum_j \frac{n_j}{t_j}$$

The second variable, incidence in the entire plant community ( $iL$ ), defined as incidence in the whole plant community including hosts and non-host is given as:

$$iL = \frac{\sum_j (i_j \times N_j)}{T}$$

Where  $i_j$  is the incidence in the  $j$ th host;  $N_j$  is the number of individuals in libraries of that host HTS2C positive for that virus OTU that were not analysed by RT-PCR; and  $T$  is the sum of all individual plant samples (analysed, non-analysed, and HTS-negative individuals) for a specific virus OTU.

### 3.3.4 Statistical analyses.

All statistical analyses were conducted with the R version 4.0.5 using packages specific for each test. The independence between HTS2C and HTS6C detection criteria and that between HTS2C and RT-PCR detections were evaluated using Kendall's rank correlation a non-parametric test appropriate for the presence of tied ranks and used to establish whether two variables are statistically dependent or not. Thus, the Kendall's rank coefficient ( $\tau$ ) was used to test the null hypothesis of independence between HTS2C and HTS6C or RT-PCR detections

Separate incidence data frames for  $ih$  and  $iL$  was compiled in a matrix with specified columns indicating the transmission mode, virus genera, OTU and habitat as predictor variables. Cases where incidence was based on a single individual plant per host species were eliminated to avoid overestimation of the incidence data. Normality of incidence data was tested using Shapiro-Wilk's test and when conditions for normality were not met, appropriate non-parametric tests were conducted. Chi-squared and regression analyses were conducted to show different aspects of the relationship between the response and predictor variables. First, the Chi-squared test was done to assess the overall association between each response variable ( $ih$ ,  $iL$ ) and predictor variables (virus genera, OTU, transmission mode, and habitat). Generalised linear model (GLM) regression analyses were used to elaborate on the specific relationships between different categories within each predictor and the response variable, considering the influence of other relevant factors. The overall fit of each model was compared to a null model by ANOVA using the likelihood ratio chi-squared test ( $G^2$ ), while the significance of coefficients in each model was analysed with Wald chi-squared test ( $W$ ).

The *glm* function from *Base R* package was used to perform logistic regressions with the underlying log likelihood (logit) function and predictor variables were classed as factors and considered as fixed effects; the level of each factor is shown in **Table 3.4**. The adjusted  $R^2$  after Nagelkerke (1991) from the package *rsq* (Zhang, 2017) was calculated to estimate the proportion of variation explained by each model in total and the proportion of variation explained by each factor. Post-hoc analyses were used to obtain estimated marginal means, conducted with the *emmeans* package, to further explore effects of predictors and their interactions in the model. Pairwise contrasts were performed with the *pairs* function from the *emmeans* package. For each model, the results from *emmeans* and pairwise comparisons provided several statistical estimates: that is, degree of freedom, odds ratios, standard error, *F*-statistic, adjusted *p*-value, and confidence intervals which were used to understand the effects of different levels within each factor. Odds ratios were log transformed to determine the probability of virus incidence, defined here as the likelihood or chance that a virus will be detected considering the influence of habitat, transmission mode, and virus genera. Probability was expressed as decimal value ranging from 0 to 1, where 0 indicates no chance of incidence and 1 indicates certainty of incidence. Thus, the odds ratios were used to quantify the strength of the levels in each factor (Kalra, 2016) while probability values were used to predict *ih* and *iL* according to level of each factor

**Table 3.4: Factor and their levels used in logistic regression analyses.**

| Factors      | Levels             |
|--------------|--------------------|
| Virus genera | <i>Anulavirus</i>  |
|              | <i>Cucumovirus</i> |
|              | <i>Polerovirus</i> |
|              | <i>Potyvirus</i>   |
|              | <i>Tobamovirus</i> |
| OTUs         | BChV               |
|              | CMV                |
|              | PPV                |
|              | PZSV               |

|                   |                |
|-------------------|----------------|
|                   | TAV            |
|                   | TuMV           |
|                   | WMV            |
|                   | YoMV           |
| Transmission mode | Contact        |
|                   | Non-persistent |
|                   | Persistent     |
|                   | Pollen         |
| Habitat           | Crop           |
|                   | Edge           |
|                   | Oakwood        |
|                   | Wasteland      |

### 3.3.5 Genetic diversity and Phylogenetic analyses.

A subset of read sequence libraries with alignments that produced a sufficient depth of coverage (**Table A26**) and satisfied the HTS2C selection criteria were used for population genomic analyses. As whole genomes may not be recovered at even depths (Gaafar and Ziebell 2020), libraries with at least 10X coverage of the genome were included in the analyses. Pilot analyses of genetic diversity of CMV in which libraries with progressively lower coverage were included, indicated that libraries with 10X coverage depth could be used for estimating genetic diversity without introducing errors or bias. Reads were aligned with their respective reference genome using the Burrows-Wheeler Alignment (BWA) algorithm (Li and Durbin 2009). Variant calling analysis was conducted with the binary alignment map (bam) files generated from the BWA output, and implemented with BCFtools (Li *et al.* 2009). The genetic diversity analyses and population genomics were performed with the R package *PopGenome* (Pfeifer *et al.* 2014).

The nucleotide diversity ( $\pi$ ), that is, the average pairwise difference between all possible pairs of individuals in a sample, was used as a measure of genetic diversity ( $p = S / L$  where  $S$  is the number of sites with more than one nucleotide variant and  $L$  is the number of nucleotides in the

sequence). Whole genome consensus sequences (WGCS) were used to estimate within-population genetic diversities using the Tamura 3-parameter model (Tamura 1992), with standard errors of each estimate based on 1000 bootstrap replicates, as implemented in Mega X (Kumar *et al.* 2018). To avoid biased estimates of  $\pi$ , missing data was allowed when it accounted for not more than 30% of the sequence and populations were defined either by habitat or study site.

A sliding window analysis was used to calculate nucleotide diversity along the genome (i.e. averaging across several variants). The approach allows for a systematic examination of localised nucleotide diversity across the genome. As the genomes of CMV, TAV, PZSV are tripartite,  $p$  was estimated over the full length of each segment. A window size of 1000 nt was used for all viruses, with a shift of 5 nt along the genome made for each estimate of  $p$ . Heatmaps were implemented with the R package *adeigenet* v2.1.5 (Jombart 2008) and used to visualise single nucleotide polymorphisms (SNPs) of genotypes in respect to the respective reference genomes of each virus.

To investigate the strength of association between CMV genotypes and habitat, level of anthropic disturbance (i.e., Crop with Edge and Oakwood with Wasteland), site, or host species observations (i.e., traits), phylogenetic inference and Bayesian tip association significance (BaTS) tests (Parker *et al.* 2008) were conducted. The BaTS analysis was used to test the observations inferred from the phylogenies and generates an association index (AI) and parsimony score (PS). The AI index measures the overall strength of association between genotypes at tips of phylogenetic trees and traits (the imbalance of internal nodes of the phylogeny), and the PS index measures the degree of homoplasy between genotypes in the phylogenetic tree and tip values. Thus, low AI values represent strong phylogeny–trait association and the PS value is inversely related to the strength of tip-trait association (Wang *et al.* 2001) (Parker *et al.* 2008). The BaTS approach repeatedly simulates associations under the null hypothesis that traits at the tips are randomly distributed across the phylogeny. The resulting  $p$ -values follow a uniform distribution, and the type 1 error of the test will be correct for all levels of statistical significance. Significant AI and PS  $p$ -values indicate that the observed association and level of homoplasy between genotypes and traits is unlikely to have occurred by chance alone. Whole genome consensus sequences were aligned with the CLUSTALW algorithm (Thompson *et al.* 1994) implemented in Mega X. Analyses were performed with MrBayes v3.2.7 (Huelsenbeck and Ronquist 2001) using

the General Time Reversible model (GTR) with gamma distributed substitution rates and invariant sites allowed. Four Markov chains were run for 20 million generations, sampling each chain every 500 trees. Posterior probabilities and mean branch lengths of Bayesian consensus phylogenies were derived from 30,000 post-burnin trees. A random sample of 2,000 posterior post-burnin trees from the Bayesian inference were used to account for phylogenetic uncertainty in the tip-association analysis. Convergence and posterior parameter distributions were assessed using the MCMC Tracer Analysis Tool v1.7.1.2 (Rambaut *et al.* 2014)

## 4. RESULTS.

### 4.1 The virome of a crop and its wild relative.

In line with understanding virus emergence and the application of HTS in ecological studies of plant-virus interactions, two approaches were proposed, as explained in the Introduction: the study of the virome of a crop and its wild relative, and the study of a multi-host multi-virus system. This section is therefore focused on the study of the virome of cultivated and wild carrots in different environments.

#### 4.1.1 Comparison of the virome of wild and cultivated carrot populations.

During the summers of 2019 and 2020, a total of 45 carrot populations including cultivated, wild and off-type carrots (carrots growing within the crop field and likely representing hybrids between cultivated and wild carrots) were sampled in France, covering a north-south gradient. Additionally, populations of five cultivated and one wild carrot were sampled in the province of Segovia (central Spain) in 2021. Illumina sequencing was performed on cDNAs obtained from highly purified dsRNA extracts from pools of 50 plants for each population. Following *de novo* assembly of reads from plant pools or from individually sequenced plants of wild and cultivated carrots, 1,672 viral contigs were identified via BLASTn and BLASTx- based contigs referencing. Analyses of the viral contigs revealed a virome comprising 45 viruses belonging to 12 different viral families (Figure 4.1) in French and Spanish carrot populations. In most cases, long contigs representing near complete genomes with average depth of coverage between 23x to 31592x were assembled, demonstrating the feasibility of assembling long contigs from pools of plants from the same species/location (Ma et al. 2020; Rivarez et al. 2022). Out of these, 20 viruses (44%) belonged to already known species, while the remaining 25 viruses (56%) belonged to novel species in nine different viral families or are unclassified subviral agents or putative novel *Riboviria* whose precise higher order taxonomic status remains to be established. This virome is dominated by (+) ssRNA viruses, most of them being aphid transmitted, followed by dsRNA viruses. It also

includes a substantial part of subviral agents including tombusvirus-like associated RNAs (tlaRNAs) and single-stranded, linear or circular satellite RNAs that are associated with aphid-transmitted viruses (see Tables 4.1 and 4.2).

Most of the known viruses identified were +ssRNA belonging to the *Betaflexiviridae*, *Bromoviridae*, *Closteroviridae*, *Secoviridae*, *Solemoviridae*, *Potyviridae* and *Tombusviridae*. The known carrot infecting viruses with dsRNA genomes were extended through the identification of novel members of the families *Amalgaviridae*, *Totiviridae*, *Spinareoviridae* as well as two novel viruses of the *Deltapartivirus* genus and one new host record for a divergent isolate of the betapartivirus DiCV2. Most of the novel viruses belonged to genera of +ssRNA viruses that are aphid transmitted (Figure 4.1) including members of the *Closteroviridae*, *Solemoviridae* and *Tombusviridae* families, as well as their associated subviral agents that are known to be involved in complex systems of mutual complementation and assistance. Umbraviruses (family *Tombusviridae*) that lack a coat protein (CP) gene, do not form virus particles and therefore rely on helper viruses, characteristically from the family *Solemoviridae* (former *Luteoviridae*), for encapsidation and aphid transmission (Ryabov et al. 2012). Details of novel and known viruses identified are presented in Tables 4.1 and 4.2 respectively.

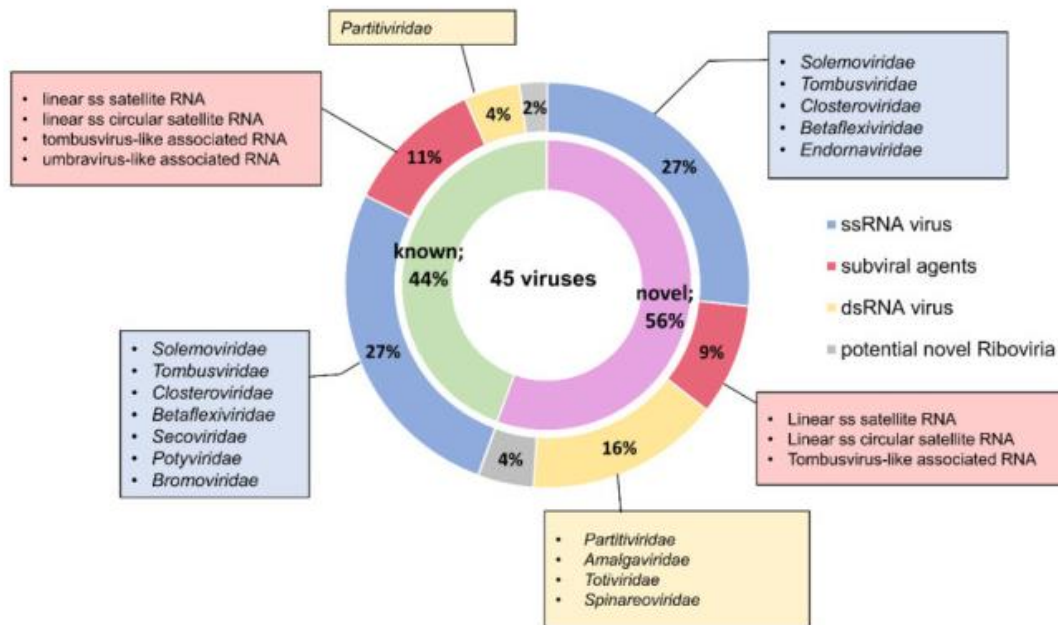


**Figure 4.1 Taxonomy and proportion of known and novel viruses detected in carrot populations from France and Spain.** The viral families are indicated, and proportions correspond to the number of viruses identified for each family (Schönegger et al. 2023).

**Table 4.1: Novel viruses identified in different carrot populations in France and Spain (adapted from Schönegger et al. 2023)**

| Virus Family                    | Virus genus              | Virus name                             | Virus acronym | Genome type <sup>b</sup> | Accession number <sup>a</sup> | Scaffold length (nt)    | Average coverage | Population type | Country |
|---------------------------------|--------------------------|----------------------------------------|---------------|--------------------------|-------------------------------|-------------------------|------------------|-----------------|---------|
| <i>Solemoviridae</i>            | <i>Polerovirus</i>       | Carrot polerovirus 1                   | CaPV1         | ssRNA+                   | OP886450*                     | 5671                    | 873x             | wild*           | France  |
| <i>Solemoviridae</i>            | <i>Polerovirus</i>       | Carrot polerovirus 2                   | CaPV2         |                          | OP886451                      | 5213                    | 324x             | wild*           | France  |
| <i>Solemoviridae</i>            | <i>Enamovirus</i>        | Carrot enamovirus 1                    | CaEV1         |                          | OP886449*                     | 5100                    | 31x              | wild            | France  |
| <i>Tombusviridae</i>            | <i>Umbravirus</i>        | Carrot umbravirus 1                    | CaUV1         |                          | OP886454                      | 3308                    | 3033             | wild            | France  |
| <i>Tombusviridae</i>            | <i>Umbravirus</i>        | Carrot umbravirus 2                    | CaUV2         |                          | OP886452*                     | 4255                    | 605x             | wild            | France  |
| <i>Tombusviridae</i>            | <i>Umbravirus</i>        | Carrot umbravirus 3                    | CaUV3         |                          | OP886453                      | 4059                    | 1421x            | wild            | France  |
| <i>Closteroviridae</i>          | <i>Closterovirus</i>     | Carrot closterovirus 2                 | CtCV2         |                          | OP886455*                     | 16302                   | 24536x           | cultivated      | France  |
| <i>Closteroviridae</i>          | <i>Closterovirus</i>     | Carrot closterovirus 3                 | CtCV3         |                          | OP886456*                     | 16291                   | 4173x            | cultivated      | France  |
| <i>Betaflexiviridae</i>         | <i>Chordovirus</i>       | Carrot chordovirus 3                   | CChV3         |                          | OP886457                      | 6881                    | 39x              | off-type        | France  |
| <i>Betaflexiviridae</i>         | <i>Chordovirus</i>       | Carrot chordovirus 4                   | CChV4         |                          | OP886458*                     | 7703                    | 23x              | wild            | France  |
| <i>Betaflexiviridae</i>         | <i>Vitivirus</i>         | Carrot vitivirus 1                     | CaVV1         |                          | OP886459*                     | 7546                    | 221x             | wild            | France  |
| <i>Endornaviridae</i>           | <i>Alphaendornavirus</i> | Carrot                                 |               |                          |                               |                         |                  |                 |         |
| linear ss                       |                          | alphaendornavirus                      | CaAEV1        |                          | OP886460                      | 14578                   | 608x             | wild            | France  |
| satellite RNAs                  | unclassified             | Carrot mottle virus                    |               |                          |                               |                         |                  |                 |         |
| tombusvirus-like-associated RNA | unclassified             | satellite RNA 2                        | CtMoVsatRNA2  |                          | OP886481                      | 323                     | 620x             | wild            | France  |
| small circular ss               | unclassified             | Carrot red leaf virus-associated RNA 2 | CtRLVaRNA2    |                          | ON603907                      | Schönegger et al. 2022a |                  |                 |         |
| satellite RNAs                  | unclassified             | Carrot red leaf virus satellite 1      | CtRLV Sat1    |                          | OM962993                      | Babalola et al. 2022    |                  |                 |         |
| small circular ss               | unclassified             | Carrot red leaf virus satellite 2      | CtRLV Sat2    |                          | OM962994.1                    | Babalola et al. 2022    |                  |                 |         |
| <i>Partitiviridae</i>           | <i>Deltapartitivirus</i> | Carrot cryptic virus 2                 | CaCV2         | dsRNA                    | OP886461-62*                  | 3329                    | 31592x           | wild            | France  |
| <i>Partitiviridae</i>           | <i>Deltapartitivirus</i> | Carrot cryptic virus 3                 | CaCV3         |                          | OP886463-64*                  | 3345                    | 13634x           | wild            | Spain   |
| <i>Amalgaviridae</i>            | <i>Amalgavirus</i>       | Carrot amalgavirus 1                   | CaAV1         |                          | OP886466                      | 3223                    | 929x             | wild            | France  |
| <i>Amalgaviridae</i>            | <i>Amalgavirus</i>       | Carrot amalgavirus 2                   | CaAV2         |                          | OP886465                      | 777                     | 19x              | wild            | Spain   |

|                          |                          |                                   |        |        |              |                         |        |      |        |
|--------------------------|--------------------------|-----------------------------------|--------|--------|--------------|-------------------------|--------|------|--------|
| <i>Reoviridae</i>        | <i>Reovirus</i>          | Carrot reovirus 1                 | CaRV1  |        | OP886467-76* | 26081                   | 12089x | wild | France |
| <i>Reoviridae</i>        | <i>Reovirus</i>          | Carrot reovirus 2                 | CaRV2  |        | OP886477-79  | 3163                    | 156x   | wild | France |
| <i>Totiviridae</i>       | <i>Totivirus</i>         | Carrot-associated toti-like virus | CaaTLV |        | OP886480*    | 5128                    | 8876x  | wild | France |
| putative novel Riboviria | putative novel riboviria | Carrot flavi-like virus 1         | CaFLV1 | ssRNA? | OM681407.1   | Schönegger et al. 2022b |        |      |        |
| putative novel Riboviria | putative novel riboviria | Carrot Ker-like virus             | CaKLV  | ssRNA? |              |                         |        |      |        |

a: viral sequences deposited in GenBank and corresponding to a complete genome or, at least, to a complete coding potential are indicated by an asterisk.

b: for viruses with divided genomes, the cumulated genome length is indicated.

**Table 4.2: Known viruses identified in different carrot populations in France and Spain (from Schönegger et al. 2023).**

| Virus Family                    | Virus genus                     | Virus name                                | Acronym    | Genome nucleic acid | Accession number of Reference |
|---------------------------------|---------------------------------|-------------------------------------------|------------|---------------------|-------------------------------|
| <i>Solemoviridae</i>            | <i>Polerovirus</i>              | Carrot red leaf virus                     | CtRLV      |                     | NC_006265                     |
| <i>Solemoviridae</i>            | <i>Polerovirus</i>              | Wild carrot red leaf virus                | WCtRLV     |                     | LT615231                      |
| <i>Tombusviridae</i>            | <i>Umbravirus</i>               | Carrot mottle virus                       | CMoV       |                     | KF533714                      |
| <i>Tombusviridae</i>            | <i>Umbravirus</i>               | Wild carrot mottle virus                  | WCMoV      |                     | LT615232                      |
| <i>Tombusviridae</i>            | <i>Umbravirus</i>               | Carrot mottle mimic virus                 | CMoMV      |                     | NC_001726                     |
| <i>Closteroviridae</i>          | <i>Closterovirus</i>            | Carrot closterovirus 1                    | CtCV1      |                     | KF533697                      |
| <i>Betaflexiviridae</i>         | <i>Closterovirus</i>            | Carrot yellow leaf virus                  | CYLV       |                     | KF533699                      |
| <i>Betaflexiviridae</i>         | <i>Chordovirus</i>              | Carrot Chordovirus 1                      | CChoV1     |                     | NC_025469                     |
| <i>Betaflexiviridae</i>         | <i>Chordovirus</i>              | Carrot Chordovirus 2                      | CChoV2     | + ssRNA             | NC_025468                     |
| <i>Secoviridae</i>              | <i>Torradovirus</i>             | Carrot Torradovirus 1                     | CaTV1      |                     | NC_025479 and NC_025480       |
| <i>Potyviridae</i>              | <i>Potyvirus</i>                | Apium virus Y                             | ApVY       |                     | NC_014905                     |
| <i>Bromoviridae</i>             | <i>Ilarvirus</i>                | Solanum nigrum ilarvirus 1                | SnIV1      |                     | MN216370                      |
| linear ss satellite RNAs        | unclassified                    | Carrot mottle virus satellite RNA         | CMoVsatRNA |                     | NC_030649                     |
| Tombusvirus like associated RNA | unclassified                    | Carrot red leaf virus associated RNA      | CtRLVaRNA  |                     | NC_003871                     |
| Tombusvirus like associated RNA | unclassified                    | Arracacha latent virus E associated RNA   | ALVEaRNA   |                     | MF136436                      |
| Tombusvirus like associated RNA | unclassified                    | Beet western yellows virus associated RNA | BWYVaRNA   |                     | KF533709                      |
| Umbra-like associated RNA       | unclassified                    | Parsley umbravirus 1                      | ParUV1     |                     | OM419177.1                    |
| <i>Partitiviridae</i>           | <i>Alphapartitivirus</i>        | Carrot cryptic virus 1                    | CaCV1      | dsRNA               | NC_038824 and NC_038823       |
| <i>Partitiviridae</i>           | <i>Betapartitivirus</i>         | Dill cryptic virus 2                      | DiCV2      |                     | NC_021147 and NC_021148       |
| <i>putative novel Riboviria</i> | <i>putative novel riboviria</i> | Carrot-associated RNA virus 1             | CaRNAV1    | ssRNA?              | OM419188                      |

#### 4.1.2 Virome structure in cultivated and wild carrots.

Of the 45 viruses detected, 25 were found only in the French carrot populations, 4 were found only in the Spanish populations while 16 were observed in both. The virome richness, defined as the number of virus species, detected in each French carrot population type over two years showed a significant difference between cultivated and wild carrots ( $t$ -test = 3.65,  $p$ -value = 0.004) with a higher virome richness in wild carrot populations ( $15.5 \pm 2.4$ ) compared to cultivated ( $12.2 \pm 2.6$ ). Off-type carrot populations had an intermediate richness ( $14 \pm 1.8$  viruses), so that the differences with cultivated or wild populations were not statistically significant. Sequence analyses of the Spanish carrot virome with mapping criteria of 95% read length and 90% identity showed that the number of known viruses detected were higher in the cultivated than wild populations (Figure 4.2). Also, for known viruses present in wild populations, the number of reads (black bars) were lower than in the cultivated populations. The conspicuous orange and yellow bars show that reads of the known viruses were higher (in millions) in cultivated populations than in the wild. Apium virus Y (ApVY) was the only case with more reads in wild Spanish carrot populations. Figure 4.2 also shows that though reads of known viruses such as *Carrot torrado virus* (CaTV1), *Carrot red leaf virus* (CtRLV), *Wild carrot red leaf virus* (WCtRLV) and *Carrot red leaf virus-associated RNA* (CtRLVaRNA) were detected in the Spanish cultivated populations, they either were relatively lower (in thousands) in the wild populations or were not detected at all.

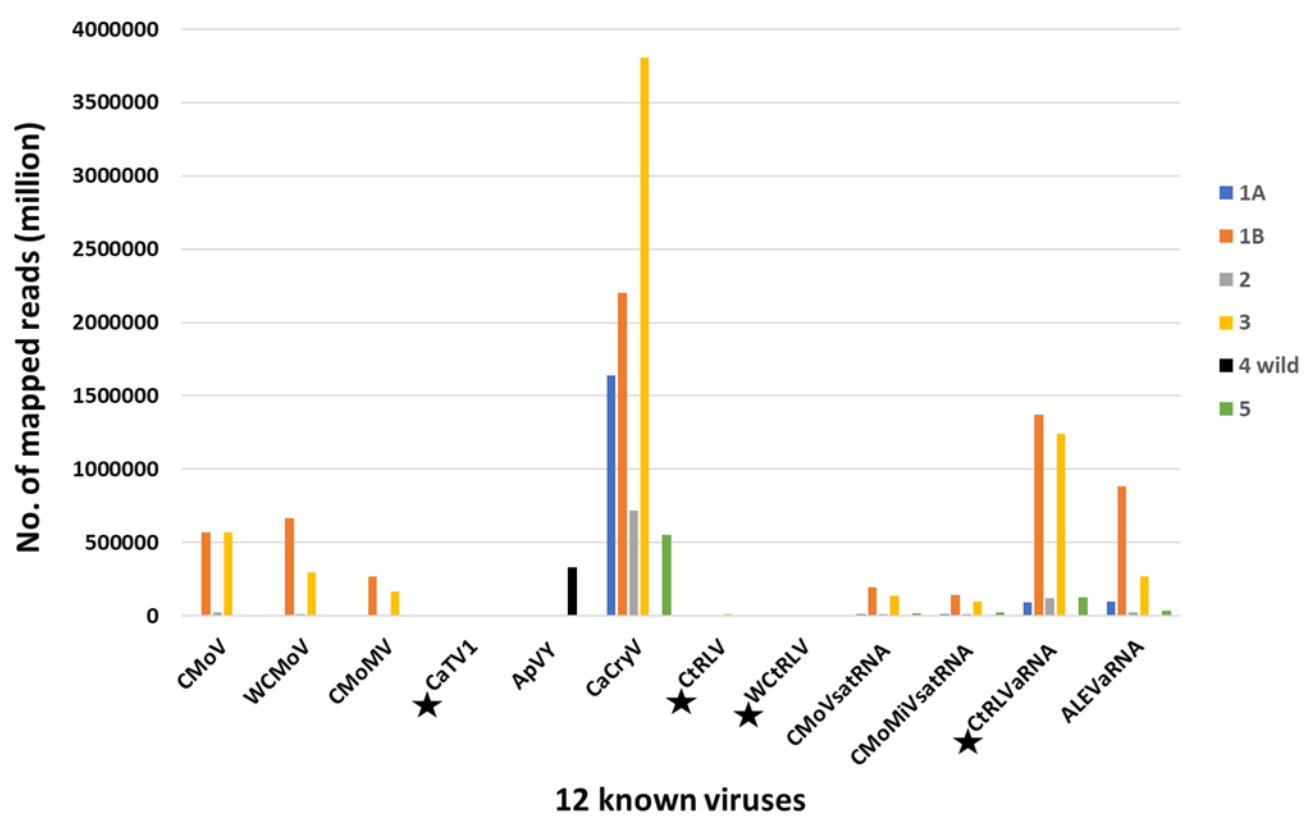


**Figure 4.2 Sequence analyses of mapped reads for 12 known viruses detected in Spanish carrot populations.** Legend shows six carrot populations 1A, 1B, 2, 3, 4 (wild) and 5. Black stars highlight cases of known viruses with few reads in cultivated and no reads found in wild population.

Of the 12 known viruses detected in Spanish carrot populations, I focused on two that had been previously reported in other European countries but not in Spain, that is, *Carrot torrado virus 1* (CaTV1) and *Apium virus Y* (ApVY). To confirm their presence in the wild and cultivated carrot populations, I performed validation by RT-PCR with specific detection primers for CaTV1 (CaTV1\_300\_F: 5'-TGTGCAACCACGAGGAATACA-3' and CaTV1\_300\_R: 5'-GATGCCTCATAGCAAAGTGTGTCAT-3' Rozado-Aguirre et al. 2016) and ApVY (ApVY\_458\_F: 5'-ACGCCATGCATTTCAAGTGTC-3' and ApVY\_458\_R: 5'-ATGTGTTCTGATCAGGGTGC-3') because they both were detected with few reads. The result obtained from the validation showed that CaTV1 was present in cultivated carrots but was not detected in wild carrots, while ApVY was mainly in wild carrots with a weak amplification in cultivated carrots (Figure 4.3). Thus, these results confirmed the presence of these viruses in Spain for the first time and contributes

information on their geographical distribution within the European Union. In addition to CaTV1 and ApVY, I also validated two novel hammerhead satellites (HhSat1 and 2) discovered with few reads in the Spanish virome that shared homology with satellite RNA of Cereal Yellow dwarf virus (CYDV-RPA satRNA). The RT-PCR validation results are shown in Figure 4.3.

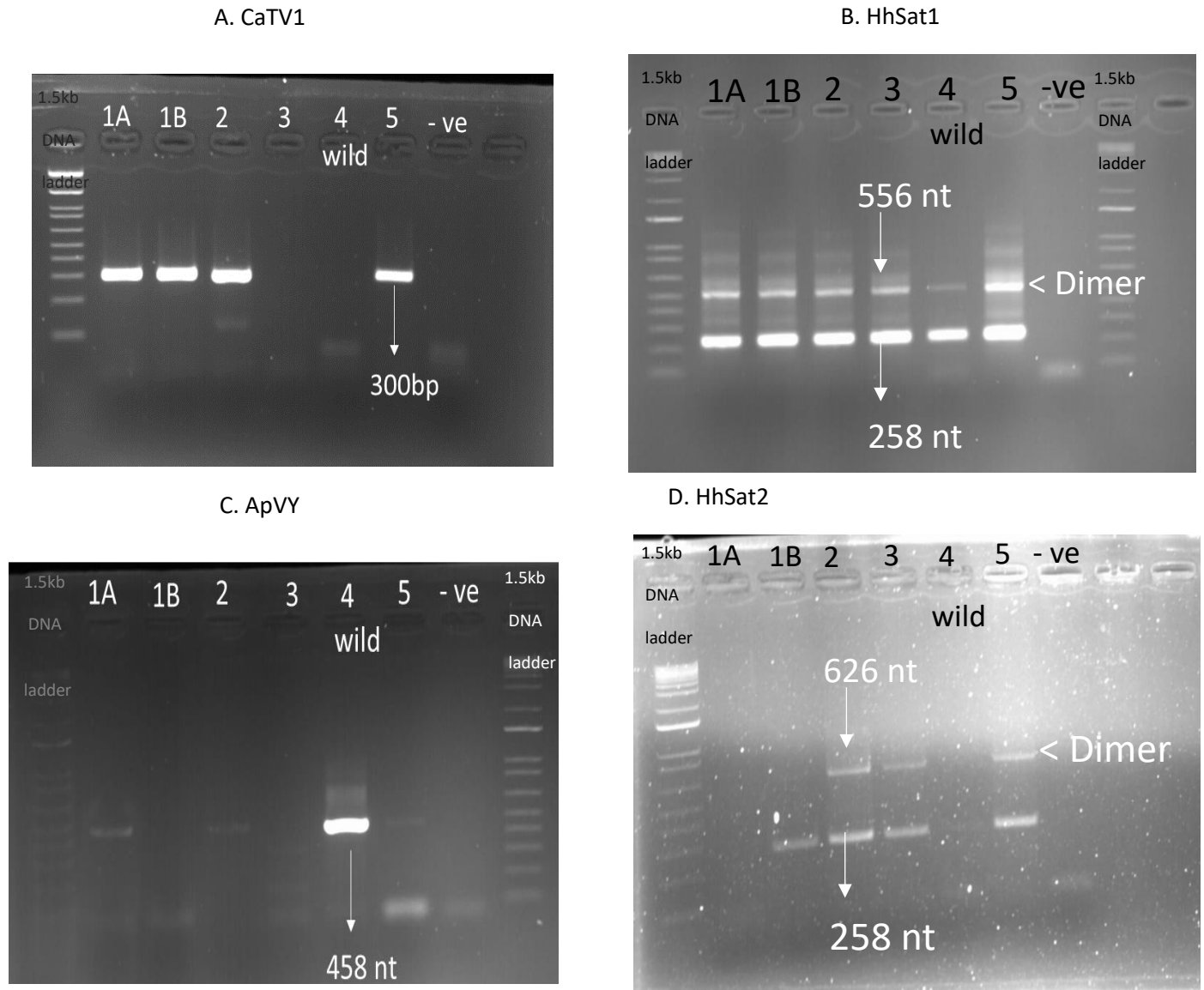


**Figure 4.3: RT-PCR validation of the detection of CaTV1, ApVY and two hammerhead satellites (HhSat 1 and 2) in Spanish carrot populations.** Left to right: in each gel, 1<sup>st</sup> lane is 1.5kb DNA ladder followed by six carrot populations: 1A, 1B, 2, 3, 4 wild and 5 in 2<sup>nd</sup> to 7<sup>th</sup> lane respectively while 8<sup>th</sup> lane contains negative control (-ve). Size of amplicon is indicated for each virus. Dimers of the HhSat 1 and 2 are replicative intermediates with self-cleaving ability.

### 4.1.3 Identification of two novel putative satellite RNAs with hammerhead structures in the virome of French and Spanish carrot samples.

We characterised two new putative hammerhead-containing sc-SatRNAs (Hhsats) discovered in the virome of carrot red leaf virus-infected carrots in Spain and France. Two groups of contigs of 293 to 368 nt with a low level of sequence similarity to CYDV-RPV satRNA (M63666) were identified in the virome of several carrot populations from Spain. Extension of these contigs by identifying and merging overlaps between contigs and remaining unassembled reads allowed us to assemble sequences of partial multimers of two different putative small satellite RNA molecules from which unit-length monomeric sequences were derived. To confirm the presence of the two satellites in sampled carrots, a two-step RT-PCR assay was performed using specific detection primers designed using the genome sequence of each of the satellites: CtRLV Sat1\_F (5'-ACAGAAAACCAACCCGAGTAA-3') and CtRLV Sat1\_R (5'-TAACCACATGGGAGTCATCCT-3'), and CtRLV Sat2\_F (5'-CCACCACACTCGTTTTGTG-3') and CtRLV Sat2\_R (5'-TCCAATTCTTCCTCGATTGAG-3'). Both primer pairs generate a 258-nt amplicon. CtRLV Sat1 was present in all six sampled carrot populations, whereas CtRLV Sat2 was found in only four cultivated carrot populations. Sanger sequencing of the amplicons for the two satellites showed them to be identical to the respective sequences obtained by assembly of Illumina reads. In addition, the presence of circular forms of these two molecules was verified by RT-PCR using pairs of specific primers: CtRLV Sat1-R1 (5'-GTCTCCTCACTTCAAAGAGTG-3') and CtRLV Sat1-F1 (5'-GCTTTACGTGTCTGTCATCAA-3'), and CtRLV Sat2-R1 (5'-TACCTCGACTGATGAGTTCAA-3') and CtRLV Sat2-F1 (5'-GCACCTCGAGACACCTTTCCT-3').

The smaller of the two molecules, which is referred to as CtRLV Sat1, is 298 nt, while the longer one, referred to as CtRLV Sat2, is 368 nt. These sequences have been deposited in the GenBank database under the accession numbers OM962993 and OM962994, respectively (Table A1, Annexes). The two molecules share 71% sequence similarity and have 84.6% (CtRLV Sat1) and 69.4% (CtRLV Sat2) similarity with CYDV-RPV satRNA. In particular, they share two regions of high sequence similarity corresponding to the conserved hammerhead ribozymes on the plus and

minus strands of their genome (corresponding to positions 1-77 and 206-260 of the CtRLV Sat2 molecule). In these regions, CtRLV Sat1 and CtRLV Sat2 are 96% and 92.3% identical and share 93.3-96% and 87.3-90.9% identity, respectively, with the corresponding regions of CYDV-RPV satRNA (Figure 4.4 and 4.5). Given the high level of sequence similarity, the predicted folding for the plus- and minus-strand ribozymes is similar to that determined for CYDV-RPV satRNA by Miller and Silver 1991. In both hammerheads, nucleotide substitutions in hairpin regions with respect to the hammerhead ribozymes of CYDV-RPV satRNA have always compensatory changes on the other hairpin strand – bases 196 and 203, 194 and 229, and 195 and 228 for the minus-strand ribozyme (Figure 4.4) and bases 1 and 16, 2 and 15, and 78 and 296 for the plus strand ribozyme (Figure 4.5) – providing strong support for the secondary structure proposed for CYDV-RPV satRNA. All other nucleotide changes as compared to CYDV-RPV satRNA hammerheads occur in loop regions (position 222 for the minus-strand ribozyme, position 21 for the plus-strand ribozyme), with the exception of mutations at positions 287-289, 291-290 and 84-85 that shorten the lateral stem 3 (naming according to Miller and Silver 1991). In the same fashion, the double hammerhead structure that has been proposed (Song et al. 1999) to be the active form for the processing of the plus-strand multimers, shows compensatory mutations so that the potential to form stem 1 is fully conserved (not shown).

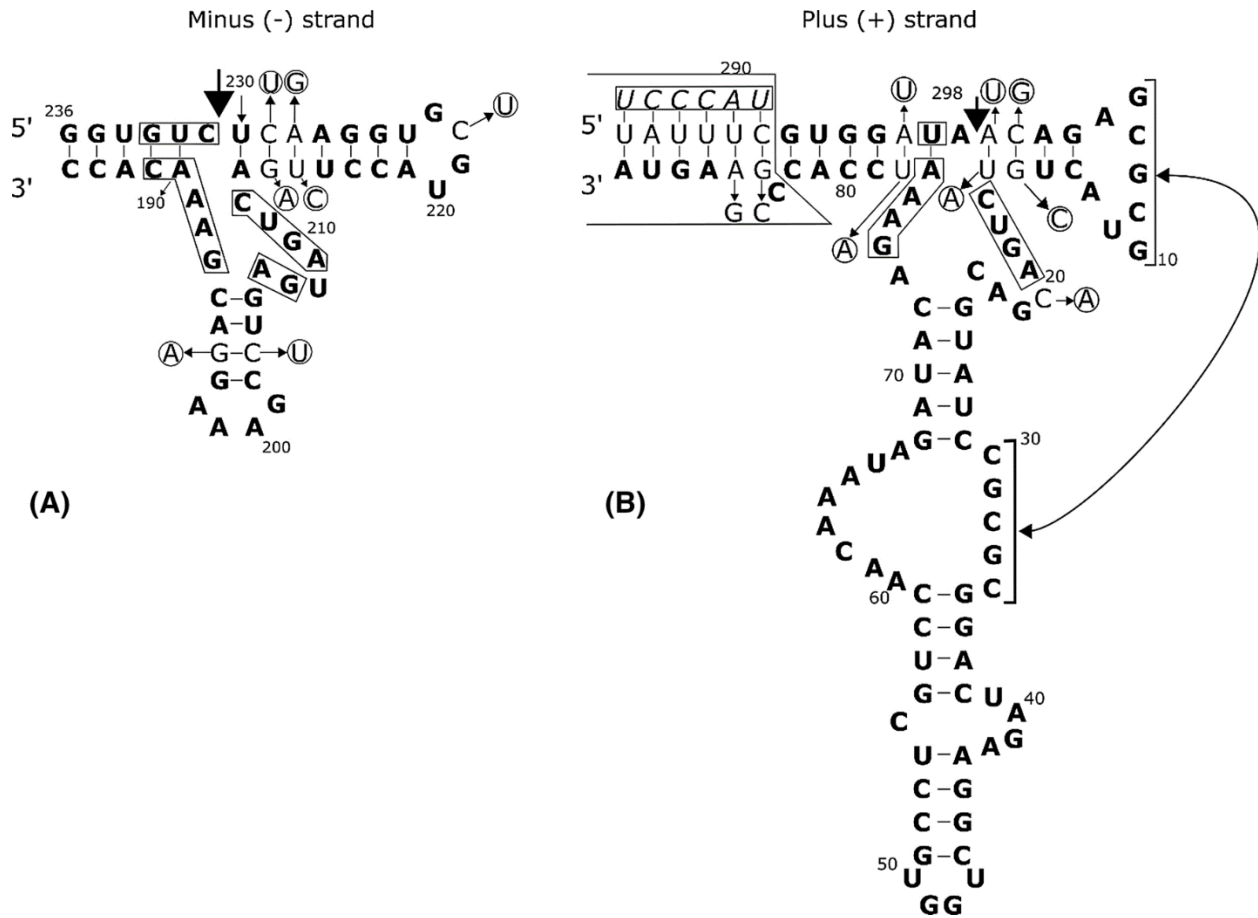


**Figure 4.4: Proposed secondary structures for the minus- and plus-strand hammerhead ribozymes of carrot red leaf virus satellite 1 (CtRLV Sat1).** The proposed structure is based on the ribozymes of CYDV-RPV satellite RNA. Nucleotide numbering is based on the full-length of CtRLV Sat1 (298 nt). (A) Minus-strand ribozyme. Bases are numbered according to the complementary plus strand. (B) Plus-strand ribozyme. Square brackets joined by a long double arrow indicate a pseudoknot structure that is also found in CYDV-RPV satRNA. In both A and B, bases enclosed in boxes are conserved among satRNA hammerhead ribozymes while large bold arrows indicate the cleavage sites. Nucleotides in bold are conserved when compared to CYDV-RPV satRNA, while small arrows and circled letters indicate mutations as compared to CYDV-RPV satRNA.

Remarkably, the region 245-310 of CYDV-RPV satRNA (Figure 4.5) that is essential for replication and has been proposed to be a replicase binding site or an origin of replication (Song and Miller 2004) is the most divergent region among the different satellites, showing both large length variation (66, 50, and 107 nt for CYDV-RPV satRNA, CtRLV Sat1, and CtRLV Sat2, respectively) and low sequence similarity (only 34% between CtRLV Sat1 and CtRLV Sat2 and 68.1% and 43.9% between CYDV-RPV satRNA and Sat1 and Sat2, respectively) (Figure 4.5). Similarly, the CYDV-RPV

satRNA 152-194 region (Figure 4.5), which has been suggested to be involved in encapsidation (Song and Miller 2004) is poorly conserved in the two carrot satellites, again showing both length variation (43, 35, and 31 nt for CYDV-RPV satRNA, CtRLV Sat1, and CtRLV Sat2, respectively) and sequence divergence (only 54.8% similarity between CtRLV Sat1 and CtRLV Sat2 and 68.6% and 58.1% between CYDV-RPV satRNA and CtRLV Sat1 and CtRLV Sat2, respectively). Taken together, the secondary structure analysis of CtRLV Sat1 (Figure 4.6) and CtRLV Sat2 (Figure 4.7) revealed three kinds of functional elements present in CYDV-RPV satRNA: (i) self-functioning ribozymes, (ii) elements that control conformational change (RNA switches, e.g., L1 and L2a, which facilitate replication), and (iii) cis-acting elements that interact with helper virus and/or host components, e.g., origins of replication and assembly (Song and Miller 2004).

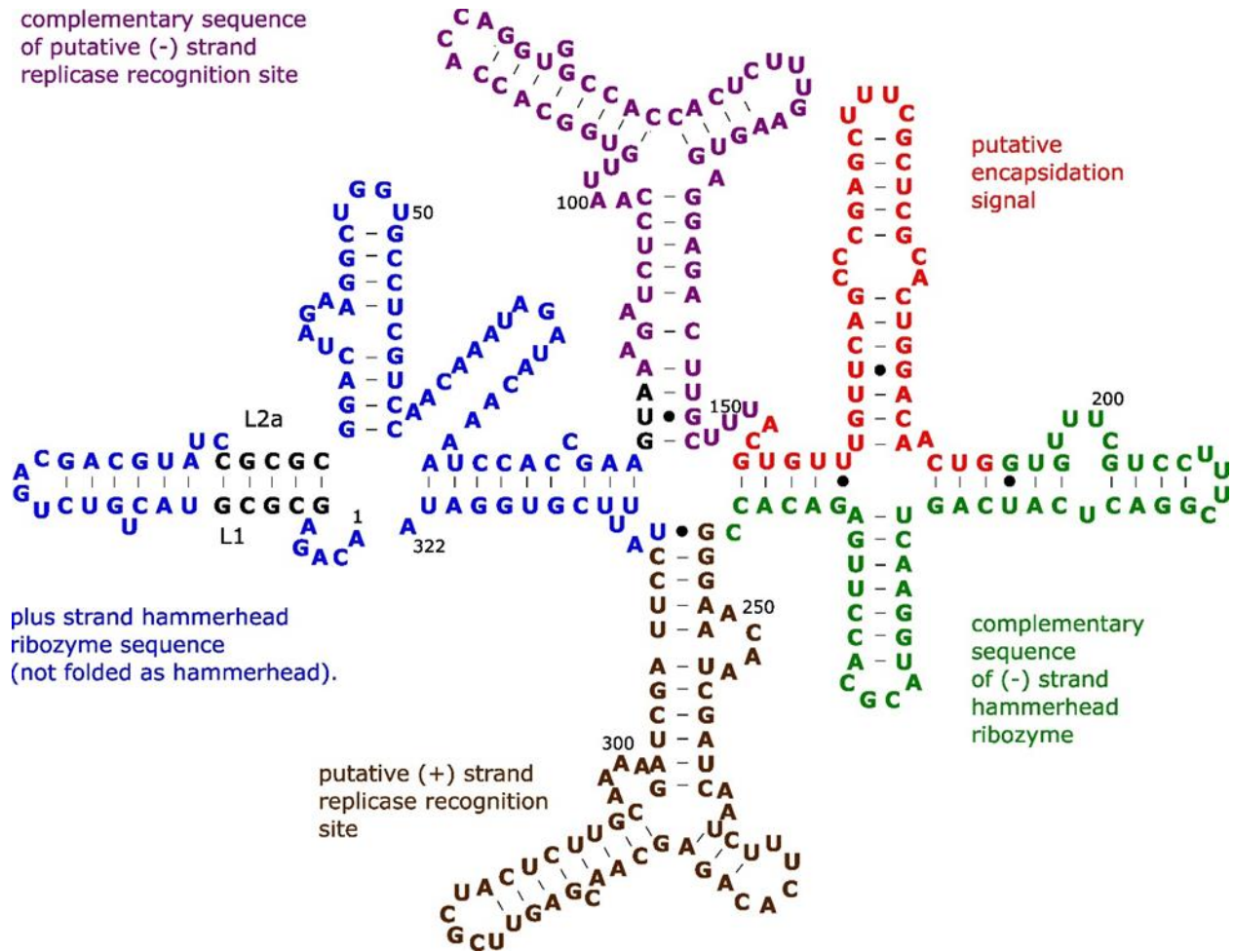


**Figure 4.5: Known and putative functional domains in the most stable secondary structure of the plus strand of CYDV-RPV satRNA adapted from Song and Miller, 2004.** Structural domains are color coded for each known or proposed function as indicated, with the L1-L2a bases shown in black.

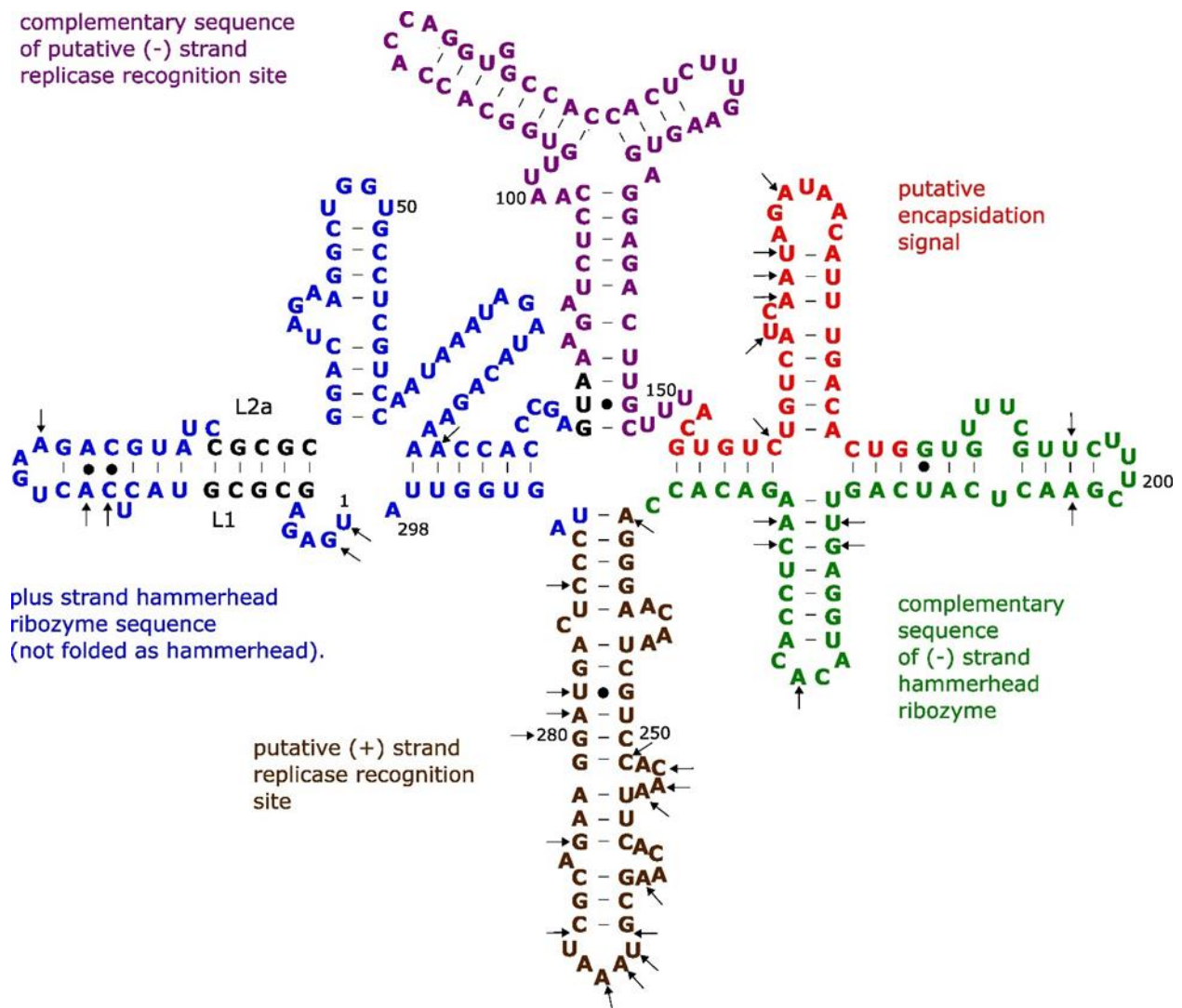


**Figure 4.6: Known and putative functional domains in the most stable secondary structure of the plus strand of CtRLV Sat1.** Structural domains are color coded for each known or proposed function as indicated, with the L1-L2a bases shown in black. Arrows indicate nucleotide differences in Sat1 and Sat2 with respect to CYDV-RPV satRNA.

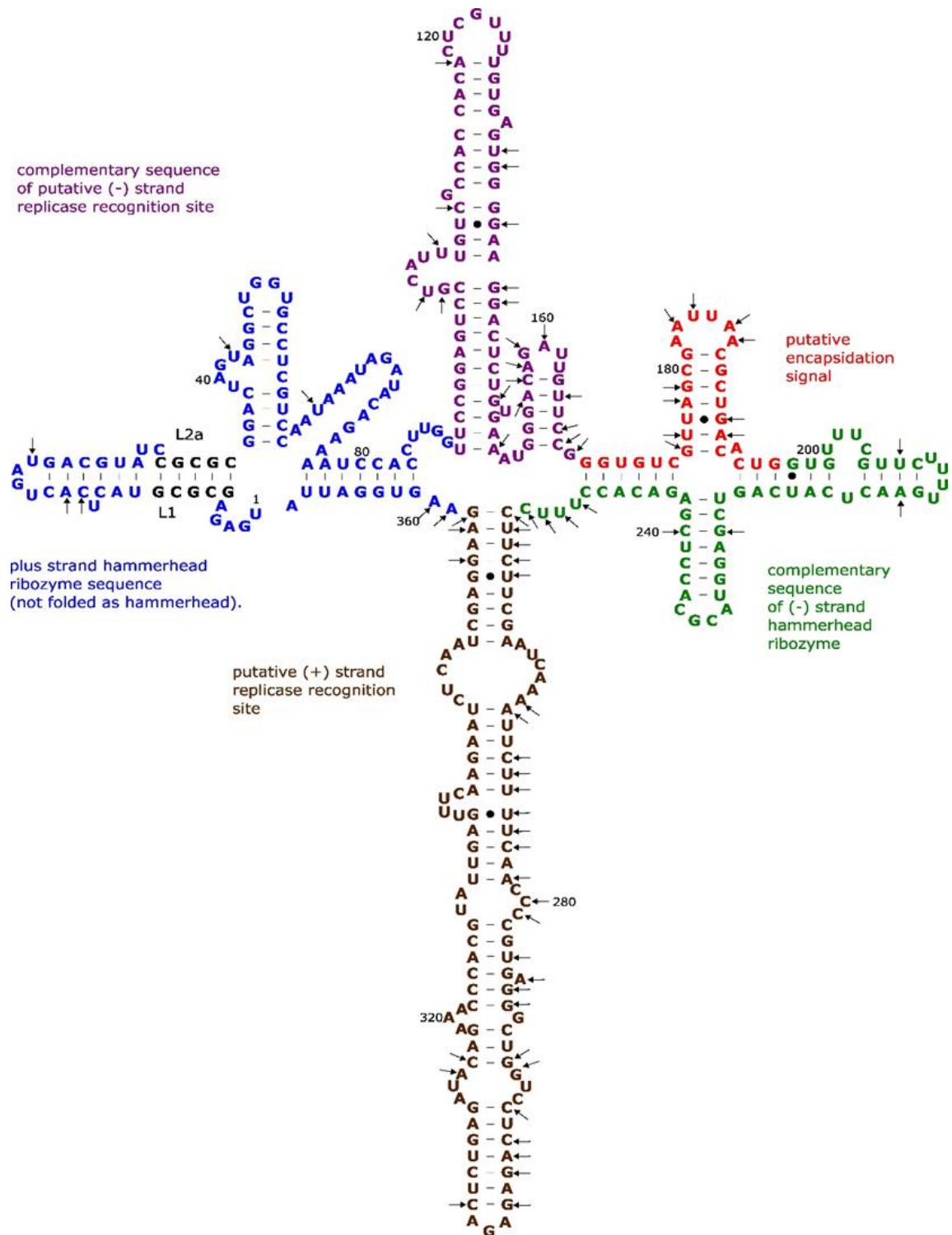


**Figure 4.7: Known and putative functional domains in the most stable secondary structure of the plus strand of CtRLV Sat2.** Structural domains are color coded for each known or proposed function as indicated, with the L1-L2a bases shown in black. Arrows indicate nucleotide differences in Sat1 and Sat2 with respect to CYDV-RPV satRNA

As judged by the number of HTS reads corresponding to each molecule, the shorter CtRLV Sat1 consistently showed much higher number of reads than the longer CtRLV Sat2. For example, in the third cultivated carrot population, CtRLV Sat1 was represented by 170,232 reads (70,155x coverage, 0.8% of total reads for the sample), while Sat2 only by 2,218 reads (739x coverage, 0.01% of total reads), which corresponds to a 95-fold higher representation for CtRLV Sat1. Given that the HTS data were generated from pools of plants and were subject to other biases during library preparations, it is not known whether the two Hhsats were present together in the same plants, and therefore, whether this observed variation in representation reflects a true difference in accumulation in infected plants, a difference in prevalence of the two molecules in the sampled carrot populations, or a combination of both factors.

Full genome sequences could be assembled for CtRLV Sat1 from HTS read sequences in each of five carrot populations but only from two populations for CtRLV Sat2 due to its lower representation. The CtRLV Sat1 genome sequences obtained for the five Spanish populations were completely identical to one another. For CtRLV Sat2, a single indel polymorphism was observed between the two populations, 123UUUU126 > UUU. Analysis of HTS datasets for carrot populations collected in France in 2020 and processed in the same way as the Spanish samples indicated the presence of CtRLV Sat2, but not CtRLV Sat1, in two populations collected in Aquitaine about 20 km apart. One of these French populations was from cultivated carrots, while the second was from wild carrots. Given the 81x-193x average coverage of CtRLV Sat2 in the HTS reads for these two populations, full-length genomic sequences could be reconstructed (GenBank OM962996). The two sequences were identical and showed a single polymorphism when compared to the Spanish sequences, a region that is also polymorphic between the two Spanish isolates (123UAUU126). The variability between the carrot satellites found in France and Spain is therefore low, while there is no evidence so far for the presence of CtRLV Sat1 in France, despite a sample of 45 carrot populations representing a total of 2250 plants. Given the differential representation of CtRLV Sat1 in Spanish samples, it seems unlikely that, if present in French samples, it could have been missed during these virome studies.

## 4.2 A multi-host-multi-virus study.

Following on from the study of plant-virus ecology in a crop and its wild relative, a multi-host virome approach was also considered. Such study of interactions between multiple hosts and viruses of species-rich communities is made practical by HTS technologies and will improve current understanding on virus host range and the influence of heterogeneous environments in shaping virus-host interactions and transmission dynamics.

### 4.2.1 Viral OTU detection by HTS and RT-PCR.

As described in section 3.2.1, field sampling conducted over two years (2015-2017) by past lab members produced multiple plant collections from different sites representing four habitats selected *a priori*. Plant sampling resulted in 78 individual collections from 23 sites, which comprised a total of 329 plant species distributed over communities with distinct cover types: the major summer (melon and maize) and winter (brassicas and barley) crops (Crop), the borders between crops (Edge), the native xerophilous forests dominated by *Quercus* spp. (Oakwood), and the successional scrubland (Wasteland), representing four habitat categories. Observed plant species richness was lowest in Crop (56 species), highest in Wasteland and Oakwood (96 and 86 species, respectively), and intermediate in Edge (73 species) (McLeish et al. 2022).

Total RNA extracted from plant tissues of the field samples was used for library preparation and subsequent sequencing. The tissue homogenisation methods and the commercial kits which together were used to optimise total RNA extracts with better concentration, purity, and integrity (see section 3.2.2), were selected to perform the RNA extraction. RNA extracts from individual samples were pooled in equimolar amounts to obtain single HTS libraries for subsequent sequencing. HTS libraries comprised extracts from samples of the same plant species, collection and site, and included RNA extracts from one or more individual field samples depending on the plant species and collection. The pooled samples used to produce the HTS libraries were prepared from rRNA-depleted RNA extracts, yielding a total of 332 libraries (122 plant species). Detection of an OTU was considered credible if a read met the criteria of having a BLAST query coverage of

100% and a query length greater or equal to 125 nucleotides (HTS2C criteria). BLAST queries of the 332 HTS libraries matched 160 OTUs reference genome (McLeish et al. 2024). Overall, the metavirome was dominated by positive-sense ssRNA viruses [(+)ssRNA, 128 OTUs], with the rest comprised of negative-sense ssRNA [(-)ssRNA, 5 OTUs], dsRNA (2 OTUs), ssDNA (16 OTUs), and reverse-transcribing double-stranded DNA viruses (9 OTUs). The field sampling, sample processing, RNA extraction and HTS of plant libraries described above were performed by Adrian Pelaez and other past lab members prior to the start of this thesis.

Of the predominant (+)ssRNA virus OTUs present, viruses that had a large host range, were detected in more than one habitat, had different genome organisation, gene expression strategy, and different transmission modes, were selected to analyse their host range and host use to address the objectives of this thesis. This study prioritised the selection of viruses transmitted by aphids in a non-persistent manner, which are the most abundant and major plant infecting RNA viruses (Agrios 2005; Hull, 2014, ICTV 2017), and for comparison, viruses from genera with other mechanisms of horizontal transmission were included. The selected OTUs belonged to the genera *Cucumovirus* (2 OTUs), *Potyvirus* (3 OTUs), *Anulavirus* (1 OTU), *Polerovirus* (1 OTU), and *Tobamovirus* (1 OTUs). Thus 5 OTUs were transmitted by aphids non-persistently, with either monopartite (*Potyvirus*) or a tripartite (*Cucumovirus*) genome organisation. One OTU (*Anulavirus*) shared the *Cucumovirus* genome organisation but was transmitted by pollen, and 1 OTU from each genus was transmitted by contact (*Tobamovirus*), or by aphids in a persistent, non-propagative manner (*Polerovirus*).

Under the HTS2C criteria, reads matched reference genomes of *cucumber mosaic virus* (CMV) that were detected from 194 libraries (88 host species), and reads from 61 libraries (38 host species) matched reference genomes of *tomato aspermy virus* (TAV), in genus *Cucumovirus*. Reads from 12, 23 and 59 libraries (11, 20 and 37 host species) matched with *plum pox virus* (PPV), *turnip mosaic virus* (TuMV), and *watermelon mosaic virus* (WMV) respectively, in the genus *Potyvirus*. Reads of 164 libraries (67 host species) matched with *pelargonium zonate spot virus* (PZSV) in the *Anulavirus* genus. *Beet chlorosis virus* (BChV, genus *Polerovirus*), was detected in 42 libraries (32 host species) while *youcai mosaic virus* (YoMV, genus *Tobamovirus*), in 48 libraries (37 host species). When the HTS6C criteria was applied, the number of libraries in which an OTU

was detected was reduced as follows: CMV, from 194 to 191; TAV, from 61 to 52; PZSV from 164 to 134, PPV from 12 to 9; TuMV from 23 to 17, BChV from 42 to 26; YoMV from 48 to 27, while that of WMV remained unchanged (Figure 4.8). Correspondence between the number of libraries in which a virus OTU was detected per HTS2C or HTS6C (Figure 4.8) showed strong positive association (Kendall's rank correlation  $z = 3.217$ ,  $\tau = 0.930$ ,  $p\text{-value} = 0.0003$ ), indicating that although the most stringent criteria (HTS6C) eliminates a proportion of detections, it did not affect the relative host range breadths.

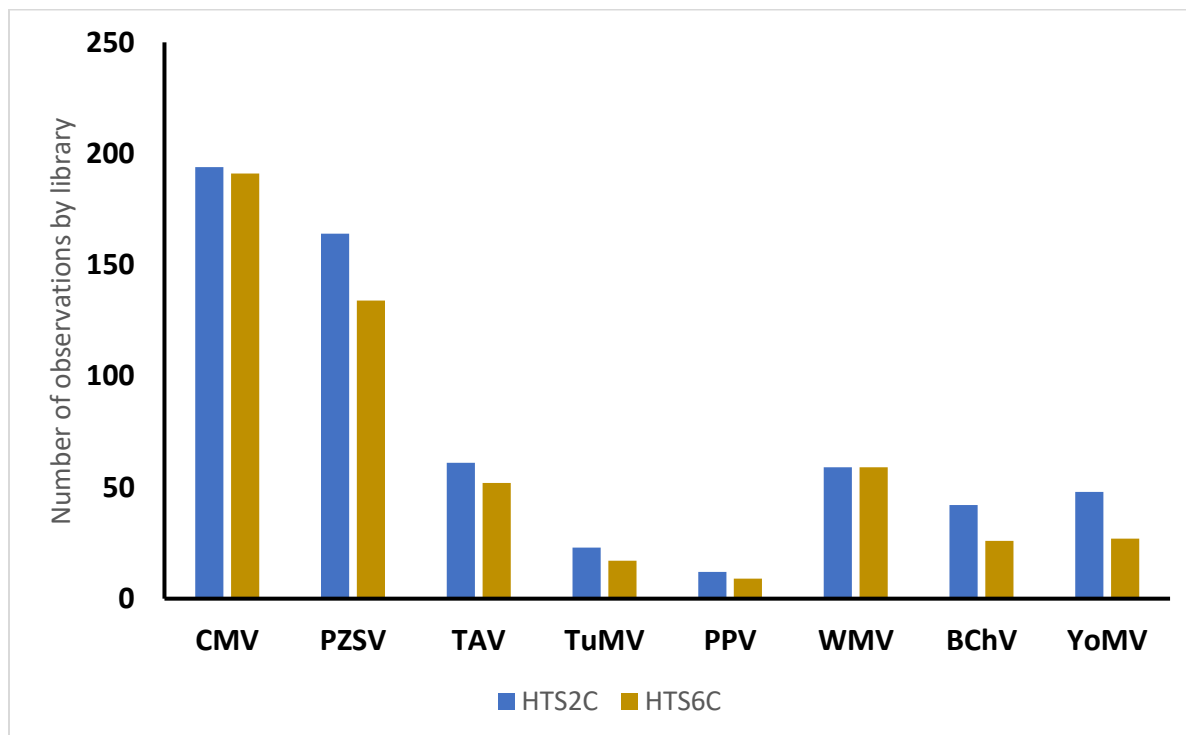


**Figure 4.8: Correspondence between number of libraries per OTU as detected by HTS2C and HTS6C criteria.**

In most libraries, the number of reads detected by the HTS2C and HTS6C criteria (Table A3, Annexes) were either the same e.g libraries PV495, PV047 or PV199 for CMV; libraries PV163, PV172, or PV173 for WMV, etc, or the reads detected by HTS6C was reduced compared to HTS2C e.g libraries PV115, PV119 or PV120 for TAV; libraries PV199, PV208, PV207 for PZSV, etc. In few libraries, some reads were only detected by HTS2C and not in HTS6C, that is libraries marked with absent '(abs)' such as PV069, PV099 or PV141 for TuMV; while others were only detected by

HTS6C and not by HTS2C that is, libraries marked by present '(p)' such as PV505, PV538 or PV556 for BChV.

As HTS detection of OTUs involves uncertainty from a number of factors that include the specificity of the detection pipeline or the proportion of false positives produced, validation of HTS detections was done by RT-PCR using a random subset of 50% of the OTU-positive libraries detected by HTS2C, with the condition that at least 50% of the libraries in which a certain OTU was detected by HTS2C of each habitat was represented. RT-PCR detection was performed with specific primer pairs for each OTU (See Table 3.3) that amplified regions of 70-160 nucleotides in their coat protein gene. Primers were designed with consensus sequences for a given taxon and RT-PCR detections were carried out on extracts of the individual plant samples that were pooled to make up each library. Table 4.3 shows a generally good correspondence between detection by HTS2C and detection by RT-PCR, with the notable exception of those in which BChV was detected by HTS. Correlation between the number of libraries in which an OTU was detected by HTS2C and RT-PCR was positive (Kendall's rank correlation  $z = 2.969$ ,  $\tau = 0.857$ ,  $p\text{-value} = 0.001$ ), showing a linear relationship between both methods. Table 4.3 indicates that the sensitivity of both detection methods was different according to OTU and habitat. It was interesting to confirm that RT-PCR was able to amplify OTUs with low number of reads (> 5 reads) detected by HTS2C and HTS6C e.g in Table A3 Annexes, see library PV386 for CMV, PV128 for TAV, PV142 for TuMV, PV028 for PPV, PV163 for WMV, PV390 for PZSV, or PV234 for BChV. In other words, irrespective of low or high read number of these libraries, RT-PCR results validated positive detections of these OTUs.

**Table 4.3: Comparison of OTU detections across habitats with RT-PCR and HTS2C (RT-PCR/HTS2C).** For each OTU, the number of libraries detected by RT-PCR/HTS2C is indicated according to habitat, while total number of libraries analysed is shown per OTU and per habitat.

| <b>Virus</b>   | <b>Crop</b> | <b>Edge</b> | <b>Oakwood</b> | <b>Wasteland</b> | <b>Total per OTU</b> |
|----------------|-------------|-------------|----------------|------------------|----------------------|
| CMV            | 10 /10      | 28 /30      | 22 /23         | 13 /15           | 73 /78 (0.936)       |
| PZSV           | 2 /2        | 46 /54      | 2 /6           | 8 /8             | 58 /70 (0.829)       |
| TAV            | 5 /6        | 40 /45      | 7 /8           | 1 /1             | 53 /60 (0.883)       |
| TuMV           | 0/2         | 9 /9        | 3 /3           | 7 /7             | 19 /21 (0.905)       |
| PPV            | 2 /2        | 7 /7        | 2 /2           | 1 /1             | 12 /12 (0.100)       |
| WMV            | 5 /5        | 11 /13      | 9 /9           | 2 /2             | 27 /29 (0.931)       |
| BChV           | 2 /3        | 11 /21      | 0/3            | 0/0              | 11 /27 (0.407)       |
| YoMV           | 2 /3        | 5 /8        | 11 /14         | 16 /20           | 34 /44 (0.773)       |
| <b>Total</b>   | 28 /33      | 157 /187    | 56 /68         | 48 /54           |                      |
| <b>per</b>     | (0.848)     | (0.840)     | (0.824)        | (0.889)          |                      |
| <b>habitat</b> |             |             |                |                  |                      |

Figure 4.9 further illustrates the comparison between by HTS2C and RT-PCR detections and showed that detections of most OTUs were in a higher number of libraries by HTS2C than by RT-PCR except for PPV in which it was the same. The 2 hosts of TuMV present in Crop and 3 hosts of BChV present in Oakwood were not confirmed by RT-PCR, but hosts for TuMV and BChV present in other habitats were similar with both detection methods. Also, across habitats, the number of hosts of PZSV was higher than TAV with HTS2C but was the same with RT-PCR. In most cases detection by RT-PCR was less sensitive and inconsistencies between methods could be attributed

to experimental factors, e.g., low titre viruses, genetic variants, or high variation in the quality of the RNA extractions; a lower sensitivity in the virus detection is expected with RT-PCR detections. In spite of these factors, we can consider a significant correspondence between HTS2C and RT-PCR detections.

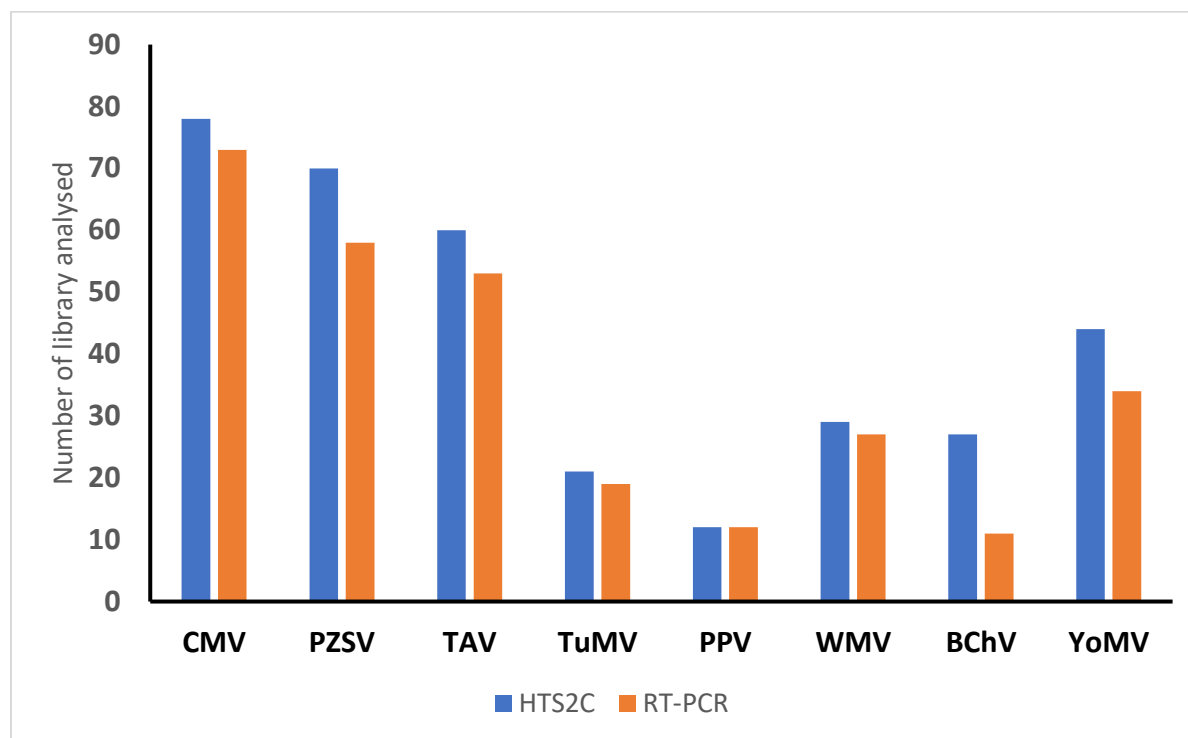


**Figure 4.9: Comparison between number of libraries where an OTU was detected by HTS2C and RT-PCR.**

For each OTU, a negative test of the specificity of RT-PCR detection was done with four to five libraries that were not positive by HTS2C. For the libraries that were positive by HTS2C and validated by RT-PCR, sanger sequencing of RT-PCR amplicons from five libraries was done to further verify the identity of the OTUs. Figure 4.10 shows primer specificity test results for different OTUs. As expected, there was no amplification with the HTS2C negative libraries for each OTU, while Table A2 shows the percent identity of nucleotide sequences of amplicons from the positive libraries, which confirmed that the sequences matched the respective OTUs from which they were derived.

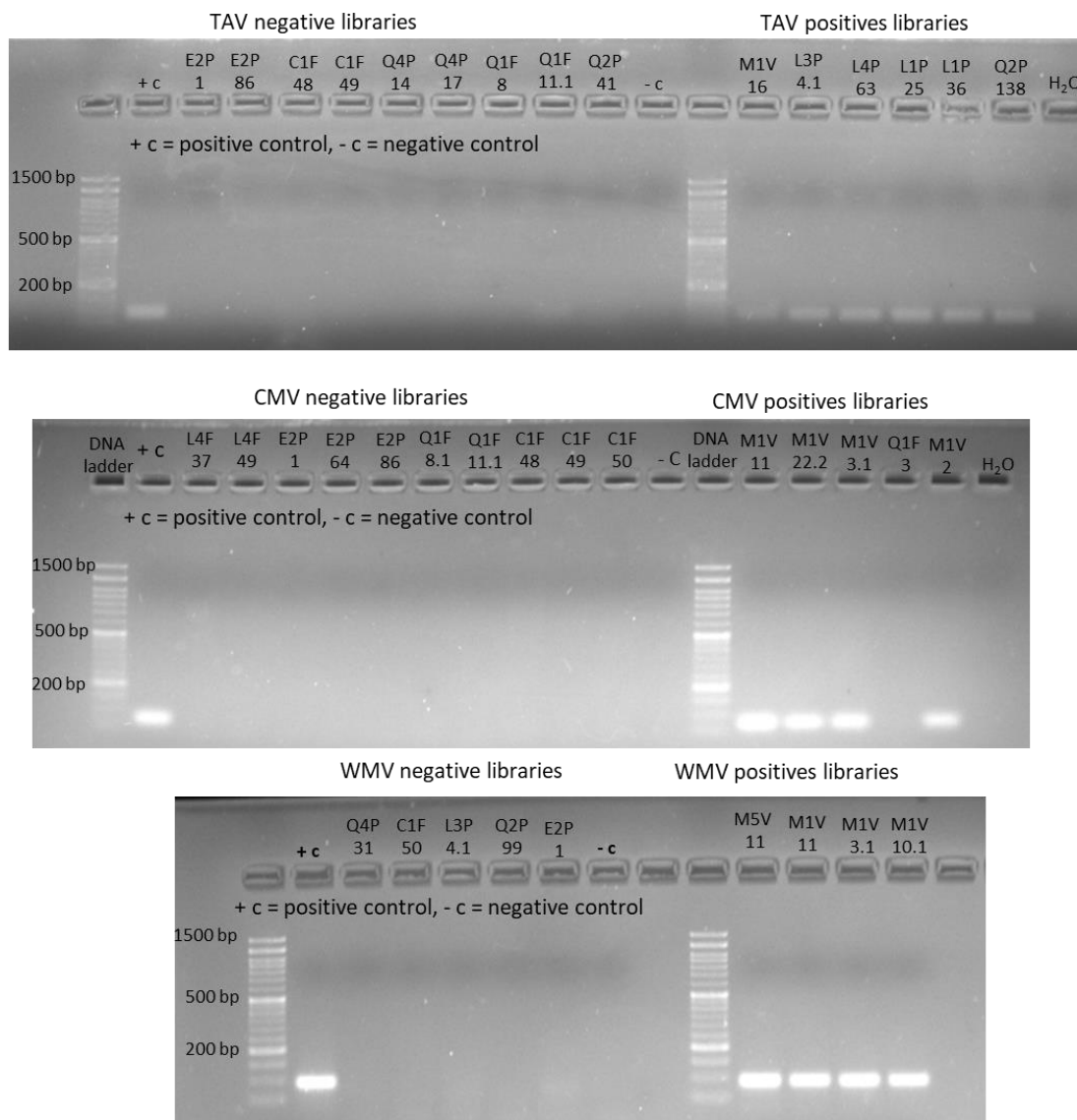


**Figure 4.10: Primer specificity test with CMV, WMV and TAV using HTS2C negative and positive libraries from four to five different plant species.** The name of the individual plant species and the library name are indicated. +c and -c represent positive and negative PCR controls, while H<sub>2</sub>O is distilled water used as negative control during cDNA synthesis.

The distribution of host species across habitats was also compared by HTS2C and RT-PCR in Figures 4.11 and 4.12 respectively. As an example, for PZSV, HTS2C detected 12 hosts each in Oakwood and Wasteland, but these were reduced to 2 and 8 respectively by RT-PCR. Despite the reduction in number of host species, both methods showed similar patterns of host range

distribution: CMV had the highest host range and PPV had the lowest host range, while TAV, WMV, BChV and YoMV showed intermediate host ranges.

The number of host species of each virus differed by habitat. For most viruses, the host species was broadest in Edge. The exception was YoMV, which had its broadest host range in the wild habitats (Oakwood and Wasteland). Across viruses, host range was narrowest in Crop, which is the habitat with lowest plant species richness. As Edge is not the habitat with the higher plant species richness (McLeish et al. 2021), it indicates that host range is determined not only by the number of potential hosts that a virus can encounter.

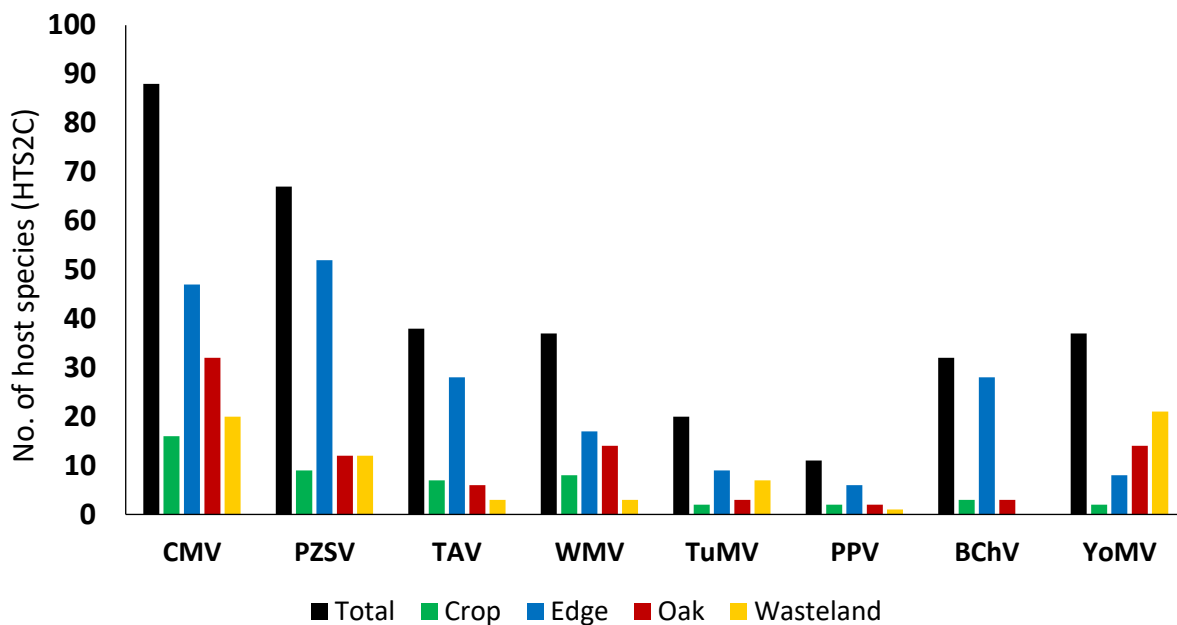


**Figure 4.11: Host range of eight OTUs in each habitat by HTS2C detection.** Black bars represent the total number of unique host species per OTU across habitats, while green, blue, red and orange bars represent number of host species in each habitat of Crop, Edge, Oakwood and Wasteland respectively.

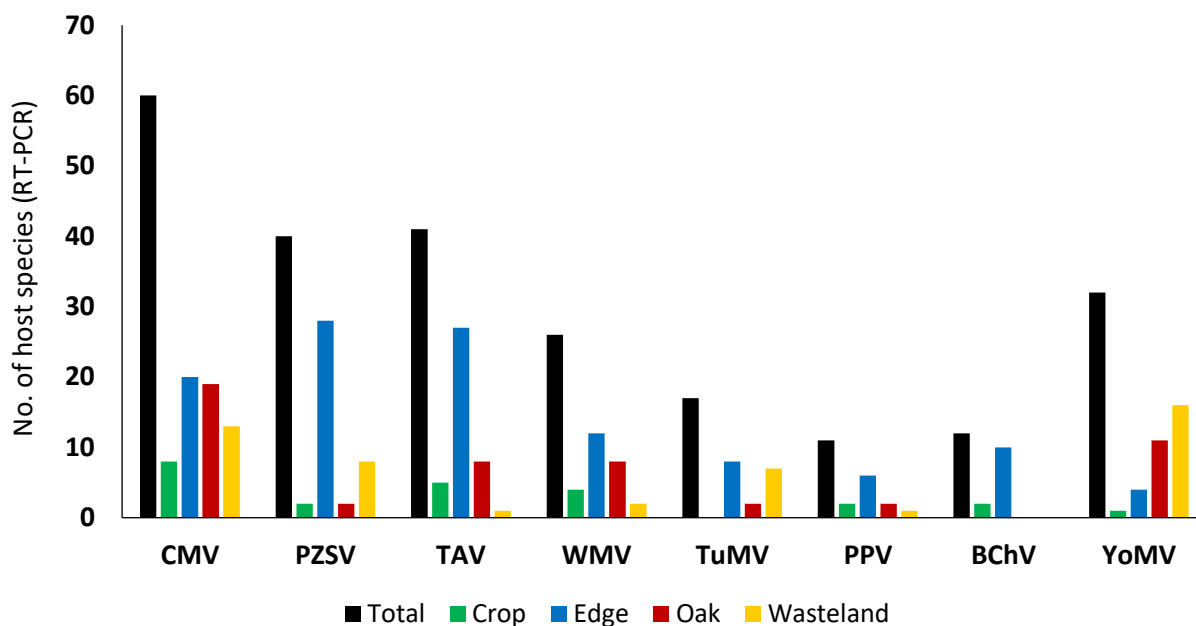


**Figure 4.12: Host range of eight OTUs in each habitat by RT-PCR detection.** Black bars represent total number of unique host species per OTU across habitats, while green, blue, red and orange bars represent number of host species in each habitat of Crop, Edge, Oakwood and Wasteland respectively.

#### 4.2.2 Host range by HTS2C detection.

Host range, defined as the number of host species used or exploited by a pathogen (García-Arenal and Fraile, 2013; McLeish et al. 2018), was determined as the number of plant species with samples that were positive by HTS2C. According to this criterion, CMV showed the largest host range (88 host species of 29 plant families), followed by PZSV (67 host species of 21 families), TAV (38 host species of 17 families), YoMV (37 host species of 13 families), WMV (37 species of 20 families), BChV (32 species of 15 families), TuMV (20 species of 15 families) and PPV (11 species of 6 families) (Tables 4.4 and 4.5). The plant families with highest host species richness were eudicots such as *Asteraceae* as well as monocots e.g. *Poaceae*. Other plant families had a lower number of host species for fewer viruses, for instance *Aristolochiaceae* had only one host species, where TAV was detected, *Cyperaceae* were only associated with CMV and WMV, while only TuMV and YoMV were found in species of *Scrophulariaceae*.

After the identity of the detected OTUs were validated by RT-PCR and Sanger sequencing of the amplicons, the OTUs are referred to as virus species from here onward.

Some host species were represented by a relatively higher number of libraries. For example, *Convolvulus arvensis* (13 libraries), *Diplotaxis eruroides* (10 libraries), while *Carduus bourgeanus*, *Cucumis melo*, *Picris echioides*, *Rubia peregrina* and *Silybum marianum* had at least 7 or 8 libraries each, showing that these plant species were abundant in the ecosystem, as the sampling procedure represents the species composition of the plant communities. These host species were present in at least 3 of 4 habitats. Abundant host species were hosts of more viruses, in fact, while *C. arvensis* and *D. eruroides* were hosts to all 8 viruses, *C. bourgeanus* was host to 7 viruses while *R. peregrina* was host to 7 viruses.

Less abundant host species with lower library number (that is 4 to 6 libraries) were hosts to at least 6 viruses and present in at least 2 or more habitats. For example, *Bromus sp.* (6 libraries) was host to 6 viruses and was infected by CMV in all habitats, while the other 5 viruses were detected in this host in 2 of 4 habitats. *Galium verum* (4 libraries) was host to 4 viruses and present in 3 of 4 habitats showing that host range is conditioned by species abundance and distribution across the ecosystem.

Therefore, while some hosts are abundant and present in various habitats, host species relative abundance does not drive virus richness at the level of habitat. Instead, the results reflect a pattern of facultative generalism, where viruses exploit available hosts opportunistically rather than being restricted to the relative abundance of specific hosts within each habitat.

**Table 4.4: Observed host range of three *Bromoviridae* species (CMV, TAV and PZSV) and one Tobamovirus specie (YoMV). OTUs**  
were detected by HTS2C and validated by RT-PCR.

| Host species                    | No. of<br>libraries<br>(HTS2C) | CMV |    |    |    | TAV |    |    |    | PZSV |    |    |    | YoMV |    |    |    |
|---------------------------------|--------------------------------|-----|----|----|----|-----|----|----|----|------|----|----|----|------|----|----|----|
|                                 |                                | CR  | ED | OA | WL | CR  | ED | OA | WL | CR   | ED | OA | WL | CR   | ED | OA | WL |
| <i>Allium sativum</i>           | 1                              | +   |    |    |    |     |    |    |    | +    |    |    |    |      |    |    |    |
| <i>Amaranthus sp.</i>           | 4                              | +   | +  |    |    | +   | +  |    |    | +    |    |    |    |      |    |    |    |
| <i>Anacyclus clavatus</i>       | 3                              | +   | +  |    |    |     | +  |    |    | +    |    | +  |    |      |    |    | +  |
| <i>Anchusa undulata</i>         | 3                              |     | +  |    | +  |     |    |    |    | +    |    |    |    |      |    |    | +  |
| <i>Andryala arenaria</i>        | 1                              |     |    |    |    |     |    |    |    |      |    |    |    |      |    | +  |    |
| <i>Anthriscus caucalis</i>      | 2                              |     | +  |    |    |     | +  |    |    | +    |    |    |    |      |    |    |    |
| <i>Aristolochia pistolochia</i> | 2                              |     | +  |    |    |     | +  |    |    | +    | +  |    |    |      |    |    |    |
| <i>Artemisia herba alta</i>     | 1                              |     |    |    |    |     |    |    |    |      |    |    |    |      |    |    | +  |
| <i>Asparagus acutifolius</i>    | 3                              |     |    | +  | +  |     |    |    |    |      |    | +  | +  |      |    |    |    |
| <i>Asphodelus aestivus</i>      | 1                              |     |    | +  |    |     |    |    |    |      |    |    |    |      |    |    |    |
| <i>Asteriscus aquaticus</i>     | 2                              |     |    |    |    |     |    |    |    |      |    | +  | +  |      |    | +  |    |
| <i>Astragalus incanus</i>       | 1                              |     |    | +  |    |     |    |    |    |      |    |    |    |      |    |    |    |
| <i>Avena sterilis</i>           | 1                              |     |    |    |    |     |    |    |    |      |    |    |    |      |    |    | +  |
| <i>Avenula bromoides</i>        | 1                              |     |    | +  |    |     |    |    |    |      |    |    |    |      |    |    |    |
| <i>Bassia scoparia</i>          | 1                              |     |    |    |    |     |    |    |    | +    |    |    |    |      |    |    |    |

|                                  |    |   |   |   |   |   |   |   |   |   |   |   |
|----------------------------------|----|---|---|---|---|---|---|---|---|---|---|---|
| <i>Brachypodium phoenicoides</i> | 2  |   | + |   |   | + |   |   | + |   | + |   |
| <i>Brachypodium retusum</i>      | 3  |   | + |   |   |   |   |   | + |   |   |   |
| <i>Bromus sp.</i>                | 6  | + | + | + | + |   | + | + | + |   | + | + |
| <i>Carduus bourgeanus</i>        | 8  | + | + | + |   |   | + | + | + | + |   | + |
| <i>Centaurea melitensis</i>      | 1  |   |   | + |   |   |   | + |   |   |   |   |
| <i>Centranthus calcitrapae</i>   | 1  |   |   | + |   |   |   | + |   |   |   |   |
| <i>Chenopodium album</i>         | 5  |   | + |   |   |   | + |   |   | + |   |   |
| <i>Cirsium arvense</i>           | 1  |   | + |   |   |   |   |   | + |   |   |   |
| <i>Convolvulus arvensis</i>      | 13 | + | + |   | + | + | + |   | + | + |   | + |
| <i>Conyza bonariensis</i>        | 2  |   | + |   |   |   | + |   |   | + |   |   |
| <i>Conyza canadensis</i>         | 3  |   | + |   |   |   |   |   |   | + |   |   |
| <i>Cucumis melo</i>              | 8  | + |   |   |   | + |   |   | + |   |   |   |
| <i>Cynodon dactylon</i>          | 4  |   | + |   |   |   | + |   |   | + |   |   |
| <i>Cyperus longus</i>            | 1  | + |   |   |   |   |   |   |   |   |   |   |
| <i>Dactylis glomerata</i>        | 1  |   |   | + |   |   |   |   |   |   |   |   |
| <i>Datura stramonium</i>         | 3  | + | + |   |   | + |   |   |   | + | + | + |
| <i>Daucus sp.</i>                | 3  |   | + |   |   |   |   |   |   | + |   |   |
| Unknown 4                        | 3  | + | + |   |   |   |   |   | + | + |   |   |
| Unknown 5                        | 1  |   |   |   |   |   |   |   |   | + |   |   |
| <i>Descurainia sophia</i>        | 2  |   | + |   | + |   |   | + |   | + |   | + |
| <i>Diploaxis erucoides</i>       | 10 | + | + |   | + |   | + |   | + | + |   | + |

|                             |   |   |   |   |   |   |   |   |   |   |   |
|-----------------------------|---|---|---|---|---|---|---|---|---|---|---|
| <i>Diplotaxis sp.</i>       | 1 | + |   |   | + |   | + |   | + |   | + |
| <i>Diplotaxis virgata</i>   | 1 |   | + |   |   |   |   |   |   |   |   |
| <i>Echium vulgare</i>       | 2 | + |   | + |   |   | + |   |   |   |   |
| <i>Erodium cicutarium</i>   | 3 | + | + |   |   |   | + | + | + |   | + |
| <i>Eruca vesicaria</i>      | 1 | + |   |   |   |   | + |   |   | + |   |
| <i>Eryngium campestre</i>   | 2 |   |   |   |   |   |   | + |   |   |   |
| <i>Festuca sp.</i>          | 1 |   |   |   | + |   |   |   |   |   | + |
| <i>Fumaria parviflora</i>   | 3 | + |   |   |   | + | + |   |   |   |   |
| <i>Galium verum</i>         | 4 | + | + | + |   | + |   |   | + |   |   |
| <i>Geranium sp.</i>         | 4 | + | + | + |   |   | + |   |   |   | + |
| <i>Helichrysum stoechas</i> | 1 |   | + |   |   |   |   |   |   |   |   |
| <i>Hieracium pilosella</i>  | 1 |   |   |   | + |   |   |   |   |   |   |
| <i>Hirschfeldia incana</i>  | 1 | + |   |   |   | + | + |   |   |   |   |
| <i>Hordeum matritense</i>   | 2 |   |   | + |   |   | + |   | + | + |   |
| <i>Hordeum vulgare</i>      | 2 | + |   |   | + |   | + |   |   |   |   |
| <i>Hypericum pubescens</i>  | 1 |   | + |   |   |   |   |   |   |   |   |
| <i>Jasminum fruticans</i>   | 1 |   | + |   |   |   |   |   |   |   |   |
| <i>Klasea pinnatifida</i>   | 2 |   |   |   |   |   | + |   |   |   |   |
| <i>Lactuca serriola</i>     | 6 | + |   |   |   | + |   | + |   | + |   |
| <i>Leontodon sp.</i>        | 4 | + | + | + |   | + | + |   |   |   |   |
| <i>Lepidium draba</i>       | 4 | + |   |   |   | + |   |   |   | + | + |
| <i>Lithospermum arvense</i> | 1 |   |   |   |   |   |   | + |   |   |   |

|                            |   |   |   |   |   |   |   |   |   |   |   |   |
|----------------------------|---|---|---|---|---|---|---|---|---|---|---|---|
| <i>Lolium perenne</i>      | 4 |   | + |   |   | + |   | + |   | + |   | + |
| <i>Lotus corniculatus</i>  | 1 |   |   | + |   |   |   |   |   |   |   |   |
| <i>Malva sylvestris</i>    | 3 |   | + |   |   | + |   | + |   |   |   |   |
| <i>Marrubium vulgare</i>   | 2 |   |   | + |   |   |   |   |   |   |   |   |
| <i>Medicago sp.</i>        | 1 |   |   |   | + |   |   |   | + |   | + |   |
| <i>Milium vernale</i>      | 2 |   | + |   |   |   |   | + |   |   |   |   |
| <i>Odontites luteus</i>    | 1 |   |   |   |   |   |   |   |   | + |   |   |
| <i>Origanum vulgare</i>    | 1 |   |   |   |   |   |   | + |   |   | + |   |
| <i>Papaver rhoeas</i>      | 4 |   | + |   | + | + |   | + | + |   |   |   |
| <i>Phalaris minor</i>      | 1 | + |   |   |   |   | + | + |   |   | + | + |
| <i>Phlomis lychnitis</i>   | 1 |   |   | + |   |   |   |   |   |   |   |   |
| <i>Picris echioides</i>    | 8 | + | + |   | + | + | + | + | + |   |   |   |
| <i>Plantago</i>            | 1 |   |   |   |   |   |   |   |   |   |   |   |
| Poaceae                    | 1 |   | + |   |   |   |   |   | + |   |   | + |
| <i>Portulaca oleraceae</i> | 1 | + |   |   |   | + |   |   |   |   |   |   |
| <i>Potentilla sp.</i>      | 1 |   | + |   |   |   | + |   | + |   |   |   |
| <i>Quercus coccifera</i>   | 4 |   |   | + |   |   |   |   | + |   |   |   |
| <i>Quercus ilex</i>        | 2 |   |   | + |   |   |   |   |   |   |   |   |
| <i>Reseda lutea</i>        | 1 |   |   | + |   |   |   |   |   |   | + |   |
| <i>Rubia peregrina</i>     | 8 |   | + |   |   | + |   | + | + |   |   |   |
| <i>Rumex pulcher</i>       | 3 |   | + |   |   |   |   | + |   |   | + | + |
| <i>Salvia verbenaca</i>    | 1 |   |   |   | + |   |   |   |   |   |   |   |

|                                   |   |   |   |   |   |   |   |
|-----------------------------------|---|---|---|---|---|---|---|
| <i>Scandix pecten-veneris</i>     | 1 | + |   | + | + |   | + |
| <i>Senecio jacobaea</i>           | 1 | + |   |   | + |   |   |
| <i>Silybum marianum</i>           | 7 | + |   | + | + |   |   |
| <i>Sisymbrium runcinatum</i>      | 1 |   |   |   |   | + |   |
| <i>Solanum nigrum</i>             | 2 | + |   |   | + |   |   |
| <i>Sonchus oleraceus</i>          | 2 | + |   |   | + |   | + |
| <i>Stachelina dubia</i>           | 1 |   | + |   |   |   |   |
| <i>Stipa parviflora</i>           | 2 |   | + |   |   | + | + |
| <i>Taraxacum officinale</i>       | 3 |   |   | + |   |   | + |
| <i>Teucrium botrys</i>            | 1 |   | + |   |   |   | + |
| <i>Teucrium capitatum</i>         | 2 |   | + |   |   | + |   |
| <i>Teucrium chamaedrys</i>        | 1 |   | + |   |   | + | + |
| <i>Teucrium pseudochamaepitys</i> | 3 |   | + | + |   |   |   |
| <i>Thapsia villosa</i>            | 2 |   | + |   |   | + |   |
| <i>Thymus vulgaris</i>            | 2 |   | + |   |   |   |   |
| <i>Torilis nodosa</i>             | 1 | + |   |   |   |   | + |
| <i>Trifolium campestre</i>        | 1 |   |   |   |   |   |   |
| <i>Verbascum sinuatum</i>         | 3 |   |   | + |   |   | + |
| <i>Vicia sp.</i>                  | 2 | + |   |   | + |   |   |
| <i>Xanthium strumarium</i>        | 1 | + |   |   |   |   |   |
| <i>Zea mays</i>                   | 1 | + |   |   |   |   |   |

|                                     |  |    |    |    |    |    |    |   |   |    |    |    |    |    |   |    |    |
|-------------------------------------|--|----|----|----|----|----|----|---|---|----|----|----|----|----|---|----|----|
|                                     |  |    |    |    |    |    |    |   |   |    |    |    |    |    |   |    |    |
| Host range by habitat<br>(HTS2C)    |  | 16 | 47 | 32 | 20 | 7  | 28 | 6 | 3 | 9  | 52 | 12 | 12 | 2  | 8 | 14 | 21 |
| Host range by ecosystem<br>(HTS2C)  |  | 88 |    |    |    | 38 |    |   |   | 67 |    |    |    | 37 |   |    |    |
| Host range by habitat<br>(RT-PCR)   |  | 8  | 20 | 19 | 12 | 4  | 25 | 6 | 1 | 2  | 26 | 2  | 8  | 1  | 4 | 11 | 16 |
| Host range by ecosystem<br>(RT-PCR) |  | 50 |    |    |    | 31 |    |   |   | 35 |    |    |    | 27 |   |    |    |

'+' detected by HTS2C only; empty grey cell detected by RT-PCR only, both '+' and grey cell detected by HTS2C and RT-PCR

**Table 4.5: Observed host range of three *Potyvirus* species (TuMV, PPV and WMV) and one *Polerovirus* specie (BChV). OTUs were detected by HTS2C and RT-PCR.**

| Host species                      | No. of libraries (HTS2C) | TuMV |    |    |    | PPV |    |    |    | WMV |    |    |    | BChV |    |    |    |
|-----------------------------------|--------------------------|------|----|----|----|-----|----|----|----|-----|----|----|----|------|----|----|----|
|                                   |                          | CR   | ED | OA | WL | CR  | ED | OA | WL | CR  | ED | OA | WL | CR   | ED | OA | WL |
| <i>Amaranthus sp</i>              | 3                        |      |    |    |    |     |    |    |    |     | +  |    |    |      | +  |    |    |
| <i>Anchusa undulata</i>           | 2                        |      |    |    | +  |     |    |    |    |     |    |    | +  |      |    |    |    |
| <i>Anthriscus caucalis</i>        | 1                        |      |    |    |    |     |    |    |    |     |    |    |    |      | +  |    |    |
| <i>Aphyllanthes monspeliensis</i> | 1                        |      |    |    |    |     |    | +  |    |     |    |    |    |      |    |    |    |
| <i>Artemisia herba alba</i>       | 2                        |      |    |    |    |     |    | +  |    |     |    | +  |    |      |    |    |    |
| <i>Asphodelus aestivus</i>        | 1                        |      |    |    |    |     |    |    |    |     |    | +  |    |      |    |    |    |
| <i>Asteriscus aquaticus</i>       | 1                        |      |    |    |    |     |    |    | +  |     |    |    |    |      |    |    |    |
| <i>Brachypodium phoenicoides</i>  | 2                        |      |    |    |    |     | +  |    |    |     |    |    |    |      | +  |    |    |
| <i>Brachypodium retusum</i>       | 3                        |      |    |    |    |     | +  |    |    |     | +  |    |    |      |    |    |    |
| <i>Bromus sp</i>                  | 2                        |      | +  |    |    |     |    |    |    | +   |    |    |    |      |    |    |    |
| <i>Carduus bourgeanus</i>         | 4                        |      |    |    |    | +   |    |    |    |     | +  | +  |    |      | +  |    |    |
| <i>Centranthus calcitrapae</i>    | 2                        |      |    |    | +  |     |    |    |    |     |    | +  |    |      |    |    |    |
| <i>Chenopodium album</i>          | 4                        |      | +  |    |    |     |    |    |    | +   | +  |    |    |      | +  |    |    |
| <i>Cirsium arvense</i>            | 1                        |      |    |    |    |     |    |    |    |     |    |    |    |      | +  |    |    |

|                             |   |   |   |   |   |   |   |   |   |
|-----------------------------|---|---|---|---|---|---|---|---|---|
| <i>Convolvulus arvensis</i> | 7 |   | + |   | + | + | + |   | + |
| <i>Conyza bonariensis</i>   | 2 |   |   |   |   |   | + |   | + |
| <i>Cucumis melo</i>         | 5 |   |   |   |   | + |   |   |   |
| <i>Cynodon dactylon</i>     | 1 |   |   |   |   |   | + |   |   |
| <i>Cyperus longus</i>       | 1 |   |   |   |   | + |   |   |   |
| <i>Dactylis glomerata</i>   | 2 |   |   | + |   |   |   | + |   |
| <i>Datura stramonium</i>    | 1 |   |   |   |   |   |   |   | + |
| <i>Daucus sp</i>            | 2 |   |   |   |   |   | + |   | + |
| <i>Desconocida 4</i>        | 2 | + |   |   | + |   |   |   |   |
| <i>Diplotaxis erucoides</i> | 4 | + | + |   |   | + | + |   | + |
| <i>Diplotaxis sp</i>        | 1 |   |   |   |   |   |   |   | + |
| <i>Echium vulgare</i>       | 1 |   |   |   |   |   |   |   | + |
| <i>Erodium cicutarium</i>   | 1 |   |   |   |   |   |   |   | + |
| <i>Eruca vesicaria</i>      | 1 |   |   |   |   |   | + |   |   |
| <i>Eryngium campestre</i>   | 1 |   |   |   |   |   |   |   | + |
| <i>Fumaria parviflora</i>   | 1 |   |   |   |   |   |   |   | + |
| <i>Galium verum</i>         | 2 |   |   |   |   |   | + | + |   |
| <i>Hirschfeldia incana</i>  | 1 |   |   |   |   |   |   |   | + |
| <i>Hordeum matritense</i>   | 1 |   |   | + |   |   |   | + |   |
| <i>Jasminum fruticans</i>   | 2 |   |   |   |   |   |   | + | + |
| <i>Lactuca serriola</i>     | 4 |   |   |   |   |   | + |   | + |
| <i>Leontodon sp</i>         | 1 |   |   | + |   |   |   |   |   |

|                               |   |   |   |   |   |   |
|-------------------------------|---|---|---|---|---|---|
| <i>Lepidium draba</i>         | 1 |   |   |   |   | + |
| <i>Lithospermum arvense</i>   | 1 |   |   |   | + |   |
| <i>Lolium perenne</i>         | 1 |   |   |   | + |   |
| <i>Lotus corniculatus</i>     | 1 |   |   |   |   | + |
| <i>Malva sylvestris</i>       | 1 |   |   |   |   | + |
| <i>Marrubium vulgare</i>      | 1 |   | + |   |   |   |
| <i>Medicago sp</i>            | 1 |   |   |   |   | + |
| <i>Papaver rhoeas</i>         | 1 | + |   |   | + |   |
| <i>Picris echioides</i>       | 5 |   |   | + | + | + |
| <i>Plantago sp</i>            | 1 |   |   | + |   |   |
| <i>Portulaca oleraceae</i>    | 1 |   |   |   | + |   |
| <i>Quercus coccifera</i>      | 3 |   | + |   | + |   |
| <i>Quercus ilex</i>           | 1 |   |   |   |   | + |
| <i>Reseda lutea</i>           | 1 |   |   |   |   | + |
| <i>Rubia peregrina</i>        | 5 | + |   | + | + | + |
| <i>Rumex pulcher</i>          | 3 | + |   |   |   | + |
| <i>Scandix pecten veneris</i> | 1 |   |   |   |   | + |
| <i>Senecio jacobaea</i>       | 1 |   |   |   |   | + |
| <i>Silybum marianum</i>       | 2 |   |   |   |   | + |
| <i>Sisymbrium runcinatum</i>  | 1 |   |   |   |   | + |
| <i>Solanum nigrum</i>         | 1 |   |   |   |   | + |
| <i>Sonchus oleraceus</i>      | 1 | + |   |   |   |   |

[illegible]

### 4.2.3 Comparison of the host range of the different viruses.

Following the host range estimation of each virus in the habitat and ecosystem, the similarity or dissimilarities in host range among virus species was analysed by hierarchical clustering analyses. For this purpose, a matrix of the number of observations per host across host species was prepared. The analyses produced a dendrogram that showed the relationships among host ranges (Figure 4.13). The host-ranges of CMV and PZSV clustered together and differed from those of the other six viruses, which formed a larger cluster. Within this larger cluster, the host ranges of the three potyviruses, WMV, TuMV and PPV clustered together, as well as those of BChV and TAV (Figure 4.13). Thus, similarity in host range was determined in part by a common evolutionary history (the three potyvirus cluster), but otherwise did not depend on virus taxonomy or mode of transmission, indicating a role of non-deterministic ecological factors in the realised host range phenotype.

Out of the 101 host species represented in the dendrogram, *D. erucoides* was the only species used by all 8 viruses, and might play an important role in the epidemiology of the analysed viruses. Hosts such as *P. echinoides*, *R. peregrina*, *Leontodon sp.*, *G. verum*, *Bromus sp.*, *C. arvensis*, *C. bourgeanus*, *L. serriola*, *Lepidium draba*, *C. album*, *H. matritense*, *H. vulgare*, were used by at least 5 viruses. Hosts like *M. vernale*, *A. sativum*, *S. runcinatum*, *E. campestre*, *Teucrium sp.*, *Plantago*, *M. vulgare*, *V. sinuatum*, *Festuca sp.*, *T. officinale*, *O. luteus*, *A. arenaria*, *A. sterilis*, *O. vulgare* were used by at least 1 virus.

As 5 of the viruses analysed share the same non-persistent transmission by aphids but have different genome organisation, another hierarchical cluster analysis was done limited to them (Figure A1, Annexes). Again, the host range of the three potyviruses clustered together, and with that of TAV, as in the analysis involving all viruses.

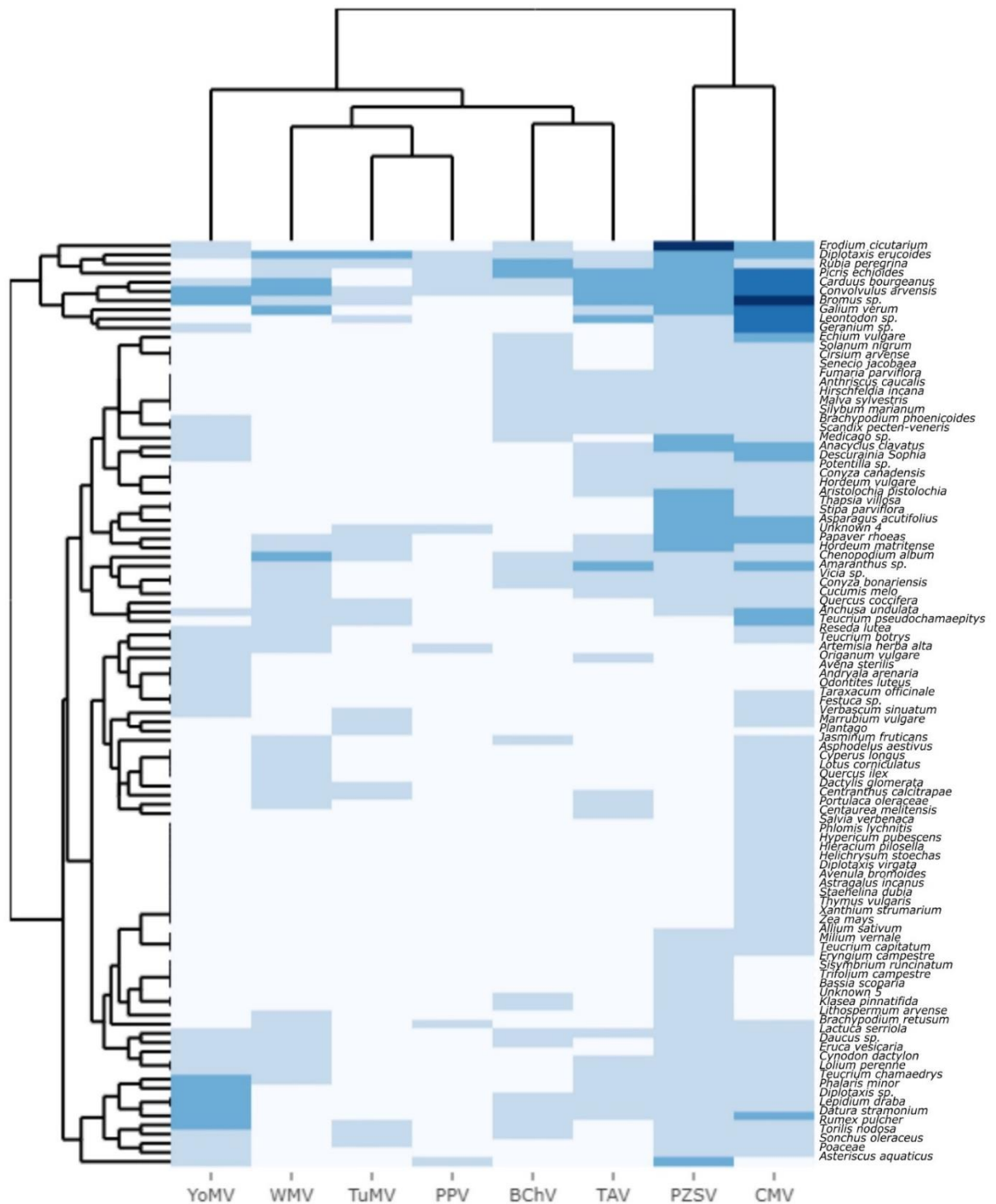


**Figure 4.13: Normalized hierarchical clustering of host ranges based on the HTS2C data used to infer dissimilarities in host use.** The heatmap was generated from a matrix of the number of species per family in the host ranges of each virus. The name of each virus species analyzed is given.

As a meta-analysis of plant virus host ranges has identified host taxonomy at the level of family as a major determinant of host range (Moury et al. 2017), a hierarchical cluster analysis of the HTS2C host ranges according to the number of species infected per plant family by each virus was performed. The clustering algorithm included a scale function to normalize the counts of rare and common interactions. Figure 4.13 shows that the hosts ranges clustered differently from Figure A1 Annexes: notably, the three potyviruses host ranges did not cluster together, and clusters were not according to virus host range or taxonomy. The plant families with more hosts were *Asteraceae*, *Poaceae* and *Brassicaceae*, which are the most abundant ones in the ecosystem (McLeish et al. 2021). This result suggest that the abundance of potential hosts plays a major role in the realised host range of a virus, by determining the probability of virus-host encounters .

The analyses of host range at the level of host species (Figure 4.13) and host family (Figure 4.14) indicates that within a plant family, viruses may use entirely different or same species. For example, within the *Asteraceae*, TuMV used *Sonchus oleraceae* and *Leontodon sp* but WMV used *Artemisia herba-alba*, *Carduus bourgeanus* and *Picris echiodes*. On the contrary, WMV and PPV used the same species within the *Asteraceae*. Such dynamics of virus-host interactions within plant families highlight the effect of species richness on host range, which can lead to niche differentiation, where some viruses may become specialists on specific hosts, while others might adapt to a broader range of hosts within the family. Thus, the analyses of host range at species and family level underscores the complementary role of ecological factors (host abundance and richness) to the influence of deterministic evolutionary factors in shaping the realised host ranges of each virus.

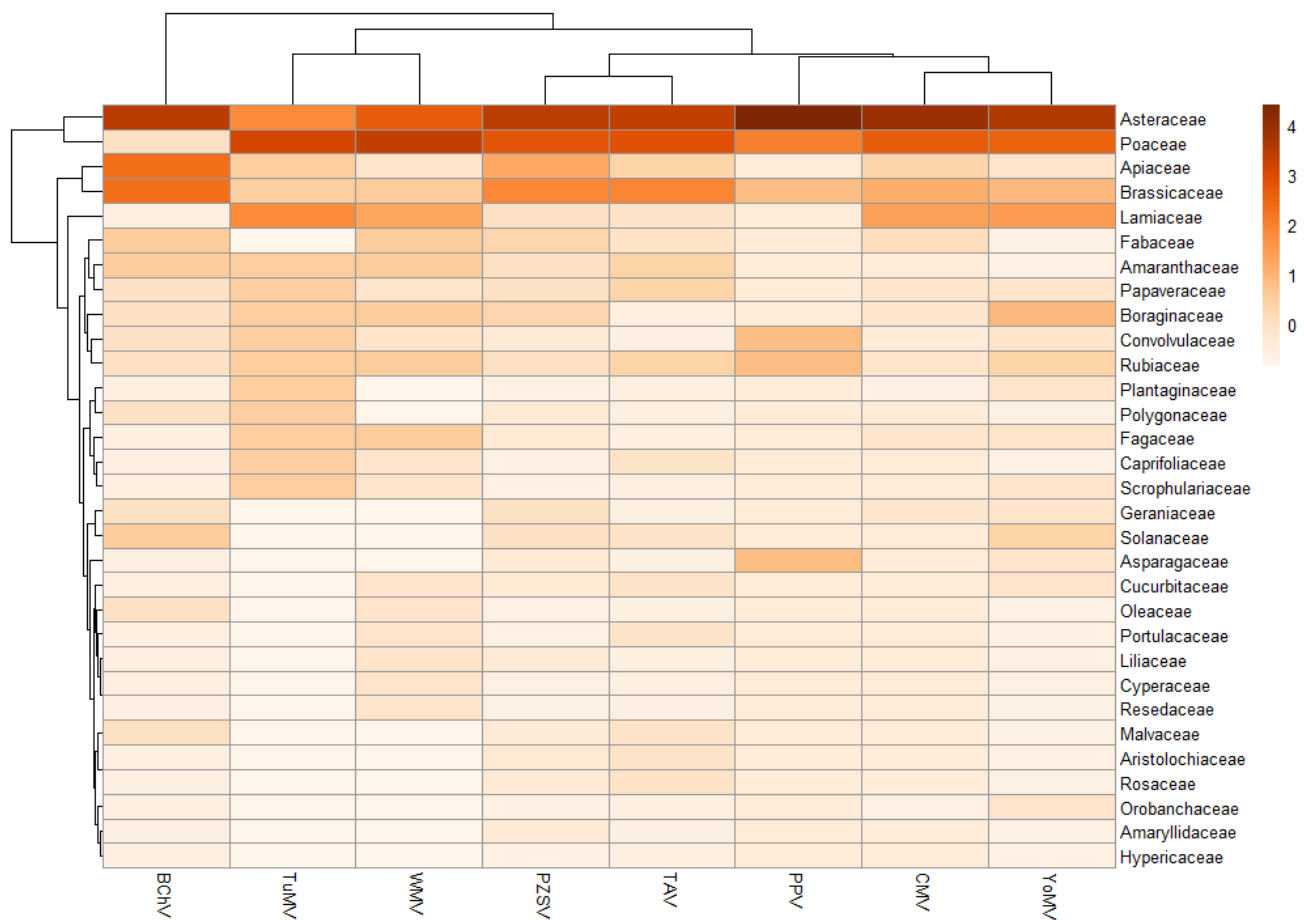


**Figure 4.14: Normalized hierarchical clustering of host ranges according to plant family based on the HTS2C data.** The heatmap was generated from a matrix of species per family in the host ranges of each virus. The names of the analyzed viruses and plant families are given.

### 4.3 Virus co-occurrence.

A common approach in ecological studies is the analysis of species co-occurrence, which is useful to understand species interactions and distribution among communities (Diamond, 1975). Patterns of co-occurrence indicate non-random (positive or negative) or random associations between virus pairs. Positive associations are observed frequencies of co-occurrence being significantly higher to the expected frequencies while negative associations are observed frequencies of co-occurrence being significantly lower to the expected frequencies. Therefore, virus co-occurrence in hosts was analysed within libraries (pools) from RT-PCR and HTS2C data, and within individual plants, to understand ecological strategies driving interaction between viruses.

The analyses at the library level with HTS2C detection (Figure 4.15) revealed 28 combinations of binary co-occurrences, of which 21 were random and seven departed significantly from random expectations, being positive associations. Positive associations occurred between CMV and TAV ( $P = 0.006$ ), TuMV and WMV ( $P = 0.012$ ), PPV and WMV ( $P = 0.009$ ), PZSV and BChV ( $P < 0.001$ ), TAV and PZSV ( $P < 0.001$ ), TAV and WMV ( $P = 0.013$ ), TAV and BChV ( $P = 0.001$ ) (Figure 4.15). YoMV was automatically excluded by the probabilistic model because it was classified as a 'truly random' species (did not differ significantly from its expected number of co-occurrences) (Griffith et al. 2016). No negative associations were identified.

Similar to the results of Figure 4.15, the analyses at the library level with RT-PCR produced 22 random and 6 positive associations between virus pairs (Figure 4.16). Here, BChV was automatically excluded by the probabilistic model. The positive associations were between YoMV and TuMV ( $P < 0.001$ ), PZSV and PPV ( $P = 0.014$ ), CMV and WMV ( $P < 0.001$ ), TAV and WMV ( $P = 0.009$ ). CMV and TAV ( $P = 0.044$ ) and PZSV and TAV ( $P = 0.047$ ). There were no negative associations observed.

The analysis performed on individual plants with RT-PCR data revealed 28 co-occurrences of which 14 were random, 9 were negative associations, and 5 were positive associations (Figure 4.17). The negative association were observed between the contact transmitted tobamovirus,

YoMV, and five viruses, PPV, BChV, PZSV, TAV and WMV with distinct transmission modes. The significant positive associations were between PPV and PZSV ( $P = 0.024$ ), CMV and WMV ( $P = 0.040$ ), CMV and TAV ( $P = 0.018$ ), TuMV and TAV ( $P = 0.018$ ), TAV and WMV ( $P < 0.001$ ) (Figure 4.17).

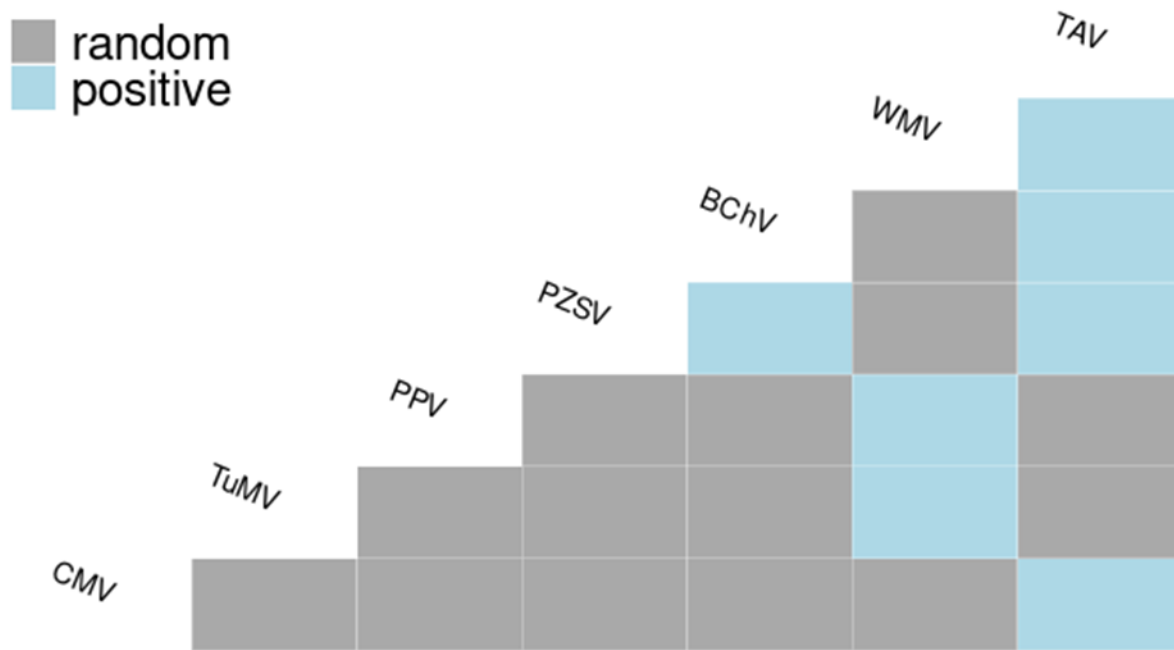


**Figure 4.15: Co-occurrence heatmap at library level with HTS2C data showing associations between pairs of virus species.** Associations are colour coded as follows: random, grey; positive, light blue.

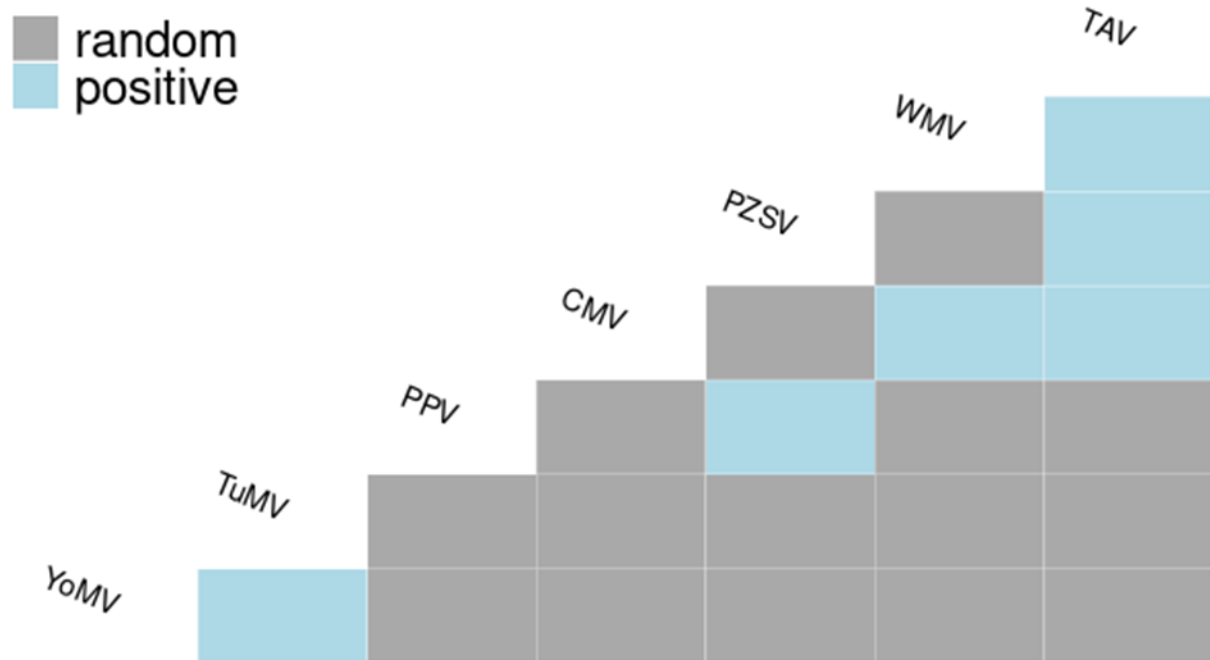


**Figure 4.16: Co-occurrence heatmap with libraries of RT-PCR data showing associations between pairs of virus species.** Associations are colour coded as follows: random, grey; positive, light blue.

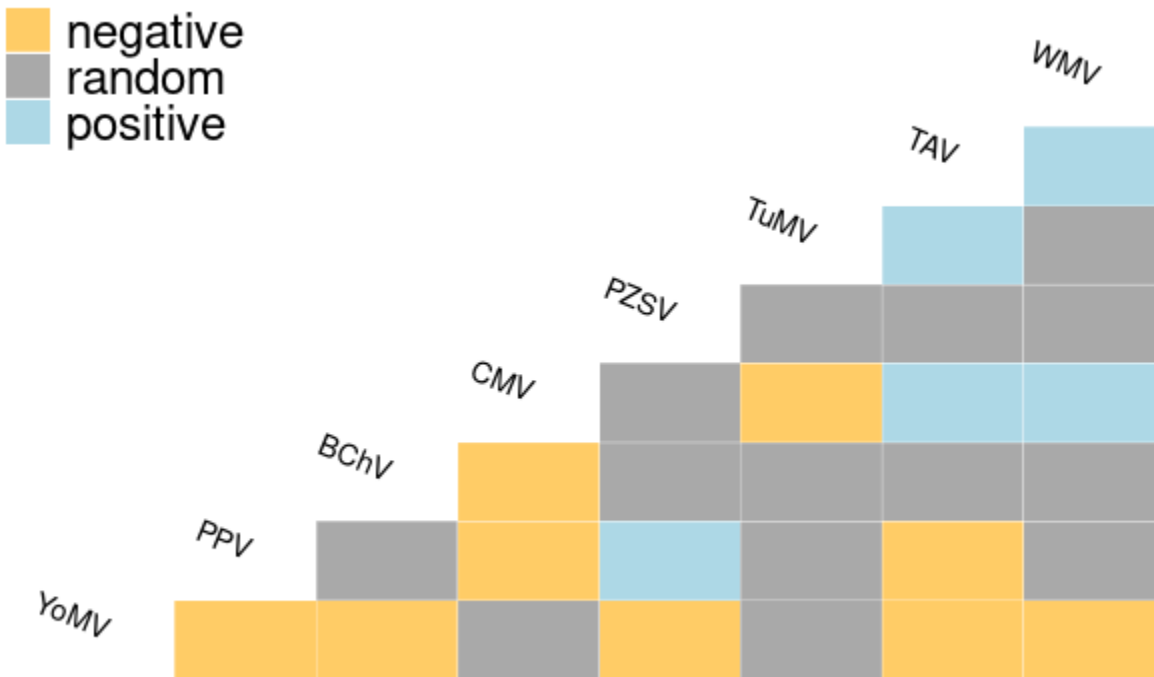


**Figure 4.17: Co-occurrence heatmap of individual plants with RT-PCR data showing associations between pairs of OTUs.** Associations are colour coded as follows: negative, yellow; random, grey; positive, light blue.

Analyses at the library level show that a large fraction of the interactions among virus species were random, and there were no negative associations, an indication of non-competitive coexistence. Analysis at the individual plant level revealed negative associations were among contact, pollen and aphid-transmitted viruses, but positive associations were mostly observed among virus species that were transmitted non-persistently by aphids, with the exception of the association TAV and PZSV with distinct transmission modes and host ranges. All three analyses of co-occurrence consistently supported positive associations between CMV and TAV and between TAV and WMV, while two of the analyses supported positive association between CMV and WMV, TAV and PZSV, and PZSV and PPV.

## 4.4 Patterns of virus incidence.

Incidence of virus infection (defined as the percentage of infected plant individuals over the total number of plant individuals analysed) was used as a proxy for host use, since it gives information on the presence and distribution of each virus among different plant species, and which host species are more frequently infected. As explained in section 3.3.3, two incidence variables were analysed: incidence in observed hosts (*ih*) and incidence in the entire plant community (*iL*). *ih* is the ratio of RT-PCR positive samples to the total sample analysed for a given host species and provides information on differential host use within the host range. *iL* provides a view of virus incidence and host range in the entire plant community, that is, considering hosts (HTS2C positive libraries) and non-hosts (HTS2C negative libraries), and may better reflect the role of the plant community on infection risk. Both *ih* and *iL* did not meet the conditions of normality, therefore non-parametric statistical tests (Kruskal-Wallis) were used for further analysis. At the ecosystem scale both *ih* and *iL* varied for each virus species, and for each virus species both variables also varied according to habitat. Tables 4.6 and 4.7 which represent *ih* and *iL* of CMV respectively are shown in this section as examples. For the other seven viruses, data is presented in Tables A4 to A17 in Annexes.

Incidence of CMV varied among host species, for example, it was high (75 – 100 %) in *B. retusum*, *Bromus sp*, *C. melo*, etc, intermediate (35 – 75 %) in *Geranium sp*, *Q. ilex*, *H. maritense*, etc and low (< 35 %) in *A. clavatus*, *C. dactylon*, *D. sophia* etc. Most hosts of CMV were present only in a single habitat while 7 hosts were present in two or more habitats, and incidence in these hosts depended on the habitat. For example, incidence in *C. arvensis* was 100 and 91 % in Crop and Edge respectively and 38 % in Wasteland, while in *Leontodon sp.* incidence was 50 % in Oakwood and 17 % Wasteland. These results for CMV show that despite being a generalist virus, CMV has a preferential host use which is dependent on habitat.

**Table 4.6: CMV incidence (*ih*) in its different hosts across all habitats and at the ecosystem scale.** Percentage (%) of infection and sample size of host (n) for each habitat is shown. The hyphen ‘–’ indicates no record.

| Host Taxon                       | Crop (%) | n  | Edge (%) | n  | Oak. (%) | n  | Wastel. (%) | n  | Total inc. (%) | Total n. |
|----------------------------------|----------|----|----------|----|----------|----|-------------|----|----------------|----------|
| <i>Amaranthus sp</i>             | 86       | 7  | -        | -  | -        | -  | -           | -  | 86             | 7        |
| <i>Anacyclus clavatus</i>        | 25       | 4  | -        | -  | -        | -  | -           | -  | 25             | 4        |
| <i>Anchusa undulata</i>          | -        | -  | -        | -  | -        | -  | 87.5        | 8  | 87.5           | 8        |
| <i>Anthriscus caucalis</i>       | -        | -  | 100      | 6  | -        | -  | -           | -  | 100            | 6        |
| <i>Asparagus acutifolius</i>     | -        | -  | -        | -  | 91.4     | 35 | -           | -  | 91.4           | 35       |
| <i>Avenula bromoides</i>         | -        | -  | -        | -  | 58.3     | 12 | -           | -  | 58.3           | 12       |
| <i>Brachypodium phoenicoides</i> | -        | -  | 100      | 1  | -        | -  | -           | -  | 100            | 1        |
| <i>Brachypodium retusum</i>      | -        | -  | 100      | 10 | -        | -  | -           | -  | 100            | 10       |
| <i>Bromus sp</i>                 | -        | -  | 80       | 5  | 80       | 5  | 100         | 5  | 86.6           | 15       |
| <i>Carduus bourgeanus</i>        | 83       | 6  | 80       | 5  | -        | -  | -           | -  | 81.8           | 11       |
| <i>Centhranthus calcitrapae</i>  | -        | -  | -        | -  | 100      | 4  | -           | -  | 100            | 4        |
| <i>Convolvulus arvensis</i>      | 100      | 6  | 90.9     | 11 | -        | -  | 37.5        | 8  | 76             | 25       |
| <i>Cucumis melo</i>              | 98.1     | 53 | -        | -  | -        | -  | -           | -  | 98.1           | 53       |
| <i>Cynodon dactylon</i>          | -        | -  | 33       | 6  | -        | -  | -           | -  | 33             | 6        |
| <i>Cyperus longus</i>            | 100      | 3  | -        | -  | -        | -  | -           | -  | 100            | 3        |
| <i>Datura stramonium</i>         | -        | -  | 100      | 1  | -        | -  | -           | -  | 100            | 1        |
| <i>Daucus sp</i>                 | -        | -  | 100      | 3  | -        | -  | -           | -  | 100            | 3        |
| <i>Descurania sophia</i>         | -        | -  | -        | -  | -        | -  | 17          | 6  | 17             | 6        |
| <i>Diplotaxis erucoides</i>      | -        | -  | 100      | 8  | -        | -  | 79          | 19 | 48.1           | 27       |

|                               |     |   |      |    |      |    |      |    |      |    |
|-------------------------------|-----|---|------|----|------|----|------|----|------|----|
| <i>Diploaxis virgata</i>      | -   | - | -    | -  | 80   | 5  | -    | -  | 80   | 5  |
| <i>Echium vulgare</i>         | -   | - | -    | -  | -    | -  | 100  | 7  | 100  | 7  |
| <i>Erodium cicutarium</i>     | -   | - | -    | -  | 67   | 3  | -    | -  | 67   | 3  |
| <i>Fumaria parviflora</i>     | -   | - | 100  | 9  | -    | -  | -    | -  | 100  | 9  |
| <i>Galium verum</i>           | -   | - | 0    | 1  | 100  | 4  | 0    | 9  | 28.6 | 14 |
| <i>Geranium sp</i>            | -   | - | -    | -  | 67   | 3  | -    | -  | 67   | 3  |
| <i>Hordeum maritense</i>      | -   | - | -    | -  | -    | -  | 54.5 | 11 | 54.5 | 11 |
| <i>Lactuca serriola</i>       | -   | - | 100  | 1  | -    | -  | -    | -  | 100  | 1  |
| <i>Leontodon sp</i>           | -   | - | 100  | 1  | 50   | 4  | 17   | 6  | 36.4 | 11 |
| <i>Lolium perenne</i>         | -   | - | 100  | 4  | -    | -  | -    | -  | 100  | 4  |
| <i>Lotus corniculatus</i>     | -   | - | -    | -  | 50   | 8  | -    | -  | 50   | 8  |
| <i>Marrubium vulgare</i>      | -   | - | -    | -  | 50   | 18 | -    | -  | 50   | 18 |
| <i>Milium vernale</i>         | -   | - | 100  | 13 | -    | -  | -    | -  | 100  | 13 |
| <i>Papaver rhoeas</i>         | -   | - | 100  | 1  | -    | -  | 100  | 7  | 100  | 8  |
| <i>Picris echoides</i>        | -   | - | -    | -  | -    | -  | 83   | 12 | 83   | 12 |
| <i>Portulaca oleraceae</i>    | 100 | 4 | -    | -  | -    | -  | -    | -  | 100  | 4  |
| <i>Quercus coccifera</i>      | -   | - | -    | -  | 10   | 20 | -    | -  | 10   | 20 |
| <i>Quercus ilex</i>           | -   | - | -    | -  | 69.4 | 36 | -    | -  | 69.4 | 36 |
| <i>Rubia peregrina</i>        | -   | - | 83.3 | 6  | -    | -  | -    | -  | 83.3 | 6  |
| <i>Rumex plucher</i>          | -   | - | 67   | 3  | -    | -  | -    | -  | 67   | 3  |
| <i>Salvia verbanaca</i>       | -   | - | -    | -  | -    | -  | 40   | 10 | 40   | 10 |
| <i>Scandix pecten_veneris</i> | -   | - | 100  | 1  | -    | -  | -    | -  | 100  | 1  |

|                             |             |           |           |            |           |            |             |            |             |            |
|-----------------------------|-------------|-----------|-----------|------------|-----------|------------|-------------|------------|-------------|------------|
| <i>Staehelina dubia</i>     | -           | -         | -         | -          | 100       | 5          | -           | -          | 100         | 5          |
| <i>Stipa parviflora</i>     | -           | -         | -         | -          | 100       | 1          | -           | -          | 100         | 1          |
| <i>Taraxacum officinale</i> | -           | -         | -         | -          | -         | -          | 60          | 5          | 60          | 5          |
| <i>Teucrium botrys</i>      | -           | -         | -         | -          | 100       | 6          | -           | -          | 100         | 6          |
| <i>Teucrium capitatum</i>   | -           | -         | -         | -          | 90        | 10         | -           | -          | 90          | 10         |
| <i>Teucrium chamaedrys</i>  | -           | -         | -         | -          | 100       | 1          | -           | -          | 100         | 1          |
| <i>Thymus vulgaris</i>      | -           | -         | -         | -          | 75        | 8          | -           | -          | 75          | 8          |
| <i>Vicia sp</i>             | -           | -         | 66.6      | 3          | -         | -          | -           | -          | 66.6        | 3          |
| <i>Zea mays</i>             | 100         | 10        | -         | -          | -         | -          | -           | -          | 100         | 10         |
| <b>Total</b>                | <b>93.5</b> | <b>93</b> | <b>88</b> | <b>100</b> | <b>63</b> | <b>188</b> | <b>58.3</b> | <b>120</b> | <b>72.5</b> | <b>501</b> |

---

**Table 4.7: CMV incidence in the whole plant community (*iL*) across all habitats and at the ecosystem scale.** Percentage (%) of infection and sample size of host (*n*) for each habitat is shown. The hyphen ‘–’ indicates no record.

| Host Taxon                       | Crop (%) | <i>n</i> | Edge (%) | <i>n</i> | Oak. (%) | <i>n</i> | Wastel. (%) | <i>n</i> | Total inc. (%) | Total <i>n</i> . |
|----------------------------------|----------|----------|----------|----------|----------|----------|-------------|----------|----------------|------------------|
| <i>Amaranthus sp</i>             | 86       | 7        | 0        | 12       | -        | -        | -           | -        | 31.7           | 19               |
| <i>Anacyclus clavatus</i>        | 25       | 4        | -        | -        | -        | -        | 0           | 5        | 11.1           | 9                |
| <i>Anchusa undulata</i>          | -        | -        | 0        | 6        | -        | -        | 87.5        | 21       | 25.9           | 27               |
| <i>Anthriscus caucalis</i>       | -        | -        | 100      | 6        | -        | -        | -           | -        | 100            | 6                |
| <i>Asparagus acutifolius</i>     | -        | -        | -        | -        | 91.4     | 35       | -           | -        | 91.4           | 35               |
| <i>Avenula bromoides</i>         | -        | -        | -        | -        | 58.3     | 12       | -           | -        | 58.3           | 12               |
| <i>Brachypodium phoenicoides</i> | -        | -        | 100      | 1        | -        | -        | -           | -        | 100            | 1                |
| <i>Brachypodium retusum</i>      | -        | -        | 100      | 10       | -        | -        | -           | -        | 100            | 10               |
| <i>Bromus sp</i>                 | -        | -        | 80       | 5        | 80       | 5        | 100         | 5        | 86.6           | 15               |
| <i>Carduus bourgeanus</i>        | 83       | 6        | 80       | 7        | -        | -        | 0           | 8        | 42.8           | 21               |
| <i>Centhranthus calcitrapae</i>  | -        | -        | -        | -        | 100      | 4        | 0           | 1        | 80             | 5                |
| <i>Convolvulus arvensis</i>      | 100      | 6        | 90.9     | 11       | -        | -        | 37.5        | 8        | 76             | 25               |
| <i>Cucumis melo</i>              | 98.1     | 99       | -        | -        | -        | -        | -           | -        | 52.5           | 99               |
| <i>Cynodon dactylon</i>          | -        | -        | 33       | 8        | -        | -        | -           | -        | 24.8           | 8                |
| <i>Cyperus longus</i>            | 100      | 3        | -        | -        | -        | -        | -           | -        | 100            | 3                |
| <i>Datura stramonium</i>         | -        | -        | 100      | 1        | -        | -        | -           | -        | 100            | 1                |
| <i>Daucus sp</i>                 | -        | -        | 100      | 3        | -        | -        | -           | -        | 100            | 3                |

|                             |     |   |      |    |      |    |      |    |      |    |
|-----------------------------|-----|---|------|----|------|----|------|----|------|----|
| <i>Descurania sophia</i>    | -   | - | -    | -  | -    | -  | 17   | 6  | 17   | 6  |
| <i>Diplotaxis eruroides</i> | -   | - | 100  | 8  | -    | -  | 79   | 19 | 48.1 | 27 |
| <i>Diplotaxis virgata</i>   | -   | - | -    | -  | 80   | 5  | -    | -  | 80   | 5  |
| <i>Echium vulgare</i>       | -   | - | -    | -  | -    | -  | 100  | 7  | 100  | 7  |
| <i>Erodium cicutarium</i>   | 0   | 1 | -    | -  | 67   | 3  | 0    | 3  | 28.7 | 7  |
| <i>Fumaria parviflora</i>   | 0   | 8 | 100  | 9  | -    | -  | -    | -  | 52.9 | 17 |
| <i>Galium verum</i>         | -   | - | 0    | 2  | 100  | 12 | 0    | 9  | 17.4 | 23 |
| <i>Geranium sp</i>          | -   | - | 0    | 1  | 67   | 3  | -    | -  | 50.3 | 4  |
| <i>Hordeum maritense</i>    | -   | - | -    | -  | 0    | 2  | 54.5 | 11 | 46.1 | 13 |
| <i>Latuca serriola</i>      | -   | - | 33.3 | 3  | -    | -  | -    | -  | 33.3 | 3  |
| <i>Leontodon sp</i>         | -   | - | 100  | 1  | 50   | 4  | 17   | 14 | 21.1 | 19 |
| <i>Lolium perenne</i>       | -   | - | 100  | 4  | -    | -  | -    | -  | 100  | 4  |
| <i>Lotus corniculatus</i>   | -   | - | -    | -  | 50   | 8  | -    | -  | 50   | 8  |
| <i>Marrubium vulgare</i>    | -   | - | -    | -  | 50   | 18 | -    | -  | 50   | 18 |
| <i>Milium vernale</i>       | -   | - | 100  | 13 | -    | -  | -    | -  | 100  | 13 |
| <i>Papaver rhoeas</i>       | -   | - | 100  | 1  | 0    | 1  | 100  | 7  | 88.9 | 9  |
| <i>Picris echoides</i>      | 0   | 2 | 0    | 8  | -    | -  | 83   | 12 | 45.3 | 22 |
| <i>Portulaca oleraceae</i>  | 100 | 4 | -    | -  | -    | -  | -    | -  | 100  | 4  |
| <i>Quercus coccifera</i>    | -   | - | -    | -  | 10   | 20 | -    | -  | 10   | 20 |
| <i>Quercus ilex</i>         | -   | - | -    | -  | 69.4 | 36 | -    | -  | 69.4 | 36 |
| <i>Rubia peregrina</i>      | 0   | 9 | 83.3 | 8  | 0    | 5  | 0    | 1  | 21.7 | 23 |

|                               |             |            |             |            |             |            |             |            |             |            |
|-------------------------------|-------------|------------|-------------|------------|-------------|------------|-------------|------------|-------------|------------|
| <i>Rumex plucher</i>          | -           | -          | 22.3        | 9          | -           | -          | -           | -          | 22.3        | 9          |
| <i>Salvia verbanaca</i>       | -           | -          | -           | -          | -           | -          | 40          | 10         | 40          | 10         |
| <i>Scandix pecten_veneris</i> | -           | -          | 100         | 1          | -           | -          | -           | -          | 100         | 1          |
| <i>Staelina dubia</i>         | -           | -          | -           | -          | 100         | 5          | -           | -          | 100         | 5          |
| <i>Stipa parviflora</i>       | -           | -          | -           | -          | 100         | 1          | 0           | 6          | 14.3        | 7          |
| <i>Taraxacum officinale</i>   | -           | -          | -           | -          | -           | -          | 60          | 5          | 60          | 5          |
| <i>Teucrium botrys</i>        | -           | -          | -           | -          | 100         | 6          | -           | -          | 100         | 6          |
| <i>Teucrium capitatum</i>     | -           | -          | -           | -          | 90          | 10         | 0           | 15         | 36          | 25         |
| <i>Teucrium chamaedrys</i>    | -           | -          | -           | -          | 100         | 1          | 0           | 3          | 25          | 4          |
| <i>Thymus vulgaris</i>        | -           | -          | -           | -          | 37.5        | 16         | -           | -          | 37.5        | 16         |
| <i>Vicia sp</i>               | -           | -          | 66.6        | 3          | -           | -          | 0           | 11         | 14.3        | 14         |
| <i>Zea mays</i>               | 100         | 10         | -           | -          | -           | -          | -           | -          | 100         | 10         |
| <b>Total</b>                  | <b>54.7</b> | <b>159</b> | <b>62.4</b> | <b>141</b> | <b>56.1</b> | <b>212</b> | <b>36.9</b> | <b>187</b> | <b>51.9</b> | <b>699</b> |

---

Similarly, incidence of the other seven viruses — TAV, TuMV, PPV, WMV, PZSV, BChV, and YoMV (see Tables A4 - A417 in Annexes) showed large variation across hosts, that is, a preferential host use. For all viruses, most hosts were specific to a single habitat and few hosts were present in two or more habitats. The differential host use across habitats observed in CMV was a common but not universal pattern for viruses and hosts in different habitats — e.g. the incidence of TAV in *Bromus sp.* was 86% in Edge and 25% in Oakwood, but for *C. bourgeanus*, TAV incidence did not differ in Edge and Oakwood, being 50 % in each habitat.

A global analysis of incidences across viruses and across habitats was performed with Kruskal-Wallis tests. Comparisons of *ih* across viruses at the ecosystem showed that *ih* for CMV was significantly higher than that of BChV ( $P=0.007$ ), PZSV ( $P=0.020$ ), and YoMV ( $P<0.001$ ). Additionally, the incidence of TAV ( $P=0.003$ ) and WMV ( $P=0.005$ ) was significantly higher than that of YoMV. In the case of *ih* across habitats considering all viruses together, it was found that the incidence was not statistically different across most habitat pairs ( $P>0.05$ ). However, the incidence in Crop ( $P=0.011$ ) and Edge habitats ( $P<0.001$ ) was significantly greater than in Wasteland habitats.

Comparisons of *iL* across viruses at the ecosystem revealed *iL* for CMV was higher than that for BChV ( $P=0.025$ ), PPV ( $P=0.000$ ), TAV ( $P=0.010$ ), TuMV ( $P=0.000$ ), WMV ( $P=0.027$ ), and YoMV ( $P=0.001$ ) but not PZSV ( $P=0.177$ ). However, *iL* of PZSV was not statistically different from other viruses and same was true of *iL* for other virus pairs ( $P>0.05$ ). For comparisons of *iL* across habitats considering all viruses together, *iL* was significantly higher in Edge than in Crop ( $P<0.001$ ), Oakwood ( $P=0.024$ ), and Wasteland ( $P<0.001$ ). *iL* did not differ for other pairs of habitats except for Oakwood which was greater than Crop ( $P=0.025$ ) and Wasteland ( $P=0.026$ ).

Kruskal-Wallis tests were used to compare mean *ih* of each virus species across habitats. Results show that *ih* of CMV, which was of 72.5% in the ecosystem, differed between habitats ( $\chi^2= 7.97$ ,  $df=3$ ,  $P=0.046$ ). Pairwise comparison revealed that *ih* in Edge was significantly greater than Wasteland ( $P = 0.030$ ). *iL* of CMV at the ecosystem scale was 51.9 % (Table 4.7). Again, *iL* of varied among habitats ( $\chi^2 = 8.18$ ,  $df=3$ ,  $P=0.042$ ), and pairwise comparisons revealed that *iL* was higher in Edge than in Wasteland ( $P= 0.017$ ) but did not differ in other habitats ( $P>0.05$ ).

*ih* of TAV at the ecosystem scale was 71% (Table A4, Annexes). Comparison of mean *ih* of TAV across habitats showed that there was no significant difference between habitats ( $\chi^2 = 7.37$ ,  $df=3$ ,  $P=0.061$ ). *iL* of TAV at the ecosystem scale was 24.8 % (174 positive samples out of 702 analyzed) (Table A5, Annexes). *iL* varied among habitats ( $\chi^2 = 28.87$ ,  $df=3$ ,  $P<0.001$ ), and pairwise comparisons revealed that *iL* was higher in Edge than in Wasteland ( $P<0.001$ ) and Crop ( $P<0.001$ ).

The analysis of *ih* for the other six virus species — TuMV, PPV, WMV, PZSV, BChV, and YoMV — across various habitats revealed distinct patterns but also similarities. WMV showed no significant differences in *ih* across habitats ( $\chi^2 = 2.55$ ,  $df = 3$ ,  $P = 0.467$ ), similar to YoMV ( $\chi^2 = 0.847$ ,  $df = 3$ ,  $P = 0.838$ ), PPV ( $\chi^2 = 1.172$ ,  $df = 3$ ,  $P = 0.759$ ), TuMV ( $\chi^2 = 3.835$ ,  $df = 3$ ,  $P = 0.147$ ) and BChV ( $\chi^2 = 2.487$ ,  $df = 1$ ,  $P = 0.1147$ ), indicating incidence of most viruses were similar across different habitats. In contrast, PZSV displayed a significant variation in incidence across habitats ( $\chi^2 = 8.157$ ,  $df = 3$ ,  $P = 0.043$ ), with a significantly higher incidence in Edge than Oakwood ( $P = 0.023$ ).

For the other six virus species *iL* showed no significant differences among habitats for PPV ( $\chi^2 = 3.121$ ,  $df = 3$ ,  $P = 0.373$ ) while for PZSV, *iL* differed across habitats ( $\chi^2 = 36.429$ ,  $df = 3$ ,  $P < 0.001$ ), and was higher in the Edge compared to Crop, Oakwood, and Wasteland ( $P < 0.001$  for all comparisons). BChV also demonstrated significant differences across habitats ( $\chi^2 = 9.034$ ,  $df = 3$ ,  $P = 0.029$ ), with a higher *iL* in Edge compared to Wasteland and Oakwood ( $P = 0.019$ ), where no incidence was recorded. TuMV followed a similar trend with significant differences across habitats ( $\chi^2 = 10.51$ ,  $df = 3$ ,  $P = 0.015$ ), showing a higher incidence in Edge compared to Oakwood ( $P = 0.041$ ). WMV displayed differences in *iL* across habitats ( $\chi^2 = 8.577$ ,  $df = 3$ ,  $P = 0.035$ ), with a significantly higher incidence in Oakwood compared to Wasteland ( $P = 0.017$ ). Similarly, YoMV showed significant difference ( $\chi^2 = 16.57$ ,  $df = 3$ ,  $P < 0.001$ ), with a higher incidence in Oakwood compared to Crop ( $P = 0.002$ ) and Edge ( $P = 0.011$ ), and a higher incidence in Wasteland compared to other habitats ( $P = 0.015$ ).

In summary, each virus species had distinct preferential host uses regardless of mode of transmission and the virus genus to which they belonged. A common pattern across viruses is that *ih* was not always higher in the most anthropic habitats - Crop and Edge.

#### 4.4.1 Ecological and genetic determinants of incidence.

Further to the analyses of incidence patterns, I also analysed the role of ecological/extrinsic factors related to the plant community in which plant-virus interactions occurred (habitat), and genetic/virus intrinsic factors (as reflected by species, genus, mode of transmission) on incidence using regression models.

Therefore, to determine the effect of ecological or genetic factors on incidence either within observed hosts (*ih*) or in the plant community (*iL*), incidence data were aggregated according to the following determinant variables – virus species, virus genus, transmission mode, and habitat. A series of fully factorial GLM regression analyses were performed with the determinant variables classed as factors and considered as fixed effects on the incidence (response variables). The overall fit of each model was compared to a null model by ANOVA using the likelihood ratio chi-square test ( $G^2$ ), while the significance of coefficients in each model was analysed with Wald chi-square test ( $W$ ). Posthoc analyses of each model were used to summarise the patterns of incidence. In GLMs, odds ratios (OR) are a key measure of association between factors and response variable: an OR > 1 indicates higher effect of a factor on *ih* while OR < 1 indicates lower effect. To further determine the probability of virus incidence, odd ratios (OR) were log-transformed and probability was expressed as decimal value ranging from 0 to 1, where 0 indicates no chance of incidence and 1 indicates certainty of incidence.

The first GLM model considered habitat and virus species as fixed effects with their interaction:  $ih \sim \text{habitat} + \text{virus species} + \text{habitat} * \text{virus species} + \varepsilon$  (Model 1), where  $\varepsilon$  is the error term that accounts for random variability and unmeasured factors affecting *ih*. Model 1 was statistically significant, ( $G^2_{28, 133} = 379.85$ ,  $P < 0.001$ ), and explained 0.737 (Nagelkerke  $R^2$ ) of the variance in incidence. The model also shows *ih* varied according to the habitat ( $W_{3, 133} = 122.84$ ,  $P < 0.001$ ), virus species ( $W_{7, 133} = 178.50$ ,  $P < 0.001$ ), and their interaction ( $W_{18, 133} = 78.49$ ,  $P < 0.001$ ). Virus species accounted for 45.6% of the explained variance of *ih*, habitat 26.5%, and their interaction 27.9%. Since interaction between habitat and virus species was significant, again, new logistic regressions consisting of only virus species as factor was analyzed for each habitat. The model for Crop showed that *ih* for virus species was statistically significant ( $W_{6, 18} = 200.45$ ,  $P < 0.001$ ), this

was same for Edge ( $W_{7,62} = 86.64$ ,  $P < 0.001$ ), Oakwood ( $W_{6,24} = 36.16$ ,  $P < 0.001$ ) and Wasteland ( $W_{6,28} = 32.20$ ,  $P < .001$ ). Pairwise comparisons of virus species across habitats revealed significant differences in the odds ratios (OR). The *ih* of CMV was significantly higher compared to YoMV across all four habitats —Crop (OR = 23.2,  $P < 0.001$ ), Edge (OR = 20.16,  $P < 0.001$ ), Oakwood (OR = 3.50,  $P < 0.001$ ), Wasteland (OR = 3.06,  $P < 0.001$ )—and PZSV across three habitats —Crop (OR = 19.7,  $P < 0.001$ ), Edge (OR = 3.66,  $P = 0.004$ ), Oakwood (OR = 6.32,  $P = 0.003$ ). *ih* of CMV was also significantly higher than PPV in both Crop (OR = 30.9,  $P < 0.001$ ) and Edge (OR = 4.2,  $P = 0.027$ ). BChV had significantly lower *ih* when compared to YoMV in the Crop habitat ( $P < 0.05$ ), and less compared to PPV, PZSV, and TAV in the edge habitat ( $P < 0.05$ ).

For the role of virus species on *ih*, post-hoc analyses (Table A18, Annexes) show that *ih* of CMV, PZSV, TAV, PPV, TuMV and WMV was significantly higher than *ih* of BChV. Apart from WMV, the *ih* of CMV was significantly higher when compared with the other six viruses: PZSV, TAV, TuMV, PPV, BChV and YoMV. Also, the *ih* of PZSV, TAV, TuMV, WMV and PPV were only higher and significantly different from YoMV. No more significant differences were found between the remaining virus species. Figure 4.18 shows that probability of *ih* was higher ( $>0.5$ ) with CMV, WMV, TAV, PPV, and, followed by PZSV and TuMV and lower ( $<0.5$ ) with YoMV and BChV.

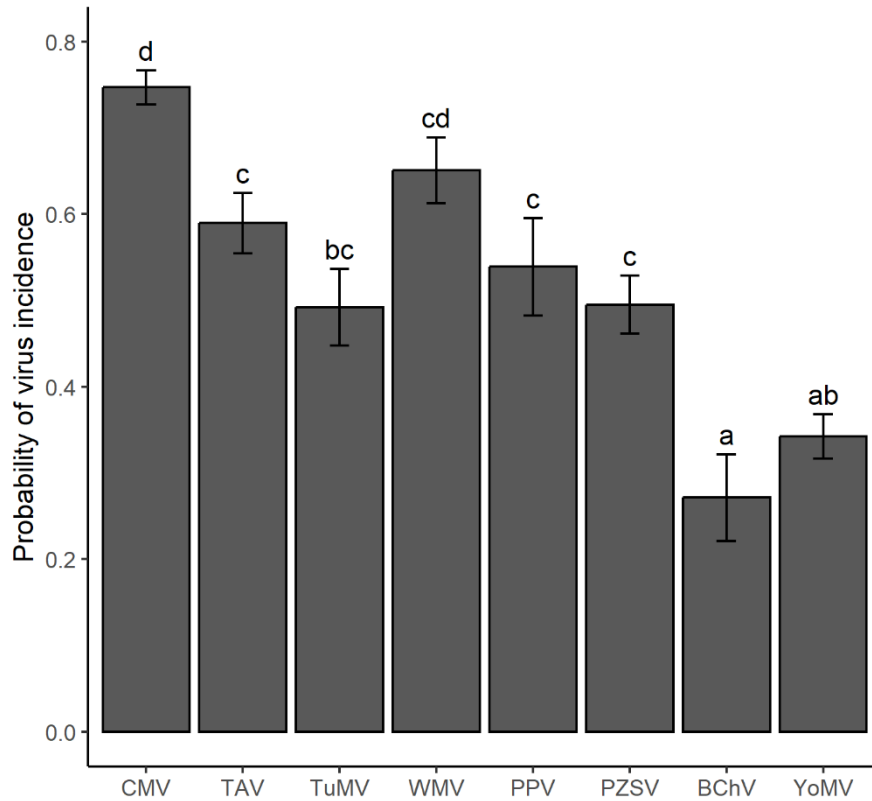


**Figure 4.18: Probability of *ih* of each virus species.** Equal letter indicates difference was not significant.

The second GLM considered habitat and virus genera as fixed effects with their interaction:  $ih \sim \text{habitat} + \text{virus genera} + \text{habitat} * \text{virus genera} + \varepsilon$  (Model 2). Model 2 was statistically significant, ( $G^2_{17,133} = 329.4$ ,  $P < 0.001$ ), and explained 0.623 (Nagelkerke  $R^2$ ) of the variance in incidence. The model also shows *ih* varied according to the habitat ( $W_{3,133} = 122.84$ ,  $P < 0.001$ ), virus genus ( $W_{4,133} = 155.09$ ,  $P < 0.001$ ), and their interaction ( $W_{10,133} = 51.45$ ,  $P < 0.001$ ). Virus genus explained 48.4% of the variance of *ih*, habitat 29.5%, and their interaction 22.1%. Since interaction between habitat and virus genus was significant, new logistic regressions consisting of only virus genera as a fixed factor was analyzed for each habitat. The model for Crop habitat showed that *ih* for virus genera was statistically significant ( $W_{4,19} = 68.33$ ,  $P < 0.001$ ), this was same for Edge ( $W_{4,62} = 79.67$ ,  $P < 0.001$ ), Oakwood ( $W_{4,24} = 30.35$ ,  $P < 0.001$ ) and Wasteland ( $W_{4,28} = 28.18$ ,  $P < 0.001$ ).

Pairwise comparisons of virus genera across habitats revealed significant differences in the odds ratios (OR). *ih* of cucumovirus was consistently higher across all habitats compared to other genera. Specifically, the *ih* of cucumovirus was higher than that of tobamovirus in Crop (OR = 10.43,  $P < 0.001$ ), Edge (OR = 11.51,  $P < 0.001$ ), Oakwood (OR = 3.05,  $P < 0.001$ ), and Wasteland (OR = 3.25,  $P < 0.001$ ). The *ih* of cucumovirus was also significantly higher than polerovirus (OR = 10.34,  $P < 0.001$ ) in Edge, and of anulavirus in Edge (OR = 2.09,  $P < 0.001$ ) and Oakwood (OR = 5.52,  $P = 0.021$ ). Additionally, tobamovirus showed significant differences when compared to other virus genera in Crop, with lower odds of incidence compared to polerovirus (OR = 28.94,  $P = 0.016$ ), and potyvirus (OR = 9.42,  $P < 0.001$ ).

To determine the role of virus genera on *ih*, post-hoc analyses (Table A19, Annexes) showed the *ih* of cucumovirus was significantly higher than that of anulavirus, polerovirus, potyvirus and tobamovirus. Also, *ih* of anulavirus was significantly higher than that of polerovirus and tobamovirus, but not potyvirus. On the other hand, *ih* of potyvirus was significantly higher than that of or tobamovirus. No significant difference was found between tobamovirus and polerovirus (Figure 4.19).

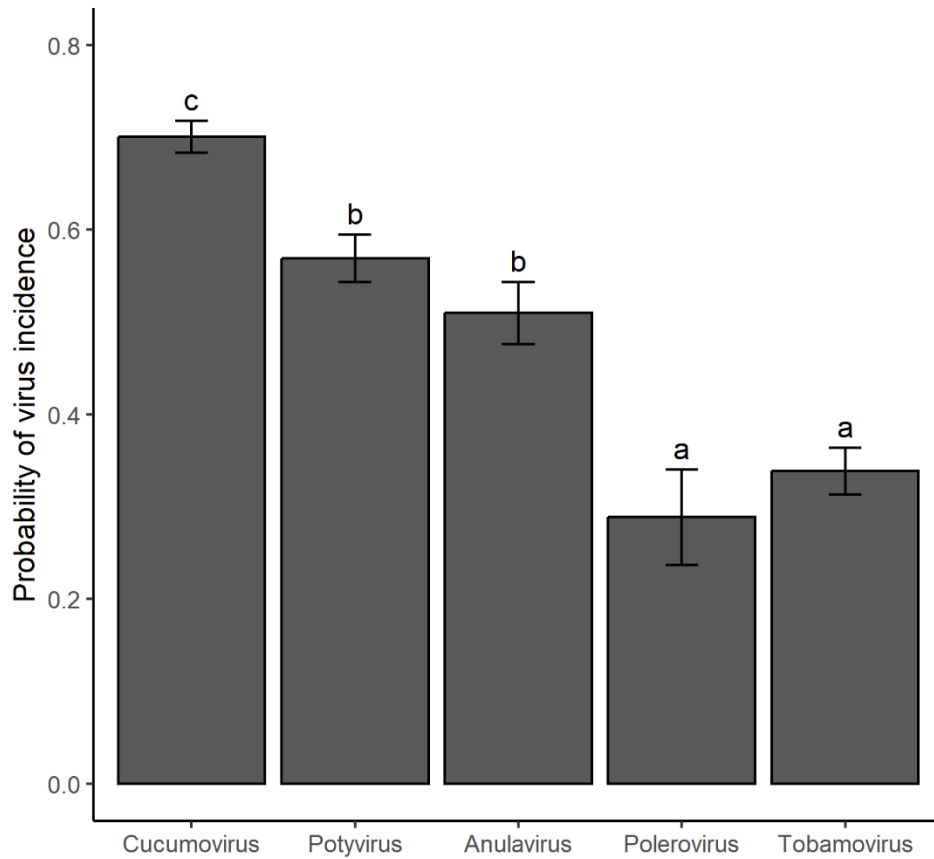


**Figure 4.19: Probability of *ih* of each virus genera.** Equal letter indicates difference was not significant.

The third GLM model considered habitat and mode of virus transmission as fixed effects with their interaction:  $ih \sim \text{habitat} + \text{virus transmission} + \text{habitat} * \text{virus transmission} + \varepsilon$  (Model 3). The model 3 was also statistically significant, ( $G^2_{28, 133}=309.33$ ,  $P < 0.001$ ), and explained 0.578 (Nagelkerke  $R^2$ ) of the variance in incidence. The model also shows *ih* varied according to the habitat ( $W_{3, 133} = 122.84$ ,  $P < 0.001$ ), virus transmission ( $W_{7, 133} = 136.15$ ,  $P < 0.001$ ), and their interaction ( $W_{7, 133} = 50.32$ ,  $P < 0.001$ ). Mode of virus transmission explained 45.9% of the variance of *ih*, habitat 31.8%, and their interaction 22.2%. Since the interaction between habitat and virus transmission was significant, new logistic regressions that consisted of virus transmission as a fixed factor was analyzed for each habitat. The model for Crop habitat showed that *ih* for virus

transmission was statistically significant ( $W_{3,19} = 68.26$ ,  $P < 0.001$ ), this was same for Edge ( $W_{3,62} = 72.64$ ,  $P < 0.001$ ), Oakwood ( $W_{2,24} = 25.71$ ,  $P < 0.001$ ) and Wasteland ( $W_{2,28} = 19.85$ ,  $P < 0.001$ ). Pairwise comparisons of transmission modes across habitats revealed significant differences in the odds ratios (OR). Viruses with non-persistent mode of transmission had significantly higher incidence than contact transmission in all four habitats —Crop (OR = 10.14,  $P < 0.001$ ), Edge (OR = 8.96,  $P < 0.001$ ), Oakwood (OR = 2.57,  $P < 0.001$ ) and Wasteland (OR = 2.37,  $P < 0.001$ ). Aphid non-persistently transmitted viruses also had significantly higher incidence compared to pollen transmitted viruses in Crop (OR = 9.63,  $P = 0.04$ ) and Oakwood (OR = 4.64,  $P = 0.003$ ) than aphid persistently transmitted viruses in the Edge (OR = 8.05,  $P < 0.001$ ). Pollen transmitted viruses had significantly higher incidence than viruses transmitted via contact and persistently by aphids in Edge (OR = 4.94,  $P < 0.001$ ). Aphid persistently transmitted viruses had significantly higher incidence than contact transmitted in Crop (OR = 10.14,  $P < 0.001$ ).

Post-hoc analyses of *ih* across habitat levels (Table A20, Annexes) showed no significant differences either between Crop and Edge, or between Oak and Wasteland. On the other hand, *ih* in Crop was significantly higher than in Oak and Wasteland. Similarly, *ih* in Edge was significantly higher than in Oak and Wasteland. This result shows a pattern of higher *ih* in anthropic than in wild habitats.

For the role of transmission modes on *ih*, post-hoc results (Table A21, Annexes) showed that *ih* of aphid non-persistent transmission was higher and significantly different from contact, aphid persistent or pollen transmission. Also, *ih* was significantly higher for pollen transmission than for contact and persistent transmission. No significant difference was found between persistent and contact transmission. Figure 4.20 shows the separate effects of transmission mode and habitat categories on *ih*.

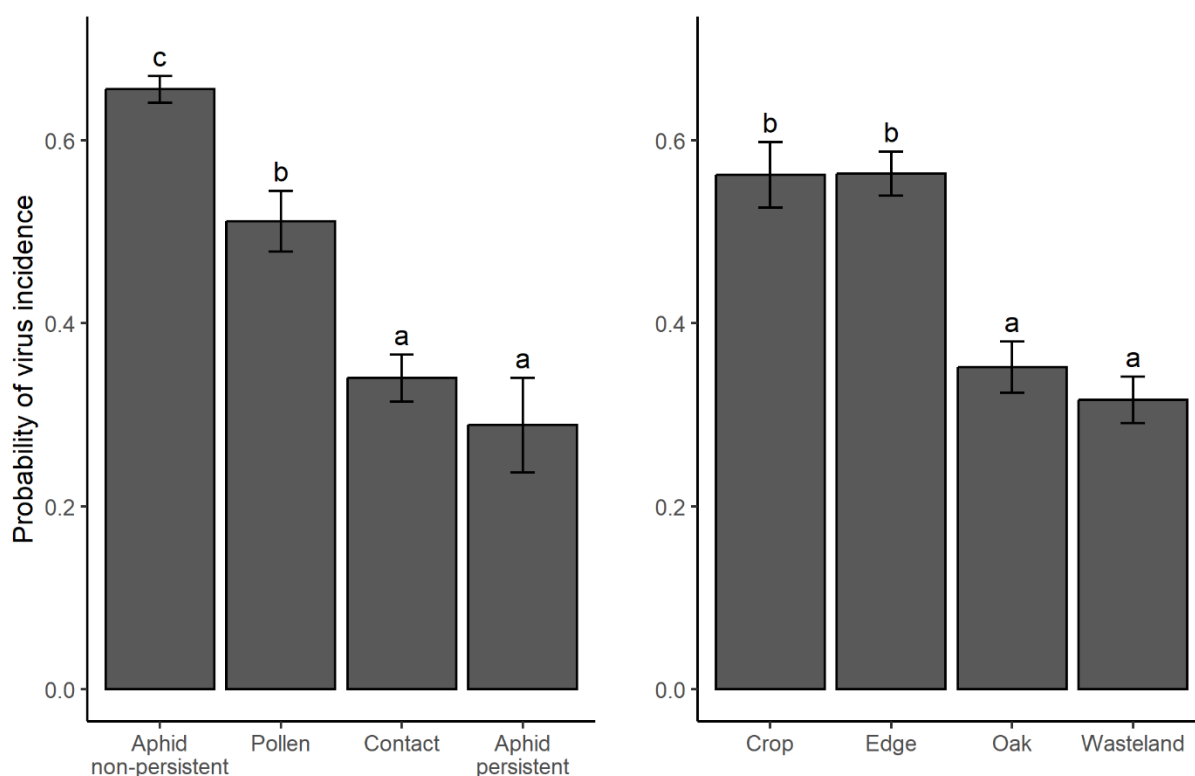


**Figure 4.20: Probability of *ih* with transmission and habitat considered separately.** Equal letter indicates difference was not significant.

Analyses of *ih* according to interaction between transmission and habitat indicate that probability of *ih* ( $P_{ih}$ ) is generally lower ( $<0.2$ ) in the non-anthropic Oak and Wasteland regardless of transmission mode and higher ( $>0.5$ ) in the anthropic Crop and Edge (Figure 4.19). Within the Crop habitat,  $P_{ih}$  is higher ( $>0.5$ ) with aphid persistent (BChV) and non-persistent viruses (CMV, TAV, PPV, TuMV, WMV) and lower ( $<0.2$ ) with pollen and contact transmitted viruses (PZSV and YoMV respectively) (Figure 4.21). The reverse is the case in Edge where  $P_{ih}$  is higher with pollen transmitted viruses and lower with non-persistent, persistent and contact transmitted viruses.

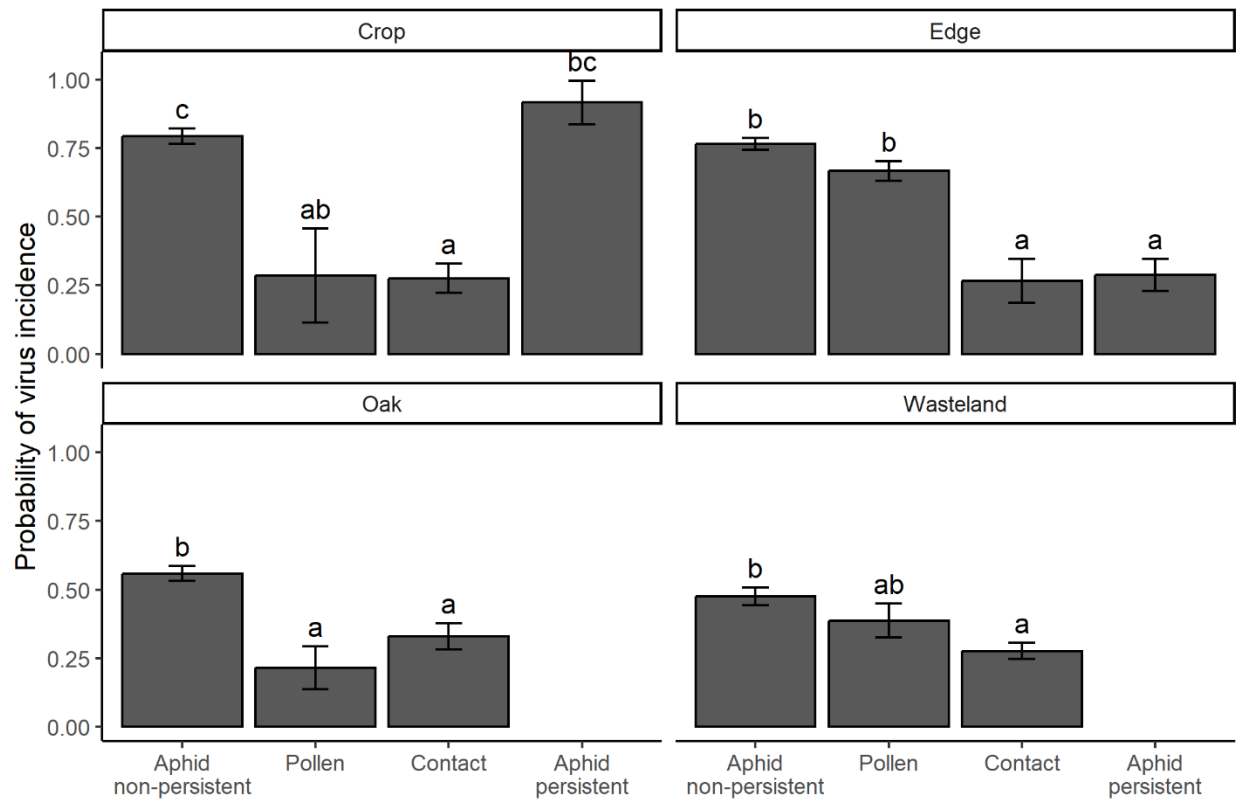


**Figure 4.21: Probability of *ih* based on interaction between transmission and habitat.** Equal letter indicates difference was not significant.

Similar to the analyses of incidence within observed hosts (*ih*), the effect of ecological or genetic factors on incidence in the whole plant community (*iL*) was also analysed. Data was aggregated according to the same determinant variables – virus species, virus genus and transmission mode, and a series of GLMs (logistic regressions) were performed with the determinant variables classed as factors and considered as fixed effects on *iL* (response variable).

The first GLM model considered habitat and virus species as fixed effects with their interaction:  $iL \sim \text{habitat} + \text{virus species} + \text{habitat} * \text{virus species} + \varepsilon$  (Model 1). Model 1 was statistically significant, ( $G^2_{28, 133} = 908.26$ ,  $P < 0.001$ ), and explained 0.716 (Nagelkerke  $R^2$ ) of the variance in incidence. The model also shows *iL* varied according to the habitat ( $W_{3, 133} = 215.30$ ,  $P < 0.001$ ), virus species ( $W_{7, 133} = 427.38$ ,  $P < 0.001$ ), and their interaction ( $W_{18, 133} = 265.59$ ,  $P < 0.001$ ). Virus

species explained 39.4% of the variance of *iL*, habitat 28.8%, and their interaction 31.8%. Since interaction between habitat and virus species was significant, again, new logistic regressions consisting of virus species as factor was analyzed for each habitat.

Pairwise comparisons of virus species across habitats revealed significant differences in the odds ratios (OR). In Crop and Edge habitats, OR of CMV and PZSV were consistently higher compared to other viruses in multiple comparisons— e.g CMV and PZSV exhibit higher odds ratios when compared to PPV (OR = 30.93) and BChV (OR = 32.93) respectively, both with significant p-values (<0.0001). A distinct pattern emerges in Oakwood and Wasteland where WMV and YoMV, show significant higher odds ratios compared to TuMV (OR=4.04, P=0.006) and PZSV (OR = 2.32, P=0.017) respectively. Across all four habitats, CMV consistently shows high odds ratios and significant p-values in comparison with other viruses like PPV, PZSV, and TuMV, highlighting its dominant likelihood of occurrence. PZSV follows closely, while WMV also shows significant contrasts, although with less odds ratios compared to CMV and PZSV.

For the role of virus species on *iL*, post-hoc analyses (Table 4.37, Annexes) show that *iL* of BChV was lower than for the other seven OTUs, same is true for PPV,. The *iL* of CMV was significantly higher than for the other seven viruses, the *iL* of PZSV and WMV were similar and higher than for the other four viruses, that of TAV was higher than for BChV and TuMV, and that of YoMV was higher than that of BChV. (Figure 4.22)

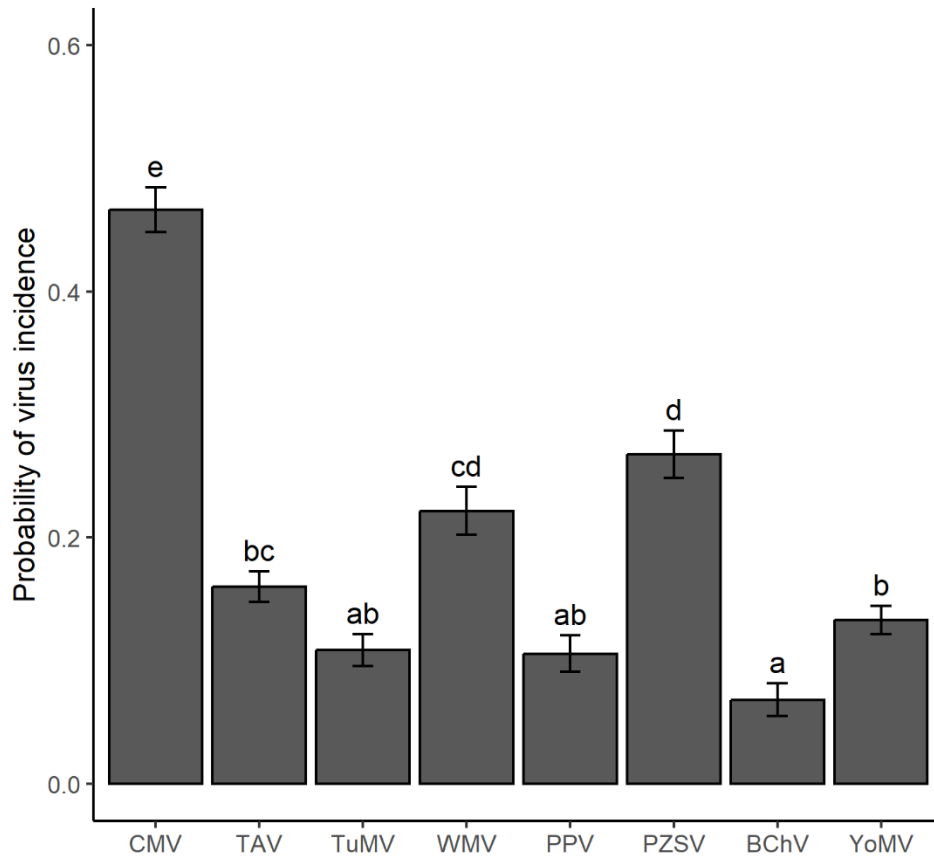


**Figure 4.22: Probability of *iL* of each virus species.** Equal letter indicates difference was not significant.

The second GLM considered habitat and virus genera as fixed effects with their interaction  $iL \sim \text{habitat} + \text{virus genera} + \text{habitat} * \text{virus genera} + \varepsilon$  (Model 2). Model 2 was significantly different from the null model, ( $G^2_{17,133} = 618.35$ ,  $P < 0.001$ ), and explained 0.384 (Nagelkerke  $R^2$ ) of the variances in incidence. The model also shows that *iL* varied according to the habitat ( $W_{3,133} = 215.30$ ,  $P < 0.001$ ), virus genus ( $W_{4,133} = 219.12$ ,  $P < 0.001$ ), and their interaction ( $W_{10,133} = 183.93$ ,  $P < 0.001$ ). Virus genus explained 33.2% of the variance of *iL*, habitat 32.2%, and their interaction 34.5%. Since interaction between habitat and virus genus was significant, new logistic regressions consisting of only virus genera as a factor was analyzed for each habitat. The model for Crop habitat showed that *iL* for virus genera was statistically significant ( $W_{4,19} = 58.74$ ,  $P < 0.001$ ), this

was same for Edge ( $W_{4,62} = 250.89$ ,  $P < 0.001$ ), Oakwood ( $W_{4,24} = 52.71$ ,  $P < 0.001$ ) and Wasteland ( $W_{4,28} = 40.71$ ,  $P < 0.001$ ).

Pairwise comparisons of virus genera across habitats revealed significant differences in the odds ratios (OR). Specifically, the OR of cucumovirus was higher than that of anulavirus in Crop (OR = 18.81,  $P < 0.001$ ), Edge (OR = 2.38,  $P < 0.001$ ), Oakwood (OR = 6.19,  $P < 0.001$ ) and Wasteland (OR = 2.54,  $P < 0.001$ ). The OR of cucumovirus was also significantly higher than that of tobamovirus in Crop (OR = 4.29,  $P < 0.001$ ), in Edge (OR = 9.51,  $P < 0.001$ ) and Oakwood (OR = 2.67,  $P < 0.001$ ). Additionally, the OR of potyvirus was higher than that of tobamovirus in Crop (OR = 3.19,  $P = 0.004$ ), and Edge (OR = 3.32,  $P < 0.001$ ).

For the role of virus genera on  $iL$ , post-hoc analyses (Table A23 Annexes) showed that  $iL$  was similar for anulavirus and cucumovirus and higher than that of tobamovirus, polerovirus and potyvirus ( $P < 0.001$  in each case), while the  $iL$  of potyvirus was significantly higher than that of polerovirus ( $P = 0.046$ ). These results show a pattern of relatively higher  $iL$  for virus genera with multipartite genome and a lower  $iL$  for the monopartite genera. Figure 4.23 shows that probability values of  $iL$  for all five genera was generally lower than  $ih$ , and that cucumovirus and anulavirus showed relatively high  $iL$ , ( $>0.25$ ) compared to polerovirus, potyvirus and tobamovirus, with  $iL$  of potyvirus and tobamovirus being similar (0.125).

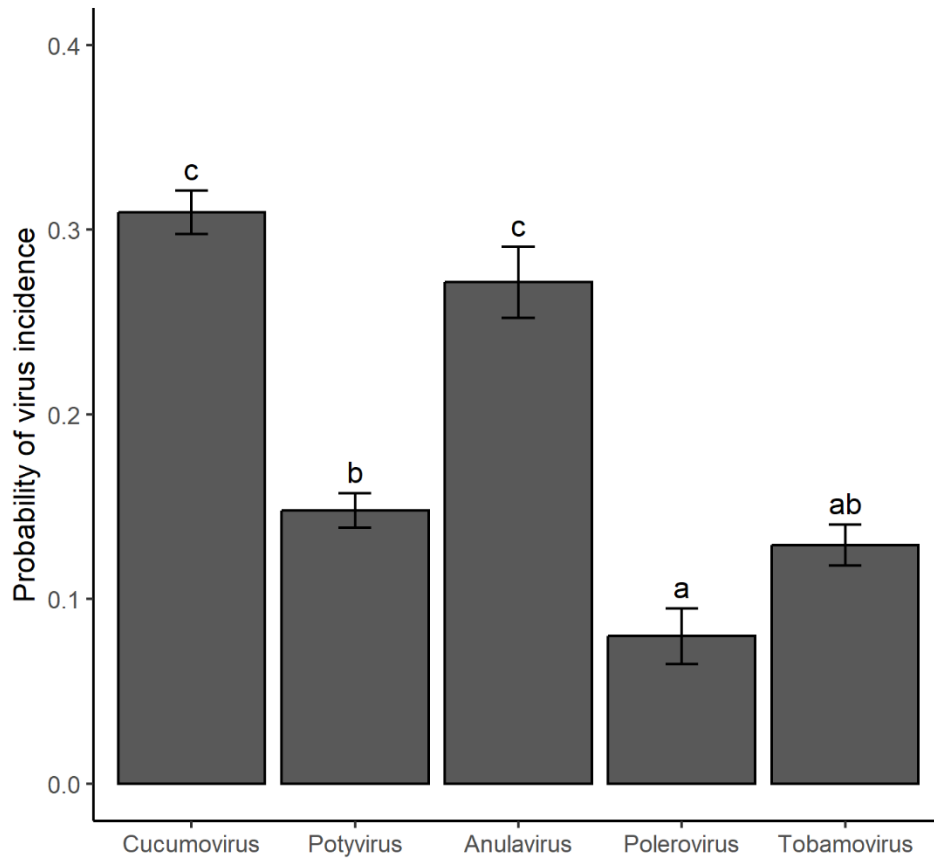


**Figure 4.23: Probability of *iL* of each virus genera.** Equal letter indicates difference was not significant.

The third GLM model considered habitat and mode of virus transmission as fixed effects with their interaction:  $iL \sim \text{habitat} + \text{virus transmission} + \text{habitat} * \text{virus transmission} + \varepsilon$  (Model 3). Model 3 was also statistically significant, ( $G^2_{13, 133} = 498.59$ ,  $P < 0.001$ ), and explained 0.317 (Nagelkerke  $R^2$ ) of the variance in incidence. The model also shows *iL* varied according to the habitat ( $W_{3, 133} = 215.30$ ,  $P < 0.001$ ), virus transmission ( $W_{3, 133} = 103.73$ ,  $P < 0.001$ ), and their interaction ( $W_{7, 133} = 179.56$ ,  $P < 0.001$ ). Mode of virus transmission explained 20.2% of the variance of *iL*, habitat 41.4%, and their interaction 38.4%. Since interaction between habitat and virus transmission was significant, new logistic regressions consisting of only virus transmission as fixed factor was analyzed for each habitat. The model for Crop habitat showed that *iL* for virus

transmission was statistically significant ( $W_{3,19} = 57.23$ ,  $P < 0.001$ ), this was same for Edge ( $W_{3,62} = 194.81$ ,  $P < 0.001$ ), Oakwood ( $W_{2,24} = 27.88$ ,  $P < 0.001$ ) and Wasteland ( $W_{2,28} = 3.37$ ,  $P < 0.001$ ).

Pairwise comparisons of transmission modes across habitats revealed significant differences in the odds ratios (OR). Viruses with aphid non-persistent transmission had significantly higher OR than contact transmitted ones in all habitats except Wasteland — Crop (OR = 3.93,  $P < 0.001$ ), Edge (OR = 6.11,  $P < 0.001$ ), Oakwood (OR = 1.94,  $P = 0.003$ ). Pollen transmitted viruses had significantly higher incidence than aphid non-persistent (OR = 17.17,  $P < 0.001$ ) or contact transmitted viruses in Crop (OR=22.58,  $P < 0.001$ ) and aphid persistently transmitted viruses in Edge (OR = 3.70,  $P = 0.003$ ). Last, aphid persistently transmitted viruses had significantly higher OR than contact transmitted in Crop (OR = 7.33,  $P = 0.003$ ).

These results highlight a pattern where contact transmission appears to be less prevalent compared to aphid non-persistent and aphid persistent modes in most habitats. Evidence suggests that pollen transmission is particularly favoured in anthropic habitats, with aphid non-persistent transmission being favored over contact and pollen-based transmission in wild habitats.

Post-hoc analyses of *iL* across habitat (Table A24, Annexes) indicates that *iL* was higher in Crop than in the other habitats, and in Edge than in Wasteland and Oakwood. *iL* did not differ between Wasteland and Oakwood. The odds ratios show that *iL* in Crop is four to five times higher than in Oak and Wasteland, while *iL* in Edge is two times higher than in Oak and Wasteland indicating an inverse relationship between the *iL* of anthropic and wild habitat groups (Figure 4.24).

Also, *iL*, post-hoc analyses (Table A25, Annexes) indicated that *iL* of aphid non-persistent viruses was 2 to 3 times higher *iL* than that of contact or aphid persistently transmitted viruses. Also, *iL* of pollen transmitted viruses were 2.5 and 4.5 times higher than that of contact and aphid persistently transmitted viruses respectively. Taken together, the results indicate that higher virus incidence is favoured by anthropic habitats and more likely first with pollen transmitted viruses and then aphid non-persistently transmitted viruses than with contact or aphid persistently transmitted viruses (Figure 4.24). Thus, transmission mode modulates infection risk of plant communities in a given habitat.

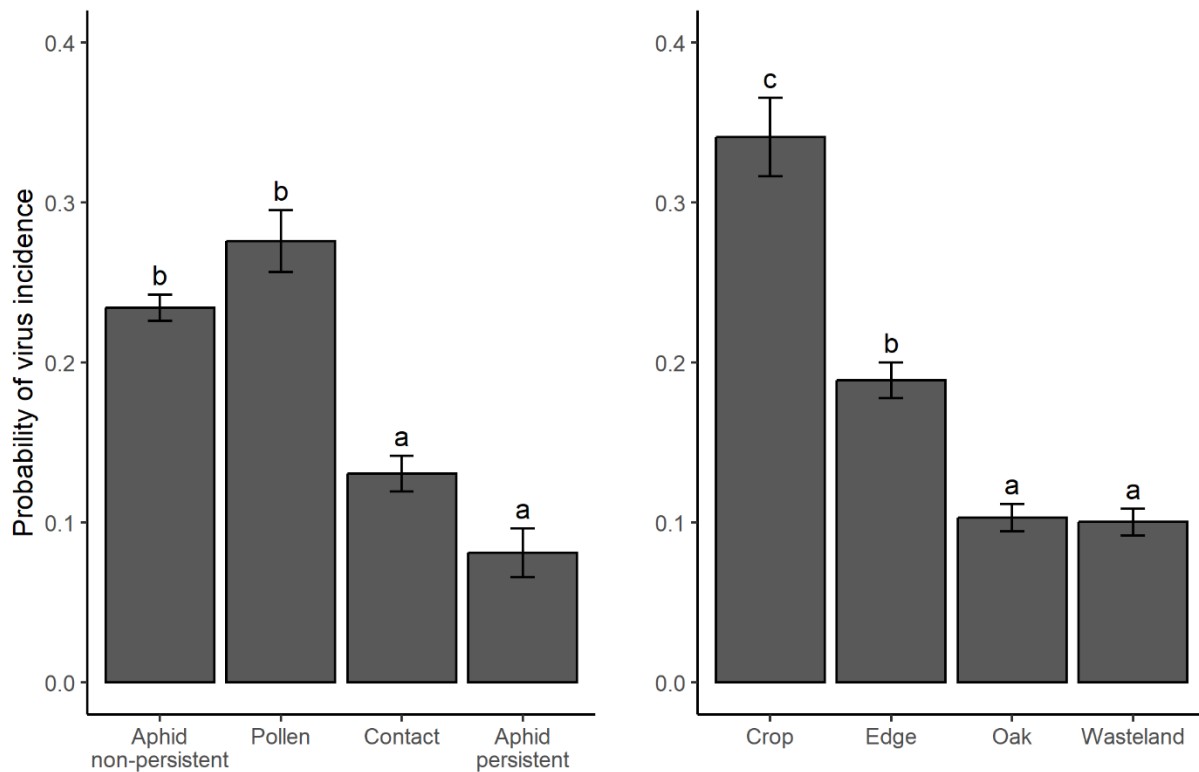


**Figure 4.24: Probability of *iL* with transmission and habitat considered separately.** Equal letter indicates difference was not significant.

Analyses of *iL* according to interaction between transmission and habitat suggests that probability of *iL* was generally lower ( $<0.25$ ) in wild habitats (Oakwood and Wasteland) regardless of transmission mode than ( $>0.5$ ) in anthropic habitats (Crop and Edge) (Figure 4.25). Within the Crop habitat *iL* is higher ( $\geq 0.5$ ) with aphid persistently (BChV) and non-persistently transmitted viruses, and lower ( $<0.2$ ) with pollen and contact transmitted viruses. In Edge *iL* was higher ( $>0.5$ ) with pollen transmitted viruses and lower with aphid non-persistent ( $\leq 0.25$ ), aphid persistent and contact transmitted viruses.

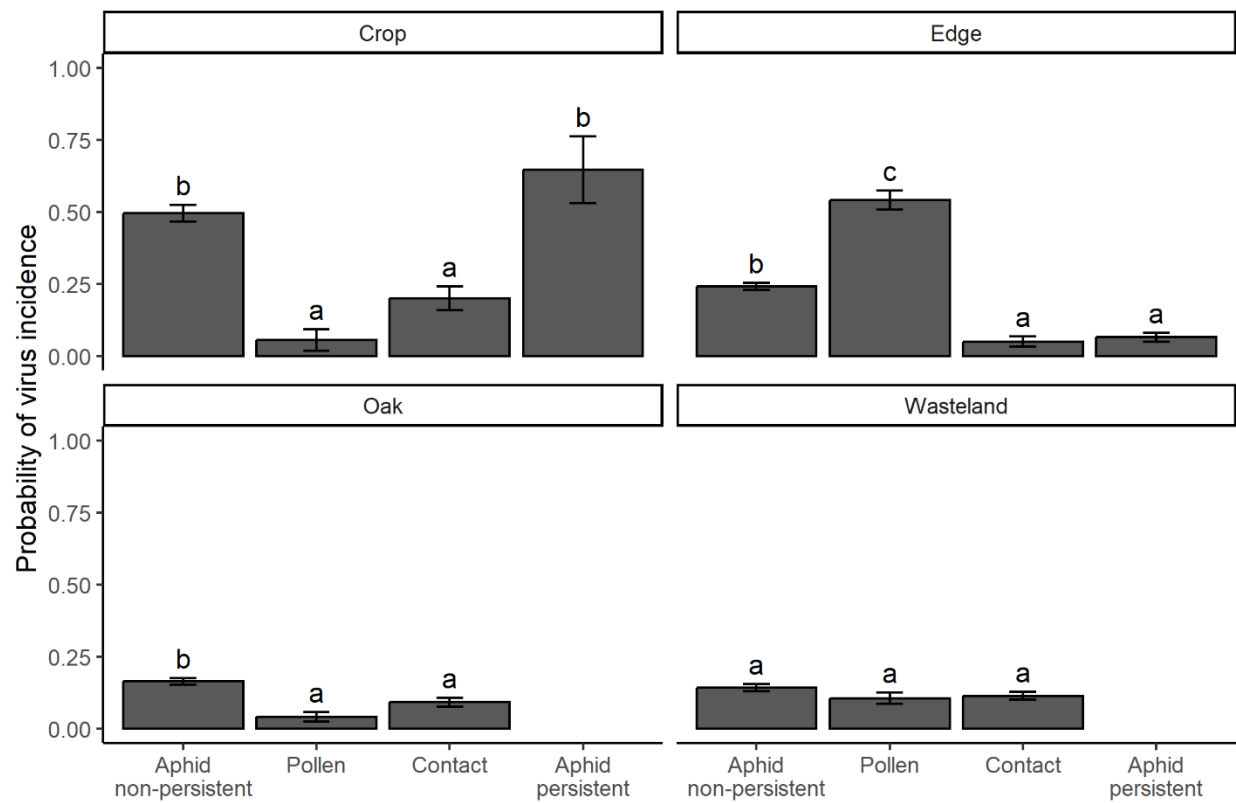


**Figure 4.25: Probability of *iL* based on interaction between transmission and habitat.** Equal letter indicates difference was not significant.

## 4.5 Genetic diversity analyses.

Analyses of genetic diversity of virus populations can provide insights into the nature of the interaction between viruses and hosts, and to whether there are adaptations which contribute to host range evolution of viruses, affecting their ability to infect multiple species of host and thus emerge in new host populations (Parrish et al. 2008). Thus, I performed nucleotide diversity analyses of viral genomes to provide insights into the mechanisms underlying the observed realised host ranges and the role of habitat in host range evolution.

Based on the BLAST queries of the HTS2C dataset, a list of unique read identifiers for each OTU was used to extract reads from the HTS libraries. The extracted reads were then aligned with their respective complete reference genome to obtain a binary alignment map (bam) file for each library. As genomic coverage differed across libraries for each virus, a limited number of HTS libraries per virus that produced suitable coverage depth (see Table A26 Annex) per genome were used for population genomic analyses. This was based on pilot analyses of CMV genetic diversity which showed that libraries with at least 10× coverage of the genome could be included (see Materials and Methods 3.3.5). Hence, diversity analyses were only performed for CMV, TAV and PZSV. Nucleotide diversity among habitats was estimated for CMV, while nucleotide diversity among selected sites was estimated for all three viruses. Variant calls were used to infer SNPs and assess nucleotide diversity for each virus genomic segment and population.

Genetic diversity of CMV among habitats was estimated with 38 libraries from all four habitats, while information from 28, 20 and 39 Edge libraries for CMV, TAV and PZSV respectively was used for comparative analyses at site level only.

Whole genome consensus sequences (WGCS) were obtained and selection of the best-fit substitution model for each alignment was done using the Akaike Information Criterion (AIC). Genetic diversity of the selected populations was estimated as the nucleotide diversity per site ( $\pi_T$ ) (Tajima, 1983). Within population genetic diversity in each WGCS data set was estimated by considering populations defined by each habitat category, or by the level of human management of the habitat (anthropisation category).

At the habitat level, CMV had  $\pi$  values that were similar across habitats and genomic segments, except for RNA3 which had a higher diversity in Crop and Wasteland than in Edge and Oakwood habitats (Table 4.8). Mean  $\pi$  at the site level (Table 4.9) was smaller than at habitat, except for RNA 3 at one of the Edge sites (L2).

**Table 4.8: Nucleotide diversity ( $\pi$ ) of CMV across habitats.**

| Segment                                  | Crop              | Edge              | Oakwood           | Wasteland         | Mean $\pm$ SE     |
|------------------------------------------|-------------------|-------------------|-------------------|-------------------|-------------------|
| RNA1                                     | 0.007             | 0.005             | 0.006             | 0.008             | 0.007 $\pm$ 0.001 |
| RNA2                                     | 0.007             | 0.006             | 0.004             | 0.007             | 0.006 $\pm$ 0.001 |
| RNA3                                     | 0.014             | 0.005             | 0.002             | 0.011             | 0.008 $\pm$ 0.005 |
| <b>Mean <math>\pm</math> SE <i>p</i></b> | 0.009 $\pm$ 0.004 | 0.005 $\pm$ 0.001 | 0.004 $\pm$ 0.002 | 0.009 $\pm$ 0.002 |                   |

**Table 4.9. Nucleotide diversity ( $\pi$ ) of CMV, TAV and PZSV at Edge sites.**

| Virus                                    | L1                | L2                | L3                | L4                | Mean $\pm$ SE     |
|------------------------------------------|-------------------|-------------------|-------------------|-------------------|-------------------|
| <b>CMV</b>                               |                   |                   |                   |                   |                   |
| RNA1                                     | 0.000             | 0.003             | 0.003             | 0.001             | 0.002 $\pm$ 0.001 |
| RNA2                                     | 0.001             | 0.004             | 0.002             | 0.001             | 0.002 $\pm$ 0.001 |
| RNA3                                     | 0.000             | 0.010             | 0.004             | 0.001             | 0.004 $\pm$ 0.003 |
| <b>Mean <math>\pm</math> SE <i>p</i></b> | 0.001 $\pm$ 0.000 | 0.006 $\pm$ 0.004 | 0.003 $\pm$ 0.001 | 0.001 $\pm$ 0.000 |                   |
| <b>TAV</b>                               |                   |                   |                   |                   |                   |
| RNA1                                     | 0.001             | 0.002             | 0.004             | 0.002             | 0.002 $\pm$ 0.001 |
| RNA2                                     | 0.002             | 0.004             | 0.003             | 0.003             | 0.003 $\pm$ 0.001 |
| RNA3                                     | 0.004             | 0.002             | 0.009             | 0.007             | 0.006 $\pm$ 0.003 |
| <b>Mean <math>\pm</math> SE <i>p</i></b> | 0.002 $\pm$ 0.002 | 0.003 $\pm$ 0.001 | 0.005 $\pm$ 0.003 | 0.004 $\pm$ 0.003 |                   |

| <b>PZSV</b>               |               |               |               |               |               |
|---------------------------|---------------|---------------|---------------|---------------|---------------|
| RNA1                      | n.a           | n.a           | 0.003         | n.a           | 0.003 ± 0.000 |
| RNA2                      | 0.006         | 0.010         | 0.006         | 0.004         | 0.007 ± 0.003 |
| RNA3                      | 0.005         | 0.005         | 0.003         | 0.005         | 0.005 ± 0.001 |
| <b>Mean ± SE <i>p</i></b> | 0.006 ± 0.001 | 0.008 ± 0.004 | 0.004 ± 0.002 | 0.005 ± 0.001 |               |

---

n.a. = libraries had insufficient information at this genomic segment.

---

For the three analysed viruses, CMV, TAV and PZSV, diversity at site level was similar, and consistently smaller for RNA1.  $\pi$  values were higher for RNA3 in CMV and TAV and for RNA2 in PZSV.

The sliding window analysis of the diversity along the genomic segments for each virus provided insights into the patterns of genetic variation within the viral genome, highlighting regions of high or low nucleotide diversity. Sliding window analysis of local nucleotide diversity ( $\pi$ ) across segments of CMV at the habitat and site level (Figure 4.26 and 4.27 respectively), and of TAV and PZSV at the level of Edge sites (Figure 4.28 and 4.29 respectively), indicated the highest diversity was observed in the RNA3 segment of CMV and the lowest in the RNA1 and RNA 3 segments of PZSV. Both CMV and TAV had higher nucleotide diversity in their RNA3 genome compared to the RNA1 and RNA2 segments. For CMV, nucleotide diversity trended in different ways across each segment for each habitat.

A Kruskal-Wallis test showed that ranks of  $\pi$  across the CMV RNA segments did not differ significantly for each habitat. The nucleotide diversity trends of TAV and PZSV across the segments for Edge sites was not as clear, particularly for RNA1 and RNA2. By comparison to that of CMV, the very low values of  $\pi$  in these instances indicate mutations at a very few sites. Together, similar trends of local variation in  $\pi$  across the genomic segments, in instances where nucleotide variation was not minimal, shows that habitat- and site-level factors do not correspond to variation in genetic diversity. The observed variation in  $\pi$  between habitats or sites of each virus may not be independent of sample size or extremely low genetic variation in general.

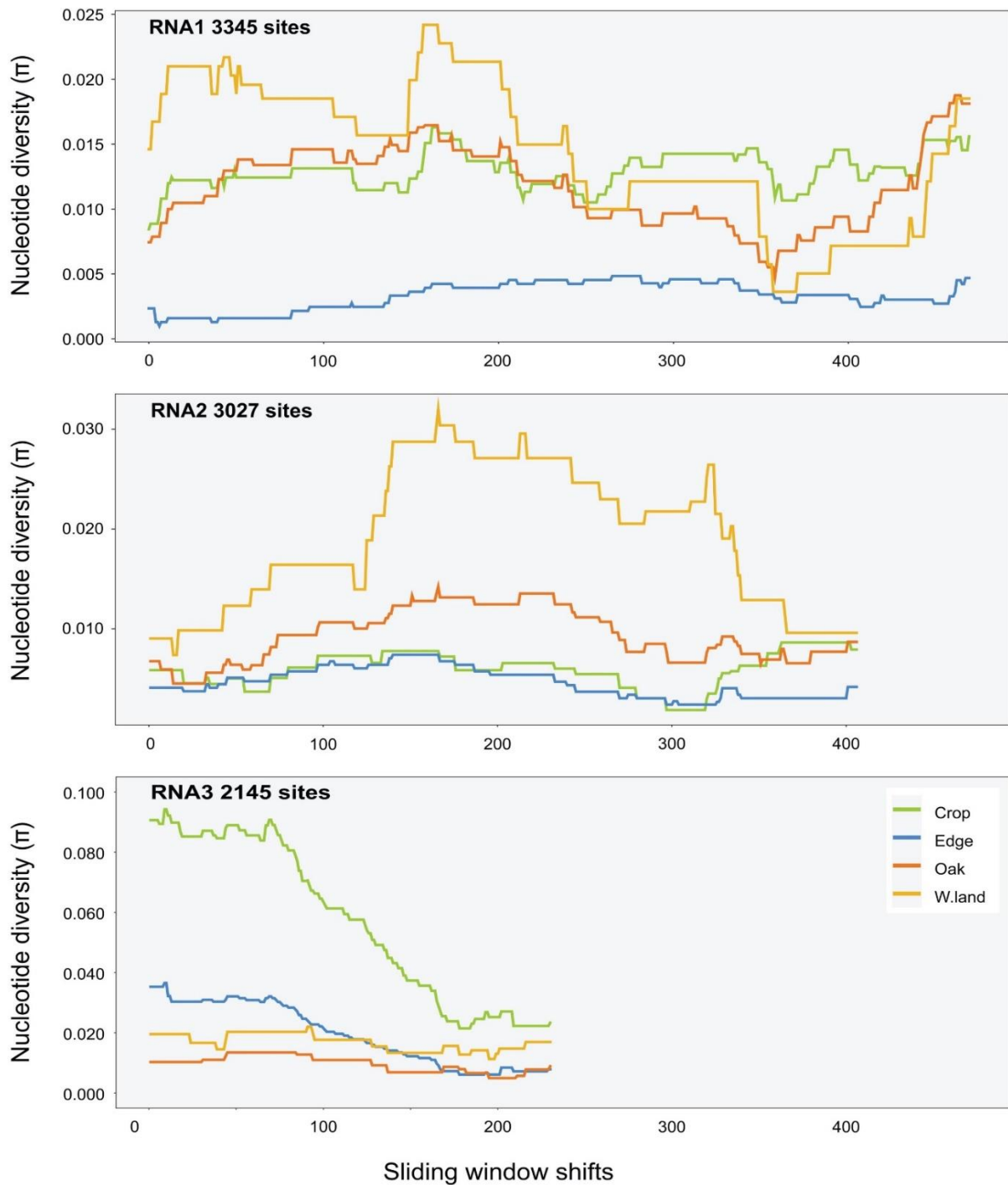


**Figure 4.26: Sliding windows analysis of nucleotide diversity ( $\pi$ ) across the consensus genomes of CMV segments at the habitat level (sliding window size was 1000 nt).**

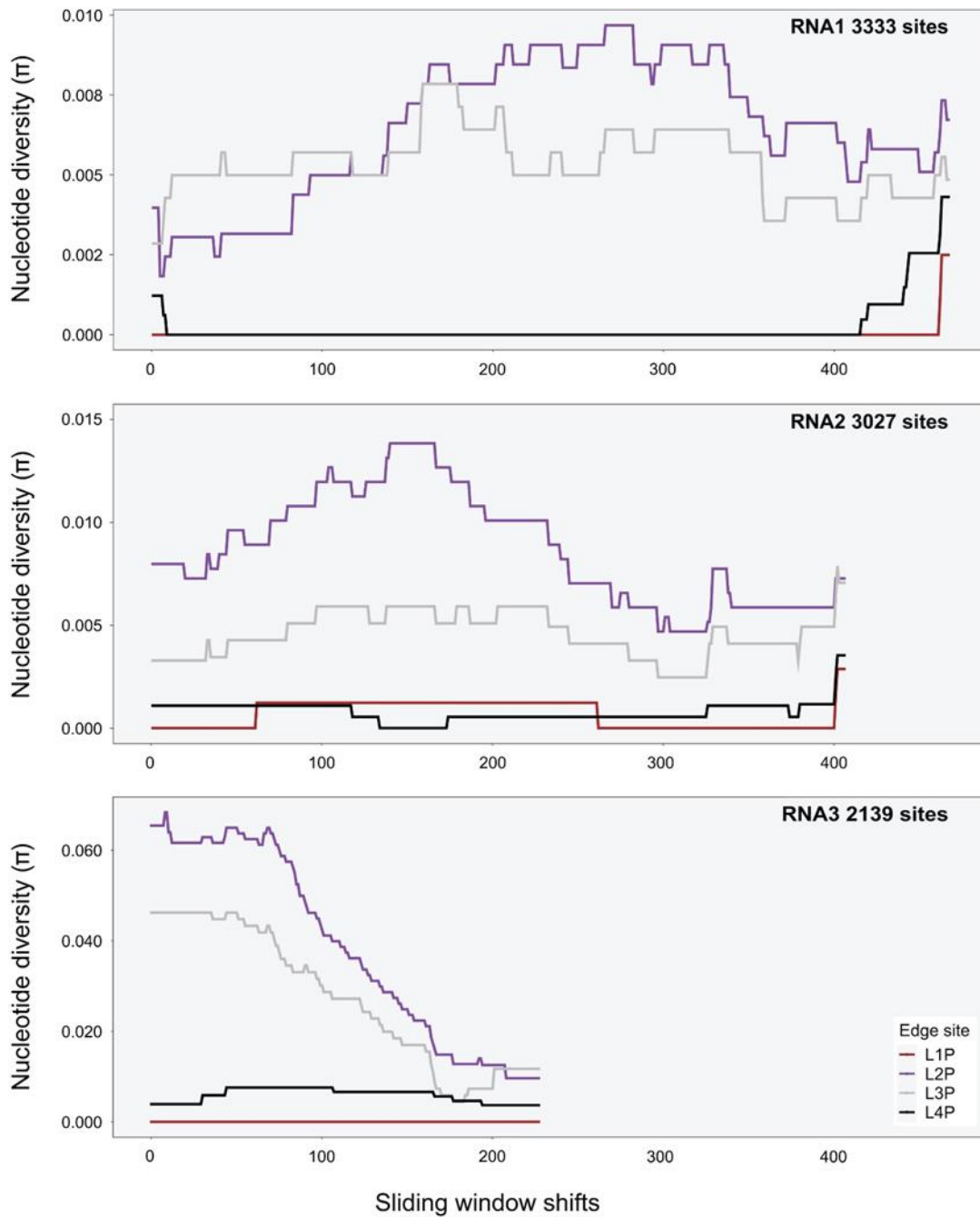


**Figure 4.7: Sliding windows analysis of nucleotide diversity ( $\pi$ ) across the consensus genomes of CMV segments at Edge sites (sliding window size was 1000 nt).**

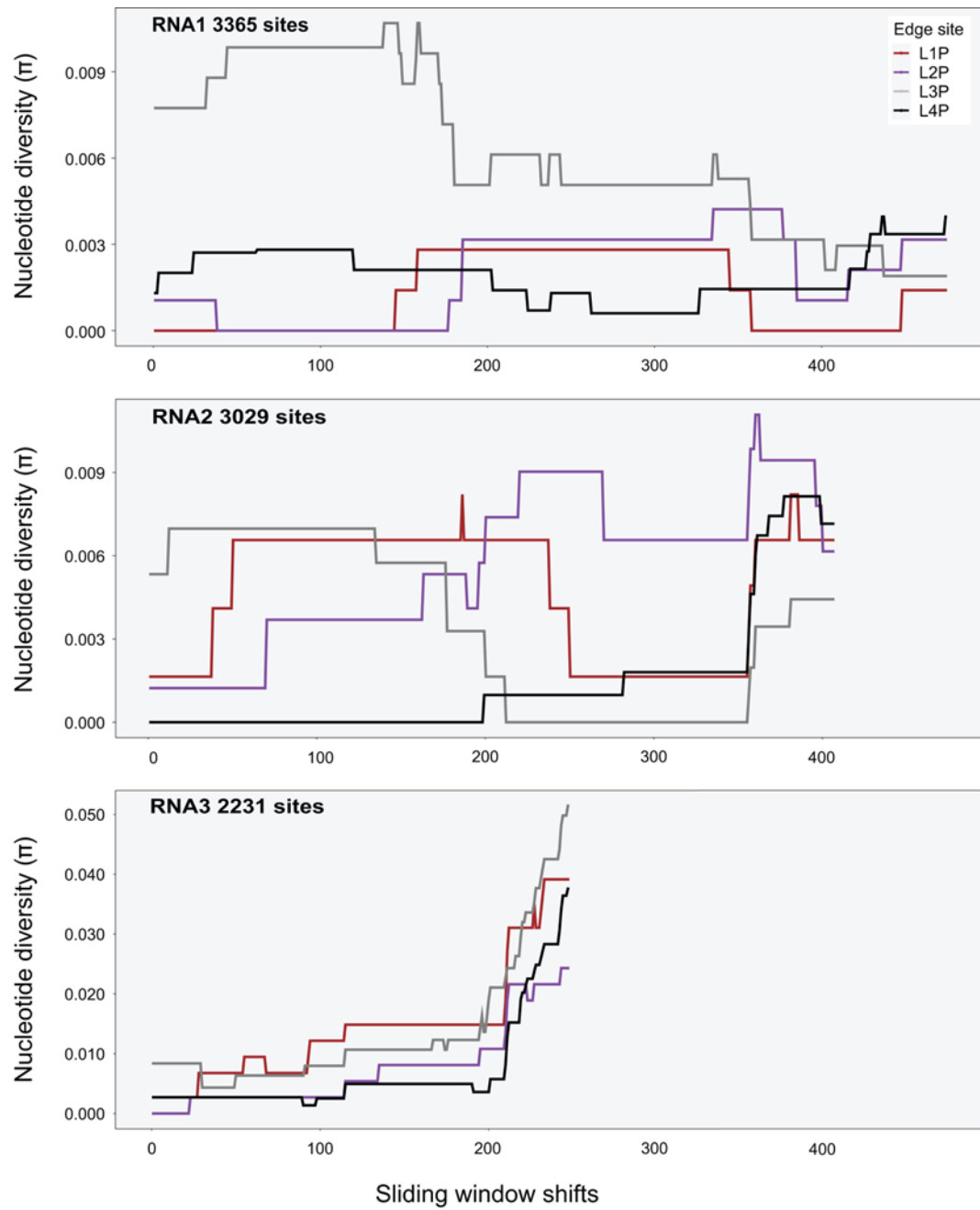


**Figure 4.28. Sliding windows analysis of nucleotide diversity ( $\pi$ ) across the consensus genomes of TAV segments at Edge sites (sliding window size was 1000 nt).**

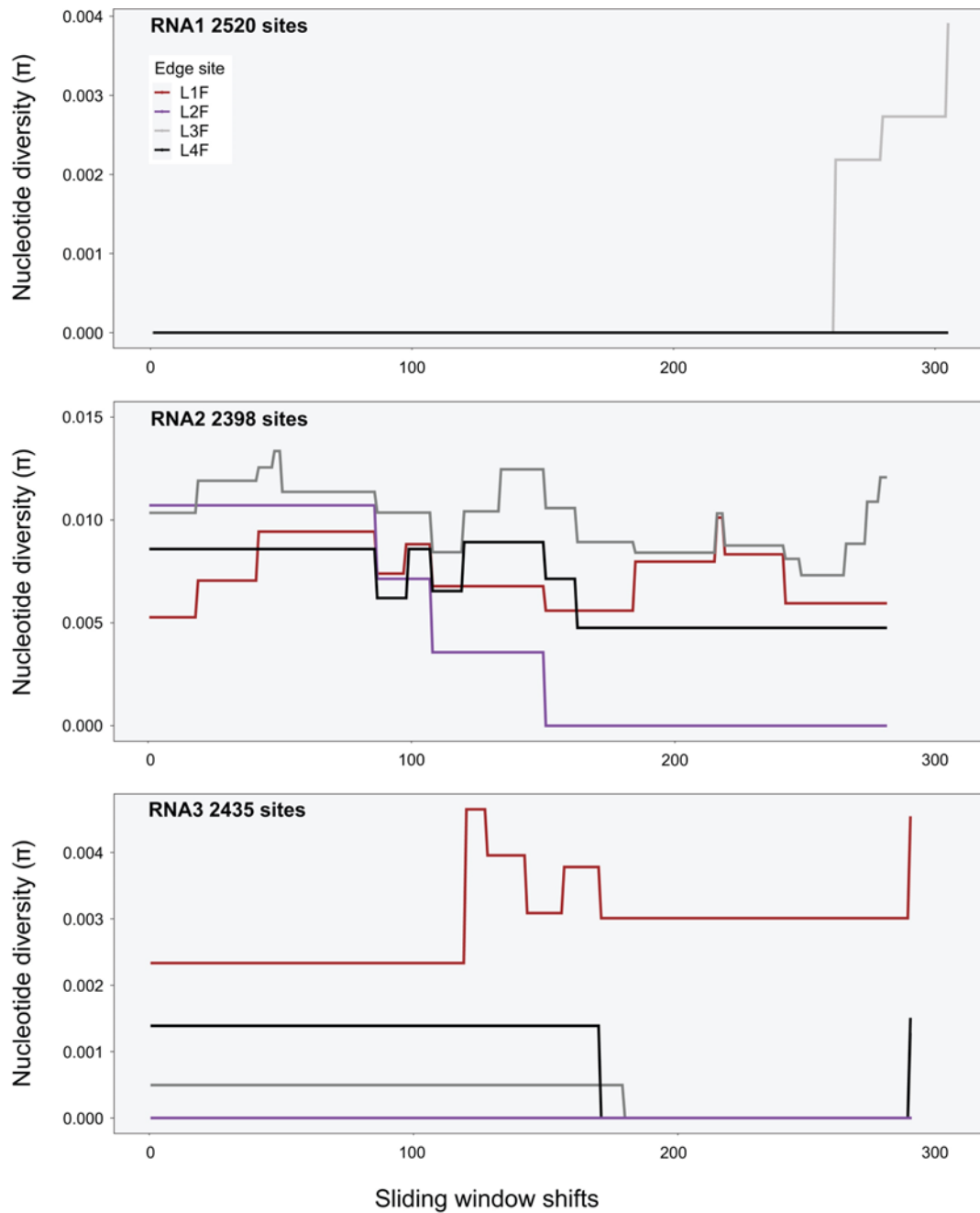


**Figure 4.29: Sliding windows analysis of nucleotide diversity ( $\pi$ ) across the consensus genomes of PZSV segments at Edge sites (sliding window size was 500 nt).**

#### 4.5.1 Phylogenetic analyses.

Phylogenetic analyses are important to reveal the genetic structure and differentiation among virus populations. By examining the branching patterns and genetic distances between populations inferred on the phylogenetic tree, the role of habitat or host species in shaping genetic variation, and population subdivision can be assessed. Phylogenies were inferred using the Maximum Likelihood method and the best fitted substitution model for each alignment. Whole genome consensus sequences from 38 libraries of CMV, were used to infer phylogenetic relationships among virus genotypes and their habitat affiliations, reads of TAV and PZSV could only be assembled for Edge sites.

For CMV, the general patterns of branching in the resulting phylogeny (Figure 4.30) was consistent with the clustering of genotypes by habitat, with most major clades including sequences from multiple host species from taxonomically distant families. In all three RNAs, isolates from Oakwood cluster separately from those of Edge plus Crop. The support of 0.9 or higher posterior probability for branching along the backbone of the phylogeny of each genomic segment, indicated Crop and Edge (and instances of Wasteland) isolates that arose from a variant that occurred in Oakwood. The next most derived clades are isolates that largely occurred in Crop and Edge. A very long branch in the RNA3 segment indicated that a relatively divergent clade of CMV infecting *C. arvensis* and *C. melo* were split between Edge and Crop habitats respectively.

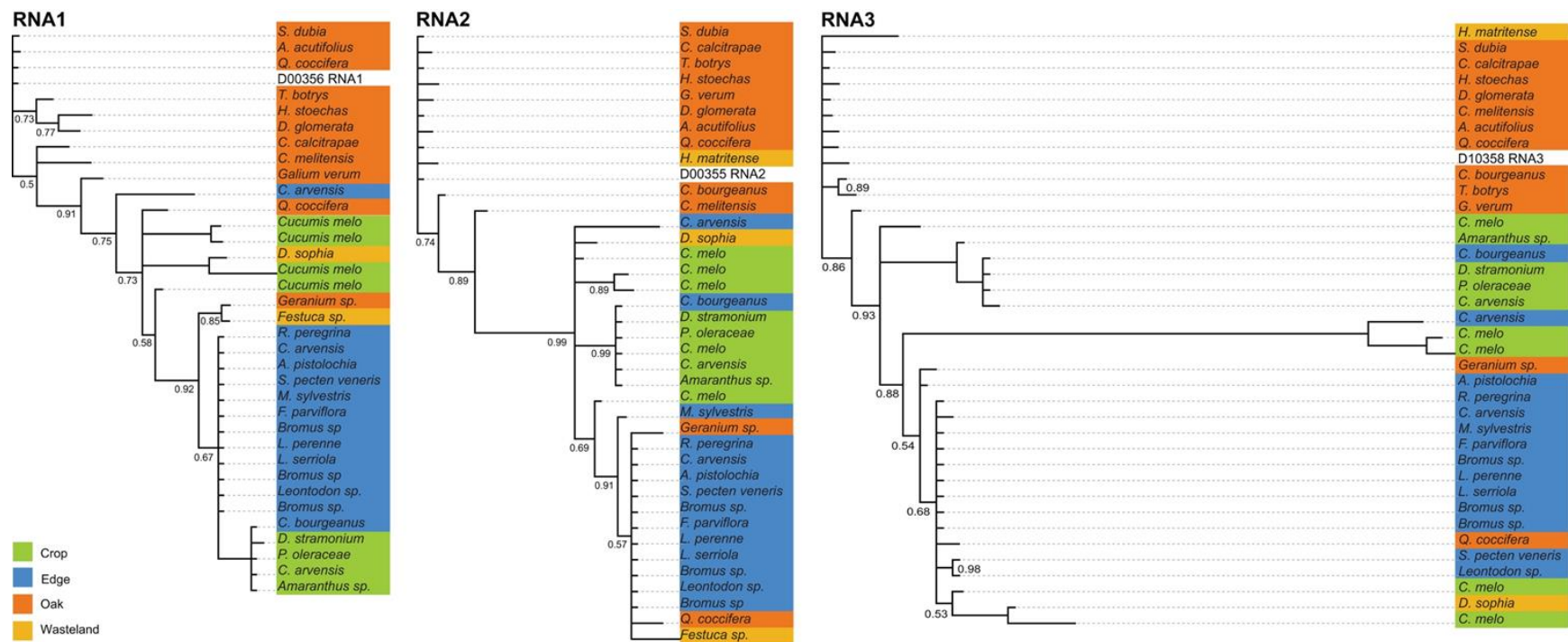


**Figure 4.30 Bayesian maximum credibility consensus phylogenetic inferences of the three genomic segments of CMV showing branches with 0.5 or higher posterior probability.** Source of genome segment according to habitat is indicated by colour: green for Crop, blue for Edge, red for Oakwood and yellow for wasteland.

Based on the virus and level of aggregation, different association indices were calculated by Bayesian Tip-association Significance tests (BaTS) (Parker et al. 2008) to analyze the association between the phylogenetic relationships of the viral populations with either host species, habitat, or degree of human management of the habitat traits. In the BaTS approach, the data are repeatedly simulated under the null hypothesis that characters at the tips are randomly distributed across the phylogeny. Theoretically, the resulting p-values follow a uniform distribution, and the type error of the test will be corrected for all levels of statistical significance (Parker et al. 2008). BaTS testing for all three RNAs of CMV showed significant ( $P < 0.0001$ , Table 4.10) associations that indicated a strong effect of anthropic disturbance, or habitat, on genetic diversity. The strong tip associations indicated by the AI statistic showed a high consistency among isolates sampled from the same habitat, which were descended from very few ancestral genotypes.

**Table 4.10: Bayesian Tip-association Significance tests (BaTS) of CMV populations at habitat.** Significant results are given in bold text. The ‘Anthropic effect’ grouped Edge and Crop to compare with Oakwood and Wasteland (i.e., non-anthropic habitats).

|                  | Statistic | Obs.<br>mean | L. 95%<br>CI | U. 95%<br>CI | Null<br>mean | L. 95%<br>CI | U. 95%<br>CI | p-value      |
|------------------|-----------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| <b>CMV RNA1</b>  |           |              |              |              |              |              |              |              |
| Habitat effect   | AI        | 0.97         | 0.56         | 1.38         | 2.46         | 2.03         | 2.85         | <b>0.000</b> |
|                  | PS        | 9.40         | 8.00         | 10.00        | 19.23        | 17.48        | 20.69        | <b>0.000</b> |
| Anthropic effect | AI        | 0.36         | 0.26         | 0.60         | 1.09         | 0.76         | 1.45         | <b>0.000</b> |
|                  | PS        | 3.00         | 3.00         | 3.00         | 7.39         | 5.91         | 8.39         | <b>0.000</b> |
| <b>CMV RNA2</b>  |           |              |              |              |              |              |              |              |
| Habitat effect   | AI        | 0.92         | 0.38         | 1.46         | 2.38         | 2.10         | 2.58         | <b>0.000</b> |
|                  | PS        | 9.05         | 9.00         | 9.00         | 20.34        | 18.38        | 21.61        | <b>0.000</b> |
| Anthropic effect | AI        | 0.44         | 0.06         | 0.77         | 1.18         | 0.93         | 1.38         | <b>0.000</b> |
|                  | PS        | 4.01         | 4.00         | 4.00         | 8.74         | 7.40         | 9.90         | <b>0.000</b> |
| <b>CMV RNA3</b>  |           |              |              |              |              |              |              |              |
| Habitat effect   | AI        | 1.17         | 0.71         | 1.62         | 2.54         | 2.08         | 2.98         | <b>0.000</b> |

|                  |    |      |      |       |       |       |       |              |
|------------------|----|------|------|-------|-------|-------|-------|--------------|
| Anthropic effect | PS | 9.43 | 8.00 | 10.00 | 19.56 | 17.83 | 21.16 | <b>0.000</b> |
|                  | AI | 0.58 | 0.22 | 0.91  | 1.21  | 0.82  | 1.54  | <b>0.010</b> |
|                  | PS | 4.36 | 4.00 | 5.00  | 8.77  | 7.08  | 10.00 | <b>0.000</b> |

Phylogeny of consensus genotypes of CMV within the Edge showed that genotypes were similar and did not cluster according to site or host species as shown in Figure 4.31. Indeed, BaTS for all three RNAs showed no significant effect of site or hosts in determining the genetic structure of CMV populations (Table 4.11). In addition to insignificant parsimony scores, the association index (AI) statistic that measures the imbalance of internal nodes of the phylogeny were also not significant.

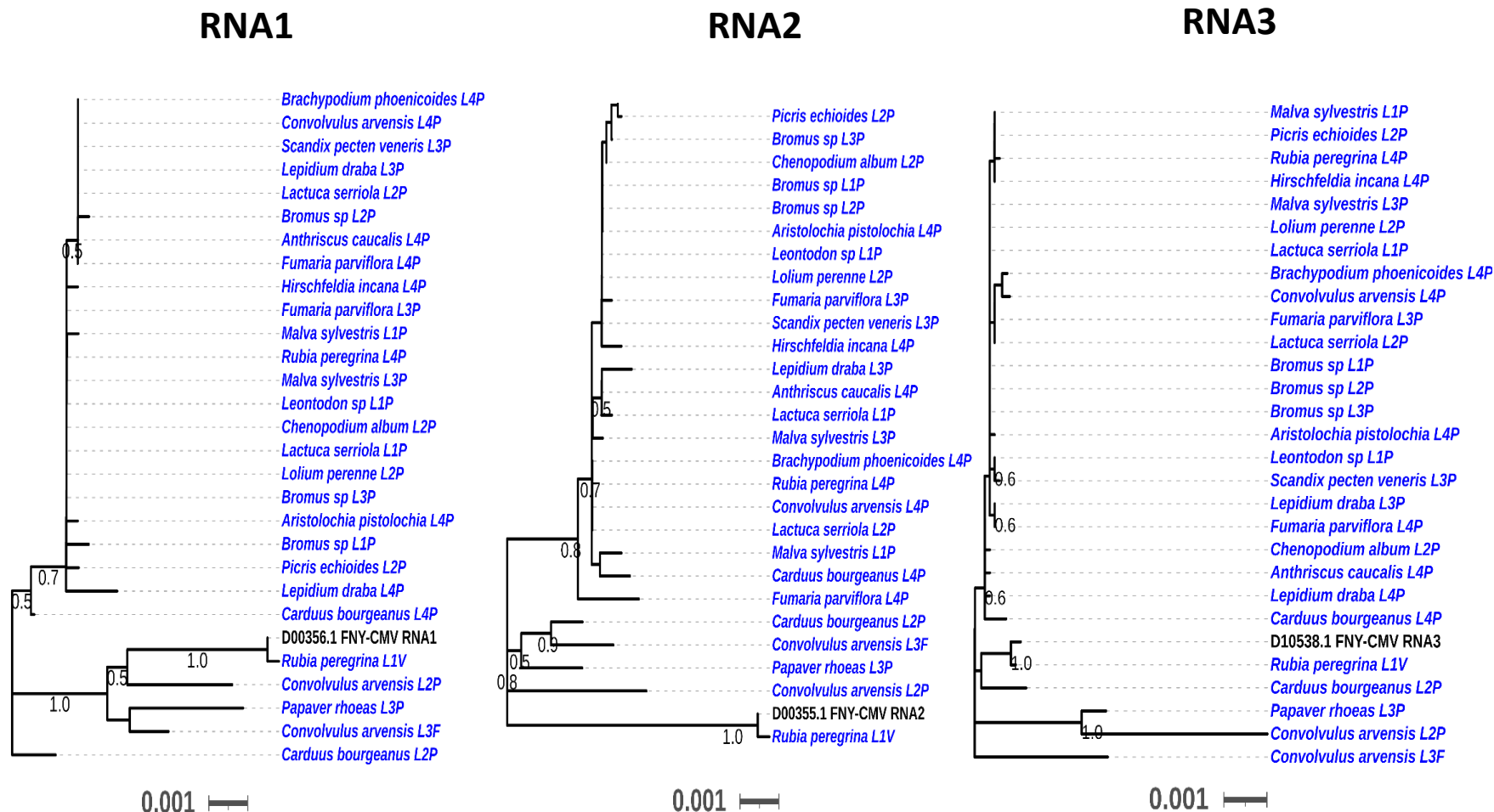


**Figure 4.31: Bayesian maximum credibility consensus phylogenetic inferences of the three genomic segments of CMV of Edge sites showing branches with 0.5 or higher posterior probability.**

**Table 4.11: Bayesian Tip-association Significance tests (BaTS) of CMV populations at Edge for the strength of site or host.**

|                 | Statistic | Obs.<br>mean | L. 95%<br>CI | U. 95%<br>CI | Null<br>mean | L. 95%<br>CI | U. 95%<br>CI | <i>p</i> -value |
|-----------------|-----------|--------------|--------------|--------------|--------------|--------------|--------------|-----------------|
| <b>CMV RNA1</b> |           |              |              |              |              |              |              |                 |
| Site effect     | AI        | 1.89         | 1.24         | 2.46         | 1.91         | 1.64         | 2.13         | 0.380           |
|                 | PS        | 13.29        | 12           | 15           | 13.78        | 12.85        | 14.63        | 0.090           |
| Host effect     | AI        | 2.54         | 2.01         | 3.06         | 2.50         | 2.33         | 2.56         | 0.960           |
|                 | PS        | 21.46        | 20           | 23           | 21.04        | 20.45        | 21.54        | 1               |
| <b>CMV RNA2</b> |           |              |              |              |              |              |              |                 |
| Site effect     | AI        | 1.89         | 1.24         | 2.46         | 1.88         | 1.59         | 2.13         | 0.450           |
|                 | PS        | 13.29        | 12           | 15           | 13.64        | 12.50        | 14.57        | 0.160           |
| Host effect     | AI        | 2.54         | 2.01         | 3.06         | 2.52         | 0.93         | 2.55         | 0.950           |
|                 | PS        | 21.46        | 20           | 23           | 21.04        | 20.45        | 21.37        | 1               |
| <b>CMV RNA3</b> |           |              |              |              |              |              |              |                 |
| Site effect     | AI        | 2.04         | 1.49         | 2.61         | 1.92         | 1.55         | 2.23         | 0.700           |
|                 | PS        | 13.84        | 12           | 15           | 13.60        | 12.09        | 14.63        | 0.680           |
| Host effect     | AI        | 2.57         | 2.06         | 3.06         | 2.54         | 2.29         | 2.62         | 0.5             |
|                 | PS        | 21.37        | 20           | 23           | 21.02        | 20.10        | 21.632       | 0.400           |

The consensus genotypes of TAV and PZSV within the Edge did neither cluster according to site or host plant as shown in **Figures 4.32 to 4.33**. Again, BaTS analyses for genotypes of TAV and PZSV within the Edge showed no significant association to site (when genotypes were assigned to their respective sites) or host species (when genotypes were grouped by plant taxon) as AI and PS had *p*-values > 0.05 (Tables A27 and A28 in Annex). Although an exception was observed in PZSV RNA3 where AI and PS had significant *p*-values for host effect, results largely suggests that there was no overall effect of site or host in determining genetic structure of the genotypes of CMV, TAV and PZSV within the Edge sites.

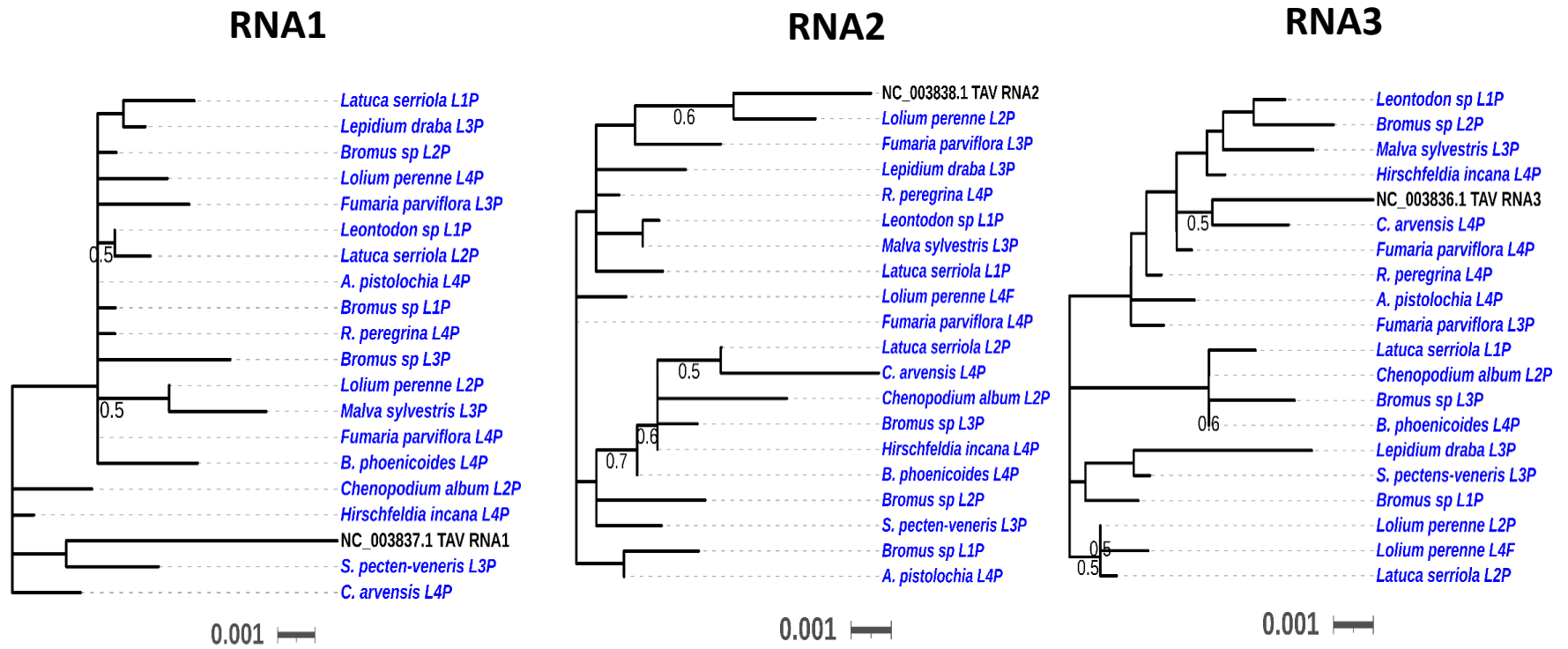


Figure 4.32: Bayesian maximum credibility consensus phylogenetic inferences of the three genomic segments of TAV of Edge sites showing branches with 0.5 or higher posterior probability.

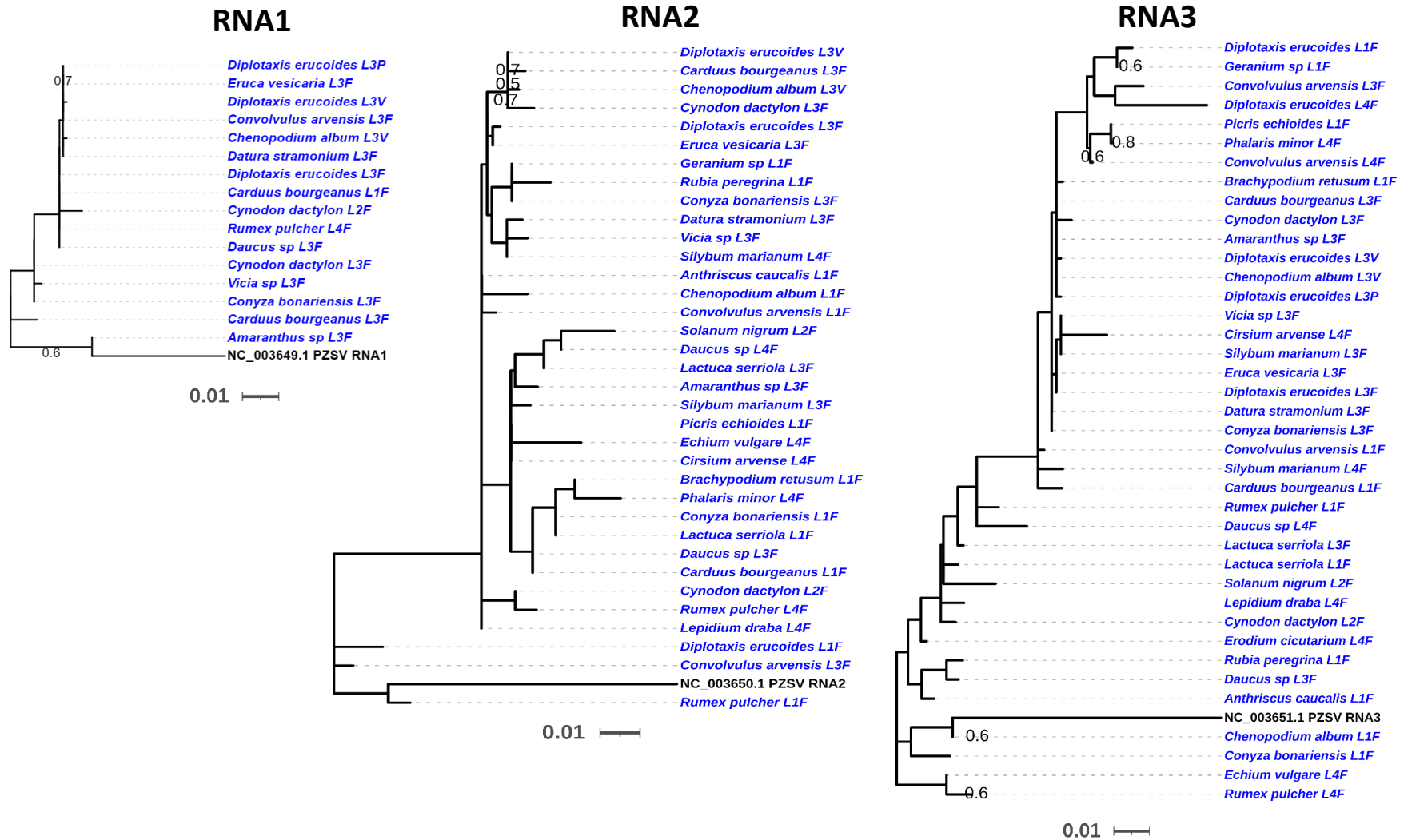


Figure 4.33: Bayesian maximum credibility consensus phylogenetic inferences of the three genomic segments of PZSV of Edge sites showing branches with 0.5 or higher posterior probability.

## 5. DISCUSSION.

Emerging viruses are an important socio-economic concern that necessitates studies to understand processes that cause emergence (Elena et al. 2014). High incidence of emergent viruses in agriculture impact plant health and productivity (Roossinck et al. 2015, Rivarez et al. 2023; Maksimović et al. 2024). Virus emergence is dependent on a range of factors intrinsic to the virus, such as genetic traits that determine its fitness in different hosts (i.e., evolutionary factors), as well as extrinsic factors (Jones, 2009; McLeish et al. 2019) that allow the virus to encounter and infect a new host population (i.e., ecological factors). Host range is an important factor in the emergence of plant viruses (Woolhouse and Gowtage-Sequeria 2005). Host range evolution is a fundamental aspect of emergence and has received attention in the recent past (Bedhomme et al. 2012; 2015; Elena, 2017; Fraile & García-Arenal 2010; McLeish et al. 2019; Seabloom et al. 2015). However, the focus on evolutionary factors that underlie emergence has been at the expense of knowledge about the effect of ecological factors on host range (Roossinck & García-Arenal, 2015, McLeish et al. 2019). In thesis, we aim at decreasing this gap in understanding by comparing the virome of wild and cultivated plant communities, determining host range and incidence of generalist viruses and analysing co-infections to uncover virus association patterns. Furthermore, the genetic diversity of virus populations in relation to host and habitat have been explored to gain knowledge for preventing or managing virus emergence across different environments.

The host range of plant viruses and other pathogens has been used for decades as a criterion for taxonomic purposes, and is typically estimated with inoculation experiments, especially for viruses (Dijkstra 1992). A virus' fundamental host range is generally unknown, that is, a virus' full potential to infect all suitable hosts. However, host range in natural settings is a plastic trait subject to the availability of suitable hosts that manifests as a realised host range – that is, a subset of the fundamental host range that a virus can successfully infect (Wells and Clark 2019) and is influenced by both intrinsic genetic traits of the virus and extrinsic factors such as habitat, species interactions, and ecological complexity (McLeish et al. 2019). Ecological factors (the distribution, abundance, and interaction of species) shape the diversity of species that a virus

comes into contact with (Arnold and Lutzoni, 2007) and so influences host range evolution. A virus' host range expands (host breadth expansion), or contracts (host breadth contraction) in response to ecological changes (Janz et al. 2001; Longdon et al. 2014; Dennehy, 2016). Furthermore, the ecology of plant viruses in alternative hosts other than crops has been recognised as relevant to understanding and predicting virus epidemics in crops (Duffus, 1971), and the biology of plant viruses in wild hosts is the object of much recent interest (García-Arenal & Zerbini, 2019; Maachi et al. 2022).

Metagenomics uses high-throughput sequencing (HTS) technology to generate large datasets, enabling the study of entire viral communities within plant populations. This high-throughput capability is essential for identifying the broad diversity of viruses across multiple plant species and environments simultaneously. Metagenomic studies are providing unprecedented insights into virus diversity (Roossinck, 2012; Rodríguez-Negrete et al. 2019, Khalili et al, 2023) and host range evolution (Babalola et al 2023; Zamfir et al. 2023) with plants in natural communities. Also important is the study of virus-virus interactions (i.e co-occurrence of virus species) within viral communities. Co-occurrence is pivotal in understanding virus evolution and resource use as it provides insights into the interactions between viruses and their hosts, as well as the ecological factors that influence these relationships (Araújo et al. 2011; McLeish et al. 2022). For example, Norberg et al. 2021 show that some virus-virus associations are mediated by either host or habitat characteristics, or the spatial structure of the host populations, suggesting that environmental heterogeneity also plays a significant role in shaping co-occurrence among viral species (Norberg et al. 2021). Therefore, the goal of this thesis was to incorporate an HTS approach with traditional techniques to study virus-host interactions across heterogeneous anthropic and non-anthropic ecosystems to better understand host range and host use of viruses. Single- and multiple-host systems were studied to gather insights into host range, virus occurrence and incidence, and population genetic structure across communities of different habitats.

## 5.1 The virome of a crop and its wild relative.

Spill-over and spill-back events between managed and unmanaged plant communities can directly influence the emergence of viruses (Roossinck and García-Arenal 2015). An HTS approach was used in a comparative study of wild and cultivated carrots to analyse virome structure and gene flow across populations. Populations of either wild carrot (*Daucus carota* ssp. *carota*) or cultivated carrot (*Daucus carota* ssp. *sativus*) were used to study ecological interactions as the two subspecies are expected to present a weak genetic barrier to virus transmission between them. Carrots were sampled in 45 populations from France (23 wild, 15 cultivated and 7 off-type) and 6 populations from Spain (1 wild and 5 cultivated). Samples were subjected to HTS of double-stranded RNA (dsRNA) extracts to capture all viruses with RNA genomes (Marais et al. 2018; Ma et al. 2020). A total of 45 virus OTUs were detected of which 20 (44%) belonged to already known species and 25 (56%) had low sequence similarity with the reference genomes Blast searches and were considered novel or tentatively novel according to the ICTV species demarcation criteria for each virus genus or family.

The 45 virus OTUs were dominated by (+) ssRNA viruses, followed by dsRNA viruses. Known viruses of the genera *Polerovirus* (family *Solemoviridae*), *Umbravirus* (family *Tombusviridae*), *Closterovirus* (family *Closteroviridae*), *Chordovirus* (family *Betaflexiviridae*), etc mostly aphid transmitted were detected both in France and Spain. The known viruses were mostly associated with cultivated carrots with a relatively small proportion (<1 %) of reads when detected in wild carrots. Novel viruses from the *Partitiviridae* family, e.g betapartitivirus (DiCV2) were detected from a higher proportion of wild carrot populations (39%) than cultivated (8%) or off type (14%), which indicated a role for wild carrots as a reservoir of viral diversity.

The viromes of wild and cultivated carrots differed in virus richness and in the viral species detected with higher virus richness in wild populations compared to cultivated populations. This suggests that wild carrots represent a reservoir of viral diversity from which novel viruses infecting carrot crops could potentially emerge. It also suggests that despite the close taxonomic relationship between cultivated and wild carrots, host adaptation or other unforeseen barriers

may limit the flow of viruses between plants belonging to these two subspecies, even at close spatial proximity.

Besides differences in virome composition, further analyses of the French carrot populations by Schönegger et al. (2023) revealed strong differences in the incidence of aphid transmitted viruses between field (that is, cultivated and off type) and wild populations. The incidence of known aphid-transmitted viruses was much higher in cultivated carrots than in wild carrots. These results are in line with other studies at landscape level showing higher virus incidence in managed or cultivated areas compared to less managed or wild areas (Pagán et al. 2012; Bernardo et al. 2018). The degree of human management has also been shown to impact aphid abundance and diversity (Al Hassan et al. 2013). Claflin et al. (2016) observed that crop dominated landscapes show a higher incidence of the vector-transmitted potato virus Y and that aphid abundance is positively correlated with virus incidence.

Additionally, the viruses of the wild and cultivated carrot populations proved to be especially rich in known and novel viruses involved in complex interactions of mutual interdependencies for aphid transmission for example, between polero- and umbraviruses. This result shows that the carrot model provides a particularly interesting pathosystem for further investigations of the poorly studied interplay between co-infecting helper and dependent viruses and of the evolutionary processes leading to synergistic or antagonistic virus-vector-host relationships in natural field conditions.

Due to the difference in sampling size, no absolute conclusions can be drawn regarding the differences in viral richness between France and Spain. However, most novel viruses (92%, 23/25) were discovered in French carrot populations especially with regard to the Nouvelle-Aquitaine region which is the main French carrot production area from where most populations were sampled. Notwithstanding, in both countries the virome of cultivated carrots was clearly distinct from that of wild carrots. Two kinds of differences were identified: first, some viruses showed differential presence as they were either exclusively or predominantly present in wild or in the cultivated compartments and second, remarkable differences in incidence level could be identified for some viruses between wild and cultivated carrots, as determined by HTS and RT-PCR analysis of individual plants within the French populations. Concerning the first point,

differentially present viruses were mostly (but not exclusively) persistent viruses such as carrot cryptic virus, which was identified in most field carrot populations but in only one wild population, or novel deltapartitiviruses and amalgaviruses found predominantly in wild carrots. These persistent viruses do not have any known vectors and are believed to be transmitted through seeds or pollen (Roossinck, 2012a), so that their differential presence can be understood as highlighting the limited genetic fluxes between cultivated and wild carrot populations. Concerning differences in prevalence of individual viruses, known aphid-transmitted viruses showed very high prevalence in field carrots while novel viruses, also expected to be aphid-transmitted showed a symmetrical pattern, with intermediate levels in wild carrots but only very sporadic detection in cultivated carrots.

## 5.2 A multi-host-multi-virus study.

### 5.2.1 Virus-Host detection by HTS and RT-PCR

For the study of multiple host- multiple virus interactions I focused on eight virus species of five different genera *Anulavirus*, *Cucumovirus*, *Potyvirus*, *Tobamovirus* and *Polerovirus* chosen among those virus OTUs detected in a previous HTS study. The criteria for the selection of these 8 OTUs were that they occurred in more than two habitats and that they had different modes of horizontal transmission. Focus was on aphid transmission, and examples of pollen and contact transmitted viruses were included for comparison.

The detection of virus OTUs from multiple host species across plant communities of four habitats was conducted by HTS of total RNA extracts from plants. The extensive sampling yielded 332 libraries (122 plant species) of RNA extracts of individual plants pooled by plant species and collection, from which read sequences were subjected to both HTS2C and HTS6C selection criteria (see Materials and Methods). The application of the most stringent read selection criteria (HTS6C) reduced the number of OTUs and the libraries from which they were detected compared to the least stringent selection (HTS2C) criteria (Figure 4.8). However, the relative host range breadths of the different virus OTUs estimated with each method remained consistent and were evident in

the patterns of virus-host associations across the habitats (Table 4.3). As HTS detection of virus OTUs involves uncertainty from different factors that include the proportion of false positives produced, or the high proportion of unknown virus species expected in wild plant communities (Bernardo et al. 2018; Zamfir et al. 2023), detection was confirmed by RT-PCR with specific primers. Percent identity of nucleotide sequences of RT-PCR-positive amplicons for each of the virus species (identity values from 86-100%, Table A2, Annexes) showed that specificity of detection by RT-PCR was good for all viruses. Munne et al. (2021) also reported the efficiency of RT-PCR for the detection of SARS-CoV-2 (a positive single-stranded RNA human virus) to be high, with its sensitivity estimated to be 70–98% and specificity is approximately 95%. The results also revealed a strong correlation between the HTS2C and RT-PCR validation (Figure 4.9). For instance, viruses like CMV and TAV showed high concordance between HTS2C and RT-PCR detections, with only minimal discrepancies in the number of libraries detected by either method. On the other hand, amplicons of BChV matched that of two related polerovirus species: *beet mild yellowing virus* (BMV) and *brassica yellows virus* (BrYV, synonyms *beet western yellows virus* (BWYV) and *turnip yellows virus* (TuYV). These virus species (BChV, BMV and BrYV) are known to be closely related serologically to one another with few successful attempts to clearly distinguish between them (Lemaire et al. 1995). The strong correlation between the methods underscores the validity of using HTS for initial broad-spectrum viral detection in diverse plant habitats and RT-PCR for specific confirmation, especially for viruses with low read numbers. In previous analyses by the group to test whether HTS detections based on few reads (2 or 4 reads) were associated with a higher frequency of false positives, the relationship between OTU detection by RT-PCR and the number of read hits was not significant (McLeish et al. 2022). This indicates that many unexpected virus–host associations are real. The large number of reads in many of these unexpected associations suggests virus infection rather than contamination. Despite the good agreement among the different detection procedures used, we cannot discard that false positives and false negatives may occur, as with any detection method.

In summary, the analyses performed with the HTS2C and RT-PCR data sets gave similar results, thus concluding that patterns in virus-host interaction can be identified despite biases in

virus detection. Also, the high concordance between the detection methods indicated reliable estimates of host range among the OTUs.

### 5.2.2 Host range evolution in heterogenous ecosystems.

The viruses with the broadest host ranges were CMV (88 species of 29 plant families) and PZSV (67 species of 21 families), followed by TAV (38 host species of 17 families), YoMV (37 host species of 13 families), WMV (37 species of 20 families) and BChV (32 species of 15 families), while TuMV (20 species of 15 families) and PPV (11 species of 6 families) had the narrowest host range (Tables 4.4 and 4.5). This evidence is consistent with previous records of host ranges for these viruses (Palukaitis and García-Arenal, 2003; Moury and Desbiez, 2020). The observations in this study have expanded on the number of known host species of the studied viruses. Apart from PPV and BChV, other viruses were present in multiple hosts species distributed across all four habitats (Figure 4.12), an indication of a generalist strategy in host use that may favour emergence (Elena et al. 2014; McLeish et al. 2018). This is consistent with previous findings that show that generalist viruses like CMV (Sacristán et al. 2005), or WMV (Peláez et al. 2021) often exploit a wide array of hosts across diverse environments. which facilitates successful transmission, survival and dispersal (Woolhouse et al. 2001; Power & Flecker, 2003). Furthermore, host range expansions are evident in CMV being able to infect over 1000 species of plant hosts in tropical, subtropical and temperate climates (Palukaitis and García-Arenal, 2003). Conversely, my study showed that viruses such as BChV and PPV that are restricted in host range, are also restricted to the anthropic Crop or Edge habitats.

The clustering analysis of host ranges (Figure 4.13) demonstrated that similarity in host range was not strictly determined by virus species identity, nor by taxonomy or transmission mode. This is particularly evident for BChV and TAV, whose host ranges clustered together despite having different transmission modes and belonging to different virus families (Figure 4.13). Also, different host use patterns among phylogenetically related viruses can translate into variation in how host resources are used. For example, CMV and PZSV which are less related taxonomically as they belong to different genera and are transmitted by different mechanisms, share a more

similar pattern of host-species use than CMV and TAV (Figure 4.13), which belong to the same genus, are both vectored by aphids, and are closely related genetically as to be able to form stable hybrids by reassortment of genomic segments (Palukaitis and García-Arenal 2003). Furthermore, sameness in transmission mode may not translate to similar pattern of host-species use as demonstrated in Figure A1 where CMV clustered separately, highlighting its broader range and unique host infection profile compared to the other aphid non-persistently transmitted viruses. Host ranges of each virus varied across habitats (Tables 4.4 and 4.5). For example, the analysis showed that host range was broader in Edge for most viruses, except for YoMV, which had a larger host range in Oakwood and Wasteland. The changes to the realised host ranges of each plant virus across the ecosystem is consistent with definitions of facultative generalism: an ecological strategy where an organism, such as a plant virus, can interact with a wide range of hosts but may also specialise on a few under certain conditions (Shipley et al. 2009). Resource exploitation strategies, like facultative generalism, imply conditional responses to environmental heterogeneity and the existence of niche differences among hosts and pathogens (McLeish et al. 2018; Seabloom et al. 2015). The distinction between a virus's fundamental and realised host range is important: while the fundamental host range includes potential host species across given environmental variation, the realised host range refers to the subset of available hosts that a virus actually uses in a specific environment (Wells & Clark, 2019). The results demonstrate that a virus's realised host range is not solely dependent on its ability to infect various species but is also constrained by ecological factors such as host availability and heterogeneity of the environment. While evolutionary history may influence host range, the ecology of host species is critical in defining the range of hosts that a virus can infect (Elena et al. 2014). Relatively high host species richness within the *Asteraceae*, *Poaceae*, and *Brassicaceae* were traits of the host range of most viruses (Figure 4.14). Host ranges of some viruses included host species of a particular family. For example, BChV only used *Asteraceae* and not *Poaceae*, CMV and YoMV used more species richness within *Asteraceae* compared to *Poaceae*, while TuMV and WMV used more species richness within *Poaceae* than *Asteraceae* (Figure 4.14). Gilbert et al. (2012) reported that plant pests and pathogens were more likely to attack closely related plant species than species separated by greater evolutionary distance. This was explained by the fact that closely related

plant species may share phylogenetically conserved traits that are important to interactions with pathogens. In tree–fungus interactions, the detected modules were also linked to host taxa but not to fungal taxa (Vacher et al. 2008). These results are consistent with the concept of host range as an overlapping continuum, with host shifts being common and easier between phylogenetically related hosts (Morris & Moury, 2019). Consequently, these studies concluded that plant taxonomy may be used as a first predictor of a pathogen host ranges. However, our results do not support this hypothesis, as most viruses were detected in hosts that belonged to distantly related plant families of Monocots, Magnoliopsida and Eudicots. Thus, plant taxonomy is not a sufficient predictor of virus host ranges, and other factors must be considered (Vienne et al. 2009; McLeish et al. 2018; Moury et al. 2017).

The present study also highlights the role of species abundance and richness in influencing host use. For instance, more abundant species, such as *Convolvulus arvensis* and *Diplotaxis erucoides*, were hosts to nearly all the viruses studied, suggesting that these plants' high abundance increases the likelihood of virus-host interactions. The broader host ranges observed for most viruses in communities of Edge cannot be attributed to a greater plant species richness of this habitat as compared with the other three, as species richness was higher for Wasteland and Oak (McLeish et al. 2021). It may be explained by the environmental conditions of Edge, which benefits from the water and fertilisers provided to adjacent crops and thus sustains plants communities along the year (Sacristán et al. 2004; McLeish et al. 2022), which facilitates the persistence of vectors and viruses (McLeish et al. 2022).

In summary, the host range patterns of the eight viruses emphasize the contribution of ecological factors, host availability, host preferences, and virus evolutionary history in the realised host range of a virus. Also, the distribution of host range according to habitats reflects the effects of ecological diversity and potential virus-host interactions specific to each habitat. The Edge in particular, had more interactions than other habitats, and seems to have the traits of a reservoir community (Babalola et al. 2023, McLeish et al. 2022).

### 5.2.3 Co-occurrence and co-infection patterns.

Co-occurrence analyses were performed to study the virus-virus interactions that shape virus community composition. It is known that interactions among viruses will be a factor in host range evolution (Malpica et al. 2006; Alcaide et al. 2020). Theoretical analyses predict that within-host dynamics of viruses in co-infected hosts have consequences in the evolution of virulence and transmission and adaptation to new hosts (May and Nowak 1995; Ganusov et al. 2002; Escribe et al. 2003, Fraile et al. 2008). Read sequence libraries were used to analyse co-occurrence while co-infection in individual plants was analysed with the RT-PCR detection data set. The results (section 4.3) revealed virus pairs that tended to co-occur in certain plant species and sites, and co-infect individual plants, as well as those that do not and highlight ecological strategies that contribute to virus community composition.

The analysis of the read sequence libraries (Figure 4.15 and Figure 4.16) indicated significant co-occurrence between CMV and TAV, TuMV and WMV, as well as PPV and WMV. These pairwise co-occurrences suggest shared ecological niches or benefits derived from co-infection, which potentially enhance transmission between or persistence in host plants. These results are consistent with the concept of "facilitation", where the presence of one species increases the likelihood of another species' survival or reproduction (Bruno et al 2003). The co-occurrence among these three virus pairs might also be influenced by their shared transmission modes, as the five viruses are transmitted non-persistently by aphids. This aligns with previous research on viruses with the same transmission vectors, which are more likely to co-infect the same hosts due to the vectors' feeding behaviour (Mauck et al. 2012). Co-infection analyses at the individual plant level revealed viruses that tended not to co-occur (Figure 4.17), particularly between contact-transmitted YoMV and either PPV, BChV, or PZSV that have other modes of transmission. Differential host use by the viruses that tended not to co-occur might arise from competition for resources within the host where the presence of one virus inhibits the infection or replication of another by triggering host-immune responses that suppress infection by other

viruses or by effects on the behaviour and performance of aphid vectors (Mauck et al. 2012; Karvonen et al. 2019).

Also, viruses with different transmission strategies might infect distinct host organs and tissues to reduce competition. For example, viruses transmitted by pollen may preferentially multiply in flower organs, while those transmitted by aphids non-persistently will rather multiply in leaves, and those transmitted by aphids persistently are limited to the phloem (Hull, 2002; García-Ordóñez and Pagán 2024). Analysis at the individual plant level also confirmed co-infection of aphid transmitted viruses such as CMV and TAV, CMV and WMV, TAV and TuMV as well as PPV and PZSV with different transmission modes. These positive associations might be explained by the same mechanisms invoked to explain the observed co-occurrences, which involved most of the co-infecting viruses. The findings suggest that while some viruses may benefit from co-infection, other viruses may be limited by competitive or antagonistic interactions based on the observed patterns of negative associations with contact transmitted tobamovirus, YoMV, and five viruses, PPV, BChV, PZSV, TAV and WMV with distinct transmission modes. Co-infection can increase the host range of co-occurring viruses by allowing them to reach hosts that they might not have otherwise infected alone. This phenomenon occurs as co-infecting viruses may exploit shared or overlapping transmission pathways, vectors, or even utilize mutual facilitation mechanisms, allowing each virus to bypass host-specific barriers. One example is seen in the co-infection dynamics of viruses such as *barley yellow dwarf viruses* (BYDV) in California grasslands, where shared vectors among co-infecting viruses facilitate their transmission to a wider range of hosts (Seabloom et al. 2009). The competitive interactions based on the observed patterns of negative associations can also shape the realised host range of each virus, e.g the presence of one virus inhibits the infection, replication or transmission of another by triggering host-immune responses, limiting their occurrence in certain hosts where competing viruses are already present. The dynamics of competitive interactions and facilitative co-infection among viruses can determine which hosts a virus can successfully infect in a given environment and shape host range evolution.

#### 5.2.4 Ecological and genetic determinants of incidence.

Regression analyses of virus incidence, transmission traits, species identity, and plant community distinctions were used to understand how extrinsic/ecological and intrinsic/virus genetic factors contribute to virus community composition. Incidence (defined as the percentage of infected plant individuals within a given sample) was estimated by considering; i) that from observed hosts (*ih*) within the realised host range of a virus; and ii) that in the entire plant community (*iL*) based on observed hosts and the other plant species in a given community. The results revealed that both ecological (habitat) and genetic factors (categorised by genus and transmission mode) significantly influence virus incidence. For each virus, both *ih* and *iL* varied significantly across habitats, with incidence rates in the anthropic habitats (Crop and Edge) higher than in the wild habitats (Oakwood and Wasteland). These patterns suggests that disturbance in anthropic communities of managed habitats is positively correlated with opportunities for virus transmission and infection. Increased incidence in the anthropic habitats is possibly due to high host plant density, reduced biodiversity, and increased vector activity (Power & Mitchell, 2004), factors that have not been analysed independently in this study.

The analyses of *ih* and *iL* across virus genera (Figures 4.19 and Figures 4.23) indicated that virus genus significantly affects incidence, with cucumoviruses and anulaviruses exhibiting highest incidence across habitats than potyviruses, tobamoviruses and poleroviruses. Despite that both potyviruses, with monopartite genomes, and cucumoviruses, with three partite genomes, are transmitted non-persistently by aphids, the incidence of cucumoviruses consistently exceeded that of the potyviruses. This suggests genomic organization and expression of multipartite viruses), is a successful strategy as predicted from a host range meta-analysis (Moury et al. 2017). The interaction between habitat and genera had significant effects on incidence. The incidence of cucumoviruses was significantly higher than the polerovirus or anulavirus in Edge and Oakwood.

Transmission mode was another factor that influenced virus incidence. In all four habitats, the aphid non-persistent transmission mode was associated with higher incidence compared to

aphid persistent, contact or pollen transmission (Figures 4.20 and 4.24). Non-persistent viruses are transmitted efficiently by vectors that frequently move between plants and their widespread occurrence across diverse plant species and habitats is favoured by (i) the polyphagous nature of many aphids, i.e non-specific aphid-plant association that enhances virus transmission across many different host plants, (Mauck et al. 2010); (ii) low virus-aphid specificity enables transmission of a virus by a broad range of aphid species, increasing the virus's incidence (Fereres and Raccah 2015); (iii) virus is acquired and inoculated during short probes (Fereres and Raccah, 2015); and (iv) virus-induced alterations in plant chemistry, such as changes in volatile emissions, which either attract aphids to infected plants or repel aphids because of poor nutritional quality causing rapid vector dispersal to healthy plants (Mauck et al. 2014; Carmo-Sousa et al. 2014). On the other hand, persistently transmitted viruses show high vector specificity, and are transmitted during phloem feeding of aphids (Fereres and Raccah, 2015), which occurs in a subset of the plants species on which these aphids probe. Interestingly, PZSV, a pollen-transmitted virus, showed high incidence in the Edge habitat. This is possibly due to the Edge supporting both plant and pollinator populations year-round (Endara and Coley 2011; Levitzky et al. 2019).

In summary, these findings underscore the interplay between ecological and genetic factors in shaping virus incidence patterns. These results emphasize the important role of habitat type and transmission mode in shaping virus incidence, offering key insights for managing plant virus outbreaks. The relatively high incidence of CMV and TAV in anthropic habitats suggests that these communities might act as hotspots for emergence, potentially due to the high density of susceptible hosts combined with aphid transmission and increased vector activity. The incidence for the other six viruses - TuMV, PPV, WMV, PZSV, BChV, and YoMV, was similar in both anthropic and wild habitats, suggesting that factors other than host plant density and ecosystem simplification influence virus incidence patterns (Roossinck & García-Arenal, 2015; Stukenbrock & McDonald, 2008). The findings that non-persistent and pollen-transmitted viruses had higher incidence rates than contact or persistently-transmitted virus highlight the need for targeted management strategies that address the specific transmission modes of viruses. For instance, managing vector populations or altering planting practices to disrupt pollinator-mediated transmission could be effective strategies (Fereres & Moreno, 2009).

### 5.2.5 Population genetic diversity and structure.

The analyses of genetic diversity and population structure were performed with the HTS2C-selected read sequences that limited the estimates of nucleotide diversity ( $\pi$ ) to the habitat and site level for CMV, and site level only for TAV and PZSV. For these three viruses, with three-partite genomes, genetic diversity varied between genomic segments (section 4.5) with variation in localised regions, but the trends along the segments were relatively consistent among habitats and sites. At the habitat level, CMV had higher  $\pi$  particularly in RNA3 in Crop and Edge habitats (Table 4.8 and Figure 4.26). At the Edge sites, the highest localised diversity (Figures 4.27-4.29) was observed in the RNA3 segment of CMV and the lowest in the RNA1 and RNA 3 segments of PZSV. Both CMV and TAV had higher  $\pi$  in their RNA3 genome compared to the RNA1 and RNA2 segments. RNA3 encodes proteins related to host colonisation and transmission (Palukaitis and García-Arenal, 2003) that are expected to show more variation across hosts than proteins involved in the replication of the genome.

The estimated value of  $\pi$  of CMV in Edge was minimal compared to that observed from the other habitats (Figure 4.26), yet the sample size of isolates used to estimate  $\pi$  from Edge was larger than from Oak or Crop, in that order. The relatively low genetic diversity of isolates of CMV that infected the melon crop, and perennial or annual host species of Edge, is consistent with early epidemiological studies carried out in France and the USA (reviewed in Tepfer and García-Arenal 2019), and more recently in Spain (Sacristán et al. 2004), that indicated wild plants growing in proximity to crops assure the year-round presence of the virus for crop infection. That Wasteland had a large variation, may be the result of a small sample with extreme pair-wise differences compared to the Edge sample. This assertion is supported by the phylogenetic inferences with isolates from Wasteland being phylogenetically over-dispersed within clades that predominate in one of the other habitats. This phylogenetic pattern (Figure 4.30) is consistent with sporadic and successful colonisation from the other habitats. Closely related isolates from Oakwood were largely early-branching lineages of each of the three RNA segments. This is

indicative of successful colonisation in all the other habitats by genotypes from Oakwood, and subsequent re-colonisation of Oakwood from the other habitats.

The Bayesian Tip-association Significance tests (BaTS) for CMV indicated significant tip associations observed for CMV in the anthropic habitats (grouping Crop and Edge). Relatively uniform selection pressures in these communities due to consistent host availability and agricultural practices (Elena et al. 2014) might promote genetic homogenization among virus populations. This finding is consistent with the notion that anthropic environments often reduce genetic diversity within species, and lead to closely related genetic variants compared to undisturbed communities with greater diversity (Malmstrom et al. 2005). There was no phylogenetic clustering of CMV, TAV, and PZSV populations from Edge based on site or host species (Figures 4.31 to 4.33), which suggests weak host-specific genetic differentiation. The negligible associations between virus populations and host species are indicative of generalist host use strategies. This is in line with studies that have shown generalist viruses tend to exhibit low levels of host-specific genetic divergence (McLeish et al. 2017). Additionally, the lack of site-specific genetic differentiation within the Edge might reflect high connectivity and gene flow, where viruses can easily spread among host species at local scales. The observed low-level genetic divergences in this habitat might suggest relatively homogeneous selection pressures (McLeish et al. 2018).

Thus, the genetic diversity and phylogenetic analyses in section 4.5 highlight the role of habitat and anthropic disturbance in shaping the genetic structure of virus populations. These findings suggest that genetic homogeneity among virus populations is promoted in plant communities subject to disturbance by human influences, such as agricultural fields, while more diverse and less disturbed communities support greater genetic diversity. Additionally, the lack of strong host-specific genetic divergences within a habitat underscores the generalist host use strategies of these viruses. These insights underscore the role of ecological factors in the genetic diversity and structure of virus populations and indicate the existence of preferential host use and habitat specificity that impacts gene flow among communities and selection within them (Fraile et al. 2017; Pagán & García-Arenal, 2018). The ecological dynamics of plant viruses can inform

strategies for managing emergence in agricultural systems and emphasizes the importance of plant community diversity in maintaining genetic variability in virus populations.

## 6. CONCLUSIONS.

This thesis aimed to understand the role of habitat heterogeneity in virus host range and host use, as it affects transmission and the process of virus emergence. Particularly, I studied virus-host interactions in a crop (carrot) and its wild relative and in a multi-host multi-virus system by using HTS to analyse host range, incidence, co-occurrence and population genetic structure. The main conclusions of this thesis are:

1. The virome of wild carrot populations had greater virus richness compared to cultivated carrots, and had a larger fraction of undescribed viruses, Thus, wild carrot populations have a role as reservoirs of viral diversity and potential sources of new viruses emerging in carrot crops.
2. Comparison of results from HTS2C and RT-PCR detections indicate that the identification of patterns in virus-host interactions is robust to biases in virus detection, and thus, the HTS2C is a reliable method of virus detection for community ecology studies.
3. Host range differed across eight analysed viruses and varied significantly among habitats, with Edge habitats being hotspots for virus interactions.
4. Host range of viruses was determined in part by an intrinsic factor of phylogenetic history, as represented by genus, and mode of transmission, and by extrinsic ecological factors.
5. Generalist virus species exhibited preferential host use patterns, revealing that local community host specialisation that arises from habitat heterogeneity, is a successful strategy for generalist viruses. Differential host use across habitats indicated these viruses behave as facultative generalists.
6. Co-occurrence and co-infection analyses revealed non-random positive and negative associations among viruses. Positive associations occurred among viruses with shared genomic structure and transmission mode, while negative associations occurred mostly among viruses with different transmission modes. Thus, transmission mode had a role in the assembly of virus communities.

7. Virus species, genus and transmission mode were important determinants of incidence, with habitat as a key factor, highlighting the importance of both evolutionary and ecological factors in host use.
8. Incidence of aphid non-persistently transmitted PZSV was higher across habitats compared to aphid persistently transmitted viruses and YoMV.
9. For all viruses, incidence was higher in the anthropic habitats of Crop and Edge than in the wild habitats of Oak and Wasteland, according to the ecosystem simplification hypothesis that anthropic intervention of habitats favours plant pathogens.
10. The genetic diversity of the analysed viruses was not structured according to host plant, which indicated weak across host fitness trade-offs and a role of ecological fitting in host range evolution.
11. Quantifying non-random patterns in plant-virus interactions at the ecosystem scale contributes to understanding virus emergence.

## 7. REFERENCES.

1. Adams I.P, Miano D.W, Kinyua Z.M, Wangai A, Kimani E et al. (2013) Use of next-generation sequencing for the identification and characterization of Maize chlorotic mottle virus and Sugarcane mosaic virus causing maize lethal necrosis in Kenya. *Plant Pathol* 62 (4): 741–749.
2. Adams IP, Skelton A, Macarthur R, Hodges T, Hinds H, Flint L, Nath PD, Boonham N, Fox A (2014) Carrot yellow leaf virus is associated with carrot internal necrosis. *PLOS One* 9, e109125. <https://doi.org/10.1371/journal.pone.0109125>
3. Adams, I. P., Glover, R. H., Monger, W. A., Mumford, R., Jackeviciene, E., Navalinskiene, M., Samuitiene, M., & Boonham, N. (2009). Next-generation sequencing and metagenomic analysis: a universal diagnostic tool in plant virology. *Molecular plant pathology*, 10(4), 537–545. <https://doi.org/10.1111/j.1364-3703.2009.00545.x>
4. Agrios GN (2005) *Plant pathology*, 5th edn. Elsevier, Amsterdam
5. Al Rwahnih, M., Daubert, S., Golino, D., & 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(2), 395–401. <https://doi.org/10.1016/j.virol.2009.02.028>
6. Alcaide, C., Rabadán, M. P., Moreno-Pérez, M. G., and Gómez, P. 2020. Implications of mixed viral infections on plant disease ecology and evolution. *Adv. Virus Res.* 106:145–169. <https://doi.org/10.1016/bs.aivir.2020.02.001>
7. Alcaide, C., Rabadán, M. P., Moreno-Pérez, M. G., and Gómez, P. 2020. Implications of mixed viral infections on plant disease ecology and evolution. *Adv. Virus Res.* 106:145–169. <https://doi.org/10.1016/bs.aivir.2020.02.001>
8. Alcaide, C., Sardanyes, J., Elena, S. F., & Gomez, P. (2021). Increasing temperature alters the within-host competition of viral strains and influences virus genetic variability. *Virus Evolution*, 7(1), veab017. <https://doi.org/10.1101/2020.07.06.190173>

9. Alizon, S., Hurford, A., Mideo, N., & Van Baalen, M. (2009). Virulence evolution and the trade-off hypothesis: history, current state of affairs and the future. *Journal of evolutionary biology*, 22(2), 245–259. <https://doi.org/10.1111/j.1420-9101.2008.01658.x>
10. Allison R., Johnston R. E., Dougherty W. G. 1986; The nucleotide sequence of the coding region of tobacco etch virus genomic RNA: evidence for the synthesis of a single polyprotein. *Virology* 154:9–20
11. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of molecular biology*, 215(3), 403-410.
12. Amari, K., Huang, C., & Heinlein, M. (2021). Potential impact of global warming on virus propagation in infected plants and agricultural productivity. *Frontiers in Plant Science*, 12. <https://doi.org/10.3389/fpls.2021.649768>
13. Araújo, M.B., Rozenfeld, A., Rahbek, C. and Marquet, P.A. (2011), Using species co-occurrence networks to assess the impacts of climate change. *Ecography*, 34: 897-908. <https://doi.org/10.1111/j.1600-0587.2011.06919.x>
14. Archie, E. A., Luikart, G., & Ezenwa, V. O. (2009). Infecting epidemiology with genetics: a new frontier in disease ecology. *Trends in ecology & evolution*, 24(1), 21–30. <https://doi.org/10.1016/j.tree.2008.08.008>
15. Armand, T., Korn, L., Pichon, E., Souquet, M., Barbet, M., Martin, J., ... & Jacquot, E. (2021). Efficiency and persistence of movento® treatment against *Myzus persicae* and the transmission of aphid-borne viruses. *Plants*, 10(12), 2747. <https://doi.org/10.3390/plants10122747>
16. Arnold, A. E., & Lutzoni, F. (2007). Diversity and host range of foliar fungal endophytes: are tropical leaves biodiversity hotspots?. *Ecology*, 88(3), 541–549. <https://doi.org/10.1890/05-1459>
17. August, P.V. (1983). The role of habitat complexity and heterogeneity in structuring tropical mammal communities. *Ecology*, 64, 1495–1507

18. Aurora Fraile, Michael J. McLeish, Israel Pagán, Pablo González-Jara, Daniel Piñero, Fernando García-Arenal (2017). Environmental heterogeneity and the evolution of plant-virus interactions: Viruses in wild pepper populations, *Virus Research*, Volume 241.
19. Babalola BM, Schönegger D, Faure C, Marais A, Fraile A, Garcia-Arenal F, Candresse T. (2022a) Identification of two novel putative satellite RNAs with hammerhead structures in the virome of French and Spanish carrot samples. *Arch. Virol.* **167**, 2287-2292. <https://doi.org/10.1007/s00705-022-05538-z>
20. Babalola, B. M., Faure, C., Marais, A., Fraile, A., Garcia-Arenal, F., Candresse & T. (2022c) First report of Apium virus Y in wild carrot (*Daucus carota* ssp. *carota*) in Spain. *New Dis Rep.* 45, e12097. <https://doi.org/10.1002/ndr2.12097>
21. Babalola, B.M., Faure, C., Marais, A. *et al.* First report of carrot torrado virus 1 (CaTV1) naturally infecting carrots in Spain. *J Plant Pathol* **104**, 1565 (2022b). <https://doi.org/10.1007/s42161-022-01196-x>
22. Bačnik, K., Kutnjak, D., Pecman, A., Mehle, N., Tušek Žnidarič, M., Gutiérrez Aguirre, I., & Ravnkar, M. (2020). Viromics and infectivity analysis reveal the release of infective plant viruses from wastewater into the environment. *Water research*, 177, 115628. <https://doi.org/10.1016/j.watres.2020.115628>
23. Barajas-Barbosa M.P, Weigelt P, Borregaard M.K, Keppel G, Kreft H. Environmental heterogeneity dynamics drive plant diversity on oceanic islands. *J Biogeogr.* 2020; 47: 2248–2260. <https://doi.org/10.1111/jbi.13925>
24. Bedhomme, S., Hillung, J., and Elena, S. F. (2015). Emerging viruses: Why they are not jacks of all trades? **Curr. Opin. Virol.** 10:1-6. <https://doi.org/10.1016/j.coviro.2014.10.006>
25. Bera, S., Fraile, A., & García-Arenal, F. (2018). Analysis of fitness trade-offs in the host range expansion of an RNA virus, tobacco mild green mosaic virus. *Journal of Virology*, 92(24), 10-1128. <https://doi.org/10.1128/jvi.01268-18>
26. Bernardo, P., Charles-Dominique, T., Barakat, M., Ortet, P., Fernandez, E., Hartnady, P., ... & Filloux, D. (2017). Geometagenomics illuminates the impact of agriculture on the distribution and prevalence of plant viruses at the ecosystem scale. *The Isme Journal*, 12(1), 173-184. <https://doi.org/10.1038/ismej.2017.155>

27. Bonnet J., Fraile A., Sacristán S., Malpica J.M., García-Arenal F. (2005). Role of recombination in the evolution of natural populations of Cucumber mosaic virus, a tripartite RNA plant virus. *Virology* 332:1(359-368).  
<https://doi.org/10.1016/j.virol.2004.11.017>.
28. Boonham N, Kreuze J, Winterc S, Van der Vlugt R, Bergervoet J, Tomlinson J, Mumford R (2014) Methods in virus diagnostics: from ELISA to next generation sequencing. *Virus Res* 186:20–31 <https://doi.org/10.1016/j.virusres.2013.12.007>
29. Branch AD, Robertson HD (1984) A replication cycle for viroids and other small infectious RNAs. *Science* 223:450–455. <https://doi.org/10.1126/science.6197756>
30. Briddon RW, Ghabrial S, Lin NS, Palukaitis P, Scholthof KBG, Vetten HJ (2012) Satellite and other virus-dependent nucleic acids. In: King AMQ, Adams MJ, Carstens EB, Lefkowitz EJ (eds) *Virus Taxonomy: classification and nomenclature of viruses*. Ninth Report of the International Committee on Taxonomy of Viruses. Elsevier, Amsterdam, pp 1211–1219
31. Brown SP, Hochberg ME, Grenfell BT (2002) Does multiple infection select for raised virulence? *Trends Microbiol.* 10: 401–405.
32. Bruening G (1989) Compilation of self-cleaving sequences from plant virus satellite RNAs and other sources. *Methods Enzymol* 180:546–558. [https://doi.org/10.1016/0076-6879\(89\)80123-5](https://doi.org/10.1016/0076-6879(89)80123-5)
33. Bruening G, Passmore BK, van Tol H, Buzayan JM, Feldstein PA (1991) Replication of plant virus satellite RNA: evidence favors transcription of circular templates of both polarities. *Mol Plant Microb Interact* 4:219–225. <https://doi.org/10.1094/mpmi-4-219>
34. Bruno, J. F., Stachowicz, J. J., & Bertness, M. D. (2003). Inclusion of facilitation into ecological theory. *Trends in Ecology & Evolution*, 18(3), 119-125.
35. Brunt, A.A. (1992). The general properties of potyviruses. In: Barnett, O.W. (eds) *Potyvirus Taxonomy*. Archives of Virology, vol 5. Springer, Vienna. [https://doi.org/10.1007/978-3-7091-6920-9\\_1](https://doi.org/10.1007/978-3-7091-6920-9_1)
36. Bujarski, J.; Gallitelli, D.; García-Arenal, F.; Pallás, V.; Palukaitis, P.; Krishna Reddy, M.; Wang, A. ICTV Virus Taxonomy Profile: Bromoviridae. *J. Gen. Virol.* 2019, 100, 1206–1207.

37. Campbell, C. L., & Neher, D. A. (1994). Estimating disease severity and incidence. *Epidemiology and management of root diseases* (pp. 117-147). Springer, Berlin, Heidelberg.
38. Canto, T., Aranda, M. A., & Fereres, A. (2009). Climate change effects on physiology and population processes of hosts and vectors that influence the spread of hemipteran-borne plant viruses. *Global Change Biology*, 15(8), 1884-1894. <https://doi.org/10.1111/j.1365-2486.2008.01820.x>
39. Carmo-Sousa, M., Moreno, A., Garzo, E., & Fereres, A. (2014). A non-persistently transmitted-virus induces a pull-push strategy in its aphid vector to optimize transmission and spread. *Virus research*, 186, 38-46 . <https://doi.org/10.1016/j.virusres.2013.12.012>.
40. Carrasco, P., de la Iglesia, F., and Elena, S. F. (2007). Distribution of fitness and virulence effects caused by single-nucleotide substitutions in tobacco etch virus. *J. Virol.* 81:12979-12984. <https://doi.org/10.1128/JVI.00524-07>
41. Carrots and Related Apiaceae Crops. (2020). India: CABI.
42. Cavender-Bares, J., Kozak, K. H., Fine, P. V., & Kembel, S. W. (2009). The merging of community ecology and phylogenetic biology. *Ecology Letters*, 12(7), 693-715.
43. Chao, A., & Jost, L. (2012). Coverage-based rarefaction and extrapolation: Standardizing samples by completeness rather than size. *Ecology*, 93(12), 2533-2547. <https://doi.org/10.1890/11-1952.1>
44. Civitello DJ, Cohen J, Fatima H, Halstead NT, Liriano J, McMahon TA, Ortega CN, Sauer EL, Sehgal T, Young S, Rohr JR (2015) Biodiversity inhibits parasites: broad evidence for the dilution effect. *Proc. Natl. Acad. Sci. USA* 112, 8667-8671. <https://doi.org/10.1073/pnas.1506279112>
45. Connor EF, McCoy ED. 2001 Species-area relationships. *Encycl. Biodivers.* 5, 397–411. [doi:10.1016/B0-12-226865-2/00132-2](https://doi.org/10.1016/B0-12-226865-2/00132-2)
46. Cooper I., Jones R.A.C. 2006. Wild Plants and Viruses: Under-Investigated Ecosystems. In: *Advances in Virus Research*, 67, 1–47. [https://doi.org/10.1016/S0065-3527\(06\)67001-2](https://doi.org/10.1016/S0065-3527(06)67001-2)

47. Cooper, J. I. 1995. Viruses and the Environment. 2 ed. Chapman and Hall, London
48. Darzi, E., Smith, E., Shargil, D., Lachman, O., Ganot, L., and Dombrovsky, A, 2018. The honeybee *Apis mellifera* contributes to Cucumber green mottle mosaic virus spread via pollination. *Plant Pathol.* 67:244-251
49. Davison AJ. *Journal of General Virology* - Introduction to 'ICTV Virus Taxonomy Profiles'. *J Gen Virol.* 2017 Jan;98(1):1. <https://doi.org/10.1099/jgv.0.000686>
50. Dennehy J. J. (2017). Evolutionary ecology of virus emergence. *Annals of the New York Academy of Sciences*, 1389(1), 124–146. <https://doi.org/10.1111/nyas.13304>
51. Diamond, J.M. (1975) Assembly of Species Communities. In: Diamond, J.M. and Cody, M.L., Eds., *Ecology and Evolution of Communities*, Harvard University Press, Boston, 342-344.
52. Dijkstra J. (1992). Importance of host ranges and other biological properties for the taxonomy of plant viruses. *Archives of virology. Supplementum*, 5, 173–176. [https://doi.org/10.1007/978-3-7091-6920-9\\_14](https://doi.org/10.1007/978-3-7091-6920-9_14)
53. Dolja, V. and Koonin, E. (2011). Common origins and host-dependent diversity of plant and animal viromes. *Current Opinion in Virology*, 1(5), 322-331. <https://doi.org/10.1016/j.coviro.2011.09.007>
54. Dolja, V. and Koonin, E. (2018). Metagenomics reshapes the concepts of rna virus evolution by revealing extensive horizontal virus transfer. *Virus Research*, 244, 36-52. <https://doi.org/10.1016/j.virusres.2017.10.020>
55. Dombrovsky, A. and Smith, E. 2017. Seed Transmission of Tobamoviruses: Aspects of Global Disease Distribution. In José C. Jiménez-López (Ed.) *Seed Biology*. IntechOpen. Pp-233-260.
56. Domier LL (2012) Luteoviridae. In: King AMQ, Adams MJ, Carstens EB, Lefkowitz EJ (eds) *Virus Taxonomy: classification and nomenclature of viruses*. Ninth Report of the International Committee on Taxonomy of Viruses. Elsevier, Amsterdam, pp 1045–1053

57. Donaire, L., Wang, Y., Gonzalez-Ibeas, D., Mayer, K. F., Aranda, M. A., & Llave, C. (2009). Deep-sequencing of plant viral small RNAs reveals effective and widespread targeting of viral genomes. *Virology*, 392(2), 203–214. <https://doi.org/10.1016/j.virol.2009.07.005>
58. Duffus, J. 1971. Role of weeds in the incidence of virus diseases. *Annu.Rev. Phytopathol.* 9:319-340.
59. Eastwell, K.C., Glass, J.R., Seymour, L.M. & Druffel, K.J. (2008) First report of infection of poison hemlock and celery by Apium virus Y in Washington State. *Plant Disease*, 92, 1710. <https://doi.org/10.1094/PDIS-92-12-1710C>
60. EFSA Panel on Plant Health (2022) Pest categorisation of Apium virus Y. *EFSA Journal*, 20, e06930. <https://doi.org/10.2903/j.efsa.2022.6930>
61. Elbeaino, T., Kontra, L., Demian, E., Jaksa-Czotter, N., Slimen, A. B., Fabian, R., Lazar, J., Tamisier, L., Digiaro, M., Massart, S., & Varallyay, E. (2020). Complete Sequence, Genome Organization and Molecular Detection of Grapevine Line Pattern Virus, a New Putative Anulavirus Infecting Grapevine. *Viruses*, 12(6), 602. <https://doi.org/10.3390/v12060602>
62. Elena Fito, SF.; Fraile, A.; García-Arena, F. (2014). Evolution and emergence of plant viruses. *Advances in Virus Research.* (88):161-191. doi:10.1016/B978-0-12-800098-4.00003-9
63. Elena S.F., Bedhomme S., Carrasco P., *et al* 2011. The evolutionary genetics of emerging plant RNA viruses. *Mol. Plant Microbe Interact.* 24: 287–93
64. Elena S.F., García-Arenal F. 2023. Plant Virus Adaptation to New Hosts: A Multi scale Approach. In *Viral Fitness and Evolution: Population Dynamics and Adaptive Mechanisms*, 167-196. Cham: Springer [https://doi.org/10.1007/978-3-031-15640-3\\_5](https://doi.org/10.1007/978-3-031-15640-3_5)
65. Elena, S. F., Carrasco, P., Daròs, J. A., & Sanjuán, R. (2014). Mechanisms of genetic robustness in RNA viruses. *EMBO Reports*, 7(2), 168-173.
66. Elena, S. F., Fraile, A., and García-Arenal, F. 2014. Evolution and emergence of plant viruses. **Adv. Virus Res.** 88:161-191. <https://doi.org/10.1016/B978-0-12-800098-4.00003-9>
67. Elena, S.F., Bedhomme S., Carrasco P., *et al* 2011. The evolutionary genetics of emerging plant RNA viruses. *Mol. Plant Microbe Interact.* 24: 287–93

68. Endara, M.J. & Coley, P.D. (2011) The resource availability hypothesis revisited: a meta-analysis. *Functional Ecology*, **25**, 389–398. <https://doi.org/10.1111/j.1365-2435.2010.01803.x>
69. Engering A., Hogerwerf, L., & Slingenbergh, J. (2013). Pathogen-host-environment interplay and disease emergence. *Emerging microbes & infections*, **2**(2), e5. <https://doi.org/10.1038/emi.2013.5>
70. Escriu F, Fraile A, Garcia-Arenal F (2003) The evolution of virulence in a plant virus. *Evolution Int J Org Evolution*, **57**: 755–65.
71. Ettema C.H. and Wardle D.A. (2002). Spatial soil ecology. *Trends in Eco. Evo* **17**, 4: 177-183. [https://doi.org/10.1016/S0169-5347\(02\)02496-5](https://doi.org/10.1016/S0169-5347(02)02496-5)
72. Fargette D, Roberts IM, Harrison BD (1991) Particle purification and properties of Cassava Ivorian bacilliform virus. *Ann Appl Biol* **119**:303–312
73. Farkas, D.H.; Holland, C.A. (2009). Overview of molecular diagnostic techniques and instrumentation. In *Cell and Tissue Based Molecular Pathology*; Tubbs, R.R., Stoler, M.H., Eds.; Churchill Livingstone: Philadelphia, CA, USA; pp. 19–32.
74. Fereres A, Raccah B. 2015. Plant virus transmission by insects, eLS. <https://doi.org/10.1002/9780470015902.a0000760.pub3>
75. Fereres, A., & Moreno, A. (2009). Behavioural aspects influencing plant virus transmission by homopteran insects. *Virus Research*, **141**(2), 158-168.
76. Fereres, A.; Perry, K.L. Movement Between Plants: Horizontal Transmission. In *Cucumber Mosaic Virus*; The American Phytopathological Society: St. Paul, MN, USA, 2019; pp. 173–184.
77. Fiallo-Olivé, E., Trenado, H. P., Louro, D., and Navas-Castillo, J. (2019). Recurrent speciation of a tomato yellow leaf curl geminivirus in Portugal by recombination. **Sci. Rep.** **9**:1332. <https://doi.org/10.1038/s41598-018-37971-z>
78. Fraile, A., Sacristán, S., and García-Arenal, F. (2008). A quantitative analysis of complementation of deleterious mutants in plant virus populations. **Span. J. Agric. Res.** **6**:195-200. <https://doi.org/10.5424/sjar/200806S1-388>

79. Frederico LA, Kunkel TA, Shaw BR. (1990) A sensitive genetic assay for the detection of cytosine deamination: determination of rate constants and the activation energy. *Biochemistry*; 29:2532–2537
80. Frederico LA, Kunkel TA, Shaw BR. (1990) A sensitive genetic assay for the detection of cytosine deamination: determination of rate constants and the activation energy. *Biochemistry*; 29:2532–2537
81. French CJ, Bouthillier M, Bernardy M, Ferguson G, Sabourin M, Johnson RC, Masters C, Godkin S, and Mumford R. 2001. First report of *Pepino Mosaic Virus* in Canada and the United States. *Plant Dis.* 85:1121.
82. Frenkel, M., Ward, C., & Shukla, D. (1989). The use of 3' non-coding nucleotide sequences in the taxonomy of potyviruses: application to watermelon mosaic virus 2 and soybean mosaic virus-n. *Journal of General Virology*, 70(10), 2775-2783. <https://doi.org/10.1099/0022-1317-70-10-2775>
83. Fuji, S., Kikuchi, M., Ueda, S., Toda, T., Furuya, H., Fukumoto, F., & Hanada, K. (2013). Characterization of a new Anulavirus isolated from Amazon lily plants. *Archives of virology*, 158(1), 201–206. <https://doi.org/10.1007/s00705-012-1467-8>
84. Gallitelli, D. Properties of a Tomato Isolate of Pelargonium Zonate Spot Virus. *Ann. Appl. Biol.* 1982, 100, 457–466.
85. Ganusov VV, Bergstrom TC, Antia R (2002) Within-host population dynamics and the evolution of microparasites in a heterogeneous host population. *Evolution* 56: 213–223
86. García-Arenal F., Fraile A. 2013. Trade-offs in host range evolution of plant viruses. *Plant Pathology* 62, 2–9. <https://doi.org/10.1111/ppa.12104>
87. García-Aréñal, F., Fraile, A., & Malpica, J. (2003). Variation and evolution of plant virus populations. *International Microbiology*, 6(4), 225-232. <https://doi.org/10.1007/s10123-003-0142-z>
88. García-Arenal, F.; Fraile, A.; Malpica, J.M. Variability and genetic structure of plant virus populations. (2001) *Annu. Rev. Phytopathol.* 39, 157–186.

89. Garcia-Ruiz H. Holste N. M., and LaTourrette K. (2021) 'Poleroviruses (Luteoviridae)'. In: Bamford, D., and Zuckerman, M. (eds) *Encyclopedia of Virology*, 4th edn. Vol. 3, pp. 594–602. Oxford: Academic Press, Elsevier.
90. Gibbs A. J. (1999) 'Evolution and Origins of Tobamoviruses', *Proceedings of the Royal Society Lond. Series B*, 354: 593–602.
91. Gibbs, A. J., Wood, J., García-Arenal, F., Ohshima, K., and Armstrong, J. S. (2015). Tobamoviruses have probably co-diverged with their eudicotyledonous hosts for at least 110 million years. **Virus Evol.** 1:vev019. <https://doi.org/10.1093/ve/vev019>
92. Gilbert, G. S., Magarey, R., Suiter, K., & Webb, C. O. (2012). Evolutionary tools for phytosanitary risk analysis: phylogenetic signal as a predictor of host range of plant pests and pathogens. *Evolutionary applications*, 5(8), 869–878. <https://doi.org/10.1111/j.1752-4571.2012.00265.x>
93. Giolitti F, Bejerman N, Nome C, Visintin G, de Breuil S, Lenardon S (2013) Biological and molecular characterization of an isolate of Pelargonium zonate spot virus infecting sunflower in Argentina. *J Plant Path.* doi:10.4454/JPP.V96I1.006
94. Gnanasekaran P, Chakraborty S (2018) Biology of viral satellites and their role in pathogenesis. *Curr Opin Virol* 33:96–105. <https://doi.org/10.1016/j.coviro.2018.08.002>
95. Gómez P, Sempere RN, and Aranda MA. 2012. *Pepino Mosaic Virus* and *Tomato torrado virus*: emerging viruses affecting tomato crops in the Mediterranean basin. *Adv. Virus Res.* 84:505–532.
96. Gómez, P., Sempere, R. N., Elena, S. F., and Aranda, M. A. 2009. Mixed infections of pepino mosaic virus strains modulate the evolutionary dynamics of this emergent virus. **J. Virol.** 83:12378-12387
97. Gonçalves, M. C., Ramos, A. D., Nascimento, T., Harakava, R., Duarte, A. P., & Lopes, J. R. S. (2020). Aphid transmission of maize yellow mosaic virus: an emerging polerovirus. *Tropical Plant Pathology*, 45(5), 544-549. <https://doi.org/10.1007/s40858-020-00374-5>
98. Gooding, G.V. 1986. Tobacco mosaic virus epidemiology and control. In: Van Regenmortel, M.H.V., Fraenkel-Conrat, H. (eds) *The Plant Viruses. The Viruses*. Springer, Boston, MA

99. Gurarie, D., & Seto, E. Y. (2009). Connectivity sustains disease transmission in environments with low potential for endemicity: modelling schistosomiasis with hydrologic and social connectivities. *Journal of the Royal Society, Interface*, 6(35), 495–508. <https://doi.org/10.1098/rsif.2008.0265>
100. Gutiérrez-Aguirre, I., Kutnjak, D., Rački, N., Rupar, M., & Ravnika, M. (2018). Monolith Chromatography as Sample Preparation Step in Virome Studies of Water Samples. *Methods in molecular biology (Clifton, N.J.)*, 1746, 63–75. [https://doi.org/10.1007/978-1-4939-7683-6\\_6](https://doi.org/10.1007/978-1-4939-7683-6_6)
101. Gutiérrez-Aguirre, I., Rački, N., Dreo, T., & Ravnika, M. (2015). Droplet digital PCR for absolute quantification of pathogens. *Methods in molecular biology (Clifton, N.J.)*, 1302, 331–347. [https://doi.org/10.1007/978-1-4939-2620-6\\_24](https://doi.org/10.1007/978-1-4939-2620-6_24)
102. Halvorsen, R., & Edvardsen, A. (2009). The concept of habitat specificity revisited. *Landscape Ecology*, 24, 851–861. <https://doi.org/10.1007/s10980-009-9363-7>.
103. Handelsman, M.R. Rondon, S.F. Brady, J. Clardy, R.M. Goodman. 1998. Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. *Chem. Biol.*, 5 (10) (1998), pp. R245–R249
104. Hanssen IM and Thomma BPHJ. 2010. *Pepino Mosaic Virus*: a successful pathogen that rapidly evolved from emerging to endemic in tomato crops. *Mol. Plant Pathol.* 11:179–189.
105. Hanssen IM, Mumford R, Blystad DR, Cortez I, Hasiów-Jaroszewska B, Hristova D, Pagán I, Pereira AM, Peters J, Pospieszny H, Ravnika M, Stijger I, Tomassoli L, Varveri C, van der Vlugt R, and Nielsen S. 2010. Seed transmission of *Pepino Mosaic Virus* in tomato. *Eur. J. Plant Pathol.* 126:145–152.
106. Hauser, S., Stevens, M., Mougél, C., Smith, H., Fritsch, C., Herrbach, E., Lemaire, O., 2000. Biological, serological and molecular variability suggest three different Beet Polerovirus species. *Phytopathology* 90 (5), 460–466.
107. Hay, J., Fellowes, A., & Timmerman, G. (1989). Nucleotide sequence of the coat protein gene of a necrotic strain of potato virus y from new zealand. *Archives of Virology*, 107(1-2), 111–122. <https://doi.org/10.1007/bf01313883>

108. Hily, J. M., Poulicard, N., Mora, M. Á., Pagán, I., & García-Arenal, F. (2016). Environment and host genotype determine the outcome of a plant-virus interaction: from antagonism to mutualism. *The New phytologist*, 209(2), 812–822. <https://doi.org/10.1111/nph.13631>
109. Hobbs, H. A., Eastburn, D. M., and D'Arcy, C. J. 2000. Solanaceous weeds as possible sources of Cucumber mosaic virus in Southern Illinois for aphid transmission to pepper. *Plant Dis.* 84:1221-1224.
110. Hodge, B. A., Paul, P. A., and Stewart, L. R. 2020. Occurrence and high throughput sequencing of viruses in Ohio wheat. *Plant Dis.*
111. Holkar, S. K., Parameswari, B., Kumar, A., Nithya, K., Shingote, P. R., Chhabra, M. L., ... & Pathak, A. D. (2020). Present status and future management strategies for sugarcane yellow leaf virus: a major constraint to the global sugarcane production. *The Plant Pathology Journal*, 36(6), 536-557. <https://doi.org/10.5423/ppj.rw.09.2020.0183>
112. Hortal, J., Triantis, K.A., Meiri, S., Thebault, E. & Sfenthourakis, S. (2009). Island species richness increases with habitat diversity. *Am. Nat.*, 174, E205–E217.
113. Hsieh, T. C., Ma, K. H., & Chao, A. (2016). INEXT: An R package for rarefaction and extrapolation of species diversity (Hill numbers). *Methods in Ecology and Evolution*, 7(12), 1451-1456. <https://doi.org/10.1111/2041-210X.12613>
114. Hu CC, Hsu YH, Lin NS (2009) Satellite RNAs and satellite viruses of plants. *Viruses* 1:1325–1350. <https://doi.org/10.3390/v1031325>
115. Huang, J., Fan, Z., Li, H. F., Tian, G., & Hu, J. (2004). First report of tomato mosaic virus on *Hibiscus rosa-sinensis* in China. *Plant Disease*, 88(6), 683-683. <https://doi.org/10.1094/pdis.2004.88.6.683c>
116. Huelsenbeck, J.P. & Ronquist, F. (2001). MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics*, 17, 754–755
117. Inoue-Nagata, A.K., Jordan, R., Kreuze, J., Li, F., López-Moya, J.J., Mäkinen, K. et al. (2022) ICTV virus taxonomy profile: *Potyviridae* 2022. *Journal of General Virology*, **103**. <https://doi.org/10.1099/jgv.0.001738>
118. Janz, N., S. Nylin, and K. Nyblom. 2001. Evolutionary dynamics of host plant specialization: a case study of the tribe Nymphalini. *Evolution* 55: 783–796

119. Jombart, T. (2008). Adegnet: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, 24, 1403–1405.
120. Jones R. A. (2009). Plant virus emergence and evolution: origins, new encounter scenarios, factors driving emergence, effects of changing world conditions, and prospects for control. *Virus research*, 141(2), 113–130. <https://doi.org/10.1016/j.virusres.2008.07.028>
121. Jones RAC, Koenig R, and Lesemann DE. 1980. *Pepino Mosaic Virus*, a new potexvirus from pepino (*Solanum muricatum*). *Ann. Appl. Biol.* 94:61–68.
122. Jones, K. E., Patel, N. G., Levy, M. A., Storeygard, A., Balk, D., Gittleman, J. L., & Daszak, P. (2008). Global trends in emerging infectious diseases. *Nature*, 451(7181), 990–993. <https://doi.org/10.1038/nature06536>
123. Jones, R. A. C. (2014). Plant virus ecology and epidemiology: historical perspectives, recent progress and future prospects. *Annals of Applied Biology*, 164(3), 320–347. <https://doi.org/10.1111/aab.12123>
124. Jones, R.A.C. 2018. Plant and insect viruses in managed and natural environments: Novel and neglected transmission pathways. *Advances in virus research* 101: 149–187. <https://doi.org/10.1016/bs.aivir.2018.02.006>
125. Karvonen, A., Jokela, J., & Laine, A. L. (2019). Importance of Sequence and Timing in Parasite Coinfections. *Trends in parasitology*, 35(2), 109–118. <https://doi.org/10.1016/j.pt.2018.11.007>
126. Katagiri, F., & Tsuda, K. (2010). Understanding the plant immune system. *Molecular plant-microbe interactions : MPMI*, 23(12), 1531–1536. <https://doi.org/10.1094/MPMI-04-10-0099>
127. Khalili, M., Candresse, T., Koloniuk, I., Safarova, D., Brans, Y., Faure, C., Delmas, M., Massart, S., Aranda, M. A., Caglayan, K., Decroocq, V., Drogoudi, P., Glasa, M., Pantelidis, G., Navratil, M., Latour, F., Spak, J., Pribylova, J., Mihalik, D., Palmisano, F., ... Marais, A. (2023). The Expanding Menagerie of *Prunus*-Infecting Luteoviruses. *Phytopathology*, 113(2), 345–354. <https://doi.org/10.1094/PHYTO-06-22-0203-R>

128. King AMQ, Adams MJ, Carstens EB, Lefkowitz EJ (2012) Ninth report of the international committee on taxonomy of viruses. Elsevier, London
129. Klein, E., Brault, V., Klein, D., Weyens, G., Lefebvre, M., Ziegler-Graff, V., ... & Gilmer, D. (2013). Divergence of host range and biological properties between natural isolate and full-length infectious cdna clone of the beet mild yellowing virus 2itb. *Molecular Plant Pathology*, 15(1), 22-30. <https://doi.org/10.1111/mpp.12061>
130. Kouyos, R. D., Silander, O. K., & Bonhoeffer, S. (2007). Epistasis between deleterious mutations and the evolution of recombination. *Trends in ecology & evolution*, 22(6), 308–315. <https://doi.org/10.1016/j.tree.2007.02.014>
131. Kreuze, J. F., Perez, A., Untiveros, M., Quispe, D., Fuentes, S., Barker, I., & Simon, R. (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), 1–7. <https://doi.org/10.1016/j.virol.2009.03.024>
132. Krueger, E. N., Beckett, R. J., Gray, S. M., & Miller, W. A. (2013). The complete nucleotide sequence of the genome of Barley yellow dwarf virus-RMV reveals it to be a new Polerovirus distantly related to other yellow dwarf viruses. *Frontiers in microbiology*, 4, 205. <https://doi.org/10.3389/fmicb.2013.00205>
133. Ksenofontov, A., Paalme, V., Arutyunyan, A., Semenyuk, P., Fedorova, N., Rumvolt, R., ... & Dobrov, E. (2013). Partially disordered structure in intravirus coat protein of potyvirus potato virus a. *Plos One*, 8(7), e67830. <https://doi.org/10.1371/journal.pone.0067830>
134. Kumar, S., Stecher, G., Li, M., Knyaz, C. & Tamura, K. (2018). MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.*, 35, 1547–1549.
135. Kumari, S. G., Makkouk, K. M., Attar, N., Ghulam, W., & Lesemann, D. (2004). First report of chickpea chlorotic dwarf virus infecting spring chickpea in syria. *Plant Disease*, 88(4), 424-424. <https://doi.org/10.1094/pdis.2004.88.4.424c>
136. Lacroix C, Jolles A, Seabloom EW, Power AG, Mitchell CE, Borer ET. 2014 Non-random biodiversity loss underlies predictable increases in viral disease prevalence. *J. R. Soc. Interface* 11, 20130947. [doi:10.1098/rsif.2013.0947](https://doi.org/10.1098/rsif.2013.0947)

137. Lapidot, M.; Guenoune-Gelbart, D.; Leibman, D.; Holdengreber, V.; Davidovitz, M.; MachBash, Z.; Klieman-Shoval, S.; Cohen, S.; Gal-On, A. Pelargonium Zonate Spot Virus Is Transmitted Vertically via Seed and Pollen in Tomato. *Phytopathology* 2010, 100, 798–804.
138. Latham, L. J., & Jones, R. A. (2004). Carrot virus Y: symptoms, losses, incidence, epidemiology and control. *Virus research*, 100(1), 89–99.  
<https://doi.org/10.1016/j.virusres.2003.12.017>
139. LaTourrette, K., Holste, N. M., & Garcia-Ruiz, H. (2021). Polerovirus genomic variation. *Virus evolution*, 7(2), veab102. <https://doi.org/10.1093/ve/veab102>
140. Laviña, A., Aramburu, J., and Moriones, E. 1996. Occurrence of tomatospotted wilt and cucumber mosaic viruses in field-grown tomato crops and associated weeds in northeastern Spain. *Plant Pathol.* 45:837-842
141. Lefeuvre, P., Martin, D. P. , Elena, S. F. , Shepherd, D. N. , Roumagnac, P. , & Varsani, A. (2019). Evolution and ecology of plant viruses. *Nature Reviews Microbiology*, 17(10), 632–644.
142. Lemaire, O., Herrbach, E., Stevens, M., Bouchery, Y. and Smith, H.G.(1995) Detection of sugar beet-infecting beet mild yellowing luteovirus isolates with a specific RNA probe. *Phytopathology*, 85, 1513–1518.
143. Levitzky, N., Smith, E., Lachman, O., Luria, N., Mizrahi, Y., Bakelman, H., Sela, N., Laskar, O., Milrot, E., and Dombrovsky, A. 2019. The bumblebee *Bombus terrestris* carries a primary inoculum of Tomato brown rugose fruit virus contributing to disease spread in tomatoes. *PLoS ONE* 14: e0210871.
144. Levitzky, N., Smith, E., Levi, A., & Pappu, H. R. (2019). Role of pollen in horizontal transmission of plant viruses: a new frontier in plant virology research. *Phytopathology*, 109(2), 183-190.
145. Li, H. & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, 25, 1754–1760
146. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, 25, 2078–2079.

147. Liu X, Chen L, Liu M, Garcia-Guzman G, Gilbert GS, Zhou S (2020) Dilution effect of plant diversity on infectious diseases: latitudinal trend and biological context dependence. *Oikos* 129, 457-465. <https://doi.org/10.1111/oik.07027>
148. Lovisolo, O., Hull, R., & Rösler, O. (2003). Coevolution of viruses with hosts and vectors and possible paleontology. *Advances in virus research*, 62, 325–379. [https://doi.org/10.1016/s0065-3527\(03\)62006-3](https://doi.org/10.1016/s0065-3527(03)62006-3)
149. García-Ordóñez, L., Pagán, I. Vertical and horizontal transmission of plant viruses: two extremes of a continuum? *npj Viruses* 2, 18 (2024). <https://doi.org/10.1038/s44298-024-00030-8>
150. Luis Alonso-Prados, J., Luis-Arteaga, M., Alvarez, J.M. *et al.* Epidemics of Aphid-transmitted Viruses in Melon Crops in Spain. *European Journal of Plant Pathology* 109, 129–138 (2003). <https://doi.org/10.1023/A:1022598417979>
151. Lundholm, J.T. 2009. Plant species diversity and environmental heterogeneity: spatial scale and competing hypotheses. *Journal of Vegetation Science* 20: 377–391.
152. Ma, Y., Marais, A., Lefebvre, M., Theil, S., Svanella-Dumas, L., Faure, C., ... & Candresse, T. (2019). Phytoviroome analysis of wild plant populations: comparison of double-stranded rna and virion-associated nucleic acid metagenomic approaches. *Journal of Virology*, 94(1). <https://doi.org/10.1128/jvi.01462-19>
153. Maachi A, Donaire L, Hernando Y, Aranda M.A.( 2022).Genetic Differentiation and Migration Fluxes of Viruses from Melon Crops and Crop Edge Weeds. *J Virol*96:e00421-22.<https://doi.org/10.1128/jvi.00421-22>
154. Maclot, F., Debue, V., Malmstrom, C. M., Filloux, D., Roumagnac, P., Eck, M., Tamisier, L., Blouin, A. G., Candresse, T., & Massart, S. (2023). Long-Term Anthropogenic Management and Associated Loss of Plant Diversity Deeply Impact Virome Richness and Composition of *Poaceae* Communities. *Microbiology spectrum*, 11(2), e0485022. Advance online publication. <https://doi.org/10.1128/spectrum.04850-22>
155. Maksimović Carvalho Ferreira, O., Lengar, Ž., Kogej, Z., Bačnik, K., Bajde, I., Milavec, M., Županič, A., Mehle, N., Kutnjak, D., Ravnika, M., & Gutierrez-Aguirre, I. (2022). Evaluation

- of Methods and Processes for Robust Monitoring of SARS-CoV-2 in Wastewater. *Food and environmental virology*, 14(4), 384–400. <https://doi.org/10.1007/s12560-022-09533-0>
156. Maksimović O, Bačnik K, Selda Rivarez MP, Vučurović A, Mehle N, Ravnikar M, Gutiérrez-Aguirre I, Kutnjak D. (2024). Virome analysis of irrigation water sources provides extensive insights into the diversity and distribution of plant viruses in agroecosystems. Volume 249, 120712, <https://doi.org/10.1016/j.watres.2023.120712>.
  157. Malmstrom, C. M., Hughes, C. C., Newton, L. A., & Stoner, C. J. (2005). Virus infection in remnant native grassland plants: interactions between human disturbance and a native virus. *Ecological Applications*, 15(6), 2022–2035.
  158. Malmstrom, C. M., Hughes, C. C., Newton, L. A., & Stoner, C. J. (2005). Virus infection in remnant native bunchgrasses from invaded California grasslands. *The New phytologist*, 168(1), 217–230. <https://doi.org/10.1111/j.1469-8137.2005.01479.x>
  159. Malpica JM, Sacristán S, Fraile A, García-Arenal F (2006) Association and Host Selectivity in Multi-Host Pathogens. *PLoS ONE* 1(1): e41. <https://doi.org/10.1371/journal.pone.0000041>
  160. Mandel JR, Brunet J (2019) Gene flow in carrot. In *The Carrot Genome. Compendium of Plant Genomes*. Simon P, Iorizzo M, Grzebelus D, Baranski, R (Eds), Springer, Cham (Switzerland), pp 59–76. [https://doi.org/10.1007/978-3-030-03389-7\\_4](https://doi.org/10.1007/978-3-030-03389-7_4)
  161. Marais A, Faure C, Bergey B, Candresse T (2018) Viral double-stranded RNAs (dsRNAs) from plants: alternative nucleic acid substrates for high-throughput sequencing. *Methods Mol Biol* 1746:45–53. [https://doi.org/10.1007/978-1-4939-7683-6\\_4](https://doi.org/10.1007/978-1-4939-7683-6_4)
  162. Marchetto, K. M., and Power, A. G. 2018. Coinfection timing drives host population dynamics through changes in virulence. **Am. Nat.** 191:173–183.
  163. Márquez, L. M., Redman, R. S., Rodriguez, R. J., & Roossinck, M. J. (2007). A virus in a fungus in a plant: three-way symbiosis required for thermal tolerance. *Science (New York, N.Y.)*, 315(5811), 513–515. <https://doi.org/10.1126/science.1136237>

164. Mauck, K. E., Bosque-Pérez, N. A., & Eigenbrode, S. D. (2012). Transmission mechanisms shape pathogen effects on host-vector interactions: evidence from plant viruses. *Functional Ecology*, 26(5), 1162-1175.
165. Mauck, K., Moraes, C., & Mescher, M. (2014). Biochemical and physiological mechanisms underlying effects of Cucumber mosaic virus on host-plant traits that mediate transmission by aphid vectors.. *Plant, cell & environment*, 37 6, 1427-39 .  
<https://doi.org/10.1111/pce.12249>.
166. May RM, Nowak MA (1995) Coinfection and the evolution of parasite virulence. *Proc. R. Soc. London B*. 261: 209–215.
167. Mayo, M. A., and D'Arcy, C. J. 1999. Family Luteoviridae: A reclassification of Luteoviruses. Pages 15-22 in: *The Luteoviridae*. H. G. Smith and H. Barker, eds. CABI Publishing, Wallingford, U.K.
168. Meentemeyer, R. K., Haas, S. E., & Václavík, T. (2012). Landscape epidemiology of emerging infectious diseases in natural and human-altered ecosystems. *Annual review of phytopathology*, 50, 379–402. <https://doi.org/10.1146/annurev-phyto-081211-172938>
169. McClelland EE, Penn DJ, Potts WK (2003) Major histocompatibility complex heterozygote superiority during coinfection. *Infect. Immun.* 71: 2079–2086.
170. McDonald B.A., Stukenbrock E.H. 2016. Rapid emergence of pathogens in agro-ecosystems: global threats to agricultural sustainability and food security. *Philosophical Transactions of the Royal Society B: Biological Sciences* 371: 20160026.  
<https://doi.org/10.1098/rstb.2016.0026>
171. McGavin, W.J.; Cock, P.J.A.; Macfarlane, S.A. (2010). Partial sequence and RT-PCR diagnostic test for the plant rhabdovirus Raspberry vein chlorosis virus. *Plant Pathol.* 60, 462–467.
172. McKern, N., Shukla, D., Barnett, O., Vetten, H., Dijkstra, J., Whittaker, L., ... & Ward, C. (1992). Coat protein properties suggest that azuki bean mosaic virus, blackeye cowpea mosaic virus, peanut stripe virus, and three isolates from soybean are all strains of the same potyvirus. *Intervirology*, 33(3), 121-134. <https://doi.org/10.1159/000150241>

173. McLeish M.J., Fraile A., García-Arenal F. (2021b). Population Genomics of Plant Viruses: The Ecology and Evolution of Virus Emergence. *Phytopathology* 111: 32–39. <https://doi.org/10.1094/PHYTO-08-20-0355-FI>
174. McLeish M.J., Sacristán S., Fraile A., García-Arenal F. 2017. Scale dependencies and generalism in host use shape virus prevalence. *Proceedings of the Royal Society B: Biological Sciences* 284: 20172066. <https://doi.org/10.1098/rspb.2017.2066>
175. McLeish M.J., Sacristán S., Fraile A., García-Arenal F. 2019b. Coinfection Organizes Epidemiological Networks of Viruses and Hosts and Reveals Hubs of Transmission. *Phytopathology* 109: 1003–1010. <https://doi.org/10.1094/PHYTO-08-18-0293-R>
176. McLeish, M. J., Fraile, A., & García-Arenal, F. (2019a). Evolution of plant virus interactions: host range and virus emergence. *Current opinion in virology*, 34, 50-55.
177. McLeish, M. J., Fraile, A., García-Arenal, F., & Gilbert, G. S. (2018). Estimation of the host range of a virus. *Biological Reviews*, 93(2), 1381-1397.
178. McLeish, M. J., Peláez, A., Pagán, I., Gavilán, R., Fraile, A., and García-Arenal, F. (2021a). Structuring of plant communities across agricultural landscape mosaics: the importance of connectivity and the scale of effect. *BMC Ecol. Evol.* 21: 173. <https://doi.org/10.1186/s12862-021-01903-9>
179. McLeish, M. J., Zamfir, A. D., Babalola, B. M., Peláez, A., Fraile, A., and García-Arenal, F. (2022). Metagenomics show high spatiotemporal virus diversity and ecological compartmentalisation: viromes of melon, Cucumis melo, crops and adjacent wild 604 communities. *Virus Evol.* 8, <https://doi.org/10.1093/ve/veac095>
180. McLeish, M., Peláez, A., Pagán, I., Gavilán, R. G., Fraile, A., & García-Arenal, F. (2024). Plant virus community structuring is shaped by habitat heterogeneity and traits for host plant resource utilisation. *The New phytologist*, 10.1111/nph.20054. Advance online publication. <https://doi.org/10.1111/nph.20054>
181. McLeish, M., Zamfir, A., Babalola, B., Peláez, A., Fraile, A., & García-Arenal, F. (2022). Metagenomics show high spatiotemporal virus diversity and ecological compartmentalisation: virus infections of melon, cucumis melo, crops, and adjacent wild communities. *Virus Evolution*, 8(2). <https://doi.org/10.1093/ve/veac095>

182. McLeish, M.; Sacristán, S.; Fraile, A.; García-Arenal, F. (2019b). Coinfection Organizes Epidemiological Networks of Viruses and Hosts and Reveals Hubs of Transmission. *Phytopathology* 109, 1003–1010. <https://doi.org/10.1094/PHYTO-08-18 0293-R>
183. McLeish, M.J., Fraile, A., García-Arenal, F. (2018). Ecological Complexity in Plant Virus Host Range Evolution. *Advances in virus research*, 101, 293-339. <https://doi.org/10.1016/bs.aivir.2018.02.009>
184. McLeish, M.J., Peláez, A., Pagán, I., Gavilán, R., Fraile, A., García-Arenal, F. 2021b. Structuring of plant communities across agricultural landscape mosaics: the importance of connectivity and the scale of effect. *BMC Ecology and Evolution*, 21, 1-12. <https://doi.org/10.1186/s12862-021-01903-9>
185. Mehle, N., Gutiérrez-Aguirre, I., Kutnjak, D., & Ravnikar, M. (2018). Water-Mediated Transmission of Plant, Animal, and Human Viruses. *Advances in virus research*, 101, 85–128. <https://doi.org/10.1016/bs.aivir.2018.02.004>
186. Mekuria, G., Ramesh, S. A., Alberts, E., Bertozzi, T., Wirthensohn, M., Collins, G., & Sedgley, M. (2003). Comparison of ELISA and RT-PCR for the detection of Prunus necrotic ring spot virus and prune dwarf virus in almond (*Prunus dulcis*). *Journal of virological methods*, 114(1), 65–69. <https://doi.org/10.1016/j.jviromet.2003.08.014>
187. Merike Sömera, Denis Fargette, Eugénie Hébrard, Cecilia Sarmiento, and the ICTV Report Consortium, [ICTV Virus Taxonomy Profile: Solemoviridae 2021](#), *Journal of General Virology* 2021;102:001707
188. Miller WA, Dinesh-Kumar SP, Paul CP. Luteovirus gene expression. *Crit Rev Plant Sci*. 1995;14(3):179. <https://doi.org/10.1080/07352689509701926>.
189. Miller WA, Hercus T, Waterhouse PM, Gerlach WL (1991) A satellite RNA of barley yellow dwarf virus contains a novel hammerhead structure in the self-cleavage domain. *Virology* 183:711–720. [https://doi.org/10.1016/0042-6822\(91\)91000-7](https://doi.org/10.1016/0042-6822(91)91000-7)
190. Miller WA, Silver SL (1991) Alternative tertiary structure attenuates self-cleavage of the ribozyme in the satellite RNA of barley yellow dwarf virus. *Nucleic Acids Res* 19:5313–5320. <https://doi.org/10.1093/nar/19.19.5313>

191. Miller, W. A., and Rasochova, L. 1997. Barley yellow dwarf viruses. *Annu. Rev. Phytopathol.* 35:167-190
192. Miller, W. A., Koev, G., and Mohan, B. R. 1997. Are there risks associated with transgenic resistance to luteoviruses? *Plant Dis.* 81:700-710.
193. Molina J.A., Martín-Sanz J.P., Valverde-Asenjo I., Sánchez-Jiménez A., Quintana J.R. 2023. Mediterranean grassland succession as an indicator of changes in ecosystem biodiversity and functionality. *Biodiversity and Conservation* 32: 95–118. <https://doi.org/10.1007/s10531-022-02481-y>
194. Moran, J., van Rijswijk, B., Traicevski, V., Kitajima, E.W., Mackenzie, A.M. & Gibbs, A.J. (2002) Potyviruses, novel and known, in cultivated and wild species of the family Apiaceae in Australia. *Archives of Virology*, 147, 1855–1867. <https://doi.org/10.1007/s00705-002-0865-8>
195. Moreno-Perez MG, Pagán Pagan I, Aragon-Caballero L, Caceres F, Fraile A, Garcia-Arenal F: Ecological and genetic determinants of Pepino mosaic virus emergence. *J Virol* 2014, 88:3359-3368. <https://doi.org/10.1128/JVI.02980-13>
196. Morris, C., & Moury, B. (2019). Revisiting the Concept of Host Range of Plant Pathogens. *Annual review of phytopathology*. <https://doi.org/10.1146/annurev-phyto-082718-100034>.
197. Morris, T.J.; Dodds, J. (1979). Isolation and analysis of double-stranded RNA from virus-infected plant and fungal tissue. *Phytopathology*, 69, 854–858.
198. Moubset, O., François, S., Maclot, F., Palanga, E., Julian, C., Claude, L., Fernandez, E., Rott, P., Daugrois, J. H., Antoine-Lorquin, A., Bernardo, P., Blouin, A. G., Temple, C., Krabberger, S., Fontenele, R. S., Harkins, G. W., Ma, Y., Marais, A., Candresse, T., Chéhida, S. B., ... Roumagnac, P. (2022). Virion-Associated Nucleic Acid-Based Metagenomics: A Decade of Advances in Molecular Characterization of Plant Viruses. *Phytopathology*, 112(11), 2253–2272. <https://doi.org/10.1094/PHYTO-03-22-0096-RVW>
199. Moury, B., & Desbiez, C. (2020). Host Range Evolution of Potyviruses: A Global Phylogenetic Analysis. *Viruses*, 12(1), 111. <https://doi.org/10.3390/v12010111>

200. Moury, B., & Simon, V. (2017). Genetic divergence and evolution in plant viruses. *Annual Review of Phytopathology*, 45, 77-100.
201. Moury, B., Fabre, F., Hébrard, E., & Froissart, R. (2017). Determinants of host species range in plant viruses. *The Journal of general virology*, 98(4), 862–873. <https://doi.org/10.1099/jgv.0.000742>
202. Munne, K., Bhanothu, V., Bhor, V., Patel, V., Mahale, S. D., & Pande, S. (2021). Detection of SARS-CoV-2 infection by RT-PCR test: factors influencing interpretation of results. *Virusdisease*, 32(2), 187–189. <https://doi.org/10.1007/s13337-021-00692-5>
203. Murrant, A.F. (1993) Complexes of transmission-dependent and helper viruses In: *Diagnosis of Plant Virus Diseases* (Matthews R.E.F., ed.), pp. 334–357. Boca Raton, FL: CRC Press.
204. Nagelkerke, N. J. D. (1991) A note on a general definition of the coefficient of determination. *Biometrika*, 78: 691-692.
205. Nakaune, R.; Nakano, M. (2006). Efficient methods for sample processing and cDNA synthesis by RT-PCR for the detection of grapevine viruses and viroids. *J. Virol. Methods*, 134, 244–249.
206. Navarro B, Rubino L, di Serio F (2017) Small Circular Satellite RNAs. In: Hadidi A, Flores R, Randles JW, Palukaitis P (eds) *Viroids and satellites*. Academic Press, pp 659–669. <https://doi.org/10.1016/B978-0-12-801498-1.00061-9>
207. Navie S. (2004). *Declared Plants of Australia. An identification and information system*. Centre for Biological Information Technology: Brisbane.
208. Nicaise, V. (2014). Crop immunity against viruses: outcomes and future challenges. *Frontiers in Plant Science*, 5. <https://doi.org/10.3389/fpls.2014.00660>
209. Norberg, A., Susi, H., Sallinen, S., Safdari, P., Clark, N., & Laine, A. (2021). Direct and indirect viral associations predict coexistence in wild plant virus communities.. <https://doi.org/10.21203/rs.3.rs-741379/v1>

210. Norton E.C, Dowd B.E (2018). Log Odds and the Interpretation of Logit Models. *Health Serv Res.* Apr;53(2):859-878. doi: 10.1111/1475-6773.12712. Epub 2017 May 30. PMID: 28560732; PMCID: PMC5867187.
211. Nouri, S., Arevalo, R., Falk, B. W., & Groves, R. L. (2014). Genetic structure and molecular variability of Cucumber mosaic virus isolates in the United States. *PloS one*, 9(5), e96582. <https://doi.org/10.1371/journal.pone.0096582>
212. Olmos, A.; Capote, N.; Bertolini, E.; Cambra, M. (2007). Molecular diagnostic methods for plant viruses. In *Biotechnology and Plant Disease Management*; Punja, Z.K., de Boer, S.H., Sanfaçon, H., Eds.; CAB International: Wallingford, UK; pp. 227–249.
213. Pagán I, Córdoba-Selles MC, Martínez-Priego L, Fraile A, Malpica JM, Jordá C, and García-Arenal F. 2006. Genetic structure of the population of *Pepino Mosaic Virus* infecting tomato crops in Spain. *Phytopathology* 96:274–279.
214. Pagán I, González-Jara P, Moreno-Letelier A, Rodelo-Urrego M, Fraile A, Piñero D, García-Arenal F (2012) Effect of biodiversity changes in disease risk: exploring disease emergence in a plant-virus system. *PLOS Pathogens*, 8, 47. <https://doi.org/10.1371/journal.ppat.1002796>
215. Pagán, I. Movement Between Plants: Vertical Transmission. In *Cucumber Mosaic Virus*; The American Phytopathological Society: St. Paul, MN, USA, 2019; pp. 185–198.
216. Palukaitis, P., & García-Arenal, F. (2003). Cucumoviruses. *Advances in virus research*, 62, 241–323. [https://doi.org/10.1016/s0065-3527\(03\)62005-1](https://doi.org/10.1016/s0065-3527(03)62005-1)
217. Palukaitis, P.; García-Arenal, F. (2019) *Cucumber Mosaic Virus*; APS Press: Washington, DC, USA.
218. Parker, J., Rambaut, A. & Pybus, O.G. (2008). Correlating viral phenotypes with phylogeny: Accounting for phylogenetic uncertainty. *Infect. Genet. Evol.*, 8, 239–246.
219. Parrish, C. R., Holmes, E. C., Morens, D. M., Park, E. C., Burke, D. S., Calisher, C. H., Laughlin, C. A., Saif, L. J., & Daszak, P. (2008). Cross-species virus transmission and the emergence of new epidemic diseases. *Microbiology and molecular biology reviews* : MMBR, 72(3), 457–470. <https://doi.org/10.1128/MMBR.00004-08>

220. Pecman, A., Adams, I., Gutiérrez-Aguirre, I., Fox, A., Boonham, N., Ravnika, M., & Kutnjak, D. (2022). Systematic Comparison of Nanopore and Illumina Sequencing for the Detection of Plant Viruses and Viroids Using Total RNA Sequencing Approach. *Frontiers in microbiology*, 13, 883921. <https://doi.org/10.3389/fmicb.2022.883921>
221. Pecman, A., Kutnjak, D., Gutiérrez-Aguirre, I., Adams, I., Fox, A., Boonham, N., & Ravnika, M. (2017). Next Generation Sequencing for Detection and Discovery of Plant Viruses and Viroids: Comparison of Two Approaches. *Frontiers in microbiology*, 8, 1998. <https://doi.org/10.3389/fmicb.2017.01998>
222. Petruzzello M (2019) List of plants in the family Apiaceae. Encyclopedia Britannica. <https://www.britannica.com/topic/list-of-plants-in-the-family-APIACEAE-2038061>. Accessed 06 April 2024
223. Phillips, P.C (2008). Epistasis — the essential role of gene interactions in the structure and evolution of genetic systems. *Nat Rev Genet* 9, 855–867. <https://doi.org/10.1038/nrg2452>
224. Poggio, S.L., Chaneton, E.J. & Ghersa, C.M. (2010). Landscape complexity differentially affects alpha, beta, and gamma diversities of plants occurring in fencerows and crop fields. *Biol. Conserv.*, 143, 2477–2486.
225. Poulin R. (2010). Network analysis shining light on parasite ecology and diversity. *Trends in parasitology*, 26(10), 492–498. <https://doi.org/10.1016/j.pt.2010.05.008>
226. Power, A. G., & Mitchell, C. E. (2004). Pathogen spillover in disease epidemics. *The American Naturalist*, 164(S5), S79–S89.
227. Que F, Hou XL, Wang GL, Xu ZS, Tan GF, Li T, Wang YH, Khadr A, Xiong AS (2019) Advances in research on the carrot, an important root vegetable in the Apiaceae family. *Hortic. Res.* 6, 69. <https://doi.org/10.1038/s41438-019-0150-6>
228. R Core Team. (2018). R: A language and environment for statistical computing. R Found. Stat. Comput. Vienna, Austria.
229. Rambaut, A., Suchard, M. A., Xie, D., & Drummond, A. J. (2014). Tracer v1.7.2. Retrieved from <http://beast.bio.ed.ac.uk/Tracer>.

230. Rasochová L, Miller WA (1996) Satellite RNA of barley yellow dwarf-RPV virus reduces accumulation of RPV helper virus RNA and attenuates RPV symptoms in oats. *Mol Plant Microb Interact* 9:646–650. <https://doi.org/10.1094/MPMI-9-0646>
231. Rasochová L, Passmore BK, Falk BW, Miller WA (1997) The satellite RNA of barley yellow dwarf virus-RPV is supported by beet western yellows virus in dicotyledonous protoplasts and plants. *Virology* 231:182–191. <https://doi.org/10.1006/viro.1997.8532>
232. Raybould, A. F., Maskell, L. C., Edwards, M. L., Cooper, J. I., and Gray, A. J. 1999. The incidence and spatial distribution of viruses in natural populations of *Brassica oleracea*. *New Phytol.* 141:265-275
233. Reinbold C, Gildow FE, Herrbach E, Ziegler-Graff V, Gonçalves MC, Van Den Heuvel J, Brault V (2001) Studies on the role of the minor capsid protein in transport of Beet western yellows virus through *Myzus persicae*. *Journal of General Virology* 82:1995–2007
234. Rhee, Y.; Tzfira, T.; Chen, M.H.; Waigmann, E.; Citovsky, V. Cell-to-Cell Movement of Tobacco Mosaic Virus: Enigmas and Explanations. *Mol. Plant Pathol.* 2000, 1, 33–39.
235. Richert-Poggeler KR, Franzke K, Hipp K, Kleespies RG (2018) Electron microscopy methods for virus diagnosis and high resolution analysis of viruses. *Front Microbiol* 9: 3255. <https://doi.org/10.3389/fmicb.2018.03255>
236. Riechmann, J. L., Laín, S., & García, J. A. (1992). Highlights and prospects of potyvirus molecular biology. *The Journal of general virology*, 73 ( Pt 1), 1–16. <https://doi.org/10.1099/0022-1317-73-1-1>
237. Rivarez, M. P. S., Pecman, A., Bačnik, K., Maksimović, O., Vučurović, A., Seljak, G., Mehle, N., Gutiérrez-Aguirre, I., Ravnikar, M., & Kutnjak, D. (2023). In-depth study of tomato and weed viromes reveals undiscovered plant virus diversity in an agroecosystem. *Microbiome*, 11(1), 60. <https://doi.org/10.1186/s40168-023-01500-6>
238. Rivarez, M. P. S., Vučurović, A., Mehle, N., Ravnikar, M., and Kutnjak, D. (2021). Global advances in tomato virome research: Current status and the impact of high-throughput sequencing. **Front. Microbiol.** 12:671925. <https://doi.org/10.3389/fmicb.2021.671925>
239. Rochow, W. F. 1969. Biological properties of four isolates of barley yellow dwarf virus. *Phytopathology* 59:1580-1589

240. Rochow, W. F. 1970. Barley yellow dwarf virus: Phenotypic mixing and vector specificity. *Science* 167:875-878
241. Rochow, W. F., and Muller, I. 1974. Mixed infections of barley yellow dwarf virus isolates in winter grains. *Plant Dis. Rep.* 58:472-475.
242. Rochow, W. F., Muller, I., and Gildow, F. E. 1983. Interference between two luteoviruses in an aphid: Lack of reciprocal competition. *Phytopathology* 73:919-922
243. Rodríguez-Negrete, E. A., Morales-Aguilar, J. J., Domínguez-Duran, G., Torres-Devora, G., Camacho-Beltrán, E., Leyva-López, N. E., Voloudakis, A. E., Bejarano, E. R., & Méndez-Lozano, J. (2019). High-Throughput Sequencing Reveals Differential Begomovirus Species Diversity in Non-Cultivated Plants in Northern-Pacific Mexico. *Viruses*, 11(7), 594. <https://doi.org/10.3390/v11070594>
244. Rojas, M.R.; Gilbertson, R.J.; Russell, D.R.; Maxwell, D.P.(1993). Use of Degenerate Primers in the Polymerase Chain Reaction to Detect Whitefly-Transmitted Geminiviruses. *Plant Dis.* 77, 340–347.
245. Roossinck M. J. (2012). Plant virus metagenomics: biodiversity and ecology. *Annual review of genetics*, 46, 359–369. <https://doi.org/10.1146/annurev-genet-110711-155600>
246. Roossinck M.J (2012a) Persistent plant viruses: molecular hitchhikers or epigenetic elements. In *Viruses: essential agents of life*, Witzany G (Ed), Springer, New York (USA), pp. 177-186.
247. Roossinck MJ (2017) Deep sequencing for discovery and evolutionary analysis of plant viruses. *Virus Res* 239: 82–86. <https://doi.org/10.1016/j.virusres.2016.11.019> PMID: 27876625
248. Roossinck MJ, Sleat D, Palukaitis P (1992) Satellite RNAs of plant viruses: structures and biological effects. *Microbiol Rev* 56:265–279. <https://doi.org/10.1128/mr.56.2.265-279.1992>
249. Roossinck, M. J., & García-Arenal, F. (2015). Ecosystem simplification, biodiversity loss and plant virus emergence. *Current opinion in virology*, 10, 56–62. <https://doi.org/10.1016/j.coviro.2015.01.005>

250. Roossinck, M. J., & García-Arenal, F. (2015). Ecosystem simplification, biodiversity loss and plant virus emergence. *Current opinion in virology*, 10, 56–62.  
<https://doi.org/10.1016/j.coviro.2015.01.005>
251. Roossinck, M. J., Martin, D. P., & Roumagnac, P. (2015). Plant Virus Metagenomics: Advances in Virus Discovery. *Phytopathology*, 105(6), 716–727.  
<https://doi.org/10.1094/PHYTO-12-14-0356-RVW>
252. Roossinck, M.J., Martin, D.P., Roumagnac P. (2015) Plant virus metagenomics: advances in virus discovery. *Phytopathology*, 105 , pp. 716-727
253. Rott M, Xiang Y, Boyes I, Belton M, Saeed H et al. (2017) Application of next generation sequencing for diagnostic testing of tree fruit viruses and viroids. *Plant Dis* 101 (8): 1489–1499. <https://doi.org/10.1094/PDIS-03-17-0306-RE>
254. Roy A. S., Smith I. M. 1994; Plum pox situation in Europe. *EPPO Bull* 24:515–523
255. Rozado-Aguirre Z, Adams I, Collins L, Fox A, Dickinson M, Boonham N (2016) Detection and transmission of carrot torrado virus, a novel putative member of the Torradovirus genus. *J Virol Methods* 235:119. <https://doi.org/10.1016/j.jviromet.2016.05.018>
256. Rubio L., Galipienso L., Ferriol I. 2020. Detection of Plant Viruses and Disease Management: Relevance of Genetic Diversity and Evolution. *Frontiers in Plant Science* 11: <https://www.frontiersin.org/journals/plant-science/articles/10.3389/fpls.2020.01092>
257. Ryabov, E. V., Fraser, G., Mayo, M. A., Barker, H., & Taliansky, M. (2001). Umbravirus gene expression helps potato leafroll virus to invade mesophyll tissues and to be transmitted mechanically between plants. *Virology*, 286(2), 363–372.  
<https://doi.org/10.1006/viro.2001.0982>
258. Sacristán, S., Fraile, A., & García-Arenal, F. (2004). Population Dynamics of Cucumber mosaic virus in Melon Crops and in Weeds in Central Spain. *Phytopathology*, 94(9), 992–998.
259. Sacristán, S., Fraile, A., and García-Arenal, F. 2004. Population dynamics of *Cucumber mosaic virus* in melon crops and in weeds in central Spain. *Phytopathology* 94:992-998. <https://doi.org/10.1094/PHYTO.2004.94.9.992>

260. Saiki, R. K., Scharf, S., Faloona, F., Mullis, K. B., Horn, G. T., Erlich, H. A., & Arnheim, N. (1985). Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science (New York, N.Y.)*, 230(4732), 1350–1354. <https://doi.org/10.1126/science.2999980>
261. Salem, N. M., Abumuslem, M., Turina, M., Samarah, N., Sulaiman, A., Abu-Irmaileh, B., and Ata, Y. 2022. New weed hosts for tomato brown rugose fruit virus in wild Mediterranean vegetation. *Plants* 11: 2287
262. Salem, N.; Mansour, A.; Ciuffo, M.; Falk, B. W.; Turina, M. (2015). "A new tobamovirus infecting tomato crops in Jordan". *Archives of Virology*. 161 (2): 503–506. doi:10.1007/s00705-015-2677-7
263. Sallinen, S., Norberg, A., Susi, H., & Laine, A. L. (2020). Intraspecific host variation plays a key role in virus community assembly. *Nature communications*, 11(1), 5610. <https://doi.org/10.1038/s41467-020-19273-z>
264. Sanjuán, R. 2010. Mutational fitness effects in RNA and single-stranded DNA viruses: Common patterns revealed by site-directed mutagenesis studies. *Philos. Trans. R. Soc. B Biol. Sci.* 365:1975-1982.
265. Santala, J., & Valkonen, J. P. T. (2018). Sensitivity of Small RNA-Based Detection of Plant Viruses. *Frontiers in microbiology*, 9, 939. <https://doi.org/10.3389/fmicb.2018.00939>
266. Scagliusi, S.M.; Basu, S.; Gouvêa, J.A.D.; Vega, J. (2009) Comparison of two diagnostic methods for evaluation of Sugarcane yellow leaf virus concentration in Brazilian sugarcane cultivars. *Funct. Plant Sci. Biotechnol.* 3, 26–30.
267. Schoeneweiss DF (1975) Predisposition, stress, and plant disease. *Annu. Rev. Phytopathol.* 13: 193–211.
268. Scholthof KBG, Adkins S, Czosnek H, Palukaitis P, Jacquot E, Hohn T, Hohn B, Saunders K, Candresse T, Ahlquist P, Hemenway C, Foster GD. 2011. [\*Top 10 plant viruses in molecular plant pathology\*](#). *Molecular Plant Pathology* 12(9):938-954
269. Schönegger D, Babalola BM, Marais A, Faure C, Candresse T (2022) Diversity of polerovirus-associated RNAs in the virome of wild and cultivated carrots. *Plant Pathol* 71, 1892– 1900. <https://doi.org/10.1111/ppa.13623>

270. Schönegger, D., Marais, A., Babalola, B. M., Faure, C., Lefebvre, M., Svanella-Dumas, L., Brázdová, S., & Candresse, T. (2023). Carrot populations in France and Spain host a complex virome rich in previously uncharacterized viruses. *PloS one*, 18(8), e0290108. <https://doi.org/10.1371/journal.pone.0290108>
271. Scott, S. W., MacFarlane, S. A., McGavin, W. J., & Fargette, D. (2014). Cassava Ivorian bacilliform virus is a member of the genus Anulavirus. *Archives of virology*, 159(10), 2791–2793. <https://doi.org/10.1007/s00705-014-2086-3>
272. Seabloom, E. W., Hosseini, P. R., Power, A. G., & Borer, E. T. (2009). Diversity and composition of viral communities: coinfection of barley and cereal yellow dwarf viruses in California grasslands. *The American naturalist*, 173(3), E79–E98. <https://doi.org/10.1086/596529>
273. Seal, S.E., Jeger, M.J., van Den Bosch, F. 2006b. Begomovirus evolution and disease management. *Adv. Virus Res.* 67, 297–316.
274. Seal, S.E., van Den Bosch, F. Jeger, M. J. 2006a. Factors influencing begomovirus evolution and their increasing global significance: implications for sustainable control. *Crit. Rev. Plant Sci.* 25, 23–46.
275. Sexton J.P, Montiel J, Shay JE, Stephens MR, Slatyer RA. 2017. Evolution of ecological niche breadth. *Annual Review of Ecology, Evolution, and Systematics* 48: 183–206
276. Shates, T., Gebiola, M., Sun, P., Helo, A., Aung, O., Kenney, J., Malmstrom, C., & Mauck, K. (2023). Non-native plant viruses prevalent in remnant natural plant communities harm native perennial hosts. *Phytobiomes Journal*. <https://doi.org/10.1094/pbiomes-05-23-0033-r>.
277. Shaw, C. and Kennedy, D. (2021). Developing an empirical model for spillover and emergence: orsay virus host range in caenorhabditis. <https://doi.org/10.1101/2021.12.10.472097>
278. Shukla D. D., Ward C. W., Brunt A. A. 1994 *The Potyviridae* Wallingford, UK: CAB International

279. Shukla, D. and Ward, C. (1988). Amino acid sequence homology of coat proteins as a basis for identification and classification of the potyvirus group. *Journal of General Virology*, 69(11), 2703-2710. <https://doi.org/10.1099/0022-1317-69-11-2703>
280. Silver SL, Rasochova L, Dinesh-Kumar SP, Miller WA (1994) Replication of barley yellow dwarf virus satellite RNA transcripts in oat protoplasts. *Virology* 198:331–335. <https://doi.org/10.1006/viro.1994.1036>
281. Simon, PW et al. in *Carrot. Handbook of Plant Breeding, Vegetables II* (eds Prohens, J & Nuez, F.) 327–357 (Springer, New York, 2008).
282. Skelton, A., Buxton-Kirk, A., Ward, R., Harju, V., Frew, L., Fowkes, A., Long, M., Negus, A., Forde, S., Adams, I.P., Pufal, H., McGreig, S., Weekes, R., Fox, A. (2019). "First report of Tomato brown rugose fruit virus in tomato in the United Kingdom". *New Disease Reports*. 40: 12. <https://doi.org/10.5197/j.2044-0588.2019.040.012>
283. Smirnova, E., Firth, A. E., Miller, W. A., Scheidecker, D., Brault, V., Reinbold, C., Rakotondrafara, A. M., Chung, B. Y., & Ziegler-Graff, V. (2015). Discovery of a Small Non-AUG-Initiated ORF in Poleroviruses and Luteoviruses That Is Required for Long-Distance Movement. *PLoS Pathogens*, 11(5), e1004868. <https://doi.org/10.1371/journal.ppat.1004868>
284. Sömera, M., Fargette, D., Hébrard, E., Sarmiento, C., & Ictv Report Consortium (2021). ICTV Virus Taxonomy Profile: *Solemoviridae* 2021. *The Journal of general virology*, 102(12), 001707. <https://doi.org/10.1099/jgv.0.001707>
285. Song SI, Miller WA (2004) *Cis* and *trans* requirements for rolling circle replication of a satellite RNA. *J Virol* 78:3072–3082. <https://doi.org/10.1128/jvi.78.6.3072-3082.2004>
286. Song SI, Silver SL, Aulik MA, Rasochova L, Mohan BR, Miller WA (1999) Satellite cereal yellow dwarf virus-RPV (satRPV) RNA requires a double hammerhead for self-cleavage and an alternative structure for replication. *J Mol Biol* 293:781–793. <https://doi.org/10.1006/jmbi.1999.3169>
287. Sparrow A. D. (1999). A heterogeneity of heterogeneities. *Trends in ecology & evolution*, 14(11), 422–423. [https://doi.org/10.1016/s0169-5347\(99\)01735-8](https://doi.org/10.1016/s0169-5347(99)01735-8)

288. Stein A., Gerstner K., Kreft H. (2014). Environmental heterogeneity as a universal driver of species richness across taxa, biomes and spatial scales. *Ecology Letters*, 17(7), 866–880. <https://doi.org/10.1111/ele.12277>
289. Stein, A., & Kreft, H. (2015). Terminology and quantification of environmental heterogeneity in species-richness research. *Biological Reviews*, 90, 815–836. <https://doi.org/10.1111/brv.12135>
290. Stolarczyk, J, Janick, J. Carrot: history and iconography. *Chronica* (2011) 51, 13.
291. Stobbe, A. H., Melcher, U., Palmer, M. W., Roossinck, M. J., & Shen, G. (2012). Co-divergence and host-switching in the evolution of tobamoviruses. *The Journal of general virology*, 93(Pt 2), 408–418. <https://doi.org/10.1099/vir.0.034280-0>
292. Susi, H., Filloux, D., Frilander, M., Roumagnac, P., & Laine, A. (2019). Diverse and variable virus communities in wild plant populations revealed by metagenomic tools. *PeerJ*, 7. <https://doi.org/10.7717/peerj.6140>.
293. Syller J. (2014). Biological and molecular events associated with simultaneous transmission of plant viruses by invertebrate and fungal vectors. *Molecular plant pathology*, 15(4), 417–426. <https://doi.org/10.1111/mpp.12101>
294. Syller, J. (2006) The roles and mechanisms of helper component proteins encoded by potyviruses and caulimoviruses. *Physiol. Mol. Plant Pathol.* 67, 119–130.
295. Syller J. (2012). Facilitative and antagonistic interactions between plant viruses in mixed infections. *Molecular plant pathology*, 13(2), 204–216. <https://doi.org/10.1111/j.1364-3703.2011.00734.x>
296. Syller, J., & Grupa, A. (2016). Antagonistic within-host interactions between plant viruses: molecular basis and impact on viral and host fitness. *Molecular plant pathology*, 17(5), 769–782. <https://doi.org/10.1111/mpp.12322>
297. Taliansky, M., Mayo, M. A., & Barker, H. A. (2003). potato leafroll virus: a classic pathogen shows some new tricks. *Molecular Plant Pathology*, 4(2), 81–89. <https://doi.org/10.1046/j.1364-3703.2003.00153.x>

298. Tamme, R., Hiiesalu, I., Laanisto, L., Szava-Kovats, R. & Pärtel, M. (2010). Environmental heterogeneity, species diversity and co-existence at different spatial scales. *J. Veg. Sci.*, 21, 796–801.
299. Tepfer, M. 2002. Risk assessment of virus-resistant transgenic plants. *Annu. Rev. Phytopathol.* 40:467-491.
300. Tepfer, M.; García-Arenal, F. (2019) Epidemiology and ecology. In *Cucumber Mosaic Virus*; The American Phytopathological Society (APS): St. Paul, MN, USA,; pp. 37–45.
301. Thomas, T., Gilbert, J., & Meyer, F. (2012). Metagenomics - a guide from sampling to data analysis. *Microbial informatics and experimentation*, 2(1), 3. <https://doi.org/10.1186/2042-5783-2-3>
302. Thrall, P., Barrett, L., Dodds, P., & Burdon, J. (2016). Epidemiological and Evolutionary Outcomes in Gene-for-Gene and Matching Allele Models. *Frontiers in Plant Science*, 6. <https://doi.org/10.3389/fpls.2015.01084>.
303. Tollenaere, C., Susi, H., & Laine, A. L. (2016). Evolutionary and Epidemiological Implications of Multiple Infection in Plants. *Trends in plant science*, 21(1), 80–90. <https://doi.org/10.1016/j.tplants.2015.10.014>
304. Tugume AK, Cuellar WJ, Mukasa SB, Valkonen JPT: Molecular genetic analysis of virus isolates from wild and cultivated plants demonstrates that East Africa is a hotspot for the evolution and diversification of Sweet potato feathery mottle virus. *Mol Ecol* 2010, 19:3139-3155.
305. Turner, P. E., Morales, N. M., Alto, B. W., & Remold, S. K. (2010). Role of evolved host breadth in the initial emergence of an RNA virus. *Evolution; international journal of organic evolution*, 64(11), 3273–3286. <https://doi.org/10.1111/j.1558-5646.2010.01051.x>
306. Turner, P. E., & Elena, S. F. (2000). Cost of host radiation in an RNA virus. *Genetics*, 156(4), 1465–1470. <https://doi.org/10.1093/genetics/156.4.1465>
307. Turney S, Gonzalez A, Millien V. 2014 The negative relationship between mammal host diversity and Lyme disease incidence strengthens through time. *Ecology* 95, 3244–3250. [doi:10.1890/14-0980.1](https://doi.org/10.1890/14-0980.1)

308. Tylianakis, J. M., Tscharntke, T., and Lewis, O. T. (2007) 'Habitat Modification Alters the Structure of Tropical Host-parasitoid Food Webs', *Nature*, 445: 202–5.
309. Urququi-Inchima, S., Haenni, A. L., & Bernardi, F. (2001). Potyvirus proteins: a wealth of functions. *Virus research*, 74(1-2), 157–175. [https://doi.org/10.1016/s0168-1702\(01\)00220-9](https://doi.org/10.1016/s0168-1702(01)00220-9)
310. Vacher C, Piou D, Desprez-Loustau ML (2008) Correction: Architecture of an Antagonistic Tree/Fungus Network: The Asymmetric Influence of Past Evolutionary History. *PLOS ONE* 3(3): <https://doi.org/10.1371/annotation/c03a4beb-a3fa-4b78-8724-b3c8ce28c415>
311. Valverde-Asenjo I, Diéguez-Antón A, Martín-Sanz JP, Molina JA, Quintana JR(2020) Soil and vegetation dynamics in a chronosequence of abandoned vineyards. *Agriculture Ecosystems and Environment* 301 (2020):107049. <https://doi.org/10.1016/j.agee.2020.107049>
312. van del Vlugt RAA, Stijger CCMM, Verhoeven JThJ, and Lesemann DE. 2000. First report of *Pepino Mosaic Virus* on tomato. *Plant Dis.* 84:103.
313. van Dijk P, Bos L (1985) Viral dieback of carrot and other Umbelliferae caused by the Anthriscus strain of parsnip yellow fleck virus, and its distinction from carrot motley dwarf. *Neth. J. Plant Pathol.* 91, 169-187. <https://doi.org/10.1007/BF02009678>
314. Velasco, L., Ruiz, L., Galipienso, L., Rubio, L., and Janssen, D. 2020. a historical account of viruses in intensive horticultural crops in the Spanish Mediterranean Arc: New challenges for a sustainable agriculture. *Agronomy* 10:860
315. Vienne, D., Hood, M., & Giraud, T. (2009). Phylogenetic determinants of potential host shifts in fungal pathogens. *Journal of Evolutionary Biology*, 22. <https://doi.org/10.1111/j.1420-9101.2009.01878.x>.
316. Villamor DEV, Ho T, Al Rwahnih M, Martin RR, Tzanetakis IE (2019) High throughput sequencing for plant virus detection and discovery. *Phytopathology* 109 (5): 716–725. <https://doi.org/10.1094/PHYTO07-18-0257-RVW> PMID: 30801236

317. Villamor, D. E. V., Ho, T., Al Rwahnih, M., Martin, R. R., & Tzanetakis, I. E. (2019). High Throughput Sequencing For Plant Virus Detection and Discovery. *Phytopathology*, 109(5), 716–725. <https://doi.org/10.1094/PHYTO-07-18-0257-RVW>
318. Walker, P. J., Siddell, S. G., Lefkowitz, E. J., Mushegian, A. R., Adriaenssens, E. M., Alfenas-Zerbini, P., Davison, A. J., Dempsey, D. M., Dutilh, B. E., García, M. L., Harrach, B., Harrison, R. L., Hendrickson, R. C., Junglen, S., Knowles, N. J., Krupovic, M., Kuhn, J. H., Lambert, A. J., Łobocka, M., Nibert, M. L., ... Zerbini, F. M. (2021). Changes to virus taxonomy and to the International Code of Virus Classification and Nomenclature ratified by the International Committee on Taxonomy of Viruses (2021). *Archives of virology*, 166(9), 2633–2648. <https://doi.org/10.1007/s00705-021-05156-1>
319. Wallis, C. M., Stone, A. L., Sherman, D. J., Damsteegt, V. D., Gildow, F. E., & Schneider, W. L. (2007). Adaptation of plum pox virus to a herbaceous host (*Pisum sativum*) following serial passages. *The Journal of general virology*, 88(Pt 10), 2839–2845. <https://doi.org/10.1099/vir.0.82814-0>
320. Walsh, N.G. & Entwisle, T.J. (1999). *Flora of Victoria*. Volume 4. Inkata Press, Melbourne.
321. Wang, Y.M.; Ostendorf, B.; Gautam, D.; Habili, N.; Pagay, V. (2022) Plant Viral Disease Detection: From Molecular Diagnosis to Optical Sensing Technology—A Multidisciplinary Review. *Remote Sens.* 14, 1542. <https://doi.org/10.3390/rs14071542>
322. Ward E, Foster SJ, Fraaije BA, McCartney HA (2004) Plant pathogen diagnostics: immunological and nucleic acid-based approaches. *Ann Appl Biol* 145 (1): 1–16.
323. Ward, C. (1993). Progress towards a higher taxonomy of viruses. *Research in Virology*, 144, 419–453. [https://doi.org/10.1016/s0923-2516\(06\)80059-2](https://doi.org/10.1016/s0923-2516(06)80059-2)
324. Wells, K., & Clark, N. J. (2019). Host Specificity in Variable Environments. *Trends in parasitology*, 35(6), 452–465. <https://doi.org/10.1016/j.pt.2019.04.001>
325. Welsh ME, Cronin JP, Mitchell CE. 2016 The role of habitat filtering in the leaf economics spectrum and plant susceptibility to pathogen infection. *J. Ecol.* 104, 1768–1777. <https://doi.org/10.1111/1365-2745.12632>

326. Western Australian Herbarium (1998-). FloraBase - The Western Australian Flora. Department of Environment and Conservation. Available at: [Taxon Profile of Daucus carota L. | Florabase \(dbca.wa.gov.au\)](http://dbca.wa.gov.au/taxon_profile_of_Daucus_carota_L)
327. Whitlock, M.C. 1996. The Red Queen Beats the Jack-Of-All-Trades: The Limitations on the Evolution of Phenotypic Plasticity and Niche Breadth. *The American Naturalist*, 148, S65–S77. <https://www.istor.org/stable/2463048>
328. Wolinsky, S. M., Korber, B. T., Neumann, A. U., Daniels, M., Kunstman, K. J., Whetsell, A. J., Furtado, M. R., Cao, Y., Ho, D. D., & Safrit, J. T. (1996). Adaptive evolution of human immunodeficiency virus-type 1 during the natural course of infection. *Science (New York, N.Y.)*, 272(5261), 537–542. <https://doi.org/10.1126/science.272.5261.537>
329. Woolhouse M.E.J., Haydon D.T., Antia R. 2005. Emerging pathogens: the epidemiology and evolution of species jumps. *Trends in Ecology & Evolution* 20: 238– 244. <https://doi.org/10.1016/j.tree.2005.02.009>
330. Woolhouse ME, Dye C. (2001) Population biology of emerging and re-emerging pathogens—preface. *Philos Trans R Soc Lond B Biol Sci.*;356:981–982.
331. Woolhouse, M. E., & Gowtage-Sequeria, S. (2005). Host range and emerging and reemerging pathogens. *Emerging infectious diseases*, 11(12), 1842–1847. <https://doi.org/10.3201/eid1112.050997>
332. Woolhouse, M. E., Taylor, L. H., & Haydon, D. T. (2001). Population biology of multihost pathogens. *Science (New York, N.Y.)*, 292(5519), 1109–1112. <https://doi.org/10.1126/science.1059026>
333. Woolhouse, M. E., Webster, J. P., Domingo, E., Charlesworth, B., and Levin, B. R. 2002. Biological and biomedical implications of the co-evolution of pathogens and their hosts. *Nat. Genet.* 32:569-577. <https://doi.org/10.1038/ng1202-569>
334. Yang, Z., Liu, X., Zhou, M. *et al.* The effect of environmental heterogeneity on species richness depends on community position along the environmental gradient (2015). *Sci Rep* 5, 15723 <https://doi.org/10.1038/srep15723>

335. Yarwood C. E. (1979). Host passage effects with plant viruses. *Advances in virus research*, 25, 169–190. [https://doi.org/10.1016/s0065-3527\(08\)60570-9](https://doi.org/10.1016/s0065-3527(08)60570-9)
336. Yarwood C. E. (1979). Host passage effects with plant viruses. *Advances in virus research*, 25, 169–190. [https://doi.org/10.1016/s0065-3527\(08\)60570-9](https://doi.org/10.1016/s0065-3527(08)60570-9)
337. Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, Madden T (2012). Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics*. 13:134.
338. Yuxin Ma, Armelle Marais, Marie Lefebvre, Chantal Faure, Thierry Candresse. 2020. Metagenomic analysis of virome cross-talk between cultivated *Solanum lycopersicum* and wild *Solanum nigrum*, *Virology* 540: 38-44, <https://doi.org/10.1016/j.virol.2019.11.009>
339. Zhang, D. (2017). A coefficient of determination for generalized linear models. *The American Statistician*, 71(4): 310-316.
340. Zhang, S., Griffiths, J. S., Marchand, G., Bernards, M. A., and Wang, A. 2022. Tomato brown rugose fruit virus: An emerging and rapidly spreading plant RNA virus that threatens tomato production worldwide. *Mol. Plant Pathol.* 23:1262-1277.
341. Zhang, S., Griffiths, J. S., Marchand, G., Bernards, M. A., and Wang, A. 2022. Tomato brown rugose fruit virus: An emerging and rapidly spreading plant RNA virus that threatens tomato production worldwide. *Mol. Plant Pathol.* 23:1262-1277.
342. Zhang, X.-S. , Holt, J. and Colvin, J. (2001) Synergism between plant viruses: a mathematical analysis of the epidemiological implications. *Plant Pathol.* 50, 732–746.
343. Zhao W, Xu Z, Zhang X, Yang M, Kang L, Liu R, Cui F (2018) Genomic variations in the 3'-termini of rice stripe virus in the rotation between vector insect and host plant. *New Phytol* 219:1085–1096. <https://doi.org/10.1111/nph.15246>
344. Zhong, Z., Norvienyeku, J., Chen, M., Bao, J., Lin, L., Chen, L., ... & Wang, Z. (2016). Directional selection from host plants is a major force driving host specificity in magnaporthe species. *Scientific Reports*, 6(1). <https://doi.org/10.1038/srep25591>

**Table A1: Deposited sequences of CtRLV- Hammerhead satellites obtained from Spanish and French carrot populations.**

| GenBank Record | Base pair | FASTA sequence                                                                                                                                                                                                                                                                                                                                                                                      |
|----------------|-----------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| OM962993       | 298       | TGAGAGCGCGTACTCACTGAAGACGTATCCGCGCGGACTAGAAGGCTGGTGCCTCGTCCAACAAATAG<br>ATACAGAAAACCAACCCGAGTAAAGATCTCCAATTGTGGCACCACCAGGTGGCCACCACTCTTTGAAGTG<br>AGGAGACTTGCTTTACGTGTCTGTCATCAATAGATAACATTTGACACTGGTGTTTCGTTCTTTGAACTCA<br>TCAGTTGAGGTACACACCTCAAGACACCAGGGAACAATCGTCCACAATTCACAAAGCGTAAATCGCAGA<br>AGGATGACTCCCATGTGGTTA                                                                          |
| OM962994       | 368       | TGAGAGCGCGTACTCACTGATGACGTATCCGCGCGGACTAGTAGGCTGGTGCCTCGTCCAATAAATAGA<br>TACAGAAAATCCACCTTGTTCCGGAGTCCGTCATTTGTCGCCACCACACTCGTTTTGTGAGTGGGGCAG<br>GACTCTGTGAAATGGGACAGATTGTTCCGGGTGTCGTTAGCGAATTAACGCTGACACTGGTGTTTCGTT<br>CTTTTGAACATCATCAGTCGAGGTACGCACCTCGAGACACCTTTCCTTCTTCAATCAAAATTCTTTTCAAC<br>CCCGTGAGGGCTGGTCCTCAGAGAGACTCTGAGATACAGGAACCCACGTATTGAGTTTCAAGAATCTCA<br>ATCGAGGAAGAAGTGGATTA |
| OM962996       | 368       | TGAGAGCGCGTACTCACTGATGACGTATCCGCGCGGACTAGTAGGCTGGTGCCTCGTCCAATAAATAGA<br>TACAGAAAATCCACCTTGTTCCGGAGTCCGTCATTTGTCGCCACCACACTCGTATTGTGAGTGGGGCA                                                                                                                                                                                                                                                       |

GGACTCTGTGAAATGGGACAGATTGTTCCGGGTGTCGTTAGCGAATTAACGCTGACACTGGTGTTCGT  
TCTTTTGAAGTCATCAGTCGAGGTACGCACCTCGAGACACCTTTCCTTCTTCGAATCAAAATTCTTTTCAA  
CCCCGTGAGGGCTGGTCCTCAGAGAGACTCTGAGATACAGGAACCCACGTATTGAGTTTCAAGAATCTC  
AATCGAGGAAGAAGTGGATTA

**Table A2: Percent identity of nucleotide sequences of amplicons from individual plant sample extracts with the reference genome for each virus.**

| Virus species | HTS Library | Host species                | Indv. Sample ID | Query length | E value  | Per. ident | Acc. Len | Accession  |
|---------------|-------------|-----------------------------|-----------------|--------------|----------|------------|----------|------------|
| CMV           | PV061       | <i>Cucumis melo</i>         | M5V 1           | 200          | 8.00E-41 | 88.16      | 2215     | AB369272.1 |
| CMV           | PV001       | <i>Amaranthus sp</i>        | M1V 2           | 200          | 2.00E-42 | 92.48      | 2216     | Y18137.1   |
| CMV           | PV003       | <i>Cucumis melo</i>         | M1V 3.1         | 91           | 3.00E-16 | 87.8       | 2215     | LC546042.1 |
| TAV           | PV002       | <i>Convolvulus arvensis</i> | M1V 16          | 51           | 2.00E-06 | 93.75      | 2389     | AJ277269.1 |
| TAV           | PV120       | <i>Leontodon sp</i>         | M5V 6           | 69           | 1.00E-21 | 98.44      | 2224     | LC380671.1 |

|      |                  |                                  |          |     |          |       |       |            |
|------|------------------|----------------------------------|----------|-----|----------|-------|-------|------------|
| TAV  | PV112            | <i>Centranthus calcitrapae</i>   | Q2P 138  | 57  | 3.00E-11 | 96.15 | 2386  | AJ277268.1 |
| TAV  | PV128            | <i>Vicia sp</i>                  | L1P 36   | 56  | 0.001    | 88.89 | 2224  | LC380674.1 |
| TAV  | PV150            | <i>Anthriscus caucalis</i>       | L4P 63   | 66  | 2.00E-10 | 97.87 | 2332  | OL472071.1 |
| TAV  | PV144            | <i>Fumaria parviflora</i>        | L3P 4.1  | 54  | 3.00E-05 | 92.5  | 2386  | EU163410.1 |
| TAV  | positive control |                                  | TAV +    | 74  | 1.00E-13 | 90.28 | 2213  | S72468.1   |
| PZSV | PV181            | <i>Diplotaxis eruroides</i>      | L3V 51   | 45  | 2.00E-04 | 94.87 | 2659  | AJ272329.1 |
| PZSV | PV214            | <i>Convolvulus arvensis</i>      | L3F 46   | 45  | 4.3      | 86.49 | 555   | KT923155.1 |
| PZSV | PV201            | <i>Convolvulus arvensis</i>      | L1F 33.1 | 37  | 5.5      | 90.32 | 2659  | AJ272329.1 |
| PPV  | PV046            | <i>Carduus bourgeanus</i>        | Z1V 22   | 97  | 3.00E-36 | 97.89 | 14938 | KY825150.1 |
| PPV  | PV048            | <i>Desconocida 4</i>             | Z2V 13   | 121 | 1.00E-35 | 96.91 | 10712 | EF569215.1 |
| PPV  | PV006            | <i>Artemisia herba alba</i>      | Q1F 13.1 | 123 | 4.00E-36 | 100   | 9784  | MK372956.1 |
| PPV  | PV031            | <i>Convolvulus arvensis</i>      | L2F 32   | 118 | 1.00E-35 | 100   | 1264  | KP198595.1 |
| PPV  | PV187            | <i>Brachypodium phoenicoides</i> | L2V 16   | 107 | 1.00E-35 | 96    | 10712 | EF569215.1 |
| PPV  | positive control | <i>Nicotiana sp</i>              | PPV +    | 110 | 3.00E-37 | 97.92 | 2854  | M21847.1   |

|      |       |                              |         |     |          |       |       |            |
|------|-------|------------------------------|---------|-----|----------|-------|-------|------------|
| TuMV | PV143 | <i>Diplotaxis eruroides</i>  | L3P 24  | 100 | 5.00E-09 | 95.65 | 867   | EF061746.1 |
| TuMV | PV142 | <i>Convolvulus arvensis</i>  | L3P 5   | 84  | 5.00E-08 | 97.56 | 9835  | AF169561.2 |
| TuMV | PV186 | <i>Sonchus oleraceus</i>     | L3V 75  | 55  | 2.00E-07 | 97.56 | 9798  | EU861593.1 |
| TuMV | PV147 | <i>Papaver rhoeas</i>        | L3P 20  | 67  | 1.00E-05 | 95.24 | 9798  | AB252114.1 |
| TuMV | PV141 | <i>Bromus sp</i>             | L3P 32  | 75  | 2.00E-12 | 100   | 9792  | MH469724.1 |
| WMV  | PV224 | <i>Vicia sp</i>              | L3F 45  | 61  | 2.00E-15 | 96.49 | 822   | FJ628395.1 |
| WMV  | PV004 | <i>Cyperus longus</i>        | M1V 10  | 50  | 9.00E-14 | 96.3  | 10025 | MN854641.1 |
| WMV  | PV002 | <i>Convolvulus arvensis</i>  | M1V 11  | 82  | 1.00E-14 | 96.43 | 10102 | KT992093.1 |
| WMV  | PV003 | <i>Cucumis melo</i>          | M1V 3.1 | 75  | 2.00E-17 | 100   | 846   | MK614763.1 |
| WMV  | PV061 | <i>Cucumis melo</i>          | M5V 11  | 81  | 2.00E-11 | 92.98 | 9965  | OM746965.1 |
| BChV | PV498 | <i>Rubia peregrina</i>       | H3P 39  | 113 | 6.00E-09 | 100   | 2004  | EU022511.1 |
| BChV | PV544 | <i>Sisymbrium runcinatum</i> | C1F 49  | 256 | 2.00E-32 | 94.9  | 5776  | AF352024.1 |
| BChV | PV143 | <i>Diplotaxis eruroides</i>  | L3P 12  | 162 | 4.00E-35 | 95.92 | 563   | AF167476.1 |
| BChV | PV149 | <i>Silybum marianum</i>      | L3P 25  | 102 | 2.00E-12 | 94.64 | 2004  | EU022511.1 |

|      |                |                           |        |     |          |      |      |            |
|------|----------------|---------------------------|--------|-----|----------|------|------|------------|
| YoMV | PV085          | <i>Avena sterilis</i>     | E4P18  | 114 | 5.00E-41 | 95.5 | 6302 | MW248354.1 |
| YoMV | PV085          | <i>Avena sterilis</i>     | E4P105 | 115 | 2.00E-46 | 99.1 | 6303 | KF137561.1 |
| YoMV | PV084          | <i>Teucrium capitatum</i> | Q1P68  | 114 | 1.00E-47 | 100  | 6303 | U30944.1   |
| YoMV | purified virus | <i>Brassica petsai</i>    | YoMV + | 115 | 2.00E-46 | 99.1 | 6303 | U30944.1   |

---

**Table A3: Comparison between number of reads detected by HTS2C, HTS6C for eight OTUs.** Also, the table indicates the status of RT-PCR detection (positive /negative) for each OTU with details of the corresponding library, plant family, habitat, site.

| Virus | Library  | Plant Taxa                  | Plant Family   | Habitat | Site | RT-PCR Detection | HTS2C (No. of reads) | HTS6C (No. of reads) |
|-------|----------|-----------------------------|----------------|---------|------|------------------|----------------------|----------------------|
| CMV   | PV495    | <i>Anacyclus clavatus</i>   | Asteraceae     | crop    | H2P  | positive         | 2                    | 2                    |
| CMV   | PV002    | <i>Convolvulus arvensis</i> | Convolvulaceae | crop    | M1V  | positive         | 391,348              | 264,417              |
| CMV   | PV011    | <i>Portulaca oleraceae</i>  | Portulacaceae  | crop    | M1V  | positive         | 288,047              | 191,452              |
| CMV   | PV003    | <i>Cucumis melo</i>         | Cucurbitaceae  | crop    | M1V  | positive         | 187,496              | 107,585              |
| CMV   | PV001    | <i>Amaranthus sp</i>        | Amaranthaceae  | crop    | M1V  | positive         | 626                  | 228                  |
| CMV   | PV004bgi | <i>Cyperus longus</i>       | Cyperaceae     | crop    | M1V  | positive         | 2                    | 2                    |
| CMV   | PV061    | <i>Cucumis melo</i>         | Cucurbitaceae  | crop    | M5V  | positive         | 2,416,188            | 1,683,996            |
| CMV   | PV063    | <i>Cucumis melo</i>         | Cucurbitaceae  | crop    | M8V  | positive         | 30                   | 26                   |
| CMV   | PV046    | <i>Carduus bourgeanus</i>   | Asteraceae     | crop    | Z1V  | positive         | 5                    | 2                    |
| CMV   | PV047    | <i>Zea mays</i>             | Poaceae        | crop    | Z2V  | positive         | 7                    | 7                    |
| CMV   | PV199    | <i>Carduus bourgeanus</i>   | Asteraceae     | edge    | L1F  | positive         | 10                   | 10                   |
| CMV   | PV032    | <i>Brachypodium retusum</i> | Poaceae        | edge    | L2F  | positive         | 2                    | 2                    |
| CMV   | PV217    | <i>Datura stramonium</i>    | Solanaceae     | edge    | L3F  | positive         | 6                    | 6                    |
| CMV   | PV219    | <i>Diplotaxis erucoides</i> | Brassicaceae   | edge    | L3F  | positive         | 34                   | 30                   |
| CMV   | PV216    | <i>Cynodon dactylon</i>     | Poaceae        | edge    | L3F  | positive         | 18                   | 16                   |
| CMV   | PV218    | <i>Daucus sp</i>            | Apiaceae       | edge    | L3F  | positive         | 7                    | 6                    |

|     |       |                               |              |      |     |          |       |       |
|-----|-------|-------------------------------|--------------|------|-----|----------|-------|-------|
| CMV | PV224 | <i>Vicia sp</i>               | Fabaceae     | edge | L3F | positive | 19    | 18    |
| CMV | PV232 | <i>Lolium perenne</i>         | Poaceae      | edge | L4F | positive | 4     | 4     |
| CMV | PV120 | <i>Leontodon sp</i>           | Asteraceae   | edge | L1P | positive | 3,595 | 3,003 |
| CMV | PV115 | <i>Bromus sp</i>              | Poaceae      | edge | L1P | positive | 1,135 | 982   |
| CMV | PV123 | <i>Rumex plucher</i>          | Polygonaceae | edge | L1P | positive | 18    | 10    |
| CMV | PV128 | <i>Vicia sp</i>               | Fabaceae     | edge | L1P | positive | 31    | 28    |
| CMV | PV136 | <i>Latua serriola</i>         | Asteraceae   | edge | L2P | positive | 1,833 | 1,611 |
| CMV | PV129 | <i>Bromus sp</i>              | Poaceae      | edge | L2P | positive | 1,286 | 1,127 |
| CMV | PV130 | <i>Carduus bourgeanus</i>     | Asteraceae   | edge | L2P | positive | 917   | 791   |
| CMV | PV137 | <i>Lolium pernne</i>          | Poaceae      | edge | L2P | positive | 859   | 734   |
| CMV | PV138 | <i>Potentilla sp</i>          | Rosaceae     | edge | L2P | positive | 83    | 68    |
| CMV | PV139 | <i>Rubia peregrina</i>        | Rubiaceae    | edge | L2P | positive | 20    | 14    |
| CMV | PV194 | <i>Papaver rhoeas</i>         | Papaveraceae | edge | L2P | positive | 15    | 14    |
| CMV | PV148 | <i>Scandix pecten_veneris</i> | Apiaceae     | edge | L3P | positive | 1,696 | 1,406 |
| CMV | PV144 | <i>Fumaria parviflora</i>     | Papaveraceae | edge | L3P | positive | 1,003 | 845   |
| CMV | PV158 | <i>Rubia peregrina</i>        | Rubiaceae    | edge | L4P | positive | 375   | 1,234 |
| CMV | PV150 | <i>Anthriscus caucalis</i>    | Apiaceae     | edge | L4P | positive | 1,465 | 314   |
| CMV | PV030 | <i>Milium vernale</i>         | Poaceae      | edge | L1V | positive | 18    | 16    |
| CMV | PV096 | <i>Milium vernale</i>         | Poaceae      | edge | L2V | positive | 28    | 24    |
| CMV | PV169 | <i>Galium verum</i>           | Rubiaceae    | edge | L2V | negative | 16    | 14    |
| CMV | PV173 | <i>Rubia peregrina</i>        | Rubiaceae    | edge | L2V | positive | 15    | 14    |

|     |             |                                  |                |         |     |          |     |     |
|-----|-------------|----------------------------------|----------------|---------|-----|----------|-----|-----|
| CMV | PV187       | <i>Brachypodium phoenicoides</i> | Poaceae        | edge    | L2V | positive | 4   | 4   |
| CMV | PV097       | <i>Convolvulus arvensis</i>      | Convolvulaceae | edge    | L2V | positive | 3   | 2   |
| CMV | PV188 (abs) | <i>Convolvulus arvensis</i>      | Convolvulaceae | edge    | L3V | positive | 1   | -   |
| CMV | PV009       | <i>Quercus coccifera</i>         | Fagaceae       | oakwood | Q1F | positive | 730 | 444 |
| CMV | PV037       | <i>Asparagus acutifolius</i>     | Asparagaceae   | oakwood | Q2F | positive | 13  | 10  |
| CMV | PV036       | <i>Quercus ilex</i>              | Fagaceae       | oakwood | Q2F | positive | 7   | 6   |
| CMV | PV041       | <i>Avenula bromoides</i>         | Poaceae        | oakwood | Q3F | positive | 10  | 8   |
| CMV | PV042       | <i>Quercus coccifera</i>         | Fagaceae       | oakwood | Q3F | positive | 27  | 20  |
| CMV | PV084       | <i>Teucrium capitatum</i>        | Lamiaceae      | oakwood | Q1P | positive | 5   | 4   |
| CMV | PV089       | <i>Marrubium vulgare</i>         | Lamiaceae      | oakwood | Q1P | positive | 4   | 4   |
| CMV | PV423       | <i>Bromus sp</i>                 | Poaceae        | oakwood | Q1P | positive | 16  | 14  |
| CMV | PV089       | <i>Marrubium vulgare</i>         | Lamiaceae      | oakwood | Q1P | positive | 4   | 4   |
| CMV | PV426       | <i>Erodium cicutarium</i>        | Geraniaceae    | oakwood | Q1P | positive | 2   | 2   |
| CMV | PV559       | <i>Stipa parviflora</i>          | Poaceae        | oakwood | Q1P | positive | 27  | 22  |
| CMV | PV095       | <i>Diplotaxis virgata</i>        | Brassicaceae   | oakwood | Q1P | positive | 142 | 92  |
| CMV | PV441       | <i>Leontodon sp</i>              | Asteraceae     | oakwood | Q2P | positive | 182 | 162 |
| CMV | PV108       | <i>Quercus ilex</i>              | Fagaceae       | oakwood | Q2P | positive | 53  | 50  |
| CMV | PV112       | <i>Centhranthus calcitrapae</i>  | Caprifoliaceae | oakwood | Q2P | positive | 757 | 668 |
| CMV | PV110       | <i>Teucrium botrys</i>           | Lamiaceae      | oakwood | Q2P | positive | 381 | 344 |
| CMV | PV114       | <i>Staelhelina dubia</i>         | Asteraceae     | oakwood | Q2P | positive | 388 | 345 |
| CMV | PV437       | <i>Geranium sp</i>               | Geraniaceae    | oakwood | Q2P | positive | 337 | 284 |

|     |            |                              |                |           |     |          |     |     |
|-----|------------|------------------------------|----------------|-----------|-----|----------|-----|-----|
| CMV | PV101      | <i>Asparagus acutifolius</i> | Asparagaceae   | oakwood   | Q2P | positive | 207 | 186 |
| CMV | PV105      | <i>Galium verum</i>          | Rubiaceae      | oakwood   | Q2P | positive | 123 | 116 |
| CMV | PV560      | <i>Teucrium chamaedrys</i>   | Lamiaceae      | oakwood   | Q2P | positive | 67  | 38  |
| CMV | PV054      | <i>Thymus vulgaris</i>       | Lamiaceae      | oakwood   | Q4P | positive | 2   | 2   |
| CMV | PV051      | <i>Lotus corniculatus</i>    | Fabaceae       | oakwood   | Q4P | positive | 2   | 2   |
| CMV | PV023(abs) | <i>Leontodon sp</i>          | Asteraceae     | Wasteland | E4F | positive | 1   | -   |
| CMV | PV065      | <i>Diplotaxis eruroides</i>  | Brassicaceae   | Wasteland | E1F | positive | 2   | 2   |
| CMV | PV070      | <i>Descurania sophia</i>     | Brassicaceae   | Wasteland | E1F | positive | 432 | 378 |
| CMV | PV071      | <i>Salvia verbanaca</i>      | Lamiaceae      | Wasteland | E1F | positive | 5   | 4   |
| CMV | PV075      | <i>Taraxacum officinale</i>  | Asteraceae     | Wasteland | E2F | positive | 2   | 2   |
| CMV | PV078      | <i>Bromus sp</i>             | Poaceae        | Wasteland | E4F | positive | 13  | 10  |
| CMV | PV079      | <i>Convolvulus arvensis</i>  | Convolvulaceae | Wasteland | E4F | positive | 13  | 12  |
| CMV | PV087      | <i>Echium vulgare</i>        | Boraginaceae   | Wasteland | E4P | positive | 16  | 14  |
| CMV | PV092      | <i>Anchusa undulata</i>      | Boraginaceae   | Wasteland | E1F | positive | 2   | 2   |
| CMV | PV098      | <i>Picris echoides</i>       | Asteraceae     | Wasteland | E2F | positive | 5   | 4   |
| CMV | PV257      | <i>Desconocida 5 (abs)</i>   | Brassicaceae   | Wasteland | E3F | positive | 4   | -   |
| CMV | PV262      | <i>Galium verum (abs)</i>    | Rubiaceae      | Wasteland | E3F | negative | 1   | -   |
| CMV | PV356      | <i>Hordeum maritense</i>     | Poaceae        | Wasteland | E3P | positive | 36  | 30  |
| CMV | PV359      | <i>Medicago sp (abs)</i>     | Fabaceae       | Wasteland | E3P | negative | 1   | -   |
| CMV | PV386      | <i>Papaver rhoeas</i>        | Papaveraceae   | Wasteland | E4P | positive | 2   | 2   |

|      |       |                               |                |      |     |          |        |       |
|------|-------|-------------------------------|----------------|------|-----|----------|--------|-------|
| BChV | PV116 | <i>Daucus sp</i>              | Apiaceae       | Edge | L1P | positive | 2      | 2     |
| BChV | PV122 | <i>Picris echioides</i>       | Asteraceae     | Edge | L1P | negative | 3      | 2     |
| BChV | PV128 | <i>Vicia sp</i>               | Fabaceae       | Edge | L1P | negative | 3      | 2     |
| BChV | PV133 | <i>Diplotaxis sp</i>          | Brassicaceae   | Edge | L2P | negative | 2      | 2     |
| BChV | PV143 | <i>Diplotaxis eruroides</i>   | Brassicaceae   | Edge | L3P | positive | 1,052  | 515   |
| BChV | PV146 | <i>Malva sylvestris</i>       | Malvaceae      | Edge | L3P | positive | 3      | 2     |
| BChV | PV148 | <i>Scandix pecten veneris</i> | Apiaceae       | Edge | L3P | negative | 4      | 2     |
| BChV | PV149 | <i>Silybum marianum</i>       | Asteraceae     | Edge | L3P | positive | 3      | 2     |
| BChV | PV156 | <i>Hirschfeldia incana</i>    | Brassicaceae   | Edge | L4P | positive | 12,110 | 5,342 |
| BChV | PV158 | <i>Rubia peregrina</i>        | Rubiaceae      | Edge | L4P | positive | 3      | 2     |
| BChV | PV200 | <i>Chenopodium album</i>      | Amaranthaceae  | Edge | L1F | negative | 2      | 2     |
| BChV | PV201 | <i>Convolvulus arvensis</i>   | Convolvulaceae | Edge | L1F | negative | 2      | 2     |
| BChV | PV211 | <i>Solanum nigrum</i>         | Solanaceae     | Edge | L2F | positive | 2      | 2     |
| BChV | PV212 | <i>Amaranthus sp</i>          | Amaranthaceae  | Edge | L3F | positive | 8,806  | 4,030 |
| BChV | PV214 | <i>Convolvulus arvensis</i>   | Convolvulaceae | Edge | L3F | negative | 5      | 4     |
| BChV | PV215 | <i>Conyza bonariensis</i>     | Asteraceae     | Edge | L3F | positive | 25     | 16    |
| BChV | PV217 | <i>Datura stramonium</i>      | Solanaceae     | Edge | L3F | negative | 2      | 2     |
| BChV | PV218 | <i>Daucus sp</i>              | Apiaceae       | Edge | L3F | negative | 2      | 2     |
| BChV | PV223 | <i>Silybum marianum</i>       | Asteraceae     | Edge | L3F | negative | 2      | 2     |
| BChV | PV230 | <i>Echium vulgare</i>         | Boraginaceae   | Edge | L4F | negative | 2      | 2     |
| BChV | PV234 | <i>Rumex pulcher</i>          | Polygonaceae   | Edge | L4F | positive | 2      | 2     |

|      |            |                                   |                |           |     |              |        |       |
|------|------------|-----------------------------------|----------------|-----------|-----|--------------|--------|-------|
| BChV | PV498      | <i>Rubia peregrina</i>            | Rubiaceae      | Crop      | H3P | positive     | 14,494 | 5,689 |
| BChV | PV505(p)   | <i>Hordeum vulgare</i>            | Poaceae        | Crop      | H3P | Not analysed | -      | 2     |
| BChV | PV538 (p)  | <i>Brassica oleracea</i>          | Brassicaceae   | Crop      | C1F | Not analysed | -      | 1,096 |
| BChV | PV544      | <i>Sisymbrium runcinatum</i>      | Brassicaceae   | Crop      | C1F | positive     | 7      | 2     |
| BChV | PV556 (p)  | <i>Brassica oleracea</i>          | Brassicaceae   | Crop      | C2F | Not analysed | -      | 1,550 |
| PPV  | PV046      | <i>Carduus bourgeanus</i>         | Asteraceae     | Crop      | Z1V | positive     | 2      | 2     |
| PPV  | PV048(abs) | <i>Desconocida 4</i>              | NA             | Crop      | Z2V | positive     | 1      |       |
| PPV  | PV028      | <i>Rubia peregrina</i>            | Rubiaceae      | Edge      | L1V | positive     | 2      | 2     |
| PPV  | PV031      | <i>Convolvulus arvensis</i>       | Convolvulaceae | Edge      | L2F | positive     | 14     | 4     |
| PPV  | PV032      | <i>Brachypodium retusum</i>       | Poaceae        | Edge      | L2F | positive     | 37     | 22    |
| PPV  | PV187      | <i>Brachypodium phoenicoides</i>  | Poaceae        | Edge      | L2V | positive     | 32     | 16    |
| PPV  | PV204      | <i>Diplotaxis eruroides</i>       | Brassicaceae   | Edge      | L1F | positive     | 2      | 2     |
| PPV  | PV029(abs) | <i>Picris echioides</i>           | Asteraceae     | Edge      | L1V | positive     | 1      | -     |
| PPV  | PV034(abs) | <i>Brachypodium retusum</i>       | Poaceae        | Edge      | L4F | positive     | 5      | -     |
| PPV  | PV006bgi   | <i>Artemisia herba alba</i>       | Asteraceae     | Oak       | Q1F | positive     | 2      | 2     |
| PPV  | PV100      | <i>Aphyllanthes monspeliensis</i> | Asparagaceae   | Oak       | Q4P | positive     | 2      | 2     |
| PPV  | PV306      | <i>Asteriscus aquaticus</i>       | Asteraceae     | Wasteland | E2P | positive     | 2      | 2     |
| PZSV | PV557(abs) | <i>Erodium_cicutarium</i>         | Geraniaceae    | Crop      | C2F | positive     | 2      | -     |
| PZSV | PV046      | <i>Carduus_bourgeanus</i>         | Asteraceae     | Crop      | Z1V | positive     | 5      | 4     |
| PZSV | PV199      | <i>Carduus_bourgeanus</i>         | Asteraceae     | Edge      | L1F | positive     | 190    | 122   |

|      |            |                             |                |      |     |          |     |     |
|------|------------|-----------------------------|----------------|------|-----|----------|-----|-----|
| PZSV | PV208      | <i>Rubia_peregrina</i>      | Rubiaceae      | Edge | L1F | positive | 161 | 104 |
| PZSV | PV207      | <i>Picris_echioides</i>     | Asteraceae     | Edge | L1F | positive | 136 | 92  |
| PZSV | PV205      | <i>Geranium_sp</i>          | Geraniaceae    | Edge | L1F | positive | 124 | 76  |
| PZSV | PV201      | <i>Convolvulus_arvensis</i> | Convolvulaceae | Edge | L1F | positive | 123 | 78  |
| PZSV | PV204      | <i>Diplotaxis_eruroides</i> | Brassicaceae   | Edge | L1F | positive | 103 | 60  |
| PZSV | PV128      | <i>Vicia_sp</i>             | Fabaceae       | Edge | L1P | positive | 29  | 20  |
| PZSV | PV120      | <i>Leontodon_sp</i>         | Asteraceae     | Edge | L1P | negative | 24  | 16  |
| PZSV | PV116      | <i>Daucus_sp</i>            | Apiaceae       | Edge | L1P | positive | 20  | 10  |
| PZSV | PV121      | <i>Lepidium_draba</i>       | Brassicaceae   | Edge | L1P | positive | 17  | 10  |
| PZSV | PV117      | <i>Diplotaxis_eruroides</i> | Brassicaceae   | Edge | L1P | positive | 15  | 6   |
| PZSV | PV126      | <i>Silybum_marianum</i>     | Asteraceae     | Edge | L1P | positive | 12  | 4   |
| PZSV | PV115      | <i>Bromus_sp</i>            | Poaceae        | Edge | L1P | positive | 9   | 6   |
| PZSV | PV123(abs) | <i>Rumex_pulcher</i>        | Polygonaceae   | Edge | L1P | positive | 5   | -   |
| PZSV | PV029      | <i>Picris_echioides</i>     | Asteraceae     | Edge | L1V | positive | 6   | 2   |
| PZSV | PV030      | <i>Milium_vernale</i>       | Poaceae        | Edge | L1V | positive | 6   | 2   |
| PZSV | PV163      | <i>Lolium_perenne</i>       | Poaceae        | Edge | L1V | positive | 5   | 4   |
| PZSV | PV028(abs) | <i>Rubia_peregrina</i>      | Rubiaceae      | Edge | L1V | positive | 1   | -   |
| PZSV | PV211      | <i>Solanum_nigrum</i>       | Solanaceae     | Edge | L2F | positive | 188 | 108 |
| PZSV | PV139      | <i>Rubia_peregrina</i>      | Rubiaceae      | Edge | L2P | positive | 30  | 14  |
| PZSV | PV138      | <i>Potentilla_sp</i>        | Rosaceae       | Edge | L2P | positive | 18  | 14  |
| PZSV | PV129      | <i>Bromus_sp</i>            | Poaceae        | Edge | L2P | positive | 17  | 6   |

|      |            |                                  |                |      |     |          |           |         |
|------|------------|----------------------------------|----------------|------|-----|----------|-----------|---------|
| PZSV | PV136      | <i>Lactuca_serriola</i>          | Asteraceae     | Edge | L2P | positive | 17        | 8       |
| PZSV | PV137      | <i>Lolium_perenne</i>            | Poaceae        | Edge | L2P | positive | 17        | 8       |
| PZSV | PV130      | <i>Carduus_bourgeanus</i>        | Asteraceae     | Edge | L2P | negative | 15        | 4       |
| PZSV | PV194      | <i>Papaver_rhoeas</i>            | Papaveraceae   | Edge | L2P | positive | 11        | 10      |
| PZSV | PV168      | <i>Conyza_canadensis</i>         | Asteraceae     | Edge | L2V | positive | 9         | 4       |
| PZSV | PV169      | <i>Galium_verum</i>              | Rubiaceae      | Edge | L2V | negative | 5         | 4       |
| PZSV | PV173(abs) | <i>Rubia_peregrina</i>           | Rubiaceae      | Edge | L2V | positive | 5         | -       |
| PZSV | PV187(abs) | <i>Brachypodium_phoenicoides</i> | Poaceae        | Edge | L2V | positive | 4         | -       |
| PZSV | PV219      | <i>Diplotaxis_erucoides</i>      | Brassicaceae   | Edge | L3F | positive | 1,438,658 | 955,669 |
| PZSV | PV224      | <i>Vicia_sp</i>                  | Fabaceae       | Edge | L3F | positive | 1,023     | 672     |
| PZSV | PV217      | <i>Datura_stramonium</i>         | Solanaceae     | Edge | L3F | negative | 947       | 677     |
| PZSV | PV216      | <i>Cynodon_dactylon</i>          | Poaceae        | Edge | L3F | positive | 417       | 266     |
| PZSV | PV214      | <i>Convolvulus_arvensis</i>      | Convolvulaceae | Edge | L3F | positive | 211       | 124     |
| PZSV | PV218      | <i>Daucus_sp</i>                 | Apiaceae       | Edge | L3F | positive | 142       | 84      |
| PZSV | PV149      | <i>Silybum_marianum</i>          | Asteraceae     | Edge | L3P | positive | 71        | 44      |
| PZSV | PV144      | <i>Fumaria_parviflora</i>        | Papaveraceae   | Edge | L3P | positive | 54        | 26      |
| PZSV | PV148      | <i>Scandix_pecten_veneris</i>    | Apiaceae       | Edge | L3P | positive | 21        | 14      |
| PZSV | PV181      | <i>Diplotaxis_erucoides</i>      | Brassicaceae   | Edge | L3V | positive | 39,249    | 25,760  |
| PZSV | PV081      | <i>Lithospermum_arvense</i>      | Boraginaceae   | Edge | L3V | positive | 60        | 39      |
| PZSV | PV188      | <i>Convolvulus_arvensis</i>      | Convolvulaceae | Edge | L3V | positive | 9         | 8       |
| PZSV | PV225      | <i>Cirsium_arvense</i>           | Asteraceae     | Edge | L4F | negative | 142       | 78      |

|      |             |                                  |              |           |     |          |     |    |
|------|-------------|----------------------------------|--------------|-----------|-----|----------|-----|----|
| PZSV | PV237       | <i>Erodium_cicutarium</i>        | Geraniaceae  | Edge      | L4F | positive | 141 | 98 |
| PZSV | PV235       | <i>Silybum_marianum</i>          | Asteraceae   | Edge      | L4F | positive | 133 | 86 |
| PZSV | PV231       | <i>Lepidium_draba</i>            | Brassicaceae | Edge      | L4F | positive | 132 | 85 |
| PZSV | PV234       | <i>Rumex_pulcher</i>             | Polygonaceae | Edge      | L4F | positive | 131 | 83 |
| PZSV | PV236       | <i>Trifolium_campestre</i>       | Fabaceae     | Edge      | L4F | negative | 96  | 54 |
| PZSV | PV034       | <i>Brachypodium_retusum</i>      | Poaceae      | Edge      | L4F | positive | 17  | 8  |
| PZSV | PV158       | <i>Rubia_peregrina</i>           | Rubiaceae    | Edge      | L4P | negative | 26  | 12 |
| PZSV | PV150       | <i>Anthriscus_caucalis</i>       | Apiaceae     | Edge      | L4P | positive | 24  | 10 |
| PZSV | PV156       | <i>Hirschfeldia_incana</i>       | Brassicaceae | Edge      | L4P | positive | 16  | 13 |
| PZSV | PV153       | <i>Carduus_bourgeanus</i>        | Asteraceae   | Edge      | L4P | negative | 8   | 6  |
| PZSV | PV152       | <i>Brachypodium_phoenicoides</i> | Poaceae      | Edge      | L4P | positive | 7   | 6  |
| PZSV | PV559       | <i>Stipa_parviflora</i>          | Poaceae      | Oakwood   | Q1P | negative | 6   | 2  |
| PZSV | PV421(abs)  | <i>Asteriscus_aquaticus</i>      | Asteraceae   | Oakwood   | Q1P | negative | 3   | -  |
| PZSV | PV430(abs)  | <i>Papaver_rhoeas</i>            | Papaveraceae | Oakwood   | Q1P | negative | 2   | -  |
| PZSV | PV426(abs)  | <i>Erodium_cicutarium</i>        | Geraniaceae  | Oakwood   | Q1P | positive | 1   | -  |
| PZSV | PV037       | <i>Asparagus_acutifolius</i>     | Asparagaceae | Oakwood   | Q2F | positive | 2   | 2  |
| PZSV | PV464       | <i>Rubia_peregrina</i>           | Rubiaceae    | Oakwood   | Q3P | negative | 4   | 4  |
| PZSV | PV277       | <i>Anacyclus_clavatus</i>        | Asteraceae   | Wasteland | E1P | positive | 3   | 2  |
| PZSV | PV298 (abs) | <i>Teucrium_capitatum</i>        | Lamiaceae    | Wasteland | E1P | positive | 1   | -  |
| PZSV | PV306       | <i>Asteriscus_aquaticus</i>      | Asteraceae   | Wasteland | E2P | positive | 44  | 32 |
| PZSV | PV257       | <i>Desconocida_5</i>             | Brassicaceae | Wasteland | E3F | positive | 4   | 2  |

|      |            |                             |                |           |     |          |      |     |
|------|------------|-----------------------------|----------------|-----------|-----|----------|------|-----|
| PZSV | PV262(abs) | <i>Galium_verum</i>         | Rubiaceae      | Wasteland | E3F | positive | 2    | -   |
| PZSV | PV356(abs) | <i>Hordeum_matritense</i>   | Poaceae        | Wasteland | E3P | positive | 1    | -   |
| PZSV | PV359(abs) | <i>Medicago_sp</i>          | Fabaceae       | Wasteland | E3P | positive | 1    | -   |
| PZSV | PV390      | <i>Thapsia_villosa</i>      | Apiaceae       | Wasteland | E4P | positive | 3    | 2   |
| TAV  | PV002      | <i>Convolvulus_arvensis</i> | Convolvulaceae | Crop      | M1V | positive | 191  | 1   |
| TAV  | PV011(abs) | <i>Portulaca_oleraceae</i>  | Portulacaceae  | Crop      | M1V | positive | 151  | -   |
| TAV  | PV003(abs) | <i>Cucumis_melo</i>         | Cucurbitaceae  | Crop      | M1V | positive | 50   | -   |
| TAV  | PV001(abs) | <i>Amaranthus_sp</i>        | Amaranthaceae  | Crop      | M1V | positive | 3    | -   |
| TAV  | PV061      | <i>Cucumis_melo</i>         | Cucurbitaceae  | Crop      | M5V | positive | 604  | 2   |
| TAV  | PV046(abs) | <i>Carduus_bourgeanus</i>   | Asteraceae     | Crop      | Z1V | negative | 1    | -   |
| TAV  | PV120      | <i>Leontodon_sp</i>         | Asteraceae     | Edge      | L1P | positive | 1051 | 680 |
| TAV  | PV115      | <i>Bromus_sp</i>            | Poaceae        | Edge      | L1P | positive | 378  | 246 |
| TAV  | PV119      | <i>Lactuca_serriola</i>     | Asteraceae     | Edge      | L1P | positive | 166  | 96  |
| TAV  | PV124      | <i>Malva_sylvestris</i>     | Malvaceae      | Edge      | L1P | positive | 120  | 74  |
| TAV  | PV128      | <i>Vicia_sp</i>             | Fabaceae       | Edge      | L1P | positive | 2    | 2   |
| TAV  | PV163(abs) | <i>Lolium_perenne</i>       | Poaceae        | Edge      | L1V | positive | 4    | -   |
| TAV  | PV162(abs) | <i>Conyza_canadensis</i>    | Asteraceae     | Edge      | L1V | positive | 2    | -   |
| TAV  | PV210      | <i>Cynodon_dactylon</i>     | Poaceae        | Edge      | L2F | negative | 9    | 6   |
| TAV  | PV136      | <i>Lactuca_serriola</i>     | Asteraceae     | Edge      | L2P | positive | 472  | 303 |
| TAV  | PV129      | <i>Bromus_sp</i>            | Poaceae        | Edge      | L2P | positive | 421  | 258 |

|     |            |                               |                |      |     |          |     |     |
|-----|------------|-------------------------------|----------------|------|-----|----------|-----|-----|
| TAV | PV137      | <i>Lolium_perenne</i>         | Poaceae        | Edge | L2P | positive | 240 | 128 |
| TAV | PV131      | <i>Chenopodium_album</i>      | Amaranthaceae  | Edge | L2P | positive | 170 | 92  |
| TAV | PV130      | <i>Carduus_bourgeanus</i>     | Asteraceae     | Edge | L2P | positive | 2   | 2   |
| TAV | PV133      | <i>Diplotaxis_sp</i>          | Brassicaceae   | Edge | L2P | positive | 2   | 2   |
| TAV | PV135      | <i>Fumaria_parviflora</i>     | Papaveraceae   | Edge | L2P | positive | 2   | 2   |
| TAV | PV138      | <i>Potentilla_sp</i>          | Rosaceae       | Edge | L2P | negative | 2   | 2   |
| TAV | PV171      | <i>Lolium_perenne</i>         | Poaceae        | Edge | L2V | positive | 55  | 30  |
| TAV | PV172      | <i>Picris_echioides</i>       | Asteraceae     | Edge | L2V | positive | 13  | 6   |
| TAV | PV170      | <i>Lactuca_serriola</i>       | Asteraceae     | Edge | L2V | positive | 12  | 6   |
| TAV | PV212      | <i>Amaranthus_sp</i>          | Amaranthaceae  | Edge | L3F | positive | 2   | 2   |
| TAV | PV213      | <i>Carduus_bourgeanus</i>     | Asteraceae     | Edge | L3F | positive | 2   | 2   |
| TAV | PV215      | <i>Conyza_bonariensis</i>     | Asteraceae     | Edge | L3F | negative | 2   | 2   |
| TAV | PV216      | <i>Cynodon_dactylon</i>       | Poaceae        | Edge | L3F | positive | 2   | 2   |
| TAV | PV148      | <i>Scandix_pecten_veneris</i> | Apiaceae       | Edge | L3P | positive | 459 | 260 |
| TAV | PV144      | <i>Fumaria_parviflora</i>     | Papaveraceae   | Edge | L3P | positive | 338 | 212 |
| TAV | PV145      | <i>Lepidium_draba</i>         | Brassicaceae   | Edge | L3P | positive | 203 | 122 |
| TAV | PV141      | <i>Bromus_sp</i>              | Poaceae        | Edge | L3P | positive | 182 | 116 |
| TAV | PV146      | <i>Malva_sylvestris</i>       | Malvaceae      | Edge | L3P | positive | 169 | 86  |
| TAV | PV147      | <i>Papaver_rhoeas</i>         | Papaveraceae   | Edge | L3P | positive | 4   | 4   |
| TAV | PV142(abs) | <i>Convolvulus_arvensis</i>   | Convolvulaceae | Edge | L3P | positive | 2   | -   |
| TAV | PV143      | <i>Diplotaxis_erucoides</i>   | Brassicaceae   | Edge | L3P | positive | 2   | 2   |

|     |            |                                  |                  |      |     |          |      |     |
|-----|------------|----------------------------------|------------------|------|-----|----------|------|-----|
| TAV | PV176      | <i>Anacyclus_clavatus</i>        | Asteraceae       | Edge | L3V | positive | 46   | 22  |
| TAV | PV182      | <i>Lactuca_serriola</i>          | Asteraceae       | Edge | L3V | negative | 29   | 14  |
| TAV | PV232      | <i>Lolium_perenne</i>            | Poaceae          | Edge | L4F | positive | 1192 | 692 |
| TAV | PV233      | <i>Phalaris_minor</i>            | Poaceae          | Edge | L4F | negative | 17   | 10  |
| TAV | PV231(abs) | <i>Lepidium_draba</i>            | Brassicaceae     | Edge | L4F | positive | 2    | -   |
| TAV | PV151      | <i>Aristolochia_pistolochia</i>  | Aristolochiaceae | Edge | L4P | positive | 673  | 452 |
| TAV | PV158      | <i>Rubia_peregrina</i>           | Rubiaceae        | Edge | L4P | positive | 574  | 328 |
| TAV | PV156      | <i>Hirschfeldia_incana</i>       | Brassicaceae     | Edge | L4P | positive | 238  | 122 |
| TAV | PV152      | <i>Brachypodium_phoenicoides</i> | Poaceae          | Edge | L4P | positive | 166  | 90  |
| TAV | PV154      | <i>Convolvulus_arvensis</i>      | Convolvulaceae   | Edge | L4P | positive | 133  | 70  |
| TAV | PV150      | <i>Anthriscus_caucalis</i>       | Apiaceae         | Edge | L4P | positive | 129  | 62  |
| TAV | PV157      | <i>Lepidium_draba</i>            | Brassicaceae     | Edge | L4P | positive | 88   | 54  |
| TAV | PV153      | <i>Carduus_bourgeanus</i>        | Asteraceae       | Edge | L4P | positive | 37   | 24  |
| TAV | PV132      | <i>Convolvulus_arvensis</i>      | Convolvulaceae   | Edge | L2P | NA       | 29   | 4   |
| TAV | PV197      | <i>Anthriscus_caucalis</i>       | Apiaceae         | Edge | L1F | positive | 28   | 18  |
| TAV | PV140(abs) | <i>Silybum_marianum</i>          | Asteraceae       | Edge | L2P | NA       | 2    | -   |
| TAV | PV231(abs) | <i>Lepidium_draba</i>            | Brassicaceae     | Edge | L4F | positive | 2    | -   |
| TAV | PV014(abs) | <i>Datura_stramonium</i>         | Solanaceae       | Crop | M2V | NA       | 1    | -   |
| TAV | PV103(abs) | <i>Centaurea_melitensis</i>      | Asteraceae       | Oak  | Q2P | NA       | 7    | -   |
| TAV | PV009(abs) | <i>Quercus_coccifera</i>         | Fagaceae         | Oak  | Q1F | positive | 1    | -   |
| TAV | PV423(abs) | <i>Bromus_sp</i>                 | Poaceae          | Oak  | Q1P | positive | 2    | -   |

|      |            |                                   |                  |           |     |          |       |       |
|------|------------|-----------------------------------|------------------|-----------|-----|----------|-------|-------|
| TAV  | PV559(abs) | <i>Stipa_parviflora</i>           | Poaceae          | Oak       | Q1P | negative | 2     | -     |
| TAV  | PV112      | <i>Centranthus_calcitrapae</i>    | Caprifoliaceae   | Oak       | Q2P | positive | 6     | 6     |
| TAV  | PV105      | <i>Galium_verum</i>               | Rubiaceae        | Oak       | Q2P | positive | 5     | 4     |
| TAV  | PV441      | <i>Leontodon_sp</i>               | Asteraceae       | Oak       | Q2P | positive | 3     | 2     |
| TAV  | PV113      | <i>Carduus_bourgeanus</i>         | Asteraceae       | Oak       | Q2P | positive | 2     | 2     |
| TAV  | PV560(abs) | <i>Teucrium_chamaedrys</i>        | Lamiaceae        | Oak       | Q2P | positive | 1     | -     |
| TAV  | PV356      | <i>Hordeum_matritense</i>         | Poaceae          | Wasteland | E3P | positive | 10    | 4     |
| TuMV | PV023      | <i>Leontodon_sp</i>               | Asteraceae       | Wasteland | E4F | positive | 9     | 6     |
| TuMV | PV024      | <i>Verbascum_sinuatum</i>         | Scrophulariaceae | Wasteland | E4F | positive | 2     | 2     |
| TuMV | PV042      | <i>Quercus_coccifera</i>          | Fagaceae         | Oak       | Q3F | positive | 5     | 2     |
| TuMV | PV069(abs) | <i>Plantago_sp</i>                | Plantaginaceae   | Wasteland | E1F | positive | 2     | -     |
| TuMV | PV089      | <i>Marrubium_vulgare</i>          | Lamiaceae        | Oak       | Q1P | positive | 2     | 2     |
| TuMV | PV099(abs) | <i>Anchusa_undulata</i>           | Boraginaceae     | Wasteland | E2F | positive | 1     | -     |
| TuMV | PV107      | <i>Quercus_coccifera</i>          | Fagaceae         | Oak       | Q2P | positive | 2     | 2     |
| TuMV | PV111(abs) | <i>Teucrium_pseudochamaepitys</i> | Lamiaceae        | Oak       | Q2P | positive | 1     | -     |
| TuMV | PV127(abs) | <i>Torilis_nodosa</i>             | Apiaceae         | Edge      | L1P | positive | 1     | -     |
| TuMV | PV131      | <i>Chenopodium_album</i>          | Amaranthaceae    | Edge      | L2P | positive | 2     | 2     |
| TuMV | PV139      | <i>Rubia_peregrina</i>            | Rubiaceae        | Edge      | L2P | positive | 2     | 2     |
| TuMV | PV141(abs) | <i>Bromus_sp</i>                  | Poaceae          | Edge      | L3P | positive | 2     | -     |
| TuMV | PV142      | <i>Convolvulus_arvensis</i>       | Convolvulaceae   | Edge      | L3P | positive | 2     | 2     |
| TuMV | PV143      | <i>Diploaxis_erucoides</i>        | Brassicaceae     | Edge      | L3P | positive | 5,245 | 3,200 |

|      |            |                                |                |           |     |          |     |     |
|------|------------|--------------------------------|----------------|-----------|-----|----------|-----|-----|
| TuMV | PV147      | <i>Papaver_rhoeas</i>          | Papaveraceae   | Edge      | L3P | positive | 42  | 26  |
| TuMV | PV154(abs) | <i>Convolvulus_arvensis</i>    | Convolvulaceae | Edge      | L4P | positive | 1   | -   |
| TuMV | PV186      | <i>Sonchus_oleraceus</i>       | Asteraceae     | Edge      | L3V | positive | 149 | 86  |
| TuMV | PV209      | <i>Rumex_pulcher</i>           | Polygonaceae   | Edge      | L1F | positive | 2   | 2   |
| TuMV | PV313      | <i>Centranthus_calcitrapae</i> | Caprifoliaceae | Wasteland | E2P | positive | 5   | 2   |
| TuMV | PV356      | <i>Hordeum_matritense</i>      | Poaceae        | Wasteland | E3P | positive | 2   | 2   |
| TuMV | PV380      | <i>Dactylis_glomerata</i>      | Poaceae        | Wasteland | E4P | positive | 4   | 4   |
| TuMV | PV514      | <i>Diplotaxis_erucoides</i>    | Brassicaceae   | Crop      | M3V | positive | 2   | 2   |
| TuMV | PV519(abs) | <i>Desconocida_4</i>           | NA             | Crop      | M4V | positive | 1   | -   |
| WMV  | PV092      | <i>Anchusa_undulata</i>        | Boraginaceae   | Wasteland | E1F | positive | 39  | 18  |
| WMV  | PV356      | <i>Hordeum_matritense</i>      | Poaceae        | Wasteland | E3P | positive | 3   | 2   |
| WMV  | PV207      | <i>Picris_echioides</i>        | Asteraceae     | Edge      | L1F | positive | 4   | 2   |
| WMV  | PV163      | <i>Lolium_perenne</i>          | Poaceae        | Edge      | L1V | positive | 2   | 2   |
| WMV  | PV130      | <i>Carduus_bourgeanus</i>      | Asteraceae     | Edge      | L2P | positive | 147 | 103 |
| WMV  | PV172      | <i>Picris_echioides</i>        | Asteraceae     | Edge      | L2V | positive | 10  | 10  |
| WMV  | PV169      | <i>Galium_verum</i>            | Rubiaceae      | Edge      | L2V | negative | 10  | 6   |
| WMV  | PV173      | <i>Rubia_peregrina</i>         | Rubiaceae      | Edge      | L2V | positive | 4   | 4   |
| WMV  | PV214      | <i>Convolvulus_arvensis</i>    | Convolvulaceae | Edge      | L3F | positive | 16  | 12  |
| WMV  | PV212      | <i>Amaranthus_sp</i>           | Amaranthaceae  | Edge      | L3F | positive | 29  | 10  |
| WMV  | PV224      | <i>Vicia_sp</i>                | Fabaceae       | Edge      | L3F | positive | 14  | 6   |
| WMV  | PV216      | <i>Cynodon_dactylon</i>        | Poaceae        | Edge      | L3F | positive | 4   | 4   |

|      |          |                                |                |      |     |          |       |       |
|------|----------|--------------------------------|----------------|------|-----|----------|-------|-------|
| WMV  | PV218    | <i>Daucus sp</i>               | Apiaceae       | Edge | L3F | positive | 2     | 2     |
| WMV  | PV147    | <i>Papaver rhoeas</i>          | Papaveraceae   | Edge | L3P | negative | 404   | 203   |
| WMV  | PV143    | <i>Diplotaxis eruroides</i>    | Brassicaceae   | Edge | L3P | positive | 31    | 16    |
| WMV  | PV003    | <i>Cucumis melo</i>            | Cucurbitaceae  | Crop | M1V | positive | 3,734 | 1,399 |
| WMV  | PV004bgi | <i>Cyperus longus</i>          | Cyperaceae     | Crop | M1V | positive | 223   | 156   |
| WMV  | PV011    | <i>Portulaca oleraceae</i>     | Portulacaceae  | Crop | M1V | positive | 36    | 16    |
| WMV  | PV002    | <i>Convolvulus arvensis</i>    | Convolvulaceae | Crop | M1V | positive | 7     | 2     |
| WMV  | PV061    | <i>Cucumis melo</i>            | Cucurbitaceae  | Crop | M5V | positive | 11266 | 4460  |
| WMV  | PV006bgi | <i>Artemisia herba alba</i>    | Asteraceae     | Oak  | Q1F | positive | 12    | 6     |
| WMV  | PV009bgi | <i>Quercus coccifera</i>       | Fagaceae       | Oak  | Q1F | positive | 5     | 4     |
| WMV  | PV112    | <i>Centranthus calcitrapae</i> | Caprifoliaceae | Oak  | Q2P | positive | 11    | 10    |
| WMV  | PV110    | <i>Teucrium botrys</i>         | Lamiaceae      | Oak  | Q2P | positive | 13    | 6     |
| WMV  | PV105    | <i>Galium verum</i>            | Rubiaceae      | Oak  | Q2P | positive | 3     | 2     |
| WMV  | PV107    | <i>Quercus coccifera</i>       | Fagaceae       | Oak  | Q2P | positive | 3     | 2     |
| WMV  | PV113    | <i>Carduus bourgeanus</i>      | Asteraceae     | Oak  | Q2P | positive | 5     | 2     |
| WMV  | PV560    | <i>Teucrium chamaedrys</i>     | Lamiaceae      | Oak  | Q2P | positive | 11    | 2     |
| WMV  | PV051    | <i>Lotus corniculatus</i>      | Fabaceae       | Oak  | Q4P | positive | 2     | 2     |
| YoMV | PV061    | <i>Cucumis melo</i>            | Cucumis        | Crop | M5V | Positive | 4     | -     |
| YoMV | PV062    | <i>Cucumis melo</i>            | Cucumis        | Crop | M7V | Positive | 5     | 4     |
| YoMV | PV096    | <i>Milium vernale</i>          | Poaceae        | Edge | L2V | Positive | 12    | 8     |
| YoMV | PV097    | <i>Convolvulus arvensis</i>    | Convolvulaceae | Edge | L2V | Positive | 7     | -     |

|      |       |                                  |                |           |     |          |       |      |
|------|-------|----------------------------------|----------------|-----------|-----|----------|-------|------|
| YoMV | PV187 | <i>Brachypodium phoenicoides</i> | Poaceae        | Edge      | L2V | Positive | 4     | -    |
| YoMV | PV217 | <i>Datura stramonium</i>         | Solanaceae     | Edge      | L3F | Positive | 4     | -    |
| YoMV | PV188 | <i>Convolvulus arvensis</i>      | Convolvulaceae | Edge      | L3V | Positive | 4     | 4    |
| YoMV | PV211 | <i>Solanum nigrum</i>            | Solanaceae     | Edge      | L2F | Negative | 4     | 4    |
| YoMV | PV219 | <i>Diplotaxis eruroides</i>      | Brassicaceae   | Edge      | L3F | Negative | 2     | 2    |
| YoMV | PV081 | <i>Lithospermum arvense</i>      | Boraginaceae   | Edge      | L3V | Negative | 2     | -    |
| YoMV | PV084 | <i>Teucrium capitatum</i>        | Lamiaceae      | Oakwood   | Q1P | Positive | 105   | 72   |
| YoMV | PV426 | <i>Erodium cicutarium</i>        | Geraniaceae    | Oakwood   | Q1P | Positive | 121   | 26   |
| YoMV | PV088 | <i>Andryala arenaria</i>         | Asteraceae     | Oakwood   | Q1P | Positive | 11    | 2    |
| YoMV | PV423 | <i>Bromus sp</i>                 | Poaceae        | Oakwood   | Q1P | Positive | 10    | -    |
| YoMV | PV089 | <i>Marrubium vulgare</i>         | Lamiaceae      | Oakwood   | Q1P | Positive | 4     | 2    |
| YoMV | PV421 | <i>Asteriscus aquaticus</i>      | Asteraceae     | Oakwood   | Q1P | Positive | 6     | 2    |
| YoMV | PV430 | <i>Papaver rhoeas</i>            | Papaveraceae   | Oakwood   | Q1P | Positive | 2     | 2    |
| YoMV | PV441 | <i>Leontodon sp</i>              | Asteraceae     | Oakwood   | Q2P | Positive | 2     | -    |
| YoMV | PV464 | <i>Rubia peregrina</i>           | Rubiaceae      | Oakwood   | Q3P | Positive | 2     | -    |
| YoMV | PV454 | <i>Klasea pinnatifida</i>        | Asteraceae     | Oakwood   | Q3P | Positive | 1     | -    |
| YoMV | PV090 | <i>Dactylis glomerata</i>        | Poaceae        | Oakwood   | Q4P | Positive | 6     | -    |
| YoMV | PV559 | <i>Stipa parviflora</i>          | Poaceae        | Oakwood   | Q1P | Negative | 4     | 2    |
| YoMV | PV108 | <i>Quercus ilex</i>              | Fagaceae       | Oakwood   | Q2P | Negative | 2     | -    |
| YoMV | PV055 | <i>Thymus vulgaris</i>           | Lamiaceae      | Oakwood   | Q4P | Negative | 2     | 2    |
| YoMV | PV064 | <i>Artemisia herba alba</i>      | Asteraceae     | Wasteland | E1F | Positive | 21536 | 5994 |

|      |       |                             |                  |           |     |          |     |     |
|------|-------|-----------------------------|------------------|-----------|-----|----------|-----|-----|
| YoMV | PV069 | <i>Plantago</i>             | Plantaginaceae   | Wasteland | E1F | Positive | 6   | 4   |
| YoMV | PV065 | <i>Diplotaxis eruroides</i> | Brassicaceae     | Wasteland | E1F | Positive | 4   | -   |
| YoMV | PV070 | <i>Descurania sophia</i>    | Brassicaceae     | Wasteland | E1F | Positive | 2   | -   |
| YoMV | PV092 | <i>Anchusa undulata</i>     | Boraginaceae     | Wasteland | E1F | Positive | 2   | -   |
| YoMV | PV071 | <i>Salvia verbanaca</i>     | Lamiaceae        | Wasteland | E1F | Positive | 2   | 2   |
| YoMV | PV277 | <i>Anacyclus clavatus</i>   | Asteraceae       | Wasteland | E1P | Positive | 8   | 2   |
| YoMV | PV298 | <i>Teucrium capitatum</i>   | Lamiaceae        | Wasteland | E1P | Positive | 1   | -   |
| YoMV | PV257 | <i>Desconocida 5</i>        | Brassicaceae     | Wasteland | E3F | Positive | 4   | -   |
| YoMV | PV262 | <i>Galium verum</i>         | Rubiaceae        | Wasteland | E3F | Positive | 4   | 2   |
| YoMV | PV078 | <i>Bromus sp</i>            | Poaceae          | Wasteland | E4F | Positive | 119 | 18  |
| YoMV | PV079 | <i>Convolvulus arvensis</i> | Convolvulaceae   | Wasteland | E4F | Positive | 16  | 16  |
| YoMV | PV023 | <i>Leontodon sp</i>         | Asteraceae       | Wasteland | E4F | Positive | 2   | -   |
| YoMV | PV085 | <i>Avena sterilis</i>       | Poaceae          | Wasteland | E4P | Positive | 878 | 166 |
| YoMV | PV380 | <i>Dactylis glomerata</i>   | Poaceae          | Wasteland | E4P | Positive | 3   | -   |
| YoMV | PV086 | <i>Verbascum sinuatum</i>   | Scrophulariaceae | Wasteland | E4P | Positive | 2   | 2   |
| YoMV | PV075 | <i>Taraxacum officinale</i> | Asteraceae       | Wasteland | E2F | Negative | 2   | 2   |
| YoMV | PV352 | <i>Eryngium campestre</i>   | Apiaceae         | Wasteland | E3P | Negative | 1   | -   |
| YoMV | PV087 | <i>Echium vulgare</i>       | Boraginaceae     | Wasteland | E4P | Negative | 2   | 2   |
| YoMV | PV386 | <i>Papaver rhoeas</i>       | Papaveraceae     | Wasteland | E4P | Negative | 1   | -   |

(abs) = libraries absent from HTS6C, (p) = libraries present only in HTS6C.

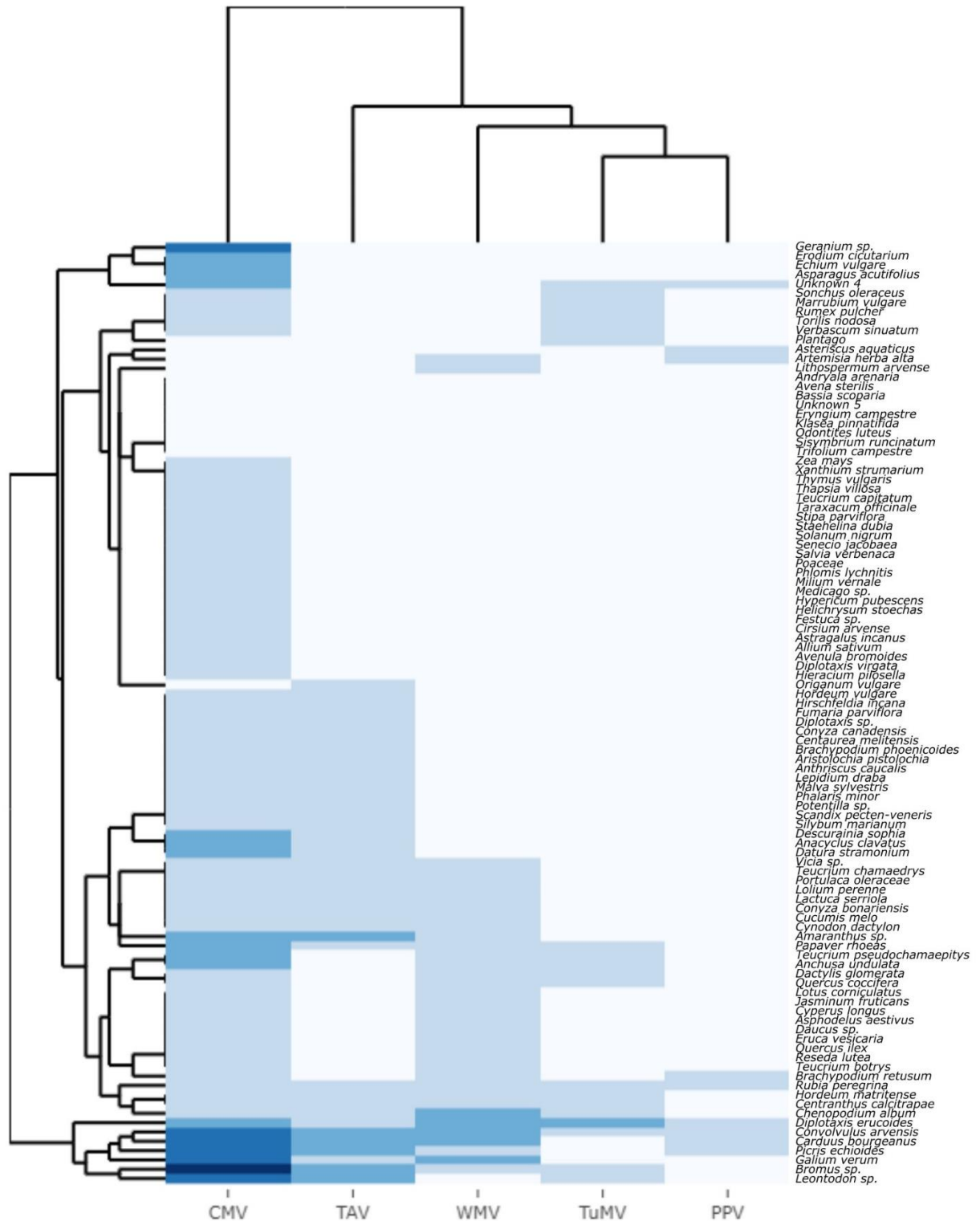


Figure A1: A heatmap and dendrograms that show distinctions in, and the degree of host-species use of five non-persistently transmitted viruses with different genomic structure.

**Table A4: *ih* of TAV indicating percentage (%) of infection and sample size of host (n) for each habitat. The hyphen ‘–’ indicates no record.**

| Host Taxon                       | Crop (%) | <i>n</i> | Edge (%) | <i>n</i> | Oak (%) | <i>n</i> | Wastel (%) | <i>n</i> | Total inc. (%) | Total <i>n</i> . |
|----------------------------------|----------|----------|----------|----------|---------|----------|------------|----------|----------------|------------------|
| <i>Amaranthus sp</i>             | 71       | 7        | 100      | 2        | -       | -        | -          | -        | 88.9           | 9                |
| <i>Anacyclus clavatus</i>        | -        | -        | 100      | 2        | -       | -        | -          | -        | 100            | 2                |
| <i>Anthriscus caucalis</i>       | -        | -        | 75       | 8        | -       | -        | -          | -        | 75             | 8                |
| <i>Aristolochia pistolochia</i>  | -        | -        | 100      | 3        | -       | -        | -          | -        | 100            | 3                |
| <i>Brachypodium phoenicoides</i> | -        | -        | 100      | 10       | -       | -        | -          | -        | 100            | 10               |
| <i>Bromus sp</i>                 | -        | -        | 85.7     | 7        | 25      | 4        | -          | -        | 63.6           | 11               |
| <i>Carduus_bourgeanus</i>        | 0        | 7        | 50       | 10       | 50      | 4        | -          | -        | 33.3           | 21               |
| <i>Centranthus calcitrapae</i>   | -        | -        | -        | -        | 87.5    | 8        | -          | -        | 87.5           | 8                |
| <i>Chenopodium album</i>         | -        | -        | 75       | 4        | -       | -        | -          | -        | 75             | 4                |
| <i>Convolvulus arvensis</i>      | 50       | 6        | 100      | 11       | -       | -        | -          | -        | 82.4           | 17               |
| <i>Conyza canadensis</i>         | -        | -        | 100      | 2        | -       | -        | -          | -        | 100            | 2                |
| <i>Cucumis melo</i>              | 66.7     | 27       | -        | -        | -       | -        | -          | -        | 66.7           | 27               |
| <i>Cynodon dactylon</i>          | -        | -        | 55.6     | 9        | -       | -        | -          | -        | 55.6           | 9                |
| <i>Diplotaxis eruroides</i>      | -        | -        | 46.2     | 13       | -       | -        | -          | -        | 46.2           | 13               |
| <i>Diplotaxis sp</i>             | -        | -        | 100      | 2        | -       | -        | -          | -        | 100            | 2                |
| <i>Fumaria parviflora</i>        | -        | -        | 91.7     | 12       | -       | -        | -          | -        | 91.7           | 12               |
| <i>Galium verum</i>              | -        | -        | -        | -        | 75      | 4        | -          | -        | 75             | 4                |

|                               |             |           |             |            |           |           |           |           |             |            |
|-------------------------------|-------------|-----------|-------------|------------|-----------|-----------|-----------|-----------|-------------|------------|
| <i>Hirschfeldia incana</i>    | -           | -         | 100         | 1          | -         | -         | -         | -         | 100         | 1          |
| <i>Hordeum matritense</i>     | -           | -         | -           | -          | -         | -         | 73        | 11        | 73          | 11         |
| <i>Lactuca serriola</i>       | -           | -         | 30.8        | 13         | -         | -         | -         | -         | 30.8        | 13         |
| <i>Leontodon sp</i>           | -           | -         | 100         | 1          | 25        | 4         | -         | -         | 40          | 5          |
| <i>Lepidium draba</i>         | -           | -         | 93.3        | 15         | -         | -         | -         | -         | 93.3        | 15         |
| <i>Lolium perenne</i>         | -           | -         | 88.9        | 9          | -         | -         | -         | -         | 88.9        | 9          |
| <i>Malva sylvestris</i>       | -           | -         | 100         | 4          | -         | -         | -         | -         | 100         | 4          |
| <i>Papaver rhoeas</i>         | -           | -         | 100         | 1          | -         | -         | -         | -         | 100         | 1          |
| <i>Picris echioides</i>       | -           | -         | 50          | 2          | -         | -         | -         | -         | 50          | 2          |
| <i>Portulaca oleraceae</i>    | 25          | 4         | -           | -          | -         | -         | -         | -         | 25          | 4          |
| <i>Rubia peregrina</i>        | -           | -         | 100         | 1          | -         | -         | -         | -         | 100         | 1          |
| <i>Scandix pecten veneris</i> | -           | -         | 100         | 1          | -         | -         | -         | -         | 100         | 1          |
| <i>Teucrium chamaedrys</i>    | -           | -         | -           | -          | 100       | 1         | -         | -         | 100         | 1          |
| <i>Vicia sp</i>               | -           | -         | 100         | 2          | -         | -         | -         | -         | 100         | 2          |
| <b>Total</b>                  | <b>54.9</b> | <b>51</b> | <b>77.9</b> | <b>145</b> | <b>60</b> | <b>25</b> | <b>73</b> | <b>11</b> | <b>70.7</b> | <b>232</b> |

**Table A5: *iL* of TAV indicating percentage (%) of infection and sample size of host (n) for each habitat. The hyphen ‘–’ indicates no record.**

| Host Taxon                       | Crop (%) | <i>n</i> | Edge (%) | <i>n</i> | Oak (%) | <i>n</i> | Wastel. (%) | <i>n</i> | Total inc. (%) | Total <i>n</i> . |
|----------------------------------|----------|----------|----------|----------|---------|----------|-------------|----------|----------------|------------------|
| <i>Amaranthus sp</i>             | 71       | 7        | 33.3     | 6        | -       | -        | -           | -        | 61.55          | 13               |
| <i>Anacyclus clavatus</i>        | 0        | 4        | 100      | 2        | -       | -        | 0           | 5        | 18.18          | 11               |
| <i>Anthriscus caucalis</i>       | -        | -        | 75       | 8        | -       | -        | -           | -        | 75.00          | 8                |
| <i>Aristolochia pistolochia</i>  | -        | -        | 100      | 3        | 0       | 1        | -           | -        | 75.00          | 4                |
| <i>Brachypodium phoenicoides</i> | -        | -        | 83.3     | 12       | -       | -        | -           | -        | 83.33          | 12               |
| <i>Bromus sp</i>                 | 0        | 1        | 85.7     | 7        | 10      | 10       | -           | -        | 38.87          | 18               |
| <i>Carduus bourgeanus</i>        | 0        | 7        | 33.3     | 15       | 50      | 4        | 0           | 8        | 20.57          | 34               |
| <i>Centranthus calcitrapae</i>   | -        | -        | -        | -        | 87.5    | 8        | 0           | 1        | 77.78          | 9                |
| <i>Chenopodium album</i>         | 0        | 7        | 18.8     | 16       | -       | -        | -           | -        | 13.04          | 23               |
| <i>Convolvulus arvensis</i>      | 50       | 6        | 20.4     | 54       | -       | -        | 0           | 8        | 20.60          | 68               |
| <i>Conyza canadensis</i>         | -        | -        | 15.4     | 13       | -       | -        | -           | -        | 15.38          | 13               |
| <i>Cucumis melo</i>              | 11.5     | 156      | -        | -        | -       | -        | -           | -        | 11.54          | 156              |
| <i>Cynodon dactylon</i>          | -        | -        | 33.3     | 15       | -       | -        | -           | -        | 33.36          | 15               |
| <i>Diplotaxis eruroides</i>      | 0        | 3        | 18.8     | 32       | -       | -        | 0           | 19       | 11.12          | 54               |
| <i>Diplotaxis sp</i>             | -        | -        | 100      | 2        | -       | -        | -           | -        | 100.00         | 2                |
| <i>Fumaria parviflora</i>        | 0        | 8        | 91.7     | 12       | -       | -        | -           | -        | 55.02          | 20               |
| <i>Galium verum</i>              | -        | -        | 0        | 2        | 25      | 12       | 0           | 9        | 13.04          | 23               |

|                               |             |            |             |            |             |           |            |            |             |            |
|-------------------------------|-------------|------------|-------------|------------|-------------|-----------|------------|------------|-------------|------------|
| <i>Hirschfeldia incana</i>    | 0           | 8          | 100         | 1          | -           | -         | -          | -          | 11.11       | 9          |
| <i>Hordeum matritense</i>     | -           | -          | -           | -          | 0           | 2         | 73         | 11         | 61.77       | 13         |
| <i>Lactuca serriola</i>       | -           | -          | 23.5        | 17         | -           | -         | -          | -          | 23.55       | 17         |
| <i>Leontodon sp</i>           | -           | -          | 100         | 1          | 25          | 4         | 0          | 43         | 4.17        | 48         |
| <i>Lepidium draba</i>         | -           | -          | 58.3        | 24         | -           | -         | -          | -          | 58.31       | 24         |
| <i>Lolium perenne</i>         | -           | -          | 66.7        | 12         | -           | -         | -          | -          | 66.68       | 12         |
| <i>Malva sylvestris</i>       | -           | -          | 66.7        | 6          | -           | -         | -          | -          | 66.67       | 6          |
| <i>Papaver rhoeas</i>         | -           | -          | 50          | 2          | 0           | 1         | 0          | 7          | 10.00       | 10         |
| <i>Picris echioides</i>       | 0           | 1          | 7.14        | 14         | -           | -         | 0          | 11         | 3.85        | 26         |
| <i>Portulaca oleraceae</i>    | 25          | 4          | -           | -          | -           | -         | -          | -          | 25.00       | 4          |
| <i>Rubia peregrina</i>        | 0           | 9          | 6.7         | 15         | 0           | 5         | 0          | 1          | 3.33        | 30         |
| <i>Scandix pecten veneris</i> | -           | -          | 100         | 1          | -           | -         | -          | -          | 100.00      | 1          |
| <i>Teucrium chamaedrys</i>    | -           | -          | -           | -          | 50          | 2         | 0          | 3          | 20.00       | 5          |
| <i>Vicia sp</i>               | -           | -          | 66.7        | 3          | -           | -         | 0          | 11         | 14.29       | 14         |
| <b>Total</b>                  | <b>12.7</b> | <b>221</b> | <b>38.3</b> | <b>295</b> | <b>30.6</b> | <b>49</b> | <b>5.9</b> | <b>137</b> | <b>23.4</b> | <b>702</b> |

**Table A6: *ih* of TuMV indicating percentage (%) of infection and sample size of host (n) for each habitat. The hyphen ‘–’ indicates no record.**

| <b>Host Taxon</b>              | <b>Crop (%)</b> | <b><i>n</i></b> | <b>Edge (%)</b> | <b><i>n</i></b> | <b>Oak. (%)</b> | <b><i>n</i></b> | <b>Wastel (%)</b> | <b><i>n</i></b> | <b>Total inc. (%)</b> | <b>Total <i>n</i>.</b> |
|--------------------------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-------------------|-----------------|-----------------------|------------------------|
| <i>Anchusa undulata</i>        | -               | -               | -               | -               | -               | -               | 38.5              | 13              | 38.5                  | 13                     |
| <i>Bromus sp</i>               | -               | -               | 100             | 2               | -               | -               | -                 | -               | 100                   | 2                      |
| <i>Centranthus calcitrapae</i> | -               | -               | -               | -               | -               | -               | 100               | 1               | 100                   | 1                      |
| <i>Convolvulus arvensis</i>    | -               | -               | 91              | 11              | -               | -               | -                 | -               | 91                    | 11                     |
| <i>Dactylis glomerata</i>      | -               | -               | -               | -               | -               | -               | 50                | 6               | 50                    | 6                      |
| <i>Diplotaxis eruroides</i>    | -               | -               | 38.5            | 13              | -               | -               | -                 | -               | 38.5                  | 13                     |
| <i>Hordeum matritense</i>      | -               | -               | -               | -               | -               | -               | 63.6              | 11              | 63.6                  | 11                     |
| <i>Leontodon sp</i>            | -               | -               | -               | -               | -               | -               | 8.3               | 24              | 8.3                   | 24                     |
| <i>Marrubium vulgare</i>       | -               | -               | -               | -               | 55.6            | 9               | -                 | -               | 55.6                  | 9                      |
| <i>Papaver rhoeas</i>          | -               | -               | 100             | 1               | -               | -               | -                 | -               | 100                   | 1                      |
| <i>Plantago sp</i>             | -               | -               | -               | -               | -               | -               | 40                | 10              | 40                    | 10                     |
| <i>Quercus coccifera</i>       | -               | -               | -               | -               | 33.3            | 18              | -                 | -               | 33.3                  | 18                     |
| <i>Rubia peregrina</i>         | -               | -               | 100             | 2               | -               | -               | -                 | -               | 100                   | 2                      |
| <i>Rumex pulcher</i>           | -               | -               | 75              | 4               | -               | -               | -                 | -               | 75                    | 4                      |
| <i>Sonchus oleraceus</i>       | -               | -               | 100             | 1               | -               | -               | -                 | -               | 100                   | 1                      |

|                           |   |   |             |           |             |           |             |           |             |            |
|---------------------------|---|---|-------------|-----------|-------------|-----------|-------------|-----------|-------------|------------|
| <i>Torilis nodosa</i>     | - | - | 33.3        | 6         | -           | -         | -           | -         | 33.3        | 6          |
| <i>Verbascum sinuatum</i> | - | - | -           | -         | -           | -         | 28.5        | 14        | 28.5        | 14         |
| <b>Total</b>              | - | - | <b>67.5</b> | <b>40</b> | <b>40.7</b> | <b>27</b> | <b>32.9</b> | <b>79</b> | <b>43.2</b> | <b>146</b> |

---

**Table A7: *iL* of TuMV indicating percentage (%) of infection and sample size of host (n) for each habitat. The hyphen ‘–’ indicates no record.**

| Host Taxon                     | Crop (%) | <i>n</i> | Edge (%) | <i>n</i> | Oak (%) | <i>n</i> | Wastel (%) | <i>n</i> | Total inc. (%) | Total <i>n</i> . |
|--------------------------------|----------|----------|----------|----------|---------|----------|------------|----------|----------------|------------------|
| <i>Anchusa undulata</i>        | -        | -        | 0        | 6        | -       | -        | 23.8       | 21       | 18.5           | 27               |
| <i>Bromus sp</i>               | 0        | 1        | 25       | 8        | 0       | 6        | -          | -        | 13.3           | 15               |
| <i>Centranthus calcitrapae</i> | -        | -        | -        | -        | 0       | 8        | 100        | 1        | 11.1           | 9                |
| <i>Convolvulus arvensis</i>    | 0        | 11       | 16.4     | 61       | -       | -        | -          | -        | 13.9           | 72               |
| <i>Dactylis glomerata</i>      | -        | -        | -        | -        | 0       | 21       | 50         | 6        | 11.1           | 27               |
| <i>Diplotaxis eruroides</i>    | -        | -        | 15.6     | 32       | -       | -        | 0          | 19       | 9.8            | 51               |
| <i>Hordeum matritense</i>      | -        | -        | -        | -        | 0       | 2        | 63.6       | 11       | 53.8           | 13               |
| <i>Leontodon sp</i>            | -        | -        | 0        | 1        | 0       | 5        | 4.7        | 43       | 4.1            | 49               |
| <i>Marrubium vulgare</i>       | -        | -        | -        | -        | 21.7    | 23       | -          | -        | 21.8           | 23               |
| <i>Papaver rhoeas</i>          | -        | -        | 50       | 2        | 0       | 1        | 0          | 7        | 10             | 10               |
| <i>Plantago sp</i>             | -        | -        | -        | -        | -       | -        | 40         | 10       | 40             | 10               |
| <i>Quercus coccifera</i>       | -        | -        | -        | -        | 12.5    | 48       | -          | -        | 12.5           | 48               |
| <i>Rubia peregrina</i>         | 0        | 9        | 13.3     | 15       | 0       | 5        | -          | -        | 6.9            | 29               |
| <i>Rumex pulcher</i>           | -        | -        | 30       | 10       | -       | -        | 0          | 4        | 10.7           | 28               |
| <i>Sonchus oleraceus</i>       | -        | -        | 33.3     | 3        | -       | -        | -          | -        | 33.3           | 3                |
| <i>Torilis nodosa</i>          | -        | -        | 33.3     | 6        | -       | -        | -          | -        | 33.3           | 6                |

|                           |          |           |             |            |             |            |             |            |             |            |
|---------------------------|----------|-----------|-------------|------------|-------------|------------|-------------|------------|-------------|------------|
| <i>Verbascum sinuatum</i> | -        | -         | -           | -          | -           | -          | 9.1         | 44         | 9.1         | 44         |
| <b>Total</b>              | <b>0</b> | <b>21</b> | <b>18.8</b> | <b>144</b> | <b>9.23</b> | <b>119</b> | <b>15.7</b> | <b>166</b> | <b>14.0</b> | <b>450</b> |

---

**Table A8: *ih* of PPV indicating percentage (%) of infection and sample size of host (n) for each habitat. The hyphen ‘–’ indicates no record.**

| Host Taxon                        | Crop (%)  | <i>n</i>  | Edge (%)    | <i>n</i>  | Oak (%)     | <i>n</i>  | Wastel (%)  | <i>n</i> | Total inc. (%) | Total <i>n</i> . |
|-----------------------------------|-----------|-----------|-------------|-----------|-------------|-----------|-------------|----------|----------------|------------------|
| <i>Aphyllanthes monspeliensis</i> | -         | -         | -           | -         | 10          | 10        | -           | -        | 10             | 10               |
| <i>Artemisia herba alba</i>       | -         | -         | -           | -         | 70.5        | 17        | -           | -        | 70.5           | 17               |
| <i>Asteriscus aquaticus</i>       | -         | -         | -           | -         | -           | -         | 66.7        | 3        | 66.7           | 3                |
| <i>Brachypodium phoenicoides</i>  | -         | -         | 100         | 1         | -           | -         | -           | -        | 100            | 1                |
| <i>Brachypodium retusum</i>       | -         | -         | 64.3        | 14        | -           | -         | -           | -        | 64.3           | 14               |
| <i>Carduus bourgeanus</i>         | 37.5      | 8         |             |           | -           | -         | -           | -        | 37.5           | 8                |
| <i>Convolvulus arvensis</i>       | -         | -         | 33.3        | 9         | -           | -         | -           | -        | 33.3           | 9                |
| <i>Desconocida 4</i>              | 85.7      | 7         | -           | -         | -           | -         | -           | -        | 85.7           | 7                |
| <i>Diploaxis erucoides</i>        | -         | -         | 100         | 3         | -           | -         |             | -        | 100            | 3                |
| <i>Picris echioides</i>           | -         | -         | 50          | 8         | -           | -         | -           | -        | 50             | 8                |
| <i>Rubia peregrina</i>            | -         | -         | 100         | 6         | -           | -         | -           | -        | 100            | 6                |
| <b>Total</b>                      | <b>60</b> | <b>15</b> | <b>63.4</b> | <b>41</b> | <b>48.1</b> | <b>27</b> | <b>66.7</b> | <b>3</b> | <b>58.1</b>    | <b>86</b>        |

**Table A9: *iL* of PPV indicating percentage (%) of infection and sample size of host (n) for each habitat. The hyphen ‘–’ indicates no record.**

| Host Taxon                        | Crop (%)    | <i>n</i>  | Edge (%)    | <i>n</i>   | Oak (%)   | <i>n</i>  | Wastel (%)  | <i>n</i>  | Total inc. (%) | Total <i>n</i> . |
|-----------------------------------|-------------|-----------|-------------|------------|-----------|-----------|-------------|-----------|----------------|------------------|
| <i>Aphyllanthes monspeliensis</i> | -           | -         | -           | -          | 10        | 10        | -           | -         | 10             | 10               |
| <i>Artemisia herba alba</i>       | -           | -         | -           | -          | 41.4      | 29        | 0           | 21        | 24.0           | 50               |
| <i>Asteriscus aquaticus</i>       | -           | -         | -           | -          | 0         | 4         | 66.7        | 3         | 28.6           | 7                |
| <i>Brachypodium phoenicoides</i>  | -           | -         | 9.1         | 11         | -         | -         | -           | -         | 9.1            | 11               |
| <i>Brachypodium retusum</i>       | -           | -         | 56.3        | 16         | -         | -         | -           | -         | 56.3           | 16               |
| <i>Carduus bourgeanus</i>         | 37.5        | 8         | 0           | 16         | 0         | 4         | 0           | 8         | 8.3            | 36               |
| <i>Convolvulus arvensis</i>       | 0           | 11        | 5.4         | 56         | -         | -         | 0           | 8         | 4.0            | 75               |
| <i>Desconocida 4</i>              | 54.5        | 11        | 0           | 1          | -         | -         | -           | -         | 50.0           | 12               |
| <i>Diplotaxis eruroides</i>       | 0           | 3         | 9.4         | 32         | -         | -         | 0           | 19        | 5.6            | 54               |
| <i>Picris echioides</i>           | 0           | 1         | 18.2        | 22         | -         | -         | 0           | 11        | 11.8           | 34               |
| <i>Rubia peregrina</i>            | 0           | 9         | 40          | 15         | 0         | 5         | 0           | 1         | 20             | 30               |
| <b>Total</b>                      | <b>20.9</b> | <b>43</b> | <b>15.4</b> | <b>169</b> | <b>25</b> | <b>52</b> | <b>2.82</b> | <b>71</b> | <b>23.3</b>    | <b>214</b>       |

**Table A10: *ih* of WMV indicating percentage (%) of infection and sample size of host (n) for each habitat. The hyphen ‘–’ indicates no record.**

| <b>Host Taxon</b>              | <b>Crop (%)</b> | <b><i>n</i></b> | <b>Edge (%)</b> | <b><i>n</i></b> | <b>Oak. (%)</b> | <b><i>n</i></b> | <b>Wastel (%)</b> | <b><i>n</i></b> | <b>Total inc. (%)</b> | <b>Total <i>n</i>.</b> |
|--------------------------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-------------------|-----------------|-----------------------|------------------------|
| <i>Amaranthus sp</i>           | -               | -               | 100             | 2               | -               | -               | -                 | -               | 100                   | 2                      |
| <i>Anchusa undulata</i>        | -               | -               | -               | -               | -               | -               | 50                | 8               | 50                    | 8                      |
| <i>Artemisia herba alba</i>    | -               | -               | -               | -               | 33.3            | 18              | -                 | -               | 33.3                  | 18                     |
| <i>Carduus bourgeanus</i>      | -               | -               | 100             | 3               | 100             | 4               | -                 | -               | 100                   | 7                      |
| <i>Centranthus calcitrapae</i> | -               | -               | -               | -               | 100             | 8               | -                 | -               | 100                   | 8                      |
| <i>Convolvulus arvensis</i>    | 100             | 6               | 71.4            | 7               | -               | -               | -                 | -               | 84.6                  | 13                     |
| <i>Cucumis melo</i>            | 85.2            | 27              | -               | -               | -               | -               | -                 | -               | 85.2                  | 27                     |
| <i>Cynodon dactylon</i>        |                 |                 | 100             | 5               | -               | -               | -                 | -               | 100                   | 5                      |
| <i>Cyperus longus</i>          | 66.7            | 3               | -               | -               | -               | -               | -                 | -               | 66.7                  | 3                      |
| <i>Daucus sp</i>               | -               | -               | 66.7            | 3               | -               | -               | -                 | -               | 66.7                  | 3                      |
| <i>Diplotaxis erucoides</i>    | -               | -               | 83.3            | 6               | -               | -               | -                 | -               | 83.3                  | 6                      |
| <i>Galium verum</i>            | -               | -               | 0               | 1               | 25              | 4               | -                 | -               | 20                    | 5                      |
| <i>Hordeum matritense</i>      | -               | -               | -               | -               | -               | -               | 45.5              | 11              | 45.5                  | 11                     |
| <i>Lolium perenne</i>          | -               | -               | 66.7            | 3               | -               | -               | -                 | -               | 66.7                  | 3                      |
| <i>Lotus corniculatus</i>      | -               | -               | -               | -               | 37.5            | 8               | -                 | -               | 37.5                  | 8                      |

|                            |             |           |             |           |             |           |             |           |             |            |
|----------------------------|-------------|-----------|-------------|-----------|-------------|-----------|-------------|-----------|-------------|------------|
| <i>Papaver rhoeas</i>      | -           | -         | <b>0</b>    | <b>1</b>  | -           | -         | -           | -         | <b>0</b>    | <b>1</b>   |
| <i>Picris echioides</i>    | -           | -         | <b>60</b>   | <b>10</b> | -           | -         | -           | -         | <b>60</b>   | <b>10</b>  |
| <i>Portulaca oleraceae</i> | <b>75</b>   | <b>4</b>  | -           | -         | -           | -         | -           | -         | <b>75</b>   | <b>4</b>   |
| <i>Quercus coccifera</i>   | -           | -         | -           | -         | <b>35.7</b> | <b>14</b> | -           | -         | <b>35.7</b> | <b>14</b>  |
| <i>Rubia peregrina</i>     | -           | -         | <b>66.7</b> | <b>3</b>  |             |           | -           | -         | <b>66.7</b> | <b>3</b>   |
| <i>Teucrium botrys</i>     | -           | -         | -           | -         | <b>66.7</b> | <b>6</b>  | -           | -         | <b>66.7</b> | <b>6</b>   |
| <i>Teucrium chamaedrys</i> | -           | -         | -           | -         | <b>100</b>  | <b>1</b>  | -           | -         | <b>100</b>  | <b>1</b>   |
| <i>Vicia sp</i>            | -           | -         | <b>100</b>  | <b>1</b>  | -           | -         | -           | -         | <b>100</b>  | <b>1</b>   |
| Total                      | <b>87.5</b> | <b>40</b> | <b>73.3</b> | <b>45</b> | <b>50.8</b> | <b>63</b> | <b>47.4</b> | <b>19</b> | <b>65.3</b> | <b>167</b> |

---

**Table A11: *iL* of WMV indicating percentage (%) of infection and sample size of host (n) for each habitat. The hyphen ‘–’ indicates no record.**

| <b>Host Taxon</b>              | <b>Crop (%)</b> | <b><i>n</i></b> | <b>Edge (%)</b> | <b><i>n</i></b> | <b>Oak (%)</b> | <b><i>n</i></b> | <b>Wastel (%)</b> | <b><i>n</i></b> | <b>Total inc. (%)</b> | <b>Total <i>n</i>.</b> |
|--------------------------------|-----------------|-----------------|-----------------|-----------------|----------------|-----------------|-------------------|-----------------|-----------------------|------------------------|
| <i>Amaranthus sp</i>           | -               | -               | 50              | 4               | -              | -               | -                 | -               | 50                    | 4                      |
| <i>Anchusa undulata</i>        | -               | -               | 0               | 6               | -              | -               | 19                | 21              | 14.8                  | 27                     |
| <i>Artemisia herba alba</i>    | -               | -               | -               | -               | 33.3           | 18              | -                 | -               | 33.3                  | 18                     |
| <i>Carduus bourgeanus</i>      | 0               | 7               | 27.3            | 11              | 100            | 4               | 0                 | 8               | 23.3                  | 30                     |
| <i>Centranthus calcitrapae</i> | -               | -               | -               | -               | 100            | 8               | 0                 | 1               | 88.9                  | 9                      |
| <i>Convolvulus arvensis</i>    | 100             | 6               | 13.2            | 38              | -              | -               | -                 | -               | 25.0                  | 44                     |
| <i>Cucumis melo</i>            | 27              | 89              | -               | -               | -              | -               | -                 | -               | 25.8                  | 89                     |
| <i>Cynodon dactylon</i>        |                 |                 | 38.5            | 13              | -              | -               | -                 | -               | 38.5                  | 13                     |
| <i>Cyperus longus</i>          | 66.7            | 3               | -               | -               | -              | -               | -                 | -               | 66.7                  | 3                      |
| <i>Daucus sp</i>               | -               | -               | 20              | 10              | -              | -               | -                 | -               | 20.0                  | 10                     |
| <i>Diplotaxis eruroides</i>    | -               | -               | 29.4            | 17              | -              | -               | 0                 | 19              | 13.9                  | 36                     |
| <i>Galium verum</i>            | -               | -               | 0               | 2               | 8.3            | 12              | 0                 | 9               | 4.3                   | 23                     |
| <i>Hordeum matritense</i>      | -               | -               | -               | -               | 0              | 2               | 45.5              | 11              | 38.5                  | 13                     |
| <i>Lolium perenne</i>          | -               | -               | 22.2            | 9               | -              | -               | -                 | -               | 22.2                  | 9                      |
| <i>Lotus corniculatus</i>      | -               | -               | -               | -               | 37.5           | 8               | -                 | -               | 37.5                  | 8                      |
| <i>Picris echioides</i>        | 0               | 1               | 27.2            | 22              | -              | -               | 0                 | 11              | 17.6                  | 34                     |

|                            |             |            |             |            |             |            |             |           |             |            |
|----------------------------|-------------|------------|-------------|------------|-------------|------------|-------------|-----------|-------------|------------|
| <i>Portulaca oleraceae</i> | 75          | 4          | -           | -          | -           | -          | -           | -         | 75          | 4          |
| <i>Quercus coccifera</i>   | -           | -          | -           | -          | 13.2        | 38         | -           | -         | 13.2        | 38         |
| <i>Rubia peregrina</i>     | 0           | 9          | 15.4        | 13         | 0           | 5          | 0           | 1         | 7.1         | 28         |
| <i>Teucrium botrys</i>     | -           | -          | -           | -          | 66.7        | 6          | -           | -         | 66.7        | 6          |
| <i>Teucrium chamaedrys</i> | -           | -          | -           | -          | 100         | 1          | 0           | 3         | 25          | 4          |
| <i>Vicia sp</i>            | -           | -          | 33.3        | 3          | -           | -          | 0           | 11        | 7.1         | 14         |
| <b>Total</b>               | <b>29.4</b> | <b>119</b> | <b>22.3</b> | <b>148</b> | <b>31.4</b> | <b>102</b> | <b>9.48</b> | <b>95</b> | <b>23.4</b> | <b>464</b> |

---

**Table A12: *ih* of BChV indicating percentage (%) of infection and sample size of host (n) for each habitat. The hyphen ‘–’ indicates no record.**

| Host Taxon                  | Crop (%) | <i>n</i> | Edge (%) | <i>n</i> | Oak (%) | <i>n</i> | Wastel (%) | <i>n</i> | Total inc. (%) | Total <i>n</i> . |
|-----------------------------|----------|----------|----------|----------|---------|----------|------------|----------|----------------|------------------|
| <i>Amaranthus sp</i>        | -        | -        | 50       | 2        | -       | -        | -          | -        | 50             | 2                |
| <i>Chenopodium album</i>    | -        | -        | 0        | 2        | -       | -        | -          | -        | 0              | 2                |
| <i>Convolvulus arvensis</i> | -        | -        | 0        | 13       | -       | -        | -          | -        | 0              | 13               |
| <i>Conyza bonariensis</i>   | -        | -        | 50       | 2        | -       | -        | -          | -        | 50             | 2                |
| <i>Datura stramonium</i>    | -        | -        | 0        | 1        | -       | -        | -          | -        | 0              | 1                |
| <i>Daucus sp</i>            | -        | -        | 16.7     | 6        | -       | -        | -          | -        | 16.7           | 6                |
| <i>Diplotaxis erucoides</i> | -        | -        | 53.8     | 13       | -       | -        | -          | -        | 53.8           | 13               |
| <i>Diplotaxis sp</i>        | -        | -        | 0        | 2        | -       | -        | -          | -        | 0              | 2                |
| <i>Echium vulgare</i>       | -        | -        | 0        | 1        | -       | -        | -          | -        | 0              | 1                |
| <i>Hirschfeldia incana</i>  | -        | -        | 100      | 1        | -       | -        | -          | -        | 100            | 1                |
| <i>Malva sylvestris</i>     | -        | -        | 100      | 1        | -       | -        | -          | -        | 100            | 1                |
| <i>Picris echioides</i>     | -        | -        | 0        | 2        | -       | -        | -          | -        | 0              | 2                |
| <i>Rubia peregrina</i>      | 88.9     | 9        | 100      | 1        | -       | -        | -          | -        | 90             | 10               |
| <i>Rumex pulcher</i>        | -        | -        | 50       | 2        | -       | -        | -          | -        | 50             | 2                |
| <i>Silybum marianum</i>     | -        | -        | 40       | 5        | -       | -        | -          | -        | 40             | 5                |

|                       |             |           |             |           |   |   |   |   |           |           |
|-----------------------|-------------|-----------|-------------|-----------|---|---|---|---|-----------|-----------|
| <i>Sisymbrium</i>     |             |           |             |           |   |   |   |   |           |           |
| <i>runcinatum</i>     | 100         | 3         | -           | -         | - | - | - | - | 100       | 3         |
| <i>Solanum nigrum</i> | -           | -         | 100         | 1         | - | - | - | - | 100       | 1         |
| <b>Total</b>          | <b>91.7</b> | <b>12</b> | <b>29.3</b> | <b>58</b> | - | - | - | - | <b>40</b> | <b>70</b> |

---

**Table A13: *iL* of BChV indicating percentage (%) of infection and sample size of host (n) for each habitat. The hyphen ‘–’ indicates no record.**

| Host Taxon                   | Crop (%)    | <i>n</i>  | Edge (%)   | <i>n</i>   | Oak (%)  | <i>n</i> | Wastel (%) | <i>n</i>  | Total inc. (%) | Total <i>n</i> . |
|------------------------------|-------------|-----------|------------|------------|----------|----------|------------|-----------|----------------|------------------|
| <i>Amaranthus sp</i>         | 0           | 8         | 6.3        | 16         | -        | -        | -          |           | 4.2            | 24               |
| <i>Conyza bonariensis</i>    | -           | -         | 12.5       | 8          | -        | -        | -          | -         | 12.5           | 8                |
| <i>Daucus sp</i>             | -           | -         | 10         | 10         | -        | -        | -          | -         | 10.02          | 10               |
| <i>Diplotaxis eruroides</i>  | 0           | 3         | 21.9       | 32         | -        | -        | 0          | 19        | 13.0           | 54               |
| <i>Hirschfeldia incana</i>   | 0           | 8         | 100        | 1          | -        | -        | -          | -         | 11.1           | 9                |
| <i>Malva sylvestris</i>      | -           | -         | 16.7       | 6          | -        | -        | -          | -         | 16.7           | 6                |
| <i>Rubia peregrina</i>       | 88.9        | 9         | 7.7        | 13         | 0        | 5        | 0          | 1         | 32.1           | 28               |
| <i>Rumex pulcher</i>         | -           | -         | 50         | 2          | -        | -        | 0          | 4         | 16.7           | 6                |
| <i>Silybum marianum</i>      | -           | -         | 12.5       | 16         | -        | -        | -          | -         | 12.5           | 16               |
| <i>Sisymbrium runcinatum</i> | 100         | 3         | -          | -          | -        | -        | 0          | 1         | 75             | 4                |
| <i>Solanum nigrum</i>        | 0           | 9         | 33.3       | 3          | -        | -        | -          | -         | 8.3            | 12               |
| <b>Total</b>                 | <b>14.1</b> | <b>40</b> | <b>8.9</b> | <b>107</b> | <b>0</b> | <b>5</b> | <b>0</b>   | <b>25</b> | <b>8.3</b>     | <b>336</b>       |

**Table A14: *ih* of PZSV indicating percentage (%) of infection and sample size of host (n) for each habitat. The hyphen ‘–’ indicates no record.**

| Host Taxon                       | Crop (%) | <i>n</i> | Edge (%) | <i>n</i> | Oak (%) | <i>n</i> | Wastel. (%) | <i>n</i> | Total inc. (%) | Total <i>n</i> . |
|----------------------------------|----------|----------|----------|----------|---------|----------|-------------|----------|----------------|------------------|
| <i>Anacyclus clavatus</i>        | -        | -        | -        | -        | -       | -        | 20          | 5        | 20             | 5                |
| <i>Anthriscus caucalis</i>       | -        | -        | 83.3     | 6        | -       | -        | -           | -        | 83.3           | 6                |
| <i>Asparagus acutifolius</i>     | -        | -        | -        | -        | 28.6    | 14       | -           | -        | 28.6           | 14               |
| <i>Asteriscus aquaticus</i>      | -        | -        | -        | -        | 0       | 4        | 66.7        | 3        | 28.6           | 7                |
| <i>Brachypodium phoenicoides</i> | -        | -        | 27.2     | 11       | -       | -        | -           | -        | 27.2           | 11               |
| <i>Brachypodium retusum</i>      | -        | -        | 57.1     | 7        | -       | -        | -           | -        | 57.1           | 7                |
| <i>Bromus sp</i>                 | -        | -        | 40       | 5        | -       | -        | -           | -        | 40             | 5                |
| <i>Carduus bourgeanus</i>        | 16.7     | 6        | 40       | 10       | -       | -        | -           | -        | 31.3           | 16               |
| <i>Convolvulus arvensis</i>      | -        | -        | 68.2     | 22       | -       | -        | -           | -        | 68.2           | 22               |
| <i>Conyza canadensis</i>         | -        | -        | 100      | 2        | -       | -        | -           | -        | 100            | 2                |
| <i>Cynodon dactylon</i>          | -        | -        | 85.7     | 7        | -       | -        | -           | -        | 85.7           | 7                |
| <i>Daucus sp</i>                 | -        | -        | 100      | 6        | -       | -        | -           | -        | 100            | 6                |
| <i>Desconocida 5</i>             | -        | -        | -        | -        | -       | -        | 11.1        | 9        | 11.1           | 9                |
| <i>Diplotaxis eruroides</i>      | -        | -        | 87.5     | 16       | -       | -        | -           | -        | 87.5           | 16               |
| <i>Erodium cicutarium</i>        | 100      | 1        | 50       | 2        | 66.7    | 3        | -           | -        | 66.7           | 6                |

|                               |              |          |              |            |              |           |              |           |           |            |
|-------------------------------|--------------|----------|--------------|------------|--------------|-----------|--------------|-----------|-----------|------------|
| <i>Fumaria parviflora</i>     | -            | -        | 44.4         | 9          | -            | -         | -            | -         | 44.4      | 9          |
| <i>Galium verum</i>           | -            | -        | 0            | 1          | -            | -         | 88.9         | 9         | 80        | 10         |
| <i>Geranium sp</i>            | -            | -        | 100          | 1          | -            | -         | -            | -         | 100       | 1          |
| <i>Hirschfeldia incana</i>    | -            | -        | 100          | 1          | -            | -         | -            | -         | 100       | 1          |
| <i>Hordeum matritense</i>     | -            | -        | -            | -          | -            | -         | 27.2         | 11        | 27.2      | 11         |
| <i>Lepidium draba</i>         | -            | -        | 77.8         | 9          | -            | -         | -            | -         | 77.8      | 9          |
| <i>Lithospermum arvense</i>   | -            | -        | 100          | 1          | -            | -         | -            | -         | 100       | 1          |
| <i>Lolium perenne</i>         | -            | -        | 100          | 6          | -            | -         | -            | -         | 100       | 6          |
| <i>Medicago sp</i>            | -            | -        | -            | -          | -            | -         | 50           | 6         | 50        | 6          |
| <i>Milium vernale</i>         | -            | -        | 100          | 6          | -            | -         | -            | -         | 100       | 6          |
| <i>Papaver rhoeas</i>         | -            | -        | 100          | 1          | 0            | 1         | -            | -         | 50        | 2          |
| <i>Picris echioides</i>       | -            | -        | 68.8         | 16         | -            | -         | -            | -         | 68.8      | 16         |
| <i>Rubia peregrina</i>        | -            | -        | 61.5         | 13         | 0            | 5         | -            | -         | 44.4      | 18         |
| <i>Rumex pulcher</i>          | -            | -        | 80           | 5          | -            | -         | -            | -         | 80        | 5          |
| <i>Scandix pecten veneris</i> | -            | -        | 100          | 1          | -            | -         | -            | -         | 100       | 1          |
| <i>Silybum marianum</i>       | -            | -        | 83.3         | 6          | -            | -         | -            | -         | 83.3      | 6          |
| <i>Solanum nigrum</i>         | -            | -        | 100          | 1          | -            | -         | -            | -         | 100       | 1          |
| <i>Teucrium capitatum</i>     | -            | -        | -            | -          | -            | -         | 33.3         | 15        | 33.3      | 15         |
| <i>Thapsia villosa</i>        | -            | -        | -            | -          | -            | -         | 25           | 4         | 25        | 4          |
| <i>Vicia sp</i>               | -            | -        | 100          | 3          | -            | -         | -            | -         | 100       | 3          |
| <b>Total</b>                  | <b>28.57</b> | <b>7</b> | <b>70.11</b> | <b>174</b> | <b>22.22</b> | <b>27</b> | <b>38.71</b> | <b>62</b> | <b>55</b> | <b>270</b> |

**Table A15: *iL* of PZSV indicating percentage (%) of infection and sample size of host (n) for each habitat. The hyphen ‘–’ indicates no record.**

| Host Taxon                       | Crop (%) | <i>n</i> | Edge (%) | <i>n</i> | Oak (%) | <i>n</i> | Wastel (%) | <i>n</i> | Total inc. (%) | Total <i>n</i> . |
|----------------------------------|----------|----------|----------|----------|---------|----------|------------|----------|----------------|------------------|
| <i>Anacyclus clavatus</i>        | 0        | 4        | -        | -        | -       | -        | 20         | 5        | 11.1           | 9                |
| <i>Anthriscus caucalis</i>       | -        | -        | 83.3     | 6        | -       | -        | -          | -        | 83.3           | 6                |
| <i>Asparagus acutifolius</i>     | -        | -        | -        | -        | 18.2    | 22       | -          | -        | 18.2           | 22               |
| <i>Asteriscus aquaticus</i>      | -        | -        | -        | -        | 0       | 4        | 66.7       | 3        | 28.6           | 7                |
| <i>Brachypodium phoenicoides</i> | -        | -        | 27.2     | 11       | -       | -        | -          | -        | 27.2           | 11               |
| <i>Brachypodium retusum</i>      | -        | -        | 57.1     | 7        | -       | -        | -          | -        | 57.1           | 7                |
| <i>Bromus sp</i>                 | -        | -        | 40       | 5        | 0       | 6        | 0          | 8        | 10.5           | 19               |
| <i>Carduus bourgeanus</i>        | 16.7     | 6        | 40       | 10       | 0       | 4        | 0          | 8        | 17.9           | 28               |
| <i>Convolvulus arvensis</i>      | 0        | 11       | 41.7     | 36       | -       | -        | -          | -        | 31.9           | 47               |
| <i>Conyza canadensis</i>         | -        | -        | 100      | 2        | -       | -        | -          | -        | 100            | 2                |
| <i>Cynodon dactylon</i>          | -        | -        | 85.7     | 7        | -       | -        | -          | -        | 85.7           | 7                |
| <i>Daucus sp</i>                 | -        | -        | 100      | 6        | -       | -        | -          | -        | 100            | 6                |
| <i>Desconocida 5</i>             | 0        | 4        | -        | -        | -       | -        | 11.1       | 9        | 7.7            | 13               |
| <i>Diplotaxis eruroides</i>      | -        | -        | 87.5     | 16       | -       | -        | 0          | 19       | 40             | 35               |
| <i>Erodium cicutarium</i>        | 100      | 1        | 50       | 2        | 66.7    | 3        | -          | -        | 66.7           | 6                |

|                               |            |           |             |            |            |           |             |            |             |            |
|-------------------------------|------------|-----------|-------------|------------|------------|-----------|-------------|------------|-------------|------------|
| <i>Fumaria parviflora</i>     | 0          | 8         | 44.4        | 9          | -          | -         | -           | -          | 23.5        | 17         |
| <i>Galium verum</i>           | -          | -         | 0           | 1          | 0          | 13        | 88.9        | 9          | 34.8        | 23         |
| <i>Geranium sp</i>            | -          | -         | 100         | 1          | 0          | 11        | 0           | 8          | 5           | 20         |
| <i>Hirschfeldia incana</i>    | 0          | 8         | 100         | 1          | -          | -         | -           | -          | 11.1        | 9          |
| <i>Hordeum matritense</i>     | -          | -         | -           | -          | -          | -         | 27.2        | 11         | 27.2        | 11         |
| <i>Lepidium draba</i>         | -          | -         | 77.8        | 9          | -          | -         | -           | -          | 77.8        | 9          |
| <i>Lithospermum arvense</i>   | 0          | 22        | 100         | 1          | -          | -         | -           | -          | 4.35        | 23         |
| <i>Lolium perenne</i>         | -          | -         | 100         | 6          | -          | -         | -           | -          | 100         | 6          |
| <i>Medicago sp</i>            | -          | -         | -           | -          | -          | -         | 50          | 6          | 50          | 6          |
| <i>Milium vernale</i>         | -          | -         | 100         | 6          | -          | -         | -           | -          | 100         | 6          |
| <i>Papaver rhoeas</i>         | -          | -         | 100         | 1          | 0          | 1         | 0           | 7          | 11.1        | 9          |
| <i>Picris echioides</i>       | 0          | 1         | 68.8        | 16         | -          | -         | 0           | 11         | 39.3        | 28         |
| <i>Rubia peregrina</i>        | 0          | 9         | 61.5        | 13         | 0          | 5         | 0           | 1          | 28.5        | 28         |
| <i>Rumex pulcher</i>          | -          | -         | 80          | 5          | -          | -         | 0           | 4          | 44.4        | 9          |
| <i>Scandix pecten veneris</i> | -          | -         | 100         | 1          | -          | -         | -           | -          | 100         | 1          |
| <i>Silybum marianum</i>       | -          | -         | 83.3        | 6          | -          | -         | -           | -          | 83.3        | 6          |
| <i>Solanum nigrum</i>         | 0          | 9         | 100         | 1          | -          | -         | -           | -          | 10          | 10         |
| <i>Teucrium capitatum</i>     | -          | -         | -           | -          | 0          | 10        | 33.3        | 15         | 20.0        | 25         |
| <i>Thapsia villosa</i>        | -          | -         | -           | -          | -          | -         | 25          | 4          | 25          | 4          |
| <i>Vicia sp</i>               | -          | -         | 100         | 3          | -          | -         | 0           | 11         | 21.4        | 14         |
| <b>Total</b>                  | <b>2.4</b> | <b>83</b> | <b>64.9</b> | <b>188</b> | <b>7.6</b> | <b>79</b> | <b>17.3</b> | <b>139</b> | <b>30.4</b> | <b>489</b> |

**Table A16: *ih* of YoMV indicating percentage (%) of infection and sample size of host (n) for each habitat. The hyphen ‘-’ indicates no record.**

| <b>Host Taxon</b>                | <b>Crop (%)</b> | <b><i>n</i></b> | <b>Edge (%)</b> | <b><i>n</i></b> | <b>Oak (%)</b> | <b><i>n</i></b> | <b>Wastel (%)</b> | <b><i>n</i></b> | <b>Total inc. (%)</b> | <b>Total <i>n</i>.</b> |
|----------------------------------|-----------------|-----------------|-----------------|-----------------|----------------|-----------------|-------------------|-----------------|-----------------------|------------------------|
| <i>Anacyclus clavatus</i>        | -               | -               | -               | -               | -              | -               | 60                | 5               | 60                    | 5                      |
| <i>Anchusa undulata</i>          | -               | -               | -               | -               | -              | -               | 12.5              | 8               | 12.5                  | 8                      |
| <i>Andryala arenaria</i>         | -               | -               | -               | -               | 55.6           | 9               | -                 | -               | 55.6                  | 9                      |
| <i>Artemisia herba alba</i>      | -               | -               | -               | -               | -              | -               | 14.3              | 21              | 14.3                  | 21                     |
| <i>Asteriscus aquaticus</i>      | -               | -               | -               | -               | 25             | 4               | -                 | -               | 25                    | 4                      |
| <i>Avena sterilis</i>            | -               | -               | -               | -               | -              | -               | 40.7              | 27              | 40.7                  | 27                     |
| <i>Brachypodium phoenicoides</i> | -               | -               | 100             | 1               | -              | -               | -                 | -               | 100                   | 1                      |
| <i>Bromus sp</i>                 | -               | -               | -               | -               | 66.7           | 6               | 60                | 5               | 63.6                  | 11                     |
| <i>Convolvulus arvensis</i>      | -               | -               | 36.4            | 11              | -              | -               | 62.5              | 8               | 47.4                  | 19                     |
| <i>Cucumis melo</i>              | 27.5            | 69              | -               | -               | -              | -               | -                 | -               | 27.5                  | 69                     |
| <i>Dactylis glomerata</i>        | -               | -               | -               | -               | 15.4           | 13              | 33.3              | 6               | 21.1                  | 19                     |
| <i>Datura stramonium</i>         | -               | -               | 100             | 1               | -              | -               | -                 | -               | 100                   | 1                      |
| <i>Desconocida 5</i>             | -               | -               | -               | -               | -              | -               | 11.1              | 9               | 11.1                  | 9                      |
| <i>Descurania sophia</i>         | -               | -               | -               | -               | -              | -               | 100               | 6               | 100                   | 6                      |
| <i>Diploaxis erucoides</i>       | -               | -               | 0               | 8               | -              | -               | 36.8              | 19              | 25.9                  | 27                     |

|                           |             |           |             |           |             |           |             |            |             |            |
|---------------------------|-------------|-----------|-------------|-----------|-------------|-----------|-------------|------------|-------------|------------|
| <i>Erodium cicutarium</i> | -           | -         | -           | -         | 66.7        | 3         | -           | -          | 66.7        | 3          |
| <i>Galium verum</i>       | -           | -         | -           | -         | -           | -         | 33.3        | 9          | 33.3        | 9          |
| <i>Klasea pinnatifida</i> | -           | -         | -           | -         | 45.5        | 11        | -           | -          | 45.5        | 11         |
| <i>Leontodon sp</i>       | -           | -         | -           | -         | 50          | 4         | 4.2         | 24         | 10.7        | 28         |
| <i>Marrubium vulgare</i>  | -           | -         | -           | -         | 11.1        | 9         | -           | -          | 11.1        | 9          |
| <i>Milium vernale</i>     | -           | -         | 28.6        | 7         | -           | -         | -           | -          | 28.6        | 7          |
| <i>Papaver rhoeas</i>     | -           | -         | -           | -         | 100         | 1         | 0           | 7          | 12.5        | 8          |
| <i>Plantago sp</i>        | -           | -         | -           | -         | -           | -         | 45.5        | 11         | 45.5        | 11         |
| <i>Rubia peregrina</i>    | -           | -         | -           | -         | 60          | 5         | -           | -          | 60          | 5          |
| <i>Salvia verbanaca</i>   | -           | -         | -           | -         | -           | -         | 30          | 10         | 30          | 10         |
| <i>Teucrium capitatum</i> | -           | -         | -           | -         | 60          | 10        | 33.3        | 15         | 44          | 25         |
| <i>Verbascum sinuatum</i> | -           | -         | -           | -         | -           | -         | 21.1        | 19         | 21.1        | 19         |
| <b>Total</b>              | <b>27.5</b> | <b>69</b> | <b>26.7</b> | <b>28</b> | <b>32.9</b> | <b>75</b> | <b>27.6</b> | <b>209</b> | <b>28.8</b> | <b>381</b> |

---

**Table A17: *iL* of YoMV indicating percentage (%) of infection and sample size of host (n) for each habitat. The hyphen ‘–’ indicates no record.**

| Host Taxon                       | Crop (%) | <i>n</i> | Edge (%) | <i>n</i> | Oak (%) | <i>n</i> | Wastel. (%) | <i>n</i> | Total inc. (%) | Total <i>n</i> . |
|----------------------------------|----------|----------|----------|----------|---------|----------|-------------|----------|----------------|------------------|
| <i>Anacyclus clavatus</i>        | 0        | 4        | 0        | 2        | -       | -        | 60          | 5        | 60             | 5                |
| <i>Anchusa undulata</i>          | -        | -        | 0        | 6        | -       | -        | 4.76        | 21       | 3.7            | 27               |
| <i>Andryala arenaria</i>         | -        | -        | -        | -        | 55.6    | 9        | -           | -        | 55.6           | 9                |
| <i>Artemisia herba alba</i>      | -        | -        | -        | -        | 0       | 31       | 14.3        | 21       | 5.77           | 52               |
| <i>Asteriscus aquaticus</i>      | -        | -        | -        | -        | 25      | 4        | 0           | 3        | 14.3           | 7                |
| <i>Avena sterilis</i>            | -        | -        | -        | -        | -       | -        | 40.7        | 27       | 40.7           | 27               |
| <i>Brachypodium phoenicoides</i> | -        | -        | 100      | 13       | -       | -        | -           | -        | 7.69           | 13               |
| <i>Bromus sp</i>                 | 0        | 1        | 0        | 8        | 66.7    | 6        | 60          | 5        | 35             | 20               |
| <i>Convolvulus arvensis</i>      | 0        | 11       | 36.4     | 39       | -       | -        | 62.5        | 8        | 15.5           | 58               |
| <i>Cucumis melo</i>              | 11.5     | 165      | -        | -        | -       | -        | -           | -        | 11.5           | 165              |
| <i>Dactylis glomerata</i>        | -        | -        | -        | -        | 9.52    | 21       | 33.3        | 6        | 14.8           | 27               |
| <i>Datura stramonium</i>         | 0        | 19       | 50       | 2        | -       | -        | -           | -        | 9.52           | 21               |
| <i>Desconocida 5</i>             | -        | -        | -        | -        | -       | -        | 11.1        | 9        | 11.1           | 9                |
| <i>Descurania sophia</i>         | -        | -        | 0        | 2        | -       | -        | 100         | 6        | 75             | 8                |
| <i>Diploaxis erucoides</i>       | -        | -        | 0        | 23       | -       | -        | 36.8        | 19       | 16.7           | 42               |

|                           |            |            |            |            |             |            |             |            |             |            |
|---------------------------|------------|------------|------------|------------|-------------|------------|-------------|------------|-------------|------------|
| <i>Erodium cicutarium</i> | 0          | 1          | 0          | 2          | 66.7        | 3          | 0           | 3          | 22.2        | 9          |
| <i>Galium verum</i>       | -          | -          | 0          | 2          | 0           | 13         | 33.3        | 9          | 12.5        | 24         |
| <i>Klasea pinnatifida</i> | 0          | 1          | -          | -          | 45.5        | 11         | -           | -          | 41.7        | 12         |
| <i>Leontodon sp</i>       | -          | -          | 0          | 1          | 50          | 4          | 2.33        | 43         | 6.25        | 48         |
| <i>Marrubium vulgare</i>  | -          | -          | -          | -          | 4.35        | 23         | -           | -          | 4.35        | 23         |
| <i>Milium vernale</i>     | -          | -          | 14.3       | 14         | -           | -          | -           | -          | 14.3        | 14         |
| <i>Papaver rhoeas</i>     | -          | -          | 0          | 2          | 100         | 1          | 0           | 7          | 10          | 10         |
| <i>Plantago sp</i>        | -          | -          | -          | -          | -           | -          | 45.5        | 11         | 45.5        | 11         |
| <i>Rubia peregrina</i>    | 0          | 9          | 0          | 15         | 60          | 5          | 0           | 1          | 12.5        | 24         |
| <i>Salvia verbanaca</i>   | -          | -          | -          | -          | -           | -          | 30          | 10         | 30          | 10         |
| <i>Teucrium capitatum</i> | -          | -          | -          | -          | 60          | 10         | 33.3        | 15         | 44          | 25         |
| <i>Verbascum sinuatum</i> | -          | -          | -          | -          | -           | -          | 8.89        | 45         | 8.89        | 45         |
| <b>Total</b>              | <b>9.0</b> | <b>211</b> | <b>6.1</b> | <b>131</b> | <b>22.7</b> | <b>141</b> | <b>23.0</b> | <b>274</b> | <b>16.4</b> | <b>745</b> |

---

**Table A18: Post hoc analyses showing pairwise comparison of *ih* across virus species.**

| OTU contrast | odds.ratio | 1/odds.ratio | std.error | statistic | adj.p.value | Lower CI | Upper CI |
|--------------|------------|--------------|-----------|-----------|-------------|----------|----------|
| CMV / BChV   | 8.712      | 0.115        | 2.555     | 7.382     | 0.000       | 3.582    | 21.188   |
| PPV / BChV   | 3.377      | 0.296        | 1.175     | 3.500     | 0.011       | 1.177    | 9.690    |
| PPV / CMV    | 0.388      | 2.579        | 0.098     | -3.753    | 0.004       | 0.180    | 0.833    |
| PZSV / BChV  | 2.939      | 0.340        | 0.857     | 3.699     | 0.005       | 1.215    | 7.111    |
| PZSV / CMV   | 0.337      | 2.964        | 0.060     | -6.124    | 0.000       | 0.197    | 0.578    |
| PZSV / PPV   | 0.870      | 1.149        | 0.230     | -0.526    | 1.000       | 0.391    | 1.938    |
| TAV / BChV   | 4.109      | 0.243        | 1.218     | 4.766     | 0.000       | 1.673    | 10.093   |
| TAV / CMV    | 0.472      | 2.120        | 0.087     | -4.064    | 0.001       | 0.269    | 0.826    |
| TAV / PPV    | 1.217      | 0.822        | 0.326     | 0.731     | 0.996       | 0.540    | 2.741    |
| TAV / PZSV   | 1.398      | 0.715        | 0.272     | 1.722     | 0.673       | 0.775    | 2.520    |
| TuMV / BChV  | 2.738      | 0.365        | 0.887     | 3.107     | 0.040       | 1.025    | 7.312    |
| TuMV / CMV   | 0.314      | 3.182        | 0.065     | -5.618    | 0.000       | 0.168    | 0.587    |
| TuMV / PPV   | 0.811      | 1.234        | 0.237     | -0.717    | 0.997       | 0.334    | 1.968    |
| TuMV / PZSV  | 0.931      | 1.074        | 0.208     | -0.318    | 1.000       | 0.474    | 1.831    |
| TuMV / TAV   | 0.666      | 1.501        | 0.156     | -1.732    | 0.666       | 0.327    | 1.356    |
| WMV / BChV   | 5.634      | 0.177        | 1.801     | 5.409     | 0.000       | 2.138    | 14.845   |
| WMV / CMV    | 0.647      | 1.546        | 0.128     | -2.198    | 0.352       | 0.355    | 1.179    |
| WMV / PPV    | 1.668      | 0.599        | 0.474     | 1.800     | 0.620       | 0.705    | 3.949    |
| WMV / PZSV   | 1.917      | 0.522        | 0.426     | 2.928     | 0.067       | 0.977    | 3.759    |
| WMV / TAV    | 1.371      | 0.729        | 0.310     | 1.398     | 0.859       | 0.692    | 2.719    |

|             |       |       |       |         |       |       |       |
|-------------|-------|-------|-------|---------|-------|-------|-------|
| WMV / TuMV  | 2.058 | 0.486 | 0.514 | 2.892   | 0.074 | 0.966 | 4.384 |
| YoMV / BChV | 1.549 | 0.645 | 0.462 | 1.470   | 0.824 | 0.628 | 3.823 |
| YoMV / CMV  | 0.178 | 5.622 | 0.027 | -11.289 | 0.000 | 0.112 | 0.283 |
| YoMV / PPV  | 0.459 | 2.180 | 0.119 | -2.998  | 0.055 | 0.209 | 1.009 |
| YoMV / PZSV | 0.527 | 1.897 | 0.097 | -3.474  | 0.012 | 0.302 | 0.922 |
| YoMV / TAV  | 0.377 | 2.652 | 0.073 | -5.058  | 0.000 | 0.210 | 0.676 |
| YoMV / TuMV | 0.566 | 1.767 | 0.119 | -2.718  | 0.117 | 0.300 | 1.068 |
| YoMV / WMV  | 0.275 | 3.636 | 0.057 | -6.234  | 0.000 | 0.147 | 0.515 |

**Table A19: Post hoc analyses showing pairwise comparison of *ih* in virus genera**

| Genera contrast              | odds.ratio | 1/odds.ratio | std.error | statistic | adj.p.value | Lower CI | Upper CI |
|------------------------------|------------|--------------|-----------|-----------|-------------|----------|----------|
| Cucumovirus /<br>Anulavirus  | 2.193      | 0.456        | 0.351     | 4.908     | 0.000       | 1.417    | 3.393    |
| Polerovirus /<br>Anulavirus  | 0.350      | 2.854        | 0.102     | -3.609    | 0.003       | 0.159    | 0.774    |
| Polerovirus /<br>Cucumovirus | 0.160      | 6.260        | 0.045     | -6.549    | 0.000       | 0.074    | 0.343    |
| Potyvirus /<br>Anulavirus    | 1.231      | 0.813        | 0.210     | 1.214     | 0.743       | 0.772    | 1.962    |
| Potyvirus /<br>Cucumovirus   | 0.561      | 1.782        | 0.075     | -4.341    | 0.000       | 0.390    | 0.807    |
| Potyvirus /<br>Polerovirus   | 3.513      | 0.285        | 1.007     | 4.381     | 0.000       | 1.607    | 7.679    |

|                              |       |       |       |         |       |       |       |
|------------------------------|-------|-------|-------|---------|-------|-------|-------|
| Tobamovirus /<br>Anulavirus  | 0.486 | 2.057 | 0.089 | -3.960  | 0.001 | 0.296 | 0.799 |
| Tobamovirus /<br>Cucumovirus | 0.222 | 4.511 | 0.031 | -10.626 | 0.000 | 0.151 | 0.326 |
| Tobamovirus /<br>Polerovirus | 1.388 | 0.721 | 0.410 | 1.109   | 0.802 | 0.620 | 3.108 |
| Tobamovirus /<br>Potyvirus   | 0.395 | 2.531 | 0.061 | -5.971  | 0.000 | 0.259 | 0.604 |

---

**Table A20: Post hoc analyses showing pairwise comparison of *ih* across habitat.**

| Habitat contrast | odds.ratio | 1/odds.ratio | std.error | statistic | adj.p.value | Lower CI | Upper CI |
|------------------|------------|--------------|-----------|-----------|-------------|----------|----------|
| Edge / Crop      | 1.009      | 0.992        | 0.169     | 0.051     | 1.000       | 0.656    | 1.551    |
| Oak / Crop       | 0.421      | 2.373        | 0.069     | -5.271    | 0.000       | 0.277    | 0.642    |
| Oak / Edge       | 0.418      | 2.393        | 0.059     | -6.154    | 0.000       | 0.29     | 0.602    |
| Wasteland / Crop | 0.358      | 2.789        | 0.058     | -6.331    | 0.000       | 0.236    | 0.544    |
| Wasteland / Edge | 0.355      | 2.813        | 0.050     | -7.363    | 0.000       | 0.248    | 0.51     |
| Wasteland / Oak  | 0.851      | 1.176        | 0.117     | -1.179    | 0.640       | 0.598    | 1.21     |

**Table A21: Post hoc analyses showing pairwise comparison of *ih* across transmission modes.**

| Transmission contrast | odds.ratio | 1/odds.ratio | std.error | statistic | adj.p.value | Lower CI | Upper CI |
|-----------------------|------------|--------------|-----------|-----------|-------------|----------|----------|
| Non-pers/ contact     | 3.636      | 0.275        | 0.480     | 9.786     | 0.000       | 2.591    | 5.102    |
| Persistent / contact  | 0.718      | 1.392        | 0.212     | -1.120    | 0.677       | 0.336    | 1.534    |
| Persistent / Non-pers | 0.198      | 5.061        | 0.054     | -5.899    | 0.000       | 0.098    | 0.400    |
| Pollen / Contact      | 2.063      | 0.485        | 0.376     | 3.978     | 0.000       | 1.293    | 3.294    |
| Pollen / Non-pers     | 0.568      | 1.762        | 0.086     | -3.759    | 0.001       | 0.385    | 0.836    |
| Pollen / Persistent   | 2.872      | 0.348        | 0.835     | 3.631     | 0.002       | 1.362    | 6.060    |

NonPers = aphid non-persistent transmission, persistent= aphid persistent transmission

**Table A22: Post-hoc analyses showing pairwise comparison of *iL* across virus species.**

| <b>Virus contrast</b> | <b>odds.ratio</b> | <b>1/odds.ratio</b> | <b>std.error</b> | <b>statistic</b> | <b>adj.p.value</b> | <b>Lower CI</b> | <b>Upper CI</b> |
|-----------------------|-------------------|---------------------|------------------|------------------|--------------------|-----------------|-----------------|
| CMV / BChV            | 12.151            | 0.082               | 2.820            | 10.761           | 0.000              | 6.014           | 24.552          |
| PPV / BChV            | 1.690             | 0.592               | 0.449            | 1.974            | 0.499              | 0.755           | 3.781           |
| PPV / CMV             | 0.139             | 7.190               | 0.025            | -11.172          | 0.000              | 0.081           | 0.238           |
| PZSV / BChV           | 5.265             | 0.190               | 1.249            | 7.003            | 0.000              | 2.565           | 10.805          |
| PZSV / CMV            | 0.433             | 2.308               | 0.053            | -6.801           | 0.000              | 0.299           | 0.629           |
| PZSV / PPV            | 3.115             | 0.321               | 0.582            | 6.086            | 0.000              | 1.769           | 5.487           |
| TAV / BChV            | 2.486             | 0.402               | 0.575            | 3.937            | 0.002              | 1.233           | 5.012           |
| TAV / CMV             | 0.205             | 4.887               | 0.025            | -13.090          | 0.000              | 0.142           | 0.295           |
| TAV / PPV             | 1.471             | 0.680               | 0.265            | 2.142            | 0.388              | 0.852           | 2.541           |
| TAV / PZSV            | 0.472             | 2.118               | 0.064            | -5.568           | 0.000              | 0.314           | 0.710           |
| TuMV / BChV           | 1.442             | 0.694               | 0.367            | 1.437            | 0.840              | 0.667           | 3.119           |
| TuMV / CMV            | 0.119             | 8.428               | 0.018            | -13.869          | 0.000              | 0.075           | 0.189           |
| TuMV / PPV            | 0.853             | 1.172               | 0.178            | -0.759           | 0.995              | 0.453           | 1.608           |
| TuMV / PZSV           | 0.274             | 3.652               | 0.045            | -7.811           | 0.000              | 0.166           | 0.453           |
| TuMV / TAV            | 0.580             | 1.724               | 0.095            | -3.322           | 0.020              | 0.353           | 0.953           |
| WMV / BChV            | 3.713             | 0.269               | 0.900            | 5.411            | 0.000              | 1.781           | 7.743           |
| WMV / CMV             | 0.306             | 3.272               | 0.041            | -8.742           | 0.000              | 0.203           | 0.461           |
| WMV / PPV             | 2.197             | 0.455               | 0.425            | 4.070            | 0.001              | 1.223           | 3.949           |
| WMV / PZSV            | 0.705             | 1.418               | 0.105            | -2.339           | 0.272              | 0.449           | 1.109           |
| WMV / TAV             | 1.493             | 0.670               | 0.215            | 2.790            | 0.097              | 0.966           | 2.309           |
| WMV / TuMV            | 2.575             | 0.388               | 0.452            | 5.386            | 0.000              | 1.512           | 4.386           |
| YoMV / BChV           | 2.195             | 0.456               | 0.529            | 3.261            | 0.025              | 1.057           | 4.558           |
| YoMV / CMV            | 0.181             | 5.535               | 0.022            | -13.968          | 0.000              | 0.125           | 0.262           |
| YoMV / PPV            | 1.299             | 0.770               | 0.245            | 1.389            | 0.863              | 0.734           | 2.299           |
| YoMV / PZSV           | 0.417             | 2.398               | 0.058            | -6.241           | 0.000              | 0.273           | 0.638           |
| YoMV / TAV            | 0.883             | 1.133               | 0.122            | -0.901           | 0.986              | 0.581           | 1.342           |
| YoMV / TuMV           | 1.523             | 0.657               | 0.256            | 2.500            | 0.195              | 0.915           | 2.535           |

|            |       |       |       |        |       |       |       |
|------------|-------|-------|-------|--------|-------|-------|-------|
| YoMV / WMV | 0.591 | 1.692 | 0.090 | -3.466 | 0.012 | 0.373 | 0.936 |
|------------|-------|-------|-------|--------|-------|-------|-------|

**Table A23: Post-hoc analyses showing pairwise comparison of *iL* across virus genera.**

| Genera contrast             | odds.ratio | 1/odds.ratio | std.error | statistic | adj.p.value | Lower CI | Upper CI |
|-----------------------------|------------|--------------|-----------|-----------|-------------|----------|----------|
| Cucumovirus/<br>Anulavirus  | 1.123      | 0.891        | 0.125     | 1.040     | 0.837       | 0.829    | 1.521    |
| Polerovirus/<br>Anulavirus  | 0.224      | 4.468        | 0.053     | -6.326    | 0.000       | 0.117    | 0.427    |
| Polerovirus/<br>Cucumovirus | 0.199      | 5.016        | 0.044     | -7.252    | 0.000       | 0.109    | 0.366    |
| Potyvirus/<br>Anulavirus    | 0.419      | 2.385        | 0.051     | -7.168    | 0.000       | 0.301    | 0.584    |
| Potyvirus/<br>Cucumovirus   | 0.373      | 2.678        | 0.034     | -10.769   | 0.000       | 0.291    | 0.479    |
| Potyvirus/<br>Polerovirus   | 1.873      | 0.534        | 0.426     | 2.758     | 0.046       | 1.007    | 3.486    |
| Tobamovirus/<br>Anulavirus  | 0.393      | 2.543        | 0.054     | -6.747    | 0.000       | 0.27     | 0.574    |
| Tobamovirus/<br>Cucumovirus | 0.350      | 2.855        | 0.040     | -9.208    | 0.000       | 0.257    | 0.478    |
| Tobamovirus/<br>Polerovirus | 1.757      | 0.569        | 0.422     | 2.344     | 0.131       | 0.912    | 3.386    |
| Tobamovirus/<br>Potyvirus   | 0.938      | 1.066        | 0.116     | -0.516    | 0.986       | 0.668    | 1.316    |

**Table A24: Post hoc analyses showing pairwise comparison of *iL* across habitat levels.**

| Habitat contrast | odds.ratio | 1/odds.ratio | std.error | Statistic | adj.p.value | Lower CI | Upper CI |
|------------------|------------|--------------|-----------|-----------|-------------|----------|----------|
| Edge / Crop      | 0.473      | 2.113        | 0.053     | -6.635    | 0.000       | 0.354    | 0.632    |
| Oak / Crop       | 0.216      | 4.619        | 0.026     | -12.744   | 0.000       | 0.159    | 0.295    |
| Oak / Edge       | 0.457      | 2.187        | 0.042     | -8.434    | 0.000       | 0.360    | 0.580    |
| Wasteland / Crop | 0.195      | 5.135        | 0.024     | -13.285   | 0.000       | 0.142    | 0.267    |
| Wasteland / Edge | 0.411      | 2.430        | 0.040     | -9.108    | 0.000       | 0.320    | 0.529    |
| Wasteland / Oak  | 0.900      | 1.112        | 0.095     | -1.005    | 0.746       | 0.687    | 1.179    |

**Table A25: Post hoc analyses showing pairwise comparison of *iL* across transmission modes.**

| Transmission contrast | odds.ratio | 1/odds.ratio | std.error | Statistic | adj.p.value | Lower CI | Upper CI |
|-----------------------|------------|--------------|-----------|-----------|-------------|----------|----------|
| NonPers/contact       | 1.886      | 0.530        | 0.205     | 5.826     | 0.000       | 1.426    | 2.495    |
| Persistent/contact    | 0.565      | 1.769        | 0.136     | -2.375    | 0.082       | 0.305    | 1.048    |
| Persistent/NonPers    | 0.300      | 3.337        | 0.066     | -5.483    | 0.000       | 0.170    | 0.527    |
| Pollen /contact       | 2.561      | 0.390        | 0.355     | 6.777     | 0.000       | 1.793    | 3.658    |
| Pollen/NonPers        | 1.358      | 0.736        | 0.144     | 2.887     | 0.020       | 1.034    | 1.783    |
| Pollen/Persistent     | 4.532      | 0.221        | 1.073     | 6.384     | 0.000       | 2.467    | 8.324    |

NonPers = aphid non-persistent transmission, persistent= aphid persistent transmission

**Table A26: Coverage depth of libraries used for genetic analyses. Coverage = (total read length x total number of reads) / (genome size).**

***CMV coverage depth***

| <b>HTS library</b> | <b>Query matches</b> | <b>RNA1<br/>(3345 nt)</b> | <b>RNA2<br/>(3050 nt)</b> | <b>RNA3<br/>(2216 nt)</b> |
|--------------------|----------------------|---------------------------|---------------------------|---------------------------|
| PV061              | 2416188              | 179936.55                 | 198048.20                 | 20299.70                  |
| PV132              | 593212               | 44177.24                  | 48623.93                  | 4983.89                   |
| PV002              | 391348               | 29144.18                  | 32077.70                  | 3287.93                   |
| PV011              | 288047               | 21451.22                  | 23610.41                  | 2420.04                   |
| PV003              | 187496               | 13963.06                  | 15368.52                  | 1575.26                   |
| PV120              | 3595                 | 267.72                    | 294.67                    | 30.20                     |
| PV136              | 1833                 | 136.51                    | 150.25                    | 15.40                     |
| PV148              | 1696                 | 126.30                    | 139.02                    | 14.25                     |
| PV158              | 1465                 | 109.10                    | 120.08                    | 12.31                     |
| PV014              | 1388                 | 103.37                    | 113.77                    | 11.66                     |
| PV129              | 1286                 | 95.77                     | 105.41                    | 10.80                     |
| PV115              | 1135                 | 84.52                     | 93.03                     | 9.54                      |
| PV144              | 1003                 | 74.69                     | 82.21                     | 8.43                      |
| PV130              | 917                  | 68.29                     | 75.16                     | 7.70                      |
| PV137              | 859                  | 63.97                     | 70.41                     | 7.22                      |
| PV112              | 757                  | 56.37                     | 62.05                     | 6.36                      |
| PV009              | 730                  | 54.36                     | 59.84                     | 6.13                      |
| PV141              | 713                  | 53.10                     | 58.44                     | 5.99                      |
| PV151              | 711                  | 52.95                     | 58.28                     | 5.97                      |

|          |     |       |       |      |
|----------|-----|-------|-------|------|
| PV001    | 626 | 46.62 | 51.31 | 5.26 |
| PV154    | 609 | 45.35 | 49.92 | 5.12 |
| PV146    | 607 | 45.20 | 49.75 | 5.10 |
| PV070    | 432 | 32.17 | 35.41 | 3.63 |
| PV114    | 388 | 28.89 | 31.80 | 3.26 |
| PV110    | 381 | 28.37 | 31.23 | 3.20 |
| PV437    | 337 | 25.10 | 27.62 | 2.83 |
| PV015bgi | 250 | 18.62 | 20.49 | 2.10 |
| PV104    | 225 | 16.76 | 18.44 | 1.89 |
| PV101    | 207 | 15.42 | 16.97 | 1.74 |
| PV021bgi | 186 | 13.85 | 15.25 | 1.56 |
| PV113    | 166 | 12.36 | 13.61 | 1.39 |
| PV103    | 156 | 11.62 | 12.79 | 1.31 |
| PV095    | 142 | 10.57 | 11.64 | 1.19 |
| PV035    | 134 | 9.98  | 10.98 | 1.13 |
| PV106    | 129 | 9.61  | 10.57 | 1.08 |
| PV105    | 123 | 9.16  | 10.08 | 1.03 |

***PZSV coverage depth***

| <b>HTS libraries</b> | <b>Query matches</b> | <b>RNA1<br/>(3383 nt)</b> | <b>RNA2<br/>(2435 nt)</b> | <b>RNA3<br/>(2659 nt)</b> |
|----------------------|----------------------|---------------------------|---------------------------|---------------------------|
| PV219                | 1264272              | 93428.32                  | 129802.05                 | 118867.24                 |
| PV220                | 581355               | 42961.50                  | 59687.37                  | 54659.18                  |
| PV143                | 224342               | 16578.63                  | 23033.06                  | 21092.70                  |
| PV178                | 36053                | 2664.28                   | 3701.54                   | 3389.71                   |
| PV181                | 34975                | 2584.61                   | 3590.86                   | 3288.36                   |
| PV212                | 1214                 | 89.71                     | 124.64                    | 114.14                    |
| PV224                | 924                  | 68.28                     | 94.87                     | 86.87                     |
| PV217                | 878                  | 64.88                     | 90.14                     | 82.55                     |
| PV213                | 841                  | 62.15                     | 86.34                     | 79.07                     |
| PV215                | 799                  | 59.05                     | 82.03                     | 75.12                     |

|       |     |       |       |       |
|-------|-----|-------|-------|-------|
| PV223 | 503 | 37.17 | 51.64 | 47.29 |
| PV216 | 351 | 25.94 | 36.04 | 33.00 |
| PV214 | 183 | 13.52 | 18.79 | 17.21 |
| PV211 | 155 | 11.45 | 15.91 | 14.57 |
| PV199 | 154 | 11.38 | 15.81 | 14.48 |
| PV208 | 135 | 9.98  | 13.86 | 12.69 |
| PV209 | 129 | 9.53  | 13.24 | 12.13 |
| PV233 | 128 | 9.46  | 13.14 | 12.03 |
| PV237 | 124 | 9.16  | 12.73 | 11.66 |
| PV206 | 116 | 8.57  | 11.91 | 10.91 |
| PV231 | 116 | 8.57  | 11.91 | 10.91 |
| PV234 | 114 | 8.42  | 11.70 | 10.72 |
| PV207 | 113 | 8.35  | 11.60 | 10.62 |
| PV197 | 109 | 8.05  | 11.19 | 10.25 |
| PV218 | 109 | 8.05  | 11.19 | 10.25 |
| PV225 | 109 | 8.05  | 11.19 | 10.25 |
| PV230 | 109 | 8.05  | 11.19 | 10.25 |
| PV202 | 108 | 7.98  | 11.09 | 10.15 |
| PV205 | 107 | 7.91  | 10.99 | 10.06 |
| PV210 | 106 | 7.83  | 10.88 | 9.97  |
| PV227 | 104 | 7.69  | 10.68 | 9.78  |
| PV235 | 102 | 7.54  | 10.47 | 9.59  |
| PV221 | 101 | 7.46  | 10.37 | 9.50  |
| PV201 | 100 | 7.39  | 10.27 | 9.40  |
| PV200 | 95  | 7.02  | 9.75  | 8.93  |
| PV229 | 94  | 6.95  | 9.65  | 8.84  |
| PV226 | 87  | 6.43  | 8.93  | 8.18  |
| PV198 | 86  | 6.36  | 8.83  | 8.09  |
| PV204 | 84  | 6.21  | 8.62  | 7.90  |

---

***TAV coverage depth***

| HTS library | Query matches | RNA1<br>(3410 nt) | RNA2<br>(3074 nt) | RNA3<br>(2386 nt) |
|-------------|---------------|-------------------|-------------------|-------------------|
| PV120       | 844           | 61.88             | 68.64             | 88.43             |
| PV232       | 816           | 59.82             | 66.36             | 85.50             |
| PV155       | 648           | 47.51             | 52.70             | 67.90             |
| PV151       | 539           | 39.52             | 43.84             | 56.48             |
| PV158       | 396           | 29.03             | 32.21             | 41.49             |
| PV136       | 335           | 24.56             | 27.24             | 35.10             |
| PV148       | 326           | 23.90             | 26.51             | 34.16             |
| PV129       | 314           | 23.02             | 25.54             | 32.90             |
| PV115       | 294           | 21.55             | 23.91             | 30.80             |
| PV144       | 263           | 19.28             | 21.39             | 27.56             |
| PV132       | 174           | 12.76             | 14.15             | 18.23             |
| PV137       | 169           | 12.39             | 13.74             | 17.71             |
| PV156       | 159           | 11.66             | 12.93             | 16.66             |
| PV145       | 148           | 10.85             | 12.04             | 15.51             |
| PV141       | 142           | 10.41             | 11.55             | 14.88             |
| PV131       | 119           | 8.72              | 9.68              | 12.47             |
| PV119       | 116           | 8.50              | 9.43              | 12.15             |
| PV154       | 106           | 7.77              | 8.62              | 11.11             |
| PV146       | 106           | 7.77              | 8.62              | 11.11             |
| PV152       | 101           | 7.40              | 8.21              | 10.58             |

**Table A27:** Bayesian Tip-association Significance tests (BaTS) of TAV populations at Edge for the strength of site or host.

|  | Statistic | Obs.<br>mean | L. 95%<br>CI | U. 95%<br>CI | Null<br>mean | L. 95%<br>CI | U. 95%<br>CI | <i>p</i> -value |
|--|-----------|--------------|--------------|--------------|--------------|--------------|--------------|-----------------|
|  | TAV RNA1  |              |              |              |              |              |              |                 |

|                 |    |       |      |      |       |       |       |       |
|-----------------|----|-------|------|------|-------|-------|-------|-------|
| Site effect     | AI | 1.30  | 0.76 | 1.83 | 1.35  | 1.17  | 1.50  | 0.290 |
|                 | PS | 10.74 | 9    | 12   | 10.83 | 9.84  | 11.73 | 0.610 |
| Host effect     | AI | 1.67  | 1.16 | 2.14 | 1.70  | 1.61  | 1.74  | 0.200 |
|                 | PS | 15.30 | 14   | 16   | 15.56 | 14.94 | 16.30 | 0.070 |
| <b>TAV RNA2</b> |    |       |      |      |       |       |       |       |
| Site effect     | AI | 1.29  | 0.71 | 1.83 | 1.33  | 1.10  | 1.48  | 0.310 |
|                 | PS | 10.57 | 9    | 12   | 10.80 | 9.75  | 11.67 | 0.620 |
| Host effect     | AI | 1.73  | 1.26 | 2.17 | 1.70  | 1.54  | 1.76  | 0.770 |
|                 | PS | 15.39 | 14   | 16   | 15.45 | 14.63 | 16.00 | 0.260 |
| <b>TAV RNA3</b> |    |       |      |      |       |       |       |       |
| Site effect     | AI | 1.44  | 0.88 | 1.96 | 1.37  | 1.24  | 1.47  | 0.910 |
|                 | PS | 10.82 | 9    | 12   | 10.76 | 9.91  | 11.43 | 0.680 |
| Host effect     | AI | 1.76  | 1.27 | 2.20 | 1.75  | 1.68  | 1.79  | 0.740 |
|                 | PS | 15.57 | 15   | 16   | 15.55 | 14.95 | 16.09 | 0.930 |

**Table A28:** Bayesian Tip-association Significance tests (BaTS) of PZSV populations at Edge for the strength of site or host.

|                  | Statistic | Obs.<br>mean | L. 95%<br>CI | U. 95%<br>CI | Null<br>mean | L. 95%<br>CI | U. 95%<br>CI | <i>p</i> -value |
|------------------|-----------|--------------|--------------|--------------|--------------|--------------|--------------|-----------------|
| <b>PZSV RNA1</b> |           |              |              |              |              |              |              |                 |
| Site effect      | AI        | 0.90         | 0.41         | 1.38         | 0.98         | 0.90         | 1.04         | 0.160           |
|                  | PS        | 5.83         | 5            | 6            | 6.37         | 5.82         | 6.83         | 0.480           |
| Host effect      | AI        | 1.53         | 1.03         | 1.96         | 1.54         | 1.53         | 1.57         | 0.830           |
|                  | PS        | 12.61        | 12           | 13           | 12.78        | 12.59        | 13.36        | 0.830           |
| <b>PZSV RNA2</b> |           |              |              |              |              |              |              |                 |
| Site effect      | AI        | 2.73         | 1.86         | 3.55         | 2.73         | 2.63         | 2.80         | 0.510           |
|                  | PS        | 21.99        | 19           | 24           | 22.10        | 21.07        | 22.73        | 0.350           |
| Host effect      | AI        | 3.57         | 2.78         | 4.32         | 3.54         | 3.50         | 3.57         | 1               |

|             |                  |       |      |      |       |       |       |              |
|-------------|------------------|-------|------|------|-------|-------|-------|--------------|
|             | PS               | 33.21 | 32   | 34   | 32.97 | 32.38 | 33.68 | 0.580        |
|             | <b>PZSV RNA3</b> |       |      |      |       |       |       |              |
| Site effect | AI               | 2.37  | 1.68 | 3.05 | 2.75  | 2.39  | 3.07  | 0.050        |
|             | PS               | 19.97 | 18   | 22   | 22.04 | 20.18 | 23.39 | <b>0.040</b> |
| Host effect | AI               | 3.26  | 2.62 | 3.87 | 3.55  | 3.39  | 3.64  | <b>0.030</b> |
|             | PS               | 31.89 | 31   | 33   | 33.08 | 32.22 | 33.82 | <b>0.040</b> |