

INEXTVIR: High throughput sequencing of the virome of agricultural crops and ecosystems across Europe

Kaoutar Daoud Hiri, Johan Rollin, Sukhorukov Grigorii, Macha Nikolski,

Sébastien Massart, Tomaž Curk, Matjaž Hren

▶ To cite this version:

Kaoutar Daoud Hiri, Johan Rollin, Sukhorukov Grigorii, Macha Nikolski, Sébastien Massart, et al.. INEXTVIR: High throughput sequencing of the virome of agricultural crops and ecosystems across Europe. ESCS-ECCB 2020, Sep 2020, Barcelona, Spain. hal-03271520

HAL Id: hal-03271520 https://hal.science/hal-03271520v1

Submitted on 25 Jun 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



INEXTVIR: High throughput sequencing of the virome of agricultural crops and ecosystems across Europe

MSCA ITN -Innovative Network for Next Generation Training and Sequencing of Virome



Kaoutar Daoud Hiri^{a,b}, Johan Rollin^{c,d}, Sukhorukov Grigorii^{e,f}, Nikolski Macha^{e,f}, Sébastien Massart^c, Tomaž Curk^b, Matjaž Hren^g

^aJožef Stefan International Postgraduate School, Ljubljana, Slovenia ^bUniversity of Ljubljana, Faculty Of Computer And Information Science, Ljubljana, Slovenia ^cPlant Pathology Laboratory, TERRA, Gembloux Agro-Bio Tech, University of Liège, Gembloux, Belgium ^dDNAVision, Gosselies, Belgium

^eUniversity of Bordeaux, CNRS, IBGC, UMR 5095, Bordeaux, France ^fUniversity of Bordeaux, Bordeaux Bioinformatics Center, Bordeaux, France ^gBioSistemika, Ljubljana, Slovenia



BACKGROUND

Plant Viruses cause 50% of the emerging plant diseases and pose an important threat to many agricultural crops worldwide, they are responsible for production losses estimated at € 15 to 45 billion per year.

INEXTVIR proposes a transdisciplinary research and training approach combining applied and fundamental plant virology, ecology, NGS platform development, computer science (bioinformatics), agronomical, socio-economic and regulatory sciences.

<u>Goal</u>: Better understand, control and exploit the virome of selected agricultural crops and ecosystems across Europe.

\$

<u>Benefits</u>: Political, economic, industrial, social and environmental.

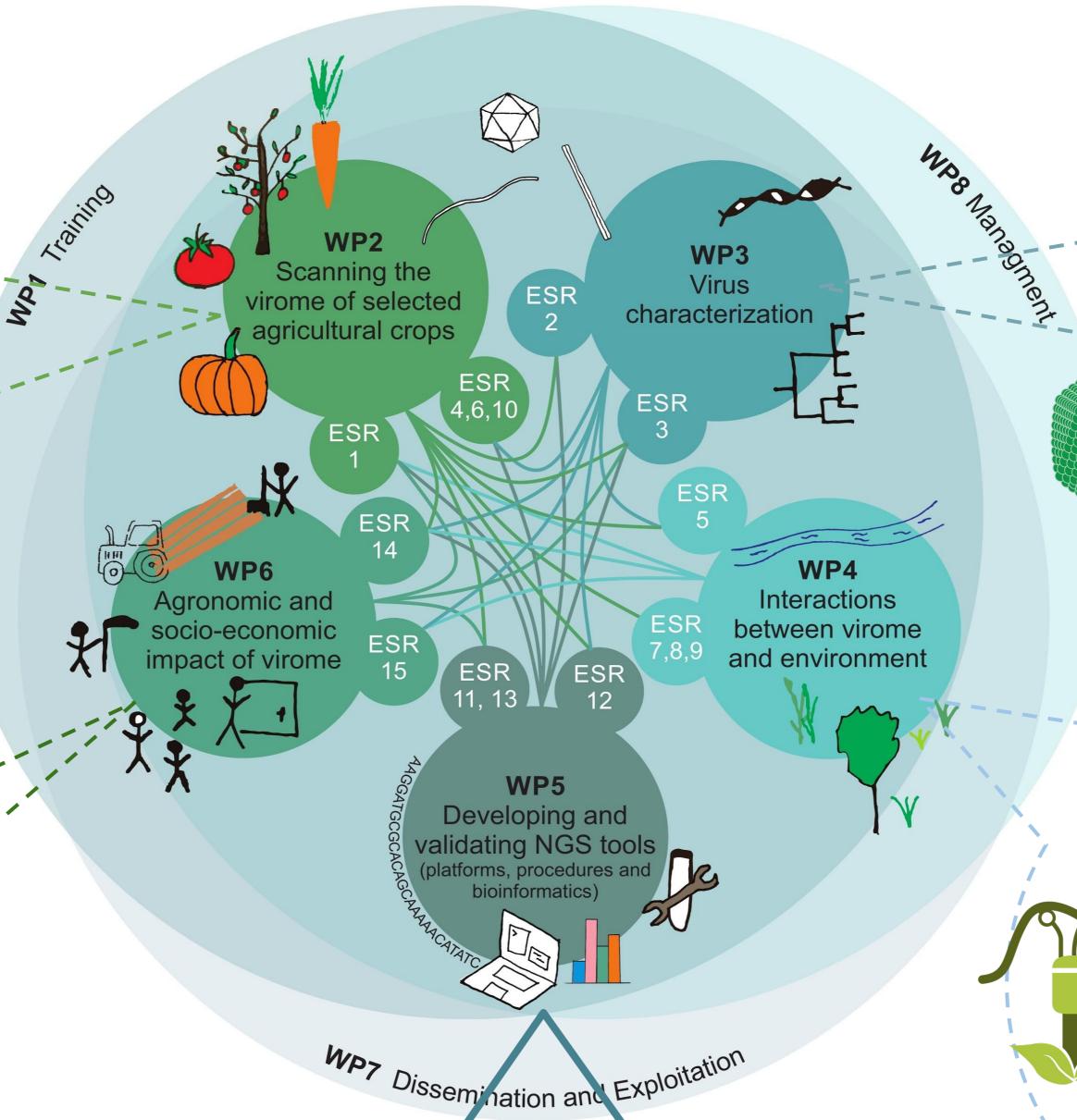
Scanning the virome of selected agricultural crops

WP2

ESR 1: Investigating the European field vegetable virome, University of Newcastle Upon Tyne, United Kingdom ESR 4: Characterisation of tomato virome, National Institute of Biology, Slovenia **ESR 6**: HTS-based viral indexing in Prunus species: virome characterization and validation of diagnostic approaches, *Institut* National de la Recherche Agronomique, France **ESR 10**: The cucurbits and lettuce viromes, ABIOPEP SL, Spain

Agronomic and socio-economic impact of virome

WP6



WP3 Virus Characterization

ESR 2: Virome characterization of European germplasm resources of pome fruits for future breeding and risk analysis, University of Liege, Belgium **ESR 3:** Survey and evaluation of the cultural practices impact on the virome of vegetable production systems and effect on sustainability, University of Liege, Belgium

WP4

Interactions between virome and environment

ESR 5: Water as an important indicator of circulating plant viruses, National Institute of Biology, Slovenia **ESR 7:** Ecological analysis of the plant virome structure and of reciprocal transfers between cultivated and wild host populations, *Institut* National de la Recherche Agronomique, France

ESR 14: Agronomic and socio-economic impact of virome, FERA Science Limited, United Kingdom **ESR 15:** Risk-benefit communication and Knowledge exchange platform, Iniversity of Newcastle Upon Tyne, United 🍃 Kingdom

WP5

Developing and validating NGS tools

ESR 11: Development and validation of innovative HTS tools for the fast and efficient detection of virus and virus-like agents in the selected crops, DNAVision, Belgium **ESR 12:** Bioinformatics pipelines development for virome classification, University of Bordeaux, France **ESR 13:** Development and integration of decision-making tools in data analysis procedures for HTS-based plant virus diagnostics BioSistemika, Slovenia

ESR 8: Effect of habitat biodiversity on the virome of wild plants and crops, Universidad Politécnica de Madrid, Spain **ESR 9:** Effect of habitat biodiversity on virus host range evolution, Universidad Politécnica de Madrid, Spain

<u>Contribution to bioinformatic and software development segments of the project</u>

Developing of own bioinformatic methods

Finding alternative methods for plant virus detection

- Viral identification in HTS data from agricultural crops and ecosystems
- Generation of artificial HTS datasets for testing purposes
- Assessment of cross-contamination in sequencing dataset
- Obtain validation criteria on plant virus identification

<u>Pipeline design</u>	
<u>Pipeline testing</u>	<u>Decision making</u>
Identify the most relevant combination of methods and tools	Propose the most appropriate pipeline depending on the use case for the user

Virus related tools benchmarking

Identify the most relevant HTS methods and tools

- Identify relevant use-cases in the project consortium (e.g. virus detection, virus classification ...)
- Identify the tools that seem relevant for the selected use-cases in our neutral study Use Docker images of tools (either existing or created) for testing and results dissemination
- Select datasets (artificial or natural) that meet test criteria
- Share the benchmark result highlighting tools performance

- Identify relevant steps in HTS data analysis (e.g. reads cleaning, identification ...)
- Create pipelines based on Docker images of the tools and test them
- Compare the pipelines performance 3)
- The most suitable pipelines for each use case will be 4) shared with project consortium in Docker images
- Document the pipeline for users and developers 5)
- Use performance comparison to make a recommendation
- Store relative performance metrics for each pipeline a)
- Access those metrics to be able to make a decision b)
- Update the performance metrics based on new tests and tools
- 2) Explain the recommendation to the users so that they make a final decision

Future consideration: Automated pipeline recommendation using a Machine Learning approach



Partner organizations

INEXTVIR is implemented by a European Consortium of universities, research institutions and companies in **Belgium**, France, Spain, Slovenia and the UK.

