



DAFNE servers tutorial

System admin: Silvia Turco, PhD*

System admin collaborator: Luca Rossini †

System Coordinator: Prof. Angelo Mazzaglia‡

September 16, 2020

*silvia.turco@unitus.it

†luca.rossini@unitus.it

‡angmazza@unitus.it

Contents

1	Introduction	3
2	Platform overview	3
3	Get an account	3
4	Platform policy	4
5	Further requirements	4
6	Troubleshooting	5
7	Tools already installed	5
8	Server connection from Windows	6
9	Server connection from Linux or Mac	8
10	Transfer files to and from the nodes (Windows users)	8
11	Transfer files to and from the nodes (Linux and Mac users)	10
12	Home sweet home	10
	12.1 Walking around	11
	12.2 Basic commands	11
	12.3 Few more hints	13
13	R studio server	13
14	Acknowledging the DAFNE server platform	14
15	Publications list	14

1 Introduction

In the frame of the SAFE-Med project, the Department of Agriculture and Forest Sciences (DAFNE), awarded as "Center of Excellence", has been equipped with a servers platform at the disposal of Professors, researchers and PhD students. The servers can be used to perform complicated and time consuming data analysis in order to answer questions coming from different scientific fields, like molecular biology, microbiology, livestock genomics, mathematical modelling, physics and so on. To have a better idea of what our servers can do, you can refer to the latest publications from the server users or ask the system administrator.

2 Platform overview

The Dafne servers platform consists of:

- **Odin**, the main HPE StoreEasy of 16TB of SAS storage.
- **Marvin** and **Bender**, the two server nodes with 96 total cores (4x Intel Xeon Gold 5118 12 cores @ 2.3 GHz), 1024 GByte of RAM memory (2x 512 GByte RAM 2.67 GHz) and 1.5 TB SAS HDD.

Each computing node runs Centos Linux 7.7 operating system, which is a specific free and Open source version of Red Hat Enterprise Linux.

The two nodes servers are exactly equivalent and defined by two different IP address. They are both connected and synchronized with Odin, the main storage centre, where your home folder is located.

3 Get an account

To get access to the platform you need a valid Unitus account and/or you should be a DAFNE member in need of computational power.

Please write to the system administrator to get your credentials, specifying which research group you are part of, which kind of jobs you want to run on the nodes and if you need to use a software that is not already installed.

Once your account is activated, you'll get an email with your username, your password, the domain name and the IP address of one of the two nodes. Please keep them private and safe.

You will also be added to the server mailing list to be updated regarding any news.

4 Platform policy

For the safety of the system and of your own private data, there are some rules you need to read before getting started and to keep in mind while you are using it.

The proper account request to the system administrator would be taken as official acceptance of the platform policy.

- Even though your account would be part of your research group, you should not share it with anyone. Each member of your group can get his own account.
- You can share your data with other members of your research group by placing them in the group home folder. To share them with someone external to your research group please contact the system admin.
- The platform is meant for scientific research. Please do not use it for anything else and do not use it to store your private data.
- Be aware that nobody other than yourself AND the system admin can have access to your home folder. This would grant your privacy and safety.
- Even if it sounds like a magical cloud in the sky, the system is not 100% out of danger. Please keep a copy of your data anywhere else as backup.
- You are kindly invited to schedule your time and resources consuming jobs during night time or weekends. Leave space for the other users.

5 Further requirements

- **Linux knowledge**

Since the nodes are running Centos, you need to have at least basics knowledge of Linux command lines and hierarchical structure. You can find a short introduction later in this

document but you are strongly invited to get more familiar with them by following some online tutorials like **Linux Command**, **Linux Tutorial**. There are also really handy **short list** with all the basic commands that you may want to print.

- **VPN connection**

Besides your credentials, to access the server you need to be connected to the Unitus network. So, If you are trying to connect from outside the university (i.e. from home), you need a vpn connection and thus, a vpn client. For this you can follow the **instructions** provided by the Unitus IT office. Note that if you have Linux, you can follow the same procedure as Mac.

- **SSH client**

If you are running Windows, once you are connected to the Unitus network, you need to log on to the platform using an SSH client to reach one of the server nodes and this would give you command-line access. However, a typical Windows environment does not come with pre-installed software to connect and run command-line executables on a HPC. If you are running Windows, you can use an SSH client like **Putty** or get a Virtual Machine running a Linux distro.

6 Troubleshooting

Whenever you are facing a problem...DON'T PANIC and remember, GIYF (Google is your friend). If you are having problem with PUTTY, check that your version is updated, your connection is working and that your options are correct. If the terminal gives you an error, double check that the command you wrote is correct, without any typing mistakes and in the proper working directory (i.e. the files you are working with should be there). If it's a system or tool related error, if some libraries are missing, please contact the system admin giving specific details about it and by reporting the error text you got.

7 Tools already installed

Abacas v1.3.1, AlighGraph, Artemis 18.1.0 Bam-readcount v0.8.0, Bandage v0.8.1, Bcftools v1.10-6, Bedtools v2.29.1, Blat v. 36x4, Bowtie2, Bwa v0.7.17-r1188, Cd-hit v4.8.1, FastQC v0.11.8,

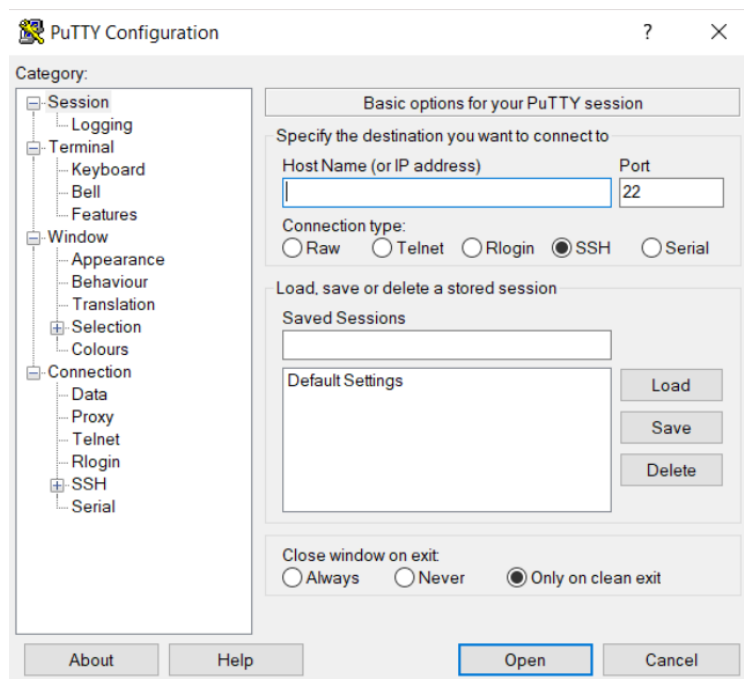
Hisat2 v2.1, IGV v2.7.2, Interproscan v5.39-77, Mauve v2.4.0, ??Mega Miniasm v0.3, Minimap2 v2.17, ??Mothur v1.43.0 MUMmer v3.23, ??Muscle v3.8.31, NCBI Blast v2.9 (with efetch and edirect), plat, Prokka v1.14.5, Qualimap v2.2.1, R 3.6, Racon v1.4.3, RaxML v8.0, Roary v3.8, Root v6.18/04, Samtools v1.9, Sickle v1.33, SPAdes v3.13.0, ??Stringtie, Trimmomatic v0.39, Velvet v1.2.10, Docker,

Please note that this list may not be updated. Ask your sysadmin for other tools.

8 Server connection from Windows

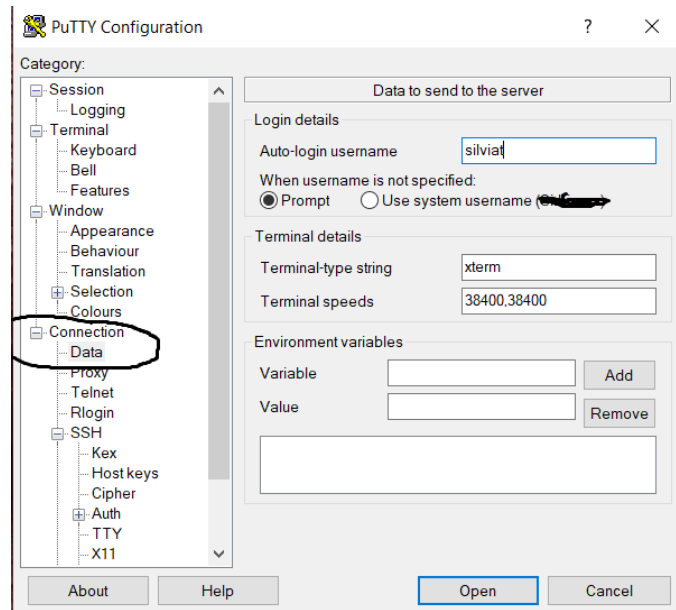
As already mention, you need to be connected to the Unitus network and only after you can connect to the servers. To do so, download PuTTY and start the executable putty.exe in your directory C:Files (x86). The configuration screen like the one below will pop up.

Within the category “Session”, in the field HostName, add the server IP address you received by email during the activation of your account.

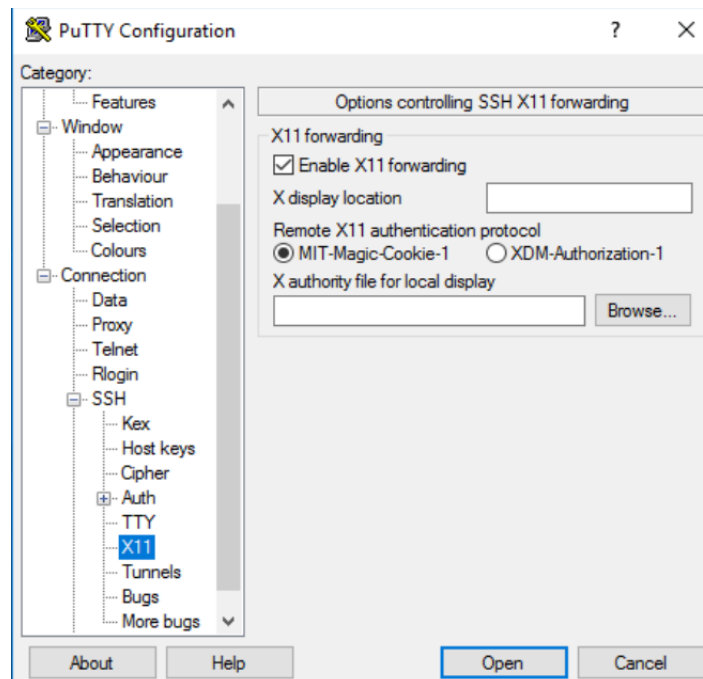


In the “Connection/Data” category, you can already write your username. This step is not manda-

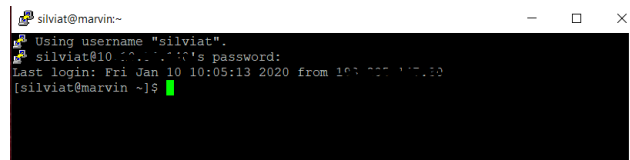
tory, you can type your username later on too.



If you need to use graphical tools like Mauve, you need to enable the X-server from Connection/SSH/X11 category as follow:



Then click “Open”. The terminal window will pop up asking for your username (if you skipped the Connection/Data step) and your password. If the connection was successful, you should see the prompt with your username, followed by @ and the node name you are connected to (either Marvin or Bender, depends on which IP you used).

A terminal window titled 'silviat@marvin:-' showing the process of logging in. The prompt is 'Using username "silviat":', followed by 'silviat@10.0.0.1:~\$' and a password prompt. The login is successful, showing 'Last login: Fri Jan 10 10:05:13 2020 from 10.0.0.10' and the prompt '[silviat@marvin ~]\$'.

You are now in your home folder (), ready to start your job.

9 Server connection from Linux or Mac

Run the following commands in a terminal, substituting username and serverIPAddress with the login name you received by email.

```
$ ssh username@serverIPAddress (or ssh username@domain_name)
```

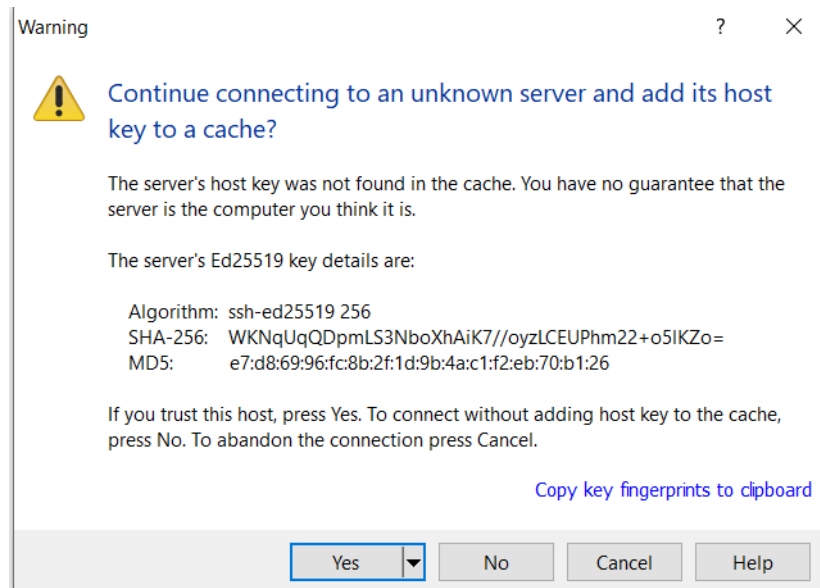
If you need to run a graphical server, be sure that you have installed XQuartz (for Mac) or Xterm (Linux) for example. Then just add -X to the previous code:

```
$ ssh -X username@serverIPAddress
```

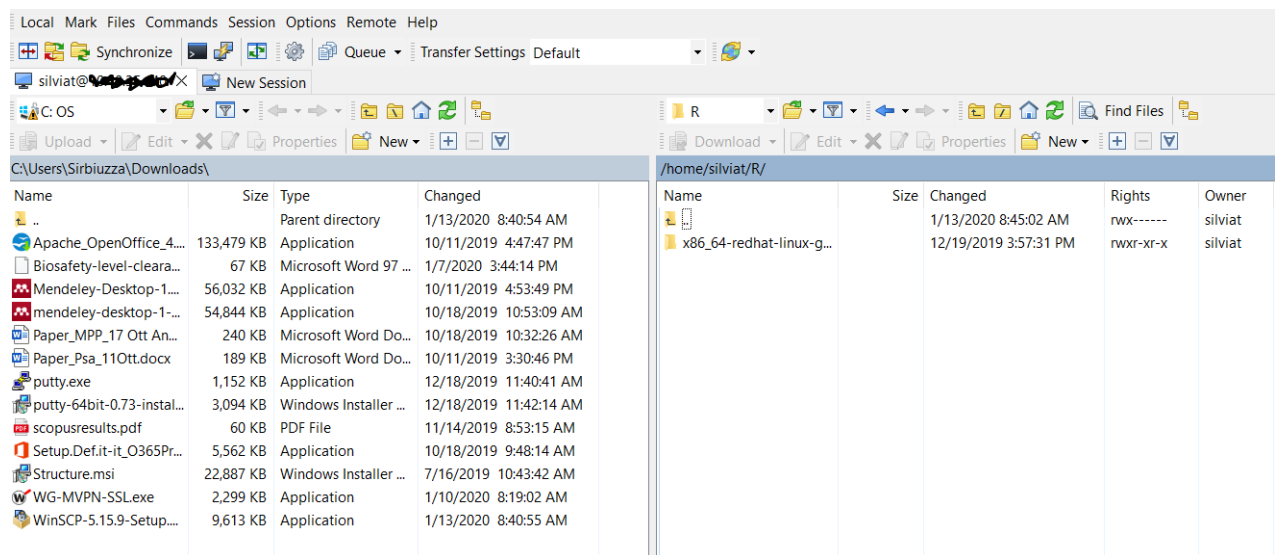
10 Transfer files to and from the nodes (Windows users)

To transfer your files to and from the nodes, you need to download WinSCP. Once you have installed it, you need to connect again to the platform, using the IP address, port 22, your username and password.

The first time connection you may get a warning message like the one below. Just click Yes and continue.



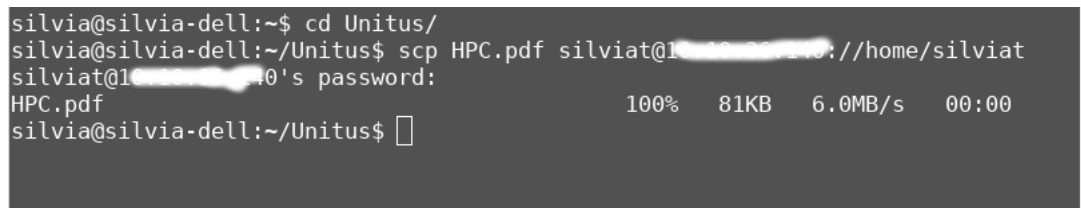
In the next window, you'll see on the left your local folder and on the right side your home folder on the platform. You can now just drag and drop your files.



11 Transfer files to and from the nodes (Linux and Mac users)

The scp command can be used to securely transferring files between a local host (i.e. your computer) and a remote host (the HPC). Open a terminal and move to the folder where the file to be transferred is located (or write the complete path). Then type the command scp followed by the filename and the destination path (if you are familiar with the cp or mv command, you see that it's the same idea).

```
$ scp filename username@IPAddress://home/userfolder
```



```
silvia@silvia-dell:~$ cd Unitus/  
silvia@silvia-dell:~/Unitus$ scp HPC.pdf silviat@10.10.10.10://home/silviat  
silviat@10.10.10.10's password:  
HPC.pdf          100%  81KB  6.0MB/s   00:00  
silvia@silvia-dell:~/Unitus$
```

12 Home sweet home

Folders on Linux are arranged in a hierarchical structure. Starting for / root, the main one we will find all the other sub-folders. As shown in the picture below, /home is where the user folder are located and where you'll find yours. Note that this folder is private and can't be accessed by others than the user and the system admin. Furthermore, /home is shared between Odin and the nodes, meaning that it does not really matter to which one you are connected, you'll always find your files there.

As the name suggested, the /share folder is also shared between Odin, Marvin and Bender and it's where all the tools have been installed, in a way that they are reachable from the nodes and readable by anyone who has access. However, some tools require libraries which need to be installed locally (meaning in each node). If you encounter any missing library in one of the two node and thus, your job cannot be run, please contact the admin. Inside the /Group folder there are the folders for each research group who asked for login access, i.e. Pathology.

If you are part of a Group and you want to share you files with your colleagues, you will find a

symbolic link of the Group folder into your own home. You all have the permission to modify the files inside that folder (if not, please contact the admin).

12.1 Walking around

You can now use some basics commands to create new folders, move around, create files or start your analysis. Keep in mind that, as well as walking around the city, when you have to move between folders, you need the address, the so called path. In Linux environment there are two path:

- The absolute path, the complete one that always starts at the top of the tree (at the root directory), for example: `/home/silviat/Unitus/tutorial.pdf`
- The relative path, is the path to the destination file from your current position and it does not start with a `“/”`. For example, `Unitus/tutorial.pdf`

This means that if I want to open or use a file that is already in the current folder, I can call it already. If the file is in another folder, I need to provide the proper path on my commands.

12.2 Basic commands

The common command structure is:

```
command options parameters
```

Usually the command is a shortcut of its function, followed by optional or mandatory parameters.

- For example, to see the contents of the current directory, type `ls` (list) at the prompt.

```
$ ls
```

- You can also add more options, for example `ls -l` to lists your files in 'long format', which contains lots of useful information, i.e. the exact size of the file, who owns the file and who has the right to look at it, and when it was last modified.

```
$ ls -l
```

- To lists all files, including the ones whose filenames begin in a dot, which you do not always want to see, use **ls -a**.

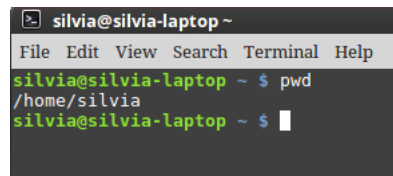
```
$ ls -a
```

- You can create a new folder by using the make directory **mkdir** command:

```
$ mkdir foldername
```

- To know your current working directory, print it with **pwd**

```
$ pwd
```

A terminal window titled 'silvia@silvia-laptop ~' with a menu bar containing 'File Edit View Search Terminal Help'. The terminal shows the command 'silvia@silvia-laptop ~ \$ pwd' followed by the output '/home/silvia' on the next line. The prompt 'silvia@silvia-laptop ~ \$' is visible on the third line.

- To change directory use "cd". The two dots means "one folder up". Note that you can access Unitus/ only if this folder is a subfolder of your current working directory, otherwise you have to provide the complete path.

```
$ cd Unitus/
```

- To remove a file, use **rm** for a file or **rmdir** to remove a directory. **Remember that there is no trash bin in Linux! Everything you remove is going to be completely deleted and cannot be restored.**

```
$ rm file
```

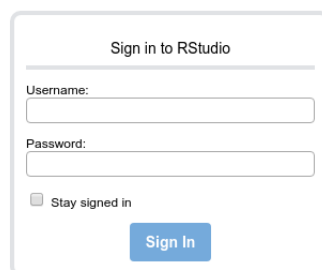
For further commands, please refer to the guide mentioned in section 4 or browse the web.

12.3 Few more hints

- `cd ../` will take you one directory level up (from `"/home/silviat/Unitus"` to `"/home/silviat"`)
- The sign `./` means current directory (`"cd Unitus"` is the same as `"cd ./Unitus"`).
- Typing `cd` anywhere will get you to your home directory.
- You can use `~` in the path to substitute for your home directory (e.g. `"cd ~/Unitus"` is the same as `cd /home/silviat/Unitus`).
- Whenever you have a long command to write, for example a long path, file name, etc, TAB will complete the name so you don't have to type everything!
- If you get an error message, it is probably write. Everything has to be typed perfectly, without any mistakes. Remember that the terminal is case sensitive.

13 R studio server

R is a language and environment for statistical computing and graphics. You can use it on our system my command line or you can use the much more easier R studio server, which it has been installed on **Bender**. To do so, once you are connected to the server, open your browser and type on the address bar the server IP address you used for the connection, followed by `:8787` A window will pop up asking your server ID and password. After the login, you will get the usual Rstudio, with your home folder showed on the fourth panel (down right).



The image shows a login form titled "Sign in to RStudio". It contains two input fields: "Username:" and "Password:". Below the password field is a checkbox labeled "Stay signed in". At the bottom right of the form is a blue button labeled "Sign In".

14 Acknowledging the DAFNE server platform

If you have been using the DAFNE servers platform to produce results you are going to publish or present in either a poster or public seminar, you are invited to mention it in the Materials and methods section: "Calculations were performed using the DAFNE server of Università degli Studi della Tuscia". In the acknowledgments and funding you may mention: "The research was carried out in the frame of the MIUR (Ministry for education, University and Research) initiative "Department of Excellence" (Law 232/2016). Please also let us know whenever you get a new publication!

15 Publications list

- Rossini, L., Contarini, M., Severini, M., Speranza, S., 2020. Reformulation of the Distributed Delay Model to describe insect pest populations using count variables. *Ecological Modelling* 436, 109286.
<https://doi.org/10.1016/j.ecolmodel.2020.109286>
- Rossini, L., Contarini, M., Severini, M., Talano, D., Speranza, S., 2020. A modelling approach to describe the *Anthonomus eugenii* (Coleoptera: Curculionidae) life cycle in plant protection: A priori and a posteriori analysis. *Florida Entomologist* 103, 259–263.
<https://doi.org/10.1653/024.103.0217>
- Rossini, L., Contarini, M., Speranza, S., 2020. A novel version of the Von Foerster equation to describe poikilothermic organisms including physiological age and reproduction rate. *Ricerche di Matematica*. <https://doi.org/10.1007/s11587-020-00489-6>
- Rossini, L., Severini, M., Contarini, M., Speranza, S., 2020. EntoSim, a ROOT-based simulator to forecast insects' life cycle: Description and application in the case of *Lobesia botrana*. *Crop Protection* 129, 105024.
<https://doi.org/10.1016/j.cropro.2019.105024>
- Rossini, L., Speranza, S., Contarini, M., 2020. Distributed Delay Model and Von Foerster's equation: Different points of view to describe insects' life cycles with chronological age and

physiological time. *Ecological Informatics* 59, 101117.

<https://doi.org/10.1016/j.ecoinf.2020.101117>

- Rahi, Y.J., Turco, S., Taratufolo, M.C., Tati, M., Cerboneschi, M., Tegli, S., Valentini, F., D'Onghia, A., Iacobellis, N.S., Balestra, G.M., Mazzaglia, A., 2020. Genetic diversity and population structure of *Pseudomonas savastanoi*, an endemic pathogen of the Mediterranean area, revealed up to strain level by the MLVA assay. *J. Plant Pathol.*

<https://doi.org/10.1007/s42161-020-00589-0>

- Mazzaglia, A., Rahi, Y.J., Taratufolo, M.C., Tati, M., Turco, S., Ciarroni, S., Tagliavento, V., Valentini, F., D'Onghia, A.M., Balestra, G.M., 2020. A new inclusive MLVA assay to investigate genetic variability of *Xylella fastidiosa* with a specific focus on the Apulian outbreak in Italy. *Sci. Rep.* 10, 1–11.

<https://doi.org/10.1038/s41598-020-68072-5>