SHORT TERM SCIENTIFIC MISSION (STSM) SCIENTIFIC REPORT

This report is submitted for approval by the STSM applicant to the STSM coordinator.

Action number: CA16107
STSM title: MALDI-TOF MS application for acurate identification and phylogenetic analysis of *Xanthomonas arboricola* pathovars
STSM start and end date: 02/12/2018 to 08/12/2018
Grantee name: Jeroen van de Bilt

PURPOSE OF THE STSM

(max. 200 words)

Several bacterial diseases in different geographic locations and host plants are attributed to bacteria belonging to the *Xanthomonas arboricola* complex. The species has been divided into seven pathovars, namely: *X. arboricola* pv. *pruni* (Xapr), *X. arboricola* pv. *corylina, X. arboricola pv. juglandis, X.arboricola* pv. *poinsetticola, X.arboricola* pv. *populi, X.arboricola* pv. *celebensis* and *X. arboricola* pv. *fragariae*. The pathovars *pruni* and *corylina* are listed as quarantine organism (EU Directive 2000/29/CE, EPPO A2 list).

Matrix-assisted laser desorption/ionization and time-of-flight mass-spectrometry (MALDI-TOF MS) is an emerging approach for routine identification of plant pathogens but still very restricted for bacteria relevant in the phytosanitary field.

The purpose of this STSM was to analyse a comprehensive collection of strains related to the *Xanthomonas arboricola* complex with MALDI-TOF MS and to obtain information to develop accurate identification and discrimination among the *Xanthomonas arboricola* pathovars by using MALDI-TOF MS technique.

The selected isolates, 64 in total, cover a significant part of the genetic diversity of the *Xanthomonas arboricola* complex as they are obtained from different orchards and fruit trees in various geographical regions in the Netherlands and in other European countries. A special focus was made on the *X. arboricola* pv. *pruni* representatives from the cherry laurel.

DESCRIPTION OF WORK CARRIED OUT DURING THE STSM

(max. 500 words)
After an introduction to the laboratory and the working rules of the host institution ZHAW, the preparation of NYGA media to grow the strains on and the inoculation of the strains for 48 hour growth (64 strains, 2 replicates, one on Monday, one on Wednesday) was carried

out.
The Shimadzu MALDI-TOF MS system of the host institution was introduced and MALDI-TOF MS data files were prepared (sample identification and comments). After an introduction on spotting of bacterial colonies from agar plates (i.e. direct smear), a first test was performed on Tuesday with *Escherichia coli* DH5α and some unknown bacterial samples.
On Wednesday and Friday, spotting of all selected strains on MALDI-TOF MS targets (64 strains in fourfold) was carried out. The system was externally calibrated using a reference strain of *Escherichia coli* DH5α. On Wednesday some issues with the calibration and some peak shifts were noticed in the routine analyses of the sample strains.
After satisfactory, but not ideal, calibration, the spectra from the 64 selected isolates were created and manually judged on quality (spectral intensity >50 mV). These spectra were matched with a commercially available database for the Shimadzu MALDI-TOF MS system (Spectral ARchive And Microbial Identification System, Anagnostec GmbH, "SARAMIS"). This database contains, so called 'super spectra', that were specially generated for identification purposes. But, as mentioned in the plan of this STSM, this kind of system specific MALDI-TOF MS databases are poorly developed for plant pathogenic bacteria and do not contain the 'super spectra' for *Xanthomonas arboricola* (this applies thus also to the database of the Home institution NVWA). With "Saramis" it is also possible to match between in-house created spectra from previous samples.

On Thursday nine reference strains (type strains and pathovar reference strains), of which some were already included in the 64 selected strains, were spotted with a formic acid extraction method and analysed with MALDI-TOF MS.

In the last years ZHAW and Mabritec AG developed a new database derived from the publicly available genome based sequence data, called PAPMID (Putatively Assigned Protein Masses for Identification), which is independent of the MALDI-TOF MS instrument and covers more than 3.800 bacterial species and does include a wide range of plant pathogens. All strains from the first run on Wednesday were matched with the PAPMID database (Table 1).

From the spectra generated in the first run a provisional phylogenetic tree was constructed with an in-house R script.

DESCRIPTION OF THE MAIN RESULTS OBTAINED

(max. 500 words)
As mentioned before the "Saramis" database does not include 'superspectra' relevant for the *Xanthomonas arboricola* complex. Thus it is not surprisingly that none of the selected strains could be identified with the "Saramis" database.

The comparison to the in-house created spectra from earlier samples was not yet completely interpreted. Results were in some cases contradictory, but it is yet unclear if this was due to the observed peak shifts mentioned above ('description of the work').

A comparison to the PAPMID database results in general in the 10 best matches found, showing the genus and species of the matched database entry and the amount of corresponding proteins between the sample and the database record. This comparison was studied in more detail in this report.
Here we found that all isolates, except two, showed the genus identification as expected (see Table 1, genus and species columns). The two contradicting isolates (PD 5802 and PD 5803) might be wrongly identified or contaminated during reviving of the cultures or inoculation on agar plates. We did not do a verification of the identification with other methods yet.
We found in general good matches on species level, none of them were really contradicting. There were seven isolates (hatched in orange) that showed results that need more investigation due to changed taxonomy (e.g. it is unclear if either for the presumed identity or the PAPMID match a nowadays outdated taxonomy was used at the time of identification). For the 35 presumed Xapr strains we found 19 with the two highest matches indication the expected pathovar. In 11 cases though, the best match was the expected pathovar but the second best match was another pathovar. In five cases the first indicated pathovar was not as expected but the second one was. Only one isolate (PD 740)

showed no Xapr in the first two best matches.
For the none Xapr pathovars no real conclusions could be made because most PAPMID matches did not indicate a pathovar level.

We created a provisional phylogenetic tree with an in-house R script. Remarkable outcome of this was that for a lot of strains one or more of the four replicates per strain ended up in different, relatively distant, branches, indicating poor reproducibility in our runs. The tree was not studied in more details yet.

Table 1: Highest matches of all selected strains with the PAPMID database. Conflicting results hatched in red, corresponding results hatched in green. The results hatched in orange need further analysis due to possibly changed taxonomy.

Strain name & match number	Presumed Identities	Identity according to PAPMID			Match in PAPMID &	
		Genus	species	pathovar	number of corresponding proteins	
PD 760	*P. syringae* pv. *syringae*	*Pseudomonas*	*syringae*	unknown	SWPmod_PL58	38
K 3402	*Xanthomonas* sp.	*Xanthomonas*	*arboricola*	pv. *celebensis*	pv_celebensis_SWPmod	38
PD 5802	*Xanthomonas* sp.	*Luteibacter*	sp.	unknown	9135	35
PD 5803	*Xanthomonas* sp.	*Luteibacter*	sp.	unknown	9135	23
PD 5807	*Xanthomonas* sp.	*Xanthomonas*	*arboricola*	pv. *juglandis*	Xaj417	39
PD 5817	*Xanthomonas* sp.	*Xanthomonas*	*arboricola*	pv. *celebensis*	pv_celebensis_SWPmod	39
PD 5859	*Xanthomonas* sp.	*Xanthomonas*	*gardneri*	unknown	SM1770	28
PD 5860	*Xanthomonas* sp.	*Xanthomonas*	*arboricola*	pv. *juglandis*	CFBP2528T	39
PD 6207	*Xanthomonas* sp.	*Xanthomonas*	*arboricola*	pv. *celebensis*	pv_celebensis_SWPmod	26
PD 6300	*Xanthomonas* sp.	*Xanthomonas*	*arboricola*	unknown	CITA44	40
PD 6702	*Xanthomonas* sp.	*Xanthomonas*	*arboricola*	pv. *celebensis*	pv_celebensis_SWPmod	32
XA 1.59	*X. fragariae*	*Xanthomonas*	*fragariae*	na.	LMG25863	33
XA 1.60	*X. translucens* pv. *translucens*	*Xanthomonas*	*translucens*	pv. *translucens*	pv._translucens DSM _SWPmod	41
PD 5808	*X. campestris*	*Xanthomonas*	*translucens*	pv. poae	LMG728	38
PD 6185	*X. campestris*	*Xanthomonas*	*campestris*	unknown	NA	27
PD 6234	*X. campestris*	*Xanthomonas*	*campestris*	unknown	8004_SWPmod	36
PD 6248	*X. campestris*	*Xanthomonas*	*campestris*	unknown	8004_SWPmod	37
PD 6426	*X. campestris*	*Xanthomonas*	*campestris*	unknown	ATCC_33913 _ NCPPB 5_SWPmod	40
XA 2.33	*X. campestris* pv. *campestris*	*Xanthomonas*	*campestris*	unknown	8004_SWPmod	42
XA 4.65	*X. campestris* pv. *fici*	*Xanthomonas*	*perforans*	unknown	9118_SWPmod	35
XA 4.57	*X. campestris* pv. *gummisudans*	*Xanthomonas*	*translucens*	pv. *translucens*	pv._translucens DSM _SWPmod	26
XA 6.37	*X. arboricola* pv. *celebensis*	*Xanthomonas*	*arboricola*	pv. *celebensis*	pv_celebensis_SWPmod	40
XA 6.39	*X. arboricola* pv. *celebensis*	*Xanthomonas*	*translucens*	pv. *phlei*	LMG730	25
PD 5255	*X. arboricola* pv. *corylina*	*Xanthomonas*	*arboricola*	pv. *celebensis*	pv_celebensis_SWPmod	37
XA 6.40	*X. arboricola* pv. *fragariae*	*Xanthomonas*	*campestris*	unknown	17	36
PD 130	*X. arboricola* pv. *juglandis*	*Xanthomonas*	*arboricola*	pv. *pruni*	pv._pruni_SWPmod	43
PD 4440	*X. arboricola* pv. *juglandis*	*Xanthomonas*	*arboricola*	pv. *juglandis* ?	Xaj417	37
XA 6.42	*X. arboricola* pv. *poinsetticola*	*Xanthomonas*	*arboricola*	unknown	CFBP7634	40
XA 1.71	*X. arboricola* pv. *pruni*	*Xanthomonas*	*arboricola*	unknown	CFBP7634	33
PD 5659	*X. arboricola* pv. *pruni*	*Xanthomonas*	*arboricola*	pv. *pruni*	IVIA2626.1	43
PD 5878	*X. arboricola* pv. *pruni*	*Xanthomonas*	*arboricola*	pv. *pruni*	pv._pruni_SWPmod	40
PD 5879	*X. arboricola* pv. *pruni*	*Xanthomonas*	*arboricola*	pv. *pruni*	pv._pruni_SWPmod	40
PD 5880	*X. arboricola* pv. *pruni*	*Xanthomonas*	*arboricola*	pv. *pruni*	pv._pruni_SWPmod	40
PD 5881	*X. arboricola* pv. *pruni*	*Xanthomonas*	*arboricola*	pv. *pruni*	pv._pruni_SWPmod	41
PD 5891	*X. arboricola* pv. *pruni*	*Xanthomonas*	*arboricola*	pv. *pruni*	pv._pruni_SWPmod	42
PD 5892	*X. arboricola* pv. *pruni*	*Xanthomonas*	*arboricola*	pv. *pruni*	IVIA2626.1	42
PD 5907	*X. arboricola* pv. *pruni*	*Xanthomonas*	*arboricola*	pv. *pruni*	pv._pruni_SWPmod	39
PD 5913	*X. arboricola* pv. *pruni*	*Xanthomonas*	*arboricola*	pv. *celebensis*	pv_celebensis_SWPmod	30
PD 5914	*X. arboricola* pv. *pruni*	*Xanthomonas*	*arboricola*	pv. *celebensis*	pv_celebensis_SWPmod	37
PD 5918	*X. arboricola* pv. *pruni*	*Xanthomonas*	*arboricola*	pv. *pruni*	pv._pruni_SWPmod	41
PD 5919	*X. arboricola* pv. *pruni*	*Xanthomonas*	*arboricola*	pv. *pruni*	pv._pruni_SWPmod	36
PD 5920	*X. arboricola* pv. *pruni*	*Xanthomonas*	*arboricola*	pv. *pruni*	IVIA2626.1	41
PD 5921	*X. arboricola* pv. *pruni*	*Xanthomonas*	*arboricola*	pv. *pruni*	pv._pruni_SWPmod	40
PD 5922	*X. arboricola* pv. *pruni*	*Xanthomonas*	*arboricola*	pv. *pruni*	pv._pruni_SWPmod	40
PD 5923	*X. arboricola* pv. *pruni*	*Xanthomonas*	*arboricola*	pv. *pruni*	pv._pruni_SWPmod	41
PD 5924	*X. arboricola* pv. *pruni*	*Xanthomonas*	*arboricola*	pv. *pruni*	pv._pruni_SWPmod	35
PD 5944	*X. arboricola* pv. *pruni*	*Xanthomonas*	*arboricola*	pv. *pruni*	pv._pruni_SWPmod	40
PD 5949	*X. arboricola* pv. *pruni*	*Xanthomonas*	*arboricola*	pv. *pruni*	IVIA2626.1	32
PD 5950	*X. arboricola* pv. *pruni*	*Xanthomonas*	*arboricola*	pv. *pruni*	pv._pruni_SWPmod	37
PD 6092	*X. arboricola* pv. *pruni*	*Xanthomonas*	*arboricola*	pv. *celebensis*	pv_celebensis_SWPmod	28
PD 6095	*X. arboricola* pv. *pruni*	*Xanthomonas*	*arboricola*	pv. *pruni*	pv._pruni_SWPmod	40
PD 6118	*X. arboricola* pv. *pruni*	*Xanthomonas*	*arboricola*	pv. *pruni*	IVIA2626.1	38
PD 6169	*X. arboricola* pv. *pruni*	*Xanthomonas*	*arboricola*	pv. *pruni*	pv._pruni_SWPmod	38
PD 6194	*X. arboricola* pv. *pruni*	*Xanthomonas*	*arboricola*	pv. *pruni*	pv._pruni_SWPmod	43
PD 6256	*X. arboricola* pv. *pruni*	*Xanthomonas*	*arboricola*	pv. *pruni*	pv._pruni_SWPmod	43
PD 6261	*X. arboricola* pv. *pruni*	*Xanthomonas*	*arboricola*	pv. *pruni*	pv._pruni_SWPmod	38
PD 6671	*X. arboricola* pv. *pruni*	*Xanthomonas*	*arboricola*	pv. *pruni*	IVIA2626.1	44
PD 6672	*X. arboricola* pv. *pruni*	*Xanthomonas*	*arboricola*	pv. *pruni*	pv._pruni_SWPmod	42
PD 6673	*X. arboricola* pv. *pruni*	*Xanthomonas*	*arboricola*	pv. *celebensis*	pv_celebensis_SWPmod	30
PD 6714	*X. arboricola* pv. *pruni*	*Xanthomonas*	*arboricola*	pv. *pruni*	IVIA2626.1	33
PD 6716	*X. arboricola* pv. *pruni*	*Xanthomonas*	*arboricola*	pv. *pruni*	IVIA2626.1	37
PD 6723	*X. arboricola* pv. *pruni*	*Xanthomonas*	*arboricola*	pv. *pruni*	pv._pruni_SWPmod	37
PD 740	*X. arboricola* pv. *pruni*	*Xanthomonas*	*arboricola*	pv. *celebensis*	NCPPB1832	26
PD 998	*X. arboricola* pv. *pruni*	*Xanthomonas*	*arboricola*	pv. *pruni*	pv._pruni_SWPmod	33

FUTURE COLLABORATIONS (if applicable)

(max. 500 words)
We will continue the work initiated in this STSM at the Home Institute (NVWA) by analyzing the created data and selecting possible strains to build up a database for identification. At least, at the species level the generated results look very promising, however, a good separation on the pathovar level will be rather difficult. We aim at the development of a database derived from genome based sequence data of all seven pathovars inside the *Xanthomonas arboricola* complex that will be complementary to the MALDI-TOF MS methodology.

Apart from the data generated, this STSM at ZHAW provided new insights on the MALDI-TOF MS cutting-edge technology for bacterial identification. This COST STSM was also a good opportunity to strengthen the long-term collaboration established between ZHAW and NVWA during the former COST 873 Action with projects aiming mainly to identify and perform subtyping of *X. arboricola* isolates.

In future the ongoing collaboration between the two institutions involved in this STSM might lead to one or more publication(s) on this topic and/or inter laboratory studies on MALDI-TOF MS, including the collaboration inside the EUPHRESCO III project entitled: "Rapid identification of plant-health related bacteria by MALDI-TOF mass spectrometry (2018-A-271)".