A socio-semantic analysis of the research domain on Xylella fastidiosa Structure and scientific dynamics

Marc Barbier, Pauline Farigoule, Astrid Cruaud, Jean-Pierre Rossi

A socio-semantic analysis of the research domain on *Xylella fastidiosa*

Structure and scientific dynamics

Marc Barbier[3], Pauline Farigoule[1,2], Astrid Cruaud[2], Jean-Pierre Rossi[2]

[1] AgroParisTech, Paris, France
[2] CBGP, INRAE, CIRAD, IRD, Montpellier SupAgro, Univ. Montpellier, Montferrier-sur-Lez, France
[3] UMR LISIS, INRAE, Marne-la-Vallée, France

Contact : Marc Barbier – marc.barbier@inrae.fr

OBJECTIVE

This poster intends to offer a socio-semantic overall view of the research domain on *Xylella fastidiosa* as a major quarantine pest with worldwide expansion. This scientometric analysis is applied on a corpus of bibliographic notices delineated in the Web of Science (WoS) database with a query based on the topic "*Xylella fastidiosa*".

The purpose of this analysis is to deliver to researchers and policy-makers a retrospective account of the constitution and development of the research domain on *Xylella fastidiosa.* This account is based on the complementary visions and knowledge of authors' discipline: computational social science, ecology and genomics.

Features of the Corpus extracted from WoS

<u>Date of extraction</u>: 15/03/2021 – Time Span: 1987-2021
<u>Topics :</u> TS=*xylella fastidiosa*
<u>Size:</u> 1840 publications
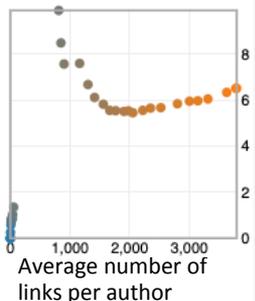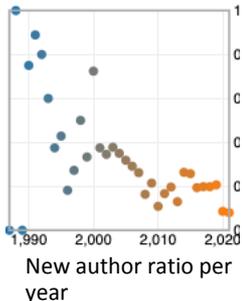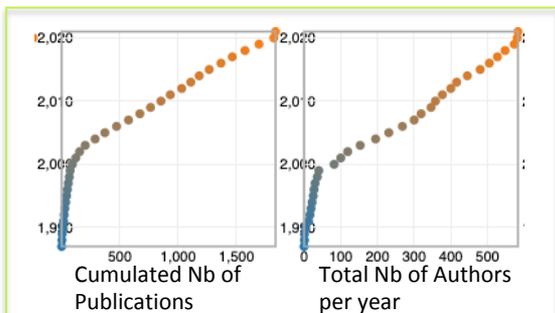<u>Export:</u> Full records and cited references

AUTHORS	Total Nb of Authors	4207
	Number of Authors	Prct of Total Pub.
With 3 publications at least	791	50%
With 4 publications at least	593	40%
With 7 publications at least	229	30%
JOURNALS	Total Nb of Journals	7763
	Number of Journals	Prct of Total Publications
Threshold 117 articles published	90	60%
Threshold 17 articles published	449	80%
CITED REFERENCES	Total Nb of Cited Ref	Prct of Total Publications
Threshold 3 citations	3993	51%
Threshold 10 citations	846	39%
Threshold 29 citations	229	20%
Threshold 70 citations	60	10%

Analysis are conducted with the CorTexT Platform, an open science capacity to run various type of textual and network analysis which belong to the RISIS European Infrastructure for science and innovation policies.
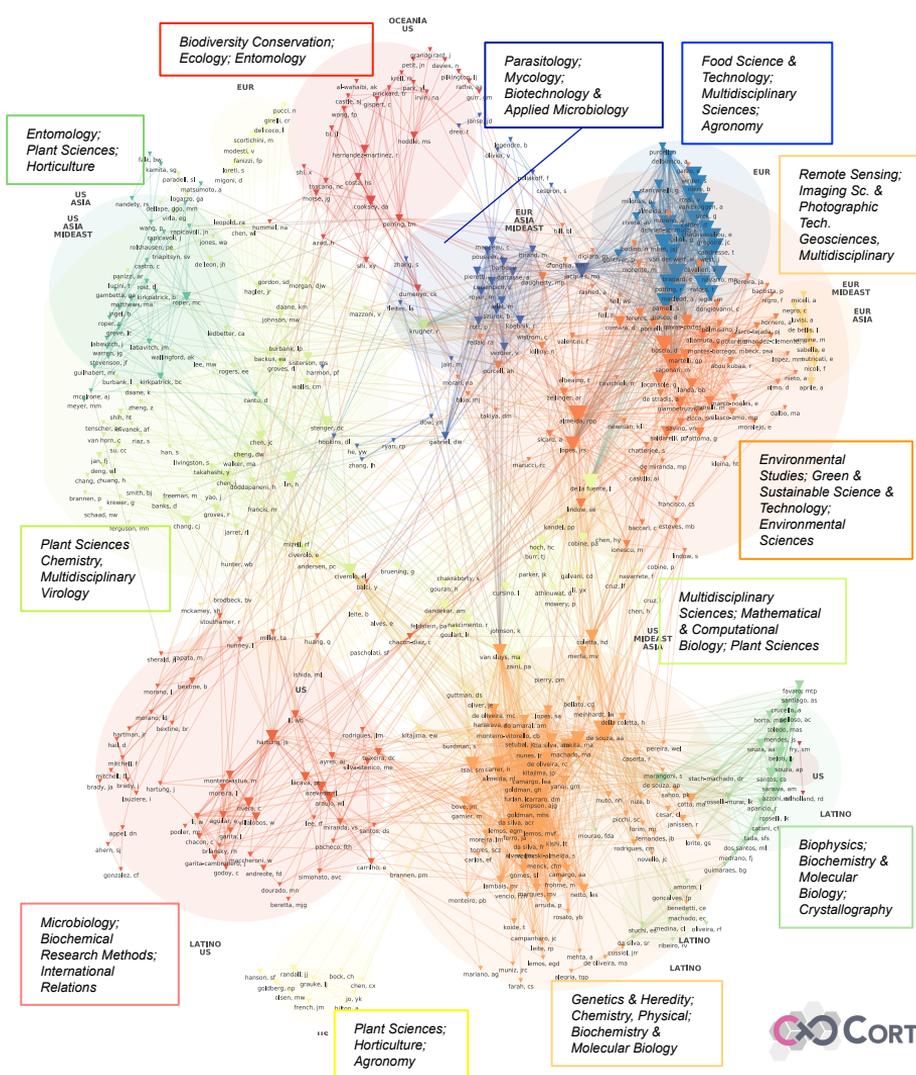
AUTHORS

A **growing population of authors** but **structurally stable** in terms of new comers and links of co-authorship


Cumulated Nb of Publications

Total Nb of Authors per year

New author ratio per year

Average number of links per author

The gradient of colours represents the years from 1985 to 2021

Biodiversity Conservation; Ecology; Entomology

Parasitology; Mycology; Biotechnology & Applied Microbiology

Food Science & Technology; Multidisciplinary Sciences; Agronomy

Entomology; Plant Sciences; Horticulture

Remote Sensing; Imaging Sc. & Photographic Tech. Geosciences, Multidisciplinary

Plant Sciences Chemistry, Multidisciplinary Virology

Environmental Studies; Green & Sustainable Science & Technology; Environmental Sciences

Multidisciplinary Sciences; Mathematical & Computational Biology; Plant Sciences

Biophysics; Biochemistry & Molecular Biology; Crystallography

Microbiology; Biochemical Research Methods; International Relations

Plant Sciences; Horticulture; Agronomy

Genetics & Heredity; Chemistry, Physical; Biochemistry & Molecular Biology

Network of co-authorship
Top 593 Authors
40% of Publications
The corpus includes
Notices of authors that have more than 4 publications
Metric Raw, Louvain viz.

Node size reflects the number of publications and **lines** show links between authors.
The landscape of co-authors clusters is calculate with louvain algo.
Two variables are projected: the **main continents** to which authors of each cluster belong (Chi2 score) and the top 3 **WoS Category** (Chi2 score).

Cortext

3

THE EPISTEMIC COMMUNITIES OF THE RESEARCH DOMAIN
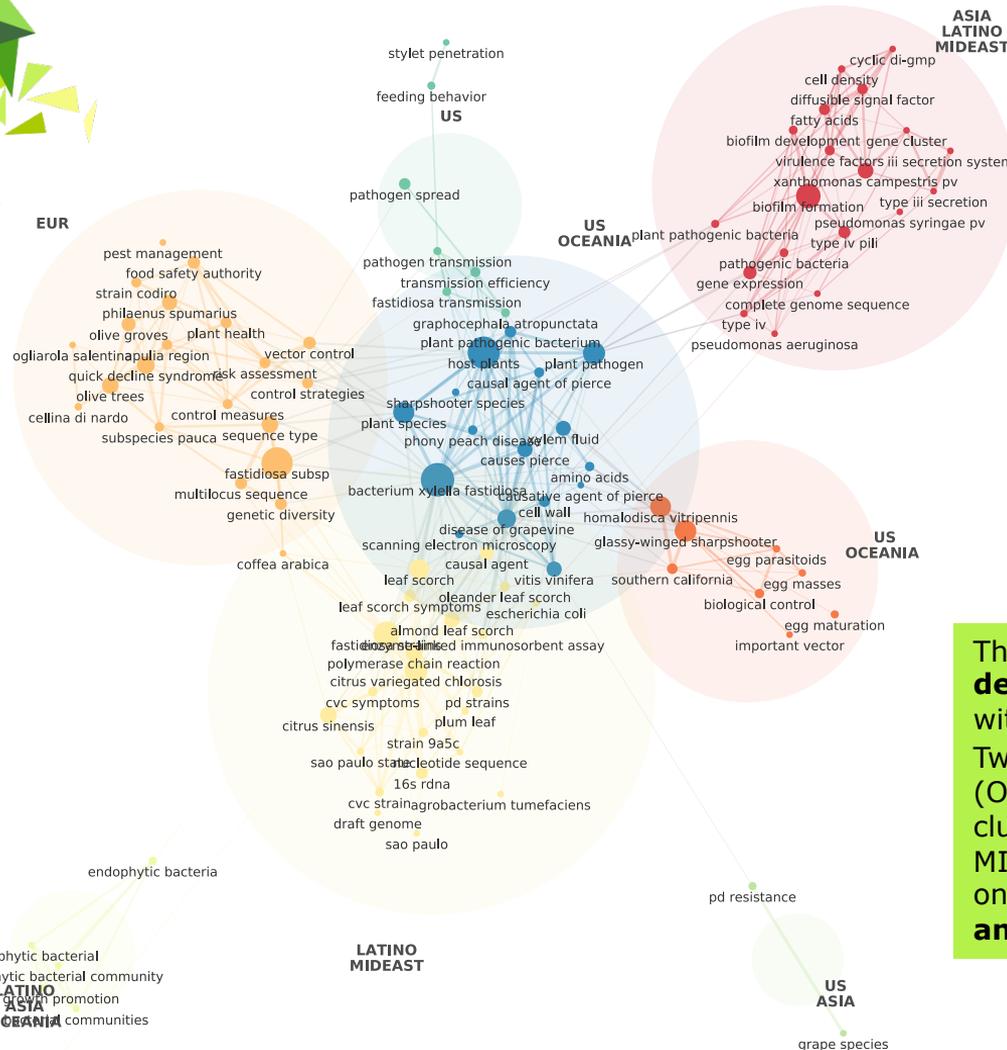
This map shows clusters of terms extracted from abstracts (top 150) built with a vectorial metric (Cosine-Heterogene) that measure the proximity of terms thanks to the calculation of the weighted contributions of Cited References (Top 500, this threshold corresponds to a score of citation above 10).

On each cluster the main Continents are projected (Top 3 with Chi2 score). Each Cluster is interpreted as an epistemic communities, meaning the proximity of concepts and objects of enquiry that emerged from authors' practices of citations.

The **landscape** shows **three clusters** that are **diseases dependant** (Pierce Dis., Blue; CVC Yellow; QDS Orange) with **clear geographical association**.

Two other main **clusters** are less disease dependant, one (Orange) on **vector life cycle** (US & OCEANIA), another cluster (Red) on **biofilm formation** (ASIA & LATINO & MIDEAST). A last cluster (Green) is associated to the Blue one and shows the research on **pathogen transmission and feeding behavior** (US)

GEO-SCIENCE OF *XYLELLA FASTIDIOSA* RESEARCH
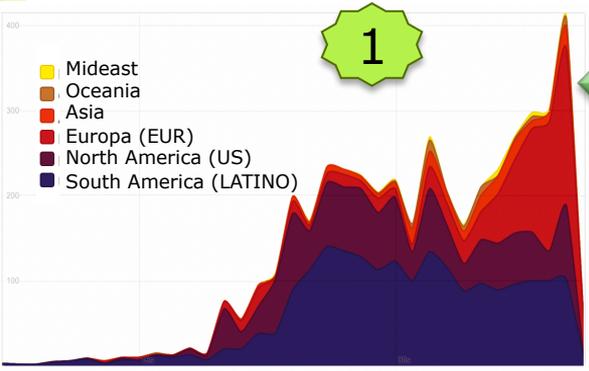
1. The **demography** of Continents shows an **expansion** of Xyl.fast. Res.Domain through years and continents. This extension is **related to disease outbreaks**.

2. A **Contingency Matrix** crossing Clusters of epistemic communities (see colours) with Continents shows a **clear specificity of association** (measured by the Nb of documents and the Chi 2 score)
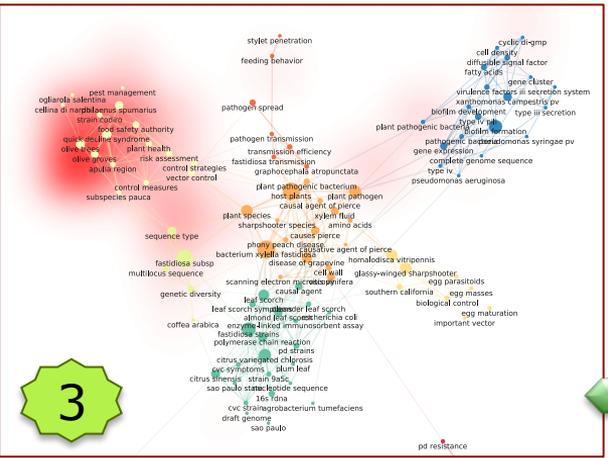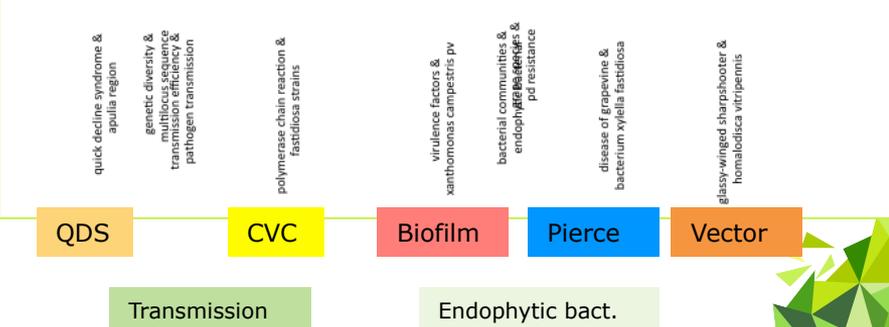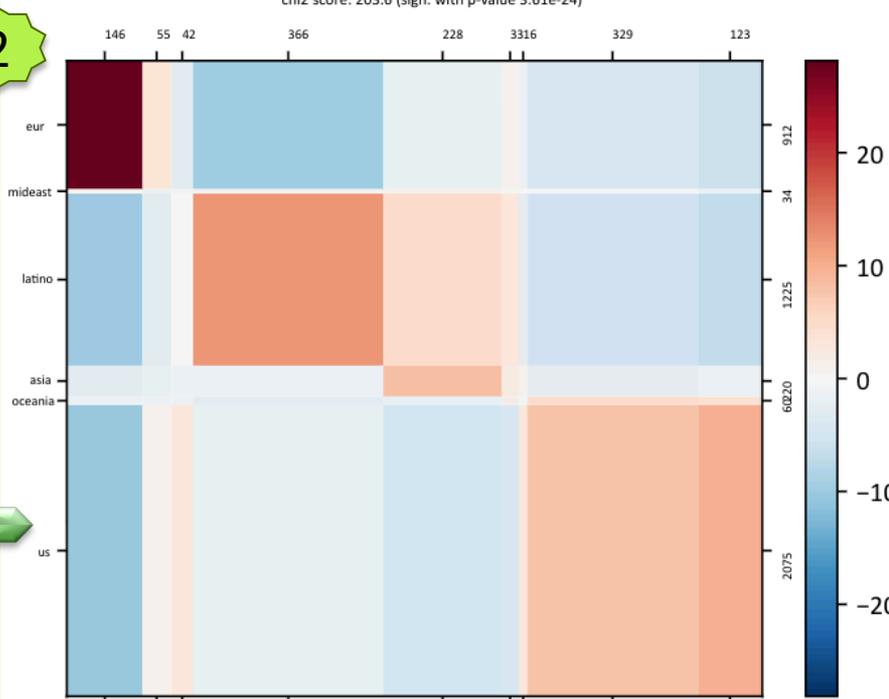
3. To better highlight the position of European countries a **heat map** (Scale of Chi2 intensity, red background) shows **how Europe impacts the global landscape with research targeted around QDS.**

5