



INRAE



3rd European Conference on
Xylella fastidiosa and XF-ACTORS final meeting



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

Structure and scientific dynamics

Marc Barbier³, Pauline Farigoule^{1,2}, Astrid Cruaud², Jean-Pierre Rossi²

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

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

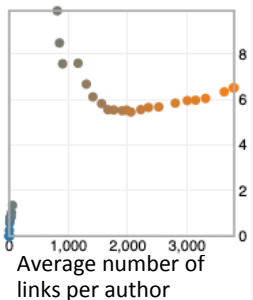
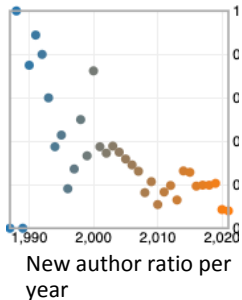
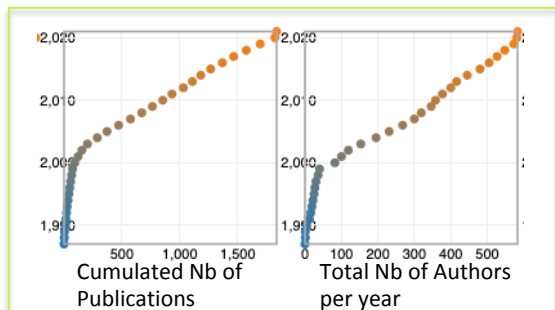
AUTHORS	Total Nb of Authors	4207
	Number of Authors	Prct of Total Pub.
With 3 publications at least	791	50%
With 4 publications at least	593	40%
With 7 publications at least	229	30%
JOURNALS	Total Nb of Journals	7763
	Number of Journals	Prct of Total Publications
Threshold 117 articles published	90	60%
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.

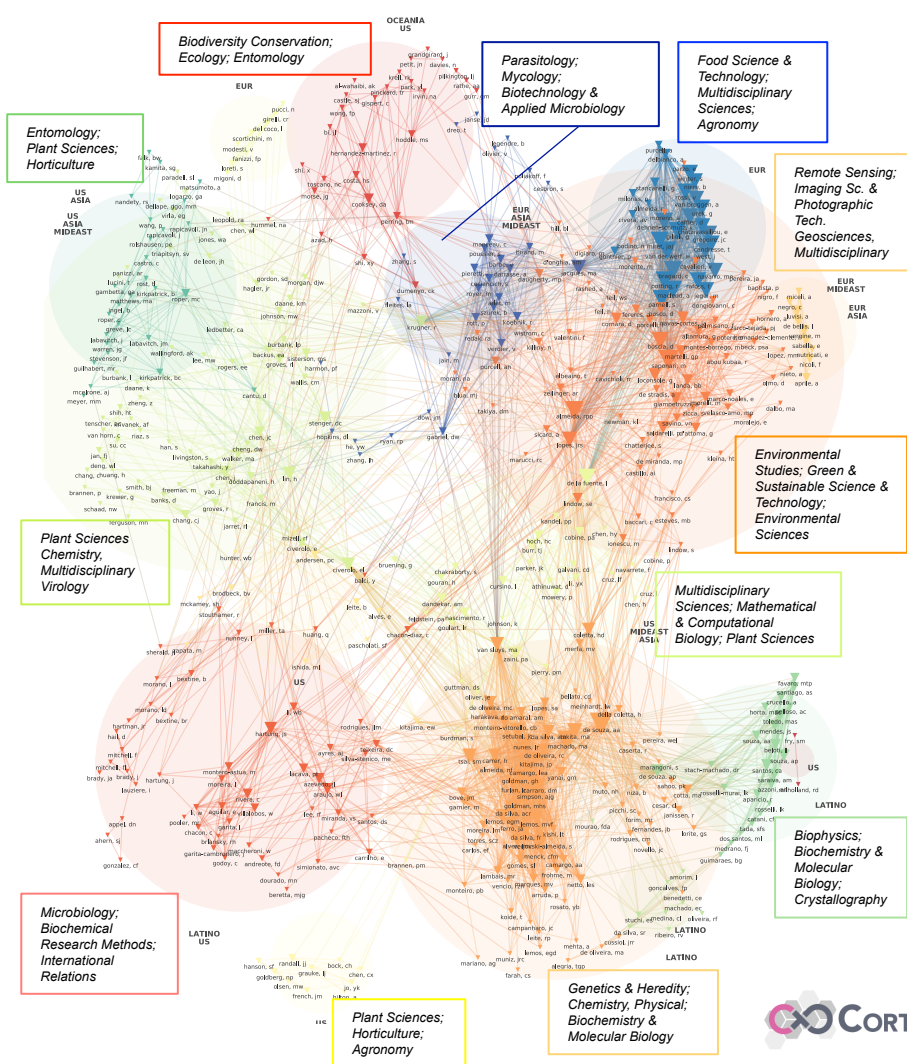
The screenshot shows the CorText dashboard for a project named 'scento_xylella_colloque_efsa'. The interface includes a navigation bar with 'dashboard' and 'project' tabs. Below the navigation bar, there are several sections: 'latest analyses' with a list of recent analyses (Structural Analysis, Network Mapping, Demography, Network Mapping); 'datasets' with a list of datasets (1recent-xylella-fastidiosa-topi-15...); 'participants' with a list of participants (Farigoule Pauline, BARBIER); and a central area for 'structural analysis' showing a file named 'statistics_growing_250_raw_0.0.htm - 3.3 kB' and a folder named '/structural_evolution/data'. The bottom of the dashboard features the CorText logo and the text 'es co-auteur top 593 avec Raw; une belle carte de sociologie des'.

AUTHORS

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



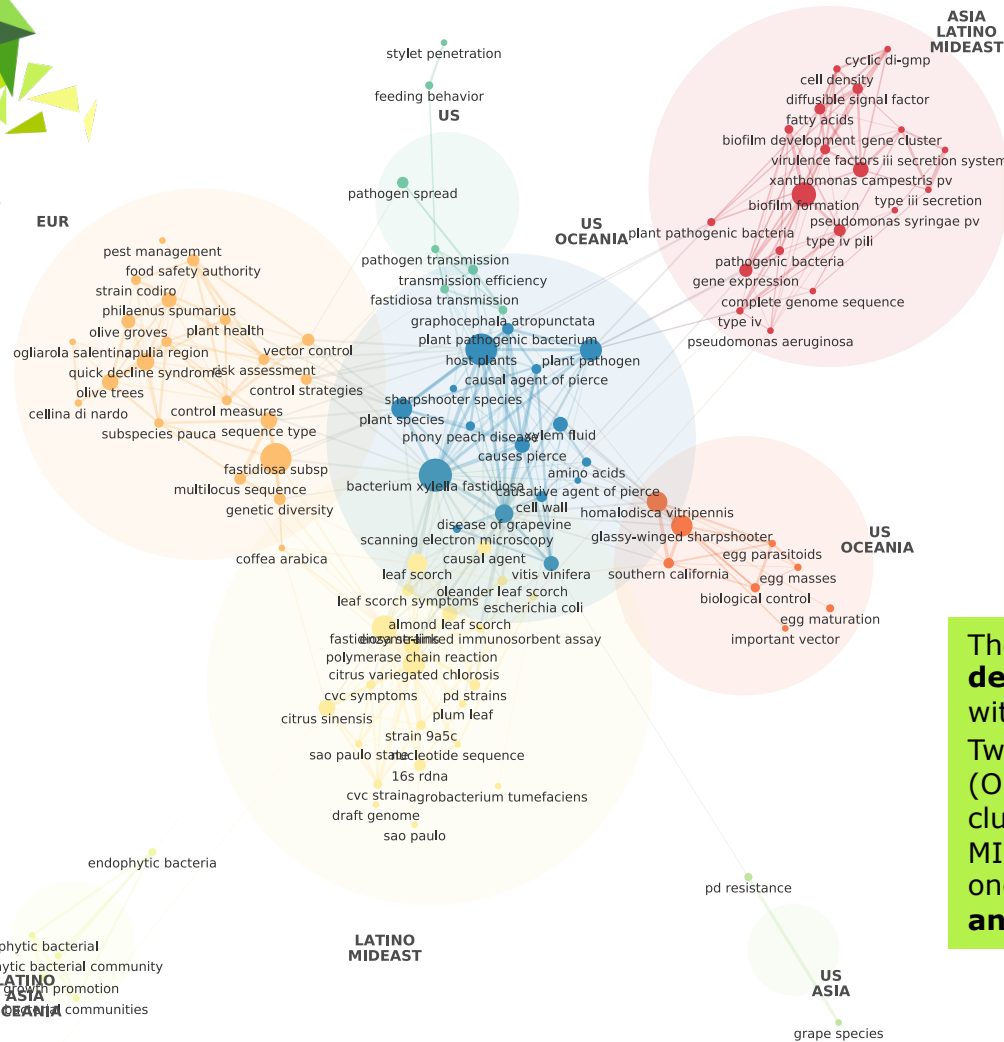
The gradient of colours represents the years from 1985 to 2021



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).

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.

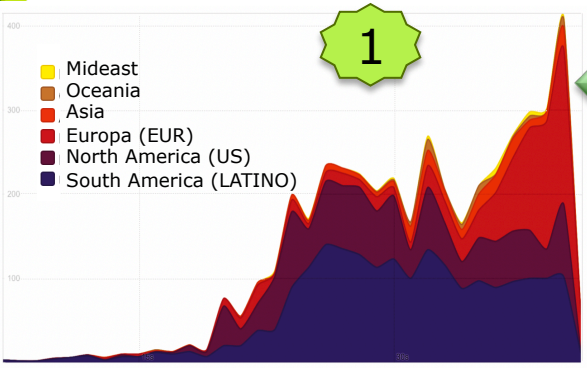
The **landscape** shows **three clusters** that are **diseases dependant** (Pierce Dis., Blue; CVC Yellow; QDS Orange) with **clear geographical association**.

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

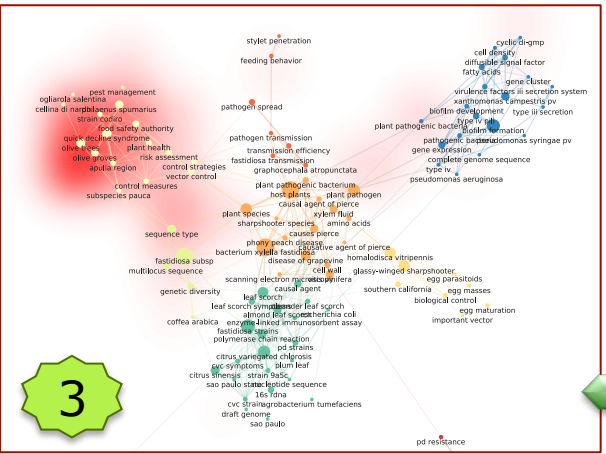
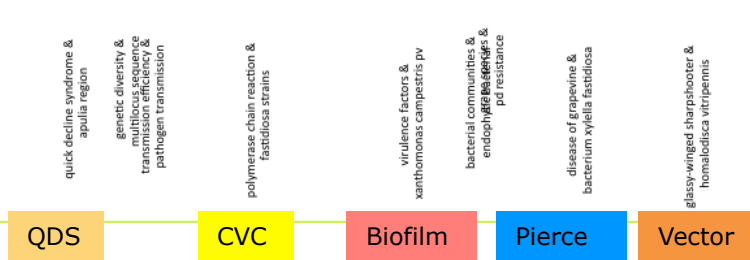
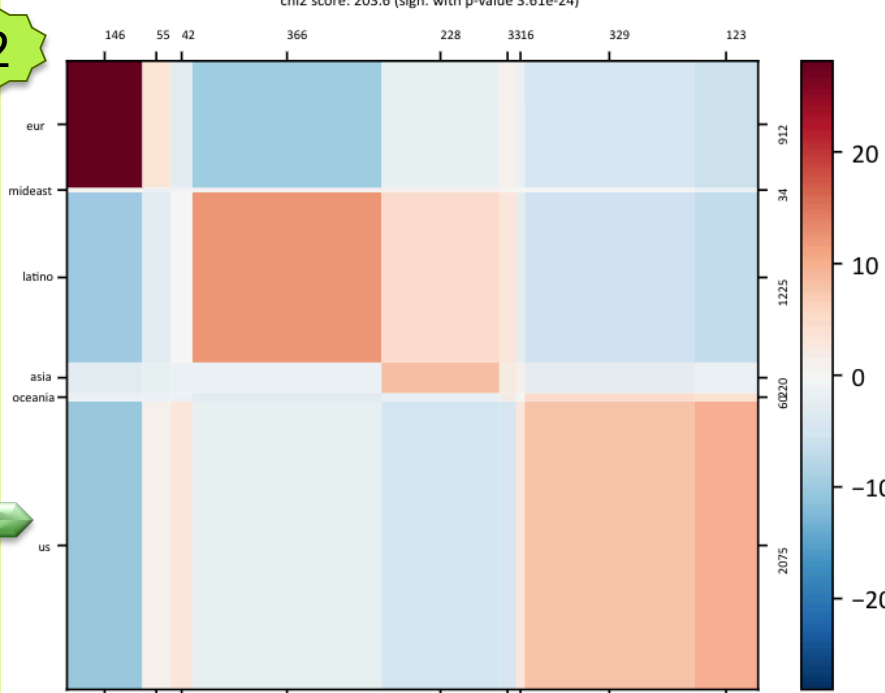
2

chi2 score: 203.6 (sign. with p-value 3.61e-24)



1

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

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**.

