

# Connecting plant phenotyping and modelling communities: lessons from science mapping and operational perspectives

Clément Saint Cast<sup>1,2,\*</sup>, Guillaume Lobet<sup>1,3</sup>, Llorenç Cabrera-Bosquet<sup>4</sup>,  
Valentin Couvreur<sup>1</sup>, Christophe Pradal<sup>5,6</sup>, François Tardieu<sup>4</sup> and Xavier Draye<sup>1</sup>

<sup>1</sup>Earth and Life Institute, University of Louvain, 1348 Louvain-La-Neuve, Belgium

<sup>2</sup>EGFV, INRAE, Bordeaux-Sciences Agro, Université Bordeaux, ISVV, 210 Chemin de Leysotte, 33 140 Villenave d'Ornon, France

<sup>3</sup>Agrosphere, IBG3, Forschungszentrum Jülich, 52428 Jülich, Germany

<sup>4</sup>INRAE, LEPSE, Montpellier SupAgro, Université de Montpellier, 2 Place Viala, 34000 Montpellier, France

<sup>5</sup>CIRAD, UMR AGAP Institut, Avenue Agropolis, 34398 Montpellier, France

<sup>6</sup>Inria & LIRMM, Université de Montpellier, CNRS, 34095 Montpellier, France

\*Corresponding author's e-mail address: [clement.saint-cast@inrae.fr](mailto:clement.saint-cast@inrae.fr)

Handling Editor: Steve Long

## ABSTRACT

Plant phenotyping platforms generate large amounts of high-dimensional data at different scales of plant organization. The possibility to use this information as inputs of models is an opportunity to develop models that integrate new processes and genetic inputs. We assessed to what extent the phenomics and modelling communities can address the issues of interoperability and data exchange, using a science mapping approach (i.e. visualization and analysis of a broad range of scientific and technological activities as a whole). In this paper, we (i) evaluate connections, (ii) identify compatible and connectable research topics and (iii) propose strategies to facilitate connection across communities. We applied a science mapping approach based on reference and term analyses to a set of 4332 scientific papers published by the plant phenomics and modelling communities from 1980 to 2019, retrieved using the Elsevier's Scopus database and the [quantitative-plant.org](http://quantitative-plant.org) website. The number of papers on phenotyping and modelling dramatically increased during the past decade, boosted by progress in phenotyping technologies and by key developments at hardware and software levels. The science mapping approach indicated a large diversity of research topics studied in each community. Despite compatibilities of research topics, the level of connection between the phenomics and modelling communities was low. Although phenomics and modelling crucially need to exchange data, the two communities appeared to be weakly connected. We encourage these communities to work on ontologies, harmonized formats, translators and connectors to facilitate transparent data exchange.

**KEYWORDS:** Bibliometric analyses; data and model integration; imaging; modelling; network; plant phenotyping.

## 1. INTRODUCTION

During the past decade, plant phenotyping platforms have generated large amounts of detailed data at different spatial and temporal scales for thousands of genotypes under controlled conditions or in the field (Tardieu *et al.* 2017). Information extracted from these datasets could be more widely used as variables or parameters of mathematical and computational models, thereby broadening the scope of information extracted from phenomics data (Muller and Martre 2019; Louarn and Song 2020). Feeding such data to process-based crop models (models

representing crop characteristics at the field scale; Jones *et al.* 2003; Holzworth *et al.* 2014) or individual-based models (models representing plants individually with various degrees of architectural realism, e.g. functional–structural plant models; Evers *et al.* 2018; Louarn and Song 2020) in *ad hoc* pipelines has the potential to predict integrated (e.g. yield) or functional traits (e.g. root system architecture) across a wide range of environments or management practices (Liu *et al.* 2017; Chen *et al.* 2019). Interestingly, such interplay between models and data could also help identify which plant traits and metadata

are needed most in model calibration and parametrization, potentially identifying new traits or environmental data of interest to the plant phenotyping community (Long 2019). Furthermore, plant phenotyping platforms supply information to improve robustness and biological soundness of plant and crop models by providing detailed data to integrate new physiological processes and genetic inputs (Gosseau et al. 2019; Muller and Martre 2019). For example, process-based crop models with genotype-dependent parameters will become more widespread as parameters can now be obtained for thousands of genotypes in phenotyping platforms (Parent and Tardieu 2014; Casadebaig et al. 2020).

A crucial question is to what extent phenomics and modelling communities exchange datasets and use them together. Bibliometric analysis and science mapping of scientific publications are relevant approaches to assess current connections and identify compatible and connectable research topics. They have been used to provide new insights into specific disciplines (e.g. reveal the historical evolution or identify the emerging topics of a research discipline) and provide a field overview by visualizing the main research areas of a community. For example, bibliometric science mapping has been used to evaluate the structure and evolution of Mediterranean forest research (Nardi et al. 2016). It has also revealed the evolution of plant phenotyping research structures (Costa et al. 2019), with a trend toward a higher diversity of phenotyped species and research in real field conditions. However, bibliometric mapping science has never been applied to reveal and evaluate the connections between different scientific communities.

Here, we carried out bibliometric analysis on a set of scientific papers related to plant phenotyping and modelling. We retrieved the papers via Elsevier's Scopus database. Then, we evaluated the connections and identified compatible research topics between the communities using the VOSviewer software (van Eck and Waltman 2010). Finally, we present strategies to facilitate connection across scientific communities in the context of the EMPHASIS research infrastructure initiative (European Infrastructure for Multi-scale Plant Phenomics and Simulation for Food Security in a Changing Climate; <https://emphasis.plant-phenotyping.eu/>).

## 2. MATERIALS AND METHODS

### 2.1 Scopus search

The Scopus database (<https://www.scopus.com/>) was used to retrieve bibliographic records related to plant phenotyping, plant image analysis software tools (called 'image analysis tools' hereafter), process-based crop models (called 'crop models' hereafter) and individual-based models (called 'plant models' hereafter) over the period 1980–2019. To identify relevant papers and retrieve the bibliographic datasets associated with each community (image analysis tool database, crop model database or plant model database), we used the following keywords in the combined field of title, abstract and keywords for:

- the plant phenotyping: phenotyping AND plant\*
- the image analysis tools: plant\* AND image\* PRE/2 analys\* AND software
- the crop models: crop PRE/2 model\* AND growth AND simulation OR crop PRE/2 growth PRE/2 model\* AND simulation

- the plant models: structural\* PRE/3 plant\* PRE/3 model\* OR functional\* PRE/3 plant\* PRE/3 model\* OR structural-functional\* PRE/2 model\* AND plant\* OR functional-structural\* PRE/2 model\* AND plant\* AND NOT structural\* PRE/1 equation\*

The proximity operator PRE/*n* specifies the proximity (*n* being the number of terms) between two terms in an exact order. For instance, 'crop PRE/2 model' will find documents in which 'crop' precedes 'model' with no more than two terms in between. The wildcard character (\*) was used to replace or represent one or more characters. For instance, 'model\*' will find documents in which 'model', 'models', 'modelling', 'modeling', 'modelled' or 'modeled' appear.

The Scopus search was conducted in June 2020. For this reason, 2020 papers were not yet completely introduced in the Scopus database and the 2020 papers were not considered in this database.

It should be noted that, due to Scopus limitations, some sources could be missing in the research. Indeed, Scopus only analyzes the citations of the journals in its index and has a bias in its coverage of European journals and Elsevier titles. The search was restricted to papers written in English; therefore, the study might exclude regionally important research published in other languages.

### 2.2 'Quantitative Plant' database

The Scopus database was completed with scientific articles from the online resource [quantitative-plant.org](https://www.quantitative-plant.org) (Saint Cast and Lobet 2019), a website referencing plant image analysis software tools, crop models and plant models. 'Quantitative Plant' extends the 'Plant Image Analysis website' (Lobet et al. 2013; Lobet 2017). One of the objectives of 'Quantitative Plant' is to develop an online portal referencing crop and plant models to raise awareness and highlight the diversity of models and their applications. This website is hand curated with image analysis tools and plant or crop models identified in a thorough review of the literature. Each tool and model is presented concisely in a consistent framework and is described by their general characteristics (e.g. plant part studied) and uses (e.g. species studied). They are referenced by one or more scientific articles presenting their characteristics or success stories. These papers were added to the corresponding Scopus datasets (image analysis tool database, crop model database or plant model database).

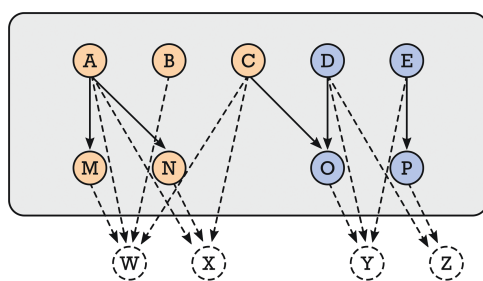
### 2.3 Bibliometric mapping

In order to visualize connections between papers from both communities, we used the VOSviewer software version 1.6.13 (freely available at <https://www.vosviewer.com/>) which was specifically developed for creating, visualizing and exploring science's bibliometric maps (van Eck and Waltman 2010). VOSviewer allows linking documents that reference the same set of cited documents (i.e. bibliographic coupling links). Two articles with at least one common reference are identified as articles with a bibliographic coupling relationship (Kessler 1963; Chang and Huang 2012). The method of bibliographic coupling can help researchers to filter out a group of articles with a particular type of connection by the creation of so-called 'bibliography coupling maps' (Boyack and Klavans 2010). A bibliography coupling map is a two-dimensional representation of a research field, in which the distance

between related papers depends on their similarities. Thus, bibliographic coupling maps provide overviews for visualizing and identifying the connection between papers and different cluster groups or communities. To display the elements on the maps, the software uses the VOS (Visualization Of Similarities) mapping technique, that is closely related to the well-known multidimensional scaling method (van Eck and Waltman 2010). The principle of the VOS mapping technique is to minimize a weighted sum of squared Euclidean distances between all pairs of papers through an optimization process. This mapping approach allows laying out papers on the map following the distance between each pair of papers that represents their similarity. In a bibliographic coupling map, similarities among papers are calculated based on the number of cited reference papers they have in common (for further explanation regarding the method, please see van Eck and Waltman 2010). The larger the number of cited reference papers have in common, the stronger the papers are related to each other. Therefore, papers citing the same references are closely located to each other on the map while less strongly related papers are located farther from each other. An example of the bibliographic coupling approach is illustrated in Fig. 1.

#### 2.4 Bibliometric clustering

To identify clusters of related papers, the software uses a weighted and parameterized variant of modularity-based clustering, that is the VOS clustering technique (Waltman *et al.* 2010; Waltman and van Eck 2013). A cluster can be understood as a research area in which one or more research topics can be identified. The assignment of two papers to the same cluster depends on the cited references the two papers have in common. Papers citing the same documents are strongly related to each other and are likely to be assigned to the same cluster. On the contrary, papers with a low number of cited references in common are likely to be assigned to different clusters. Papers without cited references in common with other papers are not assigned to a cluster and are absent from the bibliometric maps and analysis.



**Figure 1.** Example of how the bibliographic coupling approach partitions a set of scientific papers. The grey box represents the documents within a Scopus dataset (e.g. scientific papers associated to the plant phenotyping community). Documents W, X, Y and Z are documents outside the set, but are referenced by documents within the set. Solid arrows represent citations to documents within the set. Dashed arrows represent citations to documents outside the set. Colour in the panel shows how the documents might be clustered by this approach.

In order for the reader of the present article to navigate all the maps with labels, VOSviewer Map and Network files are available as [Supporting Information—Figure S1](#). It should be noted that a bibliographic coupling map represents a simplified version of reality on a subject, owing to the loss of information and partial representation of the investigated field (van Raan 2014). This limitation should be considered when interpreting results.

#### 2.5 Identification of the research topics in each community

To identify the research topics of each cluster, an analysis of the most common terms used by the papers of each cluster was performed. Terms occurring in titles, abstracts and keywords were extracted from papers and analysed to identify the frequency distribution of the key terms associated to the papers. The most frequent terms of the cluster papers were used to characterize the research topics of each group and identify compatibilities between groups (i.e. proportion of common research topics shared).

This analysis was performed in R environment (R Development Core Team 2019) using the `biblioAnalysis` function of the `bibliometrix` package (Aria and Cuccurullo 2017).

Before starting with the analysis, a thesaurus file was created to ensure consistency for different term spelling and synonyms (e.g. leaf area index is often termed 'LAI'). We also cleaned the data by omitting terms considered not relevant for analyses: terms related to time, publishers' names or geographical locations (e.g. names of cities or countries).

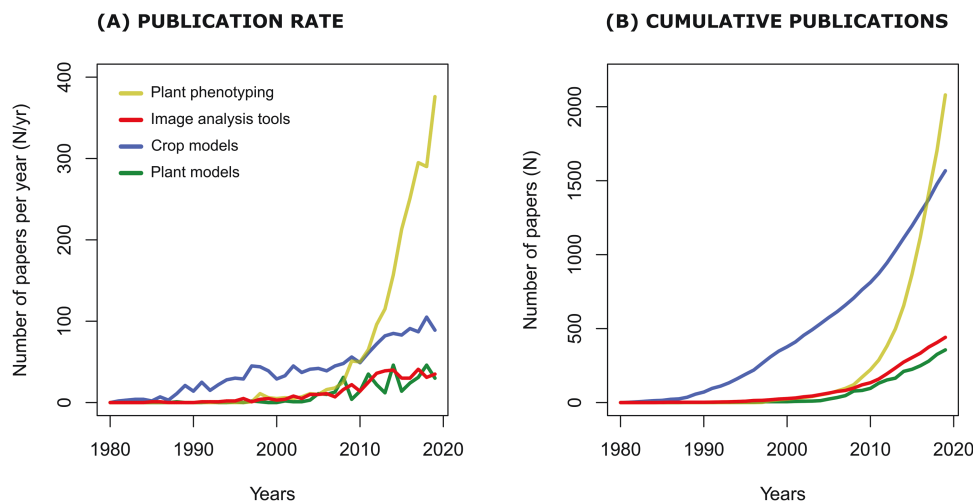
### 3. RESULTS

#### 3.1 Evolution of the trends of publication

A total of 4173 scientific papers were retrieved from the Scopus database, completed by 159 scientific papers from the [quantitative-plant.org](#) website. Approximately, 88.6 % of the papers were research papers, 6.9 % review papers, 3.4 % book chapters and the remaining 1.1 % were books, letters, conference papers or notes. Top journals were *Frontiers in Plant Science* ( $n = 224$ ; 5.2 %), *Field Crop Research* ( $n = 163$ ; 3.8 %), *Annals of Botany* ( $n = 144$ ; 3.3 %), *Plant Methods* ( $n = 132$ ; 3.2 %) and *Journal of Experimental Botany* ( $n = 120$ ; 2.8 %).

These papers were then clustered manually in four categories that belong to either phenomic or modelling communities, namely plant phenotyping, image analysis tools, crop models and plant models (numbers of papers:  $n = 2074$ ,  $n = 363$ ,  $n = 1567$  and  $n = 328$ , respectively).

The crop model community produced the highest cumulative number of papers from 1980 to 2010 (222, 134, 812 and 96 papers in 21 years for the four aforementioned categories, respectively; Fig. 2). Between 2011 and 2019, the paper rate per year increased in all four communities (+1853, +435, +117 and +532 %, or 1858, 307, 755 and 260 additional papers in only 9 years). The trend was close to exponential for the plant phenotyping community, whose paper rate per year increased by about +2600 % in 2015 to 2019 relative to the 1980–2010 reference (1425 cumulative papers, or more than 68 % of all papers in 5 years). Similarly, more than 35 % of the image analysis tool and plant model papers were published during the last 5 years (i.e. 38 and 41 %, respectively).



**Figure 2.** (A) Number of papers per year and (B) cumulative number of papers from the plant phenotyping (yellow), plant image analysis tool (red), process-based crop model (blue) and individual-based model (green) communities from 1980 to 2019.

$n = 167$  and  $n = 145$ , during the period 2015–19 for the image analysis tools and plant models, respectively).

### 3.2 Research topics of the phenomics and modelling communities

Most of the scientific papers were grouped into six clusters within each community (99, 74, 86 and 99 % of the papers, for the plant phenotyping, image analysis tools, crop models and plant models, respectively; Fig. 3). The clustering procedure discriminated papers according to the plant parts (e.g. above- or belowground compartments), the species (e.g. wheat, maize, rice or *Arabidopsis*) or the subject area (e.g. biochemistry, genetics, engineering, computer science, environmental science or mathematics). For example, the blue cluster of the plant model community is mainly represented by terms related to the shoot part of the plant (i.e. 'Light', 'Photosynthesis' and 'Plant leaf'; Fig. 3D), whereas the orange cluster is mainly represented by terms related to the roots (i.e. 'Plant root', 'Root system' and 'Soil'; Fig. 3D). The complete lists of terms identified in each cluster are given in Supporting Information—Table S1.

Not surprisingly, the communities shared common research topics, i.e. high compatibility. For example, the research topics associated to the shoot and root in the plant model community (the blue and orange clusters, for the shoot and root parts, respectively; Fig. 3D) were also observed in the communities of plant phenotyping (the orange and green clusters, for the shoot and root parts, respectively; Fig. 3A) and image analysis (the yellow and the red clusters for the shoot and root parts, respectively; Fig. 3B). On the contrary, most of the research topics identified in the crop model clusters (e.g. 'Crop yield', 'Climate Change' or 'Climate effect'; Fig. 3C) were not observed in the others communities.

### 3.3 Connection between phenomics and modelling communities

The coupling map of the four communities combined (4332 articles) is presented in Fig. 4. The plant phenotyping community is positioned at the left of the map (yellow group in the Fig. 4). It is mingled with the

image analysis tool papers (red group in Fig. 4) and characterized by a wide distribution. This community is characterized by a low number of links (i.e. cited references in common with all other scientific papers) per paper ( $249 \pm 252$ ; Table 1). The links occur largely between plant phenotyping and image analysis tool papers and extend in a lesser proportion, to the crop or plant models (42.1, 32.7, 10.6 and 14.6 % for the plant phenotyping, image analysis tools, crop models and plant models, respectively; Table 1).

The image analysis tool community is characterized by a lower number of links per paper ( $190 \pm 221$ ). A high proportion involved the image analysis tool and plant phenotyping papers and, in a lesser proportion, the crop or plant models (27.5, 50.2, 7.3 and 15.1 % for the plant phenotyping, image analysis tools, crop models and plant models, respectively).

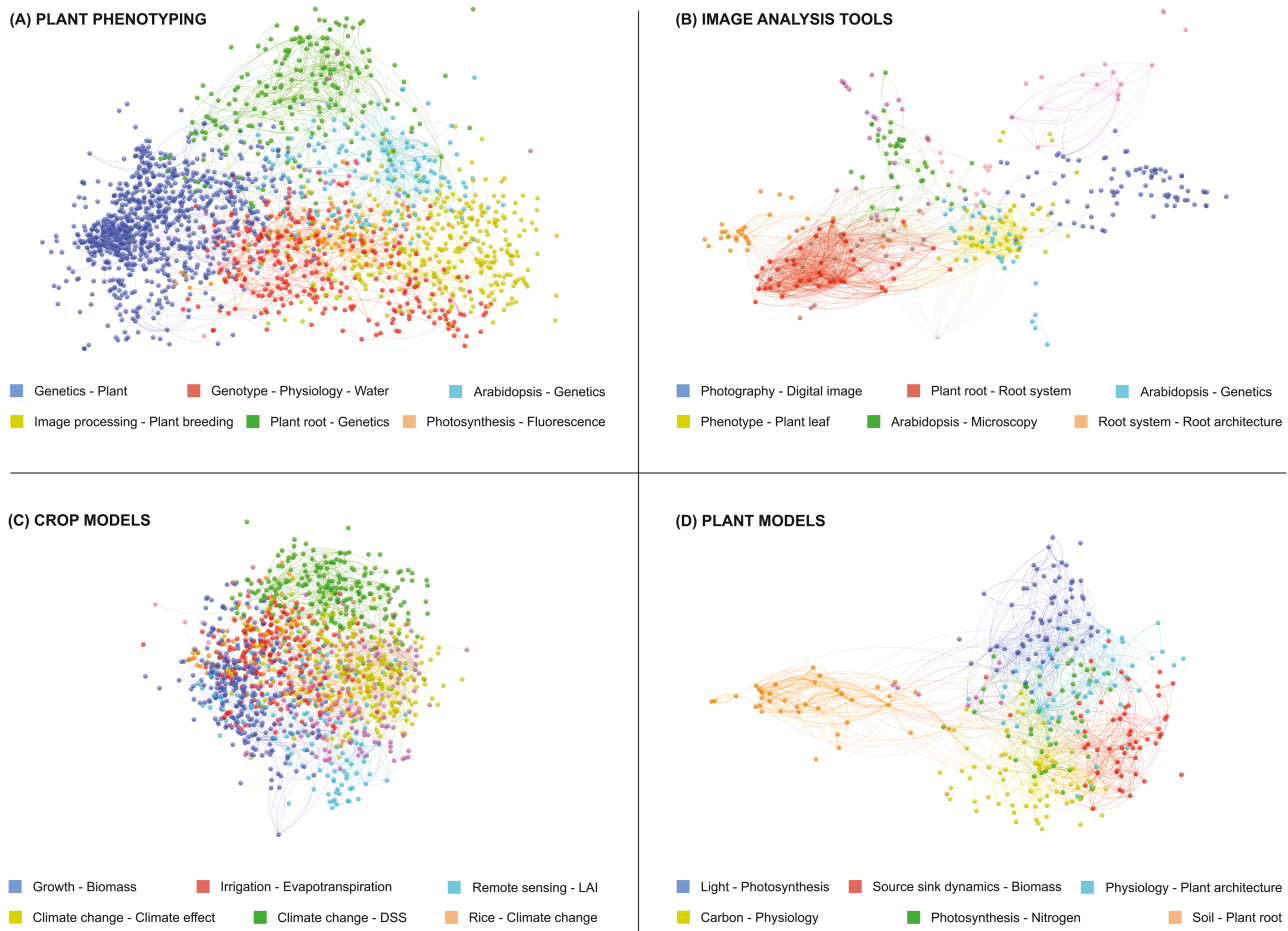
The crop model community is characterized by a narrow dispersion and a low number of links per paper ( $227 \pm 183$ ). A low proportion of links involved the plant phenotyping and the image analysis tool communities (7.0 and 2.4 % for the plant phenotyping and the image analysis tools, respectively), but a higher proportion involved the plant model community (21.7 %). The strong proportion of links between crop model papers (69.0 %) highlights a high number of common references and a strong connection between papers of this community.

The plant model community is positioned between the plant phenotyping and the crop model communities (green group at the center of the map in the Fig. 4). It is characterized by a narrow dispersion and by a higher number of links compared to the other communities ( $479 \pm 268$ ). The proportion of the links with the other communities is very low (7.2, 7.1 and 12.2 %).

## 4. DISCUSSION

### 4.1 History and research structure of the communities

**4.1.1 The crop modelling community.** During the period 1980–84, 13 publications were associated to the crop modelling community, but the first paper was referenced by Scopus in 1971. We acknowledge that



**Figure 3.** Bibliography coupling map based on (A) plant phenotyping papers, (B) plant image analysis software tool papers, (C) process-based crop model papers and (D) individual-based model papers from the time slice 1980–2019. Dots of different colours represent articles belonging to different clusters. The connecting lines indicate the bibliography coupling links between articles. In general, the closer two articles are located to each other, the stronger their relation. The two most frequent terms of each paper cluster are given below each subplot.

papers indexed in the survey provides a partial representation of the crop modelling activity as our research may be biased by the Scopus database (i.e. the publications and the journals indexed) and the field (i.e. title, abstract and keywords) or the keyword used in Scopus queries. However, the above-mentioned period seems to correspond to the early stage of the crop modelling community described by Passioura (1996), Jones *et al.* (2017) and Keating and Thorburn (2018). This early establishment led to the development of a well-organized community. Several groups (e.g. APSIM, DSSAT, EPIC or STICS; Jones *et al.* 1991, 2003; Brisson *et al.* 1998; Holzworth *et al.* 2014), symposia (e.g. International Crop Modelling Symposium [iCROPM]; <https://www.icropm2020.org/>) and international consortia (e.g. Agricultural Model Intercomparison Project [AgMIP]; Rosenzweig *et al.* 2013) were initiated to develop, improve and evaluate models. For example, AgMIP was designed to improve the capacity of models to describe the potential impacts of climate change on agriculture systems (Rosenzweig *et al.* 2013) and involved the collaboration of diverse crop modelling groups (e.g. APSIM, CropSyst, EPIC or WOFOST).

An important outcome of this project was the development of a platform that facilitates researcher collaborations from many organizations, across many countries (Porter *et al.* 2014).

**4.1.2 The plant phenotyping community.** During the last decade, research activity on plant phenotyping increased exponentially. This can be largely attributed to the emergence of the phenotyping platforms, the appearance of more complex technologies and the increasing availability of powerful sensors to address the urgent need for structural, physiological and performance-related plant traits to ensure food security in the coming decades (Tardieu *et al.* 2017). Indeed, major improvements in crop yield are needed to maintain suitable levels of agricultural production in spite of soil degradation and climate change. Over the past decade, the improvement of crop resistance and resilience to biotic and abiotic stresses has benefited from advances in genomic technologies (e.g. low-cost genome sequencing). Unfortunately, the characterization of the structure and function of the plant associated with its genetic and environmental

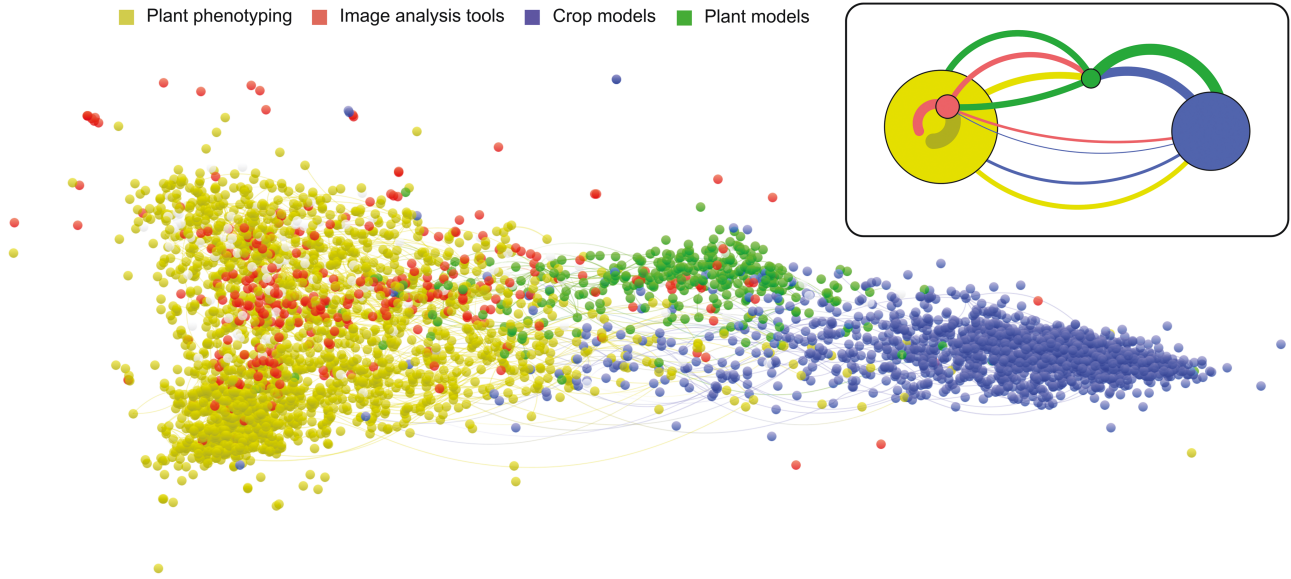


Figure 4. Combined bibliography coupling map based on plant phenotyping papers, plant image analysis software tool papers, process-based crop model papers and individual-based model papers from the time slice 1980–2019. Different colours represent the papers belonging to different communities. The connecting lines indicate the 1000 strongest bibliography coupling links between articles. In general, the closer two articles are located to each other, the stronger their relation. The black box at the top right summarizes the cross-links between communities, where the size of the disks and width of the lines stand for the total number of papers and the mean bibliography coupling links between communities, respectively. Connecting lines have the colour of the citing community.

Table 1. Mean characteristics of the papers and their associated proportion with the other papers.

Communities	Plant phenotyping	Image analysis tools	Crop models	Plant models
<b>Number of papers</b>	<b>2035</b>	<b>422</b>	<b>1405</b>	<b>343</b>
<b>Mean link per papers</b>	<b>248.90 ± 252.09</b>	<b>190.33 ± 221.09</b>	<b>227.12 ± 182.74</b>	<b>478.60 ± 267.65</b>
Proportion of link per community (%):				
Plant phenotyping	42.12	27.46	6.95	7.16
Image analysis tools	32.70	50.17	2.36	7.10
Crop models	10.55	7.29	68.97	12.19
Plant models	14.63	15.09	21.72	73.55

components remained one of the main technical challenges in research programs (Coppens et al. 2017).

The urgency to address the need for adaptation of agricultural systems to environmental challenges requires collaborative efforts between communities communicating efficiently. Faced with this challenge, several national (e.g. German Plant Phenotyping Network, French Plant Phenotyping Network or North American Plant Phenotyping Network) and international (e.g. International Plant Phenotyping Network [IPPN], European Plant Phenotyping Network 2020 [EPPN2020] or EMPHASIS) infrastructures were initiated to foster the development of novel scientific concepts, sensors and integrated models or phenotyping platforms (Roy et al. 2017).

**4.1.3 The image analysis tool community.** Plant scientists have produced massive datasets involving billions of images during the last decade (Furbank and Tester 2011; Fiorani and Schurr 2013). Indeed,

images provide information about the structure and the physiological status of the plant for the scientists (e.g. shape, colour, growth, transpiration or light received) at different spatial and temporal scales (e.g. leaf expansion rate of an individual organ or canopy expansion of a population; Dhondt et al. 2013; Coppens et al. 2017). Moreover, they can be produced for a large diversity of species (e.g. annual or perennial species) in experiments performed in controlled conditions or in the field, using automatic image recording (Tardieu et al. 2017; Neveu et al. 2019). This massive and diverse amount of images produced during the last decade called for the development of a variety of image analysis software tools dedicated to data extraction and analyse (e.g. quantify the morphological shoot traits; Lobet et al. 2013).

**4.1.4 Plant modelling community.** According to the survey, plant models emerged in 1988 and the publication of papers remained stable until 2005. The first structural plant models aimed to simulate the

diversity of the shoot and root architectures (e.g. Pagès *et al.* 1989; Prusinkiewicz *et al.* 1996; Lynch *et al.* 1997; Godin and Caraglio 1998). Their development was based on newly recognized botanical knowledge (Hallé 1986; Fitter 1987; Atger 1991). After this initial period, the plant models became more and more complex, describing the physiological processes and the endogenous (e.g. the interactions between the different organs of the plant) and exogenous (i.e. the interactions between the plant and its abiotic and biotic environment) environments of the plant (Godin and Sinoquet 2005; Prusinkiewicz and Runions 2012; Dunbabin *et al.* 2013; Sievänen *et al.* 2014). This increased complexity has been allowed by the advent of computers and the availability of means of rapid computation (DeJong *et al.* 2011; Long *et al.* 2018). Several plant modelling groups have emerged around diverse plant parts (e.g. above- or belowground compartments) or modelling approaches (e.g. purely descriptive static representations of plants or highly mechanistic dynamic simulations). Despite this diversity, the plant modelling community has initiated collaborative efforts to reinforce connectivity between modellers in order to improve and evaluate models (e.g. a collaborative benchmarking of functional–structural root architecture models; Schnepf *et al.* 2020). In the same way, the organization of international conferences (e.g. ICPMA2021 or FSPM2020) and the development of modelling platforms (e.g. GroIMP or OpenAlea; Hemmerling *et al.* 2008; Pradal *et al.* 2008; Long *et al.* 2018) allowed promotion of model use, reuse and integration of model components designed by others (Sievänen *et al.* 2014). Moreover, reviews are often published to illustrate the relevance of this modelling approach at various scales in the fields of developmental biology and promote the interest of plant modelling to other researchers (e.g. Fourcaud *et al.* 2007; DeJong *et al.* 2011; Sievänen *et al.* 2014; Evers *et al.* 2018; Passot *et al.* 2019; Louarn and Song 2020).

#### 4.2 Assessment of research topics compatibility and connection between the communities

The analysis of the most common terms used by the communities highlights a high compatibility between the plant phenotyping and image analysis tool communities. Most of the research topics identified in the image analysis tool community are also observed in the plant phenotyping community, in accordance with the connection observed between these communities. This connection is observed throughout the positions of the papers in the bibliographic coupling map (i.e. the image analysis tool and plant phenotyping papers are largely mingled; Fig. 4) and the number and proportion of links shared between these communities (Table 1). The common use of image analysis tools by the plant phenotyping community could explain these positions and the links with the plant phenotyping. Current phenotyping pipelines often rely on imaging techniques, becoming the major tool for phenotypic trait measurement (Dhondt *et al.* 2013).

In another way, compatibilities of research topics differ with the connection between communities. Despite the large range of topics addressed by plant models and their compatibilities with the plant phenotyping community, the number and proportion of links shared between communities are low. This result is likely due to the different scientific goals of these communities. On the one hand, the plant phenotyping community aims to identify structural, functional and

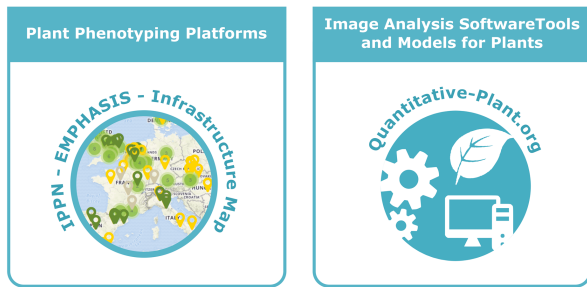
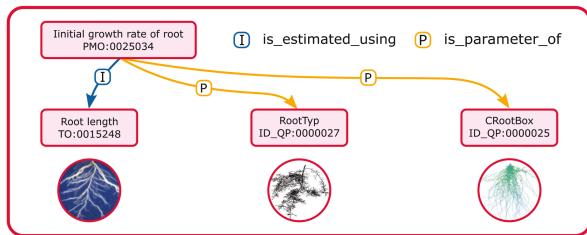
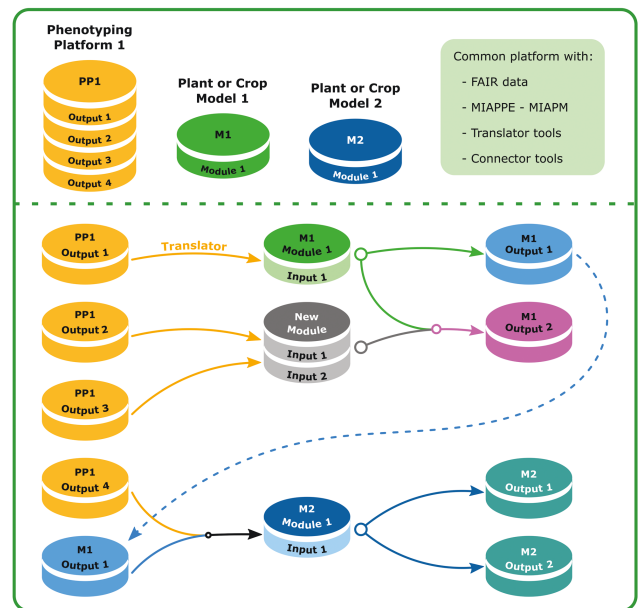
genetic traits for plant breeding purposes. On another side, the plant modelling community aims to describe and understand plant development and its interactions with the environment.

An interesting result relates to the position of the plant model community, which occupies a central position in the bibliographic coupling map. It is positioned between the plant phenotyping and the crop model communities. This finding is not trivial as also discussed by Louarn and Song (2020). During the last decade, the plant modelling community has expanded greatly its research area. Plant scientists have designed plant models simulating structural and functional processes at different scales (i.e. from the cell to the plant communities), for annual and perennial species (e.g. *Arabidopsis*, maize, wheat, mango or palm plantation; Barillot *et al.* 2016; Boudon *et al.* 2020; Perez *et al.* 2019) with the possibility to consider the endogenous and the exogenous environment of the plant. Consequently, plant models are positioned at the crossroads of plant phenotyping and crop model communities. The future challenge could be to improve its connection with the other communities to design a well-connected network.

The compatibility between research topics of the crop model community and the other communities is low. This result can be attributed to the scientific goals and spatial scale of this community. Compared to other communities, crop models aim to predict yield and potential impacts of climate change on agriculture systems. This community represents crop characteristics at the field scale compared to plant models or phenotyping platforms where each plant is considered at various levels of architectural realism. However, the application of plant phenotyping in the field is under rapid development (Costa *et al.* 2019). This emerging topic was not identified in our bibliographic analysis but is likely to be compatible and connectable with the research topics identified in the crop model community. Moreover, terms extracted from papers to identify research topics represent a simplified version of reality and a partial representation of the investigated field. For example, the terms indexed in the keywords of papers depend on the paper objectives. Most often, terms extracted from crop model papers highlighted the context of the study (e.g. climate change or climate effect; applied sciences; Fig. 3) and less the calibration or physiological processes integrated by the models, as observed for the plant model papers (e.g. source sink dynamics or photosynthesis; fundamental sciences; Fig. 3). In this way, the difference in term indexation can bias the identification of compatibility between the crop model community and the other communities.

#### 4.3 Towards an interoperable phenotyping–modelling framework: the EMPHASIS guidelines

Previously, we identified a well-established connection between the plant phenotyping and the image analysis tool communities. However, although phenomics and modelling crucially need to exchange data, the two communities appeared as weakly related. We hypothesize that this lack of connection is attributed to the lack of awareness of the benefits promoted by each community, the heterogeneous terminology used by the communities and the lack of common platforms to enable transparent data exchange. Here, we present a strategy to move towards better connection and collaboration between phenotyping and modelling communities. The framework, composed of three strategic axes, is presented in Fig. 5.

**(A) Knowing what exists: Awareness****(B) Improve semantic to find compatibilities****(C) Combining what is compatible: Interoperability**

**Figure 5.** Strategies suggested to facilitate transparent data exchange from phenotyping platform experiments to models and vice versa. (A) Promoting the interest and raising the awareness of the diversity of models, phenomics datasets and phenotyping platforms by maintaining and advertising the IPPN and EMPHASIS databases, and the [quantitative-plant.org](https://www.quantitative-plant.org) web-based repository. (B) Improving the lexical and semantic interoperability between communities by designing a structured controlled vocabulary for the plant and crop modelling communities arranged in a new ontology (Plant and Crop Modelling Ontology [PCMO]). (C) Developing a common hosting platform considering (i) phenomics data in a harmonized format, (ii) phenomics data with their associated metadata and (iii) model with their associated translators and connectors to allow the connection between phenomics data and other models.

**4.3.1 Promote the interest and raise awareness.** It is currently challenging for phenomics researchers to become aware of the diversity of models and their applications (Fig. 5A). To solve this first challenge, we developed the ‘Quantitative Plant’ online portal ([quantitative-plant.org](https://www.quantitative-plant.org)) allowing the exploration of the diversity of >100 plant and crop simulation models and their applications. This web-based repository extends the ‘Plant Image Analysis’ website (an online database for plant image analysis software tools; Lobet et al. 2013). It helps researchers in search for image analysis software for their phenotyping experiments to find out potential game-changing model applications on the associated crop or plant model webpage. From now on, the objectives will be to maintain and update this online portal and promote the interest for modelling approaches by advertising this website within the phenomics community (e.g. at conferences, workshops and symposia).

A second challenge is for the modelling community to become aware of the phenomics datasets and phenotyping platforms. To address this second challenge, a mapping exercise was carried out by IPPN ([https://www.plant-phenotyping.org/infrastructure\\_map](https://www.plant-phenotyping.org/infrastructure_map)) and EMPHASIS ([https://emphasis.plant-phenotyping.eu/emphasis\\_infrastructure\\_map](https://emphasis.plant-phenotyping.eu/emphasis_infrastructure_map)) thereby increasing the visibility of phenotyping platforms, in collaboration with national plant phenotyping communities. Surveys were carried out to extract detailed information on existing and upcoming infrastructures. In addition, platform characteristics

were inventoried throughout workshops organized in different regions of Europe and the World. Phenotyping platforms were described by their general characteristics (e.g. installation category: high-throughput phenotyping facility) and uses (e.g. trait measurements: root properties). In this way, the IPPN and EMPHASIS databases provide an overview of available plant phenotyping platforms and their associated characteristics enabling users to identify available solutions for their project.

**4.3.2 Improve lexical and semantic interoperability.** Currently, the terminology (e.g. objects, variables) used by phenotyping and modelling communities can be quite heterogeneous depending on the research discipline, scale, objectives and even between research groups (Fig. 5B). This limits the ability to accurately relate information within and across communities. A solution to facilitate the connection and the exchange of information is the use of a controlled and standardized dictionary of common and internationally recognized terms that can be shared among the communities (Walls et al. 2012). The phenomics community has tackled these issues by adopting semantic web technologies including the use of ontologies (e.g. Plant Ontology [PO] or Plant Trait Ontology [TO]; Cooper et al. 2013, 2016, 2018). However, no ontology describing the variable inputs, the variable outputs and the parameters of the plant and crop models exists. Such ontology could be used to facilitate exchange in and across communities, like



has been done in the bioinformatic community through the Elixir project (e.g. the EDAM ontology; Ison *et al.* 2013). One solution to this problem involves the development of structured controlled vocabularies for the plant and crop modelling communities arranged in a new ontology (Plant and Crop Modelling Ontology [PCMO]). The goal of the PCMO would be to produce structured controlled vocabularies of the variables and parameters used by mathematical and computational models (plant and crop models) with clear definitions and relations with the existing phenomics ontologies (e.g. description of the phenomics variables used in the parameter estimations). In addition of the benefit to find compatibilities between phenomics datasets and models using ontologies, the PCMO would facilitate the connection between the models themselves, promoting the design of modular models (Christensen *et al.* 2018; Passot *et al.* 2018; Benes *et al.* 2020; Peng *et al.* 2020) or the intercomparison of models (Athanasiadis *et al.* 2009; Porter *et al.* 2014; Schnepf *et al.* 2020).

**4.3.3 Simplify the translation.** In the future, a long-term cooperation between the phenomics and modelling communities towards the development of common platforms could be designed to enable transparent data exchange from models to experiments and *vice versa* (Fig. 5C). However, designing such platform is particularly challenging due to the diversity and volume of model and data involved. Indeed, the development of this platform involves connecting plant or crop models from different fields of research (i.e. with different syntax, semantics and inputs), integrating the huge amount of data generated by phenotyping platforms from different sensors (e.g. laser scanning systems, x-ray micro-computed tomography, magnetic resonance imaging or hyperspectral cameras) at different scales (e.g. individual plant in controlled conditions or plant population in the field) and levels of organization (e.g. cell, tissues, organ, plant and population), and analyzing and evaluating the newly designed system with numerical experiment. To face these challenges, a common hosting platform should:

1. Find and store phenomics data with their associated metadata. Recent and collaborative efforts were made by the plant community to provide access to phenomics and genotypic data. A common application programming interface (i.e. the public plant Breeding Application Programming Interface [BrAPI]; Selby *et al.* 2019) was created to improve interoperability between heterogeneous data repositories for breeding data. The implementation of this API enables data search in different systems and facilitates the integration of data from different disciplines. Moreover, the correct interpretation, replicability, comparability and interoperability of data rely on an adequate set of integrated metadata standards, which list the fields required for interpreting the data from a given experiment (Ćwiek-Kupczyńska *et al.* 2016). To address this problem, the plant community implemented a metadata model in BrAPI, facilitating the metadata implementation from different disciplines (e.g. phenotyping or genotyping data). Next to BrAPI, the MIAPPE initiative (Krajewski *et al.* 2015; Papoutsoglou *et al.* 2020), i.e. 'Minimum Information About a Plant Phenotyping Experiment', had produced a document

which specifies default metadata requirements for phenomics experiments and a Phenotyping Configuration for the ISA-Tab format, which allows one to practically organize this information within a dataset.

2. Store phenomics data in a harmonized format. Several phenotyping data management systems have been proposed for integrating, managing and sharing multi-source and multi-scale data in plant phenomics experiments for both controlled and field conditions (Köhl and Gremmels 2015; Reynolds *et al.* 2019; Honecker *et al.* 2020). For example, the Phenomics Ontology Driven Data (PODD) repository is an information system designed to support phenomics research in Australia. It aims to enable efficient storage and retrieval of plant phenotyping data and metadata generated by the Australian facilities (Li *et al.* 2013). Similarly, a central web interface and database was developed in Belgium, the Plant Systems Biology Interface for Plant Phenotype Analysis (PIPPA; [https://pippa.psb.ugent.be/pippa\\_nav/home/](https://pippa.psb.ugent.be/pippa_nav/home/)). PIPPA enables management of different types of plant phenotyping robots (e.g. WIWAM; <https://www.wiwam.be/>) and analyzing the huge datasets generated (Coppens *et al.* 2017). More recently, the open-source Phenotyping Hybrid Information System (PHIS) was developed and made available to the phenomics community (Neveu *et al.* 2019; [www.phis.inra.fr](http://www.phis.inra.fr)). Compared to most information systems, PHIS has been designed for integrating and sharing multi-source and multi-scale data from various phenotyping categories of installations (e.g. field or greenhouse).
3. Define a list of variables that might be necessary for modelling. Most plant and crop models are based on similar concepts, although they currently lack the ability to share similar variable and parameter inputs. While the various models may implement different algorithms, the driving variables are generally similar. A basic description of the variables used in model calibration and parametrization could support modelling applications and connection between the phenomics and the modelling community. This work will promote recommendations about the general information for modelling by defining a list of variables that might be necessary to use models. A document defining these variables should rather be considered as a checklist and recommendations, and consulted by a phenotyping researcher interested in modelling to ensure the inclusion of important variables for modelling purposes.
4. Promote the development of input translator tools to allow the generation of model-ready input files from harmonized phenomics datasets. Despite the fact that most plant and crop models are based on similar concepts and driving data, variables and parameters can be different and can have quite heterogeneous specifications, as similar processes can have different representations and mathematical implementations according to the models (Boote *et al.* 2013). In order to facilitate the connection between phenomics variables and model inputs, translating data to variable and parameter inputs is required for each model. As it was done by the AgMIP project, each model could develop a model-specific input

translator, which allows harmonized phenomics dataset to directly be translated to model inputs (Porter et al. 2014).

5. Develop connector tools to allow the connection between models. More than hundred plant and crop models have been created in the last two decades. All these models have the potential to be reused and combined, broadening the scope of their original uses. Connecting such models into integration networks has the potential to integrate more complex with isolated models (Passot et al. 2019; Long 2019). The integration of several models should be used to generate new outputs (e.g. integrated or more complex plant traits), improve model predictions or design new strategies (e.g. a multi-scale [from gene to globe] crop modelling framework; Benes et al. 2020; Peng et al. 2020). However, the technological barriers introduced by differences in language, data formats, spatial and temporal scales, and units have slowed this progress. AgMIP has provided a necessary important first step in bringing disparate models of each major crop together (Rosenzweig et al. 2013). The plant model community has also designed modelling platforms (e.g. OpenAlea; Pradal et al. 2008) that make sharing of models increasingly feasible. Moreover, recent efforts have been made by the plant science and crop modelling communities, such as the Crop in silico project (Cis; Marshall-Colon et al. 2017) or the collaboration between OpenAlea and GroIMP modelling platforms (Long et al. 2018). For example, the yggdrasil framework (Lang 2019) was developed to facilitate asynchronous connection among models written in different languages. It operates at different scales, resolving the historical problems associated with integrative and multi-scale modelling (Kannan et al. 2019). More recently, a centralized framework (Crop2ML; Midingoyi et al. 2021) and a new derived language (CyML; Midingoyi et al. 2020) were created throughout the Agricultural Model Exchange Initiative (AMEI) for exchanging, reusing and assembling models and model components. These types of developments will speed up model construction and the creation of application-oriented models and facilitate linkage of different types of models. These recent efforts are promising and should be encouraged to facilitate the connection between plant and crop models.

## 5. CONCLUSION

In the present analysis, we identified a well-established connection between the plant phenotyping and the image analysis tool communities. However, the connection between phenomics and modelling communities was low despite their research topics compatibilities. We hypothesize that this lack of connection is attributed to a lack of awareness of the benefits promoted by each community, the heterogeneous terminology used by the communities and the lack of common platforms to enable transparent data exchange.

In the framework of the EMPHASIS project, strategies were proposed to move towards better communication and collaboration between phenotyping and modelling communities. Firstly, we suggest raising the awareness of the diversity of models, phenomics

datasets and phenotyping platforms that exist by maintaining and advertising online databases. Secondly, we suggest improving the lexical and semantic interoperability between communities by designing a structured controlled vocabulary for the plant and crop modelling communities arranged in a new ontology (PCMO). Thirdly, we suggest the development of a common hosting platform considering (i) phenomics data with their associated metadata, (ii) phenomics data in a harmonized format and (iii) model with their associated translators and connectors to allow the connection between phenomics data and other models.

## SUPPORTING INFORMATION

The following additional information is available in the online version of this article –

**Figure S1.** VOSviewer Map and Network files.

**Table S1.** Complete lists of term identified in each cluster.

## ACKNOWLEDGEMENTS

We thank the European Commission, which supported this work through the grant numbers 739514 (EMPHASIS) and 731013 (EPPN2020). We thank the editor and the two anonymous reviewers for their valuable comments and suggestions. We also thank Sarah Cookson for her helpful comments and English revision.

## SOURCES OF FUNDING

This work was financially supported by the European Commission, through the grant numbers 739514 (EMPHASIS) and 731013 (EPPN2020). V.C. was funded as Research Fellow by the Belgian Fonds de la Recherche Scientifique (F.R.S.-FNRS, grant number: 1208619F).

## CONTRIBUTIONS BY THE AUTHORS

C.S.C., X.D., G.L. and V.C. contributed to the conception and design of the study. C.S.C. collected the data and performed the bibliometric analysis. C.S.C., X.D., G.L., V.C., C.P., L.C.-B. and F.T. contributed to data analysis and interpretation, and writing the manuscript.

## CONFLICTS OF INTEREST

None declared.

## LITERATURE CITED

- Aria M, Cuccurullo C. 2017. Bibliometrix: an R-tool for comprehensive science mapping analysis. *Journal of Informetrics* **11**:959–975.
- Atger C. 1991. L'architecture racinaire est-elle influencée par le milieu? In: Edelin C, ed. *L'arbre: biologie et développement*. Montpellier: Naturalia Monspeliensia, 71–84.
- Athanasiadis IN, Rizzoli A-E, Janssen S, Andersen E, Villa F. 2009. Ontology for seamless integration of agricultural data and models. In: Sartori F, Sicilia MÁ, Manouselis N, eds. *Metadata and semantic research*. Berlin and Heidelberg, Germany: Springer Berlin Heidelberg, 282–293.
- Barillot R, Chambon C, Andrieu B. 2016. CN-Wheat, a functional-structural model of carbon and nitrogen metabolism in wheat

- culms after anthesis. I. Model description. *Annals of Botany* **118**:997–1013.
- Benes B, Guan K, Lang M, Long SP, Lynch JP, Marshall-Colón A, Peng B, Schnable J, Sweetlove LJ, Turk MJ. 2020. Multiscale computational models can guide experimentation and targeted measurements for crop improvement. *The Plant Journal* **103**:21–31.
- Boote KJ, Jones JW, White JW, Asseng S, Lizaso JJ. 2013. Putting mechanisms into crop production models. *Plant, Cell & Environment* **36**:1658–1672.
- Boudon F, Persello S, Jestin A, Briand AS, Grechi I, Fernique P, Guédon Y, Léchaudel M, Lauri PE, Normand F. 2020. V-Mango: a functional–structural model of mango tree growth, development and fruit production. *Annals of Botany* **126**:745–763.
- Boyack KW, Klavans R. 2010. Co-citation analysis, bibliographic coupling, and direct citation: which citation approach represents the research front most accurately? *Journal of the American Society for Information Science and Technology* **61**:2389–2404.
- Brisson N, Mary B, Ripoche D, Jeuffroy MH, Ruget F, Nicoulaud B, Gate P, Devienne-Barret F, Antonioletti R, Durr C, Richard G, Beaudoin N, Recous S, Tayot X, Plenet D, Cellier P, Machet JM, Meynard JM, Delécolle R. 1998. STICS: a generic model for the simulation of crops and their water and nitrogen balances. I. Theory and parameterization applied to wheat and corn. *Agronomie* **18**:311–346.
- Casadebaig P, Debaeke P, Wallach D. 2020. A new approach to crop model calibration: phenotyping plus post-processing. *Crop Science* **60**:709–720.
- Chang Y-W, Huang M-H. 2012. A study of the evolution of interdisciplinarity in library and information science: using three bibliometric methods. *Journal of the American Society for Information Science and Technology* **63**:22–33.
- Chen TW, Cabrera-Bosquet L, Alvarez Prado S, Perez R, Artzet S, Pradal C, Coupel-Ledru A, Fournier C, Tardieu F. 2019. Genetic and environmental dissection of biomass accumulation in multi-genotype maize canopies. *Journal of Experimental Botany* **70**:2523–2534.
- Christensen AJ, Srinivasan V, Hart JC, Marshall-Colon A. 2018. Use of computational modeling combined with advanced visualization to develop strategies for the design of crop ideotypes to address food security. *Nutrition Reviews* **76**:332–347.
- Cooper L, Walls RL, Elser J, Gandolfo MA, Stevenson DW, Smith B, Preece J, Athreya B, Mungall CJ, Rensing S, Hiss M, Lang D, Reski R, Berardini TZ, Li D, Huala E, Schaeffer M, Menda N, Arnaud E, Shrestha R, Yamazaki Y. 2013. The plant ontology as a tool for comparative plant anatomy and genomic analyses. *Plant and Cell Physiology* **54**:e1.
- Cooper L, Meier A, Laporte MA, Elser JL, Mungall C, Sinn BT, Cavaliere D, Carbon S, Dunn NA, Smith B, Qu B, Preece J, Zhang E, Todorovic S, Gkoutos G, Doonan JH, Stevenson DW, Arnaud E, Jaiswal P. 2018. The Planteome database: an integrated resource for reference ontologies, plant genomics and phenomics. *Nucleic Acids Research* **46**:D1168–D1180.
- Cooper M, Technow F, Messina C, Ghossein C, Radu Totir L. 2016. Use of crop growth models with whole-genome prediction: application to a maize multi-environment trial. *Crop Science* **56**:2141–2156.
- Coppens F, Wuyts N, Inzé D, Dhondt S. 2017. Unlocking the potential of plant phenotyping data through integration and data-driven approaches. *Current Opinion in Systems Biology* **4**:58–63.
- Costa C, Schurr U, Loreto F, Menesatti P, Carpentier S. 2019. Plant phenotyping research trends, a science mapping approach. *Frontiers in Plant Science* **9**:1933.
- Ćwiek-Kupczyńska, H., Altmann, T., Arend, D., Arnaud, E., Chen, D., Cornut, G., Fiorani, F., Frohberg, W., Junker, A., Klukas, C., Lange, M., Mazurek, C., Nafissi, A., Neveu, P., Oeveren, J., Pommier, C., Poorter, H., Rocca-Serra, P., Sansone, S.-A., Scholz, U., Schriek, M., Seren, Ü., Usadel, B., Weise, S., Kersey, P., Krajewski, P. 2016. Measures for interoperability of phenotypic data: minimum information requirements and formatting. *Plant Methods* **12**:44. doi:10.1186/s13007-016-0144-4.
- DeJong TM, Da Silva D, Vos J, Escobar-Gutiérrez AJ. 2011. Using functional–structural plant models to study, understand and integrate plant development and ecophysiology. *Annals of Botany* **108**:987–989.
- Dhondt S, Wuyts N, Inzé D. 2013. Cell to whole-plant phenotyping: the best is yet to come. *Trends in Plant Science* **18**:428–439.
- Dunbabin VM, Postma JA, Schnepf A, Pagès L, Javaux M, Wu L, Leitner D, Chen YL, Rengel Z, Diggler AJ. 2013. Modelling root–soil interactions using three-dimensional models of root growth, architecture and function. *Plant and Soil* **372**:93–124.
- Evers JB, Letort V, Renton M, Kang M. 2018. Computational botany: advancing plant science through functional–structural plant modelling. *Annals of Botany* **121**:767–772.
- Fiorani F, Schurr U. 2013. Future scenarios for plant phenotyping. *Annual Review of Plant Biology* **64**:267–291.
- Fitter AH. 1987. An architectural approach to the comparative ecology of plant root systems. *New Phytologist* **106**:61–77.
- Fourcaud T, Zhang X, Stokes A, Lambers H, Korner C. 2007. Plant growth modelling and applications: the increasing importance of plant architecture in growth models. *Annals of Botany* **101**:1053–1063.
- Furbank RT, Tester M. 2011. Phenomics—technologies to relieve the phenotyping bottleneck. *Trends in Plant Science* **16**:635–644.
- Godin C, Caraglio Y. 1998. A multiscale model of plant topological structures. *Journal of Theoretical Biology* **191**:1–46.
- Godin C, Sinoquet H. 2005. Functional–structural plant modelling: Commentary. *New Phytologist* **166**:705–708.
- Gosseau F, Blanchet N, Varès D, Burger P, Campergue D, Colombet C, Gody L, Liévin J-F, Mangin B, Tison G, Vincourt P, Casadebaig P, Langlade N. 2019. Heliaphen, an outdoor high-throughput phenotyping platform for genetic studies and crop modeling. *Frontiers in Plant Science* **9**:1908.
- Hallé F. 1986. Modular growth in seed plants. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences* **313**:77–87.
- Hemmerling R, Kniemeyer O, Lanwert D, Kurth W, Buck-Sorlin G. 2008. The rule-based language XL and the modelling environment GroIMP illustrated with simulated tree competition. *Functional Plant Biology* **35**:739.
- Holzworth DP, Huth NI, deVoil PG, Zurcher EJ, Herrmann NI, McLean G, Chenu K, van Oosterom EJ, Snow V, Murphy C, Moore AD, Brown H, Whish JPM, Verrall S, Fainges J, Bell LW, Peake AS, Poulton PL, Hochman Z, Thorburn PJ, Gaydon DS, Dalgliesh NP, Rodriguez D, Cox H, Chapman S, Doherty A, Teixeira E, Sharp J, Cichota R, Vogelger I, Li FY, Wang E, Hammer GL, Robertson MJ, Dimes JP, Whitbread AM, Hunt J, van Rees H, McClelland T, Carberry PS, Hargreaves JNG,

- MacLeod N, McDonald C, Harsdorf J, Wedgwood S, Keating BA. 2014. APSIM—evolution towards a new generation of agricultural systems simulation. *Environmental Modelling & Software* **62**:327–350.
- Honecker A, Schumann H, Becirevic D, Klingbeil L, Volland K, Forberig S, Jansen M, Paulsen H, Kuhlmann H, Léon J. 2020. Plant, space and time—linked together in an integrative and scalable data management system for phenomic approaches in agronomic field trials. *Plant Methods* **16**:55.
- Ison J, Kalas M, Jonassen I, Bolser D, Uludag M, McWilliam H, Malone J, Lopez R, Pettifer S, Rice P. 2013. EDAM: an ontology of bioinformatics operations, types of data and identifiers, topics and formats. *Bioinformatics* **29**:1325–1332.
- Jones CA, Dyke PT, Williams JR, Kiniry JR, Benson VW, Griggs RH. 1991. EPIC: an operational model for evaluation of agricultural sustainability. *Agricultural Systems* **37**:341–350.
- Jones JW, Hoogenboom G, Porter CH, Boote KJ, Batchelor WD, Hunt LA, Wilkens PW, Singh U, Gijsman AJ, Ritchie JT. 2003. The DSSAT cropping system model. *European Journal of Agronomy* **18**:235–265.
- Jones JW, Antle JM, Basso B, Boote KJ, Conant RT, Foster I, Godfray HCJ, Herrero M, Howitt RE, Janssen S, Keating BA, Munoz-Carpena R, Porter CH, Rosenzweig C, Wheeler TR. 2017. Brief history of agricultural systems modeling. *Agricultural Systems* **155**:240–254.
- Kannan K, Wang Y, Lang M, Challa GS, Long SP, Marshall-Colon A. 2019. Combining gene network, metabolic and leaf-level models shows means to future-proof soybean photosynthesis under rising CO<sub>2</sub>. *In Silico Plants* **1**. doi:10.1093/insilicoplants/diz008
- Keating BA, Thorburn PJ. 2018. Modelling crops and cropping systems—evolving purpose, practice and prospects. *European Journal of Agronomy* **100**:163–176.
- Kessler MM. 1963. Bibliographic coupling between scientific papers. *American Documentation* **14**:10–25.
- Köhl K, Gremmels J. 2015. A software tool for the input and management of phenotypic data using personal digital assistants and other mobile devices. *Plant Methods* **11**:25.
- Krajewski P, Chen D, Ćwiek H, van Dijk ADJ, Fiorani F, Kersey P, Klukas C, Lange M, Markiewicz A, Nap JP, van Oeveren J, Pommier C, Scholz U, van Schriek M, Usadel B, Weise S. 2015. Towards recommendations for metadata and data handling in plant phenotyping. *Journal of Experimental Botany* **66**:5417–5427.
- Lang M. 2019. yggdrasil: a Python package for integrating computational models across languages and scales. *In Silico Plants* **1**. doi:10.1093/insilicoplants/diz001
- Li Y-F, Kennedy G, Ngoran F, Wu P, Hunter J. 2013. An ontology-centric architecture for extensible scientific data management systems. *Future Generation Computer Systems* **29**:641–653.
- Liu G, Zhao Y, Mirdita V, Reif JC. 2017. Efficient strategies to assess yield stability in winter wheat. *Theoretical and Applied Genetics* **130**:1587–1599.
- Lobet G. 2017. Image analysis in plant sciences: publish then perish. *Trends in Plant Science* **22**:559–566.
- Lobet G, Draye X, Périlleux C. 2013. An online database for plant image analysis software tools. *Plant Methods* **9**:38.
- Long Q, Kurth W, Pradal C, Migault V, Pallas B. 2018. An architecture for the integration of different functional and structural plant models. In: *Proceedings of the 7th International Conference on Informatics, Environment, Energy and Applications—IEEA'18*. Beijing, China: ACM Press, 107–113.
- Long SP. 2019. Making our plant modelling community more than the sum of its parts: a personal perspective. *In Silico Plants* **1**. doi:10.1093/insilicoplants/diy002
- Louarn G, Song Y. 2020. Two decades of functional–structural plant modelling: now addressing fundamental questions in systems biology and predictive ecology. *Annals of Botany* **126**:501–509.
- Lynch JP, Nielsen KL, Davis RD, Jablono AG. 1997. SimRoot: modelling and visualization of root systems. *Plant and Soil* **188**:139–151.
- Marshall-Colon A, Long SP, Allen DK, Allen G, Beard DA, Benes B, von Caemmerer S, Christensen AJ, Cox DJ, Hart JC, Hirst PM, Kannan K, Katz DS, Lynch JP, Millar AJ, Panneerselvam B, Price ND, Prusinkiewicz P, Raila D, Shekar RG, Shrivastava S, Shukla D, Srinivasan V, Stitt M, Turk MJ, Voit EO, Wang Y, Yin X, Zhu X-G. 2017. Crops *in silico*: generating virtual crops using an integrative and multi-scale modeling platform. *Frontiers in Plant Science* **8**:786.
- Midingoyi CA, Pradal C, Athanasiadis IN, Donatelli M, Enders A, Fumagalli D, Garcia F, Holzworth D, Hoogenboom G, Porter C, Raynal H, Thorburn P, Martre P. 2020. Reuse of process-based models: automatic transformation into many programming languages and simulation platforms. *In Silico Plants* **2**. doi:10.1093/insilicoplants/diaa007
- Midingoyi CA, Pradal C, Enders A, Fumagalli D, Raynal H, Donatelli M, Athanasiadis IN, Porter C, Hoogenboom G, Holzworth D, Garcia F, Thorburn P, Martre P. 2021. Crop2ML: an open-source multi-language modeling framework for the exchange and reuse of crop model components. *Environmental Modelling & Software* **142**:105055. doi:10.1016/j.envsoft.2021.105055
- Muller B, Martre P. 2019. Plant and crop simulation models: powerful tools to link physiology, genetics, and phenomics. *Journal of Experimental Botany* **70**:2339–2344.
- Nardi P, Matteo GD, Palahi M, Mugnozsa GS. 2016. Structure and evolution of Mediterranean forest research: a science mapping approach. *PLoS One* **11**:e0155016. doi:10.1371/journal.pone.0155016.
- Neveu P, Tireau A, Hilgert N, Nègre V, Mineau-Cesari J, Brichet N, Chapuis R, Sanchez I, Pommier C, Charnomordic B, Tardieu F, Cabrera-Bosquet L. 2019. Dealing with multi-source and multi-scale information in plant phenomics: the ontology-driven Phenotyping Hybrid Information System. *New Phytologist* **221**:588–601.
- Pagès L, Jordan MO, Picard D. 1989. A simulation model of the three-dimensional architecture of the maize root system. *Plant and Soil* **119**:147–154.
- Papoutsoglou EA, Faria D, Arend D, Arnaud E, Athanasiadis IN, Chaves I, Coppens F, Cornut G, Costa BV, Ćwiek-Kupczyńska H, Drosbeke B, Finkers R, Gruden K, Junker A, King GJ, Krajewski P, Lange M, Laporte M, Michotey C, Oppermann M, Ostler R, Poorter H, Ramírez-Gonzalez R, Ramšak Ž, Reif JC, Rocca-Serra P, Sansone S, Scholz U, Tardieu F, Uauy C, Usadel B, Visser RGF, Weise S, Kersey PJ, Miguel CM, Adam-Blondon A, Pommier C. 2020. Enabling reusability of plant phenomic datasets with MIAPPE 1.1. *New Phytologist* **227**:260–273.

- Parent B, Tardieu F. 2014. Can current crop models be used in the phenotyping era for predicting the genetic variability of yield of plants subjected to drought or high temperature? *Journal of Experimental Botany* **65**:6179–6189.
- Passioura JB. 1996. Simulation models: science, snake oil, education, or engineering? *Agronomy Journal* **88**:690–694.
- Passot S, Moreno-Ortega B, Moukouanga D, Balsera C, Guyomarc'h S, Lucas M, Lobet G, Laplaze L, Muller B, Guédon Y. 2018. A new phenotyping pipeline reveals three types of lateral roots and a random branching pattern in two cereals. *Plant Physiology* **177**:896–910.
- Passot S, Couvreur V, Meunier F, Draye X, Javaux M, Leitner D, Pagès L, Schnepf A, Vanderborght J, Lobet G. 2019. Connecting the dots between computational tools to analyse soil–root water relations. *Journal of Experimental Botany* **70**:2345–2357. doi:10.1093/jxb/ery361.
- Peng B, Guan K, Tang J, Ainsworth EA, Asseng S, Bernacchi CJ, Cooper M, Delucia EH, Elliott JW, Ewert F, Grant RF, Gustafson DI, Hammer GL, Jin Z, Jones JW, Kimm H, Lawrence DM, Li Y, Lombardozzi DL, Marshall-Colon A, Messina CD, Ort DR, Schnable JC, Vallejos CE, Wu A, Yin X, Zhou W. 2020. Towards a multiscale crop modelling framework for climate change adaptation assessment. *Nature Plants* **6**:338–348.
- Perez RPA, Fournier C, Cabrera-Bosquet L, Artzet S, Pradal C, Bricchet N, Chen T, Chapuis R, Welcker C, Tardieu F. 2019. Changes in the vertical distribution of leaf area enhanced light interception efficiency in maize over generations of selection. *Plant, Cell & Environment* **42**:2105–2119.
- Porter CH, Villalobos C, Holzworth D, Nelson R, White JW, Athanasiadis IN, Janssen S, Ripoche D, Cufi J, Raes D, Zhang M, Knapen R, Sahajpal R, Boote K, Jones JW. 2014. Harmonization and translation of crop modeling data to ensure interoperability. *Environmental Modelling & Software* **62**:495–508.
- Pradal C, Dufour-Kowalski S, Boudon F, Fournier C, Godin C. 2008. OpenAlea: a visual programming and component-based software platform for plant modelling. *Functional Plant Biology* **35**:751.
- Prusinkiewicz P, Hammel M, Hanan J, Mech R. 1996. L-systems: from the theory to visual models of plants. In: *Proceedings of the 2nd CSIRO Symposium on Computational Challenges in Life Sciences*. Citeseer, 1–32.
- Prusinkiewicz P, Runions A. 2012. Computational models of plant development and form: Tansley review. *New Phytologist* **193**:549–569. doi:10.1111/j.1469-8137.2011.04009.x
- R Development Core Team. 2019. *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org> (29 April 2022).
- Reynolds D, Ball J, Bauer A, Davey R, Griffiths S, Zhou J. 2019. CropSight: a scalable and open-source information management system for distributed plant phenotyping and IoT-based crop management. *GigaScience* **8**:giz009. doi:10.1093/gigascience/giz009.
- Rosenzweig C, Jones JW, Hatfield JL, Ruane AC, Boote KJ, Thorburn P, Antle JM, Nelson GC, Porter C, Janssen S, Asseng S, Basso B, Ewert F, Wallach D, Baigorria G, Winter JM. 2013. The Agricultural Model Intercomparison and Improvement Project (AgMIP): protocols and pilot studies. *Agricultural and Forest Meteorology* **170**:166–182.
- Roy J, Tardieu F, Tixier-Boichard M, Schurr U. 2017. European infrastructures for sustainable agriculture. *Nature Plants* **3**:756–758.
- Saint Cast C, Lobet G. 2019. Quantitative plant. <https://quantitative-plant.org/> (29 April 2022).
- Schnepf A, Black CK, Couvreur V, Delory BM, Doussan C, Koch A, Koch T, Javaux M, Landl M, Leitner D, Lobet G, Mai TH, Meunier F, Petrich L, Postma JA, Priesack E, Schmidt V, Vanderborght J, Vereecken H, Weber M. 2020. Call for participation: collaborative benchmarking of functional–structural root architecture models. The case of root water uptake. *Frontiers in Plant Science* **11**:316.
- Selby P, Abbeloos R, Backlund JE, Salido MB, Bauchet G, Benites-Alfaro OE, Birkett C, Calaminos VC, Carceller P, Cornut G, Costa BV, Edwards JD, Finkers R, Gao SY, Ghaffar M, Glaser P, Guignon V, Hok P, Kilian A, Lagare JEB, Lange M. 2019. BrAPI—an application programming interface for plant breeding applications. *Bioinformatics* **35**:4147–4155.
- Sievänen R, Godin C, DeJong TM, Nikinmaa E. 2014. Functional–structural plant models: a growing paradigm for plant studies. *Annals of Botany* **114**:599–603.
- Tardieu F, Cabrera-Bosquet L, Pridmore T, Bennett M. 2017. Plant phenomics, from sensors to knowledge. *Current Biology* **27**:R770–R783.
- van Eck NJ, Waltman L. 2010. Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics* **84**:523–538.
- van Raan AF. 2014. Advances in bibliometric analysis: research performance assessment and science mapping. In: Blockmans W, Engwall L, Weaire D, eds. *Bibliometrics: Use and Abuse in the Review of Research Performance*. Wenner-Gren International Series, Vol. 87. Chapter: Advances in bibliometric analysis: research performance assessment and science mapping. London: Portland Press Ltd, 17–28.
- Walls RL, Athreya B, Cooper L, Elser J, Gandolfo MA, Jaiswal P, Mungall CJ, Preece J, Rensing S, Smith B, Stevenson DW. 2012. Ontologies as integrative tools for plant science. *American Journal of Botany* **99**:1263–1275.
- Waltman L, van Eck NJ, Noyons ECM. 2010. A unified approach to mapping and clustering of bibliometric networks. *Journal of Informetrics* **4**:629–635.
- Waltman L, van Eck NJ. 2013. A smart local moving algorithm for large-scale modularity-based community detection. *The European Physical Journal B* **86**:471.