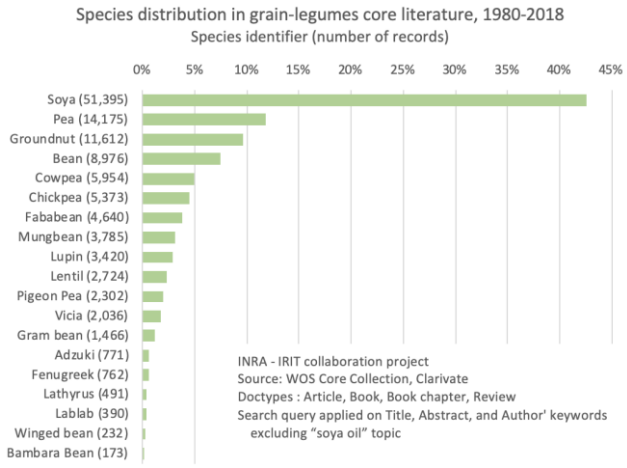
Worldwide Scientific Knowledge on Grain-legumes: to what extent does science contribute to agricultural diversity? A bibliometric method and analysis (1980-2018)

In 2016, the United-Nations created an international year for pulses to raise awareness about the considerable benefits of these crops for the sustainability transition of agrofood systems, hoping to favor their development. Yet pulses are subject to strong lock-in; and among grain-legumes, soya is the leading worldwide crop developed mainly for oil and feed uses. To reverse lock-in and to foster more diversity in the cultivation and uses of grain-legumes, scientific research is essential. Both in agriculture and food, the sciences provide new knowledge that leads to new practices and innovations. Our study examines whether, and how, science addressed the diversity of grain-legume species at the global scale, with particular interest in comparing the share of studies devoted to soya vs. pulses. We built a dataset of 107,823 scientific publications between 1980 and 2018 retrieved from Clarivate's Web of Science. These records cover 10 scientific themes of interest running the gamut of agrofood research (e.g., genetics, agronomy, nutrition). We found that focus of research has been unbalanced: soya has been – and still is – over-researched compared to other legumes. Science itself, therefore, contributes to the lack of grain-legumes diversity, raising questions for the future about how to tackle agricultural diversity.

The main species mentioned in publications on temperate climate grain-legumes over the period 1980-2018 in the WoS ("soya oil" topic being excluded).



Within the same crop family – grain-legumes – pulses and soya have developed very differently since the 1960s. Soya is the main worldwide protein crop grown mainly for feed use (with soya oil increasingly becoming a by-product). Nowadays soya crop amounts to more than 300 million metric tons, while other main grain-legumes such as pulses (pea, lentils, lupins, fababeans, chickpea, etc.) account for less than 100 million metric tons (FAOStats in Table A). Pulses production encountered few progresses, whether for food or for feed.

Table A. Global production of main pulses, soya and cereals, trienniums ending 1971, 1981, 1991, 2001, 2011, and 2017 (million tonnes). Source: FAOstat.

Year	Bean (dry)	Chick-pea	Pea (dry)	Faba broadbean	Lentil	Pigeon pea	Cowpea	Vetches	Lupin	Bambara Bean	Other pulses	Total Pulses	Soya	Cereals
1971	12	6	9	8	1	2	1	1	0,3	0,03	3	42	45	1 229
1981	15	5	7	8	1	2	1	1	1	0,03	2	41	88	1 632
1991	18	8	12	6	2	2	2	1	1	0,08	4	56	102	1 890
2001	18	6	10	8	3	3	3	1	1	0,08	3	56	177	2 104
2011	24	11	10	8	4	4	4	1	1	0,14	3	69	261	2 588
2017	31	14	16	8	7	6	7	1	1	0,18	4	95	352	2 980

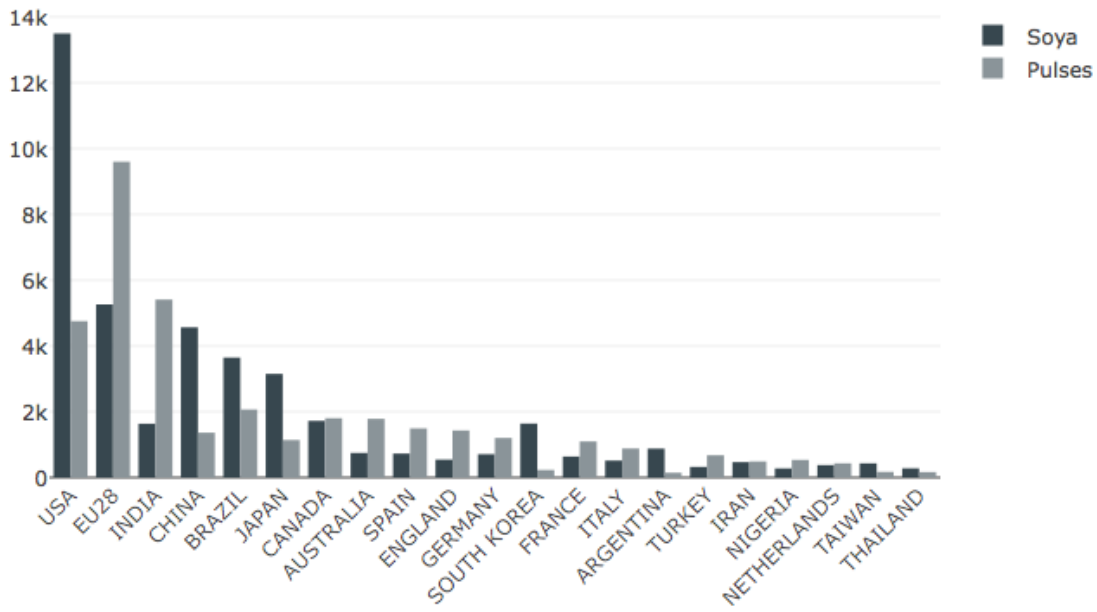
Considering that scientific research and innovation are essential drivers for developing crops, the purpose of our investigation was to uncover the landscape of grain-legume research, with particular focus on temperate climate species (that is, the grain-legume species grown in most Western countries[1]). In analyzing the current research on grain-legumes, we wanted to assess the relative occurrences of various grain-legume species in the literature and which countries are the most involved in this research. We performed a quantitative analysis of the worldwide scientific literature on these species over the last four decades, by considering ten themes of interest in agrofood system research: genetics, agronomy, ecophysiology, bioaggressors, feeding, processing, nutrition, allergy, acceptability, and socioeconomics. No bibliometric study to date has managed to tackle so many themes of interest together from agricultural and food sciences, whatever the crop species considered. In addition, to our knowledge, no study to date has sought to perform a large-scale analysis of the scientific literature on grain-legumes.

This original and ambitious bibliometric study involved 26 scientific experts on legumes, working in the 10 aforementioned themes of interest. Their expertise was coupled with database and scientometrics experts to create relevant search queries on the Web of Science and appropriate software to process and analyze the resulting bibliographic records. Our dataset merged several corpora retrieved by thematic search queries addressing the *title, abstract,* and *authors' keywords*. Our findings are based on analyzing a core corpus totaling 107,823 scientific publications (i.e., records) published between 1980 and 2018 and retrieved from the Clarivate's Web of Science (WoS). Since soya is a major crop used for oil, unlike most other grain-legumes, we have excluded records referring to the subject of "soya oil", in order to have more relevant comparisons in the corpus created.
Over the period investigated (1980-2018), soya is mentioned in 43% of the records, groundnut in near 10% and all other pulses combined, 47%. The analyses we performed revealed a strong imbalance within grain-legume species research, with soya dominant over all other grain-legume species[2]. This trend has grown even stronger in recent years. We also observed that the breakdown of themes researched were not the same for soya and pulses. PROCESSING and NUTRITION were themes of research much more often for soya than for pulses. For pulses, research mainly focused on 'upstream' themes linked to GENETICS or ECOPHYSIOLOGY. This imbalance questions the capacity of research to develop knowledge that would enable more

[1] but many of them are also important grain-legumes for semi-arid or tropical areas.
[2] If "soya oil" had been included in keyword searches, this percentage would be even greater.

food outlets for pulses, as expected by the United Nations during the IYP. In addition, while some geographical areas had more research on pulses than on soya, like the EU28 and India, newcomers such as China clearly made the choice to invest more in research on soya like American countries, reinforced by important collaborations between the USA and China on soya. (Figure B).

Figure B. Soya and Pulses records across countries over 1980-1990.
The 20 highest frequencies are based on total records by country, a group count done for the EU28. Proportional count for international collaboration records is applied. The country ranking is based on the total records number by country.



Our main conclusion is that path-dependency has strongly affected the sciences as well as agriculture itself; thus, greater awareness of this situation may help to change the trajectory of research. Those findings could foster further discussion and reflection among the scientific community about grain-legumes, especially regarding the challenge of greater crop diversity. In addition, our study makes a significant methodological contribution about building a multidisciplinary, bibliometric dataset; this method can be adopted to assess the evolution of scientific knowledge in any field.

Further investigation are still under progress to deeply analyse the knowledge dynamics on legumes, especially within food sciences.

Contact : Marie-Benoît MAGRINI (INRA)
WP2. Task T2.2 Analysis of knowledge dynamics on legumes. Forthcoming academic paper in an open-access review.