

# Tools for producing indicators from a bibliometric study of scientific production: the case of fruit and vegetable publications by the French National Institute for Agricultural Research (INRA)

Hugues LEISER<sup>1</sup>, Pascal AVENTURIER<sup>2</sup>, Dominique FOURNIER<sup>3</sup>, Françoise DOSBA<sup>4\*</sup>, Benoît JEANNEQUIN<sup>5</sup>

<sup>1</sup> INRA, Unités CSE PSH Serv. Doc., Domaine Saint-Paul, F-84914 Avignon, France

<sup>2</sup> INRA, Unité Doc. ERIST, Domaine Saint-Paul, F-84914 Avignon, France

<sup>3</sup> INRA, Unité Doc. ERIST, 2, place Viala, F-34060 Montpellier Cedex 1, France

<sup>4</sup> Montpellier SupAgro, UMR DAP, Dév. Amélior. Plantes, 2, place Viala, F-34060 Montpellier Cedex 1, France

<sup>5</sup> INRA, Domaine Exp. Inra SAD, Le Mas Blanc, F-66200 Alénia, France

## Tools for producing indicators from a bibliometric study of scientific production: the case of fruit and vegetable publications by the French National Institute for Agricultural Research (INRA).

**Abstract — Introduction.** The fruit and vegetable sector group of the French National Institute for Agricultural Research (INRA) undertook a bibliometric analysis of INRA scientific academic literature on fruits and vegetables for the 2000–2006 period. This study made it possible to produce quantitative and qualitative indicators that will shed light on the institute's scientific position in relation to the fruit and vegetable sectors. **Materials and methods.** Different steps were taken to constitute and process a corpus of INRA researcher publication data. The citation corpus gathering bibliographic data was constituted using the Web of Science. The Sphinx data processing program allowed the constituted databases to be structured and manipulated. **Results.** After the data were cleaned, there was a total of 1463 articles published by INRA authors between 2000 and 2006; this constituted the “fruit and vegetable corpus” studied, which represented 8% of the total number of INRA publications during this period. The typology of fruits and vegetables studied at INRA, the INRA research centres involved, the main subjects addressed by INRA research divisions and the national and international partnerships of INRA authors were analysed. **Conclusion.** The analysis presented highlights the overall magnitude of research dedicated to fruits and vegetables at INRA. This research is undertaken in the principal research centres of horticulture and/or fruit regions, with a high level of involvement of Southeast centres of France. Partnerships with other French and foreign organisations are numerous. This assessment allows an understanding of the rapid evolution of research that has occurred over the past several years in the fruit and vegetable field, notably the growing involvement of nutrition and consumer sciences. The development of multi-disciplinary and integrated approaches must also be emphasised.

**France / fruits / vegetables / research institutions / public research / scientific production / data collection / information science / information processing / bibliometrics**

## Outils pour produire des indicateurs à partir d'une étude bibliométrique de la production scientifique : cas des publications sur les fruits et légumes par l'Institut national de la recherche agronomique (INRA).

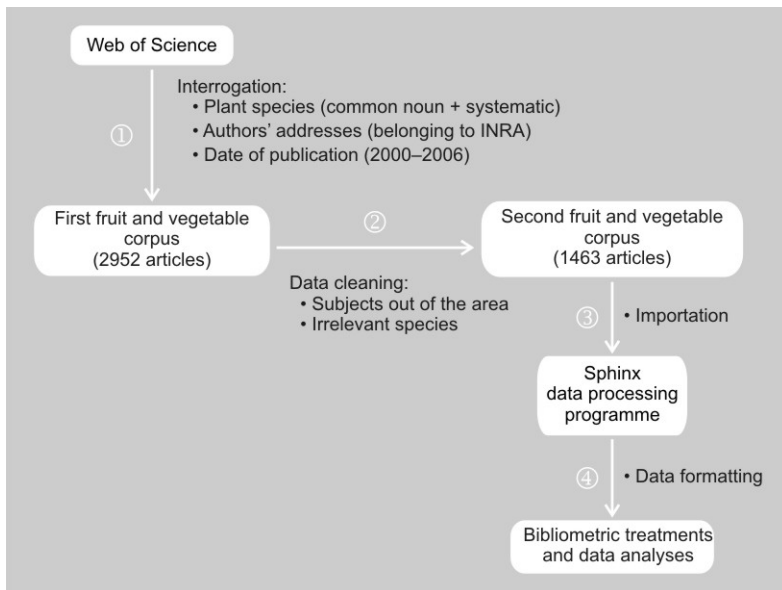
**Résumé — Introduction.** Le groupe filières des fruits et légumes de l'Institut national de la recherche agronomique (INRA) a entrepris une analyse bibliométrique des publications académiques de l'institut sur la thématique des fruits et légumes pour la période 2000–2006. Cette étude a permis de produire des indicateurs quantitatifs et qualitatifs aptes à identifier la position scientifique de l'INRA dans le secteur des fruits et légumes. **Matériel et méthodes.** Différentes mesures ont été prises pour constituer et traiter un corpus de données des publications des chercheurs de l'INRA. Le corpus de citations a été constitué en utilisant la base bibliographique internationale *Web of Science*. Le logiciel *Le Sphinx* a permis de structurer et de traiter la base de données constituée. **Résultats.** Après nettoyage des données, un total de 1463 articles publiés par des auteurs de l'INRA sur la thématique des fruits et légumes entre 2000 et 2006 a été collecté ; le « corpus des fruits et légumes » ainsi constitué représente 8 % du nombre total d'articles publiés par l'INRA au cours de cette période. Les analyses ont porté sur la typologie des fruits et légumes étudiés par l'institut, les centres des recherches INRA impliqués, les principaux sujets abordés par ses départements de recherche et les partenariats nationaux et internationaux. **Conclusion.** L'analyse présentée met en évidence l'importance globale des recherches consacrées aux fruits et légumes à l'INRA. Ces recherches sont entreprises dans les centres de recherches INRA localisés dans les principales régions de production, avec une participation élevée des centres du sud-est de la France. Les collaborations avec d'autres organismes français et étrangers sont nombreuses. Cette évaluation permet de comprendre l'évolution rapide de la recherche apparue au cours des dernières années dans le domaine des fruits et légumes, notamment avec l'émergence des sciences de la nutrition et de la consommation. Le développement d'approches globales multidisciplinaires doit également être souligné.

**France / fruits / légume / institution de recherche / recherche publique / publication scientifique / collecte de données / science de l'information / traitement de l'information / bibliométrie**

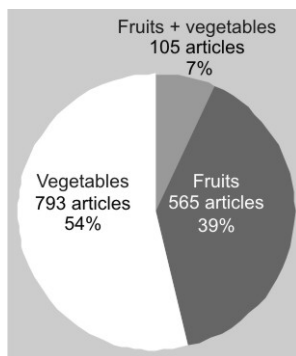
\* Correspondence and reprints

*Fruits*, 2009, vol. 64, p. 305–312  
© 2009 Cirad/EDP Sciences  
All rights reserved  
DOI: 10.1051/fruits/2009025  
www.fruits-journal.org

**RESUMEN ESPAÑOL**, p. 312



**Figure 1.** Fruit and vegetable data processing chain for identifying the articles on fruits and vegetables present in the Web of Science and published by INRA authors between 2000 and 2006.



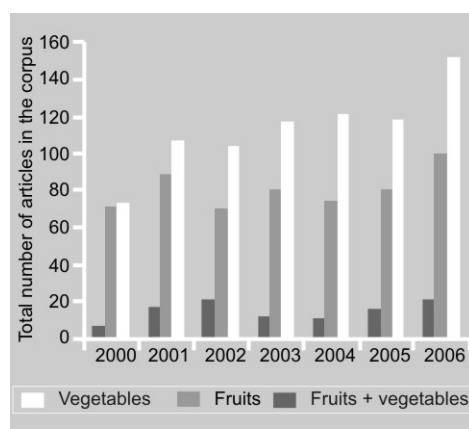
**Figure 2.** Distribution of articles related to fruits and/or vegetables and published by INRA authors between 2000 and 2006.

**Figure 3.** Evolution of the number of articles related to fruits and/or vegetables and published by INRA authors between 2000 and 2006.

## 1. Introduction

The task of the fruit and vegetable sector group of the French National Institute for Agricultural Research (INRA) is to analyse the organisation of INRA's scientific resources related to these sectors. To achieve this objective, the sector group undertook a bibliometric analysis of all of INRA's scientific output on fruits and vegetables.

Our study aimed to produce quantitative and qualitative indicators that will shed light on INRA's scientific position in relationship to the fruit and vegetable sectors. This inventory covers the period between 2000 and 2006.



For the two groups of species concerned, the synthesis of this collective effort compared the relative results to show: (I) the research centres and divisions of INRA involved; (II) the principal themes studied; (III) the quality of the scientific output and interactions taking place within INRA; (IV) partners and other organisations in France; and (V) international partnerships.

## 2. Methodology

Different steps were taken to constitute and process the corpus of data (figure 1).

The citation corpus gathering bibliographic data was constituted using the Web of Science (WOS), an international, multi-disciplinary citation database covering over 9000 peer-reviewed, international journals.

The Sphinx data processing program allowed the constituted databases to be structured and manipulated.

## 3. Main findings

After the data were cleaned, there was a total of 1463 articles published by INRA authors between 2000 and 2006; this constituted the “fruit and vegetable corpus” studied. The number of articles published on fruits and vegetables represented 8% of the total number of INRA publications during this period.

### 3.1. The typology of fruits and vegetables studied at INRA

In the “fruit and vegetable corpus” constituted, 54% of the articles concern only vegetable species and 39% only fruit species, while 7% are dedicated simultaneously to species of both groups (figure 2).

The annual number of publications increased by more than 80 % over the 2000–2006 period (figure 3), principally due to an increase in the number of articles concerning vegetable species, particularly the tomato.

The diversity in the number of species studied is approximately the same for fruits (21 species) as for vegetables (24 species).

Regarding vegetables, the thirteen main species (or groups of species) studied represent 66% of the “vegetable” corpus (*table I*). Four major species account for 44.6% of the total number of articles: tomato, which has become a model species; peas, which benefit from joint studies with protein peas; cabbages, with half including cauliflower; and green beans.

Regarding fruits, nine main species (or groups of species) represent 80% of the « fruit » corpus, with four groups of major species: apple, peach, dessert grapes, which benefit from joint studies undertaken on wine grapes, and citrus fruits, principally clementines and rootstocks (*table II*).

These data also show a fairly strong relationship between the economic importance, ranked by production, estimated by the SCEES<sup>1</sup> in 2006, and scientific output, notably for the main species. Nevertheless, some widely cultivated species in France do not figure prominently in INRA academic literature. The examination underway of non-academic literature should throw some light on this point.

### 3.2. INRA fruit and vegetable research network: the research centres involved

The analysis of publications according to the authors' affiliations shows that all INRA regional centres are involved to varying degrees with the fruit and vegetable sector. Nine sites account for over 90% of the corpus (*figure 4*).

In the research centres, units are specialised in the subjects of the research divisions to which they are attached.

<sup>1</sup> SCEES: Central Service of the economic and statistical studies of the French Ministry for agriculture.

**Table I.**

Analysis of the number of articles per vegetable species as a function of their importance in the “fruit and vegetable corpus” of articles published by INRA<sup>1</sup> authors between 2000 and 2006, and considered compared with their rank in the French production.

Vegetable species	Number of articles	Total articles (%)	Rank <sup>2</sup> / France
Tomato	192	24.1	1
Pea	67	8.4	9
Cabbage	50	6.3	5
Bean	46	5.8	7
Melon	30	3.8	8
Pepper	27	3.4	-
Lettuce	27	3.4	3
Onion	17	2.1	6
Cucumber	16	2.0	14
Radish	15	1.9	-
Carrot	14	1.8	2
Turnip	13	1.6	-
Spinach	10	1.3	-

<sup>1</sup> INRA: French National Institute for Agricultural Research.

<sup>2</sup> Rank based on production in tons in 2006 (SCEES : Service Central des Études Économiques et Statistiques, France).

**Table II.**

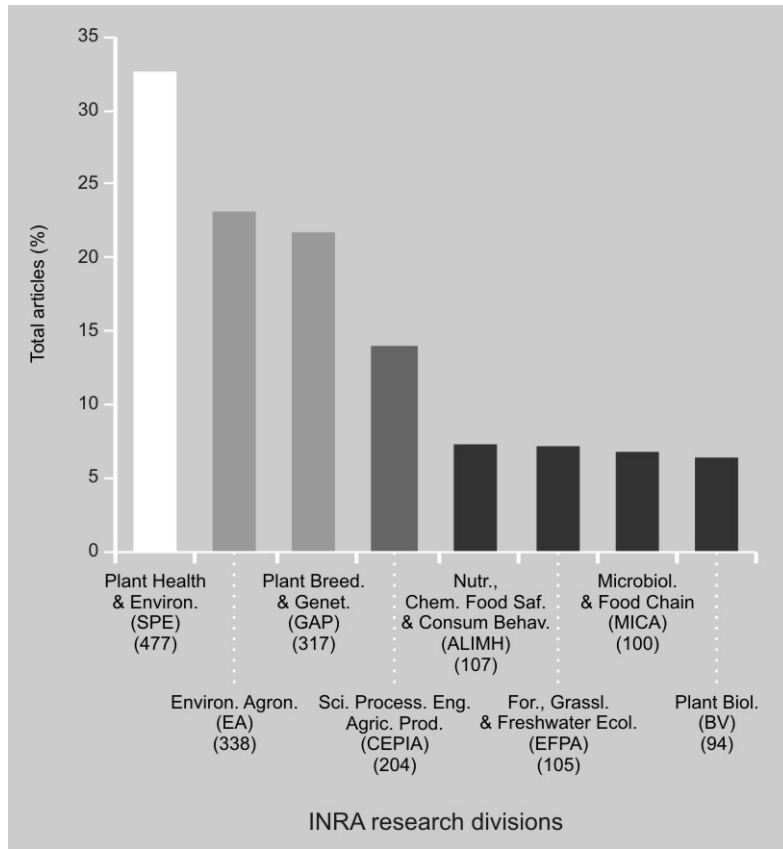
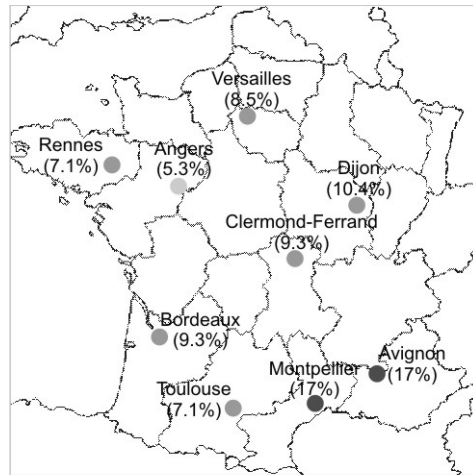
Analysis of the number of articles per fruit species and genera as a function of their importance in the “fruit and vegetable corpus” of articles published by INRA<sup>1</sup> authors between 2000 and 2006, and considered compared with their rank in the French production.

Fruit species	Number of articles	Total articles (%)	Rank <sup>2</sup> / France
Apple	110	19.0	1
Peach	75	13.0	2
Citrus	60	10.4	-
Grape	49	8.5	-
Walnut	41	7.0	7
Prune	41	7.0	3
Olive	39	6.7	-
Apricot	27	4.7	5
Pear	25	4.3	4

<sup>1</sup> INRA: French National Institute for Agricultural Research.

<sup>2</sup> Rank based on production in tons in 2006 (SCEES : Service Central des Études Économiques et Statistiques, France).

**Figure 4.** Identification of INRA research centres representing at least 5% of the corpus of fruit and vegetable articles published by INRA authors between 2000 and 2006.



**Figure 5.** Number of articles related to fruits and vegetables and published by INRA authors between 2000 and 2006, presented per INRA research division.

### 3.3. The main subjects addressed by INRA research divisions

The main subjects addressed by INRA research divisions principally concern:

- plant material and variety innovation: the “Plant Biology” (BV) and “Plant Breeding and Genetics” (GAP) divisions;
- cropping techniques and regulatory and environmental constraints: the “Environment and Agronomy” (EA), “Forest, Grassland and Freshwater Ecology” (EFPA), and “Science for Action and Sustainable Development” (SAD) divisions;
- control of bio-aggressors: the “Plant Health and Environment” (SPE) division;
- qualities of products and food: the “Science and Process Engineering of Agricultural Products” (CEPIA), “Nutrition, Chemical Food Safety and Consumer Behaviour” (ALIMH), and “Microbiology and Food Chain” (MICA) divisions;
- socio-economic approaches<sup>2</sup>: the “Science for Action and Sustainable Development” (SAD) and “Social Sciences, Agriculture and Food, Rural Development and Environment” (SAE2) divisions.

It should be noted that most divisions combine pure and applied research. Furthermore, generic and systemic approaches fostering interaction between research scientists from several divisions are often employed in fruit and vegetable research.

### 3.4. Involvement of INRA research divisions

According to the data gathered, all INRA research divisions involved in plant production conduct research on fruits and vegetables; however, the volume of articles produced by each division varies greatly (figure 5).

There are four main divisions (SPE, EA, GAP and CEPIA) which account for 77% of the fruit and vegetable corpus. This figure rises to 95% when the ALIMH, EFPA, MICA and BV divisions are included.

<sup>2</sup> This subject is little represented in this study because the majority of journals in which its research is published are not listed in the WOS. A specific bibliometric analysis is underway.

### 3.5. Partnerships

The researchers of the French National Institute for Agricultural Research collaborate with various other research teams of their own institute or with teams of other French or foreign countries' organisations.

#### 3.5.1. Collaboration between INRA research divisions

This collaboration was identified through the co-signing of articles by research scientists belonging to several divisions (*figure 6*). On average, the rate of co-signing with one or several other divisions is 34% for the ensemble of divisions.

Other observations can be mentioned (*figure 6*):

– MICA (Microbiology and Food Chain) division co-signs over 90% of its articles;

– three divisions, GAP (Plant Breeding and Genetics), EA (Environment and Agronomy) and CEPIA (Science and Process Engineering of Agricultural Products), co-sign nearly 50% of their articles with other divisions;

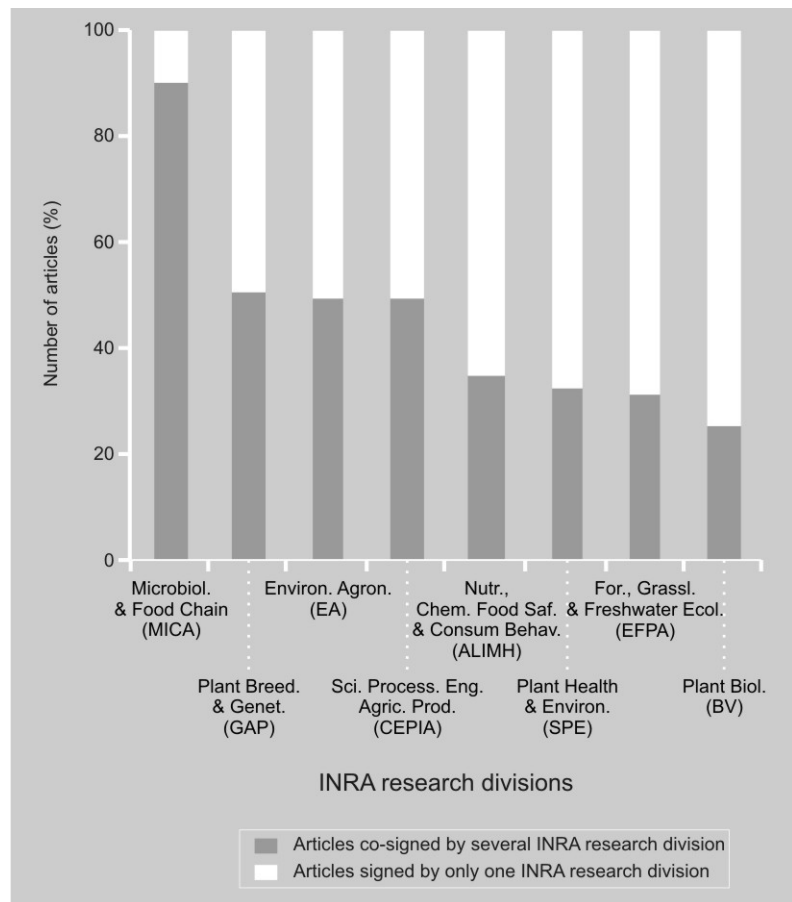
– four divisions, SPE (Plant Health and Environment), ALIMH (Nutrition, Chemical Food Safety and Consumer Behaviour), EFPA (Forest, Grassland and Freshwater Ecology) and BV (Plant Biology), co-sign approximately one-quarter of their articles.

The network between the eight research divisions representing 70% of co-signatures highlights two groups: on one hand, [CEPIA / MICA] focusing on product quality and post-harvest operations; on the other, [GAP / EA / SPE] focusing on fruit and vegetable production.

During the period under study, there was a net increase in collaboration between INRA research divisions, principally between the EA and GAP divisions.

#### 3.5.2. Collaboration with French partners outside INRA

A typology of partners established that 55% of articles are co-signed with a non-INRA French partner (*figure 7*).



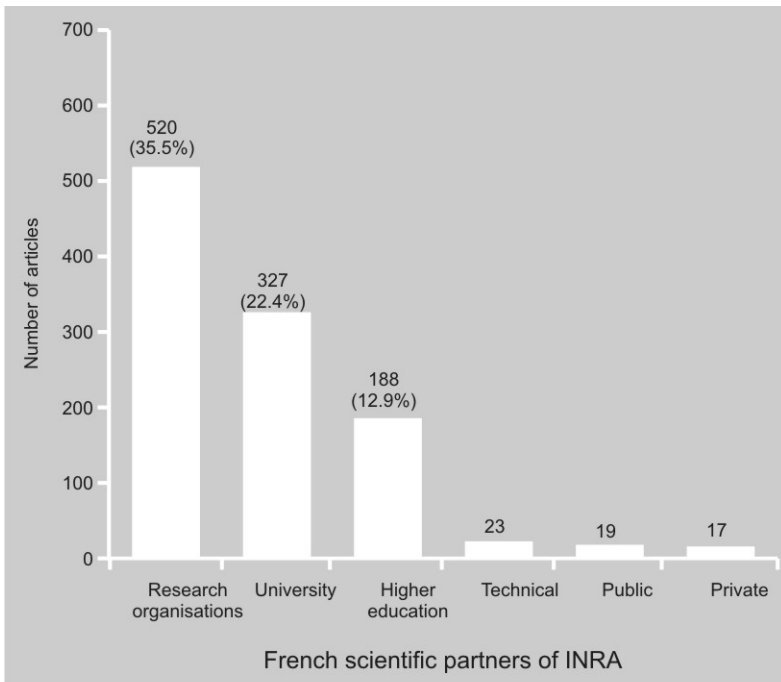
The principal types of French scientific partners are, for academic publications, primarily research organisations, followed by universities and higher-level agronomy education institutions.

#### 3.5.3. International partnerships

International partnerships were identified first at the level of co-signing countries, then aggregated into geographic zones. Articles co-signed by at least one foreign partner represent 38.7% of the fruit and vegetable corpus.

The principal foreign partnerships are with countries in northern Europe (Germany, United Kingdom, The Netherlands), southern Europe (Italy, Spain, Greece) and North America (United States) (*table III*). Since 2000, there has been a steady increase in partnerships with Asian countries, notably with China, and Oceania (multiplied by 3).

**Figure 6.** Co-signature rate of each INRA division with other divisions for articles related to fruits and vegetables and published by INRA authors between 2000 and 2006.



**Figure 7.** French partners outside of INRA who co-signed articles related to fruits and vegetables published by INRA authors between 2000 and 2006.

The international partnerships identified through the fruit and vegetable corpus are comparable with those described at the level of the whole of INRA, with, nevertheless, a strong presence of southern Europe (+ 2%) for the fruit and vegetable corpus.

#### 4. Conclusion

The bibliometric analysis concerning the INRA publications during the 2000–2006 period is composed of two stages: an analysis of academic literature (listed in the WOS) presented in the present paper and an analysis of non-academic literature, cur-

rently underway, whose results will be available during 2009.

The results of this analysis primarily highlight the overall magnitude of research dedicated to fruits and vegetables at INRA. The amount of research is significant, even if one notes that there is a concentration on the major species. The volume of research conducted at INRA thus matches the economic importance of the two most cultivated vegetable and fruit species, namely the tomato and apple, as well as other species of global significance.

Research on fruits and vegetables at INRA involves the majority of research divisions working in the plant domain. This research is undertaken in the principal research centres of horticulture and/or fruit regions, with a high level of involvement of Southeast centres of France, yet a relatively balanced distribution over mainland France. Partnerships with other French and foreign organisations are numerous within this academic valorisation framework.

This assessment, which, it must be remembered, does not cover economic or social sciences, allows an understanding of the rapid evolution of research that has occurred over the past several years in the fruit and vegetable field, notably the growing involvement of nutrition and consumer sciences. The development of multi-disciplinary and integrated approaches must also be emphasised.

#### Acknowledgements

This work was achieved by INRA researchers with the participation of CREBI (Resource and Expertise Centre in Bibliometry in Jouy-en-Josas, INRA, France).

**Table III.**

INRA<sup>1</sup> international partners in relation to the corpus of articles on fruits and vegetables published by INRA<sup>1</sup> authors between 2000 and 2006.

Continent and countries	Publications of the fruit and vegetable corpus co-signed with foreign partners (%)
Northern Europe	13
Belgium	
Germany	
The Netherlands	
United Kingdom	
Southern Europe	10
Greece	
Italy	
Spain	
North America	8
Canada	
United States	
Northern Africa	4
Morocco	
Tunisia	
Oceania	3
Australia	
New Zealand	
Asia	2.5
China	
South Korea	
South America	2
Brazil	
Chile	
Middle-East (Israel)	2

<sup>1</sup> INRA: French National Institute for Agricultural Research.

**Herramientas para producir indicadores a partir de un estudio bibliométrico de la producción científica: caso de publicaciones sobre las frutas y verduras por el Instituto nacional de investigación agronómica de Francia (INRA).**

**Resumen — Introducción.** El grupo filial de las frutas y verduras del Instituto nacional de investigación agronómica de Francia (INRA) inició un análisis bibliométrico de las publicaciones académicas del instituto sobre la temática de las frutas y verduras del periodo 2000–2006. Dicho estudio permitió producir indicadores cuantitativos y cualitativos aptos para identificar la posición científica del INRA en el sector de las frutas y verduras. **Material y métodos.** Con el fin de constituir y de tratar el cuerpo de datos de las publicaciones de los investigadores del INRA, se tomaron diversas medidas. El cuerpo de citas se constituyó gracias al empleo de la base bibliográfica internacional *Web of Science*. El software *Le Sphinx* permitió estructurar y tratar la base de datos constituida. **Resultados.** Se recopilaron, tras selección de datos, un total de 1463 artículos publicados por autores del INRA sobre la temática de las frutas y verduras entre 2000 y 2006; el « cuerpo de las frutas y verduras », constituido de este modo, representa el 8 % del número total de artículos publicados por el INRA en el transcurso de este periodo. Los análisis se centraron en la tipología de las frutas y verduras estudiadas por el instituto, en los centros de investigación INRA implicados, en los principales temas tratados por sus departamentos de investigación, así como en las cooperaciones nacionales e internacionales. **Conclusión.** El análisis presentado pone de manifiesto la importancia global de las investigaciones dedicadas a las frutas y verduras en el INRA. Dichos estudios se emprenden en centros de investigación INRA localizados en las principales regiones de producción, con una elevada participación de los centros del sudeste de Francia. Las colaboraciones con otros organismos franceses y extranjeros son numerosas. Dicha evaluación permite comprender la rápida evolución de la investigación surgida a lo largo de los últimos años en el ámbito de las frutas y verduras, sobre todo con el surgimiento de las ciencias de la nutrición y del consumo. Asimismo es importante el desarrollo de sendos acercamientos globales multidisciplinares.

**Francia / frutas / hortalizas / instituciones de investigación / investigación oficial / publicaciones científicas / colección de datos / ciencia de información / procesamiento de información**

