OST STUDIES : THEMES

BIBLIOMETRIC ANALYSIS OF RESEARCH IN GENOMICS DURING THE 1990'S

October 2003

Study produced by OST in collaboration with the LTSI,
at the request of the CNRG and the RNG

OST Observatoire des Sciences et des Techniques

LTSI

réseau national genopole®

A word about methodology

The bibliographic database of the Institute for Scientific Information (ISI) offers in addition to the standard functions the possibility of knowing the «impact», that is the citation rate of publications included in the database. Knowing how many times an article has been cited provides an estimation of the article's contribution to knowledge production in its field. In the absence of any measure of the quantity or quality of the knowledge contained in an article, impact measurement provides an invaluable indication of the «use-value» of an article as well as of the prestige of its authors.

In certain research fields – of which genomics is one – it is commonplace for scientists to use this database to observe on an individual level the impact of their work or the work of colleagues.

Nonetheless, as soon as the question turns to analyzing, comparing and tracking over time the research activity of entities of a certain size (a country, an institution, a region) or complexity (a multidisciplinary institute, a site), more elaborate – and suitably normalized – bibliometric indicators are needed. To take a simple example, in order to seize the significance of an observed increase in the number or the impact of the articles produced by a particular country it must be known if the production or the impact of most other countries or of some reference set also rose and how much.

Although access to the full text of articles included in the ISI database is not yet possible, an abstract or summary of most articles is available, and it is possible to perform textual analysis on these summaries. A number of methods of data mining can then be used to define a coherent set of articles based on the results of such analysis. This approach is especially valuable as a way to identify a set of articles representative of particular research themes or emerging research fields. Normalized, thematic, bibliometric indicators can in this way be generated to compare, to track, and to analyze research activities in a specific thematic area or new field.

OST has been producing to international standards such normalized indicators for a number of years. This know-how has been put to use developing thematic indicators such as are needed to follow-up and evaluate research policies.

Bibliometric Analysis of Research in Genomics during the 1990's

Ghislaine FILLIATREAU, Suzy RAMANANA-RAHARY,
Vincent BLANCHARD, Nelson TEIXEIRA,
Observatory of Science and Technology
and
Michel KERBAOL, Jean-Yves BANSARD,
Laboratory for the treatment of signals and images, Inserm/University of Rennes I

This study was produced at the request of the Genomics Program of the Ministry of Research and New Technologies and of the National Network of Genopoles, a component of the National Genomics Research Consortium (Prof. Jacques Haiech, Director)

Translated from the French version by Timothy Carlson

SUMMARY

Synopsis

Research articles published in genomics between 1993 and 2000 and figuring in the Science Citation Index database are an important source of information about the production and visibility of the principal actors in this research field in their French, European and world context. This information also provides a view of the trends in scientific production in genomics during this period of the 1990's.

Concerning Europe, the study pinpoints the principal regions of production, while for France it provides a means to rank the regions which have been singled out for funding by the Genomics Program of the Research Ministry, as put in place in 1999.

The principal findings of this study are as follows:

- Genomics is a research field which is active, rapidly expanding, and highly internationalized, characterized by a large number of research teams working either in synergy or competition. 53% of genomics articles are in fundamental biology, 33% are in biomedical research.

- The United States plays a major role in the field, accounting for 40% of worldwide publication with an impact index well above the world average.

- The European Union holds second place in world production with 33% of all articles published worldwide. The 10 candidate countries account for a mere 1%, but their production is increasing rapidly.

- Worldwide, only five countries can be found producing more than 5% of world publication: the United States, Japan, the United Kingdom, Germany and France. Of these, only two – the US and the UK – can boast an impact index above the world average.

- Among the top 25 producer countries, 10 appear to give pride of place to genomics among life science disciplines: China, South Korea, France, Japan, Switzerland, the United States, Germany, Finland, Australia, and Taiwan.

- Virtually all top 25 producer countries experienced an increase in impact of their production during the period studied. At the same time, the production share of certain countries, such as the United States, Canada, Italy, Australia and Denmark diminished while that of other top 25 nations rose sharply, such as the Russian Federation, Austria, Taiwan, India, and China.

- Within Europe, the top three producers – UK, Germany, and France - accounted for 60% of the Europe's production.

- Intra-European collaborations increased strongly throughout the period under study.

- Out of 226 European regions, the 20 most productive regions accounted for 48% of all publication and 54% of all citations accrued by European articles.

Concerning France, the study yielded several findings:

- France is one of the countries whose genomics research effort compares favourably to the size of its effort in either fundamental biology or biomedical research.

- Worldwide, France ranks fifth in publication in genomics, with 5,9% of world production, and ninth in terms of impact, with an impact index of 0.87 for the period under study (compared to the world average set at 1).

- During the 1990's France's world share of publication in genomics moved from 6.1% to 5.6% and its relative impact index from 0.85 to 0.88. The improvement in the impact figure occurred during the first half of this period.

- The total number of articles resulting from collaboration between a French laboratory and at least one foreign laboratory increased during this period. International collaboration by French scientists, however, shifted increasingly towards European partners, with France/US collaborative production decreasing in volume. Meanwhile, French collaboration in genomics with rapidly emerging-science countries such as China, Taiwan, South Korea, is weak.

- Within the European area, France ranks third among the 15 Member States in genomic production, with 18% of all publication, and holds 6th place in relative impact. During the period under study, France's share of EU15 publication decreased from 18.7% to 16.7%, the largest share loss after that of the Netherlands. As for relative impact, France's figures rose sharply in the first half of this period then decreased and reached in 2000 the initial level of 0.98.

- The total number of articles co-authored by a French laboratory in collaboration with at least one European partner rose markedly during this period, while collaborations with its five principal partners (UK, Germany, Italy, Belgium, and Netherlands) showed solid continuity.

- Five French regions figure among the top 30 European regions in production while two of these – Ile-de-France (Paris region) and Rhône-Alpes – are in the top 20. Of the five regions, three boast an impact index higher than the average European relative impact.

- Of France's non-overseas regions, only 12 can claim a share of national production greater than 1% and only five produce more than 5% of total national publication while one region, Ile-de-France, accounts by itself for 48% of national production. Five French regions have an impact index greater than the national average. Only Ile-de-France, Alsace, and Lorraine showed a drop in production over the period, while all other regions increased their production share, with Nord-Pas-de-Calais, Brittany, and Provence-Alpes-Côte-d'Azur showing rates of increase ranging from 20 to 30 %. With the exception of Alsace, Nord-Pas-de-Calais and Pays-de-la-Loire, the 12 producer regions showed an increase in their impact factor.

- The main indicators of the study were also calculated for France's 10 regional genopoles, for the period 1999-2000. At that time 6 genopoles each produced 5 % or more of total national publication in genomics. Four poles boasted an impact factor greater than the national average; four others showed an impact factor between 0.90 and 1.00, and 2 poles which were admitted in trial phase had impact factors of less than 0.90.

Part I – Genomics Research: An international perspective

I. Context and methods

I.1. The context for this study

In recent years, most industrialized countries have begun energetically supporting genomics research, either directly or indirectly. (http://forum.europa.eu.int/irc/rtd/cogene/info/data/pub/home.htm).

In France, the "Genomics" program was established in 1999 for the twin purpose of accelerating the development of a first-rate national infrastructure (National Genotyping Center, National Sequencing Center, various technological platforms) and encouraging the growth of a national network of regional "Genopoles". At the same time, concerted action was undertaken to stimulate technology transfer while a specific program of bio-informatics activity and development got off the ground as well (http://www.recherche.gouv.fr/recherche/aci/genob.htm).

The Genomics program was evaluated twice during the year 2002. The first evaluation, carried out by the European Molecular Biology Laboratory (EMBL, Heidelberg), took a look at the scientific strategy espoused by the program. The second evaluation, performed by the auditing firm Ernst and Young, assessed the organization and functioning of the Genopole network. (http://rng.cnrg.fr).

In conjunction with these assessments, a bibliometric analysis which would document the effects of the Genomics program on French scientific production in genomics was requested of the Observatory of Science and Technology. Bibliometrics, with its quantitative analysis of scientific publication, would make it possible to characterize the position of various actors (countries, regions, institutions, etc.) by a series of indicators, and especially to track trends in these indicators in a reliable fashion over time. Bibliometrics would seem then to be an ideal tool for tracing the effects of a particular policy on its target. Nonetheless, bibliometric analyses of this type are relatively rare, since the data are formatted to generate indicators on a geographic or disciplinary base[1], rather than for a certain research area or theme.

The first step therefore was to develop a method for building bibliometric indicators by research area or theme, which were then labelled as "thematic" indicators. Once generated, these indicators made it possible to describe the principal bibliometric characteristics of genomics research, to compare the position of the main producer countries on the European and world stage, and to identify at the French and European level those regions where strong production in genomics research occurred.

The study, which was carried out on the basis of publications dated from 1993 to 2000, reflects for the situation in genomics as it evolved during the decade of the 1990's. The study terminates at the time when the Genomics program was getting off the ground in France, and therefore

[1] Disciplines are taken to mean stable research fields, serving as a basis for university curricula and possessing specialized scientific journals in which most of the articles published in a particular discipline appear. The analysis is based on this set of journals.

represents in some ways a baseline report for use afterwards in evaluating the effects of this program.

I.2. Methods and indicators

I.2.1. Selection of articles in genomics research

The most commonly used data base in bibliometrics is the Science Citation Index (SCI). The SCI permits the user to define the discipline to which an article belongs either by using the key words which are assigned to each article from a pre-determined list or by the principal discipline to which the journal publishing the article belongs. Both methods yield satisfactory results when the discipline under study is a relatively stable, conventional academic field, but neither is satisfactory when it comes to selecting articles which belong to rapidly evolving fields of research.

For this reason alternative methodological approaches have arisen, including ones based on textual analysis of article content to determine an article's field by its vocabulary.

The current study used this type of approach, based on correspondence analysis. This choice of method is based on the hypothesis that given the very standardized language of the publications in research, there will be closely similar scientific content among articles which all contain certain word combinations (the same words with the same frequency of association with other words). Simply put, each document is coded according to the word count it yields, and then this code is translated into distances between pairs of words, with two words presumably more often associated the closer they are (the notion of co-occurrence) in the totality of documents analysed. Taking this one step further, the distance between two documents can be established by assuming that the greater the number of (the same) co-associated word pairs occurring in both documents the closer they are in subject matter. Scientific documents can thus be grouped on the basis of constellations of associated words they have in common and which presumably denote similar scientific content.

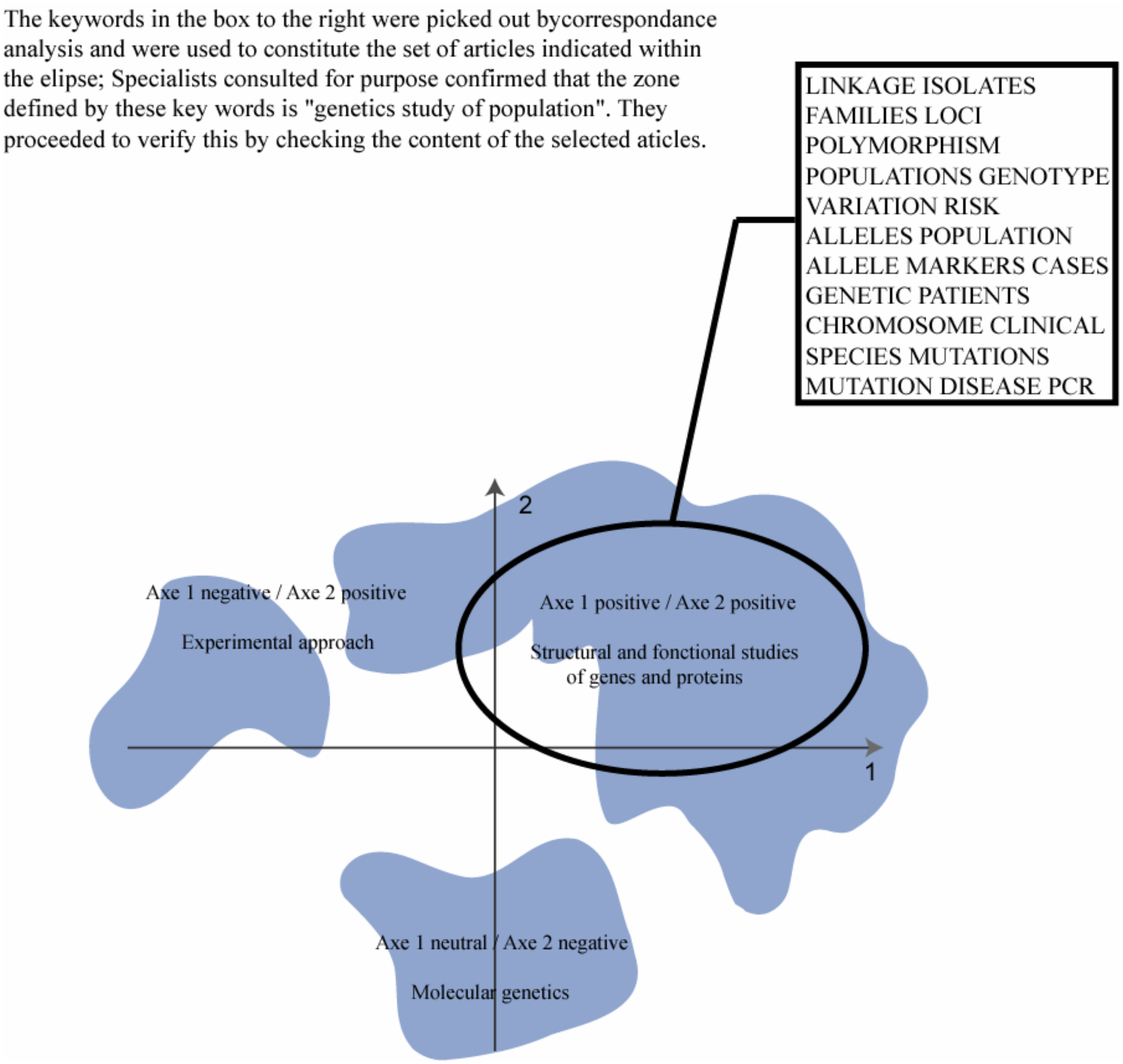
The groups of words selected to associate articles as well as the groups of articles thus produced must be validated by experts in the field under examination. This method therefore requires close cooperation between statisticians and researchers. The current study facilitated such interaction by using on line graphic representations of word proximities and article groupings.

In the first phase of the study, a set of articles was selected from all the articles in the data base published between 1993 and 2001, the selection being made on the basis of the appearance in their title or abstract of the words 'gene' or 'genes' or 'genet* or 'genom* or 'chromoso* or 'allel* or 'PCR'. Using this method, 670,458 articles utilising a vocabulary characteristic of genomic research were identified.

Since it is extremely difficult to analyse such a volume of text, this study resorted to a characteristic of factorial analysis of correspondences by which an initial set of words and documents established using this method, if sufficiently large, can be used to generate a supplementary set of documents having the same structure. In this way, an initial batch of documents can be used to generate an analytic grid to apply to further documents.

In this study, we began by analysing titles and abstracts from the 89,273 articles from the initial selection which were published in 1999. The analytic grid thus generated comprised 887 frequent words, associated "characteristic constellations", which were presented to researchers along with the full set of articles identified through the use of this grid. Once validated in this manner, the grid was used to analyze the 580,000 articles of the selection published between 1993 and 2001.

The groupings thus generated were in turn validated by experts. The validation process was accomplished using graphic representations such as the simplified one provided as an example below.



This effort resulted in the establishment of seven research "subthemes[2], which corresponded to seven sets of articles as constituted graphically by seven zones visible in the overall set of words and documents. These seven subthemes overlap, as one article can (and many do) belong to several ones, and together the seven subthemes constitute the research field of "genomics" examined in this study.

The overall set of 399,058 articles thus retained can be considered as representative of the field of genomics research, as substantiated by several tests. On one hand, experts carried out random sampling of the articles and found that the number of articles not in fact belonging to this field was about 1%, while on the other hand a certain number of tests were run which verified the representativity of the selected set – distribution analysis of the journals in which the articles were published, percentage of articles by researchers considered as representative of the field appearing in the selected set, comparing the latter to the ISI's selected list of "highly cited scientists" in the category "molecular biology genetics".

Additional information on journals in which the articles were published is presented in the Annexe ("Rapport Génomique: tableaux annexes", in French).

I.2.2. Calculating the indicators

After having matched the sample set with the overall OST data base, we calculated bibliometric indicators which measured the scientific production and visibility of various actors in genomics research (countries, groups of countries, regions) as well as indicators to measure their collaborations.

The main indicators utilised in this study are the following:

Publication share of an actor (region, zone, country) represents the portion of all articles published in a given reference set and attributable to a certain actor, the reference set being for example the world or the EU for country studies, or the nation for regional studies[3]. The part of publications directly reflects the contribution of the actor under study to the overall production set to which this actor belongs (French, European, world);

Citation share after two years of an actor is calculated similarly (but without distinguishing the origin of the citation by geographical reference set). The part of citations accumulated by

---

[2] The seven subthemes are:
Clinical approach and genomics : clinical and experimental aspects of illness related to genomics.
Genomic aspects: experimental approach; molecular genomics; structural and functional studies of genes and proteins;
Oncology: Clinical and experimental approach to the implications of the genome for oncogenesis.
Predictive genetics: genomic variations and pathological manifestations or predispositions.
Transcriptomics: expression and mutation of genes, experimental research, biology of development.
Functional genetics: cytogenetics, expression and regulation. Functional genetics and phylogenetics.
Cartography: chromosomic aberrations, caryotypes, mitotic anomalies, alterations to DNA, mapping techniques.

[3] These shares are calculated by "contribution", that is, each participation in the production of an article is counted as 1/n of a participation, n being the number of laboratories involved. An article authored by four laboratories, one each in Finland, Korea, and two in the United States would count as ¼ of a publication for Finland and Korea, and ½ of a publication for the US.

countries, or regions, for articles published, constitutes a means of evaluating their visibility and getting an idea of their impact on other producers;
Relative impact after two years of an actor is the ratio of the two preceding indicators[4] : it enables the direct comparison of the citation rate after two years of articles published by the actor with the overall average or reference rate, which by design is 1.

The relative impact is therefore a small number which allows an easy rank ordering of actors according to their scientific visibility.

Collaborations among actors are examined in this study by means of the number of articles which they have co-produced[5], which is used as a direct indicator of the "intensity" of their collaboration.

For this study, indicators were calculated:

- for the whole period, that is, from 1993 to 2000, which provides a robust average value and allows for reliable comparison among various actors;

- for three time windows within the period, which makes possible an evaluation of trends and variations in the indicators. The figures for these points [1995], [1998], [2000] are in fact average values derived from data from a 2-3 year window[6]. In this way, [1995] is by convention the average value for 1993-1995, [1998] is the average of 1997-1998, and [2000] is the average of 1999-2000.

---

[4] This impact index differs therefore from the direct impact after two years, which is calculated – for a scientist, for example – by dividing the number of citations received overall by the author by the total number of publications counted (average direct impact of articles by an author).
[5] Contrary to other indicators in this study, collaboration indicators are calculated from participation, that is, each co-signed article between two actors is counted for each of the two partners regardless of the number of total authors of the publication.
[6] These are known as "averaged, or smoothed years", corresponding to calculations averaged over three successive years in order to avoid the bias inherent in the fact that the number of articles serving as a base varies each year in function of the inevitable variations in the composition of the journal list (see methodological note B-10, Science & Technologie, Indicateurs 2002, OST, édition Economica). Changes in the base are noted each year on the site of the ISI: www.isinet.com.

II. Bibliometric Characteristics of Genomics Research

II.1. The place of genomics research in the life sciences as a whole

A simple way of getting at the quantitative volume of genomics research is to count the number of articles produced each year in this field and to compare it with the same number for each of the three other major SCI life science disciplines: fundamental biology, biomedical research, and applied-biology/ecology.

Table 1: Average number of research articles in genomics and life sciences in the SCI database

	[1995]	[1998]	[2000]	Change over the period
Genomics	40,110	50,853	60,112	50%
Fundamental biology	92,370	96,526	99,615	8%
Medical research	156,517	160,161	167,593	7%
Applied biology-ecology	35,860	37,405	38,359	7%
Life sciences	284,747	294,092	305,567	7%

Source ISI, Traitements INSERM-OST

It is clear that genomics as a research field is expanding rapidly within the area of life sciences; across the period studied, the average annual number of articles in genomics rose from 40,000 to 60,000, or a 50% increase, a sizeable number especially when compared to a rate of progression in the rest of the life sciences over the same period (7-8%). Likewise, the share of genomics studies in the life sciences overall went from less than 15% to nearly 20% by the end of the 1990's.

II.2. Composition of genomics research by discipline

By examining the disciplines in which genomic research articles were published, the "disciplinary composition" of this field of research can be identified (Table 2).

Table 2: Share of genomic articles in each discipline

Discipline	[1995]	[1998]	[2000]	Change over the period
Fundamental biology	54.8	52.7	49.9	52.3
Medical research	30.1	32.5	35.1	32.7
Applied biology-ecology	9.0	9.4	9.3	9.3
Multidisciplinary journals	4.4	3.5	3.2	3.7
Astro- and Geo- sciences	0.0	0.9	0.9	0.9
Chemistry	0.4	0.6	0.8	0.6
Engineering	0.3	0.3	0.5	0.4
Physics	0.1	0.1	0.2	0.1
Mathematics	0.1	0.1	0.1	0.1
Total	100	100	100	100

Source ISI, Traitements INSERM-OST

Overall across the period under study, more than half of all genomics publications belonged to the discipline of "fundamental biology", and one third to the "biomedical research". All in all, 85% of all genomics articles are produced by one of these two disciplines.

Applied biology is well behind with only 10% of genomics articles springing from this discipline. The category "multidisciplinary" (which corresponds to articles appearing in multidisciplinary *journals*) produced less than 4% of genomics research, while sciences of the universe generated not quite 1% of genomics production (due most likely to fields close to ecology in this discipline), and engineering sciences provided 0.5%.

A dynamic perspective will observe a clear and regular equilibrium between fundamental biology and biomedicine. The regular progression in the number of genomics articles (still very small) springing from chemistry and the slight increase in engineering articles pertaining to genomics are perhaps indications of the fact that these fields are beginning to take up questions of interest to genomics research.

II.3. Direct impact of genomics research articles

The average direct impact of a set of articles is represented by the number of citations on average each article receives during a lapse of time, given as two and five years in this study. It is well known that this bibliometric characteristic varies widely according to publishing habits and citation practices of each discipline or research community[7].

Table 3 displays the average number of citations received over two and five years by genomics articles in comparison to that of the other three major fields of the life sciences.

[7] There are many reasons for citing articles, in most cases depending on the interest of the results and the scientific prestige of the authors of the cited article.

Table 3: Average impact over two and five years of articles in genomics and in life sciences for the entire study period

	Impact over two years for the period	Impact over five years for the period
Genomics	4.36	15.23
Fundamental biology	3.25	11.56
Medical research	1.76	7.06
Applied biology-ecology	1.10	4.44
Life sciences	2.04	8.19

Source ISI, Traitements INSERM-OST

Genomics research articles, which as we have seen stem by and large (85%) from either fundamental biology or medical research, are more frequently cited than the average for the two other disciplines[8].

This suggests that genomics research is a particularly lively field of research, in which numerous laboratories and teams work on closely related topics. If we consider that the average lapse of time between the carrying out of an experiment and the publication of the findings is about two years, we see that a citation lapse of two years is quite short indeed, and therefore it can be supposed that an article is only cited by teams having rapid access to the cited article's authors' results. This demonstrates the existence of a large and rapid flow of information in this research field, and of the extent to which the field is well organised (participation in and organisation of international conferences, training workshops, and the like). It can furthermore be pointed out that by this same analysis fundamental biology itself appears to be a research field where research teams are very active in competitive and or synergistic relations to one another.

[8] This figure gives an order of magnitude or the functional equivalent in terms of annual citations per article for the relative impact, knowing that the reference point for the relative impact is 1and that the relative impact for France, 0.87, is the equivalent of 13 cit./art./year.

III. The comparison between Europe and the United States

Three sets of nations on the European continent with various European status are examined: the European Union (EU15), the ten EU candidate countries, and the EFTA (European Free Trade Association)[9].

These three geopolitical groupings are taken together for comparison with the United States[10] which in 1999 accounted for 30% of world scientific production, all disciplines taken together, as compared to 34% for the EU15.

III.1. Comparison of production

III.1.1. Publication share and change over the period under study

Table four displays a comparison among different European zones according to their production in genomics, fundamental biology, and medical research, showing also changes during the period of study.

The indicators for the whole period give a reliable picture of the position of each actor, while the indicators at several points allow a more dynamic picture to emerge.

Table 4 : Share of world publication in genomics, fundamental biology, and medical research

Zone/Country	Genomics					Fundamental biology					Medical research				
	Annual average	[1995]	[1998]	[2000]	**Change over the period**	**Annual average**	[1995]	[1998]	[2000]	**Change over the period**	**Annual average**	[1995]	[1998]	[2000]	**Change over the period**
US	**40.0**	42.7	40.6	37.4	**-13%**	**37.6**	39.1	37.7	36.2	**-7%**	**34.8**	36.8	34.6	32.9	**-11%**
EU 15	**33.3**	32.5	33.3	33.8	**4%**	**33.8**	33.2	33.9	34.0	**2%**	**38.3**	37.7	38.7	38.4	**2%**
Candidate EU	**1.0**	0.8	0.9	1.3	**66%**	**1.8**	1.7	1.7	2.0	**19%**	**1.6**	1.2	1.6	2.1	**68%**
EFTA	**2.0**	2.1	2.0	2.0	**-3%**	**1.9**	2.0	1.9	1.9	**-3%**	**2.1**	2.0	2.1	2.1	**5%**

For the overall period, the United States accounted for 40% of world production in genomics research, as compared to slightly more than 33% for the EU15. The EFTA's 2% and the candidate country's 1% reveal the weak production of these two geographic zones.
The differential between the US and the EU15 is smaller in fundamental biology where the US is less than 4 points stronger than the EU15, while in medical research it is the EU15 that leads the US by 3.5 points.
A clear decrease in production can be observed for the US in genomics as well as in fundamental biology and biomedical research. At the same time, the EU15's share has risen slightly, while the candidate countries show the most notable progress even if their absolute levels of production remain low.

[9] The 15 member states of the European Union are Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, the Netherlands, Portugal, Spain, Sweden and the United Kingdom. The 10 candidate countries are The Czech Republic, Cyprus, Estonia, Hungary, Latvia, Lithuania, Malta, Poland, Slovakia, and Slovenia. The four member countries of the EFTA are Iceland, Lichtenstein, Norway and Switzerland.
[10] The data necessary for generating indicators by state for the United States are not at present available in the OST database.

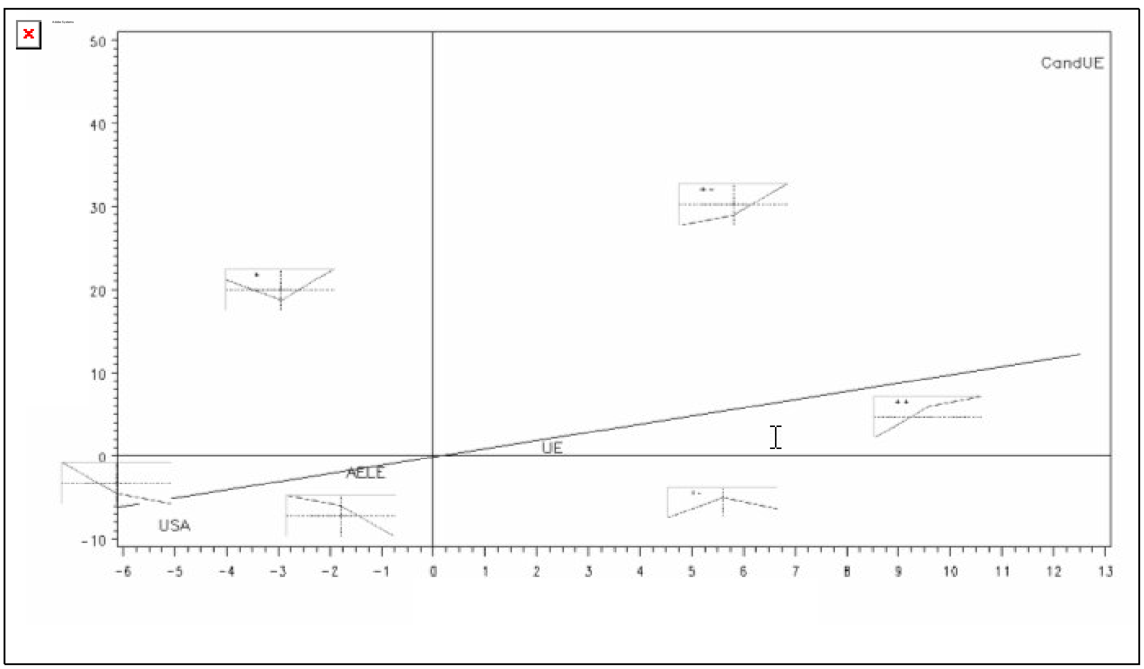
By the end of the period of study, the total of EU member plus candidate country production drew close to the level of American production; in genomics the difference is less than 2 points, in fundamental biology practically zero, and in medical research the European total is 7 points ahead of the US figure.

III.1.2. Graphic representations of changes in production

In order to make the dynamic comparison between production shares of various actors easier to grasp, the data can be displayed graphically. For example, in figure 1, below, change in production share between 1995 and 1998 is shown on the x-axis and change between 1998 and 2000 on the y-axis. The difference in duration of the two periods is neutralized by tracing an isocurve through the points at which the change is the same for the two periods. For countries lying above this line, production has changed more sharply in the second period than the first (acceleration occurring in the second period). Similarly, for countries lying below the isocurve it can be said that change was sharper during the first period (slowdown in the second period). [11]

To make this clear we have placed on the graph at the point for each actor an inset graph showing change over the two periods for that actor, which is represented by an icon.

Figure 1: Change in production share for the US and three European zones: UE15; candidate countries; EFTA



[11] In this figure, the axes divide the plane into four areas corresponding to the four possible outcomes: positive growth in both periods; first period positive growth followed by second period negative growth; negative growth in both periods; and first period negative growth followed by second period positive growth. The isocurve shows the values for which growth would be at the same rate for both periods and also divides the plane into an upper field where share growth was stronger in the second period than the first ("acceleration"), and a lower field where growth was less in the second period than the first ("slowdown"). All in all, the dynamics of total publication share across both periods can be visualized with the help of these six zones.

This figure shows first of all the usefulness of a graphic interpretation of Table 4; the figure brings out clearly the acceleration in the second period of what was already strong growth in the first period by candidate countries (CandUE) as well as the total share decrease by US publication across both periods. It also shows the first period share increase by EU15 slowing slightly in the second period, and the steady loss of share experienced by the EFTA throughout the entire two periods.

III.2. Collaborations

Table 5 shows the number of articles produced by international collaboration. This indicator represents an initial, direct measurement of the intensity of the ties between two actors, a measurement which can be refined by calculating the share of one actor's total co-publication constituted by its co-publications with one other actor. By convention, this type of calculation may count one article several times, and the values recorded therefore do not sum to the real total.

Table 5: Collaboration between the United States and various groupings of European countries

Country/Zone under study	Partner Country/Zone	Number of co-publications for the period		Share in the co-publications of the country under study	
		Annual average	Change over the period	Annual average	Change over the period
US	*US*	*17,416*	*40%*	*87.30%*	*-2%*
	UE 15	3,119	74%	15.60%	21%
	EFTA	324	67%	1.60%	20%
	Candidate countries	161	138%	0.80%	67%
EU 15	*EU 15*	*13,937*	*69%*	*84.60%*	*1%*
	US	3,119	74%	19.00%	4%
	EFTA	601	83%	3.60%	9%
	Candidate countries	335	152%	2.00%	50%
Candidate countries	Candidate countries	*348*	*230%*	*50.10%*	*26%*
	EU 15	335	152%	48.10%	-3%
	US	161	138%	23.20%	-9%
	EFTA	27	60%	3.80%	-41%
EFTA	*EFTA*	*758*	*65%*	*56.20%*	*1%*
	EU 15	601	83%	44.50%	11%
	US	324	67%	24.00%	2%
	Candidate countries	27	60%	2.00%	-5%

Source ISI, Traitements INSERM-OST

Considering the example of the EU15, the table can be read as follows: between 1995 and 2000 nearly 14,000 articles were produced annually as a result of a collaboration between at least two countries of the EU; at the same time, more than 3,000 articles were the result of collaboration between at least one EU country and at least one state in the US; 600 articles were produced by collaborative efforts involving at least one EU country and at least one of the four countries of the EFTA; and 335 articles resulted from at least one EU country and one of the ten candidate countries. The articles produced by collaboration between the EU15 and the US represent over the whole period 16% of all American co-publication and 19% of all EU15 collaboration. The US' share of total EU co-publication rose by 4% while during the same period the EU15 share of US co-publication rose by 21%.

Of particular noteworthiness in this table:

- In absolute value, the number of co-published articles rose everywhere during the period.
- Endogenous co-publication (intra-zone) in 2000 constituted 87% of US co-publication, 85% of EU15 co-publication, 55% of candidate country co-publication, and 56% of EFTA co-publication.
- US collaboration is largely endogenous (87%) but nevertheless their collaboration with the EU accounts for 15% of their co-publication.
- EU15 collaboration is equally internal (85%) but EU collaboration with the US accounts for 19% of their co-publication.
- In the case of the US, endogenous collaborations have decreased in relative value in the course of the period at the same time that collaborations with European zones have risen, indicating that US genomics research is progressively opening to the rest of the world.
- By contrast, the European zone's share of their total co-publication that resulted from collaboration with the US rose only slightly in the case of the EU15 and the EFTA and decreased in the case of the candidate countries.
- Meanwhile co-publication between EU15 countries and candidate countries rose in absolute value. It accounts for 48% of total candidate country collaboration or twice the share of co-publication between candidate countries and the US.
- Candidate countries increased considerably their collaboration with one another during the period, both in absolute terms and relative to their overall collaboration. The share of co-publication with the EU15 decreased slightly.

IV. A World View of Genomics Research: An analysis by country

IV.1. The top 25 producer countries

IV.1.1. Countries' world rank order

Table 6 displays the principal indicators showing rank ordering of the top 25 producer countries for the overall period, with the basis being worldwide production. The countries are ordered by their rank in share of total world publications (columns 1 and 4), as well as by their share of world citations (column 5), and their relative impact compared to world average (columns 2 and 6).

Table 6: World share of publications, citations, and relative world impact of the top 25 producer countries in genomics, for the overall period

Legend:

- share ≥10 %
- 5 % ≤ share < 10 %
- 1 % ≤ share < 5 %
- 0,5 % ≤ share < 1 %
- share <0,5

Rank by publication	Rank by impact	Country	Publication share	Citation share	Relative impact
1	1	UNITED STATES	40.0	52.5	1.31
2	17	JAPAN	9.9	7.1	0.71
3	3	UNITED KINGDOM	7.6	7.9	1.05
4	5	GERMANY	6.8	6.3	0.93
5	9	FRANCE	5.9	5.1	0.87
6	6	CANADA	4.0	3.6	0.9
7	16	ITALY	3.1	2.2	0.72
8	15	AUSTRALIA	2.4	1.8	0.73
9	4	NETHERLANDS	2.3	2.1	0.94
10	20	SPAIN	2.0	1.2	0.6
11	12	SWEDEN	1.8	1.5	0.81
12	2	SWITZERLAND	1.6	1.8	1.15
13	8	BELGIUM	1.0	0.8	0.88
14	11	FINLAND	0.9	0.7	0.83
15	10	ISRAEL	0.9	0.7	0.85
16	37	RUSSIAN FEDERATION	0.8	0.2	0.25
17	14	DENMARK	0.8	0.6	0.76
18	26	SOUTH KOREA	0.7	0.3	0.41
19	27	TAIWAN	0.7	0.3	0.41
20	7	AUSTRIA	0.6	0.6	0.88
21	38	INDIA	0.6	0.1	0.24
22	34	CHINA	0.6	0.2	0.31
23	32	BRAZIL	0.5	0.2	0.36
24	19	NORWAY	0.4	0.3	0.65
25	21	NEW ZEALAND	0.4	0.2	0.56
		WORLD	*100*	*100*	*1*

Source ISI, Traitements INSERM-OST

Taking the period as a whole, genomics research seems to be concentrated in a limited number of countries; the five top producer countries account for 70% of world production, the top 20 countries account for 84% and the top 25 – which include India, China, and Brazil – account for virtually all production as is clear from the fact that the 25th, New Zealand, contributed only 0.4% of world production.

In terms of production volume, the United States accounted for 40% of world production, and it received 52.5% of all citations. Its influence clearly dominates the field. This can be seen also in its relative impact which at 1.3 is largely superior to the world average (1 for this indicator, by construction). For all other countries the impact index is less than 1 with the exception of the United Kingdom (1.05) and Switzerland (1.15).

Japan, second in world production with a 10% share, receives only 7% of world citations (which translates to an impact index of 0.71). This low relative visibility, which is also true for Japan -- if to a lesser degree -- in fundamental biology and biomedical research (see tables in Annex, displayed in French), reflects the fact that this country is weakly connected to world scientific networks. Its collaboration figures attest to this as well.

Two newly industrialised countries in Asia make their presence felt in the top 25 (Korea and Taiwan), as do three emerging scientific powers (China, India, Brazil).

France figures fifth worldwide in production with a share of 5.9% while garnering 5.1% of all citations (giving it an impact index of 0.87), and it holds the same rank in genomics as it does in life sciences overall. Its genomic production nevertheless is superior to its production in fundamental biology and biomedical research (5.5% and 4.8% respectively – see table in Annex). Its impact index in genomics, by contrast, puts it in 9th place worldwide.

IV.1.2. A "specific effort" in genomics

In order to grasp the "specific effort" made in genomics by one of these countries, the country's share of world genomics production can be compared to its production share in fundamental biology or biomedical research, the two disciplines which comprise the major part of the field of genomics.

Does a particular country make a specific effort in genomics and if so how much? The answer can be derived by calculating the proportion worldwide of publications in genomics to publications in fundamental biology and medical research (reference value) and then comparing this proportion to that of the individual country under study (observed value). This procedure results in a "specific effort" index for genomics which also stands for the degree of specialization within the field of life sciences, for each country[12].

In table 7, the relative levels of production of genomics, fundamental biology and medical research for selected countries are shown in columns 3, 4, 5 while the "specific effort" for genomics for each country is shown in column 6.

[12] The "world" reference value is derived from the overall share worldwide of genomics articles in fundamental biology on one hand and medical research on the other.

Table 7: "Specific Effort" in genomics by country, for the overall period

Legend:

- share ≥10 %
- 5 % ≤ share < 10 %
- 1 % ≤ share < 5 %
- 0,5 % ≤ share < 1 %
- share <0,5

Rank by publication	Country	World share of publications for the period			Specific effort in genomics
		Genomics	Fundamental biology	Medical research	Observed/reference value
1	UNITED STATES	40	37.6	34.8	**1.09**
2	JAPAN	9.9	9.2	8.1	**1.13**
3	UNITED KINGDOM	7.6	8.0	11.3	0.82
4	GERMANY	6.8	6.6	6.3	**1.04**
5	FRANCE	5.9	5.5	4.8	**1.13**
6	CANADA	4.0	4.5	3.7	0.97
7	ITALY	3.1	3.3	4.0	0.87
8	AUSTRALIA	2.4	2.3	2.6	**1.02**
9	NETHERLANDS	2.3	2.2	2.6	0.94
10	SPAIN	2.0	2.3	2.2	0.88
11	SWEDEN	1.8	1.9	2.1	0.92
12	SWITZERLAND	1.6	1.5	1.4	**1.07**
13	BELGIUM	1.0	1	1.1	0.91
14	FINLAND	0.9	0.8	1.1	**1.03**
15	ISRAEL	0.9	0.9	1.1	0.85
16	RUSSIAN FEDERATION	0.8	1.5	0.5	0.71
17	DENMARK	0.8	0.9	1.0	0.84
18	SOUTH KOREA	0.7	0.7	0.4	**1.19**
19	TAIWAN	0.7	0.6	0.8	**1.01**
20	AUSTRIA	0.6	0.6	1.0	0.88
21	INDIA	0.6	1.0	0.9	0.64
22	CHINA	0.6	0.5	0.5	**1.2**
23	BRAZIL	0.5	0.7	0.6	0.8
24	NORWAY	0.4	0.4	0.6	0.86
25	NEW ZEALAND	0.4	0.4	0.5	0.86
	WORLD	100	100	100	**1**

Source ISI, Traitements INSERM-OST

This exercise is useful for grasping and comparing the scientific priorities of different countries concerning the field of genomics. By this means we can observe that:

- Two countries, China and South Korea, show a clear propensity for investing in genomics (relative ratio of 1.2).
- Not far behind, two other countries -- France and Japan -- are likewise committed to genomic research (1.13).
- In the range between 1 and 1.1 can be found the United States, Switzerland, Germany, Finland, Australia, and Taiwan.

- All other countries show values below the reference value, which would indicate a lack of any particular engagement in this field compared to others during the period under study.
- The Russian Federation and India appear to be under-involved in genomic research.

IV.1.3. Publication share and impact change during the period of study

Table 8 and Figure 2 present change over the period of study in the publication share and in the impact index of the top 25 producer countries in genomics. (The table corresponding to the Figure is given below, and a table showing the same for fundamental biology and medical research is provided in the Annex in French.)

Table 8: Change in world publication share and relative impact of the top 25 producer countries for the overall period

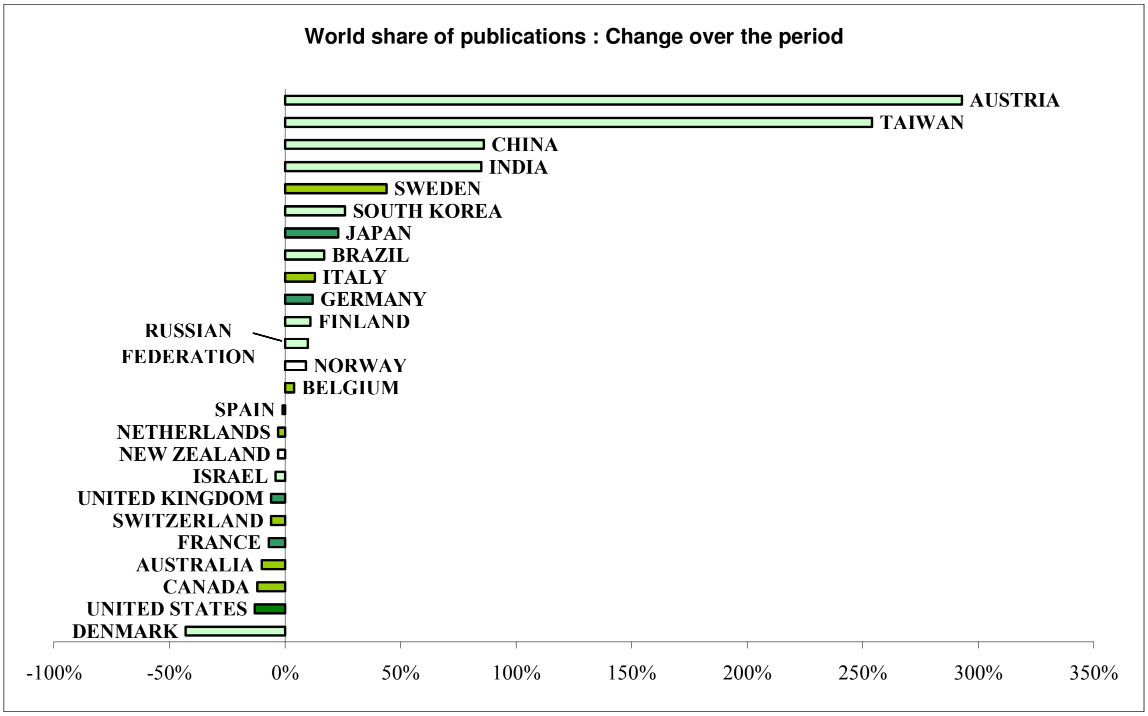
Rank by impact	Rank by publication	Country	World share of publications				Relative impact			
			[1995]	[1998]	[2000]	Change over the period	[1995]	[1998]	[2000]	Change over the period
1	1	UNITED STATES	42.7	40.6	37.4	**-13%**	1.3	1.31	1.33	**3%**
2	17	JAPAN	8.8	9.8	10.8	**23%**	0.7	0.69	0.73	**5%**
3	3	UNITED KINGDOM	7.9	7.5	7.4	**-6%**	1.02	1.03	1.08	**7%**
4	5	GERMANY	6.3	6.9	7.1	**12%**	0.92	0.94	0.93	**2%**
5	9	FRANCE	6.1	6	5.6	**-7%**	0.85	0.89	0.88	**3%**
6	6	CANADA	4.4	4.1	3.8	**-12%**	0.85	0.86	0.96	**12%**
7	16	ITALY	2.9	3.1	3.2	**13%**	0.65	0.74	0.74	**14%**
8	15	AUSTRALIA	2.4	2.3	2.1	**-10%**	0.7	0.7	0.8	**14%**
9	4	NETHERLANDS	2.5	2.5	2.4	**-3%**	0.93	0.93	0.93	**1%**
10	20	SPAIN	1.8	1.8	1.8	**-1%**	0.53	0.58	0.64	**21%**
11	12	SWEDEN	1.6	2	2.3	**44%**	0.8	0.79	0.82	**2%**
12	2	SWITZERLAND	1.6	1.6	1.5	**-6%**	1.18	1.18	1.1	**-7%**
13	8	BELGIUM	1	0.9	1	**4%**	0.91	0.88	0.87	**-5%**
14	11	FINLAND	0.8	0.9	0.9	**11%**	0.8	0.8	0.84	**5%**
15	10	ISRAEL	0.9	0.8	0.9	**-4%**	0.77	0.87	0.92	**20%**
16	37	RUSSIAN FEDERATION	0.8	0.8	0.8	**10%**	0.21	0.23	0.31	**48%**
17	14	DENMARK	1	0.8	0.6	**-43%**	0.68	0.76	0.86	**26%**
18	26	SOUTH KOREA	0.6	0.6	0.7	**26%**	0.38	0.35	0.45	**20%**
19	27	TAIWAN	0.3	0.4	1.1	**254%**	0.37	0.39	0.44	**17%**
20	7	AUSTRIA	0.3	0.6	1.1	**293%**	0.87	0.79	0.96	**11%**
21	38	INDIA	0.5	0.7	0.8	**85%**	0.21	0.23	0.27	**29%**
22	34	CHINA	0.4	0.4	0.7	**86%**	0.24	0.32	0.34	**38%**
23	32	BRAZIL	0.6	0.6	0.7	**17%**	0.35	0.39	0.35	**2%**
24	19	NORWAY	0.4	0.4	0.5	**9%**	0.62	0.63	0.68	**11%**
25	21	NEW ZEALAND	0.4	0.4	0.4	**-3%**	0.57	0.49	0.53	**-6%**
		WORLD	100	100	100		1	1	1	

Source ISI, Traitements INSERM-OST

Legend:

- share ≥10 %
- 5 % ≤ share < 10 %
- 1 % ≤ share < 5 %
- 0,5 % ≤ share < 1 %
- share <0,5

Figure 2: Change in world share of publications for the top 25 countries over the period



A certain number of countries, including the United States, the United Kingdom, Canada, Australia, Denmark, Israel, and Spain, experienced a decrease in their production share at the same time as an increase in their impact index – in some cases a sizeable increase – over the period.

The decrease in US production share over the period is noticeable. This may in part be tied to the emergence into world science of "new" countries like Taiwan, India, China, and Brazil, whose production share increased rapidly, as shown in the table. For certain countries like the US and the UK, another factor may come into play: a change in the evaluation criteria for research projects by funding institutions. In these cases, evaluation has been increasingly based on impact, causing researchers to emphasize the importance of the content of each article rather than the total number of publications. Finally, certain smaller and/or more recent producers like Spain and the Russian Federation have come to favour a policy of international collaboration with high-quality laboratories abroad as an effective means of rapidly increasing their own visibility.

Japan, whose propensity to emphasize genomic research within the life sciences is made clear by Table 7, noticeably increased its production share while also improving slightly its impact index. Germany improved its position over the whole period but less sharply during the second half of the period. France and the UK both experienced a decrease in production share, particularly in the first half of the period in the case of the UK and the second half in the case of France.

The strong commitment by China and South Korea to genomics coincides with a sharp improvement in their production share during the period.

Switzerland, Belgium and New Zealand are the only countries to experience a decrease in their impact index over the whole period, and in the case of Switzerland this drop in impact was accompanied by a decrease in production share.

France maintained its position as fifth in terms of world production share throughout the 1990's despite a decrease in its share. Its stable impact index hides a clear gain in impact during the first half of the period followed by a decrease in impact during the second half.

IV.1.4. Collaborations among countries

Table 9 establishes a list of the top six partners for a series of countries selected among the top 25 world producers. Partner countries are determined on the basis of the number of articles co-published with the country under study (columns 3 and 4). The share of total co-publication of the country under study which is accounted for by co-publication with each partner is shown in columns 5 and 6. Finally, columns 7 and 8 recall the world rank of each country in both production share and impact index. These two figures presumably indicate a country's attractiveness since they show, respectively, its potential for co-production and its international visibility.
This partnership table presents only certain of the top 25 producers in world genomic research (the complete table is included in the Annex). Articles may be counted several times, therefore the figures do not sum to real totals.

Table 9: Table of partnerships for selected countries among the top world producers

(a) : partner country's world rank by publication

(b): partner country's world rank by relative impact

Country under study	Partner country	Number of co-publications for the period		Share in the co-publications of the country under study		(a)	(b)
		Annual average	Change over the period	Annual average	Change over the period		
US	*US*	*17416*	*40%*	*87%*	*-2%*	*1*	*1*
	UK	846	75%	4%	22%	3	3
	Germany	739	90%	4%	33%	4	5
	Japan	738	87%	4%	30%	2	17
	Canada	721	47%	4%	2%	6	6
	France	537	56%	3%	9%	5	9
Total of co-publications			43%				
Japan	*Japan*	*4157*	*98%*	*86%*	*1%*	*2*	*17*
	US	738	87%	15%	-5%	1	1
	UK	117	115%	2%	10%	3	3
	Germany	99	105%	2%	5%	4	5
	Canada	69	106%	1%	5%	6	6
	France	68	110%	1%	7%	5	9
Total of co-publications			96%				
UK	*UK*	*2939*	*55%*	*69%*	*-3%*	*3*	*3*
	US	846	75%	20%	10%	1	1
	Germany	292	92%	7%	21%	4	5
	France	273	81%	6%	13%	5	9
	Italy	172	129%	4%	44%	7	16
	Netherlands	171	76%	4%	10%	9	4
Total of co-publications			59%				
Germany	*Germany*	*2675*	*88%*	*70%*	*3%*	*4*	*5*
	US	739	90%	19%	5%	1	1
	UK	292	92%	8%	6%	3	3
	France	218	80%	6%	-1%	5	9
	Switzerland	176	81%	5%	-1%	12	2
	Netherlands	151	87%	4%	3%	9	4
Total of co-publications			82%				
France	*France*	*2500*	*51%*	*74%*	*-2%*	*5*	*9*
	US	537	56%	16%	1%	1	1
	UK	273	81%	8%	17%	3	3
	Germany	218	80%	6%	16%	4	5
	Italy	142	108%	4%	35%	7	16
	Switzerland	117	89%	4%	23%	12	2
Total of co-publications			54%				
Italy	*Italy*	*1390*	*84%*	*74%*	*2%*	*7*	*16*
	US	407	81%	22%	1%	1	1
	UK	172	129%	9%	27%	3	3
	France	142	108%	8%	16%	5	9
	Germany	133	108%	7%	15%	4	5
	Netherlands	60	133%	3%	30%	9	4
Total of co-publications			80%				

Netherlands	*Netherlands*	*940*	*49%*	*67%*	*-4%*	*9*	*4*
	US	272	63%	19%	5%	1	1
	UK	171	76%	12%	13%	3	3
	Germany	151	87%	11%	20%	4	5
	France	97	74%	7%	12%	5	9
	Belgium	77	53%	6%	-1%	13	8
Total of co-publications			55%				
SPAIN	*SPAIN*	*860*	*138%*	*76%*	*4%*	*10*	*20*
	US	182	138%	16%	4%	1	1
	UK	89	99%	8%	-13%	3	3
	France	81	118%	7%	-5%	5	9
	Germany	66	105%	6%	-10%	4	5
	Italy	49	232%	4%	46%	7	16
Total of co-publications			128%				
Sweden	*Sweden*	*736*	*68%*	*67%*	*1%*	*11*	*12*
	US	221	61%	20%	-4%	1	1
	UK	92	97%	8%	18%	3	3
	Germany	81	107%	7%	24%	4	5
	Finland	62	157%	6%	53%	14	11
	France	62	72%	6%	3%	5	9
Total of co-publications			67%				
Taïwan	*Taïwan*	*290*	*199%*	*84%*	*6%*	*19*	*27*
	US	72	127%	21%	-20%	1	1
	Japan	12	303%	4%	42%	2	17
	UK	6	73%	2%	-39%	3	3
	Canada	5	159%	2%	-8%	6	6
	China	5	150%	1%	-12%	22	34
Total of co-publications			183%				
India	*India*	*218*	*119%*	*73%*	*9%*	*21*	*38*
	US	53	81%	18%	-10%	1	1
	UK	19	75%	6%	-13%	3	3
	Japan	14	206%	5%	52%	2	17
	Germany	14	146%	5%	23%	4	5
	France	7	150%	2%	25%	5	9
Total of co-publications			101%				
China	*China*	*217*	*773%*	*56%*	*75%*	*22*	*34*
	US	107	268%	28%	-26%	1	1
	Japan	35	308%	9%	-18%	2	17
	UK	29	262%	8%	-27%	3	3
	Germany	17	317%	4%	-16%	4	5
	France	15	76%	4%	-65%	5	9
Total of co-publications			398%				
BRAZIL	*BRAZIL*	*217*	*220%*	*62%*	*21%*	*23*	*32*
	US	117	154%	33%	-4%	1	1
	UK	32	50%	9%	-43%	3	3
	France	24	125%	7%	-15%	5	9
	Germany	19	176%	5%	4%	4	5
	Canada	13	193%	4%	11%	6	6
Total of co-publications			165%				

Source ISI, Traitements INSERM-OST

Taking the example of France, the table reads as follows: 74% of its co-publications, or 2,500 average annual articles, are authored by at least two French laboratories. At the same time, France leading international partner is the US (ranked first in production and impact) with an average annual co-production of 537 articles, or 16% of France's total co-publications. France's second international partner is the UK (ranked third in production and impact) with 273 articles or 8% of France's co-publication. From a dynamic perspective, the number of articles co-published by two French laboratories increased 51% over the period, in contrast, for example, to an 80% increase for articles co-published with at least one British laboratory.

The table confirms that laboratories everywhere tend to favor collaborations with same-country laboratories; endogenous collaboration accounted for anywhere from 67% (Netherlands) to 87% (US) of a country's total co-publication. China with a rate of only 57% endogenous co-publication is the notable exception.

Another fact to emerge from this table is that the attractiveness of a country – defined as its capacity to engage in international scientific relations with countries figuring among the world's top producers – is clearly correlated to its world production share and perhaps even more to its relative impact (compare columns 7 and 8). The US position as first international partner for most countries should therefore come as no surprise. By the same reasoning Japan, which has not been able to establish as many international partnerships as its number-two world ranking would lead one to expect, suffers a relatively modest impact internationally. Finally, the UK displays a greater attractiveness than its production share might indicate and, concomitantly, an excellent relative impact worldwide.

European countries prefer European countries, and the effect of "affinity by geographical proximity" is perceptible. Examples include German and French relations with Switzerland, the Netherlands, and Belgium; Sweden with Finland; or Taiwan with China.

France's attractiveness is fairly high, especially compared to its world rank in production and impact, judging from its position as third international partner for most European Union countries, as well as for Canada and Switzerland. It occupies a solid position as well with Brazil but is last on the list of partners of both China and India. It does not figure among the five top partners of South Korea or Taiwan, both of which are leading representatives of Asia's emerging dynamic presence in international science.

As far as noticeable trends are concerned, co-publication has increased for all countries with the exception of the US. For certain countries where genomic research is still getting off the ground, endogenous collaboration has increased at the same rate or faster than total collaboration (Japan, Spain, Taiwan, India, China, and Brazil). In countries with a longer tradition of genomic research, international collaboration grew faster than endogenous collaboration (UK, France, and the Netherlands).

Certain partnership links were strengthened during the period, for example Germany/US or Italy/UK.

France's total co-publications both international and endogenous shrunk between 1995 and 1998 but then increased rapidly from 1998 to 2000. The result over the whole period is a modest increase in co-publication, a virtual stagnation in endogenous collaboration and a slight increase in international collaboration. The share of co-publications with France's principal European partners increased: +35% for Italy, +23% for Switzerland, and +16% for Germany and the UK.

IV.2. International visibility by country

To check the hypothesis that researchers themselves tend to judge potential partners on the basis of impact more than on publication share, it is instructive to rank those countries present on the international genomics scene by their relative impact.

This exercise is complicated, however, by the fact that calculating the relative impact (a ratio of a citation share to a production share) for countries with very low production share leads to meaningless conclusions. For example, the world leader in impact would be Macao (3.6 relative impact) followed by Haiti (1.4). Therefore Table 10 has been constructed by eliminating countries whose low production share renders their relative impact unusable.

Table 10: World publications share, citations share and relative impact of the top 20 countries as ranked by relative impact

Legend:

- share ≥10 %
- 5 % ≤ share < 10 %
- 1 % ≤ share < 5 %
- 0,5 % ≤ share < 1 %
- share <0,5

Rank by publication	Rank by impact	Country	Publication share	Citation share	Relative impact
1	1	UNITED STATES	40	52.5	1.31
12	2	SWITZERLAND	1.6	1.8	1.15
3	3	UNITED KINGDOM	7.6	7.9	1.05
9	4	NETHERLANDS	2.3	2.1	0.94
4	5	GERMANY	6.8	6.3	0.93
6	6	CANADA	4	3.6	0.9
20	7	AUSTRIA	0.6	0.6	0.88
13	8	BELGIUM	1	0.8	0.88
5	9	FRANCE	5.9	5.1	0.87
15	10	ISRAEL	0.9	0.7	0.85
14	11	FINLAND	0.9	0.7	0.83
11	12	SWEDEN	1.8	1.5	0.81
33	13	IRELAND	0.2	0.1	0.78
17	14	DENMARK	0.8	0.6	0.76
8	15	AUSTRALIA	2.4	1.8	0.73
7	16	ITALY	3.1	2.2	0.72
2	17	JAPAN	9.9	7.1	0.71
34	18	SINGAPORE	0.2	0.1	0.66
24	19	NORWAY	0.4	0.3	0.65
10	20	SPAIN	2	1.2	0.6
		WORLD	*100*	*100*	*1*

Source ISI, Traitements INSERM-OST

World relative impact, which is by design 1, is based on a calculated total of 399,026 articles and 1,741,131 citations for the whole of the period under study.

The cut-off point for this table has been set at the first 20 countries since the 20^{th}, Spain, has an impact index of only 0.6. The world distribution of citations is even more concentrated than that of production share.

The rank ordering displayed by this table makes it easier to grasp the visibility of publications in genomics by certain small countries and to confirm – by comparing columns 1 and 2 – that there is no automatic correlation between production share and impact. The two countries whose impact clearly surpasses 1 are the US and Switzerland. By the same token, countries such as Austria, Israel, or Finland, whose contribution to world publication is less that 1%, enjoy very respectable impact index numbers.

Countries like Korea, Taiwan, Brazil, China, India, and the Russian Federation, whose impact is either low or very low (see Table 6), do not make it onto this list.

France is only 9^{th} in visibility for the overall period although it is the world's 5^{th} producer, a sizeable differential and one which puts it well down the list from both the UK and Germany.

To conclude, the preceding data and the light shed on them by this last table lead to the hypothesis that any "cognitive" hierarchy which genomic researchers may tend to establish among various producer countries is likely to be based on a mix of the size of a country's output and its international scientific visibility. Moreover, partnership data show that any such "international hierarchy" among countries must be adjusted, country by country, to take into account historical affinities of geographical proximity, language, or global region.

Part II – The European Union: Countries and regions

In this part of the study we take a look at similar questions but within the framework of the European Union, composed of 15 members (EU15). All of the indicators in this section are calculated on the basis of these 15 countries.

This shift in the terms of reference makes it easier to grasp the situation of the 15 Member States of the Union comprising a European research area, and to observe what goes on among Europe's regions[13].

For the period 1993-2000, the EU15 accounted for 33.3% of world genomics research, 33.8% of world research in fundamental biology, and 38.3% of all medical research published in the world.

V. Genomics research in European Union Member States

V.1. The relative position of the fifteen Member States

V.1.1. Relative position for the whole period of study

The relative performance in genomics research of the countries of EU15 within the European research area is portrayed in Table 12. A table comparing their performance in genomics with that of fundamental biology and medical research is provided in the Annex (see Annex in French).

[13] The US share of the world scientific publications and citations in genomics represents 40% and 50%, respectively. It is therefore appropriate to exclude the US in order to apprehend the shifts for the other countries.

Table 11: European share of publications, European citations share and relative impact for EU15 countries for the overall period

Legend:
- share ≥10 %
- 5 % ≤ share < 10 %
- 1 % ≤ share < 5 %
- 0,5 % ≤ share < 1 %
- share <0,5

Rank by publication	Rank by impact	Country	European publication share	European citation share	Relative impact
1	1	UK	22.8	27	1.18
2	3	GERMANY	20.4	21.3	1.05
3	6	FRANCE	17.8	17.5	0.98
4	11	ITALY	9.3	7.5	0.81
5	2	NETHERLANDS	6.8	7.2	1.07
6	12	SPAIN	6	4	0.67
7	8	SWEDEN	5.4	5	0.92
8	5	BELGIUM	2.9	2.9	1
9	7	FINLAND	2.7	2.5	0.93
10	10	DENMARK	2.4	2	0.86
11	4	AUSTRIA	1.9	1.9	1
12	15	GREECE	0.8	0.4	0.54
13	9	IRELAND	0.6	0.5	0.88
14	13	PORTUGAL	0.4	0.3	0.62
15	ns	LUXEMBURG	0	ns	ns
		EU 15	*100*	*100*	*1*

Source ISI, Traitements INSERM-OST

The United Kingdom, Germany, and France each account for approximately one-fifth of total EU production. Italy, the Netherlands, Spain, and Sweden all fall within between 5 and 10%, while the eight others are all at less than 3%.

The European impact index for the UK, Germany, and the Netherlands are noticeably higher than the EU15 average, while Austria, Belgium and France are located around the average, and the others lie below this point.

For a certain number of countries, the differential between production share and impact rank stands out; Italy and Spain are 4th and 6th in production share and 11th and 12th in impact index, while Austria (which like the UK is oriented more towards medical research) is 11th in production but 4th in impact.

France is 3rd in production but only 6th in relative impact index, quite far behind the UK and Germany.

V.1.2. Change in position over the period of study

From 1993 to 2000 EU15 countries increased their total world share from 32.5% to 33.8% (from 13,000 to 20,300 articles annually). Its share of world production of fundamental biology research rose from 33.2% to 34% (31,000 to 34,000 articles per year), and from 37.7% to 38.4% in biomedical research (59,000 to 65,000 articles per year).

Table 12 takes a look at change in production and impact in genomics research by country. Changes for fundamental biology and biomedical research are provided in the Annex.

Table 12: Change in the principal indicators over the period of study

Legend:

- share ≥10 %
- 5 % ≤ share < 10 %
- %1 ≤ share < 5%
- 0,5 % ≤ share < 1 %
- share <0,5

Rank of publication	Rank of impact	Country	Share of publications				Relative impact			
			[1995]	[1998]	[2000]	Change over the period	[1995]	[1998]	[2000]	Change over the period
1	1	UK	24.2	22.5	21.9	-9%	1.18	1.16	1.2	2%
2	3	GERMANY	19.4	20.7	20.9	8%	1.06	1.07	1.04	-2%
3	6	FRANCE	18.7	18.1	16.7	-11%	0.98	1.01	0.98	-1%
4	11	ITALY	8.8	9.3	9.6	9%	0.75	0.84	0.82	9%
5	2	NETHERLANDS	7.3	6.7	6.4	-13%	1.07	1.06	1.04	-3%
6	12	SPAIN	4.9	5.9	6.8	39%	0.61	0.66	0.71	17%
7	8	SWEDEN	5.5	5.5	5.3	-5%	0.93	0.89	0.91	-2%
8	5	BELGIUM	2.9	2.7	2.9	0%	1.05	1	0.96	-8%
9	7	FINLAND	2.6	2.7	2.8	7%	0.93	0.9	0.94	1%
10	10	DENMARK	2.4	2.3	2.5	6%	0.78	0.86	0.95	21%
11	4	AUSTRIA	1.8	1.9	2.1	21%	1.01	0.89	1.07	6%
12	15	GREECE	0.7	0.8	0.9	25%	0.47	0.52	0.58	24%
13	9	IRELAND	0.5	0.5	0.6	23%	0.77	0.84	0.96	25%
14	13	PORTUGAL	0.3	0.4	0.6	102%	0.61	0.53	0.63	3%
15	14	LUXEMBOURG	0	0	0	ns	ns	ns	ns	ns
		EU 15	*100*	*100*	*100*		*1*	*1*	*1*	

Source ISI, Traitements INSERM-OST

When Europe is taken as the overall reference instead of the world, comparison among European countries yields more legible results.

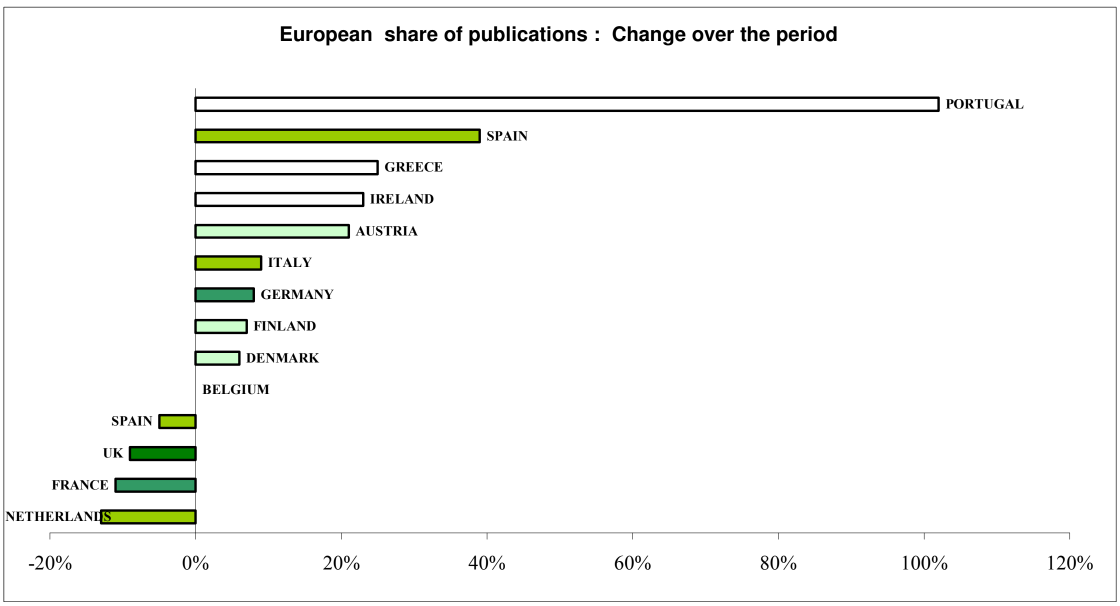
Over the period, the impact of the three largest countries of production remained stable, but France and the UK lost production share, while Germany's share rose. Among the group of countries lying between 5-10% shares of European production, Spain's strong progress stands out, as does Italy's positive growth and the erosion of the Swedish and Dutch shares.

All of the nations possessing production share less than 3% showed progress; Denmark improved its relative impact, Austria, Greece, and Ireland improved both production and relative impact, while Portugal showed a big increase in its production.

V.1.3. Graphic representations of change in production share and impact

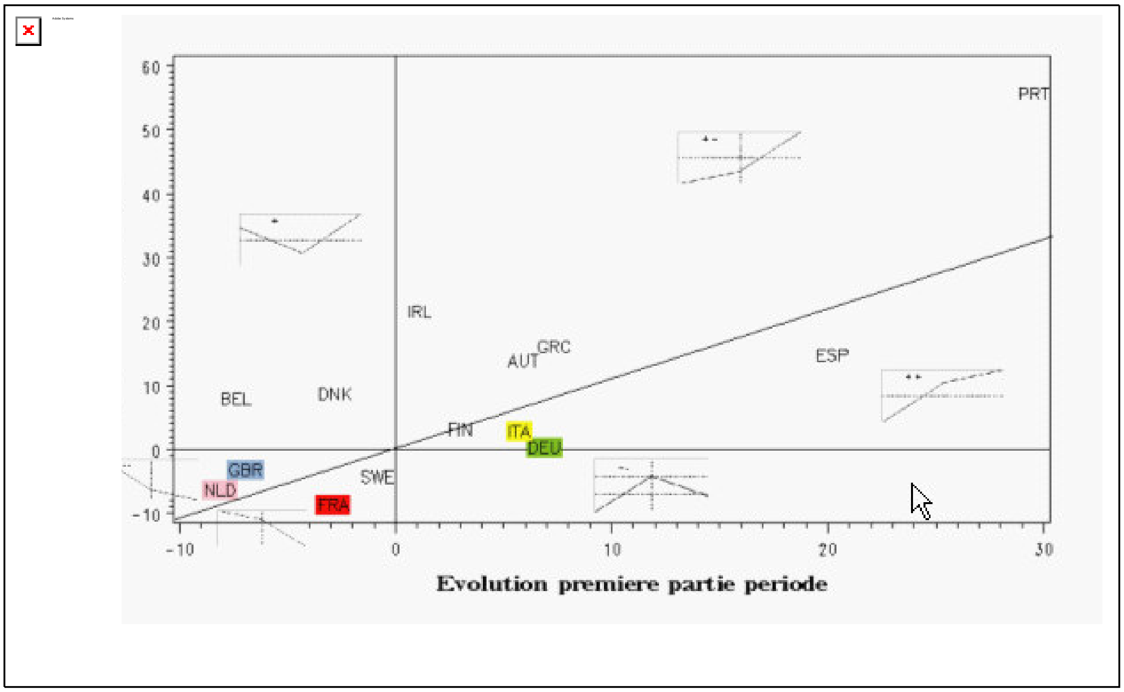
Charting these changes over time, as below, displays in more precise terms the sizeable differences among European countries.

Figure 3: Change in European share of publication of EU member countries over the period



Trends over the period can be compared visually using graphing techniques by which the x-axis shows change in production share (Figure 4) or impact index (Figure 5) between 1995 and 1998 while the y-axis shows the same type of change for the period 1998-2000. As in Figure 1, the straight line or isocurve which connects equal points on the two axes makes readily apparent whether the change over the second period is greater (upper section of the graph) or less (lower section) than the change that occurred over the first period. The icons stand for the type of change corresponding to each section of the graph.

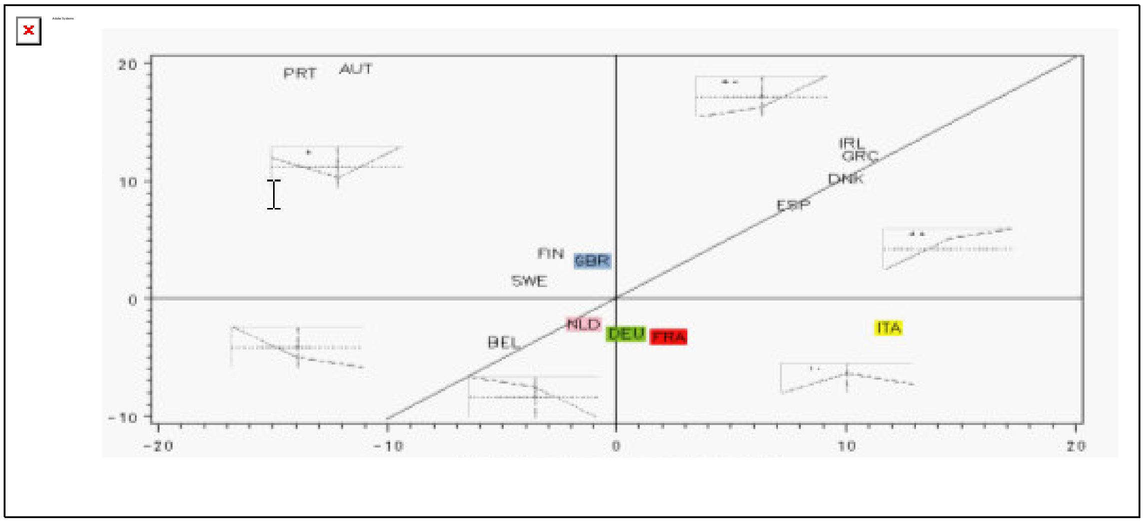
Figure 4: Details of the change in European share of publication in EU member countries



According to this figure, Portugal's strong positive change is confirmed, as are the other observed cases of positive change, in Ireland, Greece and Austria. Finland's slow steady progress can be noted, as well as Spain's rapid rate of growth over the first period followed by a slowdown in the second period. Similarly, strong positive change in Italy and Germany occurred mostly in the first period.

Finally, the four countries that lost share over the whole period can be differentiated from each other, with share loss in Sweden and France accelerating in the second period while the Netherlands and the UK slowed their loss of ground in the second period. Therefore, at the end of the second period, the Netherlands and the UK can be seen as righting their losses at the same time as share loss in France is accelerating.

Figure 5: Details of the change in European impact in EU member countries



Graphing changes in European country impact brings out the differences among rates of change of individual countries, and the fact that France, for example, which experienced a positive rate overall, nevertheless was strongly positive in impact change during the first half while showing negative growth in impact over the second period. In contrast, the UK and Sweden show loss of impact at first, only to rebound positively in the second period.

V.2. Collaboration among European Union member countries

The six first partners of the seven top producing countries have been assembled in a table displaying patterns of collaboration (table 13; a systematic table showing all 15 countries is provided in the Annex).

Table 13: Table of partnerships for the top 7 producer countries of the EU15

(a) : partner country's European rank by publications share

(b): partner country's European rank by impact

Country under study	Partner country	Number of co-publications for the period		Share in the co-publications of the country under study		(a)	(b)
		Annual average	Change over the period	Annual average	Change over the period		
UK	*UK*	*2 939*	*55%*	*69,30%*	*-3%*	*1*	*1*
	Germany	292	92%	6,90%	21%	2	3
	France	273	81%	6,40%	13%	3	6
	Italy	172	129%	4,00%	44%	4	11
	Netherlands	171	76%	4,00%	10%	5	2
	Sweden	92	97%	2,20%	24%	7	8
Total co-publications			59%				
Germany	*Germany*	*2 675*	*88%*	*69,80%*	*3%*	*2*	*3*
	UK	292	92%	7,60%	6%	1	1
	France	218	80%	5,70%	-1%	3	6
	Netherlands	151	87%	3,90%	3%	5	2
	Italy	133	108%	3,50%	14%	4	11
	Austria	110	112%	2,90%	16%	11	4
Total co-publications			82%				
France	*France*	*2 500*	*51%*	*73,90%*	*-2%*	*3*	*6*
	UK	273	81%	8,10%	17%	1	1
	Germany	218	80%	6,40%	16%	2	3
	Italy	142	108%	4,20%	35%	4	11
	Belgium	114	90%	3,40%	23%	8	5
	Netherlands	97	74%	2,90%	13%	5	2
Total co-publications			54%				
Italy	*Italy*	*1 390*	*84%*	*74,30%*	*2%*	*4*	*11*
	UK	172	129%	9,20%	27%	1	1
	France	142	108%	7,60%	16%	3	6
	Germany	133	108%	7,10%	15%	2	3
	Netherlands	60	133%	3,20%	30%	5	2
	SPAIN	49	232%	2,60%	84%	6	12
Total co-publications			80%				
Netherlands	*Netherlands*	*940*	*49%*	*67,00%*	*-4%*	*5*	*2*
	UK	171	76%	12,20%	13%	1	1
	Germany	151	87%	10,70%	20%	2	3
	France	97	74%	6,90%	12%	3	6
	Belgium	77	53%	5,50%	-1%	8	5
	Italy	60	133%	4,30%	50%	4	11
Total co-publications			55%				
Spain	***Spain***	*860*	*138%*	*75,90%*	*4%*	*6*	*12*
	UK	89	99%	7,80%	-13%	1	1
	France	81	118%	7,10%	-5%	3	6
	Germany	66	105%	5,80%	-10%	2	3
	Italy	49	232%	4,30%	46%	4	11
	Netherlands	28	200%	2,50%	32%	5	2
Total co-publications			128%				
Sweden	*Sweden*	*736*	*68%*	*67,10%*	*1%*	*7*	*8*
	UK	92	97%	8,30%	18%	1	1
	Germany	81	107%	7,30%	24%	2	3
	Finland	62	157%	5,70%	53%	9	7
	France	62	72%	5,60%	3%	3	6
	Denmark	51	34%	4,70%	-20%	10	10
Total co-publications			67%	*Source ISI*	*Traitements Inserm-OST*		

Endogenous collaboration is slightly higher in the three "Latin" countries of the selected set – France, Italy, Spain, (about 75% compared to 70% average).

The United Kingdom, leading country in both production share and impact, is the first partner of all the countries studied and, more widely, of the entire EU15 set with the exception of Austria, Finland, Belgium, and Luxemburg (see complete table in Annex). Its importance as a scientific partner can be seen in the fact that for all six other top countries it accounts for an average 10% of the partner's European copublication (from 7% for Germany to 12% for the Netherlands), which makes it the pivotal country in EU genomics research.

France appears often as the second most important partner, which can be accounted for by a strong production combined with geographical and linguistic affinities with certain other countries. Germany, which occupies a stronger position in both production and impact is a larger partner than France only in the case of Sweden and the Netherlands.

The number of articles copublished by individual pairings of countries in no case accounts for more than a small fraction of total copublication within the EU15. A few pairings produce more than 1% of the European total: Germany/UK, France/UK, France/Germany, France/Italy, UK/Italy, UK/Netherlands, Germany/Netherlands. Even taking into consideration that countries producing the most volume will also be leaders in copublication, it is still worth noting that the UK is the European pivot for copublication in genomics, and that France has also a predominant position in copublication.

Observing trends, one notices that total intra-European copublication increased over the period of study, and that in most countries the increase in total copublication was greater than simply the increase in endogenous copublication. In most cases it is also true that there was consolidation over the period in individual partnerships between countries.

Of all EU countries, France increased its copublications the least over the period of study. When examined more closely, the French data reveal that the growth in copublication with the UK, Germany, Italy, Belgium and Netherlands was nearly offset by slow growth and share loss of endogenous copublication. While total French copublication grew 54% over the period, endogenous copublication only increased 51% whereas copublication with Italy rose 108%, with the UK it rose 81%, and Germany 80%.

Turning to the share of other countries' copublication accounted for by copublication with France, one finds a relative decrease in the case of certain countries (Spain, Italy, Germany) or an increase which occurred primarily in the first period (UK, Italy, Sweden). An interesting exception is that of the Netherlands, for whom growth in French copublication occurred almost entirely in the second period.

Summing up, it can be said that measuring copublication among EU member countries makes manifest the real existence of a European Research Area in the field of genomics. Although the endogenous collaboration of all countries grew over the period of study, for most countries the growth in European copublication grew even faster, which suggests that EU15 countries actively sought research collaboration with the other members of the Union.

VI. Genomics Research in Europe's Regions

Much of research activity in genomics requires an infrastructure of specialized technical skills and equipment which are most often organised around easily identified regional points.

To facilitate a look at Europe by region, this study makes use of a scheme of regional division established by Eurostat which is based on regional administrative reality but takes into account the need for comparability by population size. The division of Europe into regions that is used in this study cuts the 15 member countries into 226 regions. It should be noted that, according to this division, all of Denmark counts as one region – since it is about that size, while certain regions such as "Ile-de-France" or "Inner London" are the equivalent in size to some small countries. France is divided according to the existing administrative Regions.

VI.1. Relative position of the top 20 producer regions in Europe and all French regions for the period of study

Table 14 shows the principal indicators for the top 20 producers among European regions as well as the regions of France (Corsica and overseas regions are not included, due to poor quality data).

A map showing the 20 top European regions is provided at the end of this chapter, and three maps are provided in the Annex to enable comparison among the top 20 regions in genomics, fundamental biology, and medical research.

Table 14: Principal indicators for the top 20 producer regions in Europe and for the regions of France over the period of study (Corsica and overseas regions excluded)

Rank by publication	Region	European publication share	European citation share	Relative impact
1	ILE-DE-FRANCE (FRA)	8.5	9.2	1.08
2	INNER LONDON (RU)	5.5	7.0	1.27
3	EAST ANGLIA (RU)	2.8	4.5	1.6
4	DENMARK (DNK)	2.4	2.0	0.86
5	BERKSHIRE&OXFORDSHIRE(RU	2.3	3.1	1.33
6	OBERBAYERN (DE)	2.2	2.5	1.16
7	STOCKHOLM (SW)	2.1	2.3	1.07
8	EASTERN SCOTLAND (RU)	2.1	2.5	1.19
9	KARLSRUHE (DE)	2.1	3.4	1.61
10	COMUNIDAD DE MADRID (ES)	2.1	1.6	0.75
11	LOMBARDIA (IT)	2.1	2.0	0.99
12	ZUID-HOLLAND (NL)	2.0	2.3	1.12
13	BERLIN (DE)	1.8	1.9	1.06
14	RHONE-ALPES (FRA)	1.6	1.4	0.89
15	LAZIO (IT)	1.6	1.3	0.8
16	NOORD-HOLLAND (NL)	1.5	1.9	1.27
17	KOELN (DE)	1.4	1.5	1.06
18	GELDERLAND (NL)	1.4	1.2	0.88
19	CATALUNA (ES)	1.4	1.0	0.74
20	UUSIMAA (FI)	1.3	1.3	1.02
	Other French regions:			
25	PROVENCE-ALPES-COTE D'AZUR	1.2	1.2	1.04
26	ALSACE	1.1	1.6	1.52
27	LANGUEDOC-ROUSSILLON	1.1	0.9	0.79
42	MIDI-PYRENEES	0.8	0.8	1.02
50	AQUITAINE	0.6	0.4	0.64
52	NORD-PAS-DE-CALAIS	0.6	0.6	1.04
57	BRETAGNE	0.5	0.3	0.68
83	CENTRE	0.3	0.2	0.49
84	PAYS-DE-LA-LOIRE	0.3	0.2	0.64
92	AUVERGNE	0.3	0.1	0.56
100	LORRAINE	0.2	0.1	0.56
108	BOURGOGNE	0.2	0.1	0.64
118	POITOU-CHARENTES	0.1	0.1	0.5
120	HAUTE NORMANDIE	0.1	0.1	0.76
138	CHAMPAGNE-ARDENNE	0.1	0.0	0.55
140	FRANCHE-COMTE	0.1	0.0	0.42
143	BASSE NORMANDIE	0.1	0.1	0.74
147	PICARDIE	0.1	0.0	0.62
149	LIMOUSIN	0.1	0.0	0.53

Sources ISI, Traitements INSERM-OST

The top 20 regions in Europe for genomic research production include 2 French regions, 4 UK regions, 4 German regions, 3 Dutch regions, 2 Spanish regions, and 2 Italian regions. Two regions only – both very unusual for their potential – account for more than 5% of European production each: Ile-de-France and Inner London. The next 10 regions lie somewhere over 2%.

By constructing a table of the top 20 regions for impact, it is possible to confirm the prestige enjoyed by the UK in genomic research: it possesses eight of the 20 most visible regions in Europe. Two among the top 20 are located in France, but not the same two as for production share; Alsace comes in third among most visible European regions while Ile-de-France finishes 19th in European impact. Two regions, East Anglia (Cambridge) and Karlsruhe, boast an exceptional impact index of 1.6, with Alsace coming in at 1.52.

This table also displays the relative position of French regions within the European research area, for genomic research.

- Ile-de-France, which is far and away the leading producer in Europe, enjoys a relative impact index slightly higher than the European average.
- France's second strongest region is Rhone-Alpes which is 14th in production share, with a relative impact noticeably lower than the European average.
- The next three French regions, Provence-Alpes-Côte d'Azur, Alsace, and Languedoc-Roussillon, are situated between the 25th and 27th place.
- The French region Midi-Pyrénées follows, in 42nd place. Its share of European production is less than 1%.
- Six other French regions come in between 50th and 99th place, another nine lie between 100th and 150th place.
- Finally, only five regions have a relative impact greater than the European average: Ile-de-France, Alsace, Provence-Alpes-Côte d'Azur, Midi-Pyrénées, Nord-Pas-de-Calais.

Taken all together, we could conclude that about five of France's regions do have a level of production and visibility that places them in the top quarter of Europe's regions.

VI.2. Change in relative position of the top 20 producer regions in Europe and three more French regions

Only those regions whose production share was greater than 1% were included in this table. Production and impact changes of the weakest producer regions ought to be considered with caution, since often the sample size is too small to give entirely reliable figures.

Table 15: Change in the principal indicators over the period of study for the top 20 European regions and 3 French regions

(a): rank by publication

(a)	Region	European publication share				Relative impact			
		[1995]	[1998]	[2000]	Change over the period	[1995]	[1998]	[2000]	Change over the period
1	ILE-DE-FRANCE	9.7	8.7	7.5	**-22%**	1.09	1.07	1.08	**-1%**
2	INNER LONDON	5.9	5.4	5.4	**-7%**	1.31	1.22	1.25	**-5%**
3	EAST ANGLIA	3	2.8	2.7	**-9%**	1.54	1.47	1.68	**9%**
4	DENMARK	2.4	2.3	2.5	**6%**	0.78	0.86	0.95	**22%**
5	BERKSHIRE & OXFORDSHIRE	2.6	2.4	2.1	**-16%**	1.3	1.34	1.41	**8%**
6	OBERBAYERN	2.3	2.2	2.1	**-7%**	1.14	1.14	1.25	**9%**
7	STOCKHOLM	2.2	2.2	2.1	**-7%**	1.12	1.01	1.05	**-7%**
8	EASTERN SCOTLAND	2.2	2	2.1	**-6%**	1.17	1.22	1.2	**3%**
9	KARLSRUHE	2.3	2.1	1.8	**-20%**	1.72	1.66	1.49	**-13%**
10	COMUNIDAD DE MADRID	1.9	2.1	2.2	**15%**	0.72	0.73	0.76	**6%**
11	LOMBARDIA	2	2.1	2.1	**5%**	0.88	0.97	1.03	**18%**
12	ZUID-HOLLAND	2.2	2	1.9	**-17%**	1.2	1.04	1.07	**-11%**
13	BERLIN	1.7	1.8	1.9	**17%**	0.97	1.08	1.13	**16%**
14	RHONE-ALPES	1.6	1.6	1.6	**-1%**	0.78	0.98	0.92	**17%**
15	LAZIO	1.6	1.5	1.5	**-5%**	0.72	0.8	0.83	**15%**
16	NOORD-HOLLAND	1.6	1.6	1.4	**-14%**	1.25	1.31	1.29	**3%**
17	KOELN	1.5	1.4	1.5	**4%**	1.16	1	1.02	**-12%**
18	GELDERLAND	1.6	1.4	1.2	**-20%**	0.88	0.81	0.85	**-4%**
19	CATALUNA	1	1.4	1.6	**58%**	0.69	0.65	0.83	**21%**
20	UUSIMAA	1.3	1.3	1.3	**-4%**	1.02	1.01	0.96	**-5%**
	Additional French regions:								
25	PROVENCE-ALPES-COTE D'AZUR	1.1	1.2	1.2	**10%**	1.01	1.12	1.04	**3%**
26	ALSACE	1.2	1.1	1	**-23%**	1.58	1.67	1.42	**-10%**
27	LANGUEDOC-ROUSSILLON	1	1.1	1.1	**5%**	0.79	0.68	0.87	**11%**
	Total UE 15 regions	*100*	*100*	*100*		*1*	*1*	*1*	

Sources ISI, Traitements INSERM-OST

Figure 6: Change in European share of publication of top 20 European regions over the period

Legend:

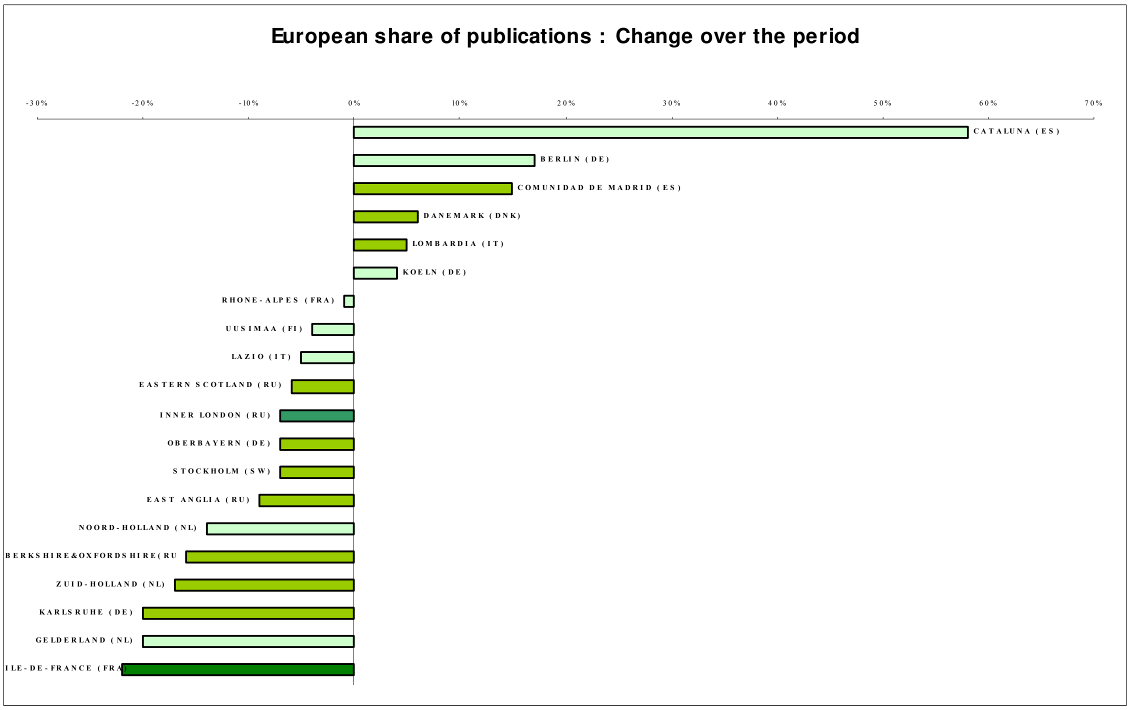
share ≥10 %

5 % ≤ share < 10 %

2 % ≤ share < 5 %

1 % ≤ share < 2 %

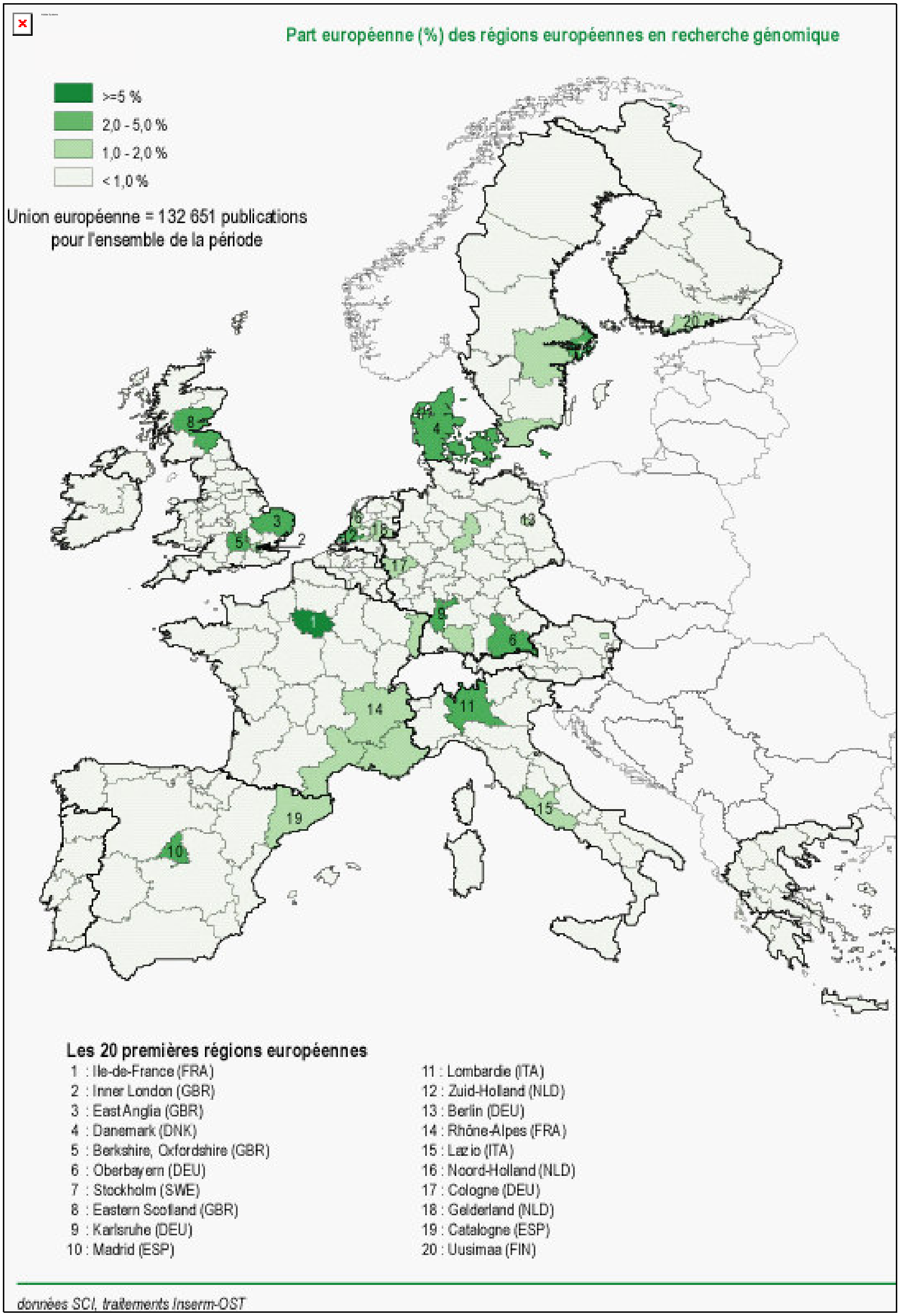
share <1%



In a certain number of regions where genomic research is well-established (Paris, London, Cambridge, Oxford, Karlsruhe, etc.), production share shows a clear drop over the period of study. This may well be due to the arrival of new regions on the genomics research forefront – as seems to be the case with Catalonia (Barcelona) or Madrid. Furthermore, there is the likelihood that a change in the criteria for evaluating projects for funding purposes is partly responsible, at least in some regions like Cambridge (East Anglia), Oxford, or Eastern Scotland where a drop in production share occurs at the same time as an increase in impact.

As it turns out, a number of regions (other than the peculiar case of the nation/region of Denmark) increased to a noticeable degree their visibility: Lombardy, Berlin, Rhône-Alpes, Lazio, and Catalonia. Several regions, however, including Karlsruhe (which is the premier region for impact over the whole period), Zuid-Holland, and Cologne, saw their impact diminish.

Ile-de-France, which experienced a decrease in its publication share over the period, maintained its relative impact at 1.08, enough to place it 10th among European regions. Both the production share and relative impact of Alsace dropped during the period, but this region nonetheless maintained a very high impact by the end of the period. Finally, the three other French regions that can claim European stature, Rhône-Alpes, Provence-Alpes-Côte d'Azur et Languedoc-Roussillon, improved their position. That being said, for both Rhône-Alpes and Provence-Alpes-Côte d'Azur the rate of change was slower in the second part of the period of study.


Part européenne (%) des régions européennes en recherche génomique
>=5 %
2,0 - 5,0 %
1,0 - 2,0 %
< 1,0 %
Union européenne = 132 651 publications
pour l'ensemble de la période
Les 20 premières régions européennes
1 : Ile-de-France (FRA)
2 : Inner London (GBR)
3 : East Anglia (GBR)
4 : Danemark (DNK)
5 : Berkshire, Oxfordshire (GBR)
6 : Oberbayern (DEU)
7 : Stockholm (SWE)
8 : Eastern Scotland (GBR)
9 : Karlsruhe (DEU)
10 : Madrid (ESP)
11 : Lombardie (ITA)
12 : Zuid-Holland (NLD)
13 : Berlin (DEU)
14 : Rhône-Alpes (FRA)
15 : Lazio (ITA)
16 : Noord-Holland (NLD)
17 : Cologne (DEU)
18 : Gelderland (NLD)
19 : Catalogne (ESP)
20 : Uusimaa (FIN)
données SCI, traitements Inserm-OST

Part III – France: Regions and Genopoles

VII. French Genomics Research and the Relative Position of Regional Genopoles

In this part of the study the chosen field of reference for calculating indicators is France, in order to analyse the production levels and dynamic trends for French regions throughout the period of study, 1993-2000 (tables 16 and 17). Also, with French production as the reference, France's Genopoles – regional clusters of public, university, and industrial research capacity, new firm incubators and technology parks, focused on genomics and related fields - can be arranged in a "virtual ranking" table showing the principal indicators of the regions or groupings of regions awarded the label "Genopole" by the Genomics program in 1999-2000 (table 18).
Maps displaying publication share in genomics, fundamental biology, and medical research are also provided in the Annex.

VII.1. Analysis by region of French production of genomics research

VII.1.1. Relative position of French regions in genomics research production

Table 16 presents production share, citation share and relative impact of those French regions whose production share is greater than 1% of total French production (or approximately 50 articles per year) for the overall period. The twelve regions thus selected are also active in fundamental biology and, often, medical research (see maps in Annex).

Table 16: Principal indicators of genomics research for the top 12 producer regions in France for the whole period

Legend:

- share ≥10 %
- 5 % ≤ share < 10 %
- 2 % ≤ share < 5 %
- 1 % ≤ share < 2 %
- share <1

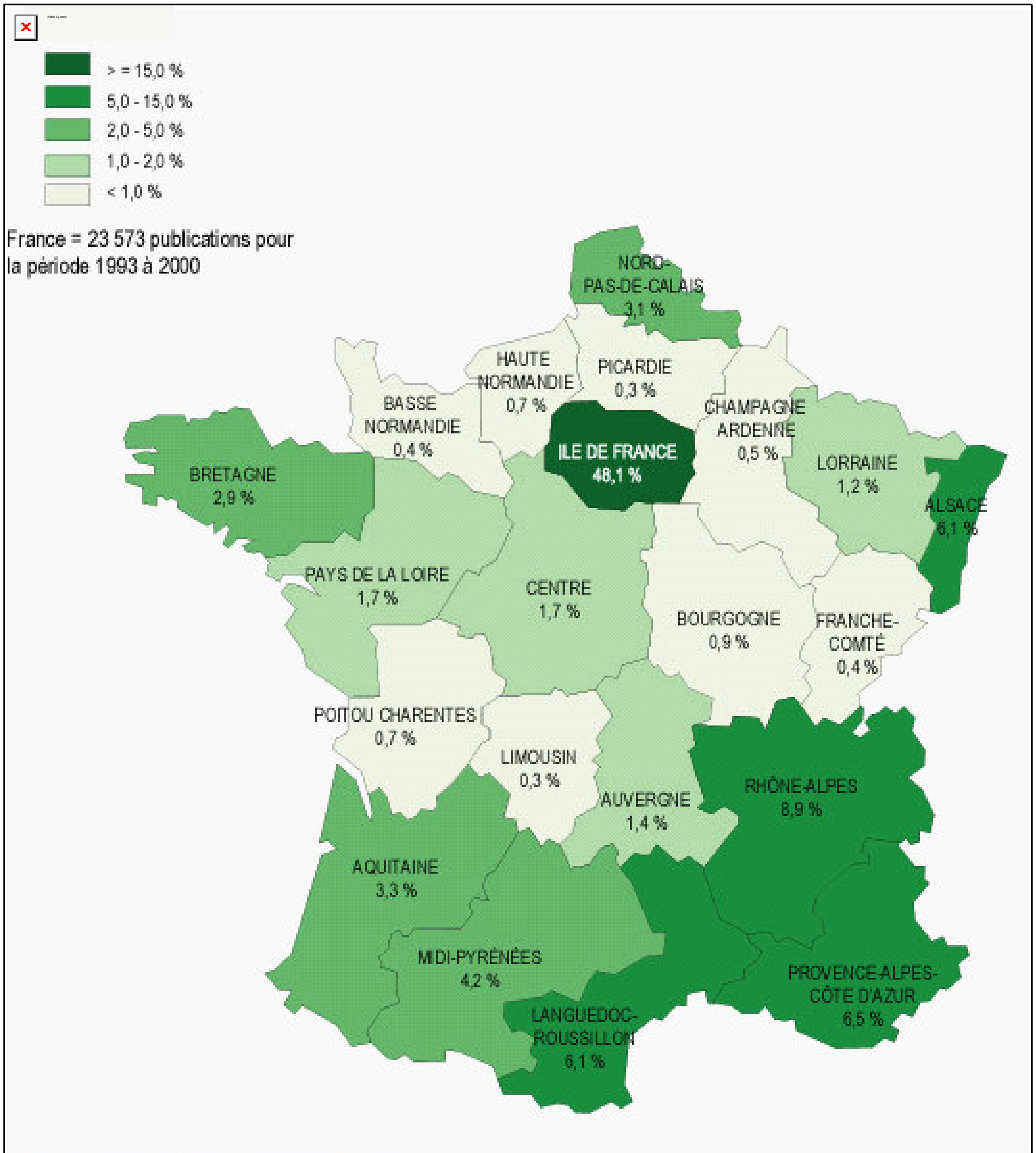
Rank by publication	Rank by impact	Region	French publication share	French citation share	Relative impact
1	2	ILE-DE-FRANCE	48.1	52.7	1.10
2	6	RHONE-ALPES	8.9	8.0	0.90
3	4	PROVENCE-ALPES-COTE D	6.5	6.9	1.05
4	1	ALSACE	6.1	9.4	1.54
5	7	LANGUEDOC-ROUSSILLON	6.1	4.9	0.80
6	5	MIDI-PYRENEES	4.2	4.4	1.04
7	10	AQUITAINE	3.3	2.2	0.65
8	3	NORD-PAS-DE-CALAIS	3.1	3.3	1.06
9	8	BRETAGNE	2.9	2.0	0.69
10	9	PAYS-DE-LA-LOIRE	1.7	1.1	0.66
11	11	AUVERGNE	1.4	0.8	0.57
12	12	LORRAINE	1.2	0.7	0.57
		Ensemble régions France	*100*	*100*	*1*

Sources ISI, Traitements INSERM-OST

In genomics research, only five French regions produce more than 5 % of the national total, one of which is the highly particular case of Ile-de-France since it concentrates nearly half of all French production. The second region of production, Rhône-Alpes, accounts for almost 10 % of French publication in genomics. The next three regions, Provence-Alpes-Côte d'Azur, Alsace, Languedoc-Roussillon, account for 6-6.5 % of national production. Then comes Midi-Pyrénées with 4 % of national production, followed by three regions around 3 %: Aquitaine, Nord-Pas-de-Calais and Bretagne. Lastly, Pays-de-la-Loire, Auvergne and Lorraine only account for 1-2 % of national production[14].

As for differences between production share rank and impact rank (comparing columns 1 and 2 of the table), noteworthy differences include a five rank difference for Nord-Pas-de-Calais (8th in production and 3rd in impact), a four rank difference for Rhône-Alpes (2nd in production but 6th in impact), and a three rank difference for both Alsace (4th in production but 1st in impact) and Aquitaine (7th in production and 10th in impact).

[14] For production ranks lower than 5th, that is for regions with less than 5% of national production, the sample size of articles is very small, thus possibly giving undue impact resonance to an article or two that happen to have particularly strong or weak citation numbers; this adverse affect on impact indicator calculation does not exist in the case of indicators of performance over the whole period of study (an eight-year average), but should be kept in mind for indicators showing change over shorter periods of two or three years.



données ISI (SCI et COMPUMATH), traitements OST

VII.1.2. Change in relative position of French regions

Table 17 reveals in detail the change in the principal indicators for the 12 regions whose production exceeds 1% of the national total. Displayed values ought to be treated with some reserve since the sample size on which they are based is in some cases rather small.

Table 17: Change in relative position of French regions

Legend:

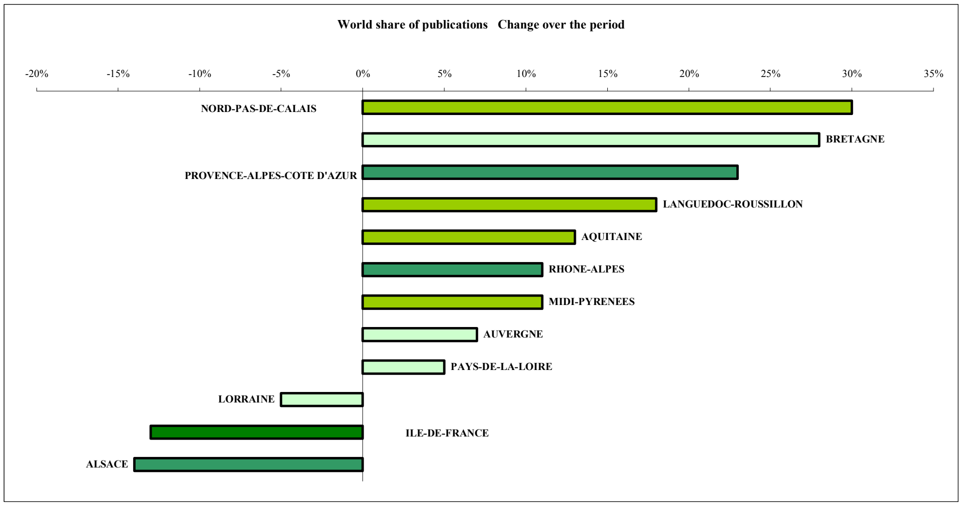
- share ≥10 %
- 5 % ≤ share < 10 %
- 2 % ≤ share < 5 %
- 1 % ≤ share < 2 %
- share <1

(a): region's French publication share rank
(b): region's French relative impact rank

(a)	(b)	Region	Share of publications				Relative impact			
			[1995]	[1998]	[2000]	Change over the period	[1995]	[1998]	[2000]	Change over the period
1	2	ILE-DE-FRANCE	51.9	47.9	45.2	**-13%**	1.11	1.07	1.11	**0%**
2	6	RHONE-ALPES	8.5	9	9.4	**11%**	0.8	0.97	0.94	**18%**
3	4	PROVENCE-ALPES-COTE D'AZUR	5.8	6.5	7.2	**23%**	1.03	1.12	1.07	**3%**
4	1	ALSACE	6.6	6	5.7	**-14%**	1.61	1.66	1.46	**-9%**
5	7	LANGUEDOC-ROUSSILLON	5.3	6.3	6.3	**18%**	0.8	0.68	0.9	**12%**
6	5	MIDI-PYRENEES	3.8	4.5	4.3	**11%**	0.92	1.24	0.98	**7%**
7	10	AQUITAINE	3.1	3.3	3.5	**13%**	0.62	0.63	0.68	**9%**
8	3	NORD-PAS-DE-CALAIS	2.7	3	3.5	**30%**	1.04	1.03	0.96	**-8%**
9	8	BRETAGNE	2.5	2.9	3.2	**28%**	0.56	0.65	0.78	**38%**
10	9	PAYS-DE-LA-LOIRE	1.5	1.6	1.6	**5%**	0.71	0.58	0.68	**-4%**
11	11	AUVERGNE	1.5	1.2	1.6	**7%**	0.54	0.49	0.61	**14%**
12	12	LORRAINE	1.2	1.3	1.2	**-5%**	0.49	0.64	0.61	**25%**

Source ISI, Traitements INSERM-OST

Figure 7: Change in share of publication in French regions over the period



Ile-de-France's publication share decreased regularly over the period of study. It's relative impact decreased over the first part of the period but between 1998 and 2000 it rose to regain its initial position. Alsace's production share decreased steadily while its extremely high impact decreased noticeably over the second half of the period.

Over the whole period, the most noteworthy positive trends in production share and impact were recorded by Bretagne, Rhône-Alpes, Languedoc-Roussillon, and Auvergne.

The region Provence-Alpes-Côte d’Azur increased its production share noticeably while maintaining a relative impact above the national average. In contrast, the region Nord-Pas-de-Calais and, to a lesser extent, Pays-de-la-Loire appear to have increased their publication share to the detriment of their publications' impact. This may well be explained as a sign of growth whereby a new actor entering the field must first focus on increasing production and then – in a maturing phase – work to improve impact.

VII.2. Relative position of French Genopoles compared to all regions

Bibliometric data can be employed in a interesting exercise which consists of cutting up regions and departments to assemble entities that correspond to the officially labeled "genopoles" as established by the Genomics program, and then calculating the principal indicators for these actors.

The Genomics program ordained seven regional zones comprising major metropolitan areas (or twinned areas): Lille for Nord-Pas-de-Calais, Lyon/Grenoble for Rhône-Alpes, Marseilles for Provence-Alpes-Côte d'Azur, Montpellier for Languedoc-Roussillon, Strasbourg/Nancy for Alsace-Lorraine, and Toulouse for Midi-Pyrénées. The "Genopole Ile-de-France" federates several Parisian locations (the Pasteur Institute, Montagne Sainte-Geneviève in Paris' university quarter, Necker-Cochin Paris hospital complex, the Gustave Roussy Institute for cancer research, Paris-Sud, CEPH-Saint-Louis Hospital), and the Genopole of Evry, itself located in the department of Essonne, south of Paris. Two other regional research areas have been designated as trial-phase genopoles: Rennes/Nantes for Bretagne/Pays-de-la-Loire and Bordeaux for Aquitaine.

The Genopoles in this way can simply be taken as geographical-political entities, for which the usual bibliometric indicators can be calculated. For six of them, the Genopole's role is to harmonize the laboratories of one entire region (Nord-Pas-de-Calais, Midi-Pyrénées, Rhône-Alpes, Languedoc-Roussillon, Provence-Alpes-Côte d'Azur, and lastly Aquitaine). Next there are Genopoles which combine two regions (Alsace/Lorraine and Bretagne/Pays-de-la-Loire). Lastly, the study cuts the Evry/Ile-de-France Genopole into the department of Essone to cover the Evry site properly speaking, and then takes the other departments of the Ile-de-France region (Paris, Val-de-Marne, Yvelines, Hauts-de-Seine, Seine-Saint-Denis, Val-d'Oise, and Seine-et-Marne) to form an entity labelled "Ile-de-France – others".

Table of genopoles

Area(s)	Region or twinned regions
Strasbourg - Nancy	ALSACE-LORRAINE
Bordeaux	AQUITAINE
Parisian sites	7 *départements* as : "ILE-DE-FRANCE-others"
Rennes - Nantes	BRETAGNE - PAYS-DE-LA-LOIRE
Evry	*département* "ESSONNE"
Montpellier	LANGUEDOC-ROUSSILLON
Toulouse	MIDI-PYRENEES
Lille	NORD-PAS-DE-CALAIS
Marseille	PROVENCE-ALPES-COTE D'AZUR
Lyon - Grenoble	RHONE-ALPES

The first observation to make is that genopole-labelled regions are those regions whose production share is greater than 1% of national production over the whole period (see table 16) with the exception of Auvergne, which ranks 11th nationally with 1.4% of production.

Exploring this observation further, it is possible to take a bibliometric snapshot of the regions selected as Genopoles at the moment of the programme's selection (the smoothed year 2000[15]), calculate the various indicators and show each region's place in the overall picture of French genomic research.

Table 18: Principal indicators for French Genopoles [2000]

Rank by publication	Rank by impact	Genopole	Share of publications	Share of citations	Relative impact
1	3	ILE-de- FRANCE-others	39.9	43.9	1.1
2	7	RHONE-ALPES	9.4	8.9	0.94
3	4	PROVENCE-ALPES-COTE D'AZUR	7.2	7.7	1.07
4	1	ALSACE-LORRAINE	6.9	9	1.31
5	8	LANGUEDOC-ROUSSILLON	6.3	5.7	0.9
6	2	ESSONNE	5.3	6.2	1.17
7	9	BRETAGNE-PAYS-de-la-LOIRE	4.8	3.6	0.74
8	5	MIDI-PYRENEES	4.3	4.2	0.98
9	6	NORD-PAS-DE-CALAIS	3.5	3.3	0.96
10	10	AQUITAINE	3.5	2.4	0.68
		FRANCE	*100*	*100*	*1*

Source ISI, Traitements INSERM-OST

This table repeats for the selected regions the data already presented for the ranking of all French regions, with several noteworthy differences.

Ile-de-France without Essonne (Evry Genopole) accounts for 40% of national production, while the Evry site taken on its own (the department of Essonne) accounts for 5% of national production, in 6th place among all genopoles. Evry's relative impact is noticeably stronger than that of "Ile-de-France – others" as well as of all the other genopoles except Alsace/Lorraine. The grouping of Alsace and Lorraine as one genopole increases the production share total for the new entity but at the cost of a weaker impact, although the new entity's impact is still noticeably higher than that of the other genopoles. The joint Bretagne/Pays-de-La-Loire genopole, which combines two regions showing strong growth over the period 1993-2000, arrives in 7th position, but with a relative impact that remains weak. Lastly, Nord-Pas-de-Calais shows the same production share as Aquitaine but its impact is clearly stronger.

Comparing these results with those of table 16, we see that five of the genopoles were able to reach a European stature at the time they were selected and labelled as genopoles. These elements provide a grasp of the situation of the regional Genopoles at the time they were first identified as such, and can serve as a baseline for a next-phase analysis which looks at bibliometric data for 2002 to observe change in genopole activity and relative position, against the national backdrop.

[15] Indicators are averaged over the two year period 1999-2000. In addition, the indicators were calculated both for windows defined by convention and for the whole period in order to confirm the results obtained for this one time window.

OST
studies : themes has as
its goal to explore R&D systems
from a research and innovation point
of view and thereby to provide a needed
tool for assessing research policies aimed
at specific thematic areas or sectors.

These studies are not based on conventional
indicators but rather on indicators constructed
within a specific perimeter, which is defined
in cooperation with experts in the specific field
under study. In the same fashion, study
contents are developed in response to the
specific characteristics of a field and
in light of the policy objectives
under evaluation.

Observatoire des Sciences et des Techniques
93, rue de Vaugirard
75006 Paris - France
www.obs-ost.fr