ELSEVIER

Research Policy 24 (1995) 959–979

research policy

Quality and efficiency of basic research in molecular biology: a bibliometric analysis of thirteen excellent research institutes

Heinrich Herbertz, Benno Müller-Hill

Institut für Genetik, Universität zu Köln, Weyertal 121, D-50931 Köln, Germany

Final version received September 1994

Abstract

We try to assess, in a systematic and objective manner, the research performance of 13 research institutes active in the field of molecular biology. For this purpose we have counted the number of scientific publications and the number of citations received during a five-year period (1980–1984). We use citations per publication as an indicator of quality and costs per citation as an indicator of efficiency of research. Peer review seems to discourage uninteresting, i.e. not cited, research. Grant systems seem to work more efficiently than funding on a permanent institutional basis.

1. Introduction

Many countries suffer these days from economic crises. The necessity to economize on public spending leads to questions about the efficiency and quality of national basic research. Basic research in science is assumed to be of potential economic use. Thus, allocation of financial support becomes the major problem of science policy. The question is raised of how to fund scientific basic research optimally, in order to preserve the potential economic advantage. In this situation, objective bibliometric indicators (for example numbers of research publications, numbers of citations, numbers of patents) gain in potential influence in science policy decisions. They are used to assess strength and weakness of scientific research in a national or international comparison (i.e. Nederhof and Noyons, 1992; Martin and Irvine, 1983; Moed et al., 1985; Healey et al., 1988). In 1988 for example, UNIPS (Unité d'indicateurs de Politique Scientifique) was established to provide quantitative information to support scientific policy decisions of the French CNRS (Bauin and Rothman, 1991). In particular, biochemistry and molecular biology, scientific fields regarded as important for future economic development, have been the subject of a number of bibliometric studies (Anonymous, 1992; Anderson, 1992; Pendlebury, 1990).

An example of a rather dubious rating can be found in a worldwide ranking of institutes doing basic research in molecular biology that was published in the journal *Science* (Anderson, 1992). Well-known institutes, like the Pasteur Institute in Paris or Stanford University, were missing, while the results of some other institutes were surprising. One of the reasons for these questionable results was the unsolved technical problem of how to identify molecular biology papers in multidisciplinary journals such as *Nature, Science* or *Proceedings of the National Academy of Sci-*

SSDI 0048-7333(94)00814-0

ences USA (*PNAS*). Therefore, these journals – with an eminent influence in the field – were excluded. The argument that this limitation would affect all institutes in a similar manner is erroneous in the case of the 13 institutes we studied. We found large differences in the proportion of the institutes' output published in *Nature* or *PNAS* – it varies between 2% and 24%. One may ask whether the *Science* articles would have been published with the advice of an active molecular biologist (Herbertz and Müller-Hill, 1992; Pendlebury and Anderson, 1992; Taubes, 1993).

Our study was designed as a collaborative effort between a student in sociology and a professor in molecular biology to evaluate the limitations of bibliometric methods used to assess research activities in molecular biology. We wanted to combine the qualitative experience-based evaluation by the expert and the objective quantitative evaluation by bibliometric methods. The obvious advantage of such a procedure is that the expert's knowledge of the field directs the search of bibliometric data. In some cases, however, it is not possible to confirm the expert's view by quantitative means. Here the evidence arises exclusively from experience.

We try first to describe an evaluation based on citations in detail. Then we will ask the question whether citation data can help to analyse the efficiency of different modes of funding and forms of organization of scientific research. We will use citations per paper as an indicator of quality, and cost per citation as an indicator of efficiency of research. We finally ask the question whether research funded mainly by grants is more efficient than research funded mainly by large and permanent institutional support (Müller-Hill, 1991).

2. Methodological remarks

2.1. The institutes

We examined 13 research institutes active in the field of molecular biology. The institutes were chosen by the expert (BMH) in consultation with other experts' views about the comparability of these institutes. The first criterion was similarity of the field of research. The second criterion was the excellence of the institute. The third criterion was variety in organizational and institutional structure. Two different research funding systems exist in Germany. On the one hand, there are several research centres funded mainly on a permanent institutional basis. On the other hand, there are universities funded mainly by grants. We chose three German research centres and four German university institutes which were known to be research intensive at the beginning of 1980s. To gain an impression of the influence of German molecular biology research in an international comparison, we chose six well-known successful international research institutes, including the European Molecular Biology Laboratory (EMBL) in Heidelberg.

The list of the institutes and the abbreviations we use is as follows:

BIOB: Biozentrum, University of Basel, Switzerland
CSHL: Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, USA
EMBL: EMBL Heidelberg (without outstations in Grenoble and Hamburg), Germany
GBFB: Gesellschaft für biotechnologische Forschung, Braunschweig, Germany
IBUF: Institute of Biology III, University of Freiburg, Freiburg, Germany
IGUC: Institute of Genetics, University of Cologne, Cologne, Germany
IVUG: Institute of Virology (Vet. Med.), University of Giessen, Giessen, Germany
MRCC: Laboratory of Molecular Biology of the MRC, Cambridge, England
MPIB: Max Planck Institute of Biochemistry, Martinsried/Munich, Germany
MPIM: Max Planck Institute of Molecular Genetics, Berlin, Germany
PAPA: Department de Biologie Moleculaire, Institut Pasteur, Paris, France
SFBK: Sonderforschungsbereich 138/156, University of Konstanz, Konstanz, Germany
YALE: Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, USA

2.2. *Quality of research and citation*

We use 'quality' of research in a social sense (Cole and Cole, 1973). We define quality as the usefulness of published research for other scientists as reflected in citations (quality = average number of citations per paper). We concentrate on published research since citation counts can only be applied to that. In this sense a paper that is cited frequently is assumed to be of more use than one that is not cited. Citations are not necessarily a measure of the intrinsic quality of work, but they are an adequate measure of the socially defined utility i.e. quality of work. There are individual cases where this assumption does not hold, for example a fundamental paper may be written in a language other than English, or may be published in a low status journal. Almost all papers we analysed were written in English. We think that the remaining cases are rare. They should not affect the analysis of large data sets such as the publication output of the institutes we studied. Many biochemists believe that methodological papers, like Lowry's, totally destroy the validity of citation counts (Lowry's paper describes a generally used method of protein determination). We will demonstrate later that this is not a major problem. Another objection against citation analysis is that wrong results may be cited excessively just because they are wrong. In biochemical basic research, the experience of the senior author tells us that this is not so. Wrong papers are not cited, they are ignored (Anderson, 1991).

This can be illustrated by two examples. In 1987 two papers were published which dealt with Alzheimer's disease (Kang et al., 1987, and Delabar et al., 1987). The first paper described the gene for the amyloid precursor protein. The second paper claimed Alzheimer's disease was caused by a duplication of the amyloid gene. The second paper, carrying the name of the Nobel prizewinner Carlton Gajdusek as senior author, was found to be wrong in the same year (Podlisny et al., 1987; Tanzi et al., 1987) or one year later (Murdoch et al., 1988). Fig. 1 shows that the second paper was cited half as often directly after publication (1987, 18 times; 1988, 50 times) as the

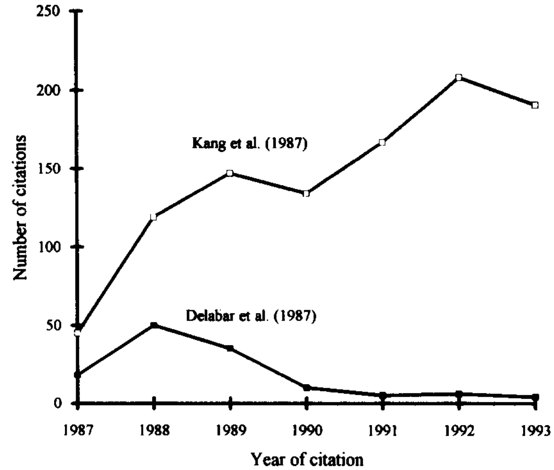

Fig. 1. Yearly distribution of citations of two articles dealing with Alzheimer's disease.

first paper. But then the number of its citations decreased to about five a year while the first paper's citation number increased to about 200 a year.

The second example can be found in the list of the 50 most cited articles in our study (McKay and Steitz, 1981). In this paper, the authors present the structure of a particular protein and a model of its DNA complex. The structure was right, but the model of the complex turned out to be wrong. The paper was cited about 200 times up to 1985. We analysed a sample of 20 citations. All 20 papers cite the correct structure of the protein, but do not even mention the wrong model.

Citation rates can only be meaningful when used for comparative purposes. To be valid a comparison must take several facts into account:

- Different scientific research fields follow their own citation practices. There are difficulties in comparing scientists working in different fields, for example in mathematics and molecular biology.
- Different kinds of publications vary in their probability of being cited. We will show that in molecular biology, methodological papers and reviews tend to be cited more frequently than papers dealing with structure or function. Abstracts and articles in books are rarely cited. Further-

more, there are publications with an educational content addressed to others than the scientists working in the same field. There might also be national differences in citing practice: there seems to be a tendency to cite papers of scientists working in the same country.

– To interpret the numbers of citations, one has to keep in mind that they do not have a strictly linear character. It does not make sense to regard a paper that received two citations to be twice as valuable as a paper that received only one citation.

– The citation–distribution function is a very skewed one. The zero- and very few-citations papers are by far the largest categories. Therefore concepts like 'averages' are not unproblematic.

– A bibliometric evaluation must always be accompanied by commentaries of experienced scientists in the field to prevent misleading interpretations of the data.

2.3. The institutes and the period studied

The institutes were chosen by the expert after discussions with other experts. A possible way of verifying the choice by bibliometric means is to analyse the journals in which the institutes pub-

Table 1
The twenty journals in which most articles were published

Articles	MPIM	IGUC	YALE	EMBL	MRCC	CSHL	BIOB	PAPA	IVUG	IBUF	MPIB	SFBK	GBFB
	449	193	417	421	505	282	501	345	102	101	874	281	393
Percentage of articles in twenty journals [a] *published by a particular institute*													
PNAS	4.7	8.8	14.3	8.5	4.7	12.0	4.8	6.4	1.0	2.9	1.8	1.1	0.8
J. Mol. Biol.	1.1	1.0	8.1	9.7	13.0	4.2	5.7	4.6	–	–	2.5	0.4	0.2
EMBO	2.7	16.0	1.4	12.8	7.3	0.3	7.3	6.1	3.8	3.8	4.2	0.4	1.5
Nucl. Ac. Res.	12.0	5.7	7.2	5.7	8.9	4.6	1.2	1.7	–	11.5	0.9	0.4	3.6
J. Biol. Chem.	9.8	–	17.9	0.7	1.0	4.9	5.9	3.5	2.9	1.0	1.8	0.7	0.8
Eur. J. Biochem.	6.7	6.7	–	3.1	3.0	–	5.7	5.2	4.8	1.0	5.2	4.6	4.8
Nature	2.2	8.2	3.1	3.5	13.8	12.0	2.6	2.6	–	3.8	1.1	0.7	1.0
Febs Lettr.	9.8	1.5	1.0	4.7	3.3	0.3	3.0	4.9	1.0	7.7	3.9	4.6	3.3
Cell	0.4	4.1	7.9	4.0	5.1	13.1	2.4	2.9	–	1.9	0.4	–	–
Mol. Gen. Gen.	13.4	7.7	2.1	3.1	0.4	0.3	3.0	1.4	–	4.8	1.4	3.2	0.8
Biochemistry	3.8	1.5	4.8	0.9	2.8	1.3	6.5	2.9	–	–	0.7	0.7	1.8
Biol. Chem. H.-S.	0.9	2.1	–	1.4	0.2	–	0.8	–	1.9	–	12.5	–	0.5
BBA	0.9	0.5	1.0	1.4	0.6	–	4.2	1.4	1.9	1.0	1.1	11.4	1.3
J. Cell Biol.	–	–	1.7	7.6	2.0	1.4	1.2	1.2	–	–	1.1	1.4	–
Virology	1.1	2.1	0.2	1.4	0.2	4.6	0.8	0.6	31.7	–	0.1	0.7	0.2
J. Bacteriol.	5.3	0.5	2.4	0.5	0.4	–	1.0	3.8	–	1.0	0.1	5.3	0.2
J. Virol.	3.3	6.2	0.5	0.9	–	6.4	1.4	0.9	1.9	1.0	–	1.1	–
Z. Naturfrschg	–	–	–	0.2	–	–	0.2	–	–	–	3.1	1.1	7.9
Developm. Biol.	–	–	–	1.9	3.9	0.3	–	2.9	–	–	0.8	0.4	–
Gene	1.3	2.1	–	0.2	1.2	3.9	1.2	1.4	–	1.9	0.1	–	0.2
Sum of articles (%) [b]	79.4	74.7	73.6	72.2	71.8	68.3	58.9	54.4	50.9	43.3	42.8	38.2	28.9
Sum of citations (%) [c]	85.7	85.9	85.8	83.7	91.3	86.9	80.5	76.2	77.7	65.9	66.9	48.3	54.7
Percentage of articles in thirty journals [d] *published by a particular institute*													
Sum of articles (%)	80.7	77.3	74.4	73.6	72.8	69.7	60.5	56.1	51.9	48.2	43.7	40.7	30.4
Sum of citations (%)	86.9	91.4	86.3	87.7	93.2	91.1	85.3	80.5	82.5	77.5	72.4	53.2	61.1

[a] Journals are ranked in descending order according to the total number of articles from all institutes analysed in this study. Reviews and abstracts are excluded.
[b] Sum of articles is the percentage of all articles of an institute.
[c] Sum of citations is the percentage of all citations received by articles in journals.
[d] Ten additional journals are included in the following order: *Exp. Cell Res.*, *BBRC*, *Cold Spring Harbor Symp. Quant. Biol.* (treated here as a journal), *Biochem. J.*, *Anal. Biochem.*, *Eur. J. Immunol.*, *Liebigs Ann. Chem.*, *J. Gen. Virol.*, *Eur. J. Cell Biol.*, and *Planta*.

lished their results. For this we inspected the 30 journals which covered most of the papers of all 13 institutes (Table 1) to see whether these journals also cover the main publication output of each single institute. The proportion of each institute's total publication output in these 30 journals varies from 30% in the case of the GBFB, to 80% in the case of the MPIM. Depending on the size and specialization of an institute, there are differences in the variety of subfields of molecular biology in which the various groups work. Furthermore, there are variations in the choice of journals based on editorial work or on national preferences. For example, the two German journals, *Zeitschrift für Naturforschung, Sektion C* and *Hoppe-Seyler's Zeitschrift für Physiologische Chemie* (since 1985, *Biological Chemistry Hoppe-Seyler*) were used almost exclusively by scientists working in German institutes.

The results for one of the institutes (GBFB) and one of the research units (SFBK) indicate that they are somewhat special. The GBFB was founded in 1965. The institute's previous name, Gesellschaft für molekularbiologische Forschung (society for research in molecular biology) indicates its purpose of basic research in molecular biology. In 1976 it was reorganized and renamed as a national research center. Part of this reorganization was the effort to integrate applied structural chemical research with basic research in molecular biology. This is the reason why the institute published many papers in journals like *Angewandte Chemie* or *Justus Liebigs Annalen der Chemie*. The SFBK often published in journals specializing in the field of membrane biology such as the *Journal of Membrane Biology* or *Journal of Histochemistry and Cytochemistry*. The MPIB published only 43% of its total publica-

Table 2
Ranking of 13 research institutes according to the citations (quality) of their publications

Institute	Citations per paper				Journal impact factors per paper			Ranking order, Citations per paper				Ranking order, Journal impact factors per paper		
	A	B	C	D	E	F	G	A	B	C	D	E	F	G
CSHL	40.4	40.0	46.2	41.6	7.6	6.9	7.8	1	1	1	1	1	1	1
MRCC	35.1	35.4	38.0	33.3	6.6	5.9	6.5	2	2	2	2	2	3	3
EMBL	26.0	25.3	30.4	25.2	5.3	4.7	5.5	3	5	3	4	5	5	5
YALE	23.9	27.7	29.3	25.3	6.5	6.3	6.6	4	3	4	3	3	2	2
IGUC	19.9	26.2	27.1	21.2	6.2	5.7	5.9	5	4	5	5	4	4	4
BIOB	19.0	20.7	23.6	19.3	4.4	4.3	4.8	6	6	6	6	6	6	7
PAPA	17.9	18.2	23.2	18.6	4.3	4.2	4.8	7	7	7	7	7	7	6
IBUF	14.2	14.7	18.6	16.0	3.8	3.6	3.9	8	8	8	8	9	9	9
MPIB	11.0	12.7	16.9	12.7	3.0	2.9	3.3	9	9	9	9	10	10	10
MPIM	10.2	12.3	15.1	10.5	4.3	4.2	4.4	10	10	10	10	8	8	8
IVUG	9.6	11.4	13.0	9.6	2.8	2.9	3.2	11	12	12	12	11	11	11
SFBK	9.2	12.2	14.3	10.2	2.8	2.8	3.1	12	11	11	11	12	12	12
GBFB	8.6	9.8	12.0	9.2	2.5	2.4	2.3	13	13	13	13	13	13	13

A: Our detailed measurement. Citations and articles are weighed in cases of collaboration. Only articles with at least one citation from outside the institute are counted. Abstracts and reviews are excluded. Self-citations are subtracted.
B: All articles mentioned in the publication lists are counted. Self-citations are not subtracted.
C: The same as A but without weighing in cases of collaboration. Self-citations are not subtracted.
D: The same as C, but self-citations are subtracted.
E: Journal impact factor (JIF) per paper of all articles published in 1980–1984 in journals found to be covered in the *Journal Citation Report* 1980–1984, published by the Institute for Scientific Information. Abstracts are excluded.
F: The same as E, but a corrected JIF of the journal *Nature* was used: 14.5 instead of 8.3. Because of the multidisciplinary character of *Nature*, the JIF combines low and high citation rates of various scientific fields and therefore underestimates the journal's impact in a scientific field with high citation rates, like molecular biology. We therefore calculate a revised JIF on the basis of our citation findings.
G: JIF per paper of all articles published during 1980–1988 in journals found to be covered in the *Journal Citation Report* 1980–1988. Abstracts are excluded.

tions in the 30 journals. But 72% of all citations obtained by papers of this institute were citations to papers in these 30 journals. This indicates that the more relevant results were published in these journals. So we conclude that 12 of the 13 institutes worked indeed predominantly in the fields of biochemistry and molecular biology, and that only one institute (GBFB) is so different that it should not be *directly* compared.

We studied the publications of 1980–1984. When we started our study in 1990 this period seemed to be adequate. Is the five-year period too short? Five years are a short period in the life of an institute. To prolong the period we also used the average journal impact factor (JIF) for the nine-year period of 1980 to 1988. We found similar ranking results when we used the average citation per paper and average JIF (Table 2). The similar ranking order of the nine-year period indicates that our results seem to be valid over an extended period.

For each publication we counted the citations during a five-year period, including the year of publication. Is this period too short? To decide this question we examined the citations of a random sample of 459 articles in journals published in 1980 over a nine-year period (Fig. 2). The average article in journals had its peak of citations two years after publication. We interpret this result as a hint that a five-year period should be adequate in the field of molecular biology.

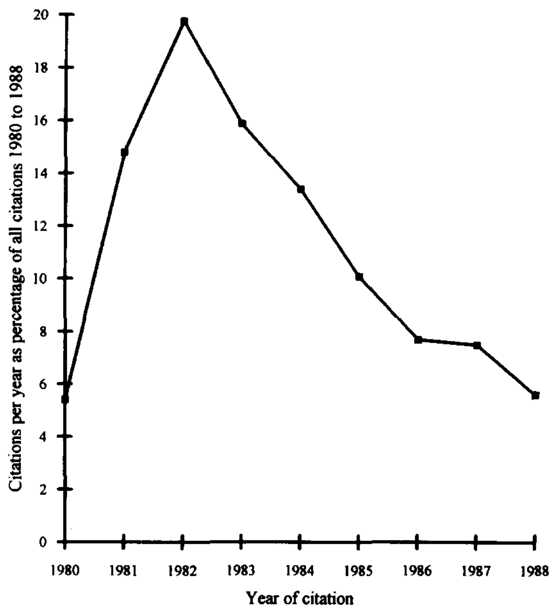

Fig. 2. Yearly distribution of citations of 459 articles published in 1980 in journals (reviews, abstracts, and obituaries are excluded).

2.4. Analysis of the data

We measure the quality of published research in terms of the average number of citations per paper. We used the publication lists (1980–1988) provided by the institutes or published in yearbooks. We assumed that the lists provided by the institutes were correct. We were not sure about the lists compiled in the yearbooks. In the cases of the MPIB and the IGUC the lists in the yearbooks were compared with the reprint collections of the directors of the various departments. They proved to be reliable and accurate. It is indeed in the self-interest of every researcher to have a maximum of papers in his list. We used the printed version of the *Science Citation Index* (SCI; cumulated volume 1980–1984 and the volumes for 1985, 1986, 1987, 1988) of the Institute for Scientific Information. The *SCI* lists under the name of the first author all publications that are cited at least once. It names the first author and the bibliographical reference of the citing publication. We did a manual count, which was certainly more time-consuming than an electronic search. On the other hand, various sources of error such as misspelt names, wrong order of the authors, or wrong or incomplete bibliographical references could be more easily corrected. Collaborations were identified according to the institutional addresses of authors given in their papers.

We tried to exclude all self-citations by using the publication lists of the various institutes. However, we were not able to correct for self-citations by collaborating groups from outside the respective institute. The publication lists provided by the institutes differ. In some cases, for instance, abstracts were included, in other cases there were not. To deal with this problem we only counted those articles with at least one citation from outside the respective group in order to exclude articles that are usually not cited, like most abstracts, obituaries or articles with an educational rather than a research content. This procedure favours some institutes (Fig. 3). Both Max Planck Institutes and the IVUG produce a large proportion, 15% and more, of uncited articles in journals (abstracts are excluded). This percentage is five or six times higher than the respective percentage of the MRCC (2%) or CSHL (3%).

We calculate citations of all articles including reviews and excluding reviews. Articles in *Methods in Enzymology* are treated as reviews. In the case of collaboration, we fractionate citations and publication values. When two groups collaborate the article is counted as 0.5, when three groups collaborate as 0.33 and so on. The citations are weighed according to the group of the first author. In molecular biology there is a rule of giving the first author's position to the scientist who did the main work, and the last author's position to the leader of the research group. When two groups collaborate the group presenting the first author gets two thirds of the citations, the other group one third. When three groups collaborate the factor is 0.5 and 0.25, respectively, and so on. This ratio is somewhat arbitrary but it seemed adequate to us. Thus we weigh the results of different forms of collaboration.

Articles in journals without any documented collaboration were cited 19.4 times on average, articles with at least two groups collaborating were cited 19.6 times on average (without weighing for collaboration). Finally, we compared the effects of the various calculations on the ranking of groups and institutes. The ranking order of the groups differed to some extent (data not shown), but the ranking order of the institutes was almost identical (Table 2).

3. Results

3.1. Publications

We counted 7102 publications in total. There were: 5291 articles in journals; 1256 articles in books; 387 reviews; 99 abstracts; 69 others. Articles in books were rarely cited even by their own authors. In the case of articles in journals we found 4.3 self-citations on average but only 1.4 self-citations of articles in books. Nearly half (45%) of the articles in books received no citations at all, 15% were cited once by other researchers. Just 10% received 10 or more citations from others within a five-year period.

On average the different kinds of publications were cited as follows: articles in books, 4 times; articles in journals, 20 times; reviews, 25 times. Within a five-year period, 9% of the articles in journals received no citation from outside the respective group, a proportion much smaller than

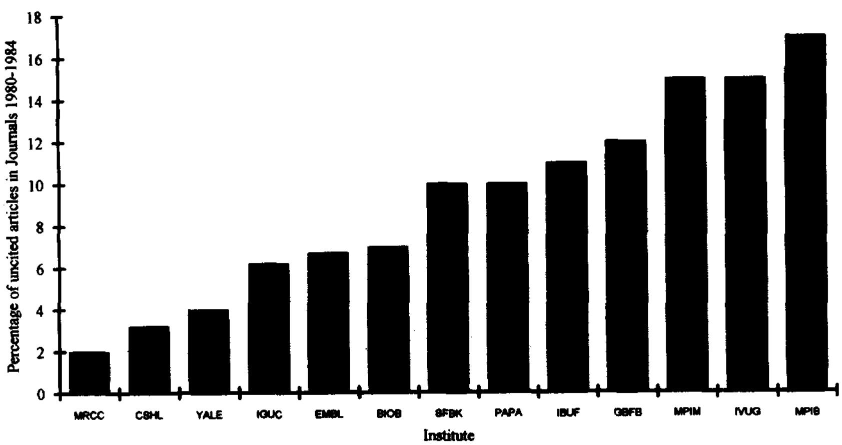

Fig. 3. Percentage of articles in journals which are not cited from outside the respective institute.

Table 3
Comparison of 13 research institutes active in basic research in the field of molecular biology

Institute [a]	CSHL	MRCC	EMBL	YALE	IGUC	BIOB	PAPA	IBUF	MPIB	MPIM	IVUG	SFBK	GBFB
Budget 1982 [b] (10^6 DM)	–	–	24.7	14.6	4.9	20.3 [f]	9.8	3.4	32.0	12.5	1.6	> 2.7 [h]	29.6
Teaching costs deducted [d]	–	–	24.7	6.6	2.2	9.1	9.8	1.5	32.0	12.5	0.7	> 1.2	11.8
Grants 1982 (10^6 DM)	29.4	–	0.5	11.3	4.4	3.5	1.1	0.9	2.9 [g]	1.3 [g]	1.3	3.2	? [i]
Total budget [c]	29.4	– [e]	25.2	25.9	9.3	23.8	10.9	4.3	34.9	13.8	2.9	> 5.9	> 29.6
Teaching costs deducted	29.4	–	25.2	17.9	6.6	12.6	10.9	2.4	34.9	13.8	2.0	> 4.4	> 11.8
Laboratory space (10^3 m^2)	–	8.2	6.0	13.7	2.9	7.7	3.0	1.3	20.4	5.4	0.8	–	19.0
Doctoral dissertations 80–84 [j]	–	29	7	30	33	60	–	17	72	43	12	6	31
Articles [k]	265.6	458.2	358.7	379.9	164.0	417.6	253.7	87.7	600.3	359.7	87.3	232.0	293.8
Reviews included	289.6	510.0	375.2	398.4	172.0	433.5	275.6	92.7	625.7	371.0	97.3	259.3	304.8
Cost/article [c]: (10^3 DM)	110.7	–	70.2	68.2	56.7	57.0	43.0	49.0	58.1	38.4	33.2	> 25.4	> 100.7
Teaching costs deducted	110.7	–	70.2	47.1	40.2	30.2	43.0	27.4	58.1	38.4	22.9	> 19.0	> 40.2
Reviews included [c]	101.5	–	64.2	65.0	54.1	54.9	39.6	46.4	55.8	37.2	29.8	> 22.7	> 97.1
Teaching costs deducted	101.5	–	64.2	44.9	38.4	29.1	39.6	25.9	55.8	37.2	20.5	> 18.1	> 38.7
Citations [l]	10733.0	16082.3	9333.4	9067.9	3260.6	7922.3	4539.7	1246.2	6619.0	3679.6	841.5	2128.7	2524.5
Reviews included	11641.1	17794.1	9648.8	10835.6	3990.6	8312.0	5307.0	1311.2	7009.5	3906.9	1167.5	2432.2	2802.2
Citations per article	40.4	35.1	26.0	23.9	19.9	19.0	17.9	14.2	11.0	10.2	9.6	9.2	8.6
Reviews included	40.2	34.9	25.7	27.2	23.2	19.2	19.3	14.1	11.2	10.5	12.0	9.4	9.2
Cost/citation [c]: (10^3 DM)	2.7	–	2.7	2.9	2.9	3.0	2.4	3.4	5.2	3.7	3.4	> 2.8	> 11.7
Teaching costs deducted	2.7	–	2.7	2.0	2.0	1.6	2.4	1.9	5.2	3.7	2.4	> 2.1	> 4.7
Reviews included [c]	2.5	–	2.6	2.4	2.3	3.0	2.1	3.3	5.0	3.5	2.5	> 2.4	> 10.6
Teaching costs deducted	2.5	–	2.6	1.6	1.6	1.5	2.1	1.8	5.0	3.5	1.7	> 1.8	> 4.2

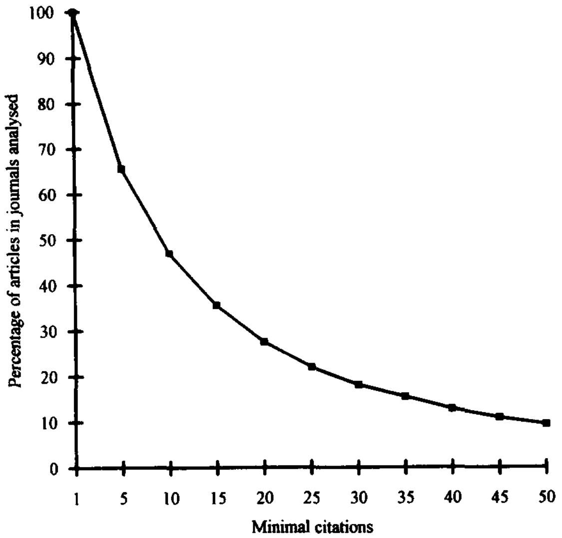

Fig. 4. Percentage of articles in journals with increasing minimal citations.

Hamilton's 19% of articles in journals in biochemistry and molecular biology that were published in 1984, and were not cited until 1988 (Hamilton, 1991). There are two possible reasons for this difference: since it is not clear what kinds of publications were counted by Hamilton, the difference may result from counting rarely cited abstracts and letters to the editors, or alternatively our sample of institutes may produce higher quality publications on average.

No more than one citation from outside was received by 16% of the articles in journals. The articles with no more than five citations collected only 4% of all citations. But those 9% of articles in journals receiving 50 citations or more collected nearly half (48%) of all citations (Fig. 4; only articles which received at least one citation were included; self-citations are subtracted and abstracts and reviews are excluded). A substantial number of articles collected no citations at all

Notes to Table 3:

[a] Data of CSHL according to Cold Spring Harbor Laboratory annual report 1983. Data of MRCC provided by A. Klug. Data of EMBL provided by L. Philipson; the articles of the department of Instrumentation and of the outstations in Grenoble and Hamburg were not included. Data of YALE provided by D.M. Engelman. Data of IGUC provided by the Dean's and Chancellor's office. Data of BIOB provided by W. Gehring, grants according to "Jahresbericht Schweizerischer Nationalfonds zur Förderung der wissenschaftlichen Forschung". Data of PAPA provided by M.M. Schwartz in form of the final budget in DM. Data of IBUF provided by R. Hertel. Data of MPIB and MPIM provided by the Generalverwaltung of the MPG; number of doctoral dissertations according to Jahrbuch der MPG. Data of IVUG provided by R. Rott. Data of SFBK according to "Sonderforschungsbereich 138, Abschlußbericht 1972–1983". Data of GBFB provided by J. Collins and management of the GBF.

[b] Yearly budgets are the budgets of 1982. They all include the salaries of all employees, research and library costs, electricity, heat, maintenance and so on unless otherwise stated. The sum does not include stipends, central administration and other central costs. The conversion value of 1982 used was 1 DM = 2.42 $ = 1.19 sFr (information provided by the Devisenabteilung of the Deutsche Bank).

[c] A: The budgets include teaching costs of the universities.

[d] B: Teaching costs are estimated to be 55% of the permanent institutional budget and are subtracted. In the case of GBFB, 60% of the permanent institutional budget is estimated to be costs for other purposes than basic research and is subtracted.

[e] MRCC records were not available.

[f] Costs of electricity and heat were estimated on the base of the Cologne institute per m^2 laboratory space.

[g] We do not know the grants given in Normalverfahren by Deutsche Forschungs Gemeinschaft.

[h] We cannot estimate the costs of electricity and heat because we have no information about the SFBK laboratory size.

[i] Number and size of grants unknown.

[j] Those finished between January 1980 and December 1984.

[k] Only articles that were cited at least once by other researchers are counted. Abstracts are not counted. The articles are weighed in the following manner: when two groups collaborated they were counted as 0.5; when three groups collaborated as 0.33; and so on.

[l] The *Citation Index* for 1980–1984 and 1985, 1986, 1987 and 1988 was consulted. Misspellings and umlauts are considered. Citations are collected for the year of appearance and for four further years. Self-citations are not counted. In the case of collaboration the citations are weighed in the following manner: when two groups collaborated, the group providing the first author gets two thirds, the other group one third of the number of citations. When three groups collaborated, the group providing the first author gets half, the other groups one quarter of the number of citations, and so on.

within five years. Does that mean that the uncited articles are of no use, and were only written for the record as argued by Hamilton (1990)? We think that this is true in many cases, although sometimes this may be misleading.

3.2. The institutes

Table 3 presents the results of the 13 institutes. The CSHL, then headed by Nobel prizewinner J.D. Watson, and the MRCC, the working place of the Nobel prizewinners M. Perutz and A. Klug, are most highly ranked. They are followed by EMBL, whose results might be even better today if one takes account of the trend of increasing citation counts between 1980 and 1988. Three university institutes, the YALE, the IGUC and the BIOB, produced somewhat lower results.

The two Max Planck Institutes in Berlin and Munich are ranked in the middle of the list. One has to be aware that the overall citation results of institutes give no information about the homogeneity or heterogeneity of the quality distribution inside the institute. For this we look at the distribution of the various research groups (Table 4). Excellent groups worked in both Max Planck Institutes and in the BIOB as well as groups with results of lower quality. We just mention the later Nobel prizewinners R. Huber, J. Deisenhofer and H. Michel at Munich and the group headed by W. Gehring in Basel as truly outstanding. The IBUF profited from the excellent results of only one group.

There are some reasons for doubting the comparability of the GBFB (see above). Its results should therefore be interpreted with particular care. We note that the Department of Cell Biology of the GBFB had good results, in contrast to the four other departments (Table 4).

3.3. The research groups

To identify research groups, we used information contained in the yearbooks and publication lists of the institutes. In two cases we were not able to identify research groups (MRCC and IVUG) since the publication lists were alphabetical without any differentiation for research groups or departments. The result of this procedure can only be an approximation of reality. There might be, for example, independent research groups in departments we cannot identify. This, and differences in the organization of the various institutes, limit the comparability of the results. The research groups vary considerably in their size and accordingly in their number of publications (Table 5). They also vary in continuity and permanent institutional support. Since all the institutes we studied are well known, we had expected the distribution of the quality (i.e. citations per paper) of the research groups to be similar to a normal distribution: a tendency for most groups to produce good results in the middle, with a few

Table 4
Groups in institutes ordered according to citations per (quality of) articles

Citations per paper	Number of Research Groups										
	GBFB	SFBK	MPIB	MPIM	IBUF	PAPA	BIOB	IGUC	YALE	EMBL	CSHL
> 32					1		1		1	2	3
> 28–32						1	1		1		
> 24–28				1					3		1
> 20–24						3		3	3		
> 16–20	1		2	1		3		1	5		1
> 12–16		1		1	1	1	2	1	1	1	
> 8–12		6	1	4	3	1			1		
> 4–8	4	2	1	7		2	2				

Only articles with at least one citation from outside the respective institute are counted. Reviews and abstracts are excluded. Citations are weighed in the case of collaboration. Self-citations are subtracted. We were unable to identify research groups of the IVUG and of the MRCC.

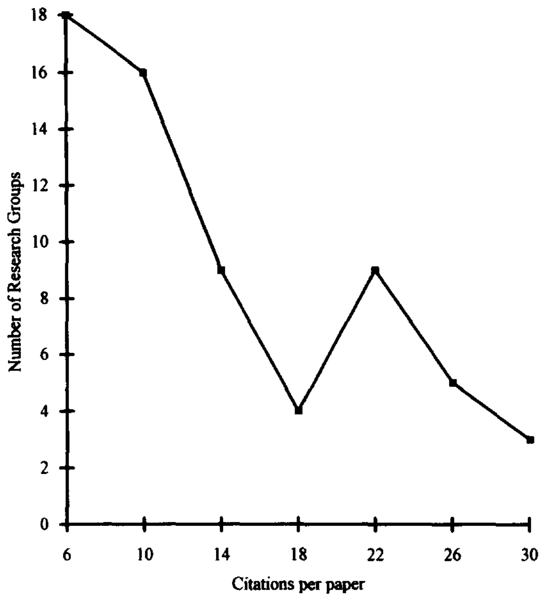

Fig. 5. Distribution of the research groups, aggregated in various classes of quality according to the average citation per article.

highly productive groups at one extreme and a few unproductive groups at the other extreme. Fig. 5 shows, however, that most groups are of low quality according to our definition.

3.4. *The 50 most cited articles*

We think that the number of citations of an article is in general a function of its quality, i.e. of its usefulness to other scientists. Therefore, methodological articles might have a better chance of being cited. The indicator, citation per article, can be biased by articles gaining an extremely high number of citations. We analysed the 50 most cited papers in order to examine whether methodological papers cause such a bias in ranking (Table 6). The senior author divided the papers into five categories according their titles and abstracts. The five categories were: method (M), structure (S), structure–function (S,F), function (F), and review (R). The 50 most cited articles belong to all five categories. They were published by nine of the thirteen institutes. The first two articles of the list are methodological. The transformation record of Hanahan was frequently used, as was the method for determining DNA sequences described by Biggin. We excluded the manual of Maniatis, Fritsch and Sambrook, *Molecular cloning: A laboratory manual* which was published by CSHL in 1982. It is a compendium of methods the authors in part had

Table 5
Ranking of 30 research groups according to citations per (quality of) articles

Institute	Research groups	Number of papers	Citations per paper
CSHL	Wigler, Fiddes	39	81.7
BIOB	Gehring, De Robertis, Noll	44	54.8
YALE	Steitz, J.A.	39	54.3
EMBL	Edström, Graf	86	41.4
CSHL	Mathews, Sambrook, Roberts	146	39.7
IBUF	Kössel	13	38.3
CSHL	Garrels, Feramisco, Helfman	52	38.2
EMBL	Simons	157	32.1
PAPA	Yaniv	37	30.7
YALE	Steitz, T.A.	33	30.2
BIOB	Burger, Schatz	77	28.6
MPIB	Huber, Deisenhofer	41	27.8
YALE	Grindley	14	27.8
YALE	Armitage	24	26.7
CSHL	Harshey, Hicks, Klar	53	26.3
YALE	Platt	28	24.9
IGUC	Starlinger	31	23.8
YALE	Shulman	27	23.8
IGUC	Müller-Hill, Beyreuther	57	23.4
PAPA	Buckingham	19	22.8
YALE	Söll	46	21.7
YALE	Lengyel	21	21.0
PAPA	Changeux	58	20.3
IGUC	Rajewsky	51	20.3
PAPA	Jacob	33	20.2
MPIM	Schuster	37	20.0
IGUC	Vielmetter, Sippel	14	19.9
PAPA	Buc	23	19.8
YALE	Macnab	14	19.7
MPIB	Kühn, Timpl, van der Mark	200	19.3

Only groups are included that published during at least four years between 1980 and 1984 in one of the institutes. Groups of IVUG and MRCC are not included. Only articles with at least one citation from outside the institute are counted. Reviews and abstracts are excluded. Self-citations are subtracted. Citations are weighed as described in Section 2.4.

Table 6
The 50 most cited articles in our study

	Institute	Citations	Content
Hanahan, D., Studies on transformation of Escherichia coli with plasmids, *J. Mol. Biol.* 166, 557 (1983)	CSHL	929	M
Biggin, M.D. et al, Buffer gradient gels and ^{35}Slabel as an aid to rapid DNA sequence determination, *PNAS* 80, 3963 (1983)	MRCC	911	M
Lerner, M.R. et al, Are snRNPs involved in splicing? *Nature* 283, 220 (1980)	YALE	552	F,R
Mount, S.M., A catalogue of splice junction sequence *Nucl. Acids Res.* 10, 459 (1982)	YALE	520	R
Anderson, S. et al, Sequence and organization of the human mitochondrial genome, *Nature* 290, 457 (1981)	MRCC	488	S
Frischauf, A.M. et al, Lambda replacement vectors carrying polylinker sequences, *J. Mol. Biol.* 170, 827 (1983)	EMBL	457	M
Müller, R. et al, Induction of c-fos gene and protein by growth factors precedes activation of c-myc, *Nature* 312, 716 (1984)	EMBL	449	F
Ruley, H., Adenovirus early region 1 A enables viral and cellular transforming genes to transform primary cells in culture *Nature* 304, 602 (1983)	CSHL	448	F
Sanger, F. et al, Cloning in singlestranded bacteriophage as an aid to rapid DNA sequencing, *J. Mol. Biol.* 143, 161 (1980)	MRCC	441	M
Deisenhofer, J. et al, X-ray structure analysis of a membran protein complex. Electron density map at 3A resolution and a model of the photosynthetic reaction center from Rhodopseudomonas viridis *J. Mol. Biol.* 180, 385 (1984)	MPIB	350	S
Sanger, F. et al, Nucleotide sequence of bacteriophage • DNA *J. Mol. Biol.* 162, 729 (1982)	MRCC	341	S
Zoller, M.J. et al, Oligonucleotide-directed mutagenesis of DNA fragments cloned into M13 vectors, *Methods Enzymol.* 100, 468 (1983)	CSHL	336	M
Dente, L. et al, pEMBL: a new family of single stranded plasmids, *Nucl. Acids Res.* 11, 1645 (1983)	EMBL	318	M
Gluzman, Y., SV-40-transformed simian cells support the replication of early SV-40 mutants, *Cell* 23, 175 (1981)	CSHL	308	F
Doerfler, W., DNA methylation and gene activity *Annu. Rev. Biochem.* 52, 93 (1983)	IGUC	297	F,R
Taparowsky, E. et al, Structure and activation of the human N-ras gene, *Cell* 34, 581 (1983)	CSHL	269	F
Anderson, S. et al, Complete sequence of bovine mitochondrial DNA. Conserved features of the mammalian mitochondrial genome *J. Mol. Biol.* 156, 683 (1982)	MRCC	267	S
Pelham, H.R.B., A regulatory upstream promoter element in the Drosophila HSP 70 heat-shock gene, *Cell* 30, 517 (1982)	MRCC	263	S,F
Baer, R. et al, DNA sequence and expression of the B95-8 Epstein-Barr virus genome, *Nature* 310, 207 (1984)	MRCC	262	S
Taparowsky, E. et al, Activation of the T24 bladder carcinoma transforming gene is linked to a single amino acid change *Nature* 300, 762 (1982)	CSHL	260	S,F
Zoller, M.J. et al, Oligonucleotide-directed mutagenesis: A simple method using two oligonucleotide primers and a single-stranded template, *DNA* 3, 479 (1984)	CSHL	260	M
Sweet, R.W. et al, The product of ras is a GTPase and the T24 oncogenic mutant is deficient in this activity *Nature* 311, 273 (1984)	CSHL	255	F

Table 6 (continued)

	Institute	Citations	Content
McGinnis, W. et al, A conserved DNA sequence in homeotic genes of the Drosophila Antennapedia and the bithorax complexes, *Nature* 308, 428 (1984)	BIOB	233	S
Galfre, G. et al, Preparation of monoclonal antibodies: strategies and procedures, *Meth. Enzymol.* 73, 3 (1981)	MRCC	233	M
Weeds, A., Acting-binding-proteins-regulators of cell architecture and motility, *Nature* 296, 811 (1982)	MRCC	233	R
Lengyel, P., Biochemistry of interferons and their actions, *Annu. Rev. Biochem.* 51, 251 (1982)	YALE	232	R
Richmond, T.J. et al, Structure of the nucleosome core particles at 7 A resolution, *Nature* 311, 532 (1984)	MRCC	223	S
Mount, S.M. et al, The U1 small nuclear RNA-protein complex selectively binds a 5′ splice site in vitro, *Cell* 33, 509 (1983)	YALE	222	F
Krämer, A. et al, The 5′ terminus of the RNA moiety of U1 small nuclear ribonucleoprotein particles is required for the splicing of messenger RNA presursors, *Cell* 38, 299 (1984)	MPIM	219	F
Westhof, E. et al, Correlation between segmental mobility and the location of antigenic determinants in proteins, *Nature* 311, 123 (1984)	MRCC	215	S
Goldfarb, M. et al, Isolation and preliminary characterization of a human transforming gene from T24 bladder carcinoma cells, *Nature* 296, 404 (1982)	CSHL	212	F
Mellon, P. et al, Identification of DNA sequences required for transcription of the human globin gene in a new SV40 host-vector system, *Cell* 27, 279 (1981)	CSHL	211	F
Changeux, J.P. et al, The acetylcholine receptor: an allosteric protein engaged in intercellular communication, *Science* 225, 1335 (1984)	PAPA	207	R
Staden, R., An interactive graphics program for comparing and aligning nucleic acid and amino acid sequences, *Nucleic Acids Res.* 10, 2951 (1982)	MRCC	205	M
Meyer, D.I. et al, Secretory protein translocation across membranes–the role of the "docking protein", *Nature* 297, 647 (1982)	EMBL	205	F
Bothwell, A.L.M. et al, Heavy chain variable region contribution to the NPb family of antibodies: somatic mutation evident in a y2a variable region, *Cell* 24, 625 (1981)	IGUC	205	S
Edgar, D. et al, The heparin-binding domain of laminin is responsible for its effects on neurite outgrowth and neuronal survival, *EMBO* 4, 1463 (1984)	MPIB	201	F
Padgett, R.A. et al, Splicing of messenger RNA precursors is inhibited by anti-sera to small nuclear ribonucleoproteins, *Cell* 35, 101 (1983)	YALE	200	F
Helenius, A. et al, On the entry of Semliki Forest virus into BHK-21 cells, *J. Cell Biol.* 84, 404 (1980)	EMBL	199	F
Shimizu, K. et al, Structure of the Ki-ras gene of the human lung carcinoma cell line Calu-1, *Nature* 304, 497 (1983)	CSHL	195	S
McGinnis, W. et al, A homologous protein-coding sequence in Drosophila homeotic genes and its conservation in other metazoans, *Cell* 37, 403 (1984)	BIOB	195	S

Table 6 (continued)

	Institute	Citations	Content
Engelman, D.M. et al, Path of the polypeptide in bacteriorhodopsin, *PNAS* 77, 2023 (1980)	MRCC	192	S
De Crombrugghe, B. et al, Role of the cyclic AMP receptor protein in activation of transcription, *Science* 224, 831 (1984)	PAPA	186	F
Pearse, B.M.F. et al, Membrane recycling by coated vesicles *Annu. Rev. Biochem.* 50, 85 (1981)	MRCC	184	R
McKay, D.B. et al, Structure of catabolite gene activator protein at 2.9 A resolution suggests binding to left-handed B-DNA, *Nature* 290, 744 (1981)	YALE	183	S
Engelman, D.M. et al, Insertion and secretion of proteins into and across membranes: the Helical Hairpin Hypothesis, *Cell* 23, 411 (1981)	YALE	178	S,F
Brigati, D.J. et al, Detection of viral genomes in cultured cells and paraffin-embedded tissue sections using biotin-labeled hybridization probes, *Virology* 126, 32 (1983)	Yale	176	M
Neupert, W. et al, How Mitochondria import proteins from the Cytoplasm, *TIBS* 6, 1 (1980)	BIOB	174	R
Schatz, G. et al, How are Proteins imported into Mitochondria *Cell* 32, 316 (1983)	BIOB	170	R
Perucho, M. et al, Common and different transforming genes are contained in human tumor derived cell lines, *Cell* 27, 467 (1981)	CSHL	165	F
Powers, S.T. et al, Genes in S. cerevisiae encoding proteins with domains homologous to the mammalian ras proteins, *Cell* 36, 607 (1984)	CSHL	165	S

The content was determined according the title and abstract of the article: method (M), predominantly structure (S), structure–function (S,F), predominantly function (F), and review (R). Citations are not weighed in case of collaboration. Self-citations are subtracted.

invented and in part had collected. One might call it a methodological review. This book was cited more than 3000 times up to 1986. The inclusion would have raised the results of the first ranked Cold Spring Harbor Laboratory from 40 to 50 average citations per paper.

We count nine methodological articles in the list of the 50 most cited articles. This is fewer than one might have expected. The methodological articles originate from four institutes. Although four of these nine articles are written by scientists from one institute, the MRCC, we are able to show that they give only a minimal advantage to this institute. We conclude that there is no bias caused by methodological papers in the case of institutes. The ranking of the research groups presented in Table 5 however is affected. One group from CSHL and one from EMBL published two very frequently cited methodological papers. Thus, both groups are ranked among the most successful of the groups.

3.5. The journal impact factor

We compared the ranking according to the average citation per paper with the ranking according to the JIF (Table 2) of the Institute of Scientific Information (ISI). The JIF is calculated by dividing the total number of citations of one- and two-year-old articles published in a journal by the total number of these articles. It is published yearly in the *Journal Citation Reports* of the ISI. We found a highly significant correlation value ($r = 0.93$) in the case of the institutes and a significant correlation value ($r = 0.77$) in the case of the research groups. Furthermore, we compared the average citation of an article published in a certain journal in our study and the respec-

Table 7
Journal impact factor of the Institute for Scientific Information and the average number of citations per article of 20 journals based on our data

Journal	A	B	C	D
Cell	15.2	68.8	76.8	174
J. Cell. Biol.	9.2	30.9	37.2	79
PNAS	8.9	42.5	49.5	298
Nature	8.3	62.7	69.2	219
J. Molecular Biology	6.2	32.1	37.2	270
Nucl. Acids. Research	6.0	31.3	35.8	246
J. Biological Chemistry	5.8	21.5	26.9	219
EMBO Journal	5.7	27.9	33.6	256
Gene	4.9	25.5	30.2	46
J. Virology	4.5	14.5	19.9	73
Biochemistry U.S.	4.3	17.1	21.5	147
Developm. Biology	3.7	18.8	24.4	47
Eur. J. Biochemistry	3.5	14.7	19.5	218
Virology	3.3	15.7	20.4	78
FEBS Lett.	3.0	9.4	13.3	202
Molecular General Genet.	2.8	9.5	13.3	149
J. Bacteriology	2.7	11.9	16.2	77
Biochim. Biophys. Acta	2.6	12.4	15.0	101
Biol. Chem. Hoppe-Seyl.	2.4	5.1	8.6	155
Ztschr. Naturforschung C	1.2	3.6	6.3	75

A: JIF averaged for the years 1980–1984 (ISI).
B: Mean citation frequency of an article per five years, without self-citations.
C: The same as B, except that self-citations are not subtracted.
D: Number of articles used in our study.

tive JIF calculated by the ISI. If we consider the 20 journals which contained most articles of the institutes, both values are significantly correlated (r = 0.92; Table 7). The mean citation rate of an article in a particular journal in our study is 4.3 times higher (with a standard deviation of 1.2) than the average JIF (1980–1984) of that journal. Curiously, in the case of *Nature*, our average citation value is 7.6 times higher than the JIF. *Nature* is a multidisciplinary journal covering articles from many scientific fields with different citation habits, for example geology and molecular biology. The JIF apparently underestimates the influence of *Nature* in the field of molecular biology. We used our data to correct the JIF of *Nature* (1980–1984) from 8.3 to 14.5 (the mean citation rate of the *Nature* papers in our study divided by 4.3). Then the above-mentioned correlation in the ranking of the institutes changes slightly, from r = 0.93 to r = 0.95.

We conclude from the data presented in Table 2 that ranking institutes in terms of the average number of citations per paper and by average JIF is similar in result if a sufficiently large number of articles per journal is available. Quantitative ranking by JIF is therefore valid for the assessment of large units, for instance research institutes. In single cases it might be misleading. The JIF depends on what is assumed to be a citable item. This can be illustrated in the case of the journal *Nucleic Acids Research*. Since 1986 the journal has published a service called "For the Record Sequences" and since 1987 so called RFLPs (restriction fragment length polymorphism). These very short 'notes' (two RFLPs per page) were rarely cited, although they were of use for human genetics. The inclusion of these services increased the number of 'citable items' from 645 in 1985 to 1942 in 1990. At the same time the JIF decreased from 6.0 in 1985 to 3.2 in 1990. Both kinds of articles were excluded from the journal in 1992. Presumably the JIF will increase again.

The JIF has been used as a 'calibrating' value representing the number of citations that articles from different fields are expected to receive (Moed et al., 1985). The actual number of citations is compared to the expected number and the difference is interpreted. The value of this approach lies in the possibility of comparing different research fields. The citations are standardized by the citing patterns of the field-specific scientific journals. For our purpose, this approach is of no use. In our interpretation, the JIF reflects to some extent the quality attributed by the peer review process to the submitted manuscripts. To get a paper published in a high-impact journal like *Nature* or *Cell* indicates quality independent of the actual number of citations this article will receive. The calibration by JIF neglects differences in the quality of the journals.

4. The different funding of research

It is difficult to determine the exact costs of research of an institute. There are various indicators that might be used. Most of them are not suitable. One may count the authors of the research publications. This procedure is misleading in two directions: it penalises research groups or institutes with a high turnover of scientists, and gives an edge to institutes with permanent scientists who do not publish. To reduce this problem one may count the number of permanent positions. But how does one compare the costs of an undergraduate student in a university laboratory with the costs of a postdoctoral fellow in a research institute? We asked for the total budget, i.e. the money spent. Money seemed to us the most trustworthy indicator. We therefore asked for detailed information about the costs of the year 1982. We assumed the costs would be the same in the other four years we analysed: this might be regarded as a shortcoming. We are not sure, for instance, whether this particular year was a year of atypical costs in one of the institutes we studied. But it is hard to imagine that the director who provides the budget information would not stress such a detail. Furthermore, to collect this information needed the substantial work and collaboration of many people. To ask for additional information from 1980 to 1984 seemed to us to be an unreasonable demand, with the possible effect that we may lose the collaboration of some of the institutes.

We asked in particular for the permanent budget: total costs of salaries of all employed, research and library costs, electricity, heat, maintenance and so on. The sum does not include costs of central administration or other central costs. We also asked for all grants, but not for stipends. In the case of the EMBL we concentrated on the laboratory in Heidelberg (the outstations in Grenoble and Hamburg were excluded). Furthermore, its division of instrumentation was excluded because the construction of apparatus is untypical of the other institutes. For more detailed information see the legend of Table 3.

The problems increase if institutes are compared that differ in organization and funding. Budgets of university institutes, for example, include the costs of teaching. Budgets of research institutes include costs of administration that might not appear in the university institutes' budget. We were unable to obtain and deduct their exact amounts. In addition, a comparison across national boundaries has to face the problems of different taxes, salaries and so on. The permanent budget of the university institutes was divided into 45% costs of research and 55% costs of teaching. The convention of the German national reports to the OECD states 45% of the permanent budget corresponds to the costs of research and 55% of the permanent budget corresponds to the costs of teaching. The results in Table 3 are calculated with and without allowing for teaching costs according to this convention. Sometimes the convention may be misleading: the permanent yearly budget of the BIOB of more than 20 million DM is not an indication of exceptionally intensive teaching, but of an effort by the local government to strengthen the research position of the institute.

The GBFB is the only national research centre in our list. One may doubt whether this institute is fully comparable with the others. On the one hand it concentrates on biotechnological methods in medicine, chemistry and environmental research (Wissenschaftsrat, 1991, p. 16), which is not different from the research done in other institutes. On the other hand, the institute is also engaged in applied research with close industrial collaboration (Gesellschaft für biotechnologische Forschung mbH, 1986, introduction). Furthermore, the institute published more articles in chemical journals than the other institutes. In another study (Irvine et al., 1990), the budget of the GBFB was fractionated: 40% of the permanent institutional budget was assumed to be costs of basic research in biological sciences. They give no reason for this procedure, they just mention discussion with officials (Irvine et al., 1990, note 17d, p. 76). If one accepts this weighing procedure, the results of GBFB would be similar to those of the MPIB.

5. Discussion

5.1. Limitations of citation analysis

Referees who evaluate proposals of scientists first evaluate their performance in the past to get an impression of the quality of the work they have done. An easy way to do this is by counting the number of publications within the last five years and then finding how many of them were published in high quality journals. Opinions differ among scientists as to which journals belong to this category. It depends on field and specialities. We asked 106 scientists working in the field of molecular biology whether they agree with a list of journals ranked by JIF (data not shown). There was little difference between the subjective evaluation by the scientists and the ranking by JIF in the case of the high quality journals like *Cell, Nature, Science, PNAS* or the EMBO journal, and their leading role in molecular biology. In this sense there is no great difference between assessing the work of scientists by considering the quality of the journals where previous work was published, or the quantitative calculations of the citation based JIF. Both rely on previous peer review.

But in a second and most important step, the referee of a research proposal will try to understand the published work to assess its content. In the end, the decision will be based on the content and not on the result of a citation analysis. It would undermine peer review if citation analysis were to replace the analysis of content more and more. Indeed it would be the end of science as science. Science would turn into some kind of art. We warn explicitly against such a development.

On the other hand, it is legitimate for politicians who oversee and determine the allocation of public money to ignore contents they cannot understand and to use citation analysis as an aid for decision-making, especially at the higher level of institutions. In the case of assessing the work of single scientists, citation analysis has only a complementary character. It should not replace the question Otto Warburg used to ask: "What has he invented?" ("*Was hat er erfunden*?").

5.2. Collaboration

There was no difference in the average citation per paper with and without collaboration. This result is in contrast to findings of Narin et al. (1991) and Münzinger and Daniel (1994), who conclude that collaborations, in particular international collaborations, lead to higher citation counts. We did not differentiate collaboration by nation, only by research groups from different research institutes. Therefore our finding of no difference may derive from the fact that we combine the results of institutes from European countries as well as two institutes from the USA in our calculation. If we examine only articles in journals produced by German research institutes (excluding the international EMBL), we find 9.6 average citations per paper without any collaboration ($N = 1279$, self-citations are subtracted), and 13.6 average citations per paper with at least two collaborating research groups ($N = 1128$). Statistically, this difference is highly significant ($t = 4.89$, $p <= 0.001$).

It is not easy to decide whether this small difference in the average citations per paper really reflects a substantial improvement in research by collaboration. For technical reasons we could not identify self-citations of the cooperating research groups. The self-citation rate of the institutes we studied, however, was 4.3 average self-citations per paper in these cases. If we assume a similar proportion of self-citations by the collaborating groups, the difference in the average citations per paper will disappear. Münzinger and Daniel (1994) studied the 300 most cited articles with at least one scientist affiliated to a German institute in the time period 1973–1987 (split into three time periods of five years each). They report 26 average citations per paper in the case of international co-authorship and 19 average citations per paper without this kind of collaboration in molecular biology. Furthermore they found that half of the 300 most cited papers were not the result of any collaboration. They also treated EMBL in Heidelberg as a German research institute. To get a comparable set of articles, we studied the 104 most cited papers by German institutes (EMBL included) in our five-year period. When we followed their method (number of citations divided by five) to get the average yearly citation per paper we found again that half of the articles (52) were not the result of any collaboration. Furthermore, there was a negligible difference in the average yearly citations per article (23.2 without any documented collaboration versus 24.2 in the case of collaboration).

We conclude that the difference in average citation per paper with and without collaboration reported by Münzinger and Daniel can be largely explained by the increase of self-citations in the case of collaboration.

Narin et al. (1991) report that articles published in 1977 in the field of biomedical research based on international collaboration received twice as many average citations per paper as those articles without collaboration. Furthermore, they state that this is a finding that holds for the biomedical and non-biomedical fields. They interpret this finding to be in part the result of a self-selection process: "the better scientists are the scientists who travel, cooperate and author papers internationally." (Narin et al., 1991, p. 322). What are the reasons for the different findings? One reason might be due to the different sample used: Narin et al. focused on research at a national level. Therefore their sample includes the work of the best, as well as the work of the mediocre, who might be the more numerous. We concentrate on a sample of the leading research institutes. Perhaps there is a minimum of international collaboration necessary to improve the research quality of the mediocre. In agreement with Narin et al., we argue that excellent research is per se international. It would be interesting to see whether there is a difference in the observed development of increasing collaborations and the reported success comparing the mediocre and the leading institutes. We conclude that a differentiation is necessary to evaluate collaboration.

One may also ask the question whether the indicator (authorship and institutional affiliation documented in the journal) is valid. There is only little knowledge about this (Moed et al., 1991). For example the indicator cannot differentiate between different forms of collaboration, such as stimulating discussions on the one hand, and the providing of material on the other hand. Furthermore, the rules of becoming an author of a particular paper might have changed in molecular biology under the pressure of 'publish or perish'. To provide material for experiments under the

Table 8
Ranking of the institutes according to cost (efficiency) of research

Institute	A	B	C	D	E	Ranking order			
						A	B	C	D
PAPA	2.40	149.1	2.40	149.1	28.8%	1	1	7	6
EMBL	2.69	188.9	2.69	188.9	37.3%	2	5	8	8
CSHL	2.74	212.0	2.74	212.0	52.2%	3	6	9	9
SFBK	> 2.77	> 170.0	> 2.07	> 126.8	15.0%	4	2	5	5
IGUC	2.85	171.3	2.02	121.5	33.1%	5	3	4	2
YALE	2.86	183.2	1.97	126.6	37.2%	6	4	3	4
BIOB	3.0	213.6	1.59	113.1	26.7%	7	7	1	1
IVUG	3.44	226.9	2.38	156.2	14.7%	8	9	6	7
IBUF	3.45	220.5	1.93	123.1	22.2%	9	8	2	3
MPIM	3.75	325.5	3.75	325.5	11.8%	10	10	10	10
MPIB	5.27	390.8	5.27	390.8	14.9%	11	11	11	11
GBFB	> 11.72	> 878.3	> 11.72	> 878.3	11.5%	12	12	12	12

A: Costs of citations (10^3 DM) per citation. Reviews and abstracts are excluded. Articles and citations are weighed as described in Section 2.4 in cases of collaboration. Costs of teaching are not subtracted.
B: Costs of articles with at least 20 citations (without self-citations), (10^3 DM) per article. Reviews are excluded. Articles are weighed in cases of collaboration. Costs of teaching are not subtracted.
C: Costs of citations (10^3 DM) per citation. Reviews and abstracts are excluded. Articles and citations are weighed in cases of collaboration. Estimated costs of teaching (= 55% of the permanent institutional support) are subtracted.
D: Costs of articles with at least 20 citations (without self-citations), (10^3 DM) per paper. Reviews are excluded. Articles are weighed in case of collaboration. Estimated costs of teaching are subtracted.
E: Articles with at least 20 citations (without self-citations) as percentage of all articles with at least one citation from outside.

condition of becoming author of the paper, for example, is nowadays usual, but it was not so in the past.

5.3. Funding by permanent budget versus funding by grants

Efficiency of research may be defined in terms of the amount of money needed to produce high quality results. In assessing the efficiency, it is of no use to count only the number of publications. Their number says nothing about the quality of the work done if the objects studied are excellent institutes. The number of publications might be useful in, for example, bibliometric studies of all institutes at a national level. We therefore use the number of citations per article as a measure of the quality (usefulness) of research. In an economic sense, citations can be seen as an indicator of demand: the information of the paper has been 'bought', i.e. used. We therefore use average costs per citation as an indicator of efficiency.

Two problems arise as a result of this assumption. First, we assume that in molecular biology only one type of basic equipment is necessary in order to solve most problems. Indeed, in most cases (but not X-ray work) there is no need for *very* expensive apparatus as there is, for example, in particle physics. This may change. Second it is assumed that 10 publications each receiving 20 citations indicate the same efficiency as one publication receiving 200 citations. This is a shortcoming since the indicator ignores the number of publications. We therefore constructed a second indicator of efficiency: the costs of producing publications of a particular quality, i.e. reaching a certain number of citations. Table 8 indicates the costs of a publication that received at least 20 citations in the five-year period. We chose a low limit of citations to handle the possible differences among the citation counts in the various subfields of molecular biology. Only one third of all publications, however, fulfilled this condition. The chosen number of minimal citations is arbitrary. Table 9 indicates that there would be only small differences if one chose another number of citations. The line 'costs per article' in Table 3 shows that the ranking would be reversed if one used number of papers as an indicator of quality. In fact the CSHL published the fewest papers with one of the highest budgets of the institutes in our comparison. But 37% of these papers were published in *Cell, PNAS* and *Nature* (Table 1).

Table 9
Percentage of the articles of the institutes with increasing minimal citations

	Percentage of all articles with at least 20, 30 and 40 citations				Ranking order		
	All	⩾ 20	⩾ 30	⩾ 40	⩾ 20	⩾ 30	⩾ 40
CSHL	100	52.2	37.5	26.6	1	1	1
MRCC	100	43.5	31.2	23.6	2	2	2
EMBL	100	37.3	26.3	20.2	3	4	3
YALE	100	37.2	26.6	17.3	4	3	4
IGUC	100	33.1	20.1	12.5	5	5	6
PAPA	100	28.8	17.9	11.9	6	6	7
BIOB	100	26.7	17.9	14.3	7	7	5
IBUF	100	22.2	14.2	10.8	8	8	8
SFBK	100	15.0	4.6	2.5	9	13	13
MPIB	100	14.9	8.0	5.3	10	9	9
IVUG	100	14.7	5.7	4.0	11	10	10
MPIM	100	11.8	5.7	3.6	12	11	11
GBFB	100	11.5	6.2	2.5	13	12	12

Abstracts and reviews are excluded and self-citations are subtracted. Articles are weighed as described in Section 2.4 in cases of collaboration.

The researchers of the CSHL seem to follow a strategy of concentrating their results in only a few highly cited papers.

Both indicators of efficiency lead to similar ranking orders. The Pasteur Institute did very well in comparison with the university institutes – even after estimating and subtracting their costs of teaching. After subtracting the estimated costs of teaching, all university institutes bar one are better placed than the research institutes. One has to keep in mind that this kind of analysis has an explorative character. In two cases, the small university institutes in Freiburg and Giessen, the chosen time period of five years leads to only a small number of publications (both about 90 weighed articles, see Table 3). The Freiburg Institute profited from one excellent scientist and his research group, which is obviously reflected in our efficiency calculation. The costs per citation, however, would only rise from 1,900 DM to 2,700 DM per citation if one excluded the outstanding group and kept the budget, an obviously unfair procedure. Nevertheless the costs would still be substantially smaller than the costs of the MPIB (5,200 DM per citation) and there would still be a remarkable difference in comparison with the MPIM (3,700 DM per citation).

In the case of the Giessen Institute one or two highly cited papers in 1985 or 1986 might change its position. The weighing for teaching costs might be misleading in the case of the institutional budget of the BIOB and the biased results of the Freiburg institute. However, the position of the BIOB without weighing for teaching costs indicates that even a smaller factor would result in a good ranking position. Concentrating on German institutes, both Max Planck Institutes appear to be the most expensive i.e. ineffective, if one disregards the GBFB.

One central question of this study was whether a funding system based on ample institutional support leads to better results than a funding system based on grants. If we consider our results this question seemed to be misdirected. The MRCC and the EMBL, both funded with ample institutional support are as highly ranked as the CSHL and YALE, both funded totally or mainly by grants. As intermediate variables, (a) the turnover of scientists, and (b) the time period a scientist is supported seem to play an important role. In the EMBL and the MRCC the number of scientists with permanent positions is small. Most appointments are limited in time. So the question may be also phrased in terms of the question how long a successful scientist should be funded with ample support: three years as it is in the "Sonderforschungsbereich" of the Deutsche Forschungsgemeinschaft, five to eight years as it is in certain American grants (NSF, Howard Hughes) or until retirement as it is in the Max Planck system. A five-year period seems to be a good compromise. In some cases even longer support may be adequate. Some of the less productive groups of both Max Planck Institutes were groups with directors close to their retirement. Was this a very special situation of just these directors, or was this a result that in general does not support the Max Planck system?

Some of the most productive research institutes are characterized by a high turnover of scientists. The Institut Pasteur is an interesting exception. There, almost all scientists have permanent positions, but not permanent funding. In a university institute where research is done by graduate students, a high turnover of scientists is a structural element. Both the turnover of scientists and the grant proposal system seem to function as a quality control mechanism: they make it more difficult to get money for less interesting research. The grant proposal system, however, has already reached a crucial point. There is not enough money to fund all the proposals found to be of high scientific interest. The Berlin and Munich Max Planck Institutes, covering both excellent and unproductive groups during the period we studied, do not emerge so favourably if we regard the institutes as a whole. This indicates that ample institutional support without regular strict, and not just formal, quality control might be the wrong way of organizing research in times when money is getting tight.

Acknowledgements

We thank H.D. Daniel, R. Ehring and Andrew Barker for helpful discussions. This work was supported by DFG Schwerpunkt Wissenschaftsforschung.

References

Anderson, A., 1991, No citation analyses please, we're British, Science 3 May, 639.

Anderson, A., 1992, Molecular biology: U.S. juggernaut overwhelms divided European elite, Science 256, 460-464.

Anonymous, 1992, Top 50 research institutions in molecular biology ranked by citation impact 1981–1991, Science Watch 3, 7.

Bauin, S. and H. Rothman, 1991, Der 'Impact' von Zeitschriften als Annäherungsmass für Zitationsraten, in: P. Weingart et al. (Eds.), Indikatoren der Wissenschaft und Technik, (Campus, Frankfurt/Main) 91–111.

Cole, J.R. and S. Cole, 1973, Social Stratification in Science (University of Chicago).

Delabar, J.-M., D. Goldgaber, Y. Lamour, A. Nicole, J.L. Huret, J. de Grouchy, P. Brown, D.C. Gajdusek, and P.M. Sinet, 1987, β amyloid gene duplication in Alzheimer's disease and karyotypically normal Down's syndrome, Science 235, 1390-1392.

Gesellschaft für biotechnologische Forschung mbH, 1986, Wissenschaftlicher Ergebnisbericht, Vorwort.

Hamilton, D.P., 1990, Publishing by- and for? -the numbers, Science 7 December 1331–1332.

Hamilton, D.P., 1991, Research papers: Who's uncited now?, Science 4 January, 25.

Healey, B.P., J. Irvine and B. Martin, 1988, Quantitative science-policy studies in the United Kingdom, Scientometrics 14, 177–183.

Herbertz, H. and B. Müller-Hill 1992, European research institutes, Science 258, 1069.

Irvine, J., B.R. Martin, and P.A. Isard, 1990, Investing in the Future: An International Comparison of Government Funding of Academic and Related Research (Edgar Elbar, Brookfield, USA) 50–78.

Kang, J., H.G. Lemaire, A. Unterbeck, J.M. Salbaum, C.L. Master, K.H. Grzeschick, G. Multhaup, K. Beyreuther and B. Müller-Hill, 1987, The precursor of Alzheimer's disease amyloid A4 protein resembles a cell surface receptor, Nature 325, 733–736.

Martin, B.R. and J. Irvine, 1983, Assessing basic research, Research Policy 12, 61–90.

McKay, D.B. and T.A. Steitz, 1981, Structure of catabolite gene activator protein at 2.9 Å resolution suggests binding to left-handed B-DNA, Nature 290, 744.

Moed, H.F., W.J.M. Burger, J.G. Frankfort and A.F.J. Van Raan, 1985, The use of bibliometric data for the measurement of university research performance, Research Policy 14, 131–149.

Moed, H.F., R.E. de Bruin, A.J. Nederhof, R.J.W. Tijssen, 1991, International scientific co-operation and awareness within the European Community: Problems and perspectives, Scientometrics 21, 291–311.

Müller-Hill, B., 1991, Funding of molecular biology, Nature 351, 11–12.

Münzinger, F. und H.D. Daniel, 1994, Biochemie und Molekularbiologie in der Bundesrepublik Deutschland: Internationale Ko-Autorschaft und deren Einfluss auf die Zitierhäufigkeit von Veröffentlichungen 1973 bis 1987, Futura 1, 20–25.

Murdoch, G.H., L. Manuelidis, J.H. Kim, and E.E. Manuelidis, 1988, β-amyloid gene dosage in Alzheimer's disease, Nucleic Acids Research 16, 357.

Narin, F., K. Stevens, and E.S. Whitlow, 1991, Scientific co-operation in Europe and the citation of multinationally authored papers, Scientometrics 21, 313–323.

Nederhof, A.J. and E.C.M. Noyons, 1992, Assessment of the international standing of university departments' research: A comparison of bibliometric methods, Scientometrics 24, 393–404.

Pendlebury, D., 1990, Cold Spring Harbor tops among independent labs, Scientist, March 19, 20.

Pendlebury, D.A. and A. Anderson, 1992, European research institutes, Science 258, 1069–1070.

Podlisny, M.B., G. Lee, and D.J. Selkoe, 1987, Gene dosage of the amyloid β precursor protein in Alzheimer's disease, Science 238, 669–771.

Tanzi, R.E., E.D. Bird, S.A. Latt and R.L. Neve, 1987, The amyloid β protein gene is not duplicated in brains from patients with Alzheimer's disease, Science 238, 664–666.

Taubes, G., 1993, Measure for measure in science, Science 14 May, 884–886.

Wissenschaftsrat, 1991, Empfehlungen zur Zusammenarbeit zwischen Grossforschungseinrichtungen und Hochschulen (Berlin, 29.1.1991) 17.