Scientometrics, Vol. 17, Nos 1–2 (1989) 127–163

CITATION CONTEXT ANALYSIS AND AGING PATTERNS OF JOURNAL ARTICLES IN MOLECULAR GENETICS

KATHERINE W. McCAIN, KATHLEEN TURNER

College of Information Studies, Drexel University, Philadelphia, PA 19104 (USA)

(Received October 27, 1988)

To compare citation history and contextual "importance," eleven highly cited articles, 4 slowly aging (Type 1) and 7 quickly aging (Type 2), were ranked using an aggregate citation context measure, the Mean Utility Index. Based on citations in late (PY 6 & 7) source articles, "methods" papers consistently ranked higher than papers cited for research results and theoretical implications, and Type 1 methods papers ranked above all Type 2 papers. A Type 1 paper representing an important theoretical concept could not be distinguished from Type 2 papers using citation context alone.

Introduction

The goal of scientific research is the production of public scientific knowledge and the creation of consensus, concerning this knowledge, in the research community.[1] The journal literature plays an important role in establishing this consensus. The scientific journal article is the major formal channel of communication. It is a validated public archival record of individual scientific activity, a vehicle for establishing priority of discovery, and a prescription for repeating or extending the reported research.[2]

The references (citations) in the scientific paper have been characterized as serving a wide variety of functions.[3–4] For instance, article references 1) demonstrate that the author understands the current consensus, 2) validate scientific claims, and 3) embed the paper in the preexisting consensus.[1] Authors may use citations to acknowledge "intellectual property rights",[5–6] as "tools of persuasion",[7] or to serve various "social" functions such as paying homage to pioneers.[8]

In general, the choice of works to be cited reflects the citing author's perceptions of how the scientific community and its knowledge base are structured and previous contributions valued. Bibliometric analysis of citation choice patterns within fields can demonstrate the aggregate perceptions within the community of publishing scientists.[9] These analyses are a useful quantitative indicator of scientific activ-

Elsevier, Amsterdam–Oxford–New York–Tokyo
Akadémiai Kiadó, Budapest

ity[10] and an unobtrusive research method for the study of scientific literatures and the scholars who produce them.[11]

One major theme in bibliometric studies of the scientific journal literature has been investigation of the aging (or obsolescence) of various literatures – by examining (1) changes in citation counts to and from journal aggregates and individual journal titles over time (diachronous studies) or (2) the age distribution of works cited at one point in time in individual articles, journals and journal aggregates (synchronous studies).[12–14] In these studies the focus has been primarily on the scientific *journal* as a set of periodically published volumes which themselves are article aggregates. Much less attention has been paid to aging patterns of individual journal articles – the primary information products of basic scientific research[2] and the vehicles for formally communicating to the scientific community information concerning research results, availability of experimental materials, and existence of useful research techniques.[15] Several recent studies have examined citation histories of individual papers or individual author's *oeuvres.*[16–22] To date, however, *Aversa*[23–24] represents the only extensive systematic investigation of aging patterns in individual scientific journal articles.

Aversa characterized the citation patterns and aging rates of 399 papers published in 1972 that were "highly cited" (30+citations in the first 5 years) in sources indexed in the print *Science Citation Index* from 1972 through 1980. She identified two distinct clusters of papers with markedly different diachronous citation patterns as described by peak citation year and rate of citation decline in the last four years: On the average, Cluster I papers received the largest fraction of total citations in the sixth year following publication, with a gradual drop-off in citations, while Cluster 2 papers showed a citation peak in the third year followed by a rapid decline in citations. Discriminant analysis identified four "citation-related" variables (proportions of total citations in publication years 6–9) as contributing to cluster membership, classifying approximately 96% of the 399 papers correctly. For a subset of 54 papers, two other variables, author prior citedness and number of coauthors, correctly classified approximately two-thirds of the papers. No text-related variables (citation context, citation content) were examined.

The present study

In the course of an investigation tracing the diffusion of methodological innovations in molecular genetics, we identified a number of early papers published between 1978 (the starting point of our study) and 1980 which were noticeable highly cited over the subsequent 6 to 8 years. We were intrigued to discover that, based on annualized citation counts from *Science Citation Index*, the citation patterns for eleven papers meeting Aversa's criterion (30+ citations in 5 years) very clearly fell into one of the two cluster pattern classes. Four papers had late (year 6+) citation peaks with an apparently slow citation drop-off (Aversa Cluster 1); seven papers peaked in year 2 or 3 with a rapid decline in citations thereafter (Aversa Cluster 2). There were no citation peaks in publication years 4–5.

These genetics papers are alike in many ways (see discussion in next section) and potentially serve the same range of communication functions. They all report (1) empirical research results in molecular genetics, (2) the use of specific recently introduced genetic engineering techniques, and (3) the creation of new experimental materials. Yet the citation histories of the papers, reflecting the subsequent use of the papers and their information content, place them clearly into two distinct groups. *Aversa*'s research showed that the citation percentages in years 6–9 were good "predictors" of citation history – in other words, that later curve slope accounted for overall curve profile and thus cluster assignment. While useful for classification, these particular variables do not add to our understanding of information use and communication processes in scientific research and publication – the mechanisms which ultimately determine an article's citation history.

In this paper we explore the possible relationship between several content-related citation variables and the apparent cluster pattern membership of these eleven "key papers". We are concerned primarily with the *use* that citing authors make of each of these papers and whether the citation history of a paper (reflected in cluster membership) can be associated with the type of use or changes in use made by citing authors early and late in a paper's life span. We assume that the choice of references in a research report or review article is based ultimately on the perceived usefulness (however defined by the author) of the cited works and their content to each author's research or scholarly exposition. Rather than polling citing authors,[25] we take the citation context and citation content[3] in the citing paper as evidence of the nature of the link(s) the author perceived between his or her work and the cited paper. Although we do not pose statistically testable research hypotheses, our working assumption is that our Cluster 1-type key papers with "staying power" (as Aversa puts it) will demonstrate a higher "perceived usefulness" than the Cluster 2-type key papers as measured on these text-based citation variables. Through our analysis, we hope to identify some generalizable content-related characteristics which influence a article's post-publication recognition, its perceived usefulness in subsequent scientific scholarship, and thus its citation history.

Methods

Key papers

In 1978, *Maniatis* et al.[26] published a paper in *Cell* which described in detail the protocols for constructing and screening a eukaryotic [non-bacterial] gene "library" – a random collection of fragments of DNA which (ideally) includes all the genetic information of that particular organism (in this paper, *Drosophila*, silkmoth, and rabbit libraries). In grossly oversimplified terms,[27] *Maniatis'* paper describes how, using recombinant genetic techniques, the DNA of a particular organism can be snipped into pieces by special enzymes and these pieces then stored (in bacterial cells) as a reproduceable collection of genes and gene fragments for future study. In principle, the researcher can isolate any gene of choice on a stored DNA fragment by "screening" the "library" (created *de novo* or acquired from someone else) with a complementary genetic "message". (The similarity to matching search term profiles and document representations is, in itself, intriguing.) A companion paper (*Lawn* et al., 1978[26]) described the creation of a gene library from human fetal liver cells and the characterization of genes for particular hemoglobin proteins. The *Maniatis* and *Lawn* papers were (and still are) very highly cited.

In the next two years, a number of other researchers used Maniatis' (and similar) techniques to create gene libraries and provide information on the structure, function, evolution, etc. of genes coding for a number of different proteins. We identified nine additional papers that reported research results, the use of particular techniques, and the availability of experimental materials and that also met Aversa's citedness criterion (30+ or more citations in the first 5 publication years). These "key papers" are listed in Table 1.

Citation counts

Initially, we searched the print *Science Citation Index* for citations to each of the key papers from their year of publication through the end of 1986. (At the end of the study we were able to include the citation counts for 1987 but these data played no part in our selection and analysis of source papers.) In *SCI*, a citation entry includes the first author's last name and initials, article publication year, abbreviated journal title, volume and first page. Citation format errors, probably originating from the reference list of the citing paper, appear in the print file as citation singletons. They are easily scanned and identified as simple page number inversions, wrong volume but right page and year (and similar errors), or "natural" typing errors (such as Slighto*n* for Slighto*m*). We found no journal *title* misassignments. We were generous in our citation tallies, including these relatively accurate as well

Table 1
Key papers in molecular genetics

Davis, M.M. et al. 1980. "An immunoglobulin heavy-chain gene is formed by at least two recombinational events." Nature 283:733-739.

Dodgson, J.B., Strommer, J. & Engel, J.D. 1979. "Isolation of the chicken β-globin gene and a linked embryonic β-like globin gene from a chicken DNA recombinant library. Cell 17:879-887.

Early, P.W. et al. 1979. "Immunoglobulin heavy chain gene organization in mice: Analysis of a myeloma genomic clone containing variable and α constant regions." Proceedings of the National Academy of Sciences USA 76:857-861.

Fritsch, E.F., Lawn, R.M. & Maniatis, T. 1980. "Molecular cloning and characterization of the human β-like globin gene cluster." Cell 19:959-972.

Kemp, D.J., Cory, S. & Adams, J.M. 1979. "Cloned pairs of variable region genes for immunoglobulin heavy chains isolated from a clone library of the entire mouse genome." Proceedings of the National Academy of Sciences USA 76:4627-4631.

Lawn, R.M. et al. 1978. "The isolation and characterization of linked δ-globin and β-globin genes from a cloned library of human DNA." Cell 15:1157-1174.

Maniatis, T. et al 1978. "The isolation of structural genes from libraries of eucaryotic DNA." Cell 15:687-701.

Sargent, T.D. et al. 1979. "The rat serum albumin gene: Analysis of cloned sequences." Proceedings of the National Academy of Sciences USA 76:3256-3260.

Slightom, J.L., Blechl, A.E. & Smithies, O. 1980. "Human fetal $^{G}\gamma$- and $^{A}\gamma$-globin genes: Complete nucleotide sequences suggest that DNA can be exchanged between these duplicated genes. Cell 21:627-638.

Tucker, P.W. et al. 1979. "Sequence of the cloned gene for the constant region of murine γ2b immunoglobulin heavy chain." Science 206:1303-1306.

Wahli, W. & Dawid, I.B. 1980. "Isolation of two closely related vitellogenin genes, including their flanking regions, from a *Xenopus laevis* gene library." Proceedings of the National Academy of Sciences USA 77:1437-1441.

as totally accurate citations as long as there was only one minor error per citation string.[28] We included self citations in our counts, both because they contribute to the overall visibility and citation history and because (practically speaking) it was impossible to eliminate non-first author self-citation without examining the full bibliographic entry for each source paper. In our analyses, however, we do distinguish among papers with matching author(s) (self-citation), matching organizational addresses (institutional self-citation), and "independent" source papers.

Source paper selection

Citation context analyses require that the citing papers be *read* to determine the context(s) in which the cited paper was used and the semantic content of the text surrounding the citation. With any substantially cited single article this can be a massive undertaking; the eleven key papers in this study received more than 3100 citations from source (citing) papers indexed in *SCI* from 1978 through 1986. To provide the greatest possible contrast among individual papers and cluster groups, while keeping the task of analysis to some reasonable size, we selected for analysis a small number of source papers published "early" and "late" in each key paper's citation history. For each key paper 10 citing papers published in years 1–2 and 10 papers published in years 6–7 were selected, by random sampling if more than 10 papers fell in these time ranges. Selected papers met the following constraints:

- Only papers in English were selected.
- Both journal articles reporting primary research and review articles were included, but not abstracts.
- Papers from earlier years (first year 5, then year 4) were included if necessary, to make up the 10 paper set in the later time period.

Citation context analysis

We were interested in discovering the general types of use made by authors citing the key papers in early and late source papers – to see whether Cluster 1 and Cluster 2 key papers had different use patterns and whether these patterns changed over time. To do this, we needed to read and characterize the *citing text* surrounding each occurrence of the key paper's citation, a citation *context* analysis. There are two inter-related approaches to this type of analysis.[3] *Context classification* approaches deal with the types or functions of references in the text and "are constructed to reflect the relationship which the citing author perceives between his/her work and the cited document"; *citation content analysis* uses "the semantic content of the citing passage to characterize the citing work." We used both approaches in our analysis.

1. Citation context classification

Small,[3] *Cronin,*[4] and *Peritz*[29] each provide a extensive comparison of citation classification schemes, and the reader is referred to these reviews for detailed discussion. Through the early 1980's citation classification schemes in the literature of the natural sciences have included published studies in physics, genetics, and chemistry.[21,30–34]

Moravcsik and *Murugesan*'s basic classification scheme[30] included four citation category pairs: conceptual/operational; evolutionary/juxtapositional; organic/perfunctory; confirmative/negational. Citation occurrences in the text were assigned to more than one of these pairs, but to only one category within the pairing. By contrast, *Chubin* and *Moitra*[31] used six mutually exclusive categories, four "affirmative" (basic, subsidiary, added information, and perfunctory) and two "negational" (partial and total). Citations were assigned to a single category. *Oppheim* and *Renn*[32] and *Spiegel-Rösing*[33] constructed classification schemes that include general characterizations of the cited work (e.g. supplying data, information, methods, etc. – for various distinct purposes) as well as additional citation functions (e.g. historical background). *Hodges'*[34] 10 "relationship indicators", used to compare citation practices in published texts with data from interviews with scholars, are a similar mix. *Ruff*[21] created a six category scheme to classify all citations to the *oeuvre* of a prominent theoretical molecular spectroscopist.

With two exceptions, these studies have been synchronous studies that classified all citations in a selected set of source papers in order to characterize the frequency and range of different citation functions in a particular literature. *Oppheim* and *Renn*[32] examined the range of citation contexts to a small set of historically important articles in physics and physical chemistry to see how they were being used in the current scholarly literature and *Ruff*[21] provided limited data on citation counts and changes in citation context over time to the set of molecular spectroscopy papers.

The scientific paper describing original research results tends to have a definite, rather formulaic structure – introduction, methods, results, and discussion – an organizational format recommended by many style manuals and required by most scientific journals.[35] This structure facilitates the reader's ability to scan and rapidly retrieve information from text.[36–37] It seems reasonable to assume that the placement of a citation within this organized text would be related to the perceived usefulness of the cited work. None of the natural science-oriented classification schemes or citation context analyses discussed above include citation *location* in their category definitions, although *Moravcsik* and *Murugesan*'s[30] category pair "conceptual/operational" distinguishes between works cited for theory or concept and those cited for tool or physical technique. *Peritz*[29] suggested that certain citation classes have their "natural place". *Voos* and *Dagaev*[38] stated that references to highly cited works are more frequently found in an article's introduction and *Ruff*[21] noted that the introductory section is likely to have more in the class "cited among others". Citation location *was* explicitly considered by *Bertram*[39] in a study relating citation location and citation "level" (the amount of the cited document represented by the citation) and by *Finney,*[40] who used a combination of cue-word types and citation location to design an automated reference classification scheme.

The citation context classification scheme developed in the present research (Table 2) represents our attempt to focus on certain fundamental aspects of information use by citing authors while making the categories as objective (and classification as replicable) as possible. For a research report, we assume that the section in which a citation occurs is a major indication of the use the author is making of the cited paper and the information it contains. (Papers without formal sections, such as short reports in *Nature*, have an introductory paragraph or two, perhaps in boldface, as well as extensive figure legends and a certain amount of text proper devoted to outlining methods and materials used.)

Table 2
Citation context classification scheme

[Research reports]

Introduction: The beginning section of the paper or the first 1-2 paragraphs in short reports published in Nature and Science.

Central: Paper cited (usually singly or paired) in one or a series of specific factual statements to set context of present research.

Peripheral: Paper cited as part of a large reference set or as part of peripherally related research.

Methods: Location varies. Usually labeled, following introduction or placed at end of article. May be explicated almost entirely in extensive figure legends (e.g. Nature, Science). Describes materials used, methods employed, adopted or modified. Characterizations of methods and

Table 2 (cont)

materials in introduction or results and discussion were coded as "Methods."

Central: All citations to methods and materials are coded as "strong" uses of previously published information.

Results & Discussion: Very often combined in research reports. Follows Introduction and (usually) Methods.

Central: Key paper cited to specifically compare/contrast reported findings with current results or to explicitly set the context of the argument (generally in the first paragraph).

Peripheral: Paper cited (usually in final paragraph) in a peripheral context – to set further research agenda or possible extention or generalization of results.

[Review articles]

Central: Key paper is cited in discussion of specific methods or materials reported [e.g. "Maniatis' (1) protocols have been used to construct eucaryotic gene libraries in a number of organisms..."]

<u>or</u> cited for specific findings or reported data [e.g. "Wahli (1) showed that the vitellogenin gene consisted of ..."]

Peripheral: Key paper is cited very generally for broad topic of research ["Wahli (1) studied the vitellogenin gene"; contrast with second example above]

<u>or</u> cited in an aggregate listing [e.g. "Many authors have used these new techniques to build genomic libraries (1-16)"]

Review articles lack this rigid structure and are treated as a fourth type of citation location. The communication role of the review article is, in many ways, similar to that of the introduction in a report of primary research. The review is both informative and tutorial, intended to provide a general orientation and bring the reader "up to speed" in some specific research area by describing the background, the previous consensus, selected recent research, and current controversies.[2,36]

Within a section of a research report, and throughout the text of a review article, a paper may be referenced in a specific, focused, detailed fashion, or peripherally, with a broad general characterization or as part of an "aggregate" reference list. Thus in our classification scheme, we distinguish between "central" and "peripheral" citation contexts – in an attempt to characterize apparent degrees of importance perceived by the author. The distinction between "central" and "peripheral" is very similar to *Moravcsik* and *Murugesan*'s[30] "organic vs. perfunctory" dichotomy. We do not consider expressed or inferred attitudes of the author (affirmative, negational) or the author's apparent motives (honorific, persuasive) in citing. In this exploratory study we wanted a small number of distinct categories which we felt might clearly demonstrate differences between the uses, over time, of different classes of papers. A overview of the classification scheme is presented in Table 2.

2. *Developing a citation "Utility Index"*

The Op. Cit. problem. Classification schemes, such as those discussed above, allow the researcher to characterize the use made of the cited work in the source paper based on the context of a given citation occurrence. Authors are not limited to one invocation of a cited work, however, and a key paper of interest may be cited multiple times in a source paper, in several different contexts. Indeed, the number of times a cited work is referenced in a source paper may be related to its perceived importance or relevance. *Voos* and *Dagaev,*[38] *Herlach,*[41] *Bonzi*[42] and *Dolman* and *Bodewicz*[16] explicitly consider the problem of multiple references, or *"op. cits"*, in citation analysis.

While it is possible to correlate multiple occurrence *counts* with judgments of relevance or overall citedness of key papers, as *Herlach*[41] and *Bonzi*[42] did, this perspective does not allow us to address our initial problem: Key paper citation histories are based on counts of source papers containing *at least one* citation occurrence (and probably more). Analysis focusing on the individual *"op cit"* context does not give an overall characterization of the usefulness of the key paper to subsequent scholarship.

It is possible to consider each of the several citation variables (expository context, citation location, authorship, *op, cits*) separately as they affect key paper Cluster assignment. It seems likely, however, that the effects of these variables may be interrelated in ways we cannot readily observe. Ideally, we would prefer a single overall measure of "perceived usefulness" that (1) appropriately combines all these variables and (2) provides a way to vary the relative importance of certain variables, to reflect different perceptions of their impact. We have attempted to do this by constructing an *index* in which all variables are included in calculating a single numeric measure, while our assumptions concerning the contribution of each variable to this index value are made explicit.

Construction of the index. In statistics and economics, an "index" is a number intended to represent the magnitude of some aggregate (e.g. indices of "real production" or the overall price level).[43] To represent the aggregate relationship between the cited key paper and the citing source paper, we developed a formula for calculating a "utility index"[44] which takes into account:

- the number of key paper citation occurrences in the source paper,
- the different locations in which citation occurrences were found (in research reports) or citation in a "state-of-the-art review",
- the expository context of each occurrence (central, peripheral),
- the explicit link(s), if any, between key paper and source paper authors and institutions (individual and institutional self-citation).

The "utility index" formula is shown below:

$$\mathrm{UI} = W_{SC}[W_i \ln(X_i + 1) + W_m \ln(X_m + 1) + W_d \ln(X_d + 1) + W_r \ln(X_r + 1)]$$

where – X_i, X_m, X_d, X_r are the citation occurrence counts in a specific location, with "central" occurrences counted as 1 and "peripheral" occurrences counted as 0.5. This, in effect, weights the sum (x) of *op. cits* in a location according to the context in which each occurs.

– W_i, W_m, W_d, W_r are weights *by location* of the *op. cits.* Four sets of weight values are used (see next section).

– W_{SC} weights the final aggregate index value based on the relationship between key paper and source paper. For source papers sharing an author with the key paper, W_{SC} is 0.10; for source papers from the same institution, W_{SC} is 0.5. "Independent" source papers have a W_{SC} value of 1.

The formula has the following properties:

- in accounting for multiple citation occurrences, it gives less than proportional weight to each incremental occurrence by taking the logarithm of the weighted *op. cit.* sum (X) in each location.[45] A count of 1 is added to the weighted sum in each location so that the log value of the sum is 0 when the occurrence count is 0.
- A source paper classified as "review" will have a non-zero value for only the last of the four terms. The converse is true for research reports.

- The formula provides a mechanism for distinguishing among citation locations and weighting by perceived importance implied by that location. For instance, the use of methods (cited in methods section) or specific research results (cited in results & discussion) may be considered, *ceteris paribus*, more important to the citing author's argument than citation in an introductory section of a research report or in a review article (see discussion below). We can represent this perceived difference by assigning a larger value to W_m or W_d than to W_i and W_r.
- The formula allows strong "penalties" for self-citation. Self-citing source papers are explicitly removed from most citation analyses, being viewed as self-serving and not representative of the paper's value as perceived by the rest of the scholarly community. However, an author's continuing publication (and self-citation) does serve to keep his or her research visible and available for others use. We prefer to "penalize" self-citation with fractional index weights, and thus allow the self-citing paper to contribute a much smaller increment to any summary measure.

Choosing numerical values for weights. Our choices of numerical values for the different weights in the Utility index formula, and our use of the log transformation for the weighted sum of citation occurrences in each location, are admittedly arbitrary and can be criticised. We have tried to make decisions consistent with previous research and assign values that appeared likely to make any existing differences in citation patterns clearly visible. Certain decisions were more straight-forward than others. Citation context (central, peripheral) seemed a clearly dichotomous variable and we weighted the sum of *op. cits* in each class accordingly, giving each "peripheral" citation occurrence half the value of one with a "central" context. The log transform of this sum incorporates our assumption that each successive occurrence of a citation adds a smaller incremental value to the perceived utility of that key paper. Self-citations are generally seen as contributing little to an "independent" assessment of quality of cited work, and we decreased the Utility Index value of such source papers by 90% (author self-citation) or 50% (institutional self-citation) accordingly.

Though we felt that making distinctions among citation locations was important, we had no strong theoretical assumptions which would guide our choice of one particular approach to weighting by location. Rather than make a single arbitrary assumption concerning the relative importance of various citation locations, we used four sets of values for location weights and calculated a separate Utility Index based on each set. These are shown below:

Utility index location weights

	Wt_{intro}	Wt_{meth}	Wt_{disc}	Wt_{rev}
Utility index #1	1	2	1	1
Utility index #2	1	3	2	1
Utility index #3	0.5	1	1	0.5
Utility index #4	1	1	1	1

In calculating all four Utility Indexes, we assume that introductory and review citations serve similar "background-related" information functions and that their placement indicates the same level of utility of the key paper to the citing author. Thus W_i and W_r always have the same value. In UI #1–#3, we vary the values assigned to "Methods" and "Results & Discussion" sections of source papers:

- Utility Index #1 assumes that the use of previously established research methods in the citing author's research is *absolutely more* important than the use of previously published experimental results or theoretical insights.
- Index #2 adds the assumption that previously published work used in explicit discussion of the citing author's research results is *more* important to the author's research and argument than that used in introductory remarks and reviews, but still *less* important than methods-related contributions.
- Index #3 simply distinguishes between any reference explicitly related to the citing author's research ("direct use") and the "background" references used in introductions and reviews. Here, "Methods" and "R&D" citations are given the same (higher) value.
- Index #4 makes no assumptions concerning relative importance of citation placement as an indicator of utility and is included for the sake of completeness. All four weights have equal value.

We recognize that this approach – looking at four measures rather than one – increases the complexity of our analysis and associated discussion. We feel this is off-set by our ability to compare our results across Utility Indexes and use them as a form of internal validation. The more similar our findings are, based on somewhat different sets of assumptions, the less likely the chance that they are simply an artifact.

The "Mean Utility Index". We can use this formula to calculate the Utility Index for a given key paper in a single source paper. More importantly, we can aggregate over a set of source papers citing the key paper and derive a *Mean Utility Index* reflecting a broader perspective on that key paper's usefulness to subsequent scholarship. We can then use the MUI to rank the key papers; higher ranking papers would be those perceived as more "useful", on the whole, by the set of citing authors. This

Mean Utility Index is similar in function to one citation-related output measure recommended by *Kochen*,[46] where "each cited paper . . . would be assigned a score composed of the number of articles that cite it, each weighted for the quality of the citation as well as for its merit and influence".

Citation content analysis

Citation content analysis deals with the semantic content of the text to which the key paper citation is linked.[3] The researcher seeks to identify the *concepts* for which the key paper is cited, rather than focusing specifically on the *context* in which the paper is used. In practice, as *Small*[3] notes, the two approaches are frequently confounded. Most citation context classifications, including this one, incorporate some generalized aspects of "theory", "data", "method", and similar concepts in definitions of types of key paper *use*. Where the two approaches differ is in the focus on citation *content* rather than the *context* of the citing-cited relationship.

The content of the citing text reflects the *perceived* content of the cited work and its role as a *concept symbol*.[47] A cited paper can represent "experimental findings, methodologies, types of data, metaphysical notations, theoretical statements or equations . . ." The degree to which highly cited papers have become "standard symbols" can be measured as the "percent uniformity" of the symbolic content in a set of citing papers.

Small[47] found the percent uniformity for 52 highly cited papers in chemistry ranged from 36% to 100%, based on an examination of 12 source documents per key paper. He suggested that one useful way to study the way works acquire these "standard" meanings would be to examine the contexts of citation in chronological sequence. Several subsequent studies have used this approach. *Cozzens*[28] analysed citation contexts of alternate citation formats of a single econometrics paper. Each format represented the paper in a separate cocited document cluster; the paper apparently stood for a distinct, different "symbolic referent" in each of the two clusters. *Small* and *Greenlee*[48] examined citations to an important techniques paper in molecular genetics and found that the percent uniformity did not increase noticeably over time. In a detailed analysis of a cocited document cluster representing research on collagen proteins, *Small*[49] reported percents of citation uniformity ranging from 58% to 100% for the 8 documents in the cluster. *Cozzens*[50] traced changes in citation content over time for two different papers – one in neuropharmacology and one in sociology of science. She distinguished very generally between citation content referring to "main knowledge claims", "method citations", and "peripheral knowledge claims" made by each paper. *Cozzens* reported a marked shift over time toward "uniform" citation of the main knowledge claim in the neuropharmacology paper, but no such trend for the sociology of science article.

Each of the eleven key papers in our study potentially represents several different "concept symbols". Each reports specific experimental findings – the identification, structure and/or characterization of the gene(s) coding for one or more related proteins. Based on these findings, several key paper author suggest broader theoretical implications for their work. In addition, each paper records the use of a specific set of methods and the creation of novel experimental materials, including entire gene libraries and specific cloned DNA fragments. A given key paper may represent one or all of these separate concepts in different source papers (*Cozzens'*[28] "split citation identity") and in separate citation contexts within the same source paper.

We identified and classed the major concepts represented by each key paper. These are listed in *Appendix 2*. We then examined the content of citation occurrences in each "early" and "late" source paper and coded the *first* occurrence of each different concept class. (In this analysis we were not interested in the *number* of times a particular key paper concept was recognized, but simply that it *was* recognized.) By aggregating concept tallies across source papers, we could identify the concepts that a key paper initially represented and any shifts or increased uniformity in citation content.

Replicability of classification

Inter-classifier consistency is a problem when applying any classification scheme. It can be a particular problem in citation context and content analyses which deal with difficult, often complex, and highly specialized subject literatures. Many such studies include one author with particular subject expertise (*Moravcsik*, for example, was trained as a physicist). Both *Chubin* and *Moitra*[31] and *Cozzens*[50] note the need for a knowledgeable subject expert in these analyses and the limitations thus placed on the research. There is little information available concerning the replicability of particular classification schemes (other than overlapping relevance judgments, not considered here). Published discussions of inter-classifier consistency deal with social sciences data and *context* rather than *content* classification. *Peritz*[29] submitted part of her demographic citations to "an expert in demography" but reports no quantitative measure of correspondence in classification. *Bonzi*[42] found 73% total agreement between her coding of library and information science citations and that of "another person". Most of the differences occurred between two adjacent (ordinal) category classes.

In the present study, the first author *(McCain)* classified and coded all citation occurrences for citation context and citation content. To replicate the citation *context* classification, the second author *(Turner)* received a detailed classification schedule (expanded from Table 2) and independently coded citation contexts in 48 randomly selected source articles (82 separate citation occurrences). *Turner* was generally familiar with biomedical terminology and writing styles but relatively unfamiliar with the particular subject matter of the key papers and source papers. Our coding matched (location and central/peripheral categories) in 88% of the cases. This suggests that the *context classification* scheme is quite replicable for empirical research reports and review articles. We did not attempt to systematically replicate the citation content analysis (Table 6; *Appendix 2*). This classification, based on *semantic content* of citation context, required more familiarity with the subject matter and the specific terminology than the second author possessed.

Results

Key paper citation patterns

The annual citation counts for each key paper and matching data for *Aversa*'s two Cluster groups[51] are shown in Table 3. Key papers are grouped by peak citation year and, within group, in decreasing order by total citations received. Data include counts for publishing year 8 (1987 for key papers published in 1980) as of the March–April 1988 issue of *Science Citation index*. (We include these data for a better comparison with *Aversa*, who used 9 years publication data in her original study. Year 8 source papers were not used in the context analysis.) The proportion of total citations received in each year is shown in parentheses; the overall citation patterns based on these proportion data are shown as plotted profiles in Appendix 1.

Aversa[23] calculated an aging rate for each Cluster as a moving average over both the last 4 years' raw citation data and proportional citations (publishing years 6–9). In Table 3 we provide two similar measures for our 8 years' worth of citation counts and comparable data from *Aversa*. Both are summary measures related to curve shape. The "Drop-off Rate" is similar to *Aversa*'s aging rate but uses citation data for years 5–8. The larger the value, the more *slowly* the citation counts decrease, and the less rapidly the key paper appears to "age". The "Peak Rate" is a moving average over the *first four years*' data. Here, larger values reflect an increasing cumulation of citations over the four year period while smaller values indicate either a slower rate of cumulation or (as in these data) a drop in citations in PY 3 and/or 4.

The data show two distinct groups of key papers, each with characteristics similar to the average paper in the corresponding Aversa Cluster. The annual citation counts of four Cluster 1 papers (*Maniatis, Lawn, Slightom, Sargent*) peak in year 6 or later, while the seven Cluster 2 papers (*Fritsch, Davis, Dodgson, Early, Tucker, Kemp, Wahli*) peak in year 2, with fractionally different counts in year 3 in several cases.

Table 3
Eight years citation data for eleven key papers in molecular genetics

Name	Peak Rate *	YR 1	YR 2	YR 3	YR 4	YR 5	YR 6	YR 7	YR 8	TOTAL	Dropoff Rate **
Maniatis	2.0	1 (<.01)	47 (.04)	128 (.12)	177 (.17)	158 (.15)	193 (.18)	176 (.17)	171 (.16)	1051	1.02
Lawn	1.5	0 (0.0)	48 (.08)	93 (.16)	70 (.12)	96 (.16)	80 (.14)	101 (.17)	103 (.17)	591	1.03
Slightom	1.45	14 (.03)	55 (.10)	75 (.14)	79 (..14)	94 (.17)	97 (.18)	80 (.15)	53 (.10)	547	0.85
Sargent	1.64	1 (.01)	23 (.12)	18 (.09)	28 (.15)	28 (.15)	34 (.18)	36 (.19)	24 (.12)	192	0.96
Aversa #1 @	1.57	2 (.02)	11 (.08)	17 (.13)	19 (.14)	20 (.15)	22 (.17)	21 (.16)	19 (.14)	131	0.98
Fritsch	1.02	35 (.10)	79 (.22)	77 (.22)	40 (.11)	48 (.14)	35 (.10)	23 (.07)	15 (.06)	352	0.69
Davis	.94	40 (.16)	56 (.23)	55 (.23)	31 (.13)	26 (.11)	15 (.06)	13 (.05)	8 (.03)	244	0.67
Dodgson	1.22	4 (.02)	38 (.21)	34 (.19)	21 (.12)	27 (.15)	16 (.09)	20 (.11)	19 (.11)	179	0.87
Early	.95	21 (.13)	50 (.32)	40 (.26)	16 (.10)	13 (.08)	4 (.03)	7 (.04)	5 (.03)	156	0.67
Tucker	1.19	1 (.01)	27 (.33)	25 (.28)	11 (.12)	11 (.12)	8 (.09)	0 (0.0)	5 (.06)	88	0.68
Kemp	1.31	0 (0.0)	22 (.34)	20 (.31)	13 (.20)	7 (.11)	1 (.02)	1 (.02)	0 (0.0)	64	0.44
Wahli	1.21	5 (.09)	11 (.21)	8 (.15)	10 (.19)	9 (.17)	1 (.02)	5 (.09)	4 (.08)	53	0.67
Aversa #2 @	1.29	4 (.04)	16 (.15)	21 (.20)	16 (.15)	15 (.14)	14 (.13)	10 (.10)	8 (.08)	104	0.82

* PEAK RATE calculated as moving average YR2 + YR3 + YR4 / YR1 + YR2 + YR3

** DROPOFF RATE calculated as moving average YR6 + YR7 + YR8 / YR5 + YR6 + YR7

@ Data from Aversa (ref. 23, p. 64). First 8 years citation counts used

The distinct and non-overlapping peak rate values reinforce this separation into two groups, based on the key papers' early performance. The drop-off rate ranges overlap, however, with *Slightom* (Cluster 1) aging more rapidly and *Dodgson* (Cluster 2) aging less rapidly than the others in their respective groups.

In summary, the data suggest that our two groups of key papers are reasonable representatives of the corresponding Aversa Clusters and thus are a useful test bed for studying internal citation-related variables which may be associated with differential aging patterns. In the remainder of this paper, we explore the differences in citation context and citation content for the 4 Cluster 1 key papers and the 7 Cluster 2 key papers.

Source paper citation context analysis

Aggregate context classification. We tallied all citation occurrences for each key paper in 10 yearly and 10 late source papers and classified the citation context of each occurrence. Table 4 shows the number of journal articles (J) and reviews (R)

Table 4
Characteristics of source papers

	Early Source Papers					**Late Source Papers**				
	J	**R**	**IC**	**SC**	**Mean * Occ/Paper**	**J**	**R**	**IC**	**SC**	**Mean * Occ/Paper**
Davis	7	3	0	2	2.1	7	3	0	0	1.4
Dodgson	9	1	1	1	1.3	9	1	0	0	1.3
Early	8	2	0	2	1.3	9	1	0	1	1.2
Fritsch	8	2	0	3	1.2	9	1	0	0	1.7
Kemp	10	0	1	5	1.5	5	5	0	1	1.4
Lawn	8	2	0	1	1.9	9	1	1	0	1.5
Maniatis	7	3	1	2	1.4	10	0	1	0	1.7
Sargent	9	1	0	0	1.2	10	0	0	1	2.0
Slightom	8	2	2	0	1.1	10	1	0	1	1.1
Tucker	8	2	0	1	1.8	9	1	0	4	1.5
Wahli	7	3	0	4	1.6	10	0	0	3	1.6

* Single outliers--source articles with 5 or more citation occurrences--were eliminated in calculating mean occurrences/paper.

in each source paper set, the number of institutional (IC) and self-citing (SC) source papers, and the mean number of citation occurrences per source paper.[52]

Based on our classification of citation occurrences, we calculated an individual Utility Index value for each key paper/source paper pair and the Mean Utility Index (MUI) for each key paper across each corresponding set of 10 yearly and 10 late source papers. Recall that the Utility Index serves as a measure of overall "perceived usefulness" of a key paper. It is a weighted sum including fixed weights for both central/peripheral context of citation occurrence and affiliation of source paper author, and four different sets of weights for citation location. Using these location weights we calculated four different MUI for each key paper in each time period. These are shown in Table 5.

In looking at differences across key papers and between key paper groups, we were primarily interested in comparing the *relative* usefulness of key papers, rather than focusing on the calculated MUI values. Our working hypothesis was that Cluster 1 papers would rank higher than Cluster 2 papers on our aggregate measure of utility. Accordingly, we used the numerical values of each MUI to establish an ordinal ranking of the eleven key papers. Fig. 1 compares key paper rank orders in early and late source paper sets. Within each MUI column, the higher the ranking of

	MUI #1		MUI #2		MUI #3		MUI #4	
Rank	Early	Late	Early	Late	Early	Late	Early	Late
1	**MAN**	**SGT** (5-1)	**MAN**	**SGT** (5-1)	**LWN**	**SGT** (6-1)	**LWN**	**SGT**
2	**LWN**	**MAN**	**LWN**	**MAN**	**MAN**	**MAN**	TUC	**MAN**
3	DAV	**LWN**	TUC	**LWN**	TUC	**LWN**	EAR	FRI (9-3)
4	TUC	DOD	DOD	DOD	DOD	DOD	**MAN**	**LWN** (1-4)
5	**SGT**	WAH (8-5)	**SGT**	WAH (9-5)	DAV	FRI (9-5)	DOD	DAV
6	DOD	FRI (10-6)	DAV	FRI (10-6)	**SGT**	WAH (10-6)	DAV	KMP (11-6)
7	EAR	DAV (3-7)	EAR	DAV	EAR	DAV (4-7)	**SGT**	DOD
8	WAH	EAR	**SLI**	EAR	**SLI**	**SLI**	**SLI**	WAH
9	**SLI**	KMP	WAH	TUC (3-9)	FRI	EAR	FRI	EAR (3-9)
10	FRI	TUC (4-10)	FRI	**SLI**	WAH	KMP	WAH	**SLI**
11	KMP	**SLI**	KMP	KMP	KMP	TUC	KMP	TUC (2-11)

Fig. 1. Rank order shifts of key papers based on Mean Utility Index values. Group 1 papers are in bold-face

Table 5
Mean utility index values

Early Source Papers (Pub. Yr. 1-2)

KEY PAPER	MUI #1	MUI #2	MUI #3	MUI #4
Maniatis	1.515	2.2	0.792	0.899
Lawn	1.253	1.941	0.881	1.074
Davis	1.089	1.378	0.634	0.646
Tucker	0.996	1.529	0.761	0.989
Sargent	0.919	1.388	0.549	0.63
Dodgson	0.874	1.39	0.637	0.758
Early	0.723	1.059	0.491	0.979
Wahli	0.571	0.833	0.348	0.434
Slightom	0.523	0.887	0.444	0.523
Fritsch	0.495	0.781	0.385	0.484
Kemp	0.293	0.468	0.222	0.268

Late Source Papers (Pub. Yr. 6-7)

KEY PAPER	MUI #1	MUI #2	MUI #3	MUI #4
Sargent	1.764	2.715	0.951	0.951
Maniatis	1.728	2.557	0.864	0.898
Lawn	1.416	2.049	0.708	0.783
Dodgson	1.402	2.048	0.701	0.756
Wahli	1.261	1.936	0.674	0.674
Fritsch	1.197	1.746	0.694	0.838
Davis	1.028	1.496	0.624	0.78
Early	0.96	1.446	0.549	0.614
Kemp	0.77	1.088	0.544	0.77
Tucker	0.765	1.187	0.489	0.497
Slightom	0.623	1.135	0.553	0.553

Group 1 key papers are in bold-face. Key papers are listed as ranked by MUI #1

Spearman Rank Order Correlations

Late Source Papers (Upper half matrix)

		MUI #1	MUI #2	MUI #3	MUI #4
Early Source	MUI #1	***	.97	.94	.80
Papers	MUI #2	.93	***	.94	.72
(Lower Half	MUI #3	.92	.97	***	.81
Matrix)	MUI #4	.77	.84	.89	***

a given key paper, the greater the usefulness of that work to the 10 source papers, considered as a group. Key papers that shifted more than three places are indicated by early and late rank positions in parentheses.

The first three MUI give very similar key paper orderings, with rank order correlations (Table 5) above 0.9 for both early and late source papers. MUI #4, which gives both "background" and "direct use" locations equal weight, is less highly correlated with any of the other three.

Several aspects of the rank order displays of MUI #1–#3 are of interest:

- in early source papers (Publication year 1–2), the two "founding" key papers are at the top of the order – *Maniatis'* paper outlined the protocols for making gene libraries, and *Lawn* described a major gene library. Of the other two Cluster 1 papers, *Sargent* is ranked fifth or sixth and *Slightom* eighth or ninth. *Tucker, Davis* and *Dodgson* are the highest ranking Cluster 2 papers. In terms of overall perceived usefulness in the early years of their citation history, the two Cluster sets appear fairly well mixed.
- In later source papers (Publication year 6–7), *Sargent*'s perceived usefulness has apparently increased markedly. This Cluster 1 paper, reporting the creation of a laboratory rat gene library, ranks at the top, followed by the *Maniatis* and *Lawn* papers. *Wahli* and *Fritsch* have replaced *Tucker* and *Davis* as high ranking Cluster 2 papers, with *Dodgson* placed at the top of this group. *Slightom*, the remaining Cluster 1 paper, continues to occupy the lower third or quarter of the rankings.
- Several key papers show rank order changes of three or more positions. *Fritsch, Sargent, Wahli* improved their ranking in the later source papers. *Davis* and *Tucker* dropped. These papers have apparently changed markedly in terms of the context in which they are perceived and cited.

MUI #4 (which does not differentially weight any citation locations) produces noticeably different early and late rankings and rank order shifts for the eleven key papers. In the early rankings, Early is placed much higher and *Maniatis* lower. Although *Sargent* and *Fritsch* improve their rankings in the later period, *Lawn drops* from rank #1 to #4 and *Dodgson* moves down to rank 7. *Kemp*, consistently ranked eleventh in the earlier period, moves up to rank #6.

Citation content analysis. For each key paper, we identified the various "concepts" for which it was cited in the 10 early and 10 late source papers *(Appendix 2)* and tallied the *first* occurrence of each concept in each source paper. Table 6 shows the percentage uniformity of citation concept class in early and late source papers. Since an individual source paper author may cite a key paper for more than one of the major concepts it represents, concept class totals frequently add up to more than 10 for a given set of 10 source papers. The key papers are listed in MUI #3 rank order, and Cluster 1 papers are in bold-face. For emphasis, we have underlined late source paper percentage uniformity values totalling 70% or more for "informational" (Results, Theory) or "methodological" (Methods, Libraries) concept symbol groups.

The table contrasts concept class counts in early and late source papers, and we can distinguish different patterns of uniformity and change in concept symbol recognition:

- Certain papers maintained a specific "identity" early and late in their citation histories – in *Small*'s[48] terms, they have a high percentage uniformity over time. *Maniatis* is *the* "techniques" paper. It is consistently cited as a token for various experimental techniques involved with creating, amplifying and screening gene libraries. It is also cited, much less frequently, to document an author's use of a gene library created in *Maniatis'* laboratory, including *Lawn*'s library. By contrast, three other key papers are most frequently cited as "informational" rather than "methodological" concept symbols. *Tucker* is highly cited in both time periods for experimental results. *Slightom* is recognised first for theoretical implications and secondarily for experimental results; *Davis* shows a similar though weaker pattern of relatively high experimental/theoretical citations in early and late source papers.
- Other key papers show a clear shift in their citation identify over time. Concept class counts for *Lawn, Sargent, Wahli* and *Dodgson* change from frequent early recognition for experimental results to later use almost exclusively as tokens for specific gene libraries – from "informational" to "methodological" concept symbols. *Early* and *Fritsch* show a less marked shift in this direction. By contrast, *Kemp*'s mixed recognition in early source papers shifts to exclusive citation for specific experimental results in the later time period.

Discussion

We assigned highly-cited journal articles in molecular genetics to one of two Cluster groups, based on their citation histories. Our working hypothesis was that a paper's citation history would reflect the perceived usefulness of its information content in subsequent research; that later-peaking, slowly-aging Cluster 1 papers would generally be perceived as more "useful" than Cluster 2 papers. We hypothesized that this perception would be demonstrated not only in the *multiplicity* of citation occurrences within a citing paper, but also in the *context* of the citation and the *nature of the concept symbol* that the key paper represented to later researchers.

Table 6
Percent uniformity of citation content

	EARLY SOURCE PAPERS					LATE SOURCE PAPERS				
	RES.	THEORY	METH.	LIBR.	TOTAL*	RES.	THEORY	METH.	LIBR.	TOTAL*
SGT	60%	NA**	10%	30%	10	15%	NA	8%	77%	13
MAN	8%	NA	75%	17%	12	0%	NA	69%	31%	13
LWN	61%	NA	23%	15%	13	8%	NA	17%	75%	12
DOD	55%	NA	NA	45%	11	0%	NA	NA	100%	10
FRI	82%	NA	18%	0%	11	46%	NA	23%	31%	13
WAH	58%	8%	17%	17%	12	17%	8%	0%	75%	12
DAV	29%	43%	14%	14%	14	10%	60%	0%	30%	10
SLI	36%	54%	9%	0%	11	20%	70%	0%	10%	10
EAR	60%	13%	7%	20%	15	40%	0%	0%	60%	10
KMP	64%	NA	18%	18%	11	100%	NA	0%	0%	10
TUC	82%	9%	NA	9%	11	73%	18%	NA	9%	11

* Key paper may represent more than one concept symbol class in a given source paper

** Key paper has no citations falling in this concept class in either time period

Our Utility Index combined citation occurrence counts, citation location, and citation context in a single measure of "perceived usefulness". The three MUI rankings which distinguished between "direct use" and "background" citation contexts were partially successful in distinguishing Cluster 1 and Cluster 2 key papers. In Figure 1, we can see the none of the MUI appear to be good *predictors* of future use, but that MUI #1–#3 do sort out three of the four longer-lived late-peaking Cluster 1 from the early-peaking, quickly-aging Group 2 key papers *later* in their citation history:

- In the first two publication years, Cluster 1 and Cluster 2 papers are mixed in the utility rankings. With the exception of the two 1978 "founding" papers in this research area (*Maniatis, Lawn*) we cannot clearly separate the two groups based on the MUI.
- In Publication Years 6–7, however, papers with Aversa Cluster 1 characteristics tend to be ranked at the top of the MUI lists. *Sargent* has moved up several positions and joined the two other high utility Cluster 1 papers, *Maniatis* and *Lawn.*

Dodgson, a Cluster 2 paper with a Cluster 1 aging rate, now ranks just below these three, followed by either *Fritsch* or *Wahli.* The remaining highly-ranked Cluster 2 papers have moved down noticeably (*Tucker*) or slightly (*Davis*) in these later MUI rankings.

- In terms of our "perceived usefulness" measure, however, the *Slightom* paper is a distinct anomaly. It ranks the lowest of the four Cluster 1 papers in both time periods, although its overall citation frequency and citation pattern place it within Cluster 1 parameters. The *Slightom* paper also has the lowest mean rate of multiple citation in both time periods – 1.1 citations/source paper. It appears that the "staying power" of this paper and its Cluster 1 status are not accounted for by the context-related variables included in the Utility Index.

The citation *content* data in Table 6 suggest reasons for both the temporal shift in MUI rank of many key papers and the relative position of the *Slightom* paper. A key paper appears to maintain its high MUI ranking (*Maniatis, Lawn*), or move up in the rank order (*Sargent, Fritsch, Wahli*), by becoming or remaining a concept symbol for important research methods or experimental materials – even in MUI # 3, in which use of methods and discussion of findings are given equal weight. By contrast, a key paper that continues to be cited over time primarily as a concept symbol for specific results and/or theoretical implications experiences the opposite fate. It will drop in the MUI rankings (*Tucker, Davis*) or remain at a lower rank if its initial perceived utility was low (*Slightom, Kemp*).

In our Utility Index, a high proportion of self citation in the source paper set will also tend to lower a key paper's rank, all other things being equal. The low early rankings of *Kemp* and *Wahli,* and *Tucker's* drop in the later source paper rankings probably reflect the high proportion of self- and institutional-citing source papers (Table 4) as well as the number and context classes of the citations contained therein.

The results from this diachronous study reinforce previous research which has shown an association between a scientific paper's "importance" (measured by citation counts at some point in time) and its identity as a "methods" paper. The citation prominence of many techniques papers is well known. In document cocitation mapping studies, *Lowry*[53] and similar "hyper-cited" methods papers may have to be removed from the citation files before analysis can proceed.[54] In lists of "citation classics," based on total citation counts over various periods of years, methods-oriented papers are among the most highly cited[55,56]. In his citation context analysis of highly cited papers in chemistry, *Small*[47] noted the "operational or procedural character of most of the contexts" and *Peritz*[29] reported that, on the average, methods papers in sociology were cited more often than either theoretical or empirical papers from the same journal and year, even when the few very highly cited "outliers" were removed.

Garfield[55] has asserted that citation counts for particular papers are really a pragmatic measure of utility or scientific activity rather than an indicator of their intellectual "significance," or "importance". Highly cited works are those "found to be useful by a relatively large number of people, or in a relatively large number of experiments". Theoretical papers achieving "classic" status are those which inspire new experiments or tests of proposed hypotheses,[57] or which have the "potential for controversial discussion.[58] Our Utility Index appears to be a partial measure of long-term utility in *Garfield's* sense: The Cluster 2 key papers which moved noticeably lower in the "late" MUI rankings (*Tucker, Davis*) or consistently are ranked low (*Early, Kemp*) were highly cited for experimental results or theoretical insights apparently perceived as centrally important to many other researchers *only immediately subsequent* to their publication. (In the case of the *Kemp* paper, useful primarily to that research group.) Later in time, this work tended to be cited (and self-cited) either by a smaller group of interested researchers or in a more peripheral, allusory fashion in an introduction or review article in which the citation functions primarily as a simple acknowledgment of the research having been done.

By contrast, technical "information", whether bench methods or experimental materials, is apparently much less likely to be subject to incorporation by "compaction"[2] or "obliteration by incorporation"[58] in the subsequent literature. Molecular biology has been described by one participant[57] as being "dependent on methodological innovations to advance the frontiers of research". A method used in subsequent research is tagged with a citation to the original publication of that technique, while etiquette requires citation of the published record as well as a printed acknowledgement when a scientist's materials are used in later published research.[15] The Cluster 2 status of the relatively highly ranked *Dodgson* and *Wahli* papers may represent a relatively smaller research community using chicken and frog (as opposed to rat and human) gene libraries, resulting in a lower rate of "methods" citations.

The *Slightom* paper confirms that long-term recognition as a broadly useful "methods and materials" concept symbol (resulting in a high Utility Index ranking) is not the only path to Group 1 status. In both periods we found a high proportion of single source paper citations to Slightom as a "theoretical" concept symbol. These citations occurred almost exclusively in the discussion section of journal articles, and were approximately evenly split between central and peripheral expository context "weight". This citation pattern is consistent with an interpretation that very broadly relevant, fundamental theoretical contributions of certain papers may be represented by high single occurrence citation counts because it is only *necessary* to the author's argument to cite the concept symbol once.

These limited data partially support our hypothesis that highly-cited papers differing in citation profile (Aversa Cluster assignment) can also be distinguished by the way they are perceived and cited several years after their initial publication. Late-peaking, slowly-aging Cluster 1 papers are likely to be cited for important, widely-useful technical contributions (methods or materials) or for highly fruitful, broadly relevant theoretical insights. Early-peaking quickly-aging Cluster 2 papers are likely to be cited for experimental results, theoretical contributions with limited applications, or technical contributions useful to a limited research population. Except for their overall prominence, reflected in higher citation counts, Cluster 2 citation profiles are similar to those of their less frequently cited counterparts in the scientific literature. Both receive the most recognition (largest number of citations) soon after publication[60] and then slip more-or-less gradually out of sight as newer work or a summary review article replace them.[61]

We can speculate concerning the dynamic process(es) generating the Cluster 1-type citation profile. Various intellectual/historical interpretations have been offered to account for "lagging" citation histories of "classic" papers, including "delayed recognition",[62] "premature discovery"[63] and "gradual acceptance of theoretical claims" [during an Kuhnian revolution].[17] These specific accounts are consistent with a more general interpretation of journal articles as tokens of technical or intellectual innovations whose diffusion and adoption[64] (signalled by citation) is spread over various time periods and across adopter pools of fixed or changing size and character. *Goffman*[65] has used similar models of the epidemic process to describe the growth of scientific specialties and their associated literatures.

We suspect that the same diffusion process is reflected in the citation history of these and similar highly-cited papers. In the peaking and dropping off of annual citation counts,[66] we may be seeing evidence of very different patterns of adoption of intellectual and technical innovations among potential users of these published research results, techniques and materials. Our Cluster 1 papers appear to represent major technological or theoretical innovations, whose use increases over a comparatively long time interval before declining, compared to the Cluster 2 papers. In a larger data set examining citation context and content for papers with Cluster 1-type profiles, we would expect to discover differences in researchers' (citing authors') *awareness* of the existence, usefulness or applicability of techniques, materials, and theory – perhaps learning of their existence through different communication channels. In the case of methods and materials innovations, even after the researcher is notified of their availability, it may take time to develop the skills to use the new technologies. Early-adopting laboratories and individuals may have been "pre-adapted" to innovate, having all the requisite skills and equipment, while others needed to acquire grant funding, equipment, and trained personnel. We might also expect to find differences in research specialization and subject orientation of earlier versus later citer/adopters, as this information spreads across disciplinary boundaries or is applied in other contexts.

We anticipate that these more extensive diachronous citation context analyses (coupled with information gathered from citing authors) will enhance our understanding of the processes by which these intellectual and technical innovations contribute to the development of research specializations and the growth of scientific knowledge.

*

We thank Roger A. *McCain*, Department of Economics, Drexel University for assistance in developing the Utility Index. Elizabeth S. *Aversa* and Jerry S. *Kidd*, College of Library and Information Service, University of Maryland, provided useful comments on an earlier draft. This research was supported by the College of Information Studies and by a Faculty Development Mini-Grant from Drexel University.

Notes and references

1. J. ZIMAN, *Public Knowledge: The Social Dimension of Science*, NY, Cambridge University Press, 1968.
2. K. SUBRAMANYAM, *Scientific and Technical Resources*, NY, Marcel Dekker, 1981.
3. H. G. SMALL, Citation context analysis, In: B. J. DERVIN, M. J. VOIGHT (Eds) *Progress in Communication Sciences*, Norwood, NJ, ABLEX, Vol. 3, 1982, pp. 287–310.
4. B. CRONIN, *The Citation Process*, London, England, Taylor Graham, 1984.
5. J. R. RAVETZ, *Scientific Knowledge and its Social Problems*, Oxford, England, Clarendon Press, 1971.
6. W. O. HAGSTROM, *The Scientific Community*, NY, Basic Books, 1965.
7. G. N. GILBERT, Referencing as persuasion, *Social Studies of Science*, 7 (1977) 113.
8. M. WEINSTOCK, Citation indexes, In: *Encyclopedia of Library and Information Science*, NY, Marcel Dekker, Vol. 5, 1971, pp. 16–40.
9. H. G. SMALL, B. C. GRIFFITH, The structure of scientific literatures I: Identifying and graphing specialties, *Science Studies*, 4 (1974) 17.
10. E. GARFIELD, M. V. MALIN, H. SMALL, Citation data as science indicators, In: ELKANA *et al.* (Eds) *Toward a Metric of Science: The Advent of Science Indicators*, NY, John Wiley & Sons, 1978, pp. 179–207.
11. L. SMITH, Citation analysis, *Library Trends*, 23 (1981) 83.
12. M. LINE, Obsolescence and changes in the use of literature with time, *Journal of Documentation*, 30 (1974) 283.
13. J. VLACHY, Citation histories of scientific publications. The data sources, *Scientometrics* 7 (1985) 505.
14. E. R. STINSON, F. W. LANCASTER, Synchronous versus diachronous methods in the measurement of obsolescence by citation studies, *Journal of Information Science*, 13 (1987) 65.
15. K. W. McCAIN, The Communication of Research-Related Information in Genetics, paper presented at the annual meeting of the Society for The Social Studies of Science, Worcester, MA. 21 Nov 1987.
16. H. DOLMAN, H. BODEWITZ, Sedimentation of a scientific concept: The use of citation data, *Social Studies of Science*, 15 (1985) 507.
17. A. P. W. HODDER, C. BALOG, A citation study of significant papers in plate tectonics, *Journal of Information Science*, 9 (1985) 221.
18. A. FEIDLER, C. D. HURT, Stratospheric aerosols: The transfer of scientific information, *Library and Information Science Research*, 8 (1986) 243.

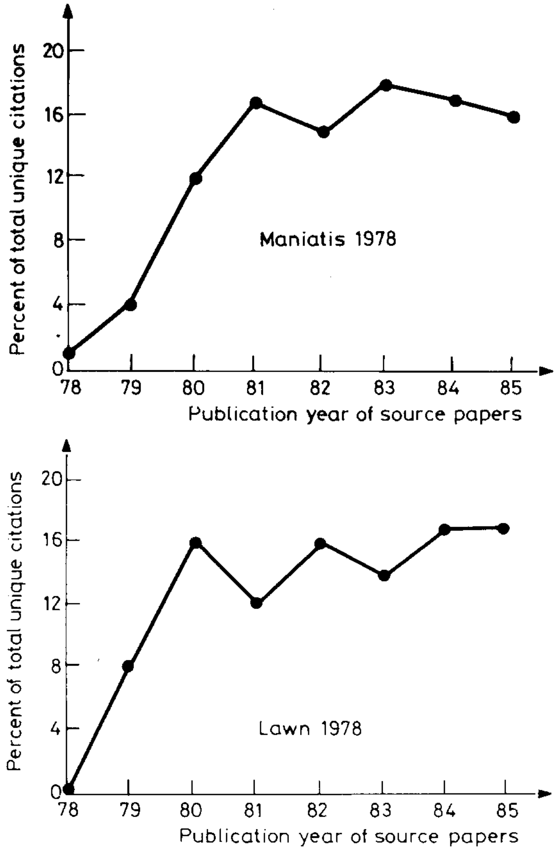
19. D. A. De STEPHANO,Citation analysis and adaptive radiation, *Scientometrics,* 11 (1987) 43.
20. S. COLE, The growth of scientific knowledge: Theories of deviance as a case study, In: *The Idea of Social Structure: Papers in Honor of Robert K. Merton,* NY, Harcourt, Brace, Jovanovich, 1975, pp. 175–220.
21. I. RUFF, Citation analysis of a scientific career: A case study, *Social Studies of Science,* 9 (1979) 81.
22. In his "Current Comments" essays, *GARFIELD* has occasionally examined the citation profiles of seminal papers written by Nobel laureates or their *ouevres.* See, for example, *Essays of an Information Scientist,* 4 (1981) 609; 6 (1983) 413; 8 (1986) 416, 432.
23. E. S. AVERSA, *Citation Patterns of 400 Scientific Papers and their Relationship to Literature Aging,* Unpublished dissertation, Philadelphia, Drexel University. 1984.
24. E. S. AVERSA, Citation patterns of highly cited papers and their relationship to literature aging: A study of the working literature, *Scientometrics,* 7 (1985) 383.
25. T. A. BROOKS, Private acts and public objects: An investigation of citer motives, *Journal of the American Society for Information Science,* 36 (1985) 223.
26. The full bibliographic description of this and other key papers used in the analysis will be found in Table 1.
27. A useful introduction to this and other genetic engineering techniques can be found in: R. W. OLD, S. B. PRIMROSE *Principles of Gene Manipulation: An Introduction to Genetic Engineering,* CA, University of California Press, 1981.
28. These types of errors have been the subject of citation analyses in their own right. See, for example, R. N. BROADUS, An investigation of the validity of bibliographic citations, *Journal of the American Society for Information Science,* 34 (1983) 132., and S. E. COZZENS, Split citation identity: A case study from economics, *Journal of the American Society for Information Science,* 33 (1982) 233.
29. B. C. PERITZ, A Classification of citation roles for the social sciences and related fields, *Scientometrics,* 5 (1983) 303.
30. M. J. MORAVCSIK, P. MURUGESAN, Some results on the function and quality of citations, *Social Studies of Science,* 5 (1975) 86. This was the first of a series of analyses in the physics literature.
31. D. E. CHUBIN, S. D. MOITRA, Content analysis of references: Adjunct or alternative to citation counting? *Social Studies of Science,* 5 (1975) 423.
32. C. OPPENHEIM, S. P. RENN, Highly cited old papers and the reasons why they continue to be cited, *Journal of the American Society for Information Science,* 29 (1978) 225.
33. I. SPIEGEL-ROSING, I. SCHWIDETSKY, Comparative bibliometric profiles of physical anthropology and human genetics, *Homo* 27 (1976) 31; I. SPIEGEL-ROSING, Science studies: Bibliometric and content analysis, *Social Studies of Science,* 7 (1977) 97.
34. T. HODGES, *Citation Indexing: Its Potential for Bibliographic Control.* Unpublished dissertation, University of California, Berkeley, 1978.
35. See, for example, the AIBS Style Manual, frequently cited in "instructions to authors" sections in Biology journals. CBE Style Manual Committee, *Council of Biology Editors Style Manual: A Guide to Authors, Editors, and Publishers in the Biological Sciences,* 4th. ed., American Institute of Biological Sciences, Arlington, VA, 1978.
36. R. A. DAY, *How to Write and Publish a Scientific Paper,* Philadelphia PA: ISI Press, 1979.
37. A. J. MEADOWS, The scientific paper as an archaeological artefact, *Journal of Information Science,* 11 (1987) 27.
38. H. VOOS, K. S. DAGAEV, Are all citations equal? or Did we *Op. Cit.* your *Idem? Journal of Academic Librarianship, 1 (1976) 19.*
39. S. J. K. BERTRAM, *The Relationship between Intra-document Citation Location and Citation Level,* Unpublished dissertation, IL, University of Illinois, 1970.

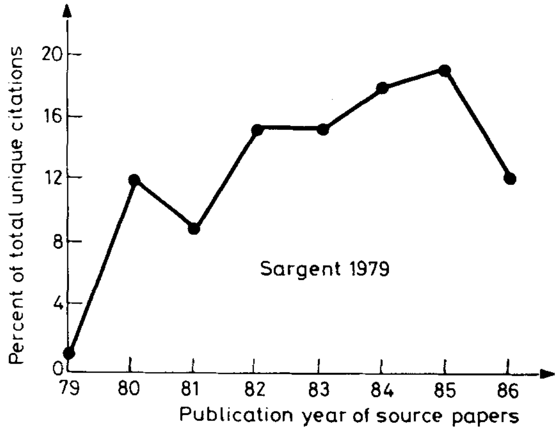
40. B. FINNEY, *The Reference Characteristics of Scientific Texts.* Unpublished dissertation, City University of London, Centre for Information Science, 1979. Cited in Ref. 8.
41. G. HERLACH, Can retrieval of information from citation indexes be simplified? *Journal of the American Society for Information Science,* 29 (1978) 308.
42. S. BONZI, Characteristics of a literature as predictors of relatedness between cited and citing works, *Journal of the American Society for Information Science,* 33 (1982) 208.
43. For information on index numbers, see entries on Index Numbers pp. 495–498; Statistical Techniques in Economics and Business, an Overview, pp. 886–893, In: *Encyclopedia of Economics,* NY: McGraw-Hill, 1982.
44. Please note that there is no intended relationship between our use of "utility" and the economist's concept of utility. For a discussion of the latter concept, see the entry on Utility pp. 934–941, in: *Encyclopedia of Economics,* NY, McGraw-Hill, 1982.
45. It seemed reasonable to assume that, while one citation is definitely better than none, two citations to a key paper do not necessarily seem to make the paper "twice as valuable. The third and subsequent citations add some additional value to the work being cited, but not proportionally to their number. Using a logarithmic transformation of our sums accomplishes this.
46. M. KOCHEN, Models of scientific output, In: Y. ELKANA *et al. (Eds), Toward a Metric of Science: The Advent of Science Indicators.* NY, John Wiley & Sons, 1978, pp. 97–136. See p. 127.
47. H. G. SMALL, Cited documents as concept symbols, *Social Studies of Science,* 8 (1978) 327.
48. H. SMALL, E. GREENLEE, Context analysis of a co-citation cluster: Recombinant-DNA, *Scientometrics,* 2 (1980) 277.
49. H. SMALL, The lives of a scientific paper, In: K. S. WARREN (Ed.), *Selectivity in Information Systems: Survival of the Fittest,* NY, Praeger Scientific, 1985, pp. 83–97.
50. S. E. COZZENS, Comparing the sciences: citation context analysis of papers from neuropharmacology and the sociology of science, *Social Studies of Science,* 15 (1985) 127.
51. See Ref. 23, p. 64.
52. Extreme outliers – typically a single self-citing paper with 5 or more citations to the key paper – were eliminated in calculating these mean values. Only one outlier at most was found for each key paper in a given time period.
53. E. GARFIELD, Citation frequency as a measure of research activity and performance, *Essays of an Information Scientist,* 1 (1977) 406.
54. B. C. GRIFFITH, H. SMALL, J. STONEHILL, S. DEY, The structure of scientific literature II: The macro- and micro-structure of science, *Science Studies,* 4 (1974) 339.
55. E. GARFIELD, *Citation Indexing: Its Theory and Application in Science, Technology, and Humanities,* NY, John Wiley & Sons, 1979.
56. E. GARFIELD, The articles most cited in 1961–1982. 2. Another 100 *Citation Classics* highlight the technology of science, *Essays of an Information Scientist,* 7 (1985) 218.
57. E. GARFIELD, The articles most cited in 1961–1982. 4. 100 Additional *Citation Classics, Essays of an Information Scientist,* 7 (1985) 306.
58. E. GARFIELD, The 'Obliteration Phenomenon' in science – and the advantage of being obliterated! *Essays of an Information Scientist,* 2 (1977) 396.
59. See quote from P. A. SHARP in ref. 55, p. 307.
60. E. GARFIELD, Uncitedness III – The importance of not being cited, *Essays of an Information Scientist,* 1 (1977) 413.
61. D. De SOLLA PRICE, Networks of scientific papers, *Science,* 149 (1965) 510.
62. See G. S. STENT, Prematurity and uniqueness in scientific discovery, *Scientific American,* 227 (1972) 84; H. V. WYATT, Knowledge and prematurity–journey from transformation to DNA, *Perspectives in Biology and Medicine,* 18 (1961) 596.

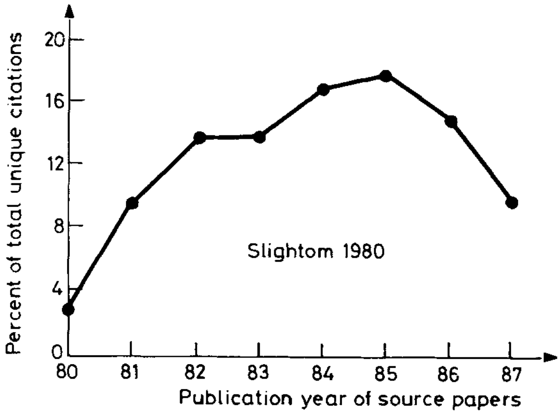
63. S. COLE, Professional standing and the reception of scientific discoveries, *American Journal of Sociology*, 76 (1970) 286.
64. E. M. ROGERS, *Diffusion of Innovations*, 3rd ed., NY, MacMillan, 1983.
65. See, for example, W. GOFFMAN, Mathematical approach to the spread of scientific ideas – The history of mast cell research, *Nature*, 205 (1966) 449.
66. If we plot the change in slope of the diffusion curve over time, we would get a curve fundamentally similar to a citation profile – first rising, then declining. There is one difficulty in drawing a parallel between "standafd" diffusion of innovation models and our citation histories. The diffusion data represent individual adoption decisions at specific points in time, while our data are generated by continued use/citation of adopted ideas. We also note that, in studying the spread of scientific ideas, the potential adopter pool is likely to grow over time, not only as new scientists enter the research arena, but as the innovations are seen as relevant in different contexts.

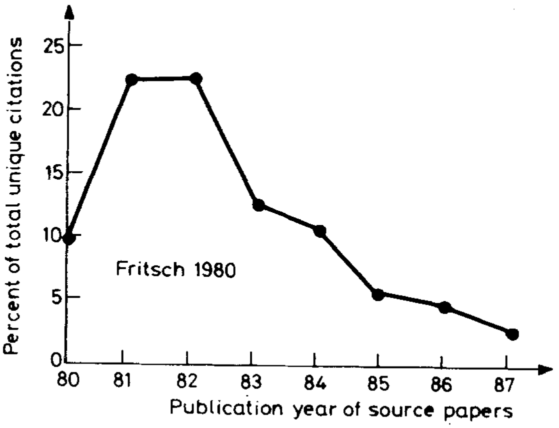
Appendix 1

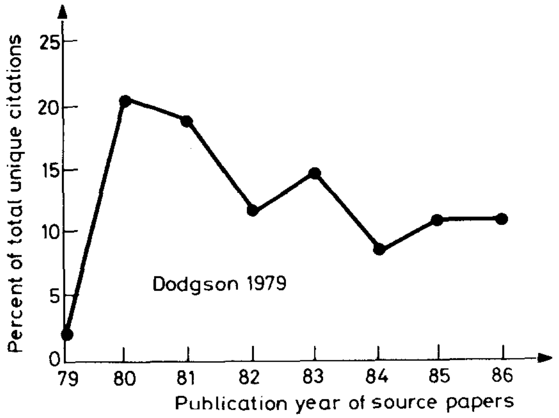
Citation profiles

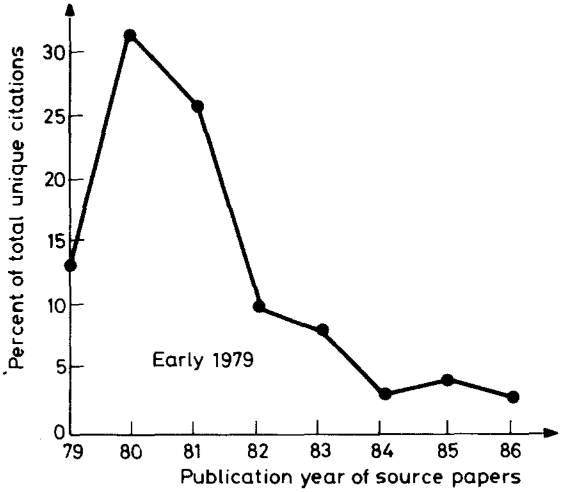


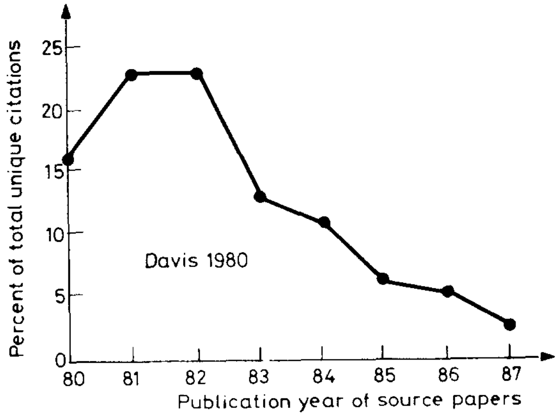

Percent of total unique citations
20
16
12
8
4
0
79
80
81
82
83
84
85
86
Publication year of source papers
Sargent 1979

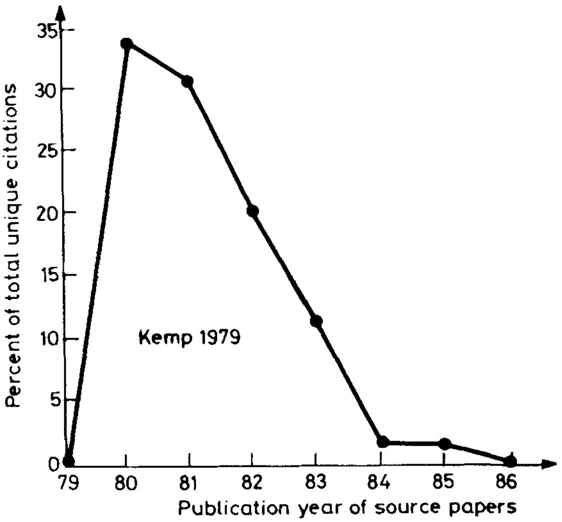

Percent of total unique citations
20
16
12
8
4
0
80
81
82
83
84
85
86
87
Publication year of source papers
Slightom 1980

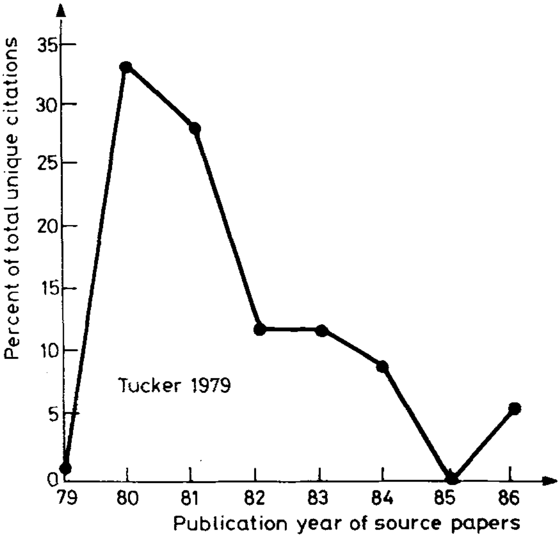

Percent of total unique citations
25
20
15
10
5
0
80
81
82
83
84
85
86
87
Publication year of source papers
Fritsch 1980

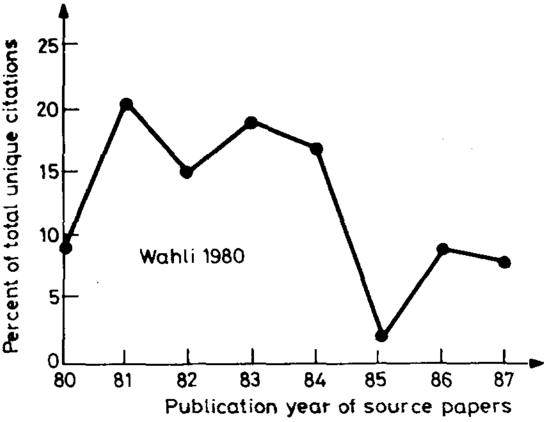

Percent of total unique citations
25
20
15
10
5
0
Dodgson 1979
79
80
81
82
83
84
85
86
Publication year of source papers

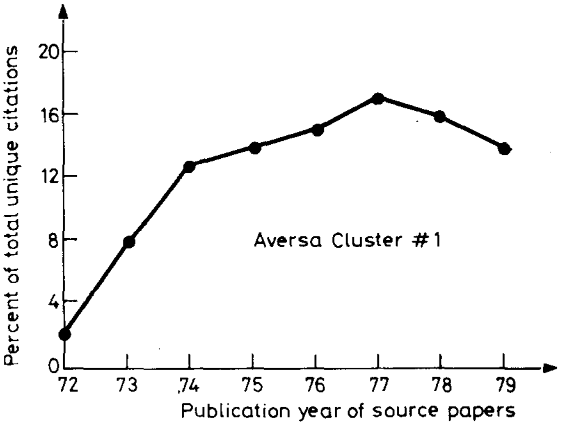

Percent of total unique citations
30
25
20
15
10
5
0
Early 1979
79
80
81
82
83
84
85
86
Publication year of source papers

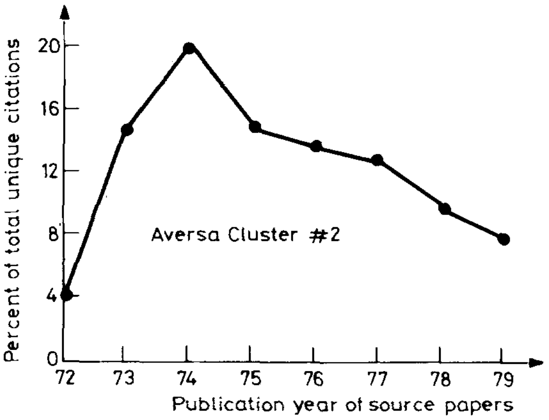

Percent of total unique citations
25
20
15
10
5
0
Davis 1980
80
81
82
83
84
85
86
87
Publication year of source papers


Percent of total unique citations
35
30
25
20
15
10
5
0
Kemp 1979
79
80
81
82
83
84
85
86
Publication year of source papers


Percent of total unique citations
35
30
25
20
15
10
5
0
Tucker 1979
79
80
81
82
83
84
85
86
Publication year of source papers


Percent of total unique citations
25
20
15
10
5
0
80
81
82
83
84
85
86
87
Publication year of source papers
Wahli 1980


Percent of total unique citations
20
16
12
8
4
0
72
73
74
75
76
77
78
79
Publication year of source papers
Aversa Cluster #1


Percent of total unique citations
20
16
12
8
4
0
72
73
74
75
76
77
78
79
Publication year of source papers
Aversa Cluster #2

Appendix 2

Key paper concept classes

Concept classes:

CC 1. Experimental results (RES)
CC 2. Broader theoretical implications (THEORY)
CC 3. Experimental methods (METH)
CC 4. Experimental materials (LIBR)

Key papers and cited concept classes

Davis, M.M. et al. 1980. "An immunoglobulin heavy-chain gene is formed by at least two recombinational events." Nature 283:733-739.

- CC 1: Organization of mouse heavy chain genes
- CC 2: Switch recombination as pathway of β-cell differentiation
- CC3: Library construction, screening techniques
- CC4: Library from BALB/c mouse sperm line DNA (new), BALB/c M603 mouse myeloma library (old), and derived clones

Dodgson, J.B., Strommer, J. & Engel, J.D. 1979. "Isolation of the chicken β-globin gene and a linked embryonic β-like globin gene from a chicken DNA recombinant library. Cell 17:879-887.

- CC 1: Structure and characterization of two chicken β-hemoglobin genes
- CC 2: None
- CC3: None
- CC4: Library from chicken DNA and derived clones

Early, P.W. et al. 1979. "Immunoglobulin heavy chain gene organization in mice: Analysis of a myeloma genomic clone containing variable and α constant regions." Proceedings of the National Academy of Sciences USA 76:857-861.

- CC 1: Organization of mouse α heavy chain genes (structural & functional elements separated by introns)
- CC 2: Mechanism of heavy chain V and C region rearrangements
- CC3: Library construction, screening techniques
- CC4: Library from BALB/c M603 mouse myeloma, and derived clones

Fritsch, E.F., Lawn, R.M. & Maniatis, T. 1980. "Molecular cloning and characterization of the human β-like globin gene cluster." Cell 19:959-972.

- CC 1: Structure and sequence of several globin polypeptides
- CC 2: None
- CC3: Library construction, screening techniques
- CC4: Library from β- thalassemia human liver cells, and derived clones

Kemp, D.J., Cory, S. & Adams, J.M. 1979. "Cloned pairs of variable region genes for immunoglobulin heavy chains isolated from a clone library of the entire mouse genome." Proceedings of the National Academy of Sciences USA 76:4627-4631.

CC 1: Characterization of paired V_H immunoglobulin genes
CC 2: None
CC3: Library construction, screening techniques
CC4: Library from mouse embryo DNA, and derived clones

Lawn, R.M. et al. 1978. "The isolation and characterization of linked δ-globin and β-globin genes from a cloned library of human DNA." Cell 15:1157-1174.

CC 1: Characterization of δ- and β-globin genes
CC 2: None
CC3: Library construction, screening techniques
CC4: Library from human fetal liver, and derived clones

Maniatis, T. et al 1978. "The isolation of structural genes from libraries of eucaryotic DNA." Cell 15:687-701.

CC 1: Structure of rabbit β-globin genes
CC 2: None
CC3: Library construction, screening techniques; number of restriction endonucleases
CC4: Library from *Drosophila*, rabbit, silkmoth, and derived clones; human fetal liver library (Lawn)

Sargent, T.D. et al. 1979. "The rat serum albumin gene: Analysis of cloned sequences." Proceedings of the National Academy of Sciences USA 76:3256-3260.

CC 1: Sequence and structure of rat serum albumin gene - has multiple intervening sequences
CC 2: None
CC3: Library construction, screening techniques
CC4: Library from laboratory rat, and derived clones

Slightom, J.L., Blechl, A.E. & Smithies, O. 1980. "Human fetal $^{G}\gamma$- and $^{A}\gamma$-globin genes: Complete nucleotide sequences suggest that DNA can be exchanged between these duplicated genes. Cell 21:627-638.

CC 1: Sequence and structure of $^{G}\gamma$ and $^{A}\gamma$- globin genes
CC 2: "Hot spot" for recombination - intergenic conversion as mechanism for coevolution of related duplicated genes.

CC3: Library construction, screening techniques
CC4: Library (shotgun collection) from human embryonic DNA, and derived clones

Tucker, P.W. et al. 1979. "Sequence of the cloned gene for the constant region of murine γ2b immunoglobulin heavy chain." Science 206:1303-1306.
CC 1: Sequence data for Cγ2b immunoglobulin heavy chain
CC 2: Evolutionary history of "hinge region"
CC3: None
CC4: Library (shotgun collection) from BALB/c mouse liver, and derived clones

Wahli, W. & Dawid, I.B. 1980. "Isolation of two closely related vitellogenin genes, including their flanking regions, from a *Xenopus laevis* gene library." Proceedings of the National Academy of Sciences USA 77:1437-1441.
CC 1: Sequence and characterization of two vitellogenin genes
CC 2: Evidence of "multigene families"
CC3: Library construction, screening techniques
CC4: Library from *Xenopus laevis*[clawed frog], and derived clones