Jérôme Pierrel

PhD, Associate Professor
in History and Philosophy of Science,
Université de Bordeaux
e-mail: jerome.pierrel@u-bordeaux.fr

УДК 577.2 (05)

Journal of Molecular Biology's growth and content analysis

Abstract: Price's law of exponential growth holds for the *Journal of Molecular Biology* (*JMB*), a journal created in 1959 and which was arguably at that time the most important of the discipline. In this paper, this growth is detailed according to countries, scientists, topics, at the early stage of the history of the journal, from 1959 to the 1970s, thanks to databases (Scopus, Pudmed) as well as manual analysis. The aim of the paper is to check whether in several subfields of molecular biology, Price's views on the relationship between *little science* (at the scientist level) and *big science* (from the statistical point of view) are fruitful. During this period, which by some scientists has been considered as the "golden age" of molecular biology, important scientific problems were solved (genetic code, DNA replication) and in the 1970s, their applications -genetic engineering technology- emerged. The growth of the journal was exponential, as expected by Price's law, but outpaced the prediction, with a shorter doubling-time than 15 years. Thus, *JMB* is a suitable case for a more detailed analysis of its scientific content and of the relationship between statistical growth and scientific trends at the bench.
Keywords: molecular biology, scientometry, Price's law, science growth.

The growth of the *Journal of Molecular Biology*

Derek J. de Solla Price's book, *Little science, big science*, pioneering statistical analysis of science, has been one a milestone for scientometry. Obviously, his approach is very different from the historian of science daily routine, who is working with archives and whose aims may be to understand scientists as human being. However, according to us, the most brilliant part of Price's book was not the chapter length analysis of the exponential growth of science *per se*, but the link he managed to build between the statistical data available about scientific articles and the daily habits of scientists. This link is especially relevant for historians of science since it means that statistical analysis of scientific journals may shed new lights on the scientists and their habits that historians are studying. Up to our knowledge, it has not been tested in the area of molecular biology, a new science born in the second part of the XXth century. This is the aim of this paper.

A quick growth (1959–1966)

Price's most important contribution to scientometry was its discovery of the exponential growth of science with a 15 years doubling time. He calls it "the fundamental law of any analysis of science" [*Price*, 1986, p. 4]. Indeed, this law has been found by Price to hold from the birth of the first scientific academies in Europe in the XVIIth century up to about the 1950s. In fact, this rate of growth seems to change around 1970 as shown by Han and colleagues for the United State (Han et al. 2010). Thus, we can take for granted that an exponential growth of molecular biology is expected for the period of time that we study in this paper, from the birth of molecular biology as an academic discipline in the 1950s, to the early 1970s, when practical tools like genetic engineering were produced from the basic research done by molecular biologists during this period.

To assess Price's law in the case of molecular biology, we have chosen to analyse the *Journal of Molecular Biology*. This scientific journal was created in 1959 by John Kendrew and was arguably at that time the most important of the discipline. Not only it was the first scientific journal to bear the name of molecular biology, even before laboratories with such a name were created in the early 1960s, but the structure of the editorial board shows that the journal was keen to include scientists from Europe, the United States, and even from the Soviet Union with soviet scientist Alexandre Spirin (USSR Academy of Sciences) as a board member as soon as the tenth volume (1964).

Our data show that the number of published pages in the *Journal of Molecular Biology* (not including tables of contents and index) grew exponentially from 1959 to 1966 with a doubling time of about 2 or 3 years, much more shorter than the expected doubling time (figure 1).

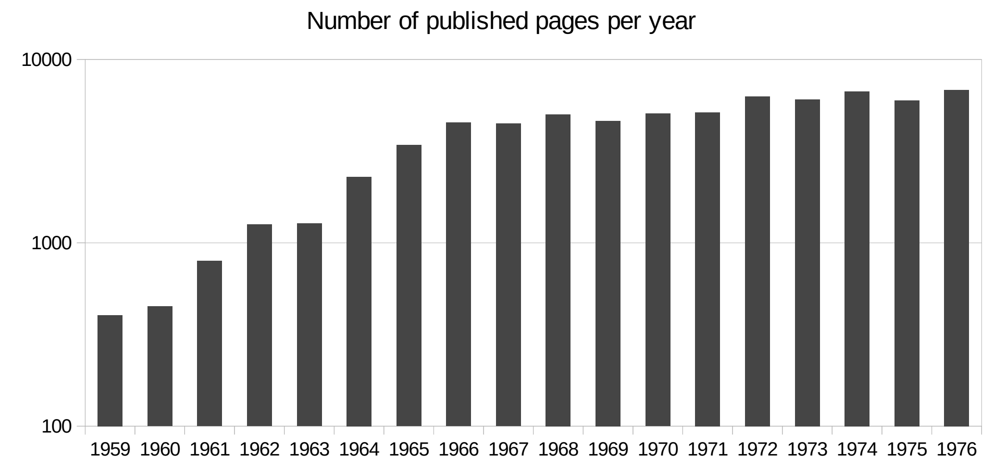

Figure 1. Number of published pages per year (not including index and tables of contents)

This confirms that molecular biology was blossoming. This fact holds also for the number of published papers in the journal and the data fit even more nicely an exponential growth. For both graphs, we rely on our own data (figure 2) since Scopus' coverage of the early years of this journal is incomplete.

The exponential growth of molecular biology was expected, but such a quick growth was not. Not only molecular biology grew faster than expected according to Price's law, but published data about molecular biology funding did not hint at such a growth. Data from historian Bruno Strasser show that the annual budget of four funding agencies which sup-

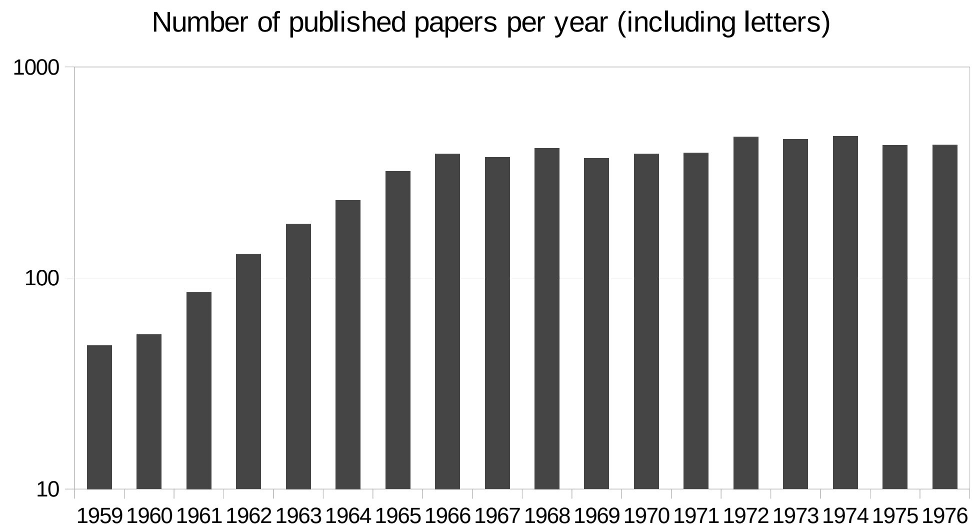

Figure 2. Number of published papers per year (including letters)

ported molecular biology (DFG in Germany, MRC in Great Britain, CNRS in France, FNRS in Switzerland) grew eight to tenfold between 1950 and 1970 [*Strasser*, 2002]. This is roughly the growth of the number of published papers in the *Journal of Molecular Biology* which grew eightfold between 1959 and 1970. However, according to Price, the cost of science grows according to the square of the number of scientists or of published papers. Thus, one would have expected that the cost of molecular biology would have a doubling time much shorter than the doubling time of molecular biology. To explain such a discrepancy, one might recall that molecular biology was only part of the budget of these agencies and this may hide its rapid growth.

How could molecular biology achieve such a high growth rate? One could mention three mechanisms that may explain this rate. According to the first one, molecular biology may rely, more than other sciences, on multiple authorship. Price understood that multiple authorship may have a particular significance:

(...) part of the social function of collaboration is that it is a method for squeezing papers out of the rather large population of people who have less than a whole paper in them [*Price*, 1986, p. 128].

Molecular biology seems to rely more on multiple authorship than chemistry (figure 3). Price original results were based on *Chemical abstracts* from 1900 to about 1960. His work shows that article with one author made more than 80 percent of *Chemical abstracts* in 1910 but less than 50 percent after 1950 with a steady decline. One can confirm such clear trends in the *Journal of Molecular Biology.* The most clear trend is the decline of the number of article written by a sole author since after 1966, they never made more than 20 percent of papers. Even before 1966, they only hit the 30 percent level once in 1963. What is relevant in our data is that clearly, most of the papers published by the *Journal of Molecular Biology* were written by several authors, more than expected according to Price's data, and this might explain part of the rapid growth of the journal.

The second mechanism which might explain such a high growth rate is that molecular biology was trendy and thus attracted not only new scientists for their PhD but also other scientists from established fields, like biochemistry and biophysics, which boosted its growth rate. This hypothesis needs further research to be quantitatively assessed. However, it fits nicely with the qualitative results of historians of molecular biology who showed that molecular biology in the United Kingdom, in France or in Switzerland grew very rapidly thanks to scientists from other fields, mostly biochemists, physicists, geneticists (de Chadarevian, 2002; Gaudillière, 2002; Morange, 2003; Strasser, 2006). This was the case of the famous 'Phage group' created by physicist turned geneticist Max Delbrück. Something similar happened into what Gunther Stent calls the 'structural school' of molecular biology, with among their members the famous physicists John Kendrew and Francis Crick (Stent, 1968).

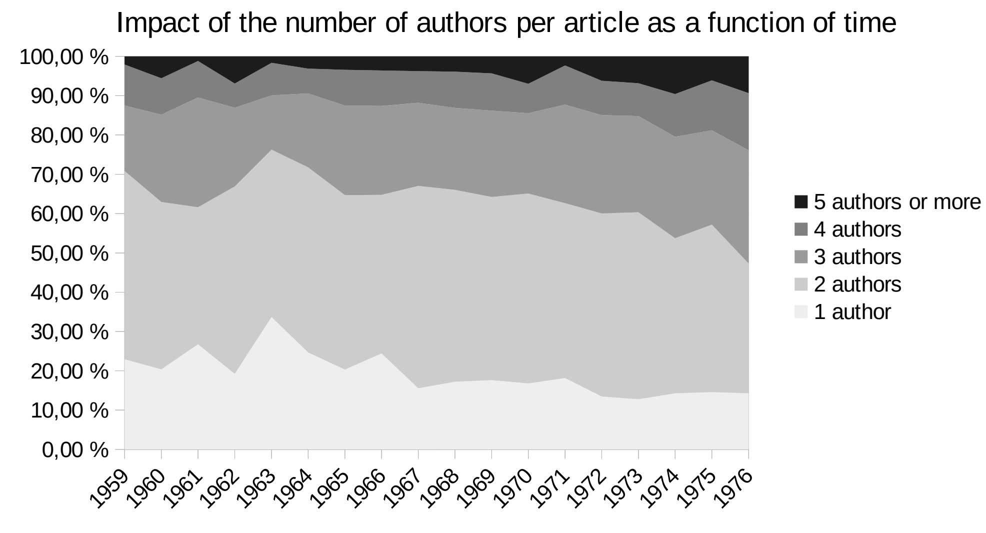

Figure 3. Impact of the number of authors per article as a function of time

The third mechanism is that one country with a very high growth rate contributes more than others to molecular biology. This is somehow confirmed by the following data. We have noticed that the growth of the budget of European agencies funding molecular biology was not enough to explain the growth of molecular biology. However, the NIH budget grew more quickly than these European agencies' budgets: from 1956 to 1963, it rose from 98 to 930 millions of dollars (Kay, 2000). Thus the growth rate of molecular biology funding in the US fits more closely with the observed growth rate of molecular biology than the growth rate of European funding does. This is confirmed by data according to country provided by the Scopus database. Figure 4 shows that the share of papers from the United States is the most important in the *Journal of Molecular Biology*, and this share is growing over the years until the early 1970s. In contrast to Price's data on *Chemical abstracts*, our data about the *Journal of Molecular Biology* do not show an important share for the Soviet Union. Indeed, most Soviet papers in this field were published in Soviet journals rather than into the *Journal of Molecular Biology.* The share of Germany is growing in the *Journal of Molecular Biology*

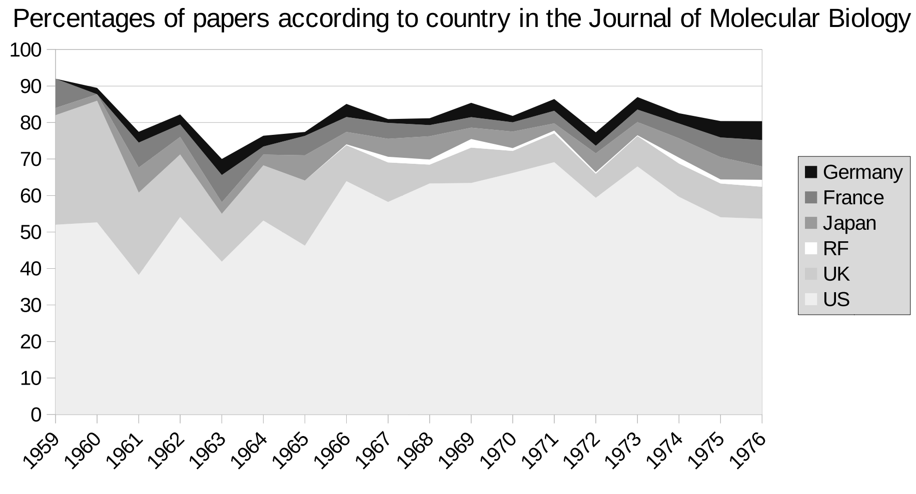

Figure 4. Percentages of papers according to country in the *Journal of Molecular Biology*. From bottom to top: United States, United Kingdom, Russian Federation, Japan, France, Germany. Data from Scopus database

whereas it was declining in *Chemical abstracts*. To conclude, the strong growth in US molecular biology seems indeed to contribute to the quick growth of *Journal of Molecular Biology* but nonetheless it is not the sole factor which explains such a rate of growth.

An unexpected break (1966–1976)

In a striking contrast with the early years of the journal, later, from 1966 to 1976, the number of published pages and papers seems to enjoy only a comparatively slow pace growth. This is an unexpected result, contradictory with Price's Law. One might even notice a small decline in published papers and pages in 1969 and 1975. Perhaps the latter was cause by the controversy about recombinant DNA with the Asilomar conference held in February 1975. A few months earlier, a moratorium about several genetic engineering experiments had been proposed in mid-1974 by Paul Berg and colleagues.

One can say for sure that this phenomenon, despite being temporary, lasted, according to Pubmed database, till the 1980s. If molecular biology had reached its saturation limit, as one might expect after such a quick exponential growth, its growth would not have resumed as it did later in the 1980s. One might also add that this phenomenon predates the observed break around 1970 by Han and colleagues (2010). Thus, this slowing down does not seem to be caused by the same mechanisms that altered the global science growth rate at the end of the 1960s and the beginning of the 1970s. What may be the cause of this event?

What strikes the historian of science is that the rapid growth of the years 1959–1966 fits with the period when the genetic code was solved. The discovery of the genetic code was, according to us, the greatest achievement of molecular biology. The genetic code has basic and applied consequences. On the theoretical side, its universality shows that life is underpinned by the same mechanisms. On the applied side, later in the 1970s, it enabled genetic engineering with its gene transfers from human to bacteria and the production of recombinant hormones like insulin. Gunther Stent, himself a molecular biologist, did not failed to notice that from 1953 to 1963, a period that he calls the "dogmatic phase", scientists elucidated both how DNA stores genetic information and how this information is translated into proteins (Stent, 1968). The latter is what the genetic code is all about. He noticed also that this "dogmatic phase" was especially attractive for scientists and by the end of 1960s:

(...) the number of working molecular geneticists had to be reckoned by the hundreds, rather than the dozens (...).

By contrast, after 1963, Stent felt that molecular biology entered a new phase, that he called the "academic phase" since "By that time many of the details of the genetic code were known" (Stent, 1968). He even add that "(...) what remained now was the need to iron out the details." Concerning the genetic code, Stent's opinion are true since a complete version of the code was obtained in the mid-1960s and the Nobel prize in medicine was awarded to Robert Holley, Marshall Nirenberg and Har Gobind Khorana in 1968 "for their interpretation of the genetic code and its function in protein synthesis". Historian of science and molecular biologist Michel Morange shares a similar view. He wrote in his textbook a chapter entitled "The years 1965–1972, or the crossing of the desert" (original title: "Les annés 1965–1972, ou la traversée du desert"). With the completion of the genetic code, Morange felt that, using a terminology proposed by Thomas S. Kuhn, molecular biology was then in a 'normal science' phase. This is exactly what Stent meant by saying that it remained "to iron out the details". While Morange went even as far as writing that few progress were made in molecular biology during this period, he nonetheless noted that progress were made, notably in technical areas like X-ray diffraction for protein structures and

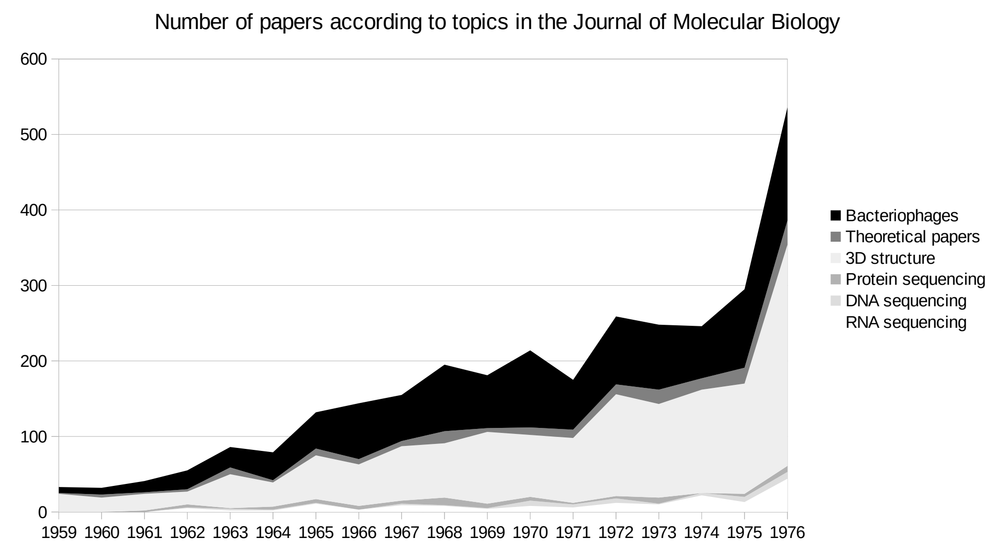

Figure 5. Number of papers according to topics in the *Journal of Molecular Biology*. Data from manual counting in title and abstract (for bacteriophages), materials and methods section (for papers about 3D structures and sequencing). Papers without a materials and methods section were counted as theoretical

nucleic acid sequencing, mostly RNA. Our data confirm this analysis (figure 5). The number of papers about three dimensional structures of molecules or about cell structures, using X-ray, NMR or electron microscopy, is on the rise as well as the number of papers about RNA and DNA sequencing (data obtained thanks to materials and methods). One can also notice that papers about the genetic code (as recognized in the index with the term 'code', 'codon', 'coding', or 'genetic code') were always very few among the published papers in the *Journal of Molecular Biology*. Their number was so small that they are not reported on the graph. Between 1959 and 1970, in the first fifty volumes of the journal, one can find only 17 such papers. The Pubmed database has only nine papers published from 1959 to 1976 in this journal that contain the expression "genetic code" in their title or abstract. This fact is intriguing because it shows that the idea that scientists were attracted towards molecular biology to solve the genetic code or to publish about it is wrong. However, it does not refute the idea that the new vocabulary used by molecular biology, which describes genetics and biochemistry in terms of code, program and information, was attractive and did attract many scientists to this field as established by historian Lily Kay [*Kay*, 2000].

Similarly, what we define as 'theoretical papers', that is to say, papers without a materials and methods section (letters to the journal do not have such a section and thus are not taken into account in our analysis), are scarce in the *Journal of Molecular Biology*. This nicely refutes the idea that molecular biology was a theoretical science. It did own a theory, consisting in the genetic code, the central dogma, etc, but its theory was only a tiny part of published papers.

Finally, papers about bacteriophages, recognized as such thanks to their title if they contain the word bacteriophage or a codename of a bacteriophage, were far more numerous. Thus, our data confirm the view of Gunther Stent, who cited the phage group and the structural school of molecular biology as important actors in this field [*Stent*, 1968].

We would like to stress that Stent and Morange, both molecular biologists who wrote about history of their own discipline, proposed timelines and views that are by and large confirmed by statistics. While history of science written by scientists has often been criticized with relevant arguments [see *Roger*, 1995], one might notice that Price's distinction between big science (that is to say, its statistical view on science) and little science (the individual scientist behaviour) does not discard history of science written by scientists but on the contrary, takes it seriously as shown by several excerpts from Price's work. Both views, statistical (from the top) and historical (from the grassroots so to speak), has to be confronted fruitfully.

In a footnote, Morange offers an interesting explanation about this stagnation period [*Morange*, 2003, p. 222]. According to him, the very quick growth of molecular biology prevented it from becoming an established discipline because it spread into others fields, "molecularizing it". More data on other scientific journals are required to test Morange's view. However, this opinion is at first sight consistent with the data introduced in this paper. Perhaps such a short doubling time of two to three years was not sustainable in the long run and would have led to a fierce competition with other biomedical fields.

Conclusion

Our data show that results from the statistical approach (Price) and from the historical approach (Morange, Stent) are consistent. The *Journal of Molecular Biology* underwent at

first a very strong growth (1959–1965) and then a stagnation period regarding the number of published papers (from 1966 to at least 1976). The idea supported by historians of science that the genetic code metaphor attracted scientists to molecular biology is confirmed by statistics. However, other mechanisms are necessary to explain the quick growth and are provided by statistics. These preliminary results show that historical approach and statistical approach complement each others.

Acknowledgments

Sergey Shalimov for comments, SPH for financial support.

References

De Chadarevian, S. (2002), *Designs for life,* Cambridge University Press, Cambridge.

Gaudillière, J.-P. (2002), *Inventer la biomédecine*, La Découverte, Paris.

Han, C.-S., Lee, S.K. and England, M. (2010), "Transition to postmodern science-related scientometric data" *Scientometrics*, vol. 84, pp. 391–401.

Kay, L. (2000), *Who wrote the book of life?* Stanford University Press, Stanford.

Morange, M. (2003), *Histoire de la biologie moléculaire*, 2nd ed., La Découverte, Paris.

Price, D. (1986), *Little science, big science... and beyond,* Columbia University Press, N.-Y.

Roger, J. (1995), *Pour une histoire des sciences à part entière*, Bibliothèque Albin Michel, Paris.

Stent, G.S. (1968), "That was the molecular biology that was", *Science*, vol. 160, pp. 390–395.

Strasser, B.J. (2002), "Institutionalizing molecular biology in post-war Europe: a comparative study", *Stud. Hist. Phil. Biol. & Biomed. Sci.,* vol. 33, pp. 533–564.

Strasser, B.J. (2006), *La fabrique d'une nouvelle science*, Leo S. Olschki Editore, Florence.