Scientometrics

Inventions shaping technological trajectories: do existing patent indicators provide a comprehensive picture?

Sam Arts,[a,b] Francesco Paolo Appio,[c] Bart Van Looy,[a,d,e]

[a]Managerial Economics, Strategy and Innovation
Faculty of Business and Economics, KU Leuven
[b]FWO, Brussels
[c]Scuola Superiore Sant'Anna
Istituto di Management
[d]Expertisecentrum O&O Monitoring (ECOOM) and Research Division INCENTIM, KU Leuven
[e]School of Management and Governance, University of Twente

Corresponding Author: F.Appio@sssup.it

Appendix 3: Illustrative examples of important inventions in the field of biotechnology and the resulting indicators:

Recombinant DNA and polymerase chain reaction

GENERAL DESCRIPTION	*Information*			
Event	The Cohen-Boyer technology for recombinant DNA is a path-breaking one by any standard, 'arguably the defining technology of modern molecular biology' (National Research Council, 1997, p.40). There are in fact three Boyer and Cohen patents: 4740470, 4468464, 4237224. All patents are continuations or continuations-in-part of patents originally filed in 1974, 1976, 1978. However, patent 4237224 is the dominant one of the three aforementioned, and has been widely adopted in the field of biotechnology.			
Patent#	US 4237224			
Title	Process for producing biologically functional molecular chimeras			
Inventor; Assignee	COHEN STANLEY N. and BOYER HERBERT W.; The Leland Stanford Junior University			
Filed	January 4, 1979			
INDICATORS (average values)	***Patent***	***Important contrib. (122)***	***Control group (83997)***	***Remarks***
Forward Citation Count	256	146.082	6.932	This invention received a high number of forward citations (256), above the average of the important contribution group (≈146).
Forward Citation Count 5y	94	36.230	2.813	This invention received a high number of forward citations within 5 years of its application (94), almost three times the average of the set of major contributions (≈36).
Forward Citation Lag	10.09	7.575	5.874	Besides serving more extensively as prior art for future generations of inventions, it also tends to remain cited for a longer time with a forward citation lag of 10.09 years, compared with the average 7.57 years of the set of the major technological inventions and the average 5.87 for the control group.
Generality	.70	.735	.510	This invention impacts on multiple different fields as shown by the high generality score (.70). Recombinant DNA has found widespread applications in biotechnology, medicine and research. Many additional practical applications of recombinant DNA are found in industry, food production, human and veterinary medicine, in agriculture, and in bioengineering: Recombinant chymosin, Recombinant human insulin, Recombinant human growth hormone (HGH, somatotropin), Recombinant blood clotting factor VIII, Recombinant hepatitis B vaccine, Diagnosis of infection with HIV, Golden rice, Herbicide-resistant crops, Insect-resistant crops, among many others.
Count Claims	14	23.336	15.421	This patent contains 14 claims, below the average of the set of major technological inventions (≈23).
Count main technology classes	3	2.475	2.195	It covers 3 main technology classes: 435 'CHEMISTRY: MOLECULAR BIOLOGY AND MICROBIOLOGY'; 530 'CHEMISTRY: NATURAL RESINS OR DERIVATIVES; PEPTIDES OR PROTEINS; LIGNINS OR REACTION PRODUCTS THEREOF'; and 536 'ORGANIC COMPOUNDS -- PART OF THE CLASS 532-570 SERIES'
Count technology subclasses	24	8.016	6.248	It covers 24 technology subclasses: 69.1 (Recombinant DNA technique included in method of making a protein or polypeptide); 183 (ENZYME (E.G., LIGASES (6.), ETC.); PROENZYMES, COMPOSITIONS THEREOF; PROCESS FOR PREPARING, ACTIVATING, INHIBITING, SEPARATING, OR PURIFYING ENZYMES); 207 (Acting on beta-galatose-glycoside bond (e.g., beta-galactosidase, etc.); 212 (Acting on

				peptide bond (e.g., thromboplastin, leucine amino-peptidase, etc. (3.4)); 231 (Acting on amide linkage in cyclic amides (e.g., penicillinase, etc. (3.5.2)); 252.33 (Escherichia (e.g., E. coli, etc.)); 320.1 (VECTOR, PER SE (E.G., PLASMID, HYBRID PLASMID, COSMID, VIRAL VECTOR, BACTERIOPHAGE VECTOR, ETC.)); 69.2 (Enzyme inhibitors or activators); 69.3 (Antigens); 69.4 (Hormones or fragments thereof); 69.5 (Lymphokines or monokines); 69.51 (Interferons); 69.52 (Interleukins); 69.6 (Blood proteins); 820 (SUBCELLULAR PARTS OF MICROORGANISMS); 849 (Escherichia coli); 91.1 (Polynucleotide (e.g., nucleic acid, oligonucleotide, etc.)); 91.4 (Modification or preparation of a recombinant DNA vector); 91.41 (By insertion or addition of one or more nucleotides); 311 (Somatostatin (SRIF); related peptides); 397 (Glycoprotein hormones); 399 (Hormones, e.g., prolactin, thymosin, growth factors, etc.); 808 (MATERIALS OR PRODUCTS RELATED TO GENETIC ENGINEERING OR HYBRID OR FUSED CELL TECHNOLOGY, E.G., HYBRIDOMA, MONOCLONAL PRODUCTS, ETC.); 23.1 (DNA or RNA fragments or modified forms thereof (e.g., genes, etc.).
First Subclass Combi Dummy	1	.656	.452	It is a patent with 263 novel pair wise combinations of technology subclasses, contributing to the trend that sees important contributions more likely to have a novel pair wise combination of technology subclasses (≈66%).
First Subclass Combi Count Re-Use	35,549	1,525.70	45.677	35,549 subsequent patents used the same component configuration which is considerably higher than the average observed for the set of important contributions (≈1,526).
Count Backward Citations	1	9.861	5.879	This invention has 1 backward citation: US 3813316 which is considerably less compared to the control group as well as the set of other major inventions
No Backward Citations	0	.172	.193	This patent does not belong to the 17.2% of patents with no backward citations.
Backward Citation Lag	6	6.063	7.506	This invention relies on more recent technical prior art with a backward citation lag of 6 years, in line with the average of its group (6.063).
Originality	0	.513	.516	This patent does not rely on prior art stemming from a broad range of technology fields as only one patent document is being cited.
Count Non-Patent References	23	43.836	22.439	23 non-patent references are cited, well below the average (≈44).
Dahlin and Behrens dissimilarity (before grant)	1	.549	.465	This patent's citation structure is dissimilar to the citation structures of past patents.
Dahlin and Behrens uniqueness (year grant)	1	.590	.668	This patent's citation structure is dissimilar to concurrent patents' citation Structures, contributing to the 59% of its group. As dissimilarity and uniqueness criteria have been fulfilled, this patent could be considered as an ex ante radical innovation.
Dahlin and Behrens adoption (after grant)	0	.361	.076	This patent's citation structure does not become replicated in the future.
Dahlin and Behrens composite	0	.139	.022	Further considerations on the composite criteria cannot be provided as only 2 criteria out of 3 have been fulfilled.

GENERAL DESCRIPTION	*Information*			
Event	Kary Mullis and colleagues at Cetus Corporation in Berkeley, California, invented a technique for multiplying DNA sequences in vitro: the polymerase chain reaction (PCR). PCR has been called the most revolutionary new technique in molecular biology in the 1980s. Cetus patented the process, and in the summer of 1991 sold the patent to Hoffman-La Roche, Inc. for $300 million. It took close to 4 years for specialists to appreciate the technology's potential, and longer still for a larger scientific community to begin practically exploiting its potential.			
Patent#	US 4683202			
Title	Process for amplifying nucleic acid sequences			
Inventor; Assignee	Kary B. Mullis; Cetus Corporation			
Filed	October 25, 1985			
INDICATORS (average values)	***Patent***	***Important contrib. (122)***	***Control group (83997)***	***Remarks***
Forward Citation Count	1,555	146.082	6.932	This invention received a high number of forward citations (1,555) well above the average (146) even of the subset of important contributions.
Forward Citation Count 5y	76	36.230	2.813	This invention received a high number of forward citations within 5 years of its application (76), well above the average (≈36).
Forward Citation Lag	12.11	7.575	5.874	Besides serving more extensively as prior art for future generations of inventions, it also tends to remain cited for a longer time with a forward citation lag of 12.11 years, high compared with the average 7.57 years of the set of the major technological inventions and the average 5.87 for the control group.
Generality	.76	.735	.510	A beyond the average generality score (.76) means an impact on multiple different fields. One need only reflect on the fact that PCR has found widespread applications in many areas of genetic analysis: medical applications (genetic testing, oncogenes, tissue typing), infectious disease applications (PCR tests for HIV), tuberculosis, disease organism), forensic applications (genetic fingerprinting, parental testing); plus a variety of research applications such as generation of hybridization probes for blotting, DNA sequencing, DNA cloning, sequence-tagged sites, phylogenic analysis of DNA from ancient sources, gene expression, and genetic mapping by studying chromosomal crossovers after meiosis.
Count Claims	21	23.336	15.421	This patent contains 21 claims, approaching the average of the set of major technological inventions (≈23).
Count main technology classes	2	2.475	2.195	It covers 2 main technology classes: 435 'CHEMISTRY: MOLECULAR BIOLOGY AND MICROBIOLOGY', and 536 'ORGANIC COMPOUNDS -- PART OF THE CLASS 532-570 SERIES'
Count technology subclasses	5	8.016	6.248	It covers 5 technology subclasses: 91.2; 317.1; 320.1; 23.1; 24.33
First Subclass Combi Dummy	1	.656	.452	It is a patent with novel pair wise combinations of technology subclasses: 435/91.2 & 435/317.1; 435/91.2 & 435/320.1; 536/23.1 & 435/91.2; 536/24.33 & 435/317.1; 536/24.33 & 435/91.2, contributing to the trend that sees important contributions more likely to have a novel pair wise combination of technology subclasses (≈66%).
First Subclass Combi Count Re-Use	4,021	1,525.70	45.677	4,021 subsequent patents used the same component configuration more than doubling the average of the set of important contributions (≈1,526).

Count Backward Citations	0	9.861	5.879	This invention has no backward citations.
No Backward Citations	1	.172	.193	This patent is one of the few (17.2%) which has not citations to technical prior art.
Backward Citation Lag	-	6.063	7.506	As this patent does not have any backward citation, this indicator cannot be calculated. There is no prior art being cited.
Originality	0	.513	.516	As this patent does not have any backward citation, this indicator being based on main technological classes of all cited patents cannot be calculated. There is no prior art.
Count Non-Patent References	5	43.836	22.439	5 non-patent references are cited, well below the average (≈44).
Dahlin and Behrens dissimilarity (before grant)	1	.549	.465	This patent shows a particular citation pattern from prior inventions as there are no backward citations – For such cases, we opted for a dissimilarity score of 1.
Dahlin and Behrens uniqueness (year grant)	1	.590	.668	This patent shows a particular citation pattern from prior inventions as there are no backward citations – For such cases, we opted for a uniqueness score of 1.. As dissimilarity and uniqueness criteria have been fulfilled, this patent could be considered as an ex ante radical innovation.
Dahlin and Behrens adoption (after grant)	0	.361	.076	As there is no citation structure, calculating an adoption indicator becomes questionable. We opted for a score of 0 in such cases.
Dahlin and Behrens composite	0	.139	.022	Further considerations on the composite criteria cannot be provided as only 2 criteria out of 3 have been fulfilled.