

Trends in the geographic distribution of human embryonic stem-cell research

Aaron Levine, M.Phil.

Doctoral Program in Science, Technology, and Environmental Policy
Woodrow Wilson School of Public and International Affairs
Princeton University
Princeton, NJ 08544
USA
adlevine@princeton.edu

ABSTRACT. Human embryonic stem-cell (hESC) research offers substantial potential benefits but has generated politically influential controversies and, in the United States, funding restrictions. Some observers fear the United States has been falling behind nations more permissive in this field, but policy debate has remained largely anecdotal. This study reports citation data indicating that the share of hESC research publications credited to the United States in the six years following the introduction of key technologies was significantly less than in five less contentious biomedical-research areas. The United States share of hESC publications fell sharply in 2003 and remained near this reduced level in 2004. Putative explanations are reviewed and several implications discussed.

Ever since James Thomson and colleagues in 1998 reported isolating them from blastocysts,¹ human embryonic stem cells (hESCs) have generated great wonder but intense controversy — and scant scientometric data.

Conditions suitable for hESC growth and differentiation have to considerable degree been elucidated^{2, 3} and methods for genetic modification described.⁴ In addition, a South Korean group has reported success in deriving hESCs through somatic-cell nuclear transfer⁵ and has refined this technique to derive new cell lines with reasonable efficiency from patients with a variety of illnesses,⁶ advances which may bring forward personalized therapies using immunologically matched stem cells. Meanwhile, in the United States, two bioethics commissions^{7, 8} and two National Academy of Sciences panels^{9, 10} have advised on hESC matters without lessening core tensions or resolving a key controversy: conscious — and in this case *conscientious* — restriction of research funding by a national government.

Federal funding is permitted for hESC research that uses a set of — by most counts — 22 cell lines derived

before 9 August 2001. Some critics have contended this policy is too restrictive and worry that the United States has been falling behind other countries in hESC research and eventually will trail in therapies. This concern has been based on the perception that several countries, such as the United Kingdom, Singapore and Israel, have fostered research environments more conducive to hESC research than the American. Anecdotal evidence for this concern includes the highly publicized departure of Roger Pedersen, a leading stem-cell researcher, from the University of California at San Francisco for the University of Cambridge, and comments from stem-cell researchers around the world.^{11, 12}

Anecdotes and comments notwithstanding, few data have been published on the actual geographic distribution of hESC research. I have assembled and in this paper report and analyze scientometric data “mapping the field” and testing the claim that the United States has been falling behind. To do so, I have taken advantage of the explicit link between two publications created by a citation, as described by Eugene Garfield, founder of the Institute for Scientific Information, half

a century ago¹³ and rendered useful for policy analysis more recently by advances in computational technologies.^{14, 15, 16}

Methods

Collection of data

I created datasets of research publications related to hESCs and to five other life-sciences technologies, and I then compared these datasets systematically. Articles were first identified through cited-reference searches using Web of Science from Thomson ISI. Web of Science was selected over other potential data sources, such as PubMed, because it allowed easy identification of all publications citing a particular paper. Searches were conducted 9 May 2005 using the Science Citation Index Expanded database, which indexed approximately 5,900 journals across 150 scientific disciplines.

Searches were restricted to document type “Articles,” a research-communication category that nominally — but not perfectly — excluded commentaries and reviews. Each dataset did include items other than research articles, but I assumed that the retrieval of these unsought items occurred at rates similar for each topic searched.

Article sets focusing on hESCs were compared with articles sets focusing on other technologies. The “other technologies” control set was developed through a broad screen of biomedical technologies that were initially demonstrated in the United States and were allowed to develop relatively free of regulation. Technologies were selected on the basis of their implications for basic research and on their potential for eventual therapeutic use. As citation patterns could vary by journal,¹⁷ technologies were limited to ones initially described in either *Science* or *Nature*. Five technologies were selected for this control set: DNA microarrays,¹⁸ polymerase chain reaction (PCR),¹⁹ yeast two-hybrid screening,²⁰ green fluorescent protein (GFP) expression tagging,²¹ and RNA interference.²² Just as for hESCs, datasets that contained all research articles citing the first publication were developed for each of these control technologies. As these technologies were introduced at different points in particular years and gained acceptance at different rates, year-one was in each case defined as the first calendar year in which an initial publication was cited at least twenty times. The United States share of other technologies was

calculated by averaging the share of research performed in the United States for each of the control technologies. This average was calculated for the first six years in the development of each technology, starting with year one, as defined above.

Statistical analysis

Confidence intervals were calculated for the percentage of hESC-related research performed in the United States by assuming the Bernoulli variable derived from the citation data was approximately normally distributed. Given this assumption, the 95-percent confidence interval could be calculated as follows:

$$95\% \text{ CI} = \pm 1.96 * \sqrt{\frac{p(1-p)}{n}}$$

where p is the fraction of articles from the United States in a given year and n is the total number of articles from that year. The confidence interval for the “other technologies” line was calculated as follows:

$$95\% \text{ CI} = \pm 1.96 * \frac{1}{5} \sqrt{\sum_{i=1}^5 \frac{p_i(1-p_i)}{n_i}}$$

where the subscript indicates the five individual technologies. P-values for the differences in means between the U.S. share of hESC-related research and the U.S. share of research related to the other biomedical technologies were calculated using the same assumptions and two-sided tests.

Results

To assess the general geographic distribution of hESC-related research, all papers citing James Thomson’s 1998 *Science* paper¹ were classified into five regions (Figure 1) on the basis of the address of the corresponding author. The share of articles from the United States was relatively constant at approximately 40 percent from 1999 to 2002, but it dropped to 30 percent in 2003 and remained at this lower level in 2004. The share of articles from Europe also fell from 34 percent in 1999 to 16 percent in 2004, with the bulk of this drop occurring between 1999 and 2000. In contrast to both the United States and Europe, the percentage of articles from Asia increased from 0 percent in 1999 to 30 percent in 2003, before declining to 22 percent in 2004. The other category grew from 9 percent in 2003 to 22 percent in 2004. This

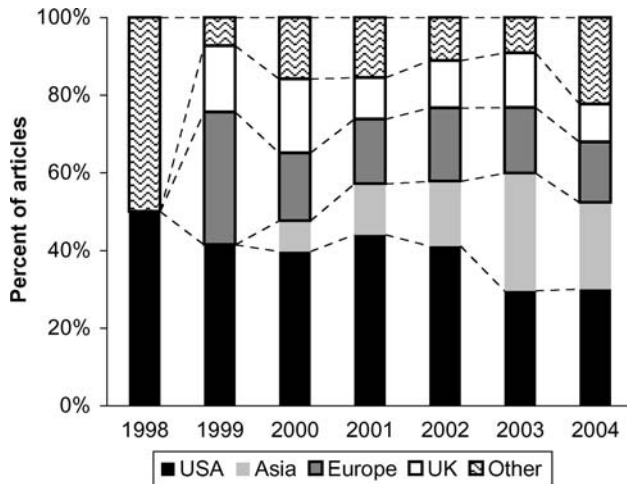


Figure 1. Geographic distribution of published research related to hESCs. Articles citing Thomson (1998) are classified based on the address of the corresponding author.

growth was driven largely by publications from Israel, Canada, and Australia.

To address how, if at all, the geographic distribution of hESC-related research differed from other less contentious and unregulated technologies, the analysis described above was repeated for the five other technologies. The U.S. share of published research related to hESCs was lower than the average U.S. share of research related to the others at comparable points in their developments (Figure 2). This difference was statistically significant across all six years (Table 1).

Discussion

Scholarly communications data have often been used to track and compare the performances of countries, institutions, and journals.¹⁶ Interpretation of such data can be difficult, however, and several factors may to some extent have confounded the citation-based scientometrics reported here. Data were used as an initial proxy for the development of the biomedical fields compared, but not all articles citing the first hESC paper could be taken as advances in the hESC field itself. Some were articles in related fields. Others were even further removed and represented researchers commenting on tangentially related discoveries. A few were policy stud-

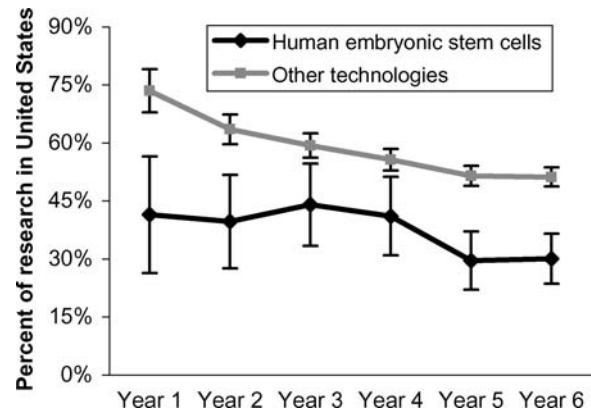


Figure 2. U.S. share of hESC-related research compared with the average of five other technologies. Year one was defined as the first calendar year in which the initial publication was cited at least twenty times. Vertical lines represent 95-percent confidence intervals.

ies or commentaries. Moreover, as technologies gain acceptance and are absorbed into everyday use, scientists, displaying what has been called the “obliteration phenomenon,” tend to stop citing initial papers.²³

More substantively, several factors may have contributed to the significant differences observed between the U.S. share of research related to hESCs and the U.S. share of research related to the other biomedical technologies.

First, hESC research may have diffused more rapidly to other countries because the seminal work was the product of an international collaboration in the first place. Although most of the authors on the first hESC publication were at the University of Wisconsin, Joseph Itskovitz-Eldor, from the Rambam Medical Center in Haifa, Israel, was a co-author.

Second, as the successful isolation of hESCs marked the endpoint of a technological race more than it did the development of a new and unexpected technology, other participants, who were already distributed around the world, would have been well positioned to pursue follow-on hESC research immediately. This research might, therefore, have occurred more quickly, and among a more dispersed group of scientists, than might otherwise have been expected.

Third, the controversial status of hESC research in the United States might have reduced the amount of research performed here and, thus, the contribution of

Trends in the geographic distribution of human embryonic stem-cell research

Table 1. Publications on six biomedical technologies, showing U.S. share and share difference, over time. Year one was defined as the first calendar year in which a new technology's initial publication was cited at least twenty times. The p-value indicates the likelihood that the observed results would be seen under the null hypothesis that the U.S. produced equal shares of research related to hESCs and research related to other selected technologies. (Differences were calculated prior to rounding.)

	Study year					
	1	2	3	4	5	6
	Yearly total publications on six biomedical technologies, all countries					
hESCs	41	63	84	90	142	193
DNA microarrays	22	57	82	128	175	231
PCR	87	614	1093	1592	1534	1429
Two-hybrid	62	115	195	310	294	268
GFP	80	173	234	275	277	236
RNAi	36	112	181	176	237	290
	Yearly total publications on six biomedical technologies, U.S. share, percent					
hESCs	41	40	44	41	30	30
DNA microarrays	82	65	68	65	65	62
PCR	69	59	57	48	42	41
Two-hybrid	84	73	62	60	57	57
GFP	66	61	59	52	44	47
RNAi	67	60	52	53	49	49
	Yearly total publications on five non-hESC biomedical technologies, average U.S. share, percent					
	74	64	59	56	52	51
	Yearly total U.S. share difference, five non-hESC biomedical technologies minus hESC technology (p-value)					
	32	24	15	15	22	21
	(<0.001)	(<0.001)	(0.007)	(0.007)	(<0.001)	(<0.001)

hESC = human embryonic stem cell

PCR = polymerase chain reaction

Two-hybrid = yeast two-hybrid screening

GFP = green fluorescent protein expression tagging

RNAi = RNA interference

U.S. scientists to the scientific commons. This reduction could have resulted from scientists leaving the country or through decisions to work in less controversial and better-funded areas. It may have been the case, for instance, that U.S. scientists focused disproportionately on adult stem-cell research, while scientists in other countries were more aggressively pursuing embryonic stem-cell research.

As no rules restrict privately funded hESC researchers in the United States, a related possibility is that the U.S. regulatory environment might have pushed hESC research into the private sector. Scientists at private firms typically have less incentive to publish than do scientists in academic posts and in some settings may even be prevented from publishing; thus, a shift to private-sector research could account for some decline in the U.S. share of research in this field. Furthermore, concerns over intellectual property-rights protections,

market conditions, and the regulatory climate may discourage private sector research from starting.²⁴

Finally, the U.S. share of hESC-related research may be lower than expected because other countries have invested heavily in the field in the hope of gaining economic benefits. Limited evidence for this theory exists, particularly in the form of statements suggesting that some Asian countries, such as China, Singapore, and Taiwan, view biotechnology in general and, in some cases, stem cells in particular, as a growth opportunity.^{12, 25, 26, 27}

These factors, as well as others, have probably contributed to the results reported here. These contributions have varied both in degree and in time, and the data available do not allow relative influences to be untangled.

In the United States, hESC research has remained controversial. At the national level, legislation has been

proposed that would loosen federal funding restrictions. States have also joined the debate. In November 2004, California voters approved a proposition that allocates \$3 billion to fund stem-cell research over ten years. Several other states, following into these unfamiliar policy waters, have also created or are considering their own funding plans.

While debate wears on, one question at least is settled by empirical test. In our new century's most closely watched race, the United States, long the global life-sciences hegemon, is indeed falling uncharacteristically behind.

The author thanks Lee Silver for guidance throughout the project and Harold Shapiro, Valerie Thomas, Mark Watson, the late David Bradford, David Wilcove, Shirley Tilghman, Jesse Rothstein, and Elana Broch and two anonymous reviewers for assistance and helpful suggestions.

Aaron Levine is a Ph.D. candidate in the Program in Science, Technology, and Environmental Policy at Princeton University's Woodrow Wilson School of Public and International Affairs, where his research focuses on assessing the impact of public policy on biomedical research. He completed his undergraduate studies in biology at the University of North Carolina at Chapel Hill and holds an M.Phil. degree from the University of Cambridge, where he studied computational biology at the Sanger Centre and developed algorithms to help analyze the human genome sequence. Prior to moving to the Woodrow Wilson School, he worked as a strategy consultant at Bain and Company and served as a Science & Technology Policy Graduate Fellow with the Board on Science, Technology, and Economic Policy at the National Academy of Sciences. He is supported by a National Science Foundation graduate research fellowship and by Princeton University.

References

1. J. A. Thomson *et al.*, "Embryonic stem cell lines derived from human blastocysts." *Science*, 1998, 282(5391): 1145–1147.
2. M. Schuldiner, O. Yanuka, J. Itskovitz-Eldor, D. A. Melton, and N. Benvenisty, "Effects of eight growth factors on the differentiation of cells derived from human embryonic stem cells." *Proc Natl Acad Sci USA*, 2000, 97(21): 11307–11312.
3. M. Richards, C. Y. Fong, W. K. Chan, P. C. Wong, and A. Bongso, "Human feeders support prolonged undifferentiated growth of human inner cell masses and embryonic stem cells." *Nat Biotechnol*, 2002, 20(9): 933–936.
4. T. P. Zwaka and J. A. Thomson, "Homologous recombination in human embryonic stem cells." *Nat Biotechnol*, 2003, 21(3): 319–321.
5. W. S. Hwang *et al.*, "Evidence of a pluripotent human embryonic stem cell line derived from a cloned blastocyst." *Science*, 2004, 303(5664): 1669–1674.
6. W. S. Hwang *et al.*, "Patient-specific embryonic stem cells derived from human SCNT blastocysts." *Science*, 2005, 308(5729): 1777–1783.
7. National Bioethics Advisory Commission, *Ethical Issues in Human Stem Cell Research: Volume I: Reports and Recommendations* (Rockville, MD, 1999).
8. President's Council on Bioethics, *Human Cloning and Human Dignity: An Ethical Inquiry* (Washington, D.C., 2002).
9. Committee on the Biological and Biomedical Applications of Stem Cell Research, *Stem Cells and the Future of Regenerative Medicine* (Washington, D.C.: National Academy Press, 2002).
10. Committee on Guidelines for Human Embryonic Stem Cell Research, *Guidelines for Human Embryonic Stem Cell Research* (Washington, DC: The National Academies Press, 2005).
11. G. Dyer, "Scientist cheered by UK stance on stem cells." *Financial Times*, August 30, 2002, p8.
12. D. Gershon, "Complex political, ethical and legal issues surround research on human embryonic stem cells." *Nature*, 2003, 422(6934): 928–929.
13. E. Garfield, "Citation indexes for science — New dimension in documentation through association of ideas." *Science*, 1955, 122(3159): 108–111.
14. F. Narin and K. S. Hamilton, "Bibliometric performance measures." *Scientometrics*, 1996, 36(3): 293–310.
15. D. J. Price, "Networks of scientific papers." *Science*, 1965, 149: 510–515.
16. C. L. Borgman and J. Furner, "Scholarly communication and bibliometrics," in *Annual Review of Information Science and Technology*, B. Cronin, ed., 2002, 36: 3–72.
17. P. O. Seglen, "The skewness of science." *Journal of the American Society for Information Science*, 1992, 43(9): 628–638.

Trends in the geographic distribution of human embryonic stem-cell research

18. M. Schena, D. Shalon, R. W. Davis, and P. O. Brown, "Quantitative monitoring of gene expression patterns with a complementary DNA microarray." *Science*, 1995, 270(5235): 467–470.
19. R. K. Saiki *et al.*, "Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase." *Science*, 1988, 239(4839): 487–491.
20. S. Fields and O. Song "A novel genetic system to detect protein-protein interactions." *Nature*, 1989, 340(6230): 245–246.
21. M. Chalfie, Y. Tu, G. Euskirchen, W. W. Ward and D. C. Prasher, "Green fluorescent protein as a marker for gene expression." *Science*, 1994, 263(5148): 802–805.
22. A. Fire *et al.*, "Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*." *Nature*, 1998, 391(6669): 806–81.
23. E. Garfield and A. Welljams-Dorof, "Citation data — Their use as quantitative indicators for science and technology evaluation and policy-making." *Science and Public Policy*, 1992, 19: 321–327.
24. C. B. Cohen. "Stem cell research in the US after the President's speech of August 2001." *Kennedy Institute of Ethics Journal*, 2004, 14(1): 97–114.
25. C. Dennis, "Stem cells rise in the East." *Nature*, 2002, 419(6905): 334–336.
26. K. Birmingham, "Singapore pushes biomedical research." *Nat Med*, 2001, 7(11): 1169–1170.
27. D. Cyranoski, "Taiwan: Biotech vision." *Nature*, 2003, 421(6923): 672–673.