

This is a report completed in time for the October 2000 [International Conference on Health Research for Development](#) in Bangkok. After the meeting, hyperlinks to the paper will be created on the project website. Until then, please do not cite, quote, or create hyperlinks to this document, or share the URL except with those involved in providing input to the International Conference. If you have comments or corrections, please contact us.

# World Survey of Funding for Genomics Research

## Final Report

to the Global Forum for Health Research and the World Health Organization

September 2000

Robert Cook-Deegan, Carmie Chan, and Amber Johnson  
Stanford-in-Washington Program  
2661 Connecticut Avenue, NW  
Washington, DC 20418  
[www.stanford.edu/class/siw198q/websites/genomics/  
bobcd@stanford.edu](http://www.stanford.edu/class/siw198q/websites/genomics/bobcd@stanford.edu)  
(01) 202-332-6235 phone; (01) 202-332-1416 fax

## Table of Contents

<i>Main Conclusions and Inferences from the Data</i> .....	3
<i>Results</i> .....	5
Government and Nonprofit Funding: Survey Results .....	5
Publicly Traded Genomics Firms: Analysis of Public Data.....	6
Aggregate data on “genomics” firms.....	6
R&D Figures from Publicly Traded Genomics Firms .....	7
Market Capitalization of the Largest Four Publicly Traded Genomics Firms .....	10
Patent Ownership.....	12
<i>Caveats and Qualifications</i> .....	13
What is Genomics? .....	13
What is a Genomics Firm? .....	15
Incomplete Data on Private Firms.....	16
Category Errors and Double Counting .....	17
<i>Genomics Outside the Developed Economies</i> .....	18
<i>Appendix I: Origins of the Survey</i> .....	21
<i>Appendix II: Methods</i> .....	23
<i>Appendix III: Cover Letter and Survey Form</i> .....	25

In May 2000 we initiated a survey of organizations that fund genomics research throughout the world. The project was funded by a grant from Burroughs Wellcome Fund to the Stanford-in-Washington program (Stanford University). The purpose was to do a one-time cross-sectional analysis of funding, and to couple that to an analysis of trends, based on analysis of publicly available data. The trends include data on private R&D funding, on patent ownership, and on market value of publicly traded firms, which give a glimpse of some underlying trends in the financial inputs and scientific outputs of genomics.

### **Main Conclusions and Inferences from the Data**

**The private sector** (pharmaceutical, biotechnology, and genomic startup firms) **is a bigger funder of genomics than the public sector** (government agencies and nonprofit organizations).

**The majority of genomics funding**, both public and private, **goes to performers in the United States.** All of the largest half dozen genomics startups are US firms. Seventy-six percent of publicly traded and 71 percent of privately held genomics firms on our list are US-based (see table). European and, to a lesser extent, Asian firms play a larger role among major pharmaceutical firms, but these remain almost exclusively in major developed economies, and much of the genomics even in foreign-owned firms is taking place in the United States.

**Ownership of patents and other intellectual property will be heavily concentrated in the United States**, and to a lesser extent other developed economies in Europe and Asia. The table in a footnote shows the patent holdings of those surveyed. It does not include some institutions with substantial numbers of patents in the DNA Patent Database. The most significant omissions are universities and nonprofit research institutes (such as Salk, Scripps, and Cold Spring Harbor Laboratories). Many universities have more DNA-based patents than major pharmaceutical firms.<sup>1</sup>

The significance of DNA-based patents is far from certain. To date, they have formed the basis for a few highly lucrative therapeutic proteins such as insulin, recombinant erythropoietin and other growth factors. Several of these patents have withstood court challenge, but the breadth and strength of DNA-based patents in general and gene-patents in particular, is still being defined within the legal regimes of individual countries. The US Patent and Trademark Office has generally been first to issue DNA-based patents, and its practices may not be followed to the same extent by the European Patent Office or other individual countries. Similarly, the U.S. Court of Appeals for the Federal Circuit has proven “patent friendly” since its establishment a decade ago a trend that may not occur abroad.

---

<sup>1</sup> The University of California, for example, has 399 patents in the database, Johns Hopkins 147, Harvard 140, MIT 127, Stanford 108, Washington University (St. Louis) 90, Caltech 61, and Yale 60; Salk holds 118 patents, Scripps 84 and Cold Spring Harbor 34. Certain biotechnology and pharmaceutical firms were not listed in our survey, because they did not have four or more listed genomic collaborations. Some of them nonetheless have substantial DNA-based patent holdings. Chiron (which acquired Cetus) holds 261 patents in the database, for example.

In terms of dollar flows, however, the United States is the largest pharmaceutical market, and company executives use a rule of thumb that 2/3 of profits derive from that market for most major therapeutic products; U.S. patent practices prevail in that most lucrative market.

Both first-mover advantages and intellectual property rights are mainly held in the United States, with the remainder concentrated in other developed economies, suggesting that future profits and resource flows will likewise concentrate there. The focus in genomics has been on creating valuable data, rather than a balanced distribution of benefits among the world's population. The sequence information is apt to consist in part of a pure public good, equally available for use by researchers throughout the world regardless of who paid for its creation. Extraction of value from the data, however, depends on substantial further research and development to realize useful products and services.

The primary value of genomics data for health research in the near future is as a tool for discovery, and the base of researchers who can use the data is overwhelmingly in the developed economies. Most value of genomics data, moreover, comes not from raw sequence information but instead depends on tacit knowledge that is difficult to replicate, from first-mover advantages, and from intellectual property—all of which will lodge with its creators. **Absent explicit attention at the international level, the initial technological fruits of genomics are likely to consist primarily of therapeutic and diagnostic applications for conditions affecting large populations in rich countries.** Even more than for biomedical research in general, the skew of research funding is heavily toward the developed economies with large pharmaceutical markets.

## Results

### Government and Nonprofit Funding: Survey Results<sup>2</sup>

#### Funding in \$US

(Listed in order of total funding for Year 2000)

	1998	1999	2000 (est.)
<b>National Human Genome Research Institute, NIH</b>	210,891,000	270,733,000	326,391,000
<b>Wellcome Trust</b>	60,256,410	100,742,942	115,777,195
<b>European Commission</b>	21,344,717	104,602,510	89,968,511
<b>US Department of Energy*</b>	85,500,000	89,800,000	88,900,000
<b>American Cancer Society</b>		50,000,000	50,000,000
<b>Knut and Alice Wallenberg Foundation</b>	5,000,000	11,000,000	35,000,000
<b>The SNP Consortium</b>		28,000,000	22,000,000
<b>Cancer Genome Anatomy, Mammalian Gene Collection, Genetic Annotation Initiative and related programs, National Cancer Institute (with cofunding from other NIH institutes)</b>	7,000,000	11,300,000	21,800,000
<b>Howard Hughes Medical Institute</b>	20,000,000	20,000,000	20,000,000
<b>Kazusa DNA Research Institute</b>	14,800,000	14,500,000	14,400,000
<b>Imperial Cancer Research Fund</b>			12,296,588
<b>Centre National de Sequencage Genoscope</b>	4,529,148	7,458,396	7,986,721
<b>Katholieke Universiteit Leuven</b>	5,000,000	5,100,000	5,200,000
<b>Fondation Jean Dausset-CEPH</b>	6,296,692	5,439,331	4,111,561
<b>Merck Genome Research Institute#</b>	3,700,000		
<b>National Institute of General Medical Sciences, NIH</b>	3,000,000	3,200,000	3,500,000
<b>Australian Genome Research Facility</b>	610,687	1,615,385	1,666,667
<b>Program in Medical Genomics, National Health and Medical Research Council (Australia)</b>	319,331	165,993	649,425
<b>Swedish Medical Research Council</b>			200,000
<b>Total</b>	<b>448,247,985</b>	<b>723,657,557</b>	<b>819,847,667</b>

\*Figures for US Department of Energy from the White House

#Figures for Merck Genome Research Institute from 1999 Corporate Philanthropy annual report

<sup>2</sup> The figures in the table report are from survey responses, except where noted. The figures were reported to us in different currencies. They have not been adjusted for inflation. For 1998 and 1999, foreign currencies were adjusted by the Purchasing Power Parity (PPP) figures from the Office of Economic Cooperation and Development, Paris (<http://www.oecd.org/>) as part of its Principal Economic Indicators series. A similar figure was not available for 2000, so figures were adjusted by currency exchange rates on August 31, 2000. The original figures will be posted on the website in the original spreadsheet, so others can use the PPP figure to recalculate once it is available.

## Publicly Traded Genomics Firms: Analysis of Public Data

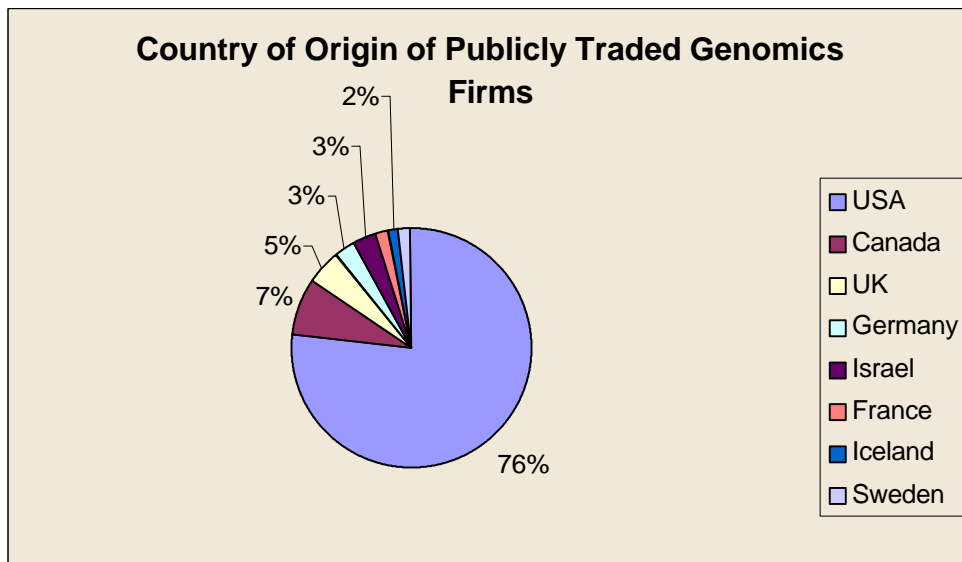
### Aggregate data on “genomics” firms<sup>3</sup>

161 firms total; 64 with publicly traded stock; 97 privately held

Firms on project website list on September 25, 2000

#### Genomics Firms with Publicly Traded Stock

Country of origin	Number of firms
USA	49
Canada	5
UK	3
Germany	2
Israel	2
France	1
Iceland	1
Sweden	1



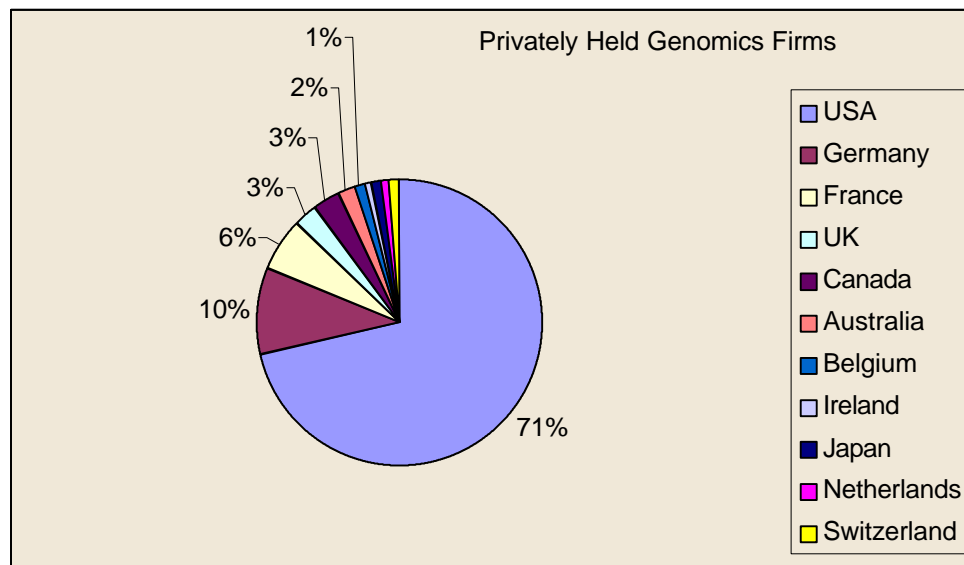
<sup>3</sup> These figures changed frequently while the project progressed. The week of completing the final report, for example (25 Sept 2000), four firms had their initial public offerings, requiring us to recalculate market capitalization and national distribution charts and tables. Please note: these firms are startup firms dedicated wholly or in part to genomics. Established biotechnology and pharmaceutical firms with known genomics collaborations are listed in a separate table on the website.

## Privately held firms

Country of origin	Number of firms
USA	68
Germany	10*
France	6
UK	3
Canada	3
Australia	2
Belgium	1**
Ireland	1
Japan	1
Netherlands	1**
Switzerland	1

\* One German/US firm (Atugen) is counted as German only.

\*\* One Belgian/Dutch (Galapagos) firm was counted as Belgian only.



## R&D Figures from Publicly Traded Genomics Firms

	(\$US millions)						
	1999	1998	1997	1996	1995	1994	1993
Total*	845.8	690.2	507.9	318.8	210.1	149.2	81.7
Big 4**	415.8	268.7	197.4	116.1	70.0	49.5	24.2

\* R&D figures reported to the Securities and Exchange Commission (or in annual reports) for Abgenix, Aclara, Affymetrix, Aurora Biosciences, Axys, Biacore, Corixa, CuraGen, Diversa, Gene Logic, Genome Therapeutics, Genomic Solutions, Genset, Hyseq, Invitrogen, Lexicon Genetics, Life Technologies, LJI Biosystems, Lynx, Magainin, Maxygen, Myriad Genetics, Pathogenesis, Protein Design Labs, and Sequenom Inc.

\*\* Celera, Human Genome Sciences, Incyte, and Millennium

Annual R&D spending reported to the US Securities and Exchange Commission by the four largest publicly traded firms that are primarily focused on genomics (Celera, Human Genome Sciences, Incyte, and Millennium) was over \$415 million in 1999, and will likely be higher in 2000. If we add R&D reported by another 25 publicly

traded firms dedicated solely or substantially to genomics,<sup>4</sup> we reach \$845 million in 1999. (Figures have not yet been reported for Year 2000.)

Spending on genomics in established biotechnology and pharmaceutical firms is likely of at least comparable magnitude. Many established biotechnology or pharmaceutical firms have substantial genomics investments.<sup>5</sup> Our data on genomics in these firms is too incomplete to report any quantitative conclusions, but the larger firms have market capitalization over \$150 billion. If the two firms reporting to us that 3 to 4 percent of their R&D was devoted to genomics are representative of industry norms (and we have no way of knowing), then genomics funding from PhRMA members alone would be in the range of \$800 million to \$1 billion.<sup>6</sup> We know of no way to estimate genomics research spending among 97 privately held startup firms that we have identified as having a substantial involvement in genomics.<sup>7</sup> Even leaving out the privately held firms, private spending is surely over \$1 billion and probably in the range of \$1.5 to \$2 billion. The publicly available data already allow an important conclusion: private annual spending for genomics is substantially higher than the public sector funding, probably in the range of twice the government and nonprofit spending.

Our data showing a predominant U.S. presence in patent holdings and ownership of genomics firms corroborate the conclusions of a report prepared for the European Commission by Sandra Thomas and Nicholas Simmonds.<sup>8</sup>

---

<sup>4</sup> Firms publicly traded in 1999 (therefore 1999 R&D spending reported): Abgenix, Aclara, Affymetrix, Aurora Biosciences, Axys, Biacore, Corixa, CuraGen, Diversa, Gene Logic, Genome Therapeutics, Genomic Solutions, Genset, Hyseq, Invitrogen, Lexicon Genetics, Life Technologies, LJI Biosystems, Lynx, Magainin, Maxygen, Myriad Genetics, Pathogenesis, Protein Design Labs, and Sequenom Inc. R&D figures in spreadsheet form available at:

<http://www.stanford.edu/class/siw198q/websites/genomics/pubgenomicsR&D.xls>.

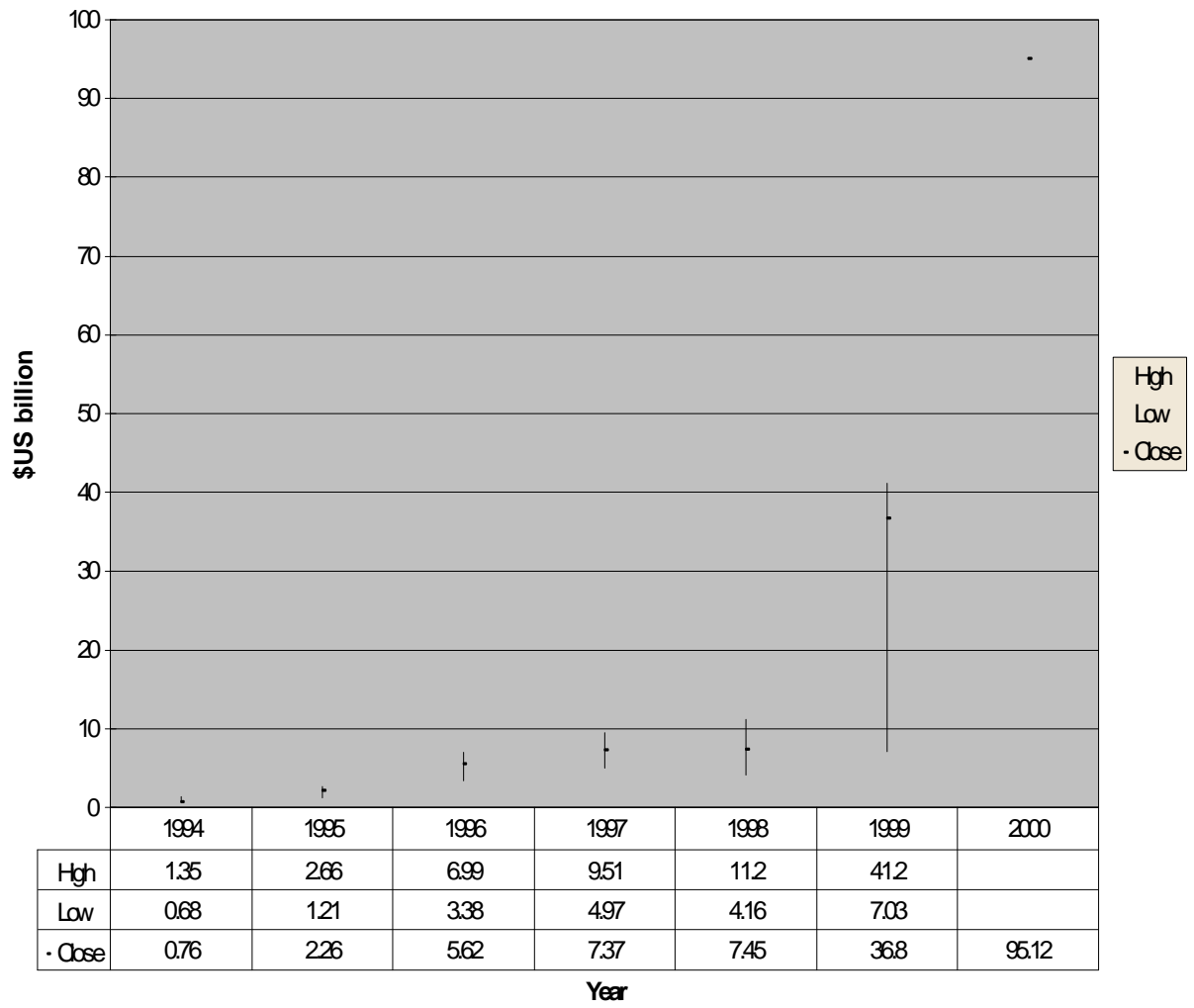
<sup>5</sup> All established biotechnology and pharmaceutical firms listed on the project website have at least four known collaborations with the genomics firms or with a major academic genomics center: Allergan, American Home Products (Incl. Genetics Institute and Wyeth-Ayerst), Amersham Pharmacia Biotech, Amgen, ArQule, Astra-Zeneca, Aventis, BASF, Bayer, Biogen, Boehringer Ingelheim, Bristol-Myers Squibb, DuPont/Dupont Merck, Genentech, Genzyme, Glaxo SmithKline, Hoffman-La Roche (Roche Holdings and others), J&J (Ortho, Janssen), Lilly, Merck, Novartis (incl. Institute of Functional Genomics), Novo-Nordisk, Otsuka, PE Corp and Applied Biosystems, Pfizer (including former Parke-Davis and Warner-Lambert), Pharmacia & Upjohn, Qiagen, and Schering-Plough (incl. Berlex). List available at: <http://www.stanford.edu/class/siw198q/websites/genomics/pharma-biotech.htm>.

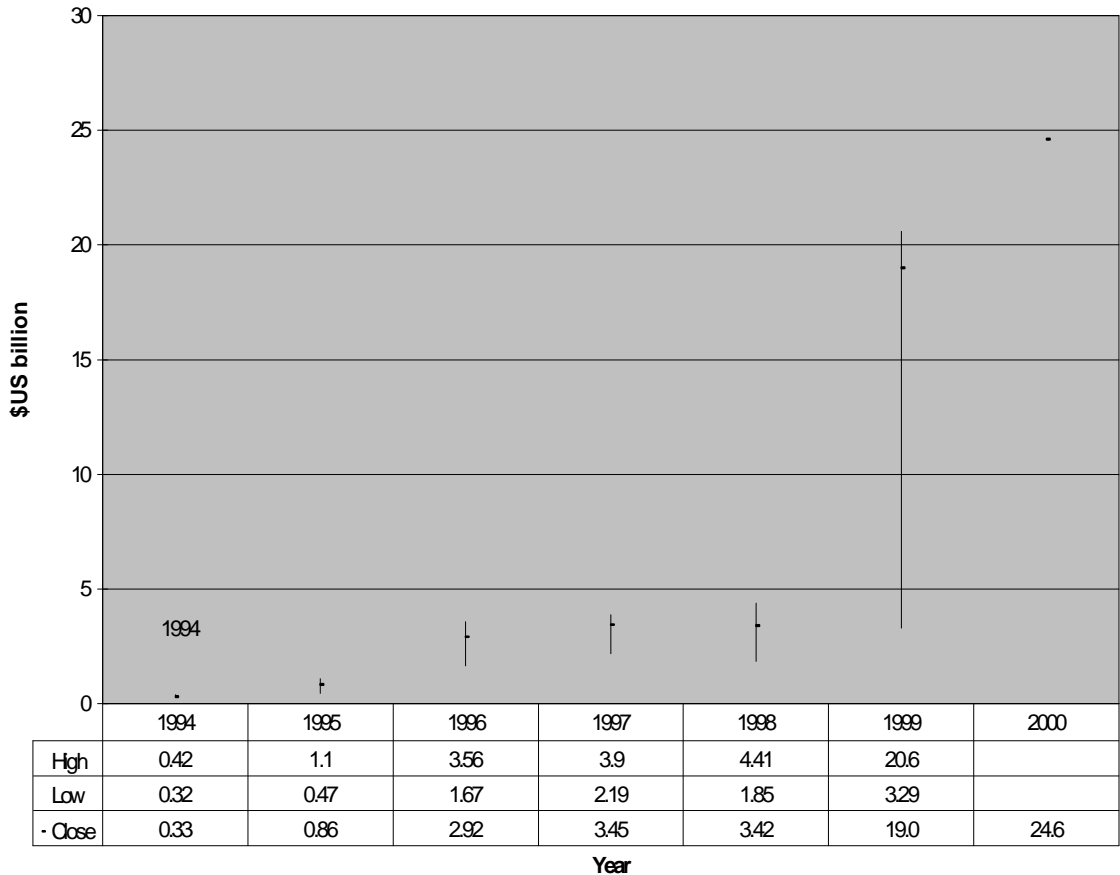
<sup>6</sup> PhRMA ([www.phrma.org](http://www.phrma.org)) includes most major research-intensive pharmaceutical firms. All other firms combined throughout the world account for comparable revenues, but have lower R&D expenditures. PhRMA reported \$26.4 billion R&D expenditures for 2000, of which \$22.5 billion was in the United States (<http://www.phrma.org/press/newsreleases//2000-01-18.59.phtml>).

<sup>7</sup> 14 of these 101 firms appear to be DNA sequencing services, genetic profiling services (for genetic testing or paternity testing or forensics), and some (e.g., Zymogenetics and Dragon Genomics) are research arms of larger firms. Most, however, appear to be startups that depend heavily on genomics in their business plans, as described on their websites, press releases, or in biotechnology and pharmaceuticals trade news articles. The entire list is available at: <http://www.stanford.edu/class/siw198q/websites/genomics/genomefirms.htm>.

<sup>8</sup> S.M. Thomas and Nicholas Simmonds, *The Industrial Use of Genome Resources in Europe*, Directorate-General for Science, Research and Development, European Commission (Brussels, Belgium) EUR 18850, 1999.







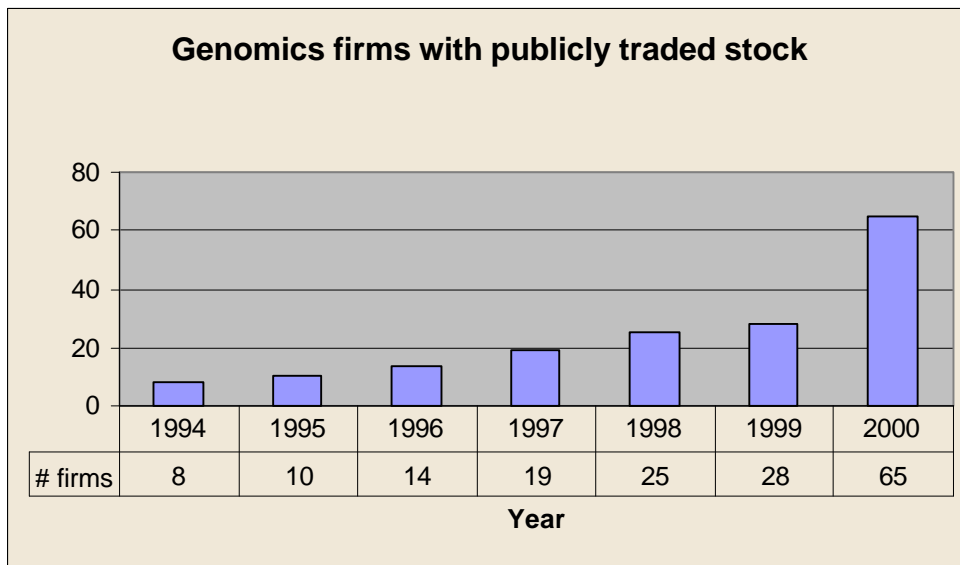
### Market Capitalization of the Largest Four Publicly Traded Genomics Firms

Interpreting the market capitalization figures must be done with some care. Two graphs show a broad range of publicly traded firms reported as “genomics” in whole or in part either by themselves or in articles mentioning them in biotechnology trade journals or scientific journals. The increases in capitalization are due to several different factors:

- An increase in the number of firms,
- An increase in stock prices, and
- An increasingly expansive definition of “genomics”.

The figures on the “big four” because these are the largest firms that are almost wholly dedicated to genomics. Their business strategies have consistently centered on large-scale, high-throughput creation and analysis of data on DNA structure. Incyte turned to genomics in 1993, Human Genome Sciences and Millennium were established that year, and Celera was created as part of PE Corp. in 1998. Tracking these four firms since 1998 permits some assessment of the effect of firm number versus market valuation.

The number of firms has consistently grown, but showed an especially marked jump 1999 to 2000.



## Patent Ownership

Number of patents assigned to organizations and firms surveyed

	DPD patents*	Comments
United States Government	472	Includes NIH, CDC, USDA, Department of Commerce, Department of Defense
Incyte	276	
Glaxo SmithKline	228	Includes Glaxo, Wellcome, Burroughs Wellcome, SmithKline Beecham, Beecham, Smith Kline & French
Genentech	224	Does not include Hoffmann-La Roche
Aventis	178	Includes Hoechst, Marion, Roussel Uclaf, Rhone-Poulenc and Rorer
Novo-Nordisk	164	Does *not* include Zymogenetics (with 80), surveyed separately; de-merger to split pharma and enzyme Cos.
American Home Products	150	Includes Genetics Institute and Wyeth-Ayerst
Eli Lilly	149	
Hoffman-La Roche	131	Includes Roche Holdings, Roche Research Institute, but not Genentech
Merck	131	Includes Merck&Co., Merck Patent GmbH, Merck Frosst, Rhone Merieux, and Merck Sharpe & Dohme
Novartis (incl. Institute of Functional Genomics)	112	Includes Sandoz, Ciba, Geigy and combinations
Human Genome Sciences Inc.	106	
Amgen	95	
Astra-Zeneca	79	
Pharmacia & Upjohn	77	Includes Pharmacia pharma, Upjohn and combinations
Schering-Plough (incl. Berlex)	69	Includes Berlex
PE Corp	63	Incl PE Corp, Perkin Elmer, Applied Biosystems, PerSeptive Biosystems (Celera had 0 patents by end of 1999)
Life Technologies	58	
Bayer	55	
Bristol-Myers Squibb	54	
Biogen	52	
Millennium Pharmaceuticals	49	
Genzyme	38	
Amersham Pharmacia Biotech	36	
Johnson & Johnson	34	Includes J&J, Ortho (various) and Janssen
Boehringer Ingelheim	32	
Lynx Therapeutics Inc.	27	
BASF	26	
Affymetrix Inc.	25	
Diversa Corp.	20	
Large Scale Biology Corp.	17	Includes BioSource Technologies
Myriad Genetics Inc.	16	
Pfizer, Inc. (see also Parke-Davis and Warner-Lambert)	14	Includes Warner-Lambert and Parke-Davis
Onyx Pharmaceuticals	13	
Otsuka Pharmaceuticals	13	

\* Patents in the DNA Patent Database ([www.genomic.org](http://www.genomic.org)), which covers DNA-based patents from 1980 through the end of 1999.

## Caveats and Qualifications

While we believe we have assembled the most comprehensive genomics funding data available in the public domain, our data nonetheless have gaps, some of them quite serious. Both our survey results and our analysis of publicly available data should be interpreted with caution, in light of the limitations noted below.

The survey had only modest aspirations, to get a rough estimate of the “public” genome funding and to show trends (but not absolute amounts) in private sector funding. We anticipated that few private firms would respond to the survey, despite promises to keep private firm responses confidential and to report such data only in aggregate.

We did not pursue all funding sources with equal vigor, but focused most on obtaining: (1) funding data from the largest government and nonprofit funding sources, and (2) R&D and market capitalization data on publicly traded genomics firms. For government and nonprofits, we list 83 nonprofit and government contacts, but we received data only from 20.<sup>9</sup> Because we focused on the largest such funders, however, we believe the reported figures may be useful as lower-bound estimates. If nonprofit and government funders submit data after this report is written, as several have indicated they will do, the new figures will be added to the spreadsheet posted on the website.<sup>10</sup>

### What is Genomics?

One of the most significant findings of the survey process was significant growth in funding for genomics. There are at least four causes of the increased funding levels:

1. Real funding increases,
2. Expansion of what counts as “genomics” (definition creep),
3. Dissemination of genomic methods into all of biomedical research (boundary blurring), and
4. Use of the term “genomics” to attract capital.

The term “genomics” was coined by mouse geneticist Tom Roderick to describe an approach to the study of DNA at the level of chromosomes, entire genomes, or large clusters of genes. The purpose of the term was to distinguish it from more traditional genetic approaches that focused on one gene or a family of functionally or structurally related genes or sequences. In addition to this scientific concept of genomics also implied large scale and a “whiz-bang” high tech approach to studying DNA structure. Implicitly, genomics implied creating and using large databases, extensive use of

---

<sup>9</sup> Most data came directly from survey responses. In a few cases, however, we secured data from alternative sources because important funders did not return surveys. Figures on the US Department of Energy, for example, were obtained from the White House. An email message from the Merck Genome Research Institute indicated it was company policy not to share its funding figures, but genomics research funding for 1998 was reported in the annual Merck corporate report on philanthropy (<http://www.merck.com/overview/philanthropy/12.htm>).

Several European and Asian national government contacts expressed interest but did not return surveys by the time of this final report, but may subsequently report data. If so, their data will be added to the project website (<http://www.stanford.edu/class/siw198q/websites/genomics/gov&nonprofitsurv.xls>). Late additions from private firms, however, cannot be incrementally added since we agreed in advance to report them only in aggregate.

<sup>10</sup> <http://www.stanford.edu/class/siw198q/websites/genomics/gov&nonprofitsurv.xls>

laboratory automation, and generally a more “capital intensive biology,” than was the norm in the mid-1980s.

Even when the term genomics first came into use, its boundaries were fuzzy, but as the genome project became less controversial and moved into the mainstream it grew even fuzzier. We addressed the definition in our survey cover letter (attached), but our brief and sketchy examples did not eliminate confusion, and indeed no one could: the definition of genomics was and remains inherently imprecise. Many respondents contacted us to get clarification, but we merely referred them to the relevant paragraph in the cover letter, which was only somewhat helpful. Some respondents contacted us to get clarification of how we defined genomics, and their vexation suggests that genomics as a category may not be useful for much longer to distinguish it from other parts of molecular biology. The usual concern was that micro-array technologies, DNA sequencing, and map data are now used in much of biomedical research, making it difficult to defend any bright lines between genomics and other research. Genomics has entered many fields. At some point invasion becomes conquest, and assimilation then follows.

At the National Institutes of Health, for example—the world’s largest single source of support for biomedical research, it once made sense to use the extramural budget for the National Center for Human Genome Research (which subsequently became the National Human Genome Research Institute, NHGRI) as a proxy for NIH’s “genome” budget. Now programs such as the Cancer Genome Anatomy Program at the National Cancer Institute (NCI), the Environmental Genome Program at the National Institute of Environmental Health Sciences, NCI’s Genetic Annotation Initiative, and the NIH-wide Mammalian Gene Collection fund work at least as “genomic” as that funded through NHGRI. Moreover, virtually all NIH institutes devote a substantial fraction of their budgets to gene hunts and programs that ten years ago would clearly have been defined as “genomics.” At the same time, the National Human Genome Research Institute intramural research program studies cancer and diseases studied in other NIH institutes and centers. NHGRI separates its intramural budget and does not count it as part of the Human Genome Project, which constitutes the lion’s share of its extramural budget.

Informal discussions with NIH institute directors and senior administrators suggested that as much as 20 percent of their “basic research” could be labeled “genomics,” but it was not reported that way and those sums were not included in our survey responses (except from the American Cancer Society and the Howard Hughes Medical Institute, which estimated their genomics expenditures as a fraction of their molecular biological research).

This definitional wobble is further destabilized by the dynamics of capital formation in the private sector. Private investment in genomics was virtually nil when the public Human Genome Project took flight. Before and during 1993, several small biotechnology firms redirected their efforts towards mapping and sequencing DNA, several new firms started up (including three of the “big four” tracked in our data: Human Genome Sciences, Incyte, and Millennium). Private funding reached rough parity with government and nonprofit funding in 1993 in the United States.<sup>11</sup> Since

---

<sup>11</sup> A point estimate at the end of 1993 was \$150 million in private investment, comparable to the \$169 million appropriated to NIH and DOE. See Robert Cook-Deegan, contract report "Survey of Genome

1994, the rate of growth in the public (government and nonprofit) sector has been substantial, but private genomics research funding has risen even faster. Private investment rose especially sharply in 1999 and 2000.

Some of this private investment was real, but some (unknown) fraction was merely semantic—it could be attributed startup firms recognizing that the term “genomics” had become a capital attractant (i.e., a buzz word). Company names are another indication that “genomics” became attractive to investors. Collaborative Research became Genome Therapeutics, and Incyte Pharmaceuticals became Incyte Genomics. Many startups put genomics or cognate terms in their names.<sup>12</sup> The magnitude of this effect is impossible to quantify, but it should be taken into account when interpreting the astounding growth rates of “genomics” in the market capitalization figures reported below.

The border between genomics and the rest of molecular biology has been stretched thin and has become porous. The term remains useful, but interpreting findings from this and other surveys needs to take definition creep and the changing meaning of genomics in capital markets into account.

### **What is a Genomics Firm?**

The congressional Office of Technology concluded in its pathbreaking 1984 report, and emphasized even more strongly in another 1991 report, that “biotechnology” is not an industrial sector, but rather a set of methods useful in many industrial sectors (usually established ones such as drugs and biologics, devices, or agriculture), but also for some entirely new applications (e.g., DNA forensics).<sup>13</sup> Many firms, almost 1500 listed by the various online services,<sup>14</sup> are called “biotechnology” firms because they are largely built around technologies developed since 1980. These firms are generally competing in established markets, however, even when they compete by using novel products, services, and technical approaches. This concept applies to genomics as well; it is not a field or set of firms but an approach. Some companies have become known as “genomics” firms because a substantial fraction or all of their business plan hinges on use of large data sets containing information about DNA structure, or depend on its interpretation.

Firms on our list differ how much they focus on genomics. Of the 13 firms that responded to our survey with funding data, 11 specified a fraction of their R&D devoted to genomics. (Another nine firms responded to the survey, but indicated they could not provide data—usually because such data were deemed proprietary—and six responded that they did not perform research and were removed from the list.) Six of the eleven reported 80 to 100 percent of their R&D was for genomics and five of

---

Research Corporations" prepared for the Office of Technology Assessment (OTA), US Congress, for its report on DNA patenting; survey December 1993; report March 1994. The contract report is available at the National Reference Center for Bioethics Literature, Georgetown University.

<sup>12</sup> Genomica, Genomics One, Cellomics, Dragon Genomics, Genometrix, GenoPlex, GenOway, Integrated Genomics, etc. See

<http://www.stanford.edu/class/siw198q/websites/genomics/genomefirms.htm>.

<sup>13</sup> Office of Technology Assessment, US Congress, *Commercial Biotechnology, An International Analysis* (1984) and *Biotechnology in a Global Economy* (1991). OTA-BA-218 and OTA-BA-494. Available online at <http://www.wvs.princeton.edu/~ota/>.

<sup>14</sup> See for example, BioSpace (<http://www.biospace.com/>), Recombinant Capital (<http://www.recap.com/>), or the Institute for Biotechnology Information (<http://biotechinfo.com/>).

eleven reported 10 to 25 percent. If our respondents represent the firms on our list (a big “if,” given the low return), then firms would be bimodally clustered with about half predominantly focused on genomics and half for whom genomics is a fourth or less of their research effort (recall that all firms on our list were identified either by themselves or in the trade press as “genomics” firms). Several smaller firms focus on platform technologies in one or a few areas within genomics (e.g., methods to create DNA sequence or gene expression data or to interpret such data).

The firms represented on the list do different things. No taxonomy is precise, but some clustering is apparent. Some firms (e.g., Genome Express, SeqWright) are service firms that do DNA sequencing or conduct DNA-based analyses sent to them by research laboratories. Some of them also do DNA forensics or genetic testing (e.g., Myriad Genetics). Some firms make instruments (e.g., Applied Biosystems). Others develop analytical software intended for whole-genome analysis, mining DNA sequence databases, or interpreting data on very large numbers of probes or gene expression arrays (e.g., Affymetrix, Sequenom, HySeq, Gene Trace, SuperArray).

The “big four” genomics firms (Incyte, Human Genome Sciences, Millennium, and Celera) themselves have four quite distinct business plans: Incyte was the first to focus on sequencing gene fragments, then moved to full-length genes, and has pursued gene expression arrays and genome informatics. It licensed access to its data to many firms nonexclusively. Human Genome Sciences also focused first on gene fragments and then full-length genes, but with a somewhat different targeting strategy. HGS initially had just one major licensee, SmithKline Beecham, but has since added other licensees. HGS emphasizes its desire to develop protein therapeutics and become a major pharmaceutical company. Millennium has engaged in a full range of genetics methods, including pedigree studies, linkage studies, gene association studies, lineage-by-descent studies, and of course, extensive mapping and sequencing. Through the use of its genomics technologies, Millennium also helps larger firms (mostly pharmaceutical manufacturers) to find and validate discovery target molecules. Finally, Celera was established in 1998 by PE Corp. to concentrate on genomic sequencing, to create massive genetic databases on sequence and sequence variations, and to develop informatic tools to interpret the huge data sets. It is filing provisional patent applications, but its business plan centers on subscribers paying for access to data and analytical tools.

This disparity of strategies means that the research funding we report might be spent on analyzing clinical data on human populations, on generating sequence data on humans or other organisms, on interpreting large genomic data sets, on developing new instruments or methods of mapping and sequencing, or on analyzing the function of genes and sequences at the organismal, cellular, or molecular level.

### **Incomplete Data on Private Firms**

We sent surveys to 192 firms (161 genomics firms and 32 established pharmaceutical and biotechnology firms). Only 13 genomics responded with data (and nine others that they could not provide data as a matter of policy). Two pharmaceutical firms returned surveys with funding data (two others reported that they would not be responding). The response rate was too low to use the figures for any aggregate sums; they do not even provide a meaningful floor. . Also, because we continued to discover firms we did not know about through the duration of our project, our list of



161 genomics firms and 32 established firms is not complete. However, it is more comprehensive than any other list we found during the project. One genomics investment analyst estimates there are 200 genomics firms, a figure that strikes us as about right, although no one can list them all.<sup>15</sup>

Our data on genomics firms rely heavily on publicly available data. While we did receive surveys from thirteen firms, this did not include any of the six largest. Our survey data for private firms are thus not useful for making aggregate estimates. (We do not report those survey data.) The data on private firms reported below are instead based on publicly available information: (1) R&D figures reported to the Securities and Exchange Commission (for publicly traded companies), (2) data from company annual reports, and (3) data on company market values.<sup>16</sup>

The data on privately held startups and established biotechnology and pharmaceutical firms are especially thin. It is quite difficult to get data on privately held firms unless they post it on their website or mail it out. We list 101 such firms, but received insufficient data from them to report any meaningful aggregate figures. The value of our survey with respect to these firms is limited to three outcomes: (1) having prepared a list (with hyperlinks) in order to conduct the survey, (2) a count of how many companies are based in different countries, and (3) number of patents assigned to the firms and catalogued in the DNA Patent Database. The patent counts, contact information and hyperlinks are available at the project website (and the country-company counts reported below).

Because most established biotechnology and pharmaceutical firms are publicly traded, R&D figures and market capitalization figures are available through SEC filings. Genomics is generally but a small fraction of the R&D efforts in such firms, however, and that fraction is regarded as proprietary by most.<sup>17</sup> It is therefore impossible to estimate genomics R&D expenditures, or even guess at trends. We were able to obtain counts of patents in the DNA Patent Database, and these are reported at the project website. The only data from the established firms that may be of interest showed a jump in genomics funding from 1998 to 1999 and 1999 to 2000 (from a low base in 1997), but even this may not reflect an industry-wide trend, but only the two responding firms.<sup>18</sup>

### **Category Errors and Double Counting**

As a check on our categories (government and nonprofit, genomics firms, and established pharmaceutical and biotechnology firms), we included questions on the survey about sources of funding. We do not report funding amounts on private firms because the response rate was too low, but if we had, it became apparent that we would have needed to adjust to avoid category errors and double counting. This is

---

<sup>15</sup> Noubar Afeyan, quoted in "Startups Climb the Human Genome to Funding" by Matthew A. DeBellis, *Red Herring*, August 17, 2000: <http://www.redherring.com/vc/2000/0817/vc-genome081700.html>.

<sup>16</sup> Data in our tables are based on calculations by Amber Johnson, but estimates are also available through the New York Times, (<http://business.nytimes.com/>), Hoover's online (<http://www.hoovers.com/>), Motley Fool (<http://www.fool.com>) and other online investment sites.

<sup>17</sup> The two firms reporting genomics funding figures also reported "percent R&D for genomics" of 3.1 and 4 percent, respectively.

<sup>18</sup> The summed figures reported were: \$29.5 million in 1998, \$41.4 million in 1999 and \$67.4 million in 2000; up from \$8.5 million in 1997 and \$1 million in 1996.

because some government and nonprofit funding goes to private firms, so would be counted in categories. This is particularly true in several European countries, such as Germany and Austria, that have explicit policies to promote biotechnology at a time when genomics firms make a significant fraction of new startups. The limited numbers of responses from such firms indicated this could be a serious problem for our survey. We mention it here because while we could have subtracted the percent government funding from such private firms and reported it as “government,” or subtracted it from the government and nonprofit category and reported it as private, it is actually a hybrid. Any future surveys will need to take this into account. Half or more of the capital for some genomics startups has been provided by governments, and it is probably a mistake to treat these as fitting neatly into any of our categories.

Another kind of error can produce double counting. Private firms are providing funding for a number of nonprofit consortia and research institutes (the SNP consortium and the Merck Genome Research Institute in the United States, and the Helix Research Institute in Japan). Several of these are quite substantial operations, and the double counting could be significant. Without adjustments, reports could include contributions to these research efforts in both the company R&D figures and in the nonprofit category. This is not a major problem in this report because we do not report the private firm R&D numbers based on our survey, and none of the large nonprofit consortia or research institutes receives funding from the publicly traded genomics firms whose R&D figures we obtained from SEC filings and annual reports. Some double counting did occur, however, because the Wellcome Trust contributes to the SNP Consortium, although reports from both are reported as though independent in our tables.<sup>19</sup>

## **Genomics Outside the Developed Economies**

The use of genomics outside the developed economies has received relatively little attention compared to the enormous media coverage of the effort to establish and sustain the publicly funded Human Genome Project, and then the race to produce a reference human genome sequence.

A few exceptions to this generalization merit mention. Brazil has a vibrant nucleus of genomics that includes a project on trypanosomal genomics. Projects to sequence and begin to characterize the genomes of major pathogenic microbes have enormous implications for populations in regions where diseases caused by such microbes are endemic or prone to epidemics erupting. Such sequencing efforts include the microbes associated with malaria, cholera, Chaga’s disease, schistosomiasis, and river blindness. Most such projects involve collaborations between investigators in Africa, Latin America, the Middle East or Asia with major genomics laboratories in the United States or Europe. The Institute for Genomic Research created a website to

---

<sup>19</sup> This would be easily corrected, because the Wellcome Trust has a sophisticated project-by-project budgeting system. We have contacted the Wellcome Trust to ask for the SNP Consortium contributions, so we can subsequently make a note of the double-count on the spreadsheet reporting survey results (on the project website). The tables in this report, however, contain this minor error, due to our methods, not the Wellcome Trust.

We included a question for private firms asking if they were reporting genomics funding that might appear in figures from other firms (e.g., subsidiaries separately surveyed, or agreements with genomics firms that might be counted by both). We now realize this problem is not restricted to firm-to-firm R&D figures, but also government and nonprofit funders as well. Thus a similar “double counting” question should be asked of *all* funders, not just private firms.

track such microbial sequencing projects (as well as others of primarily scientific or industrial interest).<sup>20</sup> WHO hosted meetings on parasite genome projects in 1995 and specifically on schistosomal genome projects in 1996. ReLab, a loose consortium of Latin American investigators interested in genomics, was formed even as the genome project was getting launched in the developed countries. And UNESCO helped fund a genomics fellowship program through the Third World Academy of Sciences, but this program has been suspended (see <http://www.ictp.trieste.it/~twas/TWAS.html>).

A limited literature touches on genomics' pertinence to people in developing countries.<sup>21</sup> First, within the modest but growing literature about patenting DNA and intellectual property rights in molecular biology, attention is turning toward genomics in private firms and its impact on developing countries. Some groups refer to "biocolonialism," "bioimperialism" and even "biopiracy." This concern has merged with issues that have lain dormant since the inception of the public debate about recombinant DNA, which was in small part also a debate about the Cohen-Boyer patent on the seminal gene-splicing technology. Second, there is attention to "benefits-sharing" throughout biomedical research. In genetics this is taking shape as concern about use of data and materials that arise from individuals who are part of genetic studies but who do not receive financial benefits from any ensuing profits.<sup>22</sup> This is a similar line of argument to the debate about exploiting genetic data on populations outside the developed world, that has led to legislation in India, China, and Brazil to oversee (and limit export of) valuable genetic data. Finally, the UNESCO declaration on the human genome included a debate about its international implications.

To our knowledge, the benefits-sharing discussions have not produced consistent policy guidance on setting research agendas or changing licensing and access provisions, or sharing intellectual property or royalties from it. There is some information about specific arrangements made with universities and nonprofit programs, and some individuals and institutions are in some cases dealing straightforwardly with benefits sharing in developing countries. The picture with respect to private firms is quite difficult to assess, as information is sparse and policy benchmarks either do not exist or are not public.

The understandable focus on completing a reference sequence may lead the media to pursue follow-up stories, and among these are likely to be some that involve transborder genomics. Several of the major funders in genomics (most notably the

---

<sup>20</sup> The Comprehensive Microbial Resource is at <<http://www.tigr.org/tigr-scripts/CMR2/CMRHomePage.spl>>.

<sup>21</sup> Some representative articles include: Allende JE: *FASEB J* 1991 Jan;5(1):6-7 "The human genome initiative. A view from the South"; Pena SD: *Trends Biotechnol* 1996 Mar;14(3):74-7 "Third World participation in genome projects"; Watson JD, Cook-Deegan RM: *JAMA* 1990 Jun 27;263(24):3322-4 "The human genome project and international health"; and Terragni F: *Cancer Detect Prev* 1993;17(2):317-21. "Biotechnology patents and ethical aspects". Perhaps the most extensive discussion is found in a draft report prepared for WHO—Daar AS and Mattei J.-F.: "Medical Genetics and Biotechnology: Implications for Public Health", prepared for WHO in consultation with the WHO Working Group on Genetic Manipulations and advice of an ad hoc expert committee, draft prepared December 1999. This report is not widely circulated, however, and merits more attention with respect to diverse issues related to implications of advances in genetics and cloning.

<sup>22</sup> See for example Kolata G: "Sharing of Profits Is Debated As the Value of Tissue Rises" *New York Times* Monday May 15, 2000, pp. A1ff.

Wellcome Trust) have a long history of concerns about health in the developing world.

There is intense discussion and effort to use genomics to create new medical products and services with private markets. Renewed attention to uses of genomics could also shed light on how the data and technologies can benefit populations other than those living in developed economies who have highly prevalent conditions. Such attention will, however, require organization and a strategy for mediating a productive discussion.

## Appendix I: Origins of the Survey

The survey was undertaken initially for three reasons. First, and most important, the data may be valuable for planning and analysis, both of genomics research and health research more generally.<sup>23</sup> Second, since 1997 the Stanford-in-Washington teaching program has included a module on genomics research as a case study in how health research policy decisions are made. Students have repeatedly asked about uses of genomics information for diseases prevalent in Africa, Latin America and the developing world when visiting genomics firms based near Washington, DC (Celera and Human Genome Sciences) and when meeting with the director of the National Institutes of Health. The survey data will be used in that course. Finally, one of the authors (RC-D) is writing a book on health research policy for the Robert Wood Johnson Foundation,<sup>24</sup> and a case study on genomics will be a section of that book. This survey will provide data for that chapter.

One underlying reason for addressing genomics research is that it starkly illustrates a trend apparent throughout health research: the private sector is becoming the dominant partner. This is not because of stagnant public funding—quite the contrary, funding for genomics has been growing faster than biomedical research in general among both nonprofits and government funders, and biomedical research has been growing faster than the physical sciences and most other components of R&D. Rather, the growth of private funding has considerably exceeded the growth of public and nonprofit funding. US-based R&D by members of the Pharmaceutical Research and Manufacturers Association (PhRMA), which includes most of the research-intensive pharmaceutical firms throughout the world, began to rise dramatically in the

---

<sup>23</sup> The idea for the survey originated in December 1999, following discussions with staff of the Global Forum for Health Research (GFHR) and the World Health Organization (WHO). Specifically, the GFHR international survey of health research was underway, spearheaded by Catherine Michaud of Harvard School of Public Health and Andrés de Francisco of GFHR. The Burroughs Wellcome Fund was approached for a mini-grant in February, and in April the BWFund generously funded the project. Work commenced late in May. The goal was to write a report in time to submit it to WHO and GFHR before an October 2000 meeting in Bangkok, Thailand.

<sup>24</sup> The project is “National Policies Governing Health Research” in the Health Policy Research Investigators Award program: [http://www.ahsr.org/rwjf/1998/cook\\_deegan.htm](http://www.ahsr.org/rwjf/1998/cook_deegan.htm).

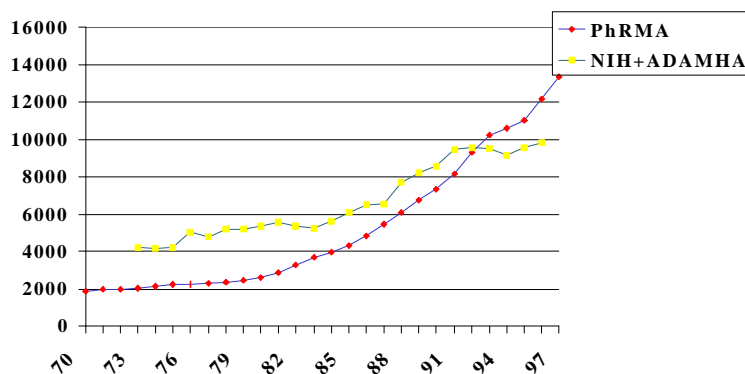
1980s and surpassed federal funding by the end of the decade.<sup>25</sup> The trend is both more sudden and more extreme in genomics.

Estimates of the market value of publicly traded genomics firms on which we have obtained information to date, show valuation in the range of \$95 billion in September 2000,<sup>26</sup> up sharply from \$37 billion at the end of 1999, \$22.6 billion in 1995, and \$760 million in 1994. Especially in 2000, the valuation of genomics firms, both individually and as a group, has proven highly volatile. For example, valuations dropped by 20 percent on a single day, following a March 15<sup>th</sup> announcement by President Bill Clinton and Prime Minister Tony Blair, about access to DNA sequence data,<sup>27</sup> although most have since recovered.

Policies intended to influence the direction of health research will have to take the massive private sector investment into account. Governments and nonprofit funders are quite important, but private pharmaceutical, biotechnology, and startup technology firms contribute an even larger share of resources in aggregate.

## Private R&D Exceeds Public

\$thousand, deflated to 1992 dollars



Sources: NSF FFRD Survey  
PhRMA Annual Survey 1997  
ADAMHA 1992 (corrected)

<sup>25</sup>

<sup>26</sup> This includes market capitalization of all 65 firms on our list of publicly traded firms. A significant fraction of these firms, however, do substantial nongenomics R&D, and so this \$95 billion market capitalization is a significant overestimate of “genomics” valuation.

<sup>27</sup> See Justin Gillis, “Clinton, Blair Urge Open Access to Gene Data,” *Washington Post* March 15, 2000, p. E1 (Business). The impact on stocks is covered in Alex Berenson and Nicholas Wade, “A Call for Sharing of Research Causes Gene Stocks to Plunge,” *New York Times* March 15, 2000.

## Appendix II: Methods

Data on genomics firms with publicly traded stock were taken from 10K and other reports to the U.S. Securities and Exchange Commission, company and organization websites, and company annual reports.

The one-page survey was fielded in June 2000.<sup>28</sup> The survey was mailed to contacts.<sup>29</sup> The initial mailing was followed by email, telephone, and in selected cases face-to-face follow-up with those who did not initially respond. The greatest efforts were expended on getting survey data from the largest sources of government and nonprofit genomics research funding.

**Government contacts** were mainly assembled by contacting known genome research administrators and scientists through email and telephone, adding contacts from public information and genome websites (such as the DOE program's portal site, Human Genome News, the Los Alamos list of most used genome websites, and the National Human Genome Research Institute), and building on a similar survey of government and nonprofit genomics research done for a book, *The Gene Wars* during 1990 and 1991. We emailed or phoned members of the Council for the Human Genome Organization, and participants in surveys of genomics done for the European Commission (one by Sandra Thomas and Nicholas Simmonds and another based on 1999 meetings under the chairmanship of Gert van Omenn).<sup>30</sup> We followed up on stories in *Science* and *Nature* on genome research in Italy, China, Germany, France, and international programs. The lists of contacts were posted on the project website, and for most of July and August were frequently revised.<sup>31</sup>

The list of “**genomics firms**” began from two sources: a December 1993 survey of the early genomics firms done by one of the authors,<sup>32</sup> and *BioWorld Report: 2000*

---

<sup>28</sup> The survey form is available online at <http://www.stanford.edu/class/siw198q/websites/genomics/survey.xls>.

<sup>29</sup> We assembled three contact lists: government and nonprofit funders, genomics firms, and established biotechnology and pharmaceutical firms with genomics collaborations. The contact lists began as five lists, with separate lists for government and private nonprofit funders and separate lists for publicly traded versus privately held genomics firms. Initial responses from several respondents on the government and nonprofit lists made clear that keeping the lists was confusing for many, in part because several organizations combine government and nonprofit funds. The publicly traded and privately held genomics firms were also combined into a single list for two reasons: (1) many firms became publicly traded while the survey was underway, and (2) it simplified the website. A new column was added to indicate whether a firm is public or privately held, and firms with publicly traded stock are now listed in red on the website.

<sup>30</sup> S. M. Thomas and Nicholas Simmonds, Science Policy Research Unit, University of Sussex, *The Industrial Use of Genomics in Europe*, 1999; and *Survey of the Current Status of 'Genomes' Research in the European Union*, European Commission, February 1999.

<sup>31</sup> The project website is <http://www.stanford.edu/class/siw198q/websites/genomics/>. Contacts were initially divided into five categories, and these were merged into three lists: government and nonprofit funders, small and startup firms with a known interest in genomics (we call these “genomics” firms), and established biotechnology and pharmaceutical firms known to have at least four collaborations with genomics firms or academic genomics centers.

<sup>32</sup> Robert Cook-Deegan, contract report "Survey of Genome Research Corporations" prepared for the Office of Technology Assessment (OTA), US Congress, for its report on DNA patenting; survey December 1993; report March 1994. The contract report is available at the National Reference Center for Bioethics Literature, Georgetown University. The draft OTA report that was never publicly released (because the US Congress defunded OTA in 1995) is also available there.



*Genomics Review*.<sup>33</sup> Our list then built on three principal sources, two web-based biotechnology services, BioSpace.com and Recombinant Capital, augmented by regular reading of scientific journals and biotechnology trade and technical publications. A few firms were also identified by membership in the Biotechnology Industry Organization or brought to our attention by scientists, stock analysts, or others.

To assemble contact information on the list, we visited the websites for each company (except the few lacking such websites), and made phone calls to clarify points of uncertainty. Our monitoring was greatly expedited by news about genomics firms in BioSpace.com's daily "Breaking News" service, Genomics Today (a news service of the Pharmaceutical Research and Manufacturers Association), and by reading scientific and trade journals (*Science* and *Nature* were the two most helpful, but we also found some contacts through *Drug Discovery*, *Nature Biotechnology*, *Nature Genetics*, *Genetic Engineering News*, *Red Herring*, and other publications and web news services). Finally, several websites for investors were essential to assembling data about R&D spending and market value of genomics firms: the EDGAR database from the Securities and Exchange Commission, the *Wall Street Journal* and *New York Times* and investor websites such as Hoover's Online<sup>34</sup> and Motley Fool.<sup>35</sup>

The list of **established biotechnology and pharmaceutical firms** began from firms with four or more collaborations listed in the BioWorld report, augmented by firms known to have genomics collaborations from news stories, scientists' referrals, or other personal contacts. Virtually all research-based pharmaceutical firms have some level of genomics investment, so the list is to some degree arbitrary because it relies on publicly disclosed genomics collaborations. We are not aware of public data that would permit a more systematic listing.

**U.S. Patent data** for each firm and organization through December 1999 were derived from patents assigned to the firms (or firms owned by those firms) noted in the DNA Patent Database. The DNA Patent Database contains US Patents starting in 1980 and extending through 1999. It is a subgroup of US patents selected by a search strategy that Jim Martinell of the U.S. Patent and Trademark Office devised in 1992 to assist the OTA study of 1993, and it has been updated each year since. Stephen McCormack and Robert Cook-Deegan read all the patents 1980-1993, and filtered out patents not actually making claims about DNA (or RNA) structure or methods. Patents since 1994 have been added, using the original Martinell search strategy, slightly modified by Stephen McCormack, as explained at the database website. Richard Burgess of OptiPat did the web implementation of the database, in collaboration with the Kennedy Institute of Ethics, Georgetown University and the Foundation for Genetic Medicine.<sup>36</sup>

---

<sup>33</sup> *2000 Genomics Review: New Technologies for the New Millennium* (Atlanta, GA: BioWorld Publishing Group, American Health Consultants, 3525 Piedmont Road, Building 6, Suite 400, Atlanta, GA 30305, USA; 2000). The BioWorld list included all firms that had been surveyed for OTA in 1993 except Mercator, which went out of business (after being purchased by Progenitor, before Progenitor closed in turn, in late 1998).

<sup>34</sup> <http://www.hoovers.com/>

<sup>35</sup> <http://www.fool.com/>

<sup>36</sup> The DNA Patent Database website is [www.genomic.org](http://www.genomic.org). It originated in a project commissioned by OTA through the National Reference Center for Bioethics Literature at Georgetown. For the current survey, we counted DNA-based patents 1980-1999 by querying the database for patents assigned to



## Appendix III: Cover Letter and Survey Form

[contact info]

September 29, 2000

Dear :

We are doing a survey of funding for genomics research throughout the world, in both the public and private sectors. The survey is being conducted by the Stanford-in-Washington Program with funding from the Burroughs Wellcome Fund. We hope and intend to produce a report that can inform the survey of health research and development throughout the world being conducted by the Global Forum for Health Research in Geneva, Switzerland. This survey is being sent to five groups of organizations: government funders, nonprofit funders, and private firms of three types—pharmaceutical firms, publicly traded genomics firms and privately held genomics firms. Our website lists the organizations being surveyed (<http://www.stanford.edu/class/siw198q/websites/genomics/>). The list was assembled from public sources: BioWorld, BioSpace.com, Recombinant Capital; journals such as *Science*, *Nature*, and *Drug Discovery*; and links on the World Wide Web. If you notice funding sources we have missed, please let us know by email ([bobcd@stanford.edu](mailto:bobcd@stanford.edu)).

If you are from a private firm, your response to the survey will be held in confidence and survey results from private firms will be reported only in aggregate. We are surveying more than 20 established pharmaceutical and biotechnology firms with known genomics interests (defined as 4 or more collaborations with genomics firms reported in the public sources, or known patent holdings based on genomics), 50 publicly traded and 50 privately held dedicated genomics firms. Results we report will not allow inferences about individual firms. We hope you do choose to respond, because the private sector is now likely funding a majority of genomics research worldwide, and estimates of funding from private firms will be of immense interest not only to government and nonprofit funders, but presumably also to you and other private respondents. These valuable data cannot be assembled reliably in any other way.

The survey form is just one page and follows this letter.

The definition of genomics is not precise. Tom Roderick coined the term, and Victor McKusick and Frank Ruddle used it to launch *Genomics* the journal in 1987. We follow their definition here, but leave interpretation to your discretion. We intend to include research that addresses all or a substantial portion of an organism's genome (including a chromosome or chromosome segment, but not a localized gene or gene family). This definition includes positional cloning if it starts from genome-wide (or chromosome-specific) marker scans. We include physical mapping and sequencing of all or a large part of a genome or chromosome. We also include array technologies that monitor expression of very large numbers of genes (hundreds or thousands), and informatic tools primarily intended to interpret DNA sequence or map information on a genomic or chromosomal scale. Software for melding high-throughput sequencing information into contigs would be included, for example, but not software for pedigree construction alone, or translation to protein sequence or simple homology

---

each firm. We adjusted for firm name changes, some of which are complicated (e.g., Collaborative Research Laboratories became Genome Therapeutics Corp, and what is now DNA Sciences Inc. was once Kiva Genetics. Its name change to DNA Sciences became possible with the purchase of the former DNA Sciences by Genetic Vectors). We also accounted for mergers and acquisitions (e.g., Genetics Institute was acquired by American Home Products). The search strategies for some pharmaceutical firms were especially complex, because of sequential mergers and acquisitions. (Some searches took as many as 9 separate steps. Smith, Kline & French merged with Beecham and Glaxo with Burroughs Wellcome, for example, and these two again are amidst a merger to become Glaxo SmithKline. Searches for individual firms were done, as well as Boolean queries to eliminate double-counting). Patent counts are reported on the "genomics firms" and "established biotechnology and pharmaceutical firm" tables on the project website. The "comments" column to the right of the patent count for each firm explains inclusion and exclusion criteria. The DNA Patent Database does not include all DNA patents, and a small, but unknown, fraction of patents 1994 and after in the database are not in fact DNA-based; the figures on patent holdings should therefore be used only as rough indicators.

comparison. We include techniques for high-throughput sequencing or genome-scale mapping, but not research directed at one or a few alleles (e.g., a single-locus diagnostic test would be excluded, even if based on DNA sequencing). We acknowledge broad gray zones, and accept therefore that genomics research is what you say it is.

In order to make our results available to the Global Forum in developing its report, we need to have responses to the survey by June 21, 2000. One of us may also follow up by phone or email to address any questions you might have.

Sincerely,

Robert Cook-Deegan, M.D.

Carmie Chan      Amber Johnson

[insert spreadsheet: survey.xls from website]