Decoding the literature on genetic variation

A survey of the scientific and patent literature on single-nucleotide variants reveals the dominance of research centers in the United States and the prolific patenting of SNP technology by a select group of biotechnology companies.

Roger Coronini, Marie-Angèle de Looze, Pierre Puget, Gérard Bley, and Shyama V. Ramani

Single-nucleotide polymorphisms (SNPs) are DNA sequence variations among individuals. Publicly funded laboratories and private businesses are attempting to associate specific SNPs (or sets of SNPs) with various medical conditions and to study the differences in SNP patterns among various human populations. Ultimately, it is hoped that knowledge of SNPs will improve medical treatment by enabling prediction of disease risk and response to therapies. To facilitate these efforts, the importance of providing publicly accessible SNP data without intellectual property restrictions prompted the formation of the SNP Consortium (http://snp.cshl.org/), a public-private initiative that, as *Nature Biotechnology* went to press, has now placed ~1.8 million SNPs in the public arena.

In this article, we present an analysis of the patent and scientific literature on SNPs (up to the end of 2001) to identify the key academic researchers and centers of excellence in the area, assess the major commercial developers of SNP technology, and understand the nature and progress of work currently underway. To assess the creation of new knowledge, we conducted a detailed study of papers published in the scientific literature. We have also measured SNP technological innovation by assessing patent applications filed at the major patent offices (see "Methodology"). The survey covers the scientific and patent literature over the period 1987–2001.

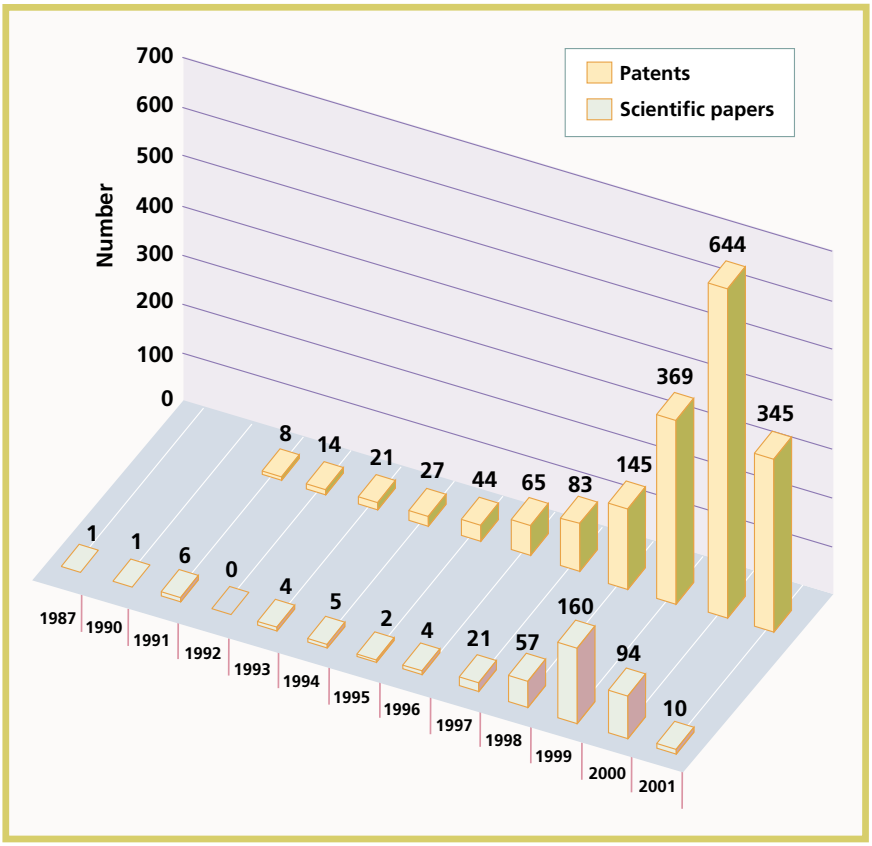

Figure 1 Scientific papers and patents relating to SNP research in the period 1987–2001.

An emerging area

From 1987 to 2001, 1,828 papers were published and 365 patents were filed on the topic of SNPs. Of these totals, 82% of papers were published and 88% of patents were filed between 1998 and 2001, suggesting that the field is still relatively young (Fig. 1). Since 1998, the number of publications has increased fourfold, and the number of patents threefold—reflecting a sudden spurt of knowledge creation quite typical of an emerging area.

As Figure 1 indicates, a sharp increase in the number of scientific papers and patent publications occurred almost simultaneously in 1998. This is noteworthy because in many high-technology sectors, spurts in scientific publication tend to precede increases in the number of patents filed by a few years. As SNPs are the latest in a long line of genetic markers (including minisatellites and microsatellites) used in efforts to map human disease and other complex traits, it is possible that familiarity with mapping technology facilitated their rapid incorporation into company R&D programs. Moreover, SNPs clearly are of interest to both university researchers and industrial researchers: for academics, they provide a means of mapping traits to genomes at higher and higher resolution; for industrial researchers, they are of use in designing SNP diagnostics (for instance, to determine disease risk or drug response) or in basic drug-discovery research to design chemical and biological entities that can effectively address all polymorphic variants of a drug target.

Patent applications are published only 18 months after deposition. Several more months pass before they are collated into databases—for example, the Derwent Biotechnology Abstracts database is

Roger Coronini, Marie-Angèle de Looze, and Shyama V. Ramani are researchers at the Institut National de la Recherche Agronomique, Université Pierre Mendès France, Grenoble, France, and Pierre Puget and Gérard Bley are researchers at the Commissariat à l'Energie Atomique, Laboratoire d'Electronique et des Technologies de l'Information, Grenoble, France (coronini@grenoble.inra.fr).

Table 1. The most prolific authors of SNP-related papers

No. papers	Name	Affiliation	Country
23	Yusuke Nakamura	Human Genome Center and Department of Cardiovascular Medicine, Medical School, University of Tokyo, Tokyo	Japan
21	Pui-Yan Kwok	School of Medicine, Division of Dermatology, Washington University, St. Louis, MO	United States
19	Giancarlo Colombo	CNR Center for Neuropharmacology, University of Cagliari, Cagliari	Italy
16	Mitsuru Emi	Department of Molecular Biology, Institute of Gerontology, Nippon Medical School, Kawasaki	Japan
16	Gian Luigi Gessa	Department of Drug Sciences, University of Sassari, Sassari	Italy
16	Eric S. Lander	Whitehead Institute, Center for Genome Research, Massachusetts Institute, of Technology (MIT), Cambridge, MA	United States
16	Katsushi Tokunaga	Graduate School of Medicine, Department of Human Genetics, Bunkyo Ku, University of Tokyo, Tokyo	Japan
15	Deborah A. Nickerson	Department of Molecular Biotechnology, University of Washington, Seattle, WA	United States
14	Peter J. Oefner	Genome Technology Center, Stanford University, Palo Alto, CA	United States
13	Toshihiro Tanaka	Institute for Medical Sciences, Center for the Human Genome, Laboratory of Molecular Medicine, Minato Ku, University of Tokyo, Tokyo	Japan

Source: Biotechnology section of the Web of Science.

updated every trimester, whereas the World Patents Index is updated every week. A similar cataloging delay occurs for scientific publications—the Biotechnology Citation Index is updated every trimester, whereas the Web of Science is updated daily. Of course, the process of peer-reviewed publication also introduces delays into the release of results. Our present study was conducted at the end of 2001, and thus the data relating to the years 2000 and 2001 are necessarily incomplete, which accounts for the lower numbers of patents between 2000 and 2001 (Fig. 1). This might also explain why some prominent biotechnology companies involved in SNP research—such as Luminex (Austin, TX) and Pyrosequencing AB (Uppsala, Sweden)—did not appear in our survey; their patents or publications may have appeared in the literature after the time period covered by our survey.

US papers predominate

We have categorized authors of SNP papers in two ways: those who publish the greatest number of papers (Table 1) and those whose work is referred to most frequently in the scientific literature (Table 2). The "most cited" authors listed may be regarded as the researchers with the greatest impact on the work of their scientific peers, and can therefore be viewed as leaders with the most influence in the creation of knowledge.

The convention that academic scientists disclose and share data in journals more readily than their corporate colleagues is supported by Table 1, which shows that the top ten "most prolific" authors of scientific papers on SNPs all come from public institutions. Unlike their industrial colleagues, academic scientists are judged and ranked according to their publication record and publish regularly to ensure future grant awards. Notably, the majority of authors in Table 1 come from research institutions in the United States and Japan, with two originating from European research centers. Although certain companies clearly opt against scientific publication, an examination of the 30 most prolific authors does, however, reveal some companies that do take this route; for example, the Canon Research Center (Kanagawa, Japan) and Perkin Elmer (now renamed Applied Biosystems; Foster City, CA) have published 10 and 8 papers on SNPs, respectively.

The pre-eminence of US researchers in the SNP field is confirmed by data presented in Table 2. Francis Collins heads the list with 219 citations: given his status as director of the US National Human Genome Research Institute (Rockville, MD), many of his papers presumably present policy and direction for SNP research efforts in the United States and are therefore highly cited. Data from Table 2 also indicate that the Whitehead Institute for Biomedical Research (Cambridge, MA) is a world-leading center for SNP research, providing a home for several top research groups. Other "most cited" authors come from laboratories elsewhere in the United States, in Japan, or in Sweden.

An examination of the 30 top-cited authors reveals that the US lead in the area of SNP research extends far beyond the top ten researchers. Of the 30 top-cited authors, 23 are located in the United States; of the other 7, 4 hail from Scandinavian countries (Sweden and Finland), and 1 each from Japan and the United Kingdom.

Ten researchers figure among the top 30 most prolific authors as well as the top 30 most cited authors. Again, most of these are from the United States; notably, Eric Lander and Deborah A. Nickerson are among the top ten in both categories. Only three European researchers are among the top 30 most cited authors: Ulf Landegren of the University of Uppsala (Uppsala, Sweden) with 75 citations, Anthony J. Brookes of the Karolinska Institute Centre (Stockholm, Sweden) with 50 citations, and Tomi Pastinen of the National Public Health Institute (Helsinki, Finland) with 73 citations. Although four Japanese researchers are among the 10 most prolific authors, they are not among the top 30 most cited.

Voracious biotech patenting

In total, we identified 164 different organizations that applied for SNP patents in the time period 1987–2001. To identify organizations that have strategically focused their intellectual property on SNPs, Table 3 lists 30 patent applicants that have deposited at least three patents each (from now on, we refer to this group of patentees as "leaders in patenting"). Their 241 patent applications account for 66% of the total number of applications on SNPs.

As one would expect, Table 3 reveals that a majority (18) of the 30 "leaders" involved

Table 2. The most cited authors of SNP-related papers

No.	Author	Laboratory	Country
219	Francis S. Collins	National Human Genome Research Institute, National Institutes of Health, Bethesda, MD	United States
185	Leonid Kruglyak	Whitehead Institute for Biomedical Research, Cambridge, MA	United States
184	David G. Wang	Whitehead Institute for Biomedical Research, Cambridge, MA	United States
171	Neil Risch	School of Medicine, Department of Genetics, Stanford University, Stanford, CA	United States
151	Michele Cargill	Whitehead Institute, Center for Genome Research, MIT, Cambridge, MA	United States
149	Masato Orita	Oncogene Division, National Cancer Center Research Institute, Tokyo	Japan
124	Eric S. Lander	Whitehead Institute, Center for Genome Research, MIT, Cambridge, MA	United States
122	Marc K. Halushka	Case Western Reserve University, School of Medicine, Department of Genetics, Cleveland, OH	United States
112	Ann-Christine Syvänen	Academic Hospital, University of Uppsala, Uppsala	Sweden
110	Deborah A. Nickerson	Department of Molecular Biotechnology, University of Washington, Seattle, WA	United States

Source: Biotechnology section of the Web of Science.

in patenting SNPs are research-based biotechnology companies. Eleven of these 18 companies are startups of less than 10 years old. Clearly, as for SNP scientific papers, the patent literature on SNPs is dominated by US-based organizations. Many of these companies are developing technology platforms to transform SNP research into medical applications either for their own in-house drug discovery programs or to provide diagnostic and/or genotyping services to others. Commercializing kits for typing SNPs is also the focus for biotechnology companies focusing on reagents and equipment, such as Qiagen and Promega.

The rest of the patentees include a smattering of large pharmaceutical firms and some not-for-profit research centers. The use of SNPs in identifying traits of agronomic importance is emphasized by the intellectual property in this area held by Monsanto and Pioneer Hi-Bred. SNP patent applications in agriculture account for 2% of the total number of applications by the top 30 patentees.

Overall, research-based biotechnology companies have a total of 184 patent applications—or around 76% of all the patent applications of the "leaders." The supplier companies, pharmaceutical companies, and five not-for-profit research centers account for 17 (7%), 13 (7%), and 21 (7%) of the total, respectively.

It must be kept in mind that for many biotechnology companies, such as Curagen, Molecular Tool (acquired in 1998 by Orchid), Orchid Biocomputer (now Orchid Bioscience), and Affymetrix (and its spin-off Perlegen; Santa Clara, CA), the recent spurt in patent filings is

Methodology

The methodology used (termed "scientometrics") is defined as the measurement of knowledge creation through an examination of publications and patents available electronically. In the research presented here, the "citations" method was applied to data drawn from the Biotechnology section of the Web of Science database produced by Thomson ISI (Philadelphia, PA), which provides both the documentary reference and the bibliography of each article identified. In addition, the Derwent Biotechnology Abstracts (DBA) database, the European and PCT Patent Application Bibliography Espace Access database from the European Patent Office, and the Derwent World Patents Index database were analyzed to identify the main areas of application and the role of firms in SNP research. The strategy is first to identify the most cited authors and leaders in scientific publications in a given time period (1987–2001) and then to identify equivalent industrial research by examining the patent applications in the same domain (through the use of similar equations for research). This strategy then allows the two databases to be compared and analyzed.

"Most prolific" and "most cited" authors. The data available in the Web of Science have a classic bibliographic structure and contain the references cited by the authors. By transferring these data onto a classical documentary database program, it is easy to establish indexes for each of the elements and to compile a list of the most cited authors. The basic documentary reference identifies the most productive authors in the field by associating each author with the number of papers published during the period under consideration (1987–2001) and then selecting the most prolific authors. The citations field identifies the "most cited" authors in the field as the authors with the highest frequency of citations.

SNP patent searching definitions. Leaders in SNP patenting were defined as organizations (companies or research centers) that have either applied for or obtained a patent in the fields under study. To formulate the research equations needed to define the subject (SNPs), we checked the quality of the corpus (relevance of references, belonging to the field of SNPs) for the period under study. Data were extracted using a search research strategy based on key words involving combinations of "single", "nucleotide", and "polymorphism". The main search strategy used the following terms: "SNP OR SNPS OR (single nucle* SAME polymorphism*) NOT (nitroprussid* OR neutron OR soluble nonreactive phosphor* OR MAX SNP OR stack node OR nanoprofil* OR special needs passengers OR tin OR donor SNP OR NO donor* OR Si nanoparticle OR stimulated nuclear OR space nuclear power OR TOPAZ OR thermal reactor OR space OR secure network* OR slow negative potential* OR simulated natural photoperiod OR aerospace OR nuclear)". The results were reformatted on a database manager and then analyzed. Data- and text-mining software was then used to tease out the main themes covered by each type of data. The software packages used were Image's (Toulouse, France) ALCESTE/ADT image and SPSS's (Chicago, IL) LexiQuest Mine and Clementine.

Table 3. Leading patentees of SNP research: main characteristics

Company	Founded	Country	Patents	Publications*
Agbiotech companies				
Monsanto (St. Louis, MO)	Before 1980	US	3	0
Pioneer Hi-Bred (Des Moines, IA)	Before 1980	US	3	0
Subtotal			*6*	*0*
Research-based biotechnology companies				
Curagen (New Haven, CT)	1993	US	42	45
Genaissance Pharmaceuticals (New Haven, CT)	1997	US	24	16
Orchid Biocomputer (now Biosciences; Princeton, NJ)	1995	US	18	27
Molecular Tool (acquired in 1998 by Orchid)	1988	US	16	0
Incyte Genomics (now Pharmaceuticals; Palo Alto, CA)	1991	US	15	25
Aclara BioSciences (Mountain View, CA)	1995	US	9	177
Epigenomics (Berlin, Germany)	2000	Germany	8	0
Genset (Paris, France)	1989	France	7	86
Nanogen (San Diego, CA)	1993	US	7	23
Affymetrix (Santa Clara, CA)	1993	US	6	703
Biogen (Cambridge, MA)	1978	US	6	132
Exact Laboratories (now Exact Sciences; Maynard, MA)	1995	US	6	7
LJL Biosystems (acquired in 2000 by Molecular Devices; Sunnyvale, CA)	1988	US	6	5
Illumina (San Diego, CA)	1998	US	4	107
RapiGene (now acquired by Qiagen)	1988	US	4	3
DzGenes (St. Louis, MO)	1998	US	3	27
Isis Innovation (Oxford, UK)	1988	UK	3	7
Subtotal			*184*	*1390*
Reagent and equipment suppliers				
Perkin Elmer (Foster City, CA)	before 1980	US	5	605
Qiagen Genomics (Bothell, WA)	1988	US	5	23
Promega (Madison, WN)	1978	US	4	6
Packard (now Hewlett Packard; Roseville, CA)	before 1980	US	3	0
Subtotal			*17*	*634*
Pharmaceutical companies				
AstraZeneca (Södertälje, Sweden)	before 1980	Sweden	9	8
Glaxo (now GlaxoSmithKline; King of Prussia, PA)	before 1980	UK	4	840
Subtotal			*13*	*848*
Not-for-profit research centers				
University of Alabama, Birmingham Research Foundation (Birmingham, AL)	1987	US	4	0
Wisconsin Alumni Research Foundation (Madison, WI)	before 1980	US	4	0
Cornell Research Foundation (New York)	before 1980	US	3	211
Whitehead Institute for Biomedical Research (Cambridge, MA)	before 1980	US	5	1
University of California system	before 1980	US	5	13
Subtotal			*21*	*225*
Total			**241**	**3097**

*Information obtained for papers from 1991–2001 from the Biotechnology Citation Index. Publications therefore cover all the fields of biotechnology and not just SNPs; the Biotechnology Citation Index compiles information only from academic or science-based journals (http://www.isinet.com/isi/products/citation/specialty/btci/) Patent information from Derwent Biotechnology Abstracts, The European and PCT Patent Application Bibliography Espace Access from the European Patent Office, and the Derwent World Patents Index patent database.

the outcome of past investment in SNP research. These companies have been applying for SNP patents since the start of the 1990s and are now poised to be strategic players in this narrow market.

Twenty-five of the 30 patenting "leaders" are affiliated with US organizations (26 if you count Glaxo as part of Glaxo SmithKline). The remainder are linked with Swedish, UK, German, or French companies. Notably, the Japanese are absent among the top 30 patentees (even though Japan has ten patents in this domain, no single Japanese organization has as many as three patents to its credit).

Also noteworthy is the emphasis on patenting as compared to publishing among the different biotechnology companies listed in Table 3. With the exception of Aclara Biosciences, Affymetrix, Biogen, Illumina, Genset, and Curagen, biotechnology companies appear to prefer patent filing to publishing papers; indeed, two leading patentees, Molecular Tool (acquired by Orchid Biosciences), and Epigenomics AG), have no publications to their name. (As Isis Innovation is the technology-transfer arm of Oxford University, it is no surprise that it produces no publications.)

To a certain extent, the lower number of papers originating from companies may reflect the fact that a database like the Web of Science (which collates about 8,500 journals) is very selective, including only mainstream academic journals. But overall, the emphasis of private enterprises (particularly startup companies) toward patents (rather than papers) reflects the importance of product focus and trade secrets for maintaining competitive edge in the commercial market. Moreover, because the technology transfer arms of universities involved in filing for patents (such as the University of Alabama, Birmingham Research Foundation or the Wisconsin Alumni Research Foundation) are separate from the university institutions where research actually takes place, they also are unlikely to be linked with academics who author significant numbers of scientific papers (Table 3).

Ailments, apparatus, and alleles

Table 4 lists the top ten public organizations and top ten companies in terms of scientific papers and patent applications. Clearly, those public organizations involved in human genome research with particular expertise in SNPs (e.g., the Stanford Human Genome Center, the University of Tokyo Human Genome Center, and the University of Washington Genome Center) have been most prolific. In contrast, very few biotechnology

Table 4. Distribution of publications and patents in all areas of SNP research among top ten public organizations and companies*

Scientific information						Patents		
Top ten public organizations	**No. papers**	**Country**	**Top ten companies**	**No. papers**	**Country**	**Top ten patentees**	**No. patents**	**Country**
University of Tokyo, Tokyo	110	Japan	PPGx (Research Triangle Park, NC)	11	US	Curagen	42	US
University of California system	106	US				Genaissance Pharmaceuticals	24	US
University of Washington (Seattle, WA)	80	US	Affymetrix	10	US	Orchid Biocomputer	18	US
University of Texas system	62	US	Third Wave Technology (Madison, WI)	9	US	Molecular Tool	16	US
Stanford University (Stanford, CA)	50	US				Incyte Genomics	15	US
Harvard University (Cambridge, MA)	48	US	Nanogen	8	US	Aclara BioSciences	9	US
Vanderbilt University (Nashville, TN)	45	US	Transgenomic (Omaha, NB)	8	US	AstraZeneca AB Pharmaceuticals	9	Sweden
Yale University (New Haven, CT)	41	US	Glaxo	7	US	Epigenomic AG	8	Germany
University of Alabama (Birmingham, AL)	39	US	Sequenom (San Diego, CA)	6	US	Genset	7	France
University of Pennsylvania system	36	US	Wakunaga Pharmaceutical (Osaka)	5	Japan	Nanogen	7	US
			Axys Pharmaceuticals (S. San Francisco, CA)	4	US			
			Curagen	4	US			
Total for top 10	**617**			**72**			**155**	
Proportion of this group in all SNP papers/patents*	**0.34**			**0.03**			**0.42**	

*A total of 1,828 papers and 365 patents focused on SNP research were found in the scientific and patent literature for the period 1987–2001.
Sources: Biotechnology section of the Web of Science and Derwent Biotechnology Abstracts, The European and PCT Patent Application Bibliography Espace Access from the European Patent Office, and the Derwent World Patents Index patent database.

companies have published reports related to SNPs indexed by the Web of Science.

In terms of patents, Curagen, Genaissance Pharmaceuticals, Orchid Biosciences (formerly Orchid Biocomputer, which also acquired Molecular Tool), and Incyte Genomics are clearly prolifically filing patents on SNPs. It is interesting to note that startup Genaissance Pharmaceuticals was founded in 1998 and originated out of Yale University, which is among the top ten public organizations publishing papers on SNPs.

For the sake of simplicity, we have further grouped our analysis of scientific publications and patents on SNPs into four main areas: associations of SNPs with human disease (Table 5), methods and techniques for scoring and discovering SNPs (Table 6), basic molecular biology of SNPs (Table 7), and allele frequencies of SNPs in populations (Table 8). In each area of SNPs, each Table presents public organizations who have published 10 or more publications on SNPs and patentees with 2 or more patent applications.

Taking into account all four areas, 1,481 papers (907 of them shown in Tables 5–8) appeared in the scientific literature in the period 1987–2001. The remaining 347 papers on SNPs focus on discussing the technology's "promise" in healthcare, agriculture, and basic biology, reflecting the fact that papers commenting on SNP research far outweigh, numerically, those detailing data of practical use.

SNP associations and disease

More than a quarter of all biotechnology papers and patents concerning SNPs deal with their associations with complex disease (particularly those afflicting much of the developed world, such as arthritis, asthma, obesity, and cancer; Table 5). Researchers hope to detect the SNPs associated with predispositions to these diseases, which could lead them to the underlying genes. There have been several papers and patents in this area that focus on applying SNP information to diagnostics and nucleic acid therapies.

The majority of biotechnology papers in this area originate from public laboratories in the United States, but Humboldt University in Germany, the University of Utrecht in the Netherlands, Kagawa University (Kagawa, Japan), McGill University (Toronto, Canada), the University of Newcastle (Newcastle, UK), the University of Manchester (Manchester, UK), the University of Oxford (Oxford, UK), and Wakunaga Pharmaceuticals (Osaka) in Japan also are publishing in this area.

The large number of patent applications in this area from Genaissance Pharmaceuticals and Curagen indicates that SNP associations with human disease form a strategic focus for these companies, with Orchid Biosciences also strong in the field. Several public organizations have been filing for patents in this area, including the Whitehead Institute for Biomedical Research, the University of Alabama Research Foundation (Birmingham, AL) and the University of California system. Interestingly, one company, PPGx (Research Triangle Park, NC), has five papers published in this area but no patents filed, according to our analysis. In addition to research-intensive biotechnology companies, well-established multinationals, such as AstraZeneca AB, have also been filing patents in regard to disease-related SNPs.

Methods and techniques

By far the greatest proportion of SNP patent applications (55% of the total number of SNP patents) focus on technologies that enable high-throughput detection of genetic variants. Research in this area is published more as patents (55%) than as papers (18% of the total number of SNP papers), emphasizing the importance of this area to companies. Patents mostly describe protocols and biochemical methods (such as amplification methods, nucleic acids hybridization methods, or enzy-

Table 5. Distribution of publications and patents focused on association of SNPs with human diseases and disorders

Scientific information						Patents		
Lead public organizations[1]	No. papers	Country	Lead companies[1]	No. papers	Country	Lead patentees[2]	No. patents	Country
University of Tokyo	36	Japan	PPGx	5	US	Genaissance Pharmaceuticals	23	US
Harvard University	30	US	Burleson Research	2	US	Curagen	18	US
University of California System	30	US	Technology (Raleigh, NC)			Orchid Biocomputer	7	US
University of Texas	28	US	Promega	2	US	Molecular Tool	7	US
Humboldt University	19	Germany	Eli Lilly (Indianapolis, IN)	3	US	Incyte Genomics	7	US
Washington University	19	US				AstraZeneca AB	7	Sweden
Vanderbilt University	18	US	GSF Forschungszentrum Umwelt & Gesundheit (Neuherberg)	2	Germany	Epigenomics AG	5	Germany
Brigham and Women's Hospital	16	US				Whitehead Institute for Biomedical Research	4	US
University of Alabama	16	US	Wakunaga Pharmaceuticals	2	Japan			
University of Utrecht (Utrecht)	16	Nether-lands				University of Alabama Research Foundation	4	US
Kagawa University	13	Japan						
University Cincinnati (Cincinnati OH)	13	US				Nanogen	3	US
						Isis Innovation	3	UK
University of Pennsylvania	13	US				DZgenes	3	US
						Biogen	3	US
McGill University (Toronto)	11	Canada				Affymetrix	3	US
						University of California System	2	US
University of Newcastle	11	UK						
China Medical College Hospital	10	China						
University of Chicago	10	US						
University of Manchester	10	UK						
University of Oxford	10	UK						
Publications in this area as a percentage of all SNP papers/ patents[3]	**23%**			**3%**			**28%**	

[1]Table displays only those public organizations with 10 or more papers on SNPs in the area.
[2]Patentees with less than two patents in the area are not shown.
[3]Note percentages provided in Tables 5 to 8 do not sum to 100% as we have not included papers on SNPs that merely discuss their promise. Furthermore, the distribution of patents over the four SNP areas does not add up to 100% because many patents are affiliated to more than one area.
Sources: Biotechnology section of the Web of Science and Derwent Biotechnology Abstracts, The European and PCT Patent Application Bibliography Espace Access from the European Patent Office, and the Derwent World Patents Index patent database.

matic methods), their implementation using innovative devices (such as microarrays or microfluidics systems), or ways to automate them (such as automated fluid handling and detection).

Leading biotechnology companies with a strategic focus on technology platforms for SNP genotyping or detection include Orchid Biosciences, Aclara Bioscience, Curagen, Epigenomics AG, Nanogen, Affymetrix, Incyte Genomics, and Illumina. Although these comprise by far the largest group of companies patenting SNPs tools, some large multinational pharmaceutical companies (such as AstraZeneca and GlaxoSmithKline) are also represented, although with many fewer applications. Some of the patentees are equipment and reagent suppliers and manufacturers, such as Perkin Elmer and Becton Dickinson; others are information technology companies that have moved into the life sciences, such as Hewlett Packard.

Public organizations active in this area include the Cornell Research Foundation, SNP consortium member the Whitehead Institute for Biomedical Research, the Wisconsin Alumni Research Foundation, the University of California system, and the University of Washington (Seattle). Notably, the Whitehead Institute and the University of Washington host several of the top ten cited authors listed in Table 2: Eric Lander, David Wang, Michele Cargill, and Leonid Kruglyak all are based at the Whitehead Institute, and Deborah Nickerson is on the faculty of the University of Washington.

The United States again is the most prominent country from which research on SNP tools is being published and patented. Nearly all (29 of 32) of the biotechnology companies working in this area—all except Epigenomics AG (Germany), Keygene (the Netherlands), and Asper Biotech (Tartu, Estonia)—are based in the United States, and only one European pharmaceutical company (AstraZeneca, Sweden) is patenting SNP tools.

Molecular biology of SNPs

A small number of papers (21% of total) and patents (19% of total) focus on the effects of point mutations or sequence alterations such as shifts in bases, insertion, or substitution of amino acids, premature stop codons, change of splicing sites, and so on in a wide variety of genes (*BRCA1* and *p53* being prominent examples).

The main research centers publishing in this area include the National Cancer Center Research Institute of the University of Tokyo, the School of Medicine at Washington University (St. Louis), and the University of California system of campuses. Authors cited in Tables 1 and 2 include Yusuke Nakamura at the Human Genome Center of the University of Tokyo, Masato Orita of the National Cancer Center Research Institute, and Pui-Yan Kwok of the School of Medicine, Washington University.

Table 6. Distribution of publications and patents focused on methods and techniques for scoring and discovering SNPs

Scientific information						Patents		
Lead public organizations[1]	**No. papers**	**Country**	**Lead companies[1]**	**No. papers**	**Country**	**Lead patentees[2]**	**No. patents**	**Country**
Washington University	34	US	Nanogen	8	US	Orchid Biocomputer	13	US
University of California System	33	US	Third Wave Technologies	8	US	Molecular Tool	13	US
						Aclara Bioscience	12	US
Stanford University	16	US	Glaxo Wellcome	6	US	Curagen	11	US
University of Pennsylvania	15	US	Affymetrix	5	US	Epigenomics AG	10	Germany
University of Uppsala	12	Sweden	Applied Biosystem	5	US	Nanogen	9	US
University of Wisconsin	11	US	Sequenom	5	US	Affymetrix	8	US
University of Wales	10	Wales	Transgenomic	5	US	Incyte Genomics	7	US
			Genometrix	4	US	AstraZeneca AB Pharmaceuticals	6	Sweden
			Asper Biotech (Tartu)	3	Estonia			
			Biosignal Packard	3	Canada	Illumina	6	US
			Bruker Dalton	3	Germany	Cornell Research Foundation (New York)	6	US
			Motorola	3	US			
			Clinical MicroSensors (Pasadena, CA)	3	US	Exact Sciences	4	US
						Glaxo	3	US
			Orchid Bioscience	3	US	Hewlett Packard	3	US
			Advion Bioscience (Ithaca, NY)	2	US	Perkin Elmer	3	US
						Promega	3	US
			Axys Pharmaceuticals	2	US	Whitehead Institute for Biomedical Research	3	US
			Caliper Technology (Mountain View, CA)	2	US			
						Wisconsin Alumuni Research Foundation	3	US
			Curagen	2	US			
			Dade Behring (Deerfield, IL)	2	US	Becton Dickison (Palo Alto, CA)	2	US
			Exelixis (S. San Francisco, CA)	2	US	DNA Sciences (Wilmington, NC)	2	US
			Incyte Genomics	2	US	Epoch Pharmaceuticals (Bothell, WA)	2	US
			Japan Science & Technology (Kyoto)	2	US			
						Hyseq (Sunnyvale, CA)	2	US
			Promega	2	US	Keygene (Wageningen)	2	Netherlands
			QIAGEN Genomics	2	US	Mosaic Technologies (Boston)	2	US
						Pangene (Fremont, CA)	2	US
						Qiagen Genomics	2	US
						Quantum Dot (Hayward, CA)	2	US
						RapiGene	2	US
						University of California system	2	US
						University of Washington (Seattle, WA)	2	US
Publications in this area as a percentage of all SNP papers/ patents[3]	**10%**			**8%**			**55%**	

[1]Table displays only those public organizations with 10 or more papers on SNPs in the area.
[2]Patentees with less than two patents in the area are not shown.
[3]Note percentages provided in Tables 5 to 8 do not sum to 100% as we have not included papers on SNPs that merely discuss their promise. Furthermore, the distribution of patents over the four SNP areas does not add up to 100% because many patents are affiliated to more than one area.
Sources: Biotechnology section of the Web of Science and Derwent Biotechnology Abstracts, The European and PCT Patent Application Bibliography Espace Access from the European Patent Office, and the Derwent World Patents Index patent database.

Leading biotechnology research companies focusing on this area include Curagen, Incyte Pharmaceuticals, and Genaissance Pharmaceuticals. They also include older, more traditional biotechnology companies, such as Biogen, Genentech, and French genetics company Genset. Agbiotech pioneer Monsanto also has patents filed in this area.

Once again, US organizations are predominant, although Japan (for instance, the University of Tokyo) and Canada (McGill University) also have academic centers that are actively publishing papers in the area.

Allele frequencies of SNPs

Papers describing differences in the frequency of certain sets of SNPs in populations either from specific geographical regions (for example, the United States) or representing specific races (for example, Caucasian) can be grouped together in a final area. From the data we analyzed, no patents have appeared in the literature corresponding to this area, perhaps because the practical implications for healthcare or agriculture are not clear or

Table 7. Distribution of publications and patents focused on the molecular biology of SNPs

Scientific information						Patents		
Lead public organizations[1]	No. papers	Country	Lead companies[1]	No. papers	Country	Lead patentees[2]	No. patents	Country
University of Tokyo	34	Japan	PPGx	3	US	Curagen	35	US
University of California system	27	US	Affymetrix	2	US	Incyte Genomics	13	US
Washington University	15	US	Curagen	2	US	Biogen	6	US
McGill University (Montreal)	13	Canada				Genaissance Pharmaceuticals	4	US
Johns Hopkins University (Baltimore, MD)	10	US				Genset	4	France
University of Cincinnati (Cincinnati, OH)	10	US				Monsanto	3	US
						Affymetrix	2	US
						Genentech	2	US
						University of California system	2	US
Publications in this area as a percentage of all SNP papers/patents[3]	**20%**			**1%**			**19%**	

[1]Table displays only those public organizations with 10 or more papers on SNPs in the area.
[2]Patentees with less than two patents in the area are not shown.
[3]Note percentages provided in Tables 5 to 8 do not sum to 100% as we have not included papers on SNPs that merely discuss their promise. Furthermore, the distribution of patents over the four SNP areas does not add up to 100% because many patents are affiliated to more than one area.
Sources: Biotechnology section of the Web of Science and Derwent Biotechnology Abstracts, The European and PCT Patent Application Bibliography Espace Access from the European Patent Office, and the Derwent World Patents Index patent database.

because assessing the genetic basis of "race" is rather controversial.

Public research organizations are much more heavily involved than private companies in studying the differences in SNPs among different populations, and the investment of the Japanese is particularly notable. For example, the group of Yusuke Nakamura, Katsushi Tokumaga, and Toshihiko Tanaka at the University of Tokyo published the most papers (12% of the total publications in this area) in our period of review (1987–2001). Public research institutions are more active than private companies in this field. As for all the other areas of SNP research, US researchers predominate in the scientific literature.

Conclusions

SNPs are an emerging field of biotechnology research and their applications are only just beginning to be documented in scientific publications and patents. US scientists and research organizations dominate this field in terms of the number of scientific papers, the most cited work, and the number of patent applications. Certainly, political initiatives taken by the US government, including the Bayh-Dole Act, have succeeded in encouraging universities to patent inventions from federally funded research. Clearly technology transfer arms of universities in the United States are far ahead of those in other parts of the world in filing patents on SNPs.

In terms of patenting and publishing, US research-based biotechnology companies are more active than any other group. Their domination is most apparent in those areas related to healthcare (including gene therapy, diagnostics, or predictive healthcare) or tools to implement such new techniques (including microarrays or microfluidics).

We found no evidence of patents or papers on this area from giant pharmaceutical firms such as Merck (Rahway, NJ), Novo Nordisk (Copenhagen, Denmark), or Aventis (Strasbourg, France). Of course, biotechnology companies are much more likely to patent their intellectual property than pharmaceutical companies, which tend to keep hold of their intellectual property as trade secrets. Thus, a lack of patent applications should not necessarily be viewed as an indication of a lack of interest in the area by the pharmaceutical sector; after all, the SNP consortium includes drug companies such as GlaxoSmithKline, AstraZeneca, Aventis, Bayer (Leverkusen, Germany, Bristol-Myers Squibb (Princeton, NJ), Novartis (Basel), Pfizer (New York, NY), Hoffman–La Roche (Nutley, NJ), and Searle (now part of Pharmacia, Skokie, IL).

Interestingly, four drug companies involved in the SNP consortium have elected not to devote efforts to publishing and filing patents in the area of SNPs: Bayer, Bristol-Myers Squibb, Pfizer, and Searle. This could be explained in several ways: by a lack of interest in investing money in controversial gene patents rather than chemical patents; by a decision to keep the information in house as a trade secret; or by an investment in generating publications or patents insufficiently large to be detectable in the specific databases we analyzed.

Our survey also reveals the activity of large companies (such as Motorola and Packard Bioscience, which is now Perkin Elmer) that specialize applying information technology to SNP research. This is because the tools of robotics, miniaturization, and information technology required for many other areas of genomics also apply to high-throughput SNP genotyping tools. Motorola (together with IBM; Armonk, NY) is collaborating with pharma companies and academic centers in the ongoing efforts of the international SNP consortium.

The presence of an established biotechnology industry in the United States and excellent research infrastructure has certainly facilitated the US lead in this area. Despite heavy investment in Europe to encourage the creation of new biotechnology startups, the data clearly show that European companies and researchers lag behind their US counterparts in SNP research. Considering the youth of the SNP field and the more encouraging recent environment for biotechnology in Europe, it is rather remarkable that the United States has established such a dominant position in this comparatively new area. Perhaps the prominent role of US centers in the human genome project also contributed to their lead in SNP research. Certainly, the meager number of European and Japanese firms in this field

Table 8. Distribution of publications focused on allele frequencies of SNPs in populations*

Lead public organizations[1]	No. papers	Country	Lead companies[1]	No. papers	Country
University of Tokyo	36	Japan	Affymetrix	4	US
Yale University (New Haven, CT)	25	US	PPGx	3	US
			Wakunaga Pharmaceuticals	3	Japan
Washington University	22	US			
Vanderbilt University	21	US			
Stanford University	17	US			
University of California system	16	US			
University of Texas system	15	US			
University of Cambridge (Cambridge, UK)	14	UK			
University of Alabama	13	US			
University of London (London)	13	UK			
University of Chicago (Chicago, IL)	12	US			
Pennsylvania State University system, PA	12	US			
University of Bonn (Bonn)	10	Germany			
University of Washington	10	US			
Publications in this area as a percentage of all SNP papers/patents[2]	**15%**			**<1%**	

[1]Table displays only those public organizations with 10 or more papers on SNPs in the area.
[2]Note percentages provided in Tables 5 to 8 do not sum to 100% as we have not included papers on SNPs that merely discuss their promise. Furthermore, the distribution of patents over the four SNP areas does not add up to 100% because many patents are affiliated to more than one area.
Sources: Biotechnology section of the Web of Science and Derwent Biotechnology Abstracts, The European and PCT Patent Application Bibliography Espace Access from the European Patent Office, and the Derwent World Patents Index patent database.

(especially compared with other areas of biotechnology) clearly indicates the poor competitiveness of Europe in this area of biotechnology.

The US biotechnology sector clearly leads in applying SNP research to healthcare and agriculture. This could be a consequence of many factors, including an established entrepreneurial culture, an excellent infrastructure for startups, world-leading expertise in SNP research, and a more favorable environment for patenting genes and gene variants at the national patenting agencies.

Further reading

Drabek, J. A commented dictionary of techniques for genotyping. *Electrophoresis* **22**, 1024–1045 (2001).

Gut, I.G. Automation I genotyping of single nucleotide polymorphisms. *Hum. Mutat.* **17**, 475–492 (2001).

De Looze, M.A., Coronini, R. & Joly, P.-B. A note on recent trends in knowledge creation and appropriation through genomics: a scientometric analysis. *Int. J. Biotechnol.* **3**, 1–2 (2001).

Saviotti, P.P., de Looze, M.A., Michelland, S. & Catherine, D. The changing marketplace of bioinformatics. *Nat. Biotechnol.* **18**, 1247–1249 (2000).

Syvanen, A.-C. Accessing genetic variation: genotyping single nucleotide polymorphisms. *Nat. Rev. Genet.* **2**, 930–942 (2001).

The International SNP Map Working Group. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature* **409**, 928–933 (2001).