Gene Meets Machine: Intellectual Property Issues in Bioinformatics

Suneeta D'Souza

Introduction

"Bioinformatics" is the new technology buzzword used to describe a rapidly developing discipline lying at the intersection of computer technology and the life sciences. These days, it is virtually impossible to pick up a popular science magazine and not find something about bioinformatics. Although there is currently no widely agreed upon definition for the term "bioinformatics", it may be defined as "the application of computing power to biological data to reveal new patterns and information below the surface of those data."[1] Bioinformatics is currently being applied to a number of scientific areas including chemistry, genomics, brain mapping, pharmacology, proteomics and structural biology.

A key aspect of these new bioinformatic technologies is the creation and maintenance of databases for the storage of biological information. In addition, stemming from the enormous growth over the past two decades in the amount of biological information available to scientists, including the success of the Human Genome Project, is the urgent need for the interpretation, integration and analysis of such information. This integrative aspect of bioinformatics is important in determining the meaning and usefulness of the data, and involves tasks such as locating particular genes in the genome sequences of a variety of organisms, creating methods to predict the structure or function of gene sequences, grouping protein sequences into families of similar or related sequences, and comparing and aligning protein sequences in order to identify evolutionary relationships.[2] A practical example of this integrative aspect of bioinformatics is using a computer program to analyze gene sequence data in order to determine how a gene related to a particular disease is turned on or off in the cell. This type information would be very useful in the realm of drug discovery and development.

The ability to use bioinformatics to organize, analyze or make predictions from data collected in the lab has had a major impact on biological research by reducing the time required to find solutions to certain biological questions.[3] With this potential to dramatically reduce the time to discover an important drug or treatment for a disease, it is not surprising that in the last few years a great deal of intellectual and financial resources, both in the private and public arenas, have been poured into the development of bioinformatic tools, particularly in the areas of genomics and proteomics.

With this influx of money and intellectual effort into the study of bioinformatics, the issue of intellectual property protection for these new technologies is an important one to consider. While existing forms of legal protection may be already be available for some bioinformatics technologies, the question arises as to whether the existing law needs to be expanded, and whether new forms of protection should be introduced, in order to ensure adequate protection for those parties involved in commercial research endeavours.

These questions become more difficult to answer when one considers the extent to which the advancement of the science of bioinformatics depends on the open and collaborative sharing of data and research tools. That is, with increased intellectual property protection, it is possible that the notion of "open science" in bioinformatics will be unattainable, and such an outcome may have many profound effects on future discoveries in bioinformatics. This paper will attempt to provide a discussion of some key intellectual property issues in bioinformatics, including a consideration of how such proprietary issues may impact the science of bioinformatics as a whole.

Intellectual Property Protection in Bioinformatics

Bioinformatics has clearly emerged as a key player in the realm of biomedical sciences and pharmaceuticals. The bioinformatics field continues to grow rapidly, and companies investing time and money in this area will likely see the need for adequate intellectual property protection for such technologies as a key issue. In fact, it is often argued that the continued advancement of the field of bioinformatics movement is highly dependent on the ability to obtain intellectual property protection, and particularly patent protection, for such innovations. That is, although intellectual property protection is often thought of in terms of the owner's ability to exclude others from using or selling a particular product or process, it has been suggested that the ultimate goal may perhaps be to attract further investment in research and development.[4] Thus, without the possibility of intellectual property protection, it has been argued that people are less likely to invest large amounts of time and money into new technologies.

Furthermore, it may be argued that intellectual property protection encourages further discoveries and advances. For instance, a patent owner gets the right to prevent others from making, using or selling an invention for the term of the patent in exchange for a full description of the invention being made available to the public, thus allowing other researchers to benefit from the advances. As such, the ability to protect intellectual property in a given area of research may be thought of as a way to maintain both the financial and intellectual momentum, which in turn stimulates further technological advances.

However, while it is possible that increased legal protections for bioinformatics inventions may result in such benefits, these may appear to come with corresponding costs as one considers the ways in which the science of bioinformatics may be impacted by such proprietary protections. It may be argued that the notion of intellectual property protection imports into the bioinformatics landscape a number of potential barriers to the notion of "open science," such as reduced access to genetic data and software tools, with the potential to seriously impede future bioinformatics discoveries. Is it possible to strike a balance between these competing interests? Before delving into such a debate, it is useful to first examine the existing modes of protection that may be used in respect of bioinformatics technologies.

Intellectual property protection for bioinformatics inventions is a largely untested area in Canadian and U.S. law.[5] However, it is clear that ability to obtain intellectual property protection in any given area of bioinformatics depends on the type of bioinformatic tool involved. Analyzing stored genomic and proteomic data requires the use of biological databases, and often involves the use of complex and sophisticated "data mining" tools such as complicated search algorithms and statistical analyses in order to sift through the mass of information. In addition, complex software applications are often used to predict the structure or interaction of biological molecules. While some of these technologies may fit into the existing framework of intellectual property law, others may fall outside of the scope of current legal protections. These issues will be discussed in the following section.

Bioinformatics Databases

Databases play a key role in bioinformatics for the collection, storage and maintenance of biological data. Databases and the information contained in them are often thought of as both products of research, as well as important precursors for future biological discovery. However, databases *per* se, existing simply as collections or arrangements of raw data, are generally not patentable subject matter. That is, a database consisting of a gene or protein sequence is patentable only to the extent that the data itself contained in the database is patentable.

Patent legislation in both Canada and the United States set out the criteria required to patent an invention. In both Canada and the U.S., the respective legislation require that for an invention to be patentable, it must be a *new* and *useful* art, process, machine, manufacture or composition of matter, or any new and useful improvement thereof.[6]

All patentable inventions, including those in the field of bioinformatics, must first fall within the statutory subject matter. The Canadian *Patent Act* states that a patent will not be granted for any mere scientific principle or abstract theorem.[7] Furthermore, the U.S. Supreme Court has identified several types of non-patentable subject matter: laws of nature and natural phenomena, and abstract ideas.[8] In addition, a mathematical formula or algorithm that is viewed in the abstract is considered unpatentable subject matter.[9]

Assuming that a given invention falls within the proper statutory subject matter and has both novelty and utility, in order to be patentable the invention must also have inventive ingenuity or "non-obviousness".[10] Generally speaking, the standard for non-obviousness requires that a patentable invention be sufficiently different from the prior art such that a person "skilled in the art" would not have considered the invention to be obvious at the time it was made. Although this test has been suggested to provide a somewhat objective standard for measuring the obviousness of a given invention, its subjectivity is inevitable and it is perhaps the most difficult test for patentability to satisfy.[11]

With this influx of money and intellectual effort into the study of bioinformatics, the issue of intellectual property protection for these new technologies is an important one to consider.

In general, matter in its naturally occurring state cannot be patented, but an isolated and purified product of nature can be.[12] Thus, it is possible for certain isolated and purified DNA sequences that are separate from the chromosomes in which they exist in nature, or DNA sequences that are created by recombinant means, to be patented if all the statutory requirements (novelty, utility and inventiveness) are met. However, raw, nonfunctional, descriptive genetic sequence information is not patentable subject matter, and the same would be the case for a database containing such information.[13] This may be considered problematic, since it is often the case that knowing the entire DNA sequence of an organism creates an information base that is more valuable than the tangible access to an isolated gene with a known function, yet the former would not be eligible for patent protection.[14]

Does the law of copyright offer protection for bioinformatics databases? In the U.S. and Canada, original compilations of data may be afforded protection under copyright law. In both the U.S. and Canada, the traditional requirement for copyrighting databases and other compilations has generally been that one need only show that the author of the work expended sufficient effort in creating the work. That is, compilations of data that merely required labour in collecting and arranging the facts contained in them have traditionally been considered eligible for copyright. This line of reasoning is referred to as the "sweat of the brow" doctrine.

However, in the leading U.S. decision of *Feist Pubs. Inc. v. Rural Tel. Serv. Co.*[15] it was held that only "creative selections and arrangements of information" are considered "compilations" for the purpose of copyright protection. A key result of this decision is that comprehensive, non-selective databases are not eligible for copyright protection in the U.S. unless they are arranged in some sort of creative manner,[16] and that no copyright exists in mere collections of data such as the alphabetical listings of the telephone book white pages.[17]

Similarly, the Canadian Federal Court of Appeal, in the decision of *Tele-Direct (Publications) Inc. v. American Business Information, Inc*., held that there was no copyright in the "in-column listings" found in the Yellow Pages directory, as the listings in question were not sufficiently inventive or creative.[18] This decision focused on the effect of the implementation in Canada of various provisions of the North American Free Trade Agreement (NAFTA) in 1993, and subsequent amendments to the *Copyright Act.* In this decision, the Court effectively interpreted such amendments as a rejection of the "sweat of the brow" doctrine, such that a degree of creativity must be present in order for a compilation to qualify for copyright. Subsequent Canadian cases have expanded on the definition of creativity.[19]

However, the case of *CCH Canadian Ltd. v. Law Society of Upper Canada*[20] has added some confusion to the issue. This case held that the decision of the Federal Court of Appeal in *Tele-Direct* did not alter the classic Anglo-Canadian test of originality, being "a work, independently created by the author and which displays at least a minimal degree of skill, judgment and labor in its overall selection or arrangement."[21] Thus, creativity was held not to be a requirement for copyright to exist in a compilation.

It should be noted, however, that the *CCH* case dealt with the issue of whether copyright exists in reported judicial decisions, head notes and case summaries, and did not specifically deal with compilations of factual data. It is a possibility that the decision in *Tele-Direct*, while not having general application in terms of the "creativity" test, may nevertheless be read as altering the traditional test for originality in the narrower area of factual compilations. This possibility was alluded to in the concurring minority reasons of Rothstein J.A.[22]

The decisions noted above are relevant to bioinformatics databases. The *Feist* decision made it clear that in the U.S., a set of facts cannot be protected merely due to the amount of resources that have been invested in the database. Although Canadian copyright law appears to afford protection to databases, it appears to be somewhat uncertain whether the "sweat of the brow" test will continue to be applied to all compilations in light of the *CCH* case, or whether some degree of "creativity" will be applied strictly to the situation of factual compilations in light of *Tele-Direct*. Since electronic databases generally have no real form of "creative" arrangement, there is a possibility that such informational databases could be left open to electronic theft. That is, there would be no protection afforded to such collections of information, irrespective of the substantial resources that may be invested in creating and maintaining such databases.

However, the "wave of the future" in bioinformatics is in data integration, or pattern recognition, and it is likely that this will be the most active area for bioinformatics patents.

In light of this potential "gap" in existing intellectual property regimes, there may be a need to look to some *new* form of protection for bioinformatics databases in order to protect the intellectual property invested in such technologies. In fact, there have been some initiatives in this regard. In 1996, a proposal regarding database protection was brought before delegates of the World Intellectual Property Organization entitled the "Treaty on Intellectual Property in Respect of Databases", however, this proposal did not receive international support at that time and further consideration will be required.[23] In the U.S., several proposals for database protection legislation have been brought before the House of Representatives.[24] These U.S. initiatives have not yet resulted in the implementation of database legislation.

However, there is currently a European Directive in place dealing with the protection of databases.[25] The European Directive is a two-tier scheme to protect both electronic and non-electronic databases. Through the Directive, copyright protection is available for databases which display the author's own intellectual creation. Furthermore, the Directive introduced a *sui generis* right protecting all databases in which a "qualitative or quantitative" investment has gone into the work, and where such investment is "substantial". The substantial investment must be made "in the obtaining, verification or presentation of the contents of the database" in order to afford protection.[26] This new database right protects database owners from the extraction or re-utilization of the contents of the database.[27] Whether similar database legislation will be implemented in Canada or the U.S. remains to be seen.

Today, there are a large number public and private online biological databases available, including a number of large databases that arose in connection with the Human Genome Project. Given the lack of formal protection available for databases in Canada and the U.S., why do so many biological databases exist? It is possible that existing means of legal protection, other than statutory protection of intellectual property, have been useful in database protection to date. For example, it is possible that contract law may be used to protect the proprietary nature of databases. That is, even where the underlying information does not qualify for statutory protection, companies may rely on user agreements and privacy agreements by which the databases may be afforded protection as trade secret or confidential information.[28]

A common strategy that has been used by bioinformatics companies, such as Incyte and Celera, has been to simply restrict access to DNA databases.[29] Encryption and physical walls have been used to restrict access to and misappropriation of data contained in databases. In fact, this type of strategy in respect of databases may be more "useful" to commercial entities than patent protection. This is because, even if one were able to meet the necessary statutory criteria for patenting the underlying genetic sequence data in a database, once a patent actually issues, there would be not be adequate protection from others seeking to use the sequence information itself, due to the statutory written description requirements which allow a full description of the invention to be made available to the public.[30]

Thus, despite the decision in *Feist,* which effectively negated the notion of statutorily protected ownership in databases, the business of bioinformatics continues to rely on license arrangements between database creators and subscribers to protect and control the access to and use of databases, and to ensure a return on investment. Such strategies, however, depend completely on secrecy, and effectively result in preventing access by users who do not have the money to subscribe for such access. Nevertheless, such licensing strategies appear to be common in the realm of bioinformatics databases.

Patentable Technologies in Bioinformatics

With the exception of databases *per se* as discussed above, bioinformatics may not have created any new form of subject matter for patent law. That is, potentially patentable computer-related technologies that have been used in the past, such as computer software, will be used in bioinformatics, but applied in new ways.[31] In addition, while databases are not themselves patentable, patent protection may be available for database-related inventions which put a sort of functional utility on the data-base itself,[32] such as new applications for databases, algorithms for extracting or mining data, and systems which include databases.[33]

In the U.S., a software invention is patentable if such invention meets the statutory requirements for all patentable inventions.[34] Historically, patents for software alone, were considered unpatentable as they were not considered to be within statutory subject matter.[35] However, the U.S. Federal Circuit Court of Appeals decision of *State Street Bank & Trust Co. v. Signature Financial Group Inc.*[36] is a case that changed the law in this area significantly, and thus may have had a substantial impact on the bioinformatics industry, even though it did not involve biology at all.

In the *State Street* case, the Court ruled that mathematic algorithms are patentable if they produce a "useful, concrete and tangible result", and that "the mere fact that a claimed invention involves inputting numbers, calculating numbers, and storing numbers, in and of itself, would not render it non-statutory subject matter unless, of course, its operation does not produce a useful, concrete and tangible result."[37] Based on this case law, software and algorithms that meet these criteria are patentable in the U.S.

In Canada, the leading court decision on the subject of software patentability was *Schlumberger Canada Ltd. v. Canada (Commissioner of Patents).*[38] In that case, it was held that software used for seismic prospecting was unpatentable due to the fact that the invention was directed at mere calculations. Unapplied algorithms and unapplied software have been considered in Canada to fall within the category of "mere scientific principle or abstract theorem" and are therefore not patentable.[39] Any tests for patentability of software have been stated in the negative, while, in contrast, in the U.S., clear positive statements on the patentability of software have been made.[40] Although the test in Canada has not been clearly stated, patent jurisprudence and Patent Office policy in Canada in respect of the patentability of algorithms and software appears to be in line with the U.S. decisions requiring a "useful, concrete and tangible result."[41]

According to this line of reasoning, software and algorithms, processes or computational techniques used to mine data from a genomic database may qualify as patentable subject matter in Canada and the U.S.. Along similar lines, algorithms used in software applications to calculate protein sequences, shapes, locations and functions may also be patentable.[42] Thus, it may be argued that current patent law generally appears to be amenable to bioinformatics patents.

What types of patentable inventions are likely to arise in the area of bioinformatics? The field of bioinformatics, in a general sense, is aimed at making sense of the wealth of many years of accumulated data, which data may come from thousands of labs located in hundreds of different countries, and all of which data may have be captured or analyzed in different ways.[43] That is, there is a large amount of *data* out there that needs to be transformed into meaningful *information*.

There is presumably great utility in being able to *store* and *integrate* data from disparate sources in order to show patterns below the surface of the data. Data storage systems *per se* are not anything new, as there are many companies who specialize in data storage, and numerous patents have issued in this respect. However, data storage in bioinformatics requires a new and different type of storage. Getting different databases to "talk" to each other (what is sometimes referred to as "interoperability") is becoming more and more essential, as users depend on different databases to fulfill different research needs.[44] Taking various data from disparate sources, and putting it all together in a coherent manner so that it can then be accessed to provide meaningful *information,* is perhaps a requirement that is unique to bioinformatics. It is not unlikely that patents will be seen in this area.[45]

However, the "wave of the future" in bioinformatics is in data integration, or pattern recognition, and it is likely that this will be the most active area for bioinformatics patents.[46] This area of bioinformatics involves pulling pieces together in order to discover something that would be impossible to see by looking at each piece on its own. To integrate genomic data in order to be able to tell something about the structure or function of a gene or protein, or how a particular gene is turned on or off, and then predict what such information will mean for a biological pathway or system and in turn for an individual's health, and finally to apply all of this information to the development of a drug or treatment, is really the key goal for bioinformatics as a whole.

Thus, while "open science" may be seen as essential in bioinformatics, intellectual property protection may be seen as inevitable.

The use of patents in the bioinformatics realm is often seen as a business tool for ensuring continued incentive for investment in bioinformatics research. In addition to protecting one's invention from copying or reverse engineering, the ability to patent a bioinformatics invention may be seen as important in providing a number of strategic benefits for those in the business of bioinformatics. Such patents may be useful in cross-licensing agreements to obtain access to a competitor's patented technology, to enhance bargaining positions in business deals, as a sales and marketing tool, and to attract funding, financing and investment.[47]

"Open Science" and Bioinformatics

While the role of intellectual property protection in bioinformatics is an important issue to consider, such consideration should not occur in a "vacuum." That is, although database protection and software patenting are not issues that are unique to the field of bioinformatics, there are some distinct aspects of the science of bioinformatics that should be recognized and considered when examining how proprietary protections should be used in bioinformatics technologies.

What makes bioinformatics unique? The growth of the bioinformatics field has been marked by a number of recent trends in the nature of biological research, including the collection and integration of data into large international databases, and the growth of research projects into large-scale endeavours beyond the scale of just a few labs.[48] As a result, bioinformatics has evolved into an inherently collaborative science. It may be argued that bioinformatics, by its very nature, requires a continuously open and collaborative process for data collection and analysis in order to subsist as a discipline.[49] Generally speaking, finding a balance between protecting proprietary interests in information, and ensuring public access to such information, has often been seen as a key issue for intellectual property law as a whole.[50] However, in the realm of bioinformatics, this balancing issue is amplified when one considers that information sharing among researchers may actually be a necessary "precursor" for future bioinformatics research and discovery.

Why is information sharing so important in bioinformatics? In the case of bioinformatics databases, it is by widespread and complete disclosure of information that data sets are likely to be enriched, verified and analyzed as more and more researchers continue to sift through such information. In the case of software and other bioinformatics research tools, it can be argued that free and open access to such tools might lead to the development of more efficient and streamlined methods of use, and to the discovery of new technologies or applications. These important aspects of bioinformatics do not appear to "fit" into the framework of proprietary research and development.

The struggle between the notion of "open science" and the notion of "proprietary" research has been strikingly apparent with the creation of large collaborative bioinformatics databases. For instance, in the case of the Human Genome Project, a vast amount of genetic sequence information was collected from a number of different sources and placed into the public arena. Subsequently, a number of commercial entities used this sequence information to obtain gene patents and create private, restricted access databases.[51]

This dichotomy of "open science versus proprietary" is often viewed as a "public versus private" debate. That is, the "proprietary" side of the coin is seen as being of interest primarily to private commercial entities, whereas the "open science" view is often associated with researchers in public institutions, universities and non-profit organizations. However, this type of characterization does not appear to accurately reflect the fact that today, due to the continuous pressure of obtaining research funding, many researchers, including those in the "public" realm, are forced to operate in a more competitive and "commercial" manner. Thus, while "open science" may be seen as essential in bioinformatics, intellectual property protection may be seen as inevitable.

In light of the above, is the concept of formal protection for bioinformatics databases a desirable one? Considering the amount of time and expense involved in creating and maintaining an online database, and the ease at which the same database can be copied by a third party, it is not surprising that there has been a recent "push" for database protection statutes. It may be argued that there is a clear link between cost recovery and database protection, in that if a database can be copied by users or competitors without any legal consequences, then the owner of the database will not be able to recover the costs of producing the database, and will have no incentive to continue to produce databases.[52] With a proper database protection scheme, it has been argued that it may be possible for science to continue to enjoy the benefit of full and free exchange of information, while at the same time maintaining the necessary incentive (that is, adequate intellectual property protection) to continue to create bioinformatics databases and other methods necessary for the progress of science.

However, there are potential problems with creating intellectual property rights in databases, many of which stem from the fact that the ordinary concept of "database" is too static to encompass the "collaborative" nature of bioinformatics databases.[53] First, if there are intellectual property rights held right down to the level of the data, downstream problems are bound to occur in relation to biologists who use and combine the data in the database, as well as for owners of technologies which underlie the database.[54] Such database protection schemes may also result in researchers in the non-commercial sector being unable to afford access to such proprietary data, or to enforce access and use restrictions on such data.[55] Furthermore, a protection scheme for databases may have the effect of causing scientists to become reluctant to share data in the same way they have done in the past, leading to a "anti-commons" problem in which vast amounts of information will be held in too many small pieces by too many people with too many rights.[56]

In addition, although database protection statutes may provide a possible solution, it is clear that any such scheme needs to be properly limited in its application and scope in order to provide a proper balance between the protection of information contained in databases, and the access by the public to such information. For example, as mentioned above, the European Directive provides protection for databases that show a "qualitative or quantitative" investment, and prohibits extraction of information from databases that is "qualitatively or quantitatively substantial." Such subjective language is bound to, and has indeed, caused problems in determining what is protected, and what is not,[57] and may lead to a form protection that is overly broad in its scope.

It may be argued that what has been created by the European Directive is a broad-based right without sufficient balancing mechanisms needed to allow for adequate access to data.[58] It has been suggested that such a balance might be achieved by "carving out" legal rights of access and use to scientific databases in the case for not-for-profit researchers and educators, and implementing mandatory licensing provisions in order to avoid the situation of such data being "monopolized."[59] It is possible that a more carefully balanced database protection scheme that incorporates such types of provisions could provide for free and open access to data, while still allowing for cost recovery through licensing in the commercial realm.

In addition to bioinformatics databases, the issue of "open science" is also relevant to the development of bioinformatics software tools. As with data collection and analysis, it may be argued that bioinformatics software development should occur in a transparent, collaborative and open access environment in order to achieve its maximum potential. While the closed-source, proprietary model is a well-known model for software development in general, such a model applied to bioinformatics would not allow for the collaborative tool sharing and development of bioinformatics software, possibly leading to a reduction in potential discoveries.[60] This is in part because, even if such software tools were liberally licensed, research institutions which have limited access to funding often cannot afford to access to such proprietary software tools. This inaccessibility by researchers could possibly lead to the stifling of successive and cumulative improvements to bioinformatics software, and may have the effect of thwarting the process of biological problem solving. [61]

In contrast to the closed, proprietary model, a number of bioinformatics collaborations over the Internet have involved the use of "free" or "open-source" software.[62] In the open-source model, the notion of "copy-left" is often used to ensure broad access to the software. "Copy-left" is a concept that uses the rights granted by copyright law in order to benefit the public, rather than the holder of the copyright, by requiring users to make all software alterations, modifications and applications available to subsequent public users. The use of copy-left has primarily been to prevent the private commercialization of derivatives of the copy-left work.[63]

It is obvious that the "open-source" model fits nicely into the "open-science" model. But is it possible that open-source software could be an option for commercial entities seeking to profit in the bioinformatics field? Perhaps the incentive for software developers to use the open-source model might arise through the use of a two-tiered licensing system, in which an open-source, copy-left license applied to non-commercial users, with the additional ability for the software developer to license those same tools to private companies seeking to protect and commercialize them.[64] Such a "dual licensing" model may be but one possible mode for achieving a balance between "open science" and intellectual property protection in bioinformatics.

Conclusion

The blending of biology and computers makes bioinformatics a unique discipline that is clearly here to stay. Bioinformatics is one of the fastest growing areas of all life science based markets. While some degree of intellectual property protection may play a role in ensuring continued investment and discovery in this area, any augmentation of the current law should recognize the need for the free and open sharing of information, which is of particular importance in the field of bioinformatics

Databases of biological data, which are not patentable *per se*, may find some protection under the law of copyright in Canada, but the same would not be the case in the U.S. Such databases may benefit from some new form of formal database protection, however, any such statutory protection should be adequately limited in its scope, and should contain provisions to ensure the liberal licensing of and access by researchers to the contents of scientific databases.

With the exception of databases *per se*, it is possible that other types of bioinformatics inventions may fit within the current framework of patent law. Patents may provide strategic benefits that may have a significant impact on future investment and technological advances in the bioinformatics realm. However, in the case of bioinformatics software and research tools, it is possible that traditional "closed" proprietary models may not prove to be as effective as alternative licensing models that incorporate the "open source" concept.

The key goal for bioinformatics will be to integrate a vast amount of biological data in order to recognize patterns which will assist in gaining important knowledge about health and disease, and which will in turn assist in finding ways to *improve* health and *treat* disease. With such an important goal in mind, it is crucial that intellectual property protections are sufficiently balanced with the need to ensure the open flow of data and preserve the collaborative nature of the science. It is this balanced approach in intellectual property regimes that will be crucial in maintaining the necessary elements required to drive the science and technology of bioinformatics forward.

Suneeta D'Souza is an associate lawyer with Peterson & Purvis, Lethbridge, Alberta. *The author acknowledge's the firm's support and comments for this paper.*

Notes

1. Tom Meyers, "Patenting and Financing Bioinformatics Inventions" (Symposium on Bioinformatics and Intellectual Property Law, April 27, 2001) (2002) 8:1 B.U. J. Sci. & Tech. L. 157, online: available at Boston University School of Law <http://www.bu.edu/law/scitech/volume8/Panel1.pdf>.
2. University of Buffalo Center of Excellence in Bioinformatics, "Introduction to Bioinformatics", online: University of Buffalo Centre of Excellence in Bioinformatics <http://www.bioinformatics.buffalo.edu/current_buffalo/primer.html>.
3. M. Chow & D. Fernandez, "Intellectual Property Strategy in Bioinformatics" (Paper presented to the First Virtual Conference on Genomics and Bioinformatics, North Dakota State University, October 15-16, 2001), online: First Virtual Conference on Genomics and Bioinformatics <http://midas-10.cs.ndsu.nodak.edu/bio/papers.html>.
4. Kate H. Murashige, "Genome Research and Traditional Intellectual Property Protection – A Bad Fit?" (1996) 7 Risk: Health, Safety & Environment 231, online: Franklin Pierce Law Centre

<http://www.piercelaw.edu/risk/vol7/summer/murashig.htm>.

5. As there has been very little case law in the general area of biotechnology and patent law in Canada, Canadian courts have considered U.S. case law and practice in the development of its patent law in the area of biotechnology, however such consideration has been cautious. See *Harvard College v. Canada (Commissioner of Patents)*, [2002] 4 S.C.R. 45, 2002 SCC 76.
6. *Patent Act*, R.S.C. 1985, c. P-4, s. 2; 35 U.S.C. § 101.
7. *Patent Act, ibid.*, s. 27(8).
8. *Diamond v. Diehr*, 450 U.S. 175 at 185 (1981).
9. *AT&T v. Excel Communications, Inc*., 172 F.3d 1352 (Fed. Cir. 1999).
10. See *Patent Act*, *supra* note 6, s. 28.3; 35 U.S.C. § 103.
11. J.R. Rudolph, "Patentable Invention in Biotechnology" (1996) 14:1 Biotechnology Advances 17 at 28.
12. Rebecca Eisenberg, "Molecules v. Information: Should Patents Protect Both?" (Symposium on Bioinformatics and Intellectual Property Law, April 27, 2001) (2002) 8:1 B.U. J. Sci. & Tech. L. 190, online: Boston University School of Law <http://www.bu.edu/law/scitech/volume8/Panel3.pdf>.
13. *Ibid*. See U.S. Patent and Trademark Office, Utility Examination Guidelines, 66 Fed. Reg. 1092, 1094 (2001), online: GPO Access <http://www.access.gpo.gov/su_docs/fedreg/a010105c.html>.
14. Eisenberg, *ibid*.
15. 499 U.S. 340 (1991) [*Feist*].
16. Dennis S. Karjala, "Data Protection Statutes and Bioinformatic Databases" (Symposium on Bioinformatics and Intellectual Property Law, April 27, 2001) (2002) 8:1 B.U. J. Sci. & Tech. L. 171, online: Boston University School of Law <http://www.bu.edu/law/scitech/volume8/Panel2.pdf>. See also James G. Silva, "Copyright Protection of Biotechnology Works: Into the Dustbin of History?" (2000) B.C. Intell. Prop. & Tech. F. 012801, online: Boston College Law School <http://www.bc.edu/bc_org/avp/law/st_org/iptf/articles/index.html>.
17. Wendy Gordon, "Data Protection Statutes and Bioinformatic Databases" *ibid*.
18. (1997), 76 C.P.R. (3d) 296, 37 B.L.R. (2d) 101 (F.C.A.) [*Tele-Direct*].
19. See Industry Canada, *Database Protection and Canadian Laws* by Robert Howell (Ottawa: Industry Canada, 1998), online: Strategis <http://strategis.ic.gc.ca/pics/ip/databe.pdf>.
20. [2002] 4 F.C. 213 (C.A.) [*CCH*].
21. *Ibid* at 250. See also David Freedman, "Revising Canadian Database Protection: What Lessons from Europe?" (2002) 82 Can. Bar Rev. 563 at 601.
22. *CCH, ibid.* at 315.
23. See Catherine Colston, "Sui Generis Database Right: Ripe for Review?" (2001) 3 J. Info. L. & T., online: Journal of Information Law & Technology <http://elj.warwick.ac.uk/jilt/01-3/colston.html>. The draft treaty can be accessed online: World Intellectual Property Organization <http://www.wipo.org/eng/diplconf/6dc_sta.htm>.
24. These include U.S., Bill H.R. 3531, *Database Investment and Intellectual Property Antipiracy Act of 1996*, 104th Cong., 1996 (Moorhead) and U.S., Bill H.R. 2652, 105th Cong., *Collections of Information Antipiracy Act*, 1997 (Coble), among others. The bills and their status are available online: U.S. House of Representatives <http://www.house.gov/house/Legproc.html>.
25. EC, *Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases* [1996] O.J. L.77/20, online: European Union <http://europa.eu.int/ISPO/infosoc/legreg/docs/969ec.html>.
26. *Ibid.*, Art. 1(2).
27. *Ibid.,* Art. 7(1).
28. Stephen M. Maurer, "Appendix C: Raw Knowledge: Protecting Technical Databases for Science and Industry" in *Proceedings of the Workshop on Promoting Access to Scientific and Technical Data for the Public Interest: An Assessment of Policy Options, National Academy of Sciences* (National Academy of Sciences, 1999), online: National Academies Press <http://www.nap.edu/html/proceedings_sci_tech/appC.html>.
29. Eisenberg, *supra* note 12.
30. *Ibid*.
31. Meyers, *supra* note 1. It should be noted that while copyright would also be an option in the case of software, this right would only protect the author from direct copying of the source code. This is considerably different than the rights conferred by a patent, in which the owner can exclude others from making, using or selling a tangible product.
32. *Ibid*.
33. T.C. Meyers *et al*., "Patent protection for protein structures and databases" (2000) 7:11 Nature Structural Biology 950 at 951.
34. Charles Vorndran & Robert L. Florence, "Bioinformatics: Patenting the Bridge Between

Information Technology and the Life Sciences" (2002) 42:1 IDEA 93 at 108.

35. *Ibid*.
36. 149 F.3d 1368, 1373 (Fed. Cir. 1998) cert.denied, 525 U.S. 1093 (1999) [*State Street*].
37. *Ibid* at 1374. In addition, this decision ended the "business method exception" by holding that business methods are indeed patentable subject to the same legal requirements for patentability that apply to any other process or method. This may be relevant to bioinformatics inventions involving business methods. See also *AT&T v. Excel Communications, Inc*., 172 F. 3d 1352 (Fed. Cir. 1999), which held that transformation of data is sufficient, but not required, for algorithms to fall within statutory subject matter.
38. (1981), 56 C.P.R. (2d) 204 (F.C. T.D.).
39. Christopher Van Barr, "Decoding software patentability in Canada" (Paper presented at Infonex Conference "Patent Practice Update" Toronto, June 18, 2001) online: Gowlings <http://www.gowlings.com/resources/publications.asp?showWhat=696>.
40. Van Barr, *ibid*., suggests that the test in Canada should be "whether the software invention has a practical application in industry, trade or commerce."
41. *Ibid*. See *Re Motorola Inc. Patent Application No. 2085228* (1998), 86 C.P.R. (3d) 71 (P.A.B.).
42. Richard Raysman & Peter Brown, "Bioinformatics 101: Basics for the Computer Lawyer" (2001) 225:88 N.Y.L.J. 3.
43. Meyers, *supra* note 1.
44. Ken Howard, "The Bioinformatics Gold Rush" (2000) 282:1 Scientific American 58 at 63. See also Teresa K. Attwood, "The Babel of Bioinformatics" (2000) 290:5491 Science 471.
45. Meyers, *supra* note 1.
46. *Ibid*.
47. James G. Gatto, "Bioinformatics Patents – Challenges and Opportunities" *Intellectual Property Advisory* (November 2001), online: Mintz Levin Cohn Ferris Glovsky and Popeo PC <http://www.mintz.com/publications/detail/152/Intellectual_Property_Advisory_Bioinformatics_Patents_Challenges_and_Opportunities/>.
48. OECD Working Group on Neuroinformatics, "Neuroscience Data and Tool Sharing: A Legal and Policy Framework for Neuroinformatics" (2003) 1 Neuroinformatics 149, online: Neuroinformatics.com <http://www.cs.mu.oz.au/~pde/writing/oecd-neuroinf.pdf>.
49. Paul A. David, "Will Building 'Good Fences' Really Make 'Good Neighbors' in Science?" (2001), online: Stanford University Economics Department <http://www-econ.stanford.edu/faculty/workp/swp01005.pdf>.
50. Freedman, *supra* note 21 at 567.
51. OECD Working Group, *supra* note 48.
52. National Research Council, *A Question of Balance: Private Rights and the Public Interest in Scientific and Technical Databases* (Washington, D.C.: National Academies Press, 1999), online: National Academies Press <http://books.nap.edu/html/question_balance/index.html>.
53. David, *supra* note 49.
54. Karjala, *supra* note 16.
55. National Research Council, *supra* note 52.
56. Karjala, *supra* note 16.
57. See P. Bernt Hugenholtz, "The New Database Right: Early Case Law from Europe" (Paper presented to the Ninth Annual Conference on International IP Law and Policy, Fordham University School of Law, New York, April 19-20, 2001), online: Institute for Information Law <http://www.ivir.nl/publications/hugenholtz/fordham2001.html>.
58. Freedman, *supra* note 21 at 612.
59. National Research Council, *supra* note 52.
60. OECD Working Group, *supra* note 48.
61. Carole Ganz-Brown, "Electronic Networks for International Research Collaboration: Implications for Intellectual Property Protection in the Early Twenty-First Century" in Orville Vernon Burton, ed., *Computing in the Social Sciences and Humanities* (Urbana: University of Illinois Press, 2002), online: University of Illinois Press <http://www.press.uillinois.edu/epub/books/burton/>.
62. For example, see <http://bioinformatics.org>.
63. OECD Working Group, *supra* note 48.
64. *Ibid*.