

CHAPTER 4

BIOTECHNOLOGY RESEARCH DATA – SCOPE OF PROTECTION UNDER COPYRIGHT REGIME AND DATABASE DIRECTIVE

4.1 INTRODUCTION

Databases have always been a subject matter of intense commercial use and social utility. However protection of databases remained a subject matter of continuing debate. The confusion is whether copyright legislation is adequate to protect database or a separate mechanism is necessary for the protection of the rights in it. Another concern regarding protection of database is the adequacy of incentives for protection and to ensure public access to the information in them. One of the significant aspects is how to protect database under copyright law. Under copyright law a database can be protected if it falls under the category of ‘compilation’. In the US copyright Act of 1976, a compilation is a work formed by the collection and assembling of pre-existing materials or data that are selected, coordinated or arranged in such a way that the resulting work as a whole constitutes an original work of authorship.¹

Research databases are the most valuable part of biotechnology research. Biotechnology, especially genomic research, attributes enormous value to the gene sequence and related databases. Protection of the research data is the most crucial element as far as an ongoing research is concerned. Researchers in biotechnology are adopting various methods of protecting their research data. Patent protection is the most common form of protection opted by most of the researchers. However the recent judicial trend is not favouring the patent protection. The patent system insists on the substantial utility criteria to establish the patentability. Further the raw data do not make the researcher entitled for protection unless the proteins are isolated and purified in order to establish novelty and substantial utility criteria.² These hurdles

¹ 17 U.S.C. Section 101 of the copyright Act provides for various definitions

² 35 U.S.C. section 101, explains patentability of gene which reads as:

Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent there for, subject to the conditions and requirements of this title.' Thus, an inventor's discovery of a gene can be the basis for

have placed researchers in a highly vulnerable situation. Application of copyright or a *sui generis* form of protection is the next option available to the researchers. The scope of protection under these two concepts is again placed at a very fragile situation because a strong mechanism to safeguard non-creative or non-original is yet to evolve internationally except the initiative by the EU Database Directive and the EU Copyright Directive.

Database constitutes one of the oldest forms of authorship protected under any law. However, databases or compilations have always held an insecure position within the copyright law. In order to attract copyright protection to a database, a work needs to be a compilation and it must be “an original work of authorship.”³ It must also involve the selection, coordination or arrangement of data as a basic criterion. In other words, the originality requirement is applied to compilations similar to other works of copyright.⁴

4.2 A BRIEF HISTORICAL PERSPECTIVE OF DATABASE PROTECTION

Advancement of technology has a direct implication on the copyright landscape. The emergence and expansion of new information and communication technologies have radically changed the way in which literary and artistic works are created and diffused. New and old creations can now be digitalised and diffused very easily. Rapid developments of the technologies help the creation of databases as well. The new technological developments help the data collection, organisation, reproduction and dissemination much easier. A brief analysis of the historical aspects of databases is necessary to understand the concepts and philosophy of intellectual property protection in databases.

There is a lack of a clear-cut characterisation of the term database. According to the European Database Directive the term “database” should be understood to

a patent on the genetic composition isolated from its natural state and processed through purifying steps that separate the gene from other molecules naturally associated with it.

³ 17 U.S.C. section 101 explaining a compilation reads as:

A “compilation” is a work formed by the collection and assembling of pre-existing materials or of data that are selected, coordinated, or arranged in such a way that the resulting work as a whole constitutes an original work of authorship. The term “compilation” includes collective works.

⁴ Ibid. also see; Dov S. Greenbaum, *The Database Debate: In Support of an Inequitable Solution*, 13 Alb. L.J. Sci. & Tech. 431., p 439(2003)

cover collections of independent works, data or other materials which are systematically or methodically arranged and can be individually accessed.⁵ Further the recital 17 of the directive further explains the database as to include; literary, artistic, musical or other collections of works or collections of other material such as texts, sound, images, numbers, facts and data.⁶

The term database was originally coined by the U.S. military to signify a collection of data shared by end-users of a computer system during 1962.⁷ The word, database has subsequently been appropriated by other industries to indicate any compilation of data independent of the storage medium. The database is usually organised and utilised by a database management system (DBMS). The DBMS is the part of the database that provides a bridge between the data and the end user through the use of software such as Structured Query Language (SQL).⁸ In simple terms a database is an organised and indexed collection of information that allows users to access and organise mixed data in an efficient manner.⁹ A simple example of a database is the white pages of a telephone directory. White pages are organized alphabetically and indexed by the last name.

The need of protection of data as such was realised and initiated in Europe. Protection of data is not only intended to cover raw data but also to standardise the handling of information about people and their practices. The rapid growth of computing technology in Europe has developed worries regarding the free flow of data. The early initiatives began in Europe during the 1950s when Europe established the Human Rights Commission.¹⁰ After that the European convention for protection

⁵ The Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases. The EC directive, 96/9/EC Article 1(2) defines a database as:
“a collection of independent works, data or other materials arranged in a systematic or methodical way and individually accessible by electronic or other means”

⁶ The EC directive, 96/9/EC Recital (17) expands on the definition of database as:
Whereas the term ‘database’ should be understood to include literary, artistic, musical or other collections of works or collections of other material such as texts, sound, images, numbers, facts, and data; whereas it should cover collections of independent works, data or other materials which are systematically or methodically arranged and can be individually accessed; whereas this means that a recording or an audiovisual, cinematographic, literary or musical work as such does not fall within the scope of this Directive;

⁷ The Oxford English Dictionary 261 (2d ed. 1989).

⁸ Dov S. Greenbaum, *The Database Debate*, p.441

⁹ Ibid.

¹⁰ Rosemary Jay and Angus Hamilton, *Data Protection - Law and Practice* (London, Sweet and Maxwell, 1999) p.1

of rights and fundamental freedom was adopted in the same year. It was suggested that the data movement might be controlled on human rights ground. While imposing restrictions it was also feared that the trade among the European countries could be hampered if information could not flow freely. This fear has led to the development of early data protection standards in Europe.¹¹

The early concern in Europe was regarding the protection of private and personal data. The 1960 s witnessed the development of various privacy bills basically to secure personal data in England. The sex discrimination Act of 1971 and the Race relation Act of 1968 manifested personal privacy. Various privacy bills were introduced to the UK parliament during 1960s and early 70s. The government appointed Kenneth Younger to chair a committee on privacy.¹² The committee filed the report in 1972. The Younger report restricted itself to computing in the private sector. Younger report highlighted a concern regarding the widespread availability of personal information about individuals via the medium of computers. The report covered both physical privacy and also information privacy. The later part of his study impacted the development of data protection standards.¹³

Three years after the younger report, the government published two white papers entitled ‘computers and privacy’ and ‘computers: safeguard for privacy’ which dealt with the use of computers in the public sector. Following the announcement of the white paper the government announced the setting up of the Data Protection committee under the chairmanship of Sir Norman Lindop in 1976. Lindop filed the ‘report of the committee on data protection’ recommending the legislation covering the private and public sectors.¹⁴ Another major initiative was by the Organisation for Economic Co-operation and Development (OECD) which adopted the guidelines for protection of privacy and transborder flow of personal data in Paris 1981. Subsequently many developments took place in the area of personal data privacy.

¹¹ Rosemary Jay and Angus Hamilton, *Data Protection - Law and Practice*, p.2

¹² Ibid.

¹³ Ibid.

¹⁴ Ibid. at p.3

Finally a comprehensive mechanism to protect the databases evolved when the European Union adopted the EU database directive in 1995.¹⁵

4.3 CLASSIFICATION OF DATABASES

Databases are basically a compilation of works, data or collection of facts arranged in a systematic way. The definition given to database is that it is a collection of independent works, data or other materials arranged in a systematic or methodical way and individually accessible by electronic or other means.¹⁶ The copyright protection of databases depends up on how the originality principle can be applied to the contents of the databases. Mere facts are not protected according to the originality concept of copyright. However the order and arrangements having certain level of creativity can be a subject matter of copyright protection. The level of creativity or originality is the deciding factor for the databases to attract copy right protection

Based on the level of creativity, the databases can be classified as creative and non-creative. In order to be entitled for copyright protection a database must be a compilation of data. Creative databases are those which meet the copyright criteria of originality and creativity. Traditional copyright requires that the compilation must meet the “modicum of originality or creativity in the selection or arrangement of data.”¹⁷ This requirement may be too stringent and may effectively exclude many of the most important commercial and scientific databases from copyright protection.¹⁸ The Developments in computer software and digital technology have led to the creation of many commercially and scientifically important databases which are excluded from copyright protection in most jurisdictions because they are “non-creative.”¹⁹

Biotechnology and bioinformatics databases are commercially and scientifically significant. However they are excluded from copyright protection because of their ‘non-creative’ and ‘non-original’ nature. The reason is the lack of

¹⁵ Rosemary Jay and Angus Hamilton, *Data Protection - Law and Practice*, p.5

¹⁶ Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases.

¹⁷ Michael J. Bastian, *Protection of "Noncreative" Databases: Harmonization of United States, Foreign and International Law*, 22 B.C. Int'l & Comp. L. Rev. 425, p.425 (1999)

¹⁸ Ibid.

¹⁹ Ibid.

minimum requirement of originality and creativity in the selection, coordination or arrangement of their contents envisaged by the copyright law.²⁰

4.3.1 ORIGINAL AND CREATIVE DATABASES

Original and creative database are eligible for copyright protection. Originality means that the selection and arrangements of the contents of the database must constitute intellectual creations. As explained earlier original and creative databases are generally protected under copyright law as compilations which mean that they must be collection and assembling of pre-existing materials and they must be selected in such a way that the work as a whole must constitutes an original work of authorship.²¹

Copyright protection of original databases is well established and harmonized through international treaties. The Berne Convention proposed to protect original databases including compilation. Another most important agreement TRIPS also protects only the structure of the databases. It does not extend to the data or material itself.

4.3.2 NON- ORIGINAL AND NON-CREATIVE DATABASE

The term noncreative and non original database is used to refer to databases that do not contain a minimum level of originality, creativity or expression of personality in the selection, coordination or arrangement of their contents and hence, are non-copyrightable. But sometimes they hold enormous commercial and scientific vale. The non creative databases may include scientific, technical, or educational application-databases. Any data including personal information, collection of information and research outputs hold enormous commercial potential. According to traditional copyright concepts, the database must have “modicum of creativity”. It is irrelevant even if the database creator has spent huge amount in making the database if the work lacks minimum amount of creativity.²²

²⁰ Michael J. Bastian, *Protection of "Noncreative" Databases*, p.429 (1999)

²¹ 17. U.S.C. Section. 101

²² Dov S. Greenbaum, *The Database Debate*, p.444

Whereas, according to the “sweat of the brow doctrine”, creativity is not an essential condition for establishing originality. A database is original if the author has invested time, money skill and labour in creating the database; and it means that if substantial investment is made to create the database, the database would be entitled to copyright protection. The earlier views, prior to *feist* case was the one recognising the “sweat of the brow” doctrine. Copyright protection was granted to industrious collection of databases. The courts interpreted the originality of a work as if it was not copied from others.²³ *Jeweler’s Circular Publishing* case puts down the justification for ‘sweat of the brow’ doctrine as:²⁴

“The right to copyright a book upon which one has expended labour in its preparation, does not depend upon whether the materials which he has collected consist or not of matters which are *publici juris*, or such materials show literary skill or originality either in thought or in language, or anything more than industrious collection. The man who goes through the streets of a town and puts down the names of each of the inhabitants, with their occupations and their street number, acquires material of which he is the author.”²⁵

The *feist* turned down the legal position recognising “sweat of the brow doctrine.” The factual data could be protected only if the selection and arrangement rendered amounts to an original creative work. This means that legal protection would only be extend to copying of selection and arrangement aspects of databases. This situation has given rise to many concerns among the database industries. The need for protection of databases remained as an unresolved issue internationally. The first positive attempt to protect non-original non-creative data base is by the European Union. The European Union Database Directive provides for protection of non creative databases, if the creator of the databases establishes that there has been

²³ Justin Hughes, *How Extra-Copyright Protection of Databases Can Be Constitutional*. 28 U. Dayton L. Rev. 159., p.163(2003)

²⁴ *Jeweler’s Circular Publg. Co. v. Keystone Pub. Co.*, 281 F. 83, 87-88 (2d Cir. 1922);

²⁵ *Ibid.*

qualitatively and quantitatively substantial investment in obtaining, verification and presentation of the contents.²⁶

4.4 PROTECTION OF DATABASES UNDER COPYRIGHT REGIME

Historically, protection of collections of information has always been considered a branch of copyright law. Under copyright law, protection is granted to any original work of authorship fixed in any tangible medium of expression. The originality is the prime requirement of copyright protection which means that the ideas or facts are not protected even if they are already expressed in another work. Protection is only granted for expression of ideas and facts. There is no ambiguity regarding the protection of traditional subject matters such as literary, artistic works and so on. However the confusion developed while applying it to more elaborate and complicated subjects like compilation of facts constituting databases. The vagueness continued till the landmark judgement by the United States Supreme Court in *Feist v. Rural*²⁷ clarified the originality and minimum quantum of creativity requirements for copyright in databases including compilation of facts.

The database protection has posed some major legal issues. If monopoly is granted to a database comprising of common facts, it will hinder the ability of others to work upon the facts contained in that database. This will stifle the research relating to those particular facts. The issue further complicates the scenario when the nature of database varies. The copyright protection of database which is creative with original compilation and the non-creative original research data are having different legal status. In the first case, the databases with original compilation are entitled for copyright protection and secondly the non-creative database holds no copyright protection over the entire work because the selection and arrangement of the

²⁶ Article 7 section 1 of the Directive 96/9/EC reads as:

“Member States shall provide for a right for the maker of a database which shows that there has been qualitatively and/or quantitatively a substantial investment in either the obtaining, verification or presentation of the contents to prevent extraction and/or re-utilization of the whole or of a substantial part, evaluated qualitatively and/or quantitatively, of the contents of that database.”

²⁷ *Feist Pubs. Inc. v. Rural Tel. Serv. Co.* 499U.S. 340 (1991)

information lacks the minimum quantum of originality requirement under traditional copyright law²⁸.

4.4.1 DATABASE LANDMARK - FEIST PUBLICATIONS V. RURAL TELEPHONE SERVICE CO.- THE DISTRICT COURT OF KANSAS

*Feist Pubs. Inc. v. Rural Tel. Serv. Co.*²⁹ set the standards for the legal protection of databases across the world. Article I, section 8, clause 8, of the US constitution is known as copyright clause.³⁰ According to the US copyright Act 1976, a compilation is not copyrightable per se,³¹ but is copyrightable only if its facts have been “selected, coordinated or arranged in such a way that the resulting work as a whole constitutes an original work of authorship.”³² Raw databases do not form a copyrightable subject matter. Even the database which is a compilation though copyrightable, receives only limited protection. The copyright does not extend to facts contained in the compilation.³³ These positions were clarified in this landmark judgment.

The case was first considered by the district court of Kansas. Rural Telephone Service Company was a certified public utility providing telephone service to several communities in Kansas district and it published a typical telephone directory, consisting of white pages and yellow pages. The white pages list, in alphabetical order, the names, addresses and telephone numbers of Rural's telephone subscribers.

²⁸ Edward J. Baba, *From Conflict To Confluence: Protection Of Databases Containing Genetic Information*, 30 Syracuse J. Int'l L. & Com. 121, p.123 (2003)

²⁹ *Rural Telephone Service Co. v. Feist Publications*, 737 F. Supp. 610,622 (Kan 1990)

³⁰ Article 1 Section 8 clause 8 of the U.S. Constitution reads as:

To promote the Progress of Science and useful Arts, by securing for limited Times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries

³¹ 17 U.S.C. Section 102(a), reads as:

Copyright protection subsists, in accordance with this title, in original works of authorship fixed in any tangible medium of expression, now known or later developed, from which they can be perceived, reproduced, or otherwise communicated, either directly or with the aid of a machine or device.

³² 17 U.S.C. Section 102(b) reads as:

In no case does copyright protection for an original work of authorship extend to any idea, procedure, process, system, method of operation, concept, principle, or discovery, regardless of the form in which it is described, explained, illustrated, or embodied in such work

³³ 17 U.S.C. Section 103(b) reads as:

The copyright in a compilation or derivative work extends only to the material contributed by the author of such work, as distinguished from the preexisting material employed in the work, and does not imply any exclusive right in the preexisting material. The copyright in such work is independent of, and does not affect or enlarge the scope, duration, ownership, or subsistence of, any copyright protection in the preexisting material.

Feist Publications Inc. was a publishing company that specialised in area-wide telephone directories covering a much larger geographic range than directories such as *Rural's*. Both the companies issued their directories free of cost to the customers. However they received majority of their income from yellow page advertisement. *Feist* planned to publish an area-wide directory covering eleven different telephone service areas and approached *Rural* for licensing of its white page listings to be included in the directory. *Rural* refused and *Feist* extracted the listings it needed from *Rural's* directory without *Rural's* consent.³⁴

Rural sued *Feist* for copyright infringement. *Rural* asserted that *Feist's* employees were obliged to travel door-to-door or conduct a telephone survey to discover the same information for them. *Feist* responded that such efforts were economically impractical and, in any event, unnecessary because the information copied was beyond the scope of copyright protection. The District Court of Kansas granted summary judgment to *Rural*, holding that telephone directories are copyrightable³⁵. Also the Court of Appeals for the Tenth Circuit affirmed the judgement given by the district court.

4.4.2 RURAL TELEPHONE SERVICE CO. V. FEIST PUBLICATIONS - SUPREME COURT

The appeal to the Supreme Court of the United States in *Rural Telephone Service Co. v. Feist Publications*³⁶ considered the case on a different perspective and discussed the issues in detail. The case was discussed on the basis of two well-established propositions. The first is that facts are not copyrightable and secondly the compilations of facts generally are. This means that there can be no valid copyright in facts that are universally understood, and as discussed in the case there can be no protection of ideas and facts. It is to be noted that many compilations consist of nothing but raw data. The whole data comprises factual information not accompanied

³⁴ *Rural Telephone Service Co. v. Feist Publications*, 737 F. Supp. 610,622 (Kan 1990)

³⁵ *Ibid.*

³⁶ *Feist Pubs. Inc. v. Rural Tel. Serv. Co.* 499 U.S. 340,378 (1991)

by any original written expression. In spite of that, copyright law contemplates that the compilations of facts are potentially within its scope of protection.³⁷

According to the Supreme Court, there was no doubt that *Feist* took from the white pages of *Rural's* directory a substantial amount of factual information including the names, towns, and telephone numbers of 1,309 of *Rural's* subscribers. Even though they did not copy all the information, it constituted a valid copyright infringement. In order to establish infringement, two elements must be proved: (1) ownership of a valid copyright, and (2) copying of original constituent elements of the work.³⁸ The question was whether *Feist*, by copying the names, towns, and telephone numbers from *Rural's* white pages, copied anything that was "original" to *Rural*. Holding that the raw data did not satisfy the originality requirement, *Rural* might have been the first to discover and report the names, towns, and telephone numbers of its subscribers, but this data did not owe its origin to *Rural*. According to the Supreme Court the information contained in white pages were uncopyrightable facts. They existed even before *Rural* had reported them and would have continued to exist if *Rural* had never published a telephone directory. By referring to the originality requirement the court ruled that, names, addresses, and telephone numbers the plaintiff could never be treated as the authors.³⁹

The question that remained is whether *Rural* selected, coordinated, or arranged these uncopyrightable facts in an original way. The standard of originality was low, but it did exist. According to the Court, the Constitution mandated some minimal degree of creativity and an author who claimed infringement must prove the existence of intellectual production of thought and conception. The selection, coordination, and arrangement of *Rural's* white pages did not satisfy the minimum constitutional standards for copyright protection.⁴⁰

The Supreme Court reversed the judgment of the appellate court and found that,⁴¹

³⁷ Ibid.

³⁸ Ibid.

³⁹ Ibid.

⁴⁰ Ibid.

⁴¹ Ibid.

“Because *Rural’s* white pages lack the requisite originality, *Feist’s* use of the listings cannot constitute infringement. This decision should not be construed as demeaning *Rural’s* efforts in compiling its directory, but rather as making clear that copyright rewards originality not effort. As this Court noted more than a century ago, “great praise may be due to the plaintiffs for their industry and enterprise in publishing this paper, yet the law does not contemplate their being rewarded in this way.””

The intellectual property rights in databases were made clear with the *Feist Publications v. Rural Telephone Service Co.* The Supreme Court held that the essence of copyright was “not to reward the labour of authors but to promote the Progress of Science and useful Arts.” To grant a monopoly over common facts would hinder the ability of others to build upon the facts contained in a work, thereby stifling the progress of science and arts.⁴² The result of the decision is that a compiler of facts in a non-creative database holds no copyright protection over the entire work since the selection and arrangement of the information lack the minimum quantum of originality.

Commending on the concept of originality, the court noted that originality, as the term is used in copyright, means, “that the work is independently created by the author (as opposed to copying from other works), and that it possesses at least some minimal degree of creativity.” further, no author may copyright his ideas or the facts he narrates, because the facts do not owe their origin to an act of authorship. Court further quoted that;⁴³

The distinction is one between creation and discovery: The first person to find and report a particular fact has not created the fact; he or she has merely discovered its existence...one who discovers a fact is not its “maker” or “originator.” The discoverer merely finds and records. Census takers, for example, do not “create” the population figures that emerge from their efforts; in a sense, they copy these figures from the world around them. The same is

⁴² Ibid.

⁴³ *Feist*, at 347-8

true of all facts scientific, historical, biographical and news of the day. They may not be copyrighted and are part of the public domain.

The idea vs. expression dichotomy asserts that an author's expression is copyrightable because it is an embodiment of her creativity. Facts are not creative hence cannot be protected even though they make up the bulk of the economic value of the database. Individual facts that are not copyrightable do not get copyright protection by simply combining in a compilation or database⁴⁴.

4.4.3 DATABASE PROTECTION - CONFLICTING PHILOSOPHIES: SWEAT OF THE BROW DOCTRINE VS. ORIGINALITY

Prior to *Feist Publications, Inc. v. Rural Telephone Service Co., Inc.*, there was confusion among courts as to how to evaluate copyright's originality requirement in database copyright cases⁴⁵. The general practice for the courts was to allow database owners to argue that their efforts, the "sweat of the brow," constituted reason enough to protect a database.⁴⁶ Another view was that the originality requirement based on whether the compiler had shown any creativity or judgment in the compilation of the database, was proved to be a determining factor to attract copy right protection.

The concept of originality was recognised as one of the fundamental aspects of copyright law, and it was made clear in two earlier Supreme Court cases, *Burrow-Giles Lithographic Co. v. Sarony*⁴⁷ and the *Trademark Cases*⁴⁸, and they introduced the concept without defining or quantifying it. The courts subsequently defined originality to mean either independent creation or creativity of the author, the so-called personality based copyright⁴⁹. Compilations that do not fulfil the originality requirement, were protected based on their 'sweat of the brow.' In *Jeweller's Circular*

⁴⁴ Ibid.

⁴⁵ Timothy Young, *Casenotes: Copyright Protection for Factual Compilations: The White Pages of the Phone Book are not Original Enough to be Copyrighted - But Why? - Feist Publication, Inc. v. Rural Telephone Service Co.*, 111 S. Ct. 1282 (interim ed. 1991), 17 U. Dayton L. Rev. 631, p. 632 (1992)

⁴⁶ James E. Schatz et al., *What's Mine is Yours? The Dilemma of a Factual Compilation*, 17 U. Dayton L. Rev. 423, p. 425 (1992).

⁴⁷ *Burrow-Giles Lithographic Co. v. Sarony* 111 U.S. 53, 57-58 (1884).

⁴⁸ *Trademark Cases* 100 U.S. 82, 94 (1879).

⁴⁹ Ginsburg, *No "Sweat"? Copyright and Other Protection of Works of Information After Feist v. Rural Telephone*, 92 Colum. L. Rev. 338, p. 346 (1992).

*Publishing Co. v. Keystone Publishing Co*⁵⁰. In *Jeweller's*, the Second Circuit court stated that an author of a compilation of facts, “produces by his labour a meritorious composition, in which he may obtain a copyright, and thus obtain the exclusive right of multiplying copies of his work.” “The sweat of the brow doctrine” gives authors awards for their skill and labour. This was basically a policy that was intended primarily to protect the investment of authors.⁵¹

Since 1976 the “sweat of the brow” doctrine as a defining factor in granting copyright started to diminish.⁵² In the 1976 Copyright Act, originality was introduced as a statutory requirement in deciding copyright protection in compilations, although the Act did not provide a clear definition. The Act provided a list of “works of authorship” and empowered the courts with flexibility to determine the degree of copyrights requirement of originality.⁵³

Originality as a primary requirement to attract copyright protection was first established in the *Feist* case. Till the judgement in the *Feist* case the courts were not sure of whether to uphold the sweat of the brow doctrine to assess the copyrightability of databases. *Feist* cleared the ambiguity and established that the databases were entitled for copyright protection if they displayed “a modicum of originality” in the “selection, coordination, or arrangement” of its contents. The “sweat of the brow” doctrine was explicitly rejected in *Feist* on the ground that, it was exceeding what was permissible under the U.S. Constitution.⁵⁴ The Supreme Court in *Feist* pointed out that the Intellectual Property Clause in the U.S. Constitution mandated originality as a prerequisite for copyright protection and the databases could be protected only to the extent provided by the Intellectual Property Clause.⁵⁵

The *Feist* decision also held that the selection and arrangement of the database was entitled to protection but underlying data would remain as a part of the public domain. The *Feist* decision also had negative effects on protection of database. Courts were prevented from exercising their authority to protect databases after the *Feist*

⁵⁰ *Jeweller's Circular Publishing Co. v. Keystone Publishing Co.*, 281 F. 83 (2d Cir. 1922).

⁵¹ Dov S. Greenbaum *The Database Debate*, p.453

⁵² *Ibid.*

⁵³ *Ibid.*

⁵⁴ Xuqiong (Joanna) Wu, *Foreign and International Law: A. E.C. Database Directive*, 17 Berkeley Tech. L.J. 571., p.592 (2002).

⁵⁵ *Ibid.*

decision. They had to find other methods to preserve databases, expanding the protection afforded to databases at the expense of the public domain⁵⁶. The new situation had given rise to many concerns among the database creators. Though emphasis was given on originality element of the databases it protected only the selection and arrangement of the database. It means that only the structure was given protection, even if the data were compiled in an original and creative fashion. The data remained free for people to copy, extract or reutilize. The result was that the copyright protection was virtually unavailable to databases⁵⁷.

The situation is far worse in the case of non creative and non original databases, because many research databases like bioinformatics and genomic research data are of the nature of raw discovery. Also they are non-creative and non-original, having tremendous future research potential. The researchers spend huge sums by way of investment in research to sequence the data, and they require enormous amount of skill and labour by way of research. Protections of these databases are at stake while emphasis is given to originality only to the structure of the database without protecting data underneath.

4.5 INTERNATIONAL INITIATIVES FOR DATABASE PROTECTION

The need for protection of databases was recognised much earlier. The copyright protection was offered to compilations in almost all regions of the world. However, the copyright philosophy as applicable to databases continued to be a subject matter of continuing debate. The protection of database compilations which lacked ‘modicum requirement of originality’ remained uncertain. The earlier copyright philosophy “sweat of the brow doctrine” and its applicability to compilations for attracting copyright in databases were subjected to legislative and judicial scrutiny. Uncertainty is still prevailing in the copyright sector as to which is the most appropriate concept, ‘modicum creativity’ or the ‘sweat of the brow’. Various international initiatives attempted to harmonise these provision relating to database protection.

⁵⁶ Dov S. Greenbaum, *The Database Debate: In Support of an Inequitable Solution*, 13 Alb. L.J. Sci. & Tech. 431., p.

⁵⁷ Peter K. Yu, *Evolving Legal Protection For Databases*, available at: <http://www.peteryu.com/gigalaw1101.pdf> last accessed on 12-09-2011

4.5.1 BERNE CONVENTION

The Berne Convention for the Protection of Literary and Artistic Works 1886 is the earliest exclusive multi-lateral copyright agreement with the objective to harmonize dissimilar national copyright regimes⁵⁸. Article 2, of the Berne Convention lists out copyrightable subject matter, which includes literary works and collections of literary works. The convention also explains original databases under Article 2(5) as collections of literary and artistic works which includes encyclopaedias and anthologies. They constitute an intellectual creation by reason of the selection and arrangements of their contents.⁵⁹ This particular provision refers to creative and original compilations.

Since 1886, the Berne Convention has played a major role in harmonizing national copyright laws. The convention is currently administered by the WIPO. Although the Berne Convention has historically played a central role in harmonizing national copyright laws, it does not cover protection of databases. Berne Convention provides for protection of creative compilations.⁶⁰ The creative compilation is different from factual data. Unfortunately, genetic information is factual data; therefore the data in these compilations are not covered under the Bern Convention. There has been much criticism against the Berne Convention due to its ineffectiveness. It does not provide any mechanism for authors to enforce any of their rights granted under the treaty. The Berne Convention does not provide any outright support for copyright of factual databases.

⁵⁸ Berne Convention, Art 2(1), reads as:

The expression "literary and artistic works" shall include every production in the literary, scientific and artistic domain, whatever may be the mode or form of its expression, such as books, pamphlets and other writings; lectures, addresses, sermons and other works of the same nature; dramatic or dramatico-musical works; choreographic works and entertainments in dumb show; musical compositions with or without words; cinematographic works to which are assimilated works expressed by a process analogous to cinematography; works of drawing, painting, architecture, sculpture, engraving and lithography; photographic works to which are assimilated works expressed by a process analogous to photography; works of applied art; illustrations, maps, plans, sketches and three-dimensional works relative to geography, topography, architecture or science.

⁵⁹ Berne Convention, Art. 2 (5) reads as:

Collections of literary or artistic works such as encyclopaedias and anthologies which, by reason of the selection and arrangement of their contents, constitute intellectual creations shall be protected as such, without prejudice to the copyright in each of the works forming part of such collections.

⁶⁰ Ibid.

4.5.2 TRIPs

The Agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPs) was adopted in 1995 as a part of the World Trade Agreement (WTA). The TRIPs set out the minimum standard of protection for intellectual property rights. However Member states are free to provide for stronger IPRs in databases. According to the TRIPs, an author may seek protection in a “compilation of data, which by reason of the selection or arrangement of their contents constitutes intellectual creations.”⁶¹ This provision provided by TRIPs essentially covers only creative databases, and it means that the agreement does not extend copyright protection to the data contained in compilations. However, the protection shall not extend to the data or material itself.⁶² This clause, though offers copyright in compilation and creative databases, does not extent to non creative databases. Though these provisions bar protection against the contents of the databases, member countries are free to change their national laws to protect databases especially non creative databases. For example, France grants protection to non-creative databases through its national copyright law.⁶³ Though TRIPs provides the greatest international protection to databases through multilateral treaty, it fails to cover the non creative databases. Biotechnology databases that are basically the collection and arrangement of genetic information, are considered to be non-creative. Hence even the TRIPs failed to offer adequate protection.

4.5.3 INITIATIVES FROM EUROPE

Europe made the primary initiative to provide exclusive protection of databases. The first attempt was the European Union Database Directive which provided copyright protection to the structure of the database and a *sui generis* protection to the contents of the databases. The European Union Copyright Directive which followed the Database Directive extended additional protection to the

⁶¹ TRIPs agreement, Article 10(2). Reads as:

“Compilations of data or other material, whether in machine-readable or other forms, which by reason of the selection or arrangement of their contents constitute intellectual creations shall be protected as such. Such protections, which shall not extend to the data or material itself, shall be without prejudice to any copyright subsisting in the data or material itself.”

⁶² Ibid

⁶³ Edward J. Baba, *From Conflict To Confluence: Protection Of Databases Containing Genetic Information*, p.145

databases. It provided technological measures to protect the databases and prohibited circumvention technology of safeguard them.

4.5.3.1 EUROPEAN UNION DATABASE DIRECTIVE

The European Union made the first initiative to develop a separate mechanism to protect the databases. The Directive on the Legal Protection of Databases was enacted in 1996 and came into force in January 1, 1998. The Database Directive widened the scope of database by defining it as a collection of data or materials systematically arranged in such a way that it can be accessed individually.⁶⁴ The Directive does not insist the data to be contained in the database to be physically stored in an organized manner. It extends copyright protection only to selection or arrangement of the contents of the database which constitute “the author’s own intellectual creation,”⁶⁵

The directive requires that member states provide two levels of protection for databases: (1) copyrights for protectable selection and arrangement through existing copyright law, and (2) *sui generis* protection for database contents.⁶⁶ The *sui generis* right is intended to protect the interest of the database owners who have made substantial investments in crating the database. The right prevents the extraction and re-utilization of the whole or of a substantial part of the contents of the database.⁶⁷ As opposed to copyright, no originality is needed for ensuring protection, but a mere condition of investment is sufficient. The Directive appears to be upholding the “sweat of the brow doctrine” and the death of the ‘originality’ concept.

The EU Database directive offers protection for a term of fifteen years, which can be further extended by an additional fifteen years if the owner makes an

⁶⁴ Directive 96/9/EC, Article 1, section 2 of the Directive defines a database as:

“a collection of independent works, data or other materials arranged in a systematic or methodical way and individually accessible by electronic or other means.”

⁶⁵ Directive 96/9/EC, Article 3 section 1 reads as:

“databases which, by reason of the selection or arrangement of their contents, constitute the author’s own intellectual creation shall be protected as such by copyright.

⁶⁶ Directive 96/9/EC, Article 7 section 1 reads as:

Member States shall provide for a right for the maker of a database which shows that there has been qualitatively and/or quantitatively a substantial investment in either the obtaining, verification or presentation of the contents to prevent extraction and/or re-utilization of the whole or of a substantial part, evaluated qualitatively and/or quantitatively, of the contents of that database.

⁶⁷ Ibid.

additional investment by accumulating more information or modifying the database⁶⁸. The directive does not require establishing the creativity or novel contribution; protection will be triggered without regard to the contents of the database⁶⁹. There are exceptions that provide for the use of information in the databases for purposes of teaching and scientific research, as long as such use does not conflict with the normal exploitation of the database by the author⁷⁰.

4.5.3.2 EUROPEAN UNION COPYRIGHT DIRECTIVE

In an effort to harmonize the copyright laws of the European Union, the E.U. Council of Ministers adopted EU Directive on copyright on April 9, 2001, in order to harmonise certain aspects of copyright and related rights in the information society. The objective was to establish uniform rules on copyright and related rights throughout the European Union⁷¹. The member countries were to amend their domestic legislations to comply with the directive before December 22, 2002. The 2001 Copyright Directive is built on six earlier directives dealing with copyright and related rights, namely: protection of databases, duration of copyright protection, satellite broadcasting, rental and lending rights, computer programs, and semiconductor topographies protection. The Copyright Directive was broader in

⁶⁸ Directive 96/9/EC, Article 10 reads as:

1. The right provided for in Article 7 shall run from the date of completion of the making of the database. It shall expire fifteen years from the first of January of the year following the date of completion.
2. In the case of a database which is made available to the public in whatever manner before expiry of the period provided for in paragraph 1, the term of protection by that right shall expire fifteen years from the first of January of the year following the date when the database was first made available to the public.
3. Any substantial change, evaluated qualitatively or quantitatively, to the contents of a database, including any substantial change resulting from the accumulation of successive additions, deletions or alterations, which would result in the database being considered to be a substantial new investment, evaluated qualitatively or quantitatively, shall qualify the database resulting from that investment for its own term of protection.

⁶⁹ Directive 96/9/EC Article 7 Section 4 reads as:

The right provided for in paragraph 1 shall apply irrespective of the eligibility of that database for protection by copyright or by other rights. Moreover, it shall apply irrespective of eligibility of the contents of that database for protection by copyright or by other rights. Protection of databases under the right provided for in paragraph 1 shall be without prejudice to rights existing in respect of their contents.

⁷⁰ Directive 96/9/EC Article 7 Section 5 reads as:

The repeated and systematic extraction and/or re-utilization of insubstantial parts of the contents of the database implying acts which conflict with a normal exploitation of that database or which unreasonably prejudice the legitimate interests of the maker of the database shall not be permitted.

⁷¹ Directive 2001/29/EC of the European Parliament and of the Council of 22 May 2001 on the harmonisation of certain aspects of copyright and related rights in the information society

scope, applying to a variety of media. The Copyright Directive allowed only one mandatory exemption and nineteen optional exemptions per member country⁷².

The *sui generis* protection of the 1996 Database Directive was entirely retained in the 2001 Copyright Directive. The EU Copyright Directive introduced in 2002 also broadens the scope of the *sui generis* protection in databases by providing legal protection to technological measures that a database owner may employ to protect his database from piracy.⁷³ The Directive provides for protection against circumvention technologies and the database owner may employ any effective technological measures to protect his database from piracy⁷⁴.

Exception to *sui generis* Right

The Copyright Directive creates an additional exception to the *sui generis* right for publicly accessible libraries, museums, or archives to reproduce data as long as there is no direct or indirect economic or commercial advantage through such use⁷⁵. This exception is especially significant in the context of databases, due to the concerns voiced by opponents of the 1996 Database Directive. There have been many criticisms against the EU database directive. Many organizations such as the European Bureau of Library, Information and Documentation Associations (EBLIDA) had criticised that the Database Directive limited public access to libraries and archival centres.⁷⁶ They argued that the directive would restrict public access and

⁷² Directive 2001/29/EC Article 5 provide a detailed description of exemptions and limitations.

⁷³ Directive 2001/29/EC Article 6 section1 reads as

Member States shall provide adequate legal protection against the circumvention of any effective technological measures, which the person concerned, carries out in the knowledge, or with reasonable grounds to know, that he or she is pursuing that objective.

⁷⁴ Copyright Directive, art. 6 section 2(c) reads as :

2). Member States shall provide adequate legal protection against the manufacture, import, distribution, sale, rental, advertisement for sale or rental, or possession for commercial purposes of devices, products or components or the provision of services which:(c) are primarily designed, produced, adapted or performed for the purpose of enabling or facilitating the circumvention of, any effective technological measures.

also see Edward J. Baba, *From Conflict To Confluence: Protection Of Databases Containing Genetic Information*, 30 Syracuse J. Int'l L. & Com. 121., p. 141 (2003)

⁷⁵ Ibid. at art. 5(2)(c).reads as:

(c) in respect of specific acts of reproduction made by publicly accessible libraries, educational establishments or museums, or by archives, which are not for direct or indirect economic or commercial advantage;

⁷⁶ Edward J. Baba, *From Conflict To Confluence*: p.142

would be contrary to good public policy. The 2001 Copyright Directive addresses this concern by permitting an exception for reproduction by libraries and archives.⁷⁷

4.5.4 THE WIPO COPYRIGHT TREATY

The WIPO Diplomatic Conference adopted the WIPO Copyright Treaty (WCT) which included certain provisions to clarify and amend the Berne Convention, establishing international norms for copyright law in a digital environment.⁷⁸ The Article 5 of the Treaty deals with copyright in databases and other compilations. According to the treaty the copyrightable compilations include data or other material, in any form, if the selection or arrangement of their contents constitutes intellectual creations. This protection does not extend to the data or the material itself and is without prejudice to any copyright subsisting in the data or material contained in the compilation.⁷⁹ The treaty also provides for adoption of technological measures to protect the copyrighted materials.⁸⁰

4.5.5 THE WIPO DRAFT DATABASE TREATY - 1996

The World Intellectual Property Organisation held a conference in Geneva to discuss the adoption of draft copyright treaty in 1996, the same year in which the Database Directive was adopted. The treaty was meant to update the copyright provisions of Berne Convention. “The Draft Treaty on Intellectual Property in Respect of Databases” was influenced by the European Database Directive.⁸¹

The draft treaty proposal required member nations to provide protection to databases that were created based on “a substantial investment in the collection, assembly, verification, organization or presentation” of data and to “prohibit the

⁷⁷ Edward J. Baba, *From Conflict To Confluence* p.142

⁷⁸ Dov S. Greenbaum, *The Database Debate*, p.467

⁷⁹ WIPO Copyright Treaty, Article 5 reads as:

Compilations of data or other material, in any form, which by reason of the selection or arrangement of their contents constitute intellectual creations, are protected as such. This protection does not extend to the data or the material itself and is without prejudice to any copyright subsisting in the data or material contained in the compilation.

⁸⁰ WIPO Copyright Treaty, Article 11 reads as:

Contracting Parties shall provide adequate legal protection and effective legal remedies against the circumvention of effective technological measures that are used by authors in connection with the exercise of their rights under this Treaty or the Berne Convention and that restrict acts, in respect of their works, which are not authorized by the authors concerned or permitted by law.

⁸¹ Mark J. Davison, *The Legal Protection of Databases*. p.218

extraction or utilization” of the contents⁸². The proposed treaty envisaged protection regardless of the creativity in the selection or arrangement of the data covering creative and non creative databases. Interestingly though the treaty was circulated among the delegates, it was never debated and it failed to turn it into an agreement.⁸³ The initial proposal of WIPO was based on the Directive under the impression that EU Directive is a sound model for protection of databases. However, it encountered stern opposition from a coalition of scientific and library organizations and the majority of developing countries.⁸⁴

4.5.6 US INITIATIVES

The protections of databases were a major concern in the US; and the US Congress considered several bills designed to address the issues of data protection including protection of non-creative databases. Between 1996 and 1999 a number of different bills were introduced into the Congress. A number of amendments were made to some of those bills at different committee stages. However, we can see that these legislative proposals were initiated after the EU had passed the directive and these new proposals were almost identical to the *sui generis* right originally suggested by the EU⁸⁵.

4.5.6.1 THE DATABASE INVESTMENT AND INTELLECTUAL PROPERTY ANTIPIRACY BILL

The first proposed legislation was the Database Investment and Intellectual Property Antipiracy Bill of 1996.⁸⁶ It was introduced soon after the adoption of the EU database Directive. It resembled the EU directive in many respects. The definition of the database and the *sui generis* right granted by the bill were similar to those in the EU directive.⁸⁷ However, the proposed bill provided for more protection compared to

⁸² WIPO Database Draft Treaty, Articles. I reads as:

Contracting Parties shall protect any database that represents a substantial investment in the collection, assembly, verification, organization or presentation of the contents of the database.

⁸³ Mark J. Davison, *The Legal Protection of Databases*. p.218

⁸⁴ Ibid.

⁸⁵ Ibid.

⁸⁶ HR 3531 of 104th Congress (1995-1996)

⁸⁷ HR 3531 Section 3 reads as:

(a) A database is subject to the Act if it is the result of a qualitatively or quantitatively substantial investment of human, technical, financial or other resources in the collection, assembly,

the EU. The bill provided for a 20 year protection period and also directly dealt with the issues of circumvention of database protection systems. The bill prohibited the importation, manufacture or distribution of circumvention devices to interfere with database protection systems.⁸⁸ There was considerable criticism of the 1996 Bill, particularly from the scientific and educational organisations.

4.5.6.2 THE COLLECTION OF INFORMATION ANTIPIRACY BILL

The primary goal of the Collection of Information Antipiracy Bill,⁸⁹ was to restore the “sweat of the brow” doctrine eliminated by the *Feist* decision. The Anti-piracy bill prohibited copying of substantial part of a collection of information gathered or maintained by a person through the investment of substantial resources, so as to cause material harm to the primary market of the database. The protection offered would apply to factual databases that were created and maintained by a substantial investment of money or other resources. A further goal of the Antipiracy Bill was to ensure protection for the U.S. database publishers similar to the E.U. Database Directive.⁹⁰

4.5.6.3 THE CONSUMER AND INVESTOR ACCESS TO INFORMATION BILL

The Consumer and Investor Access to Information Act.⁹¹ provides for a narrower protection scheme than the Collection of Information Antipiracy Act. The new Act prohibited duplication or distribution of data from the first database only to an extent of leading to a head to head competition. The bill created liability only if information was extracted from a database to create a competing substantially similar database. This was rather a narrower approach and found support only among

verification, organization or presentation of the database contents, and (i) the database is used or reused in commerce; or (ii) the database owner intends to use or reuse the database in commerce.

⁸⁸ Mark J. Davison, *The Legal Protection of Databases*, p.192

⁸⁹ H.R. 354 (1999)

⁹⁰ Sec. 1402. of H.R. 354 (1999) on Prohibition against misappropriation reads as:

‘Any person who extracts, or uses in commerce, all or a substantial part, measured either quantitatively or qualitatively, of a collection of information gathered, organized, or maintained by another person through the investment of substantial monetary or other resources, so as to cause harm to the actual or potential market of that other person, or a successor in interest of that other person, for a product or service that incorporates that collection of information and is offered or intended to be offered for sale or otherwise in commerce by that other person, or a successor in interest of that person

⁹¹ H.R. 1858 (1999).

educational and research communities. The new Act had only a limited scope of protection and the reduced possibility of liability.⁹²

Both the bills attempted to provide a protection that would benefit genomic databases. However, The Collection of Information Antipiracy would provide a higher level of protection by prohibiting the extraction of a substantial amount of information. The bill also gave the database compiler a broader scope of protection⁹³. Both bills survived House committees. However after some initial negotiations, the committees failed to reach a compromise and the bills died at the close of the 106th Congress⁹⁴.

Following the Failure of the Collection of Information Antipiracy bill and the Consumer and Investor Access to Information bill, the next major initiative to protect Databases began when a Supreme Court decision in *International News Service and N.B.A. v. Motorola Inc.*⁹⁵ came in. The case seriously addressed factors which constituted misappropriation of databases. In the instant case, the basketball league NBA sued the manufacturer of hand-held pagers Motorola, that provided real-time information about professional basketball games. In 1996, the NBA won the exclusive rights to transmit scores from federal Manhattan court. They alleged copyright infringement and commercial misappropriation of “hot news” data (“hot news” misappropriation) under New York law. The Court of Appeals while ruling against the NBA made following observations. The court identified five factors supporting a “hot news” misappropriation claim⁹⁶:

We hold that the surviving “hot-news” [International News Service]-like claim is limited to cases where : (i) a plaintiff generates or gathers information at a cost; (ii) the information is time-sensitive ; (iii) a defendant's use of the

⁹² H.R. 1858 (1999). SEC. 102. explaining prohibition against distribution of duplicates reads as:

It is unlawful for any person or entity, by any means or instrumentality of interstate or foreign commerce or communications, to sell or distribute to the public a database that:-(1) is a duplicate of another database that was collected and organized by another person or entity; and (2) is sold or distributed in commerce in competition with that other database

⁹³ Edward J. Baba, *From Conflict to Confluence*., p.139

⁹⁴ *Ibid* at p.140

⁹⁵ *International News Service and N.B.A. v. Motorola Inc.* 105 F.3d 841 (2d Cir . 1997).

⁹⁶ Robin Jeweler et al, *Protecting Noncreative Databases : Bills Before the 108th Congress, CRS Report for Congress*, p.CRS-4., available at : http://ipmall.info/hosted_resources/crs/RS21662_040406.pdf.

information constitutes free riding on the plaintiff's efforts; (iv) the defendant is in direct competition with a product or service offered by the plaintiffs ; and (v) the ability of other parties to free ride on the efforts of the plaintiff or others would so reduce the incentive to produce the product or service that its existence or quality would be substantially threatened.⁹⁷

4.5.6.4 DATABASE AND COLLECTIONS OF INFORMATION MISAPPROPRIATION ACT

Following *International News Service and N.B.A. v. Motorola Inc.*⁹⁸ another bill called “Database and Collections of Information Misappropriation bill” was introduced by 108th Cong. on October 8, 2003.⁹⁹ The full Judiciary Committee approved the bill on January 21, 2004. It created a new federal right of private action against misappropriation of databases, enforceable through a civil suit by a database owner against an infringer¹⁰⁰. The Act provides for permitted and prohibited acts relating to databases. The section three of the Acts provides for prohibition against misappropriation of databases.¹⁰¹ Further the Act also provides for permitted act under section four.¹⁰²

⁹⁷ *International News Service and N.B.A. v. Motorola Inc.* 105 F.3d 841 (2d Cir . 1997).

⁹⁸ Ibid.

⁹⁹ H.R. 3261.

¹⁰⁰ Robin Jeweler et al, *Protecting Noncreative Databases : Bills Before the 108th Congress*, CRS Report for Congress, p.CRS-3.

¹⁰¹ H.R. 3261, Section 3 reads as:

(a) LIABILITY- Any person who makes available in commerce to others a quantitatively substantial part of the information in a database generated, gathered, or maintained by another person, knowing that such making available in commerce is without the authorization of that other person (including a successor in interest) or that other person's licensee, when acting within the scope of its license, shall be liable for the remedies set forth in section 7 if--

(1) the database was generated, gathered, or maintained through a substantial expenditure of financial resources or time; (2) the unauthorized making available in commerce occurs in a time sensitive manner and inflicts injury on the database or a product or service offering access to multiple databases; and (3) the ability of other parties to free ride on the efforts of the plaintiff would so reduce the incentive to produce or make available the database or the product or service that its existence or quality would be substantially threatened.

¹⁰² H.R. 3261, Section 4 reads as:

- (a) independently generated or gathered information- This Act shall not restrict any person from--(1) independently generating or gathering information obtained by means other than extracting it from a database generated, gathered, or maintained by another person; and (2) making that information available in commerce.
- (b) Acts of making available in commerce by non profit scientific or research institutions
- (c) hyperlinking
- (d) news reporting

The newly created protection against misappropriation would not extend to databases gathered and maintained by a federal, state or local employee pursuant to their jobs, or to those required to be maintained pursuant to federal law or regulation. Also excluded from protection are computer programs used in the manufacture, production, operation or maintenance of a database¹⁰³.

4.5.6.5 CONSUMER ACCESS TO INFORMATION ACT OF 2004

Due to disagreement from earlier Congresses over the most desirable approach to database protection, the House Committee on Energy and Commerce declined to endorse “the Database and Collections of Information Misappropriation bill”, citing concerns about the bill’s constitutionality and its potential chilling effect on the use of information.” The Committee reported “Consumer Access to Information Act” on March 16, 2004.¹⁰⁴

The “Consumer Access to Information Act” offers a more limited protection to databases while preserving consumer access to and use of factual information. Like “Database and Collections of Information Misappropriation bill”, “Consumer Access to Information bill” is based on the principles of misappropriation and unfair competition established in the *International News Service and N.B.A. v. Motorola* cases. The “Consumer Access to Information bill” incorporates the entire criterion, including the requirement, absent from “Database and Collections of Information Misappropriation bill”, In order to be a misappropriation, the taking and use must be effected by ‘one in direct competition with the infringed database’.¹⁰⁵ Interactive

¹⁰³ H.R. 3261, Section 5(a) reads as:

(a) GOVERNMENT INFORMATION-

(1) IN GENERAL- Except as provided in paragraph (2), protection under this Act shall not extend to--
 (A) a database generated, gathered, organized, or maintained by a Federal, State, or local governmental entity, or by an employee or agent of such an entity, acting within the scope of such employment or agency; or(B) a database generated, gathered, or maintained by an entity pursuant to and to the extent required by a Federal statute or regulation requiring such a database.

(2) EXCEPTION- Nothing in this subsection shall preclude protection under this Act for a database gathered, organized, or maintained by an employee or agent of an entity described in paragraph (1) that is acting outside the scope of such employment or agency, or by a Federal, State, or local educational institution, or its employees or agents, in the course of engaging in education, research, or scholarship.

¹⁰⁴ H.R.3872 -- Consumer Access to Information Act of 2004

¹⁰⁵ H.R.3872 section 2 (b) defines misappropriation and reads as:

(b) Definition- For purposes of subsection (a), the term ‘misappropriation of a database’ means that--(1) a person (referred to in this section as the ‘first person’) generates or collects the information in the database at some cost or expense; (2) the value of the information is highly

computer service providers are exempt from liability for making available information provided by others. The major difference between the bills is the nature of infringement action provided. “Consumer Access to Information bill” does not create a private right of action. It directs the Federal Trade Commission (FTC) to treat misappropriation of databases as a violation of its rules defining unfair or deceptive practices.¹⁰⁶

4.6 BIOTECHNOLOGY RESEARCH AND RELEVANCE OF DATABASE PROTECTION

Biotechnology and related research have been recognised to be the most revolutionary area in contributing to the health sector. The advancement of technology has changed the very nature of scientific research in recent days. Biotechnology research has changed from hypothesis-driven to discovery-driven which is made possible with the help of growth in information technology. Research Databases have become a critical part of science today. Developments of modern technologies along with high input of experimental methodologies have resulted in extensive data generation. Biotechnology research with the help of information technology has developed a new area of research data management called bioinformatics. Bioinformatics process, curate, manage, and mine the biological data available in the databases¹⁰⁷. Bioinformatics does not produce its own raw data; it examines other researchers' data and relies on research outcomes to attain the information or data.

The biotechnology research database due to its very nature had been a subject matter of many controversies. Genetic databases in the form of sequence databases and gene expression databases are the core component of the biotechnology research and drug industry. The genetic database is essentially a string of alternating letters (A,

time-sensitive;(3) another person's (referred to in this section as the 'other person') use of the information constitutes free-riding on the first person's costly efforts to generate or collect it; (4) the other person's use of the information is in direct competition with a product or service offered by the first person; and (5) the ability of other parties to free-ride on the efforts of the first person would so reduce the incentive to produce the product or service that its existence or quality would be substantially threatened.

¹⁰⁶ Robin Jeweler et al, *Protecting Noncreative Databases : Bills Before the 108th Congress, CRS Report for Congress*, p.CRS-3.

¹⁰⁷ Almut Schulze and Julian Downward, *Analysis of Gene Expression by Microarrays: Cell Biologist's Gold Mine or Minefield?*, 113 J. Cell Sci. 4151., p.4151 (2000),

G, T, C) representing a DNA sequence. In order to analyse these sequence data, the entries must be organized in a logical manner such that they can be retrieved by computer search protocols. However, the integrity of the information is essential and an ordered non-creative organization is essential for researchers to work with these kinds of genomic databases¹⁰⁸. The genomic databases with non-creative arrangements cannot be protected under traditional copyright regime as copyrightable compilations due to the lack of sufficient originality.

The genomic databases which are bare, raw and unorganised make them unoriginal compilations. It does not mean that it must not be protected. The databases that contain genetic information must be unadulterated and organized in a particular manner so as to make it commercially valuable databases for further research. These data require tremendous amount of investment, research efforts, resources and work in compiling the collection. It is unfortunate that the data as result of the research are raw and non-creative; yet hold enormous value as new discovered data.¹⁰⁹ The genomic databases fall under the category of non-creative databases.

4.6.1 BIOINFORMATICS DATABASES

Bioinformatics otherwise called computational biology is recognised as a separate data management technology for genomic research. It can be broadly explained as the use of computers and computer allied analytical techniques in collecting and managing biological information. Briefly, bioinformatics is conceptualising biology in terms of molecules by applying “informatics techniques”, which is derived from disciplines such as applied mathematics, computer science and statistics in order to understand and organize the information associated with these molecules on a large scale.¹¹⁰ Bioinformatics with the help of Information Technology manages the biological information by data collection, warehousing, data mining, database searches, analysis, interpretation, modeling and product design.

¹⁰⁸ Edward J. Baba, *From Conflict To Confluence*, p.137

¹⁰⁹ Ibid at, p.138

¹¹⁰ Nicholas M Luscombe, Dov Greenbaum & Mark Gerstein, *What is Bioinformatics? An Introduction and Overview*, available at http://www.eecis.udel.edu/~lliao/cis841s06/bioinformatics_overview_gerstein.pdf last accessed on 12-8-2011

Bioinformatics also involves discovery, development and implementation of computational algorithms and software tools to understand the biological process.¹¹¹

The bioinformatics helps in storing the biological data organized in the form of a database. The databases allow the researchers an easy access to existing information and submit new entries. Bioinformatics also develops tools and resources that aid in the analysis of research data. It develops computational tools to analyze the biological data and interpret the results in a biologically meaningful manner.¹¹²

4.6.1.1 TYPES OF BIOINFORMATICS

Based on the nature of activities bioinformatics can be classified as computational bioinformatics and application bioinformatics

4.6.1.1.1 COMPUTATIONAL BIOINFORMATICS

Computational bioinformatics develops the software algorithms and tools for biological data analysis and also develops application for the research. The software algorithms and tools are developed with the help of software engineers, mathematicians, statisticians and biotechnology scientists. Computational bioinformatics also performs database construction and curation. The data generated are stored in to databases so that they can be retrieved in future¹¹³.

4.6.1.1.2 APPLICATION BIOINFORMATICS

User level application of bioinformatics is generally considered to be application bioinformatics. Bioinformatics software applications are the cornerstone of biotechnology and can be used in the analysis of sequences of data. The bioinformatics tools are also required to analyse the structure of DNA, RNA, and protein and to understand their interaction with in the cell structure. A whole new branch termed Structural Bioinformatics is devoted to predict the structure and possible roles of these structures of Proteins or RNA. Further, bioinformatics

¹¹¹ <http://raunakms.wordpress.com/2010/06/05/what-is-bioinformatics-%E2%80%93-a-general-perspective/> ; last accessed on 16-09-2011

¹¹² Ibid.

¹¹³ Ibid.

applications help analyze the function sequences and help predict the functional interaction between various proteins or genes¹¹⁴.

4.6.2 BIOTECHNOLOGY RESEARCH DATABASES.

Earlier bioscience researched to isolate single genes and proteins. But biotechnology research today is a more complex interaction between many levels of biological information such as DNA, RNA, proteins and so on. It is also about understanding how they work together. Modern day technologies help collect the data of thousands of genes simultaneously. Technologies such as DNA micro arrays and Affymetrix Gene Chips, enable the collection of this data¹¹⁵. Once the data are retrieved, they can be effectively processed, stored and analyzed to understand every relation¹¹⁶. Laboratories of molecular biologists bring out gigabytes of information and the bioinformatics techniques turn this information into useful and practical databases.

The genomic databases are basically different from other types of databases. These databases help in increasing the rate of discovery by providing for future research. Significantly, most of the new scientific data, such as nucleic acid sequence, cannot be published by conventional means such as journals. The publication is possible only by way of electronic cataloguing using bioinformatics techniques. The electronic cataloguing of the sequence of information within a database facilitates the emerging need for computational analysis of genetic information. Based on the nature, the genomic databases can be classified as follows.

4.6.2.1 DNA SEQUENCES

Biotechnology research basically comprises of exploration of DNA to understand the genetic mark-ups. Biotechnology researchers sequence the DNA to understand the genetic variations. The DNA sequence databases consist of combinations of nucleotides sequenced to express the genes. Four nucleotides, adenine, guanine cytosine and thymine, represented by their initial letters, A, C, G and T, form the particular DNA make-up of genes. A particular DNA molecule can be

¹¹⁴ Ibid.

¹¹⁵ Dov S. Greenbaum, *The Database Debate*, p.447

¹¹⁶ Ibid. at p.449

graphically represented by listing the nucleotide sequences that makes DNA molecule. These listing of nucleotide sequences forms the DNA databases¹¹⁷.

4.6.2.1.1 ESTs

DNA sequences are generally complete sequences and partial sequences. Partial sequences are also called expressed sequence tags (EST s). Another very crucial data are single nucleotide polymorphisms (SNPs). SNPs are different alleles (tiny variations) of the same gene, and exploring nucleotides of DNA sequences requires a series of biochemical and chemical reactions that are then analyzed by a computer. The organization of the sequences into an accurate and discernable collection of information requires many hours of labour on the part of bench researchers, the data compilers, and analysis groups. EST and SNP databases may be used to identify previously unknown genetic sequences or as templates characterising proteins and have tremendous potential for further research¹¹⁸.

4.6.2.2 SNP DATABASES

Another set of genetic information compiled into databases are Single Nucleotide Polymorphisms (SNPs), which are small, single nucleotide, variations in the sequence of a gene that may affect the protein production. These nucleotide variations called SNPs may change the nature of protein production and the modified protein may then exhibit a visible change in the organism. For example, SNPs have been implicated as the cause for the variation in eye colours. These are alterations in the genomic sequence that are not necessarily fatal or harmful; they simply may express a different phenotype and many SNPs, perhaps the majority, do not produce a physical change¹¹⁹.

In order to identify these single nucleotide variations in genes, DNA sequences from many different individuals must be compared to understand the small differences or the variations. The sequences must then be compared to multiple versions of the same sequence to find the single nucleotide differences. This process

¹¹⁷ Molly A. Holman and Stephen R. Munzer, *Intellectual Property Rights in Genes and Gene Fragments: A Registration Solution for Expressed Sequence Tags*, 85 Iowa L. Rev. 735, p.742(2000)

¹¹⁸ Matthew Rimmer, *The New Conquistadors: Patent Law and Expressed Sequence Tags*, (16) JLIS, p. 12 (2005)

¹¹⁹ Ibid.

may involve enormous amount of labour, money and time. Creation of the databases of SNPs may involve many scientists, researchers, computers, and a considerable amount of funding¹²⁰.

SNPs are highly significant in biotechnology research. They themselves may not change protein expression but some times may cause harmful mutations that will cause diseases; therefore, the SNP can be used as a screening marker for the possibility of disease development.

4.6.2.3 GENOME DATABASES

Genome databases forms a catalogue for the entire genetic information for individual organisms. Genome databases also attempt to compile the information about genes and their location on an organism's chromosome and the positions of related genes. These databases are larger in size and the costs of these databases are quite high due to the substantive volume of information carried by them. It is practically difficult for individuals to compile the volume of information needed for a chromosomal database. The compilation of the information for a chromosomal sequence database requires massive collaborative efforts among academic institutions and corporate research institutes.¹²¹ Human Genome Project (HGP) the largest collaborative research from institutions across the world which successfully sequenced the human genome. HGP database is the best example of genomic database.

Genomic database of an organism will assist researchers studying the particular organism in many ways. If a scientist is attempting to isolate the sequence of a particular gene, he may find it helpful to search the database for possible sequences that could facilitate the cloning of the gene. If the gene has already been isolated, the database may help in identifying the location of the gene and its role in causing disease.¹²²

¹²⁰ Matthew Rimmer, *The New Conquistadors: Patent Law and Expressed Sequence Tags*, p. 12

¹²¹ Edward J. Baba, *From Conflict To Confluence*, p.131

¹²² Ibid.

4.6.2.4 PATHWAY DATABASES

Increasingly powerful technologies, including genome-wide molecular measurements, have accelerated progress toward a complete map of molecular interaction networks in cells and between cells of many organisms. The growing scales of these maps are processed and stored in the databases¹²³. This newly emerging type of database incorporates the genetic information of an organism with the biochemical pathways and the component reactions of the enzymes and substrates inside the cell. Pathway databases include genes, proteins and metabolites data and the mechanisms by which they interact to complete the biological processes. The pathway databases are helpful in analysing and signaling, metabolic, molecular interaction, genetic interaction and gene regulatory pathway information.¹²⁴

A pathway is a series of biochemical reactions that are linked in the sense that the product of one reaction is involved in a subsequent reaction, either as a catalyst substrate or as an enzyme. The total volume of pathway data mapped by biologists stored in databases is rapidly growing. The molecular profiling¹²⁵ methods, such as RNA profiling using microarrays, or protein quantification using mass spectrometry,¹²⁶ provide large amounts of information about the dynamics of cellular pathway components. They enhance the pathway analysis techniques helping to pathway databases. These pathway databases have incredible future uses as they can be very helpful in future probes.¹²⁷

One of the potential uses of the database is to analyse the relations within a gene, for example the database connects enzymes with their genes and to their genomic sequence, and then links the genes and sequences to a description of the reaction catalyzed by each enzyme. The reaction information normally contains a

¹²³ Emek Demir et al., *The BioPAX community standard for pathway data sharing*, Nature Biotechnology, Vol. 28, No. 9, p.935 (2010)

¹²⁴ Ibid.

¹²⁵ Molecular profiling methods provide for a parallel analysis of thousands of genes or proteins. molecular profiling includes a variety of technologies which enables comprehensive screening of DNA, RNA protein and so on.

¹²⁶ Mass spectrometry is an analytical technique that measures the mass-to-charge ratio of charged particles and is used for evaluating the chemical structures of molecules, such as peptides and other chemical compounds

¹²⁷ Emek Demir et al., *The BioPAX community standard for pathway data sharing*, p.936

wide range of substrates¹²⁸ that which enzymes may accept. The data may also contain information about the chemicals that are known to inhibit or activate it. It may also contain the enzyme's subunit structure.¹²⁹

4.7 JUSTIFICATIONS FOR DATABASE RIGHT IN GENOMIC DATABASES

Biotechnology researchers normally followed patent protection in their research findings. The landmark case *Diamond v. Chakrabarty*¹³⁰, began the era of biotechnology patents for genetically modified organism. However the subsequent events were not that favourable for biotechnology research in patent landscape. The later case laws and judgements were totally negating the patents for biotechnology researches. These decisions are due to the common belief that the biotechnology researches are basically discoveries. They are the product of the nature and lack substantial utility. These views by the courts proved to be disastrous for genomic research especially genomic databases. The law is still unsettled with regard to patenting of genomic databases.

The very nature of biotechnology research has placed researchers in an uncertain situation causing worry about how to protect their intellectual property rights in research results. Since Biotechnology research is relying on information technology for the research data management, the next choice is to opt for a copyright protection to the research databases. However the applicability of copyright protection in non-creative database again remained a subject of debate and the copyright regime did not recognise copyright protection for non creative databases. Then the next available option is to opt for a protection under *sui generis* database protection similar to EU database directive or an appropriate provision under copyright law.

The protection of biotechnology research data like DNA sequences under copyright regime has many justifications. Following are the significant among these justifications.

¹²⁸ A substrate is a molecule upon which an enzyme acts. Enzymes catalyze chemical reactions involving the substrate

¹²⁹ Edward J. Baba, *From Conflict To Confluence*, p. 131

¹³⁰ *Diamond v. Chakrabarty*, 447 U.S. 303 (1980)

4.7.1 DNA AS A COPYRIGHTABLE SUBJECT MATTER

The fundamental copyright criterion is that a work must be an original work of authorship fixed on a tangible medium. When a scientist creates a new DNA from a new sequence of nucleotides, he creates a work which can be treated as an authorship. The rationale is that the new DNA sequence, designed by a scientist, qualifies as a work of original authorship, because the specific order or sequence of the base pairs in DNA is fixed in a tangible medium of expression, which is the DNA itself. In simple terms, the copyrighted expression is the actual DNA sequence itself, regardless of whether the sequence is fixed on paper or in a DNA strand¹³¹. DNA, RNA, or protein sequence can be expressed only in limited ways; these biological codes have become standard techniques for describing molecules and are therefore not “creative expressions.”¹³²

The best example is the artificial human insulin, where a biologist can genetically engineer a plasmid by embedding the new DNA which can replicate and produce human insulin. Genetically engineered DNA acts as a copy for the expression of the nucleotide sequence for producing a protein. Normally there can be only one way of expressing these methods, which is standard to the genetic engineering techniques adopted by the researcher.

4.7.2 DNA - AS LITERARY WORKS AND COMPUTER PROGRAMS

One attempt is to bring the genomic database as copyrightable literary work similar to computer programs, According to the US copyright Act, a literary work can be any thing that is expressed in words, numbers or other verbal or numerical symbols or indicia, regardless of the nature of the material objects. This may include books, periodicals, manuscripts, phonorecords, film, tapes, disks, or cards, in which they are embodied.”¹³³ Further computer programs are also deemed to be copyrightable

¹³¹ Leslie G. Restaino, et.al., *Patenting DNA-Related Inventions in the European Union, United States and Japan: A Trilateral Approach or a Study in Contrast?*. UCLA J. L. Tech. 2., P.9 (2002)

¹³² Scott McBride, *Bioinformatics And Intellectual Property Protection*, Berkeley Technology Law Journal., Volume 17:4 p16 (2002)

¹³³ 17 U.S.C. section 101 defines as literary work as:

"Literary works" are works, other than audiovisual works, expressed in words, numbers, or other verbal or numerical symbols or indicia, regardless of the nature of the material objects, such as books, periodicals, manuscripts, phonorecords, film, tapes, disks, or cards, in which they are embodied.

literary works under us copyright law. Hence a DNA expressed in a tangible medium is also eligible protection.

DNA sequences or genetic sequences are simply strings of nucleotides symbolised using the first letter of the nucleotides of the DNA. (adenine, thymine, guanine, and cytosine, symbolized as “A”, “T”, “G” and “C”). A biologist arrange a string of specified symbols in order to genetically engineer a DNA. These DNA sequences act as a program to produce specific proteins. These arrangements can easily fit into the category of literary work under copyright subject especially when a computer program has been recognised to be included in this category.¹³⁴

4.7.3 DNA AS A COMPILATION

Some of the supporters of copyright in DNA sequences view DNA sequences as valid compilation works. The justification is that the combining of various bits of DNA constitute a compilation; for example, combining the gene for human insulin, with a promoter from *E. coli* with a pX02 plasmid and inserting it into a Bacillus cell, is a valid compilation. In spite of the originality of the individual components, the selection and arrangement are original and copyrightable in such compilations of DNA¹³⁵. Thus a plasmid that codes for human insulin inserted in a bacterium may not be copyrightable; however the scientist’s exact selection of the human gene with the *E. coli* promoter and the pX01 plasmid within the Bacillus cell should be copyrightable. Here in this case, the compilation is sufficiently original to fit in as copyrightable subject matter. Hence, the arranged DNA sequences must be sufficiently original¹³⁶.

Though many attempts were made to establish originality in DNA databases, the fact which remains is that there still exist many uncertainties regarding the applicability of originality concepts to DNA databases. A major section of scientists and intellectual property experts believe that raw biotechnology research databases are non creative and non original. Though there are lots of intellectual and financial

¹³⁴ Irving Kayton, *Copyright in Living Genetically engineered Works*, 50 GEO. WASH. L. REV. 191., p.201 (1981)

¹³⁵ Ibid.

¹³⁶ James G. Silva, *Copyright Protection of Biotechnology Works: Into the Dustbin of History?*, B.C. Intell. Prop. & Tech. F. 012801, p 20(2000)

involvements in developing them, adequate protection is still not available to these databases as they are non original and non creative¹³⁷.

4.7.4 DNA AND THE “SWEAT OF THE BROW” DOCTRINE

Even though it is a common view that the biotechnology research databases are facts lacking originality, some proponents suggest that ‘sweat of the brow doctrine’ must be allowed in copyrighting of biotechnology databases. A researcher who spends considerable time and labour to separate the sequence of a gene should have a copyright in the sequence by virtue of his hard work. He may not have creatively conceived the sequence, by himself. Originality lies in the industriousness of the act of collection of the DNA sequence, and not in the presentation and expression of the sequence. Hence it is most appropriate to recognise the skill and labour of the researcher and also his financial investment in carrying out the research that has to be rewarded¹³⁸.

4.8 PROBLEMS WITH COPYRIGHT PROTECTIONS IN GENOMIC DATABASES

The inherent problem with protecting genetic information through copyright is the lack of originality possible within the selection, coordination or arrangement of the data themselves. According to the U.S. Copyright Act, the selection, coordination and arrangement of a database are able to be protected as original; but that protection does not extend to the underlying data. The human genome expressed as a database has similar limitations. Approximately 99.9% of genomes between any two humans are identical, with the variations forming an extremely small amount. Hence the creativity that can be employed is very limited, when evaluating the selection of the data that are already predetermined. The only creativity element that can be expressed is the reporting of genetic information resulting from sequencing. but it covers only the expression of the data, and not the data themselves.¹³⁹

¹³⁷ James G. Silva, *Copyright Protection of Biotechnology Works*, P21

¹³⁸ Edward J. Baba, *From Conflict to Confluence*., p.134

¹³⁹ Julie D. Cromer, *It's Hard to Find a Good Pair of Genes: So Why Make Them Free for the Taking?* 76 UMKC L. Rev. 505., p.512 (2007)

The Copyright Act mandates that “in no case does the copyright protection for an original work of authorship extend to any idea, procedure, process, system, method of operation, concept, principle, or discovery, regardless of the form in which it is described, explained, illustrated, or embodied in such work.” The gene sequencing could be classified as a “discovery;” it may fall outside the purview of copyright altogether.¹⁴⁰

4.8.1 FACTS LACK ORIGINALITY,

Originality is a condition precedent for copyright protection which means “that the work was independently created by the author and that it possesses at least some minimal degree of creativity.” The author cannot copyright his ideas or the facts he narrates, because facts do not owe their origin to an act of authorship. Further the concept of ‘sweat of the brow’ has been disengaged after the decision in *Feist* case. The originality is the fundamental element of copyright protection. The DNA sequence is a fact that normally lack originality.¹⁴¹ Gene sequences obtained from nature are merely discovery of facts. The biologist who sequences a gene is not doing an independent job. They merely copy it from nature; hence there is no minimum creativity envisaged by the copyright law. So long as a DNA sequence is based on a sequence discovered from nature, there is no independent creation, and the criteria of modicum creativity for originality is not met with.¹⁴² Hence the DNA or Gene sequences fall outside the ambit of copyrightable subject matter.

4.8.2 DOCTRINE OF MERGER AND THE NON-EQUIVALENCE OF DNA AND COMPUTER PROGRAMS.

Similar to originality, there is no copyright protection for ideas. The US copyright act provides that ideas, procedure, system, method of operation, concept, principle, or discovery are not entitled to copyright protection.¹⁴³ Like *Feist* on

¹⁴⁰ James G. Silva, *Copyright Protection of Biotechnology Works: Into the Dustbin of History?*, B.C. Intell. Prop. & Tech. F. 012801., P.25 (2000)

¹⁴¹ *Ibid.* at P. 26

¹⁴² *Ibid.*

¹⁴³ 17 U.S.C. section 102(b) reads as:

(b) In no case does copyright protection for an original work of authorship extend to any idea, procedure, process, system, method of operation, concept, principle, or discovery, regardless of the form in which it is described, explained, illustrated, or embodied in such work.

originality *Baker v. Seldon*,¹⁴⁴ established that there can be no copyright protection for ideas or procedures. Some times the idea often merges with the expression. The judgement in *Baker* eventually developed the “merger” doctrine. When there are very few ways of expressing a particular idea, the expression merges with the idea¹⁴⁵.

The situation is squarely applicable in the case of DNA sequences. The earlier illustration of combining the gene for human insulin with a promoter from *E. coli*, with a pX02 plasmid and inserting into a *Bacillus* cell can be justified for copyright protection. However, the selection can be argued to be original if the researcher independently creates the selection. Here the selection has minimal creativity because the researcher is only considering what is scientifically known and necessary. Even if we assume that the selection is original, still it may clash with the merger doctrine. The idea of combining promoters, plasmids, genes and bacteria can only be expressed in limited ways. Therefore there is merger and no copyright protection¹⁴⁶.

Another view is that the DNA is comparable to computer programs. DNA acts like computer programs. They are basically instructions to produce proteins coded using nucleotides similar to computer programs using programming language. This observation is also not correct, because there is only one way to express a “genetic program.” The combinations of the four nucleotide bases can be programmed only in one way to produce a single protein. Whereas the computer programs can have a single instruction expressed in numerous ways via different programming languages.¹⁴⁷

In *Apple Computer*¹⁴⁸ the case the Court stated that computer programs are subject to the doctrine of merger. The court reasoned that, “If other programs can be written or created which perform the same function...then that program is an expression of the idea and hence copyrightable. Otherwise the idea merges with the expression. However a DNA sequence in the form of a gene is the same as a computer program that can be expressed only in one way; hence there is a merger. As such, the merger doctrine precludes the copyrighting of DNA sequences that actually code for a protein.

¹⁴⁴ *Baker v. Seldon*, 101 U.S. 99 (1879),

¹⁴⁵ James G. Silva, *Copyright Protection of Biotechnology Works*, P.29

¹⁴⁶ *Ibid.* at. P. 31

¹⁴⁷ *Ibid.* at. P. 31

¹⁴⁸ *Apple Computer, Inc. v. Franklin Computer Corp.*, 714 F. 2d 1240, 1253 (3d Cir. 1983).

4.8.3 UTILITARIAN PROHIBITIONS.

According to the copyright law, useful articles are precluded from copyright protection. The Copyright Act defines a “useful article” as an, “article having an intrinsic utilitarian function that is not merely to portray the appearance of the article or to convey information.”¹⁴⁹ Thus, purely utilitarian objects are not copyrightable. DNA sequences basically have utilitarian functions that are neither physically nor conceptually separable. DNA sequences are coded to produce specific proteins and they cannot be physically removed from the DNA. Hence all the elements of a DNA coding are entirely utilitarian and not separable¹⁵⁰.

4.9 TRADE SECRET

Another option available for researchers to protect genome sequences is by keeping the discovered sequences as trade secrets. Trade secret protection has a much more limited application to databases than copyright law. Trade Secret Law protects the value of information kept out of the public domain through secrecy and obligations of confidence. The trade secrecy protects valuable commercial information that is not particularly novel or creative. When compared to patent law, trade secrecy (a) protects a potentially broader array of non-creative information. Examples are customer lists and marketing plans. Further the trade secret does not require patent law’s high standards of inventiveness. Because of these advantages, some businesses prefer to rely on trade secret protection rather than on patent protection or copyright protection¹⁵¹.

However, some times, when the statutory protection such as a patent or copyright fails to provide adequate protection to databases, another standard commercial strategy available for preserving the value of data and databases is secrecy or restriction on access. According to the US Trade secret Act, any “information, including a formula, pattern, or compilation... that derives independent

¹⁴⁹ 17 U.S.C. section 101a useful is illustrated as:

A “useful article” is an article having an intrinsic utilitarian function that is not merely to portray the appearance of the article or to convey information. An article that is normally a part of a useful article is considered a “useful article”.

¹⁵⁰ James G. Silva, *Copyright Protection of Biotechnology Works*, P.34

¹⁵¹ Jacqueline Lipton, *Balancing Private Rights And Public Policies: Reconceptualizing Property In Databases*, Berkeley Technology Law Journal Vol. 18: No.773 p. 816 (2003)

economic value from not being generally known to others who could profit from its use and is the subject of efforts that are reasonable under the circumstances can maintain its secrecy”.¹⁵²

Since trade secret protection exists until disclosure, it is possible for researchers to maintain part of the genomic sequence as a trade secret for some time. However, Protection of genetic information as a trade secret has some drawbacks. Some researchers believe that trade secrets restrict access to genomic data and the dissemination of information is curtailed due to non disclosure¹⁵³. The social value of the data is diminished as opposed to enhancement of it through disclosure. This would lead to duplication of research when other researchers explore the same area without knowing the present research. Further the duplicated effort may negate the economic value of the trade secret if an identical discovery is made¹⁵⁴.

4.10 CONCLUSION

Database protection is one the controversial areas in intellectual property regime. Copyright protection for compilation of data is limited to original and creative databases. All databases hold enormous value for the creator in terms of commercial potential and future research. Genomic databases which normally fall in the category of non creative databases have enormous scope of future research.

The protection of database under copyright regime is still a researchable area with enormous potential. Protection of creative and non-creative databases under the copyright regime has emerged as a major challenge for copyright regime. The non-creative databases may be requiring less creativity; however they involve considerable amount research and investment and have tremendous commercial potential. This scenario is still very delicate even after *Feist v. Rural* where it was held that

¹⁵² Uniform Trade Secrets Act Section 1(4) (amended 1985). Defines trade secret as:
[I]nformation, including a formula, pattern, compilation, program device, method, technique, or process, that: (i) derives independent economic value, actual or potential, from not being generally known to, and not being readily ascertainable by proper means by, other persons who can obtain economic value from its disclosure or use, and (ii) is the subject of efforts that are reasonable under the circumstances to maintain its secrecy.

¹⁵³ Rebecca S. Eisenberg and Arti K. Rai, *Harnessing and Sharing the Benefits of State-Sponsored Research: Intellectual Property Rights and Data Sharing in California's Stem Cell Initiative*, 21 Berkeley Tech. L.J. 1187, p.1196(2006)

¹⁵⁴ Ibid.

originality was the basic requirement though it required only a ‘modicum of creativity’. The *Feist* decision invalidated the ‘sweat of the brow’ and set originality’ as a basic copyright criterion.

In *Feist*, The Court took the view that the “original” used in Copyright Act of 1976 intended to overrule the “sweat of the brow doctrine” by legislation. Prior to *Feist*, there was strong view that the skill and labour involved in creating the databases must be recognised. The conflict was due to the copyright doctrine of ‘sweat of the brow’ which gave emphasis on the value of skill and labour involved. However the post *Feist* era has seen major challenges for database owners especially non-creative databases.

The first initiative came up from the European Union in the form of the EU Database Directive which required the member states to provide two levels of protection for databases. Copyright protection for the selection and arrangement of the database and a *sui generis* protection for the contents. Subsequently in 2001, the EU Copyright Directive was introduced which again broadened the scope of the *sui generis* protection in databases by providing legal protection to technological measures to protect databases. It also provided for the protection against circumvention technologies. the new *sui generis* right provides for exclusive protection to databases in which substantial investment is involved. The directive recognized the skill and labour involved in the production of the databases which involved substantial investment of time and money. As opposed to copyright, no originality is needed for ensuring protection of databases, but a mere condition of investment is sufficient to attract protection of databases. This move in a way restored the “sweat of the brow” doctrine and gave less emphasis on the originality requirement.

Biotechnology research and the results being placed in a highly vulnerable position are struggling to attract an appropriate intellectual property protection. There have been attempts to protect the research databases under the realm of copyright. Copyright protection of DNA sequences, at first, seems unusual. However DNA sequences engineered by scientists can be compared to artistic works. This comparison is an attempted logic for attracting copyright protection. Nevertheless,

copyright protection of DNA sequences is a realistic concept when one considers the copyright protection for computer programs.

But when a careful analysis of the Copyright Act, copyright case laws, and the actual science of DNA is done, the illusion of copyright protection of DNA sequences disappear. The copyright law requires that every copyrightable work must be original and expressed in a tangible medium. While considering the originality aspect, a DNA sequence cannot be considered as original, because DNA sequences exist in nature and are copied from the nature. Hence we cannot take DNA as a subject matter of copyright protection as such. Further the merger doctrine is also applicable preventing the DNA sequence from being truly a copyrightable subject matter. The idea merges with the expression because DNA sequences can be expressed in only one manner and with the same nucleotides.

The Biotechnology research data undoubtedly have enormous value as a subject matter of future research. Even though it falls outside the ambit of true copyright protection, there is still scope to protect it as a subject of database protection. There is substantive investment of skill, labour, time and money in sequencing the genomic data. Here again, the analysis of the nature of biotechnology research data shows that it forms only a non-creative database. Unlike creative compilations protected under copyright regime biotechnology research data are only having the nature of a discovery. It forms only a non-creative non-original database under the existing copyright regime. The Initiative like the EU database Directive and the EU Copyright Directive can resolve the uncertainty prevailing around protection of biotechnology research databases. A properly tailored *sui generis* protection similar the EU database protection can prove to be helpful in the effective management of Intellectual property issues in biotechnology.