

Patents and Bioinformatics

Patrick Lloyd

Reddie & Grose

UK & European Patent Attorneys

London & Cambridge

September 28-29, 2004

Reddie & Grose

Contents

- o What is Bioinformatics?
- o How does it differ from other aspects of patent law?
- o What are the issues?
- o Reach Through Claims
- o Examples

What is bioinformatics?

- o Blending of information technology and biology
- o Information management for biology
- o Software tools for data mining

How does it differ

- o Very similar issues to Software and business methods.
- o Thus, easier to protect claims in US
- o Need to show technical effect in EP
- o BUT – problems with Industrial Application

Industrial Application

- o Gene Sequence Patents
- o Patents for full length genes
- o Patents for Expressed Sequence Tags (ESTs) – the part of a gene an organism expresses, which is not dormant
- o ESTs generated by automated sequencing technologies
- o Early attempts to protect were withdrawn in face of criticism

Industrial Application

- o US Patent Office has developed a test of utility that must be credible, specific and substantial.
- o Applications that disclose only theoretical utilities will be rejected
- o Much more difficult to protect

Industrial Application

- o Biotechnology directive 2002
- o European Patent Convention Amended
- o Rule 23e(3) states:
- o The industrial application of a sequence or a partial sequence of a gene must be disclosed in the patent application

Issues

- o Bioinformatics Industry anticipated to grow to \$1.7 billion in next 2- 3 years
- o But most of the revenue will be in the therapeutic products which are discovered using bioinformatics tools
- o How to capture that market in bioinformatics patents

Reach-Through

- o Attempts to claim the product of the bioinformatics process
- o Thus, the end result, the therapeutic product attracts royalties under the patent

Reach Through Claims

- o Patents with reach through claims have been granted, particularly in the US
- o They are regarded now as invalid and are no longer granted
- o In essence, both the US and EP Patent offices take the view that reach-through claims are invalid for lack of disclosure

Rochester v Searle

- o United States CAFC Case 2004
- o Rochester developed a screening assay for use in determining whether a particular drug displayed such selectivity and received U.S. Patent No. 5,837,479 in 1998.
- o Rochester also obtained U.S. Patent No. 6,048,850 directed to methods for selectively inhibiting COX-2 activity in a human host by administering a non-steroidal compound that selectively inhibits activity of the COX-2 gene product to a human host in need of such treatment. However, no compounds to be used in the method were discussed.

The Rochester Claim

- o A method for selectively inhibiting PGHS-2 activity in a human host, comprising administering a non-steroidal compound that selectively inhibits activity of the PGHS-2 gene product to a human host in need of such treatment.

What the Court said

- o No examples of inhibitors were identified in the patent, and there was no suggestion of how an inhibitor could be made other than by trial and error.
- o The court agreed that the patent disclosed "nothing more than a hoped-for function for an as-yet-to-be-discovered compound, and a research plan for trying to find it."

Some Other Problems

- o Many Bioinformatics Patents relate to the use of data stored in databases, e.g to screen for a certain characteristic
- o Where does that take place?
- o Might it be outside the jurisdiction?
- o Imagine database in US and screening by Internet in India – would the US patent be infringed?
- o UK Case of *Menashe v William Hill*

Who has been filing

- o Incyte Pharma 764 issued US Patents
 - o Over 3000 entries on worldwide databases
- o Human Genome Sciences 382 issued US Patents
 - o Over 3500 entries on worldwide databases

Some Examples of Bioinformatics Patents

US 6,023,659 Incyte Pharma

A relational database system for storing biomolecular sequence information in a manner that allows sequences to be catalogued and searched according to one or more protein function hierarchies.

- o The hierarchies allow searches for sequences based upon a protein's biological function or molecular function. Also disclosed is a mechanism for automatically grouping new sequences into protein function hierarchies.

US 6,023,659 Incyte Pharma

- o Claim 1
- o A computer system comprising:
 - a database containing records pertaining to a plurality of biomolecular sequences;
 - a first hierarchy of protein function categories into which at least some of said biomolecular sequences are grouped, said protein function categories specifying biological functions of proteins corresponding to said biomolecular sequences and said first hierarchy including
 - (i) a first set of protein function categories specifying biological functions at a cellular level, and
 - (ii) a second set of protein function categories specifying biological functions at a level above the cellular level; and
 - a user interface allowing a user to selectively view information regarding said plurality of said biomolecular sequences as it relates to said first hierarchy.

US 6,023,659 Incyte Pharma

- o Claim 14. A method of using a computer system to present information pertaining to a plurality of biomolecular sequence records stored in a database...
- o Claim 35 In an internal database, a method of using a computer system to automatically categorize biomolecular sequence records into protein function categories...

US 5,953,727 Incyte Pharma

- o A computer system comprising:
a database having sequence records
containing information identifying one or more projects to which each of said sequence records belong, each of said projects grouping one or more biomolecular sequences generated during work to obtain a full-length gene sequence from a shorter sequence; and
a user interface allowing a user to selectively view information regarding said one or more projects.

US 6,300,078 Rosetta Inpharmatics, Inc

- o Methods and systems for characterizing the actions of drugs in cells. In particular, methods for identifying multiple primary targets through which a drug, drug candidate, or other compound of interest acts on a cell. Thus, the invention also relates to methods for drug development based on the disclosed methods for identifying multiple primary targets of a drug.

US 6,300,078 Claim 1

- o A computer system for determining a number of primary targets of a drug composition in a cell type comprising:
 -and
 - (c) one or more programs encoded by the memory, the one or more programs causing the processor to determine the number of expression sets wherein
 - (i) the expression sets each comprise a plurality of cellular constituents having similar inflection concentrations of the drug composition;
 - (ii) the inflection concentration of the drug composition for a cellular constituent is the level of exposure to the drug composition at which the cellular constituent is increased or decreased by the drug composition in a drug response; and
 - (iii) the drug response comprises measured amounts of the pluralities of cellular constituents in a cell of the cell type at a plurality of levels of exposure of said cell type to the drug composition,

US 6,300,078 Claim 42

- o A computer program product for directing a computer in a computer-aided determination of a number of primary targets of a drug composition in a cell type, said computer program product comprising

US 6,225,076 Rutgers Univ.

- o A method of identifying an agent for use as an inhibitor of bacterial RNA polymerase comprising:
 - (a) obtaining a set of atomic coordinates defining the three-dimensional structure of the core RNA polymerase
 - o (b) selecting a potential agent by performing rational drug design with the atomic coordinates obtained in step (a), wherein said selecting is performed in conjunction with computer modeling;
 - (c) contacting the potential agent with a bacterial RNA polymerase; and
 - (d) measuring the activity of the bacterial RNA polymerase;

US 6,047,109 SmithKline

- o A computational method maximizing open reading frame length in an assembly consensus sequence

US 6,047,109 SmithKline

- o A computer system for maximizing open reading frame length in an assembly consensus sequence comprising:
 - a. means for identifying frameshift ambiguities in the consensus sequence wherein the number of frameshift ambiguities is N;
 - b. means for dividing the consensus sequence into N+1 fragments split by the identified frameshift ambiguities;
 - c. means for generating a matrix of score values for each frameshift ambiguity in each frame from the fragment immediately following the frameshift ambiguity, where the matrix has N columns and three rows;
 - d. means for applying a scoring metric that allocates a value to a translation product of a sequence from one or more fragments wherein the value reflects the likelihood that the sequence is a real coding sequence;
 - e. means for repeating steps a through d for the reversed complement of the consensus sequence;
 - f. means for rebuilding the consensus sequence wherein the highest scoring reading frame in the highest scoring direction is optimized; and
 - g. means for displaying the results of step (f) on an output device.