Freedom-to-operate patent searches for sequences of drug targets

Heahyun Yoo

Patent Analysis group
Information and Knowledge Integration Informatics
Pharmaceutical Research Institute
Bristol-Myers Squibb Company

February 12, 2007
PIUG Biotech meeting, Boston, MA

Agenda

- ***Targets and freedom-to-operate (FTO) searches***
- Criteria/scope of sequence analysis for FTO searches
- Challenges in sequence based FTO searches
- Suggested ways to alleviate problems associated with the challenges
- Summary

Freedom-to-operate search for validated targets

- **Technically…. The role of target sequences**

  In drug discovery process, high throughput compound (drug) screening is most commonly performed using recombinant cell which express target genes.

- **Legally…. The intellectual property issue**

  Evaluate intellectual property (IP) for using sequences of validated targets in drug discovery process - to determine whether target genes/proteins have been claimed in issued patents.

Agenda

- Targets and freedom-to-operate (FTO) searches
- ***Criteria/scope of sequence analysis for FTO searches***
- Challenges in sequence based FTO searches
- Suggested ways to alleviate problems associated with the challenges
- Summary

Scope of sequence analysis for FTO searches

- **Claim status**: claimed sequences only (in patents granted in the country of operation and applications that might be granted)

- **Sequence databases**: select all relevant patent sequence databases

- **Publication dates**: all publications for the past 20 or so years with exceptions associated with possible extension of patents.

- **Typical criteria** of assessing relevancy of hits from the sequence analysis results (continued in the next slide)

Scope of sequence analysis for FTO searches (continued)

Criteria of selecting relevant sequence hits from Blast results

- Percent identity and/or the length of amino acids/nucleotides sequences matching between the query and hits
- Sequences claimed in association with particular biological functions, genetic variants, recombinant cells, sequences with artificial modified residues,
- Implicitly claimed sequences, i.e., sequences are claimed but the actual data are not disclosed in patents hence not indexed in sequence databases – Challenges in searching patent sequences.

Agenda

- Targets and freedom-to-operate (FTO) searches
- Criteria/scope of sequence analysis for FTO searches
- ***Challenges in sequence based FTO searches***
- Suggested ways to alleviate problems associated with the challenges
- Summary

Challenges of FTO search

1. How patent claims are drafted: typical examples

Nucleic acid Sequences

- ……wherein said first nucleic acid molecule hybridizes to a second nucleic acid molecule shown in SEQ ID NO: 1 using hybridization conditions of 40-50% formamide, 37-42degrees…
- An isolated nucleic acid molecule comprising a polynucleotide encoding the complete polypeptide encoded by the cDNA in ATCC Deposit No. 12345.
- The compound of claim 1 wherein the modified nucleobase is a 5-methylguanosine....

Challenges of FTO search

1. How patent claims are drafted: typical examples

Protein Sequences

- …. wherein Conantokin-G has the formula Gly-Glu-Xaa1 –Xaa2 -Leu-Gln-Xaa1 -Asn-Gln-Xaa2 -Leu-Ile-Arg-Xaa1 -Lys-Ser-Asn (SEQ ID NO:1) where Xaa1 and Xaa2 are each g-carboxy-glutamic acid and wherein said ……

- The non-naturally occurring protein according to claim 1, wherein said substitutions are selected from the group of substitutions consisting of …

- An amino acid sequence encoding an enzyme that has at least 75% identity over length of 410 amino acids based on the CLUSTALW method of alignment when compared to a polypeptide having the sequence as set forth in SEQ ID NO:1.

Challenges of FTO search

1. How patent claims are drafted: typical examples

Would one have FTO in using the following sequence* for compound screening?

```
   1 gctcggcggg tcactcagct atggagcgga agtcgaatct gtctctgctt ctcctactgc
  61 tggtcctggg catgcccctg gtgcggggct ccagccctct gcccctggtc gtcaacactt
 121 ggccttttaa gaatgccact gaagcagcgt ggtggacatt gctatctgga ggttctgccc
 181 tggatgcagt ggagaacggc tgtgctgtgt gtgagaagga gcagtgtgat gggactgtag
 241 gctttggagg aagtcctgat gaaggtggcg aaaccaccct ggatgccatg ataatggatg
 301 gcactgccat ggatgtggga gcagtgggag gccttagaag aattaaaaac……
```

*Sequence data: GenBank entry: S81393

Challenges of FTO search

1. How patent claims are drafted: typical examples

A hit from a PCT application…… document_location=Claim 2; SEQ ID NO 1459; 201pp;English /pub_date=13-MAR-2005, mitochondrial DNA sequence SEQ ID NO:1459.

```
Query: 15  tcagctatggagcggaagtcgaatctgtctctgcttctcctactgctggtcctgggcatg 74
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 1   tcagctatggagcggaagtcgaatctgtctctgcttctcctactgctggtcctgggcatg 60

Query: 75  cccctggtgcggggctccagccctctgcccctggtcgtcaacacttggccttttaagaat 134
           |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Sbjct: 61  cccctggtgcggggctccggccctctgcccctggtcgtcaacacttggccttttaagaat 120

Query: 135 gccactgaagcagcgtggtggacattgctatctggaggttctgccctggatgcagtggag 194
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 121 gccactgaagcagcgtggtggacattgctatctggaggttctgccctggatgcagtggag 180

Query: 195 aacggctgtgctgtgtgtgagaaggagcagtgtgatgggactgtaggctttggaggaagt 254
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 181 aacggctgtgctgtgtgtgagaaggagcagtgtgatgggactgtaggctttggaggaagt 240

Query: 255 cctgatgaaggtggcgaaaccaccctggatgccatgataatggatggcactgccatggat 314
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 241 cctgatgaaggtggcgaaaccaccctggatgccatgataatggatggcactgccatggat 300

Query: 315 gtgggagcagtgggaggccttagaagaattaaaaac 350…….
           |||||||||||||||||||||||||||||||||||
Sbjct: 301 gtgggagcagtgggaggccttagaagaattaaaaac 336  …….
```

Note: No US nor EP granted patents claimed the hit sequence.

Challenges of FTO search

1. How patent claims are drafted: typical examples

A hit from a US granted patent

```
 …..phosphate aminotransferase VF5/HPA DNA.
/section="Claim 1; Fig 7; 95pp; English.
/pub_date=1997-08-14 /Subset=gs_claim_nuc
 Length = 1065
 Score = 40.1 bits (20), Expect = 0.49
 Identities = 20/20 (100%)
 Strand = Plus / Plus


Query: 316 ggccttagaagaattaaaaa 335
           ||||||||||||||||||||
Sbjct: 117 ggccttagaagaattaaaaa 136
```

Challenges of FTO search

1. How patent claims are drafted: typical examples

What is claimed is:

1. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide encoding an enzyme as set forth in SEQ ID NOS: 25–32;

(b) a polynucleotide which is complementary to the polynucleotide of (a); and

(c) a polynucleotide comprising at least 15 consecutive bases of the polynucleotide of (a) or (b) and which hybridize under stringent conditions to a polynucleotide encoding an enzyme as set forth in SEQ ID NOS: 25–32.

2. The polynucleotide of claim **1** wherein the polynucleotide is DNA.

3. The polynucleotide of claim **1** wherein the polynucleotide is RNA.

⋮

Challenges of FTO search

2. Patent sequence databases coverage

- GeneSeq: Basic patents/applications; 43 authorities; From 1981 to the present; update > 2 month (faster update by GeneSeq Alert).
- Registry: Basic patents/applications and US equivalents mainly from GenBank; 38 authorities; From around 1990 to the present (at the time of Genbank's inaugural release); updated about once a month; complete annotation records only for those post 1990.
- PCTGEN: WIPO filings from August 2001 to the present; weekly update.
- Gene-IT: Basic and non-basic patents
- USGene: Sequences from US issued patents and applications

Challenges of FTO search:

3. Sequence analysis tools

BLAST and FASTA

- Most commonly used tools currently
- Designed for retrieving sequences based on homology
- Sequence analysis for FTO is based on % identity, even for short regions, rather than overall homology.

Agenda

- Targets and freedom-to-operate (FTO) searches
- Criteria/scope of sequence analysis for FTO searches
- Challenges in sequence based FTO searches
- ***Suggested ways to alleviate problems associated with the challenges***
- Summary

Suggested approaches to alleviate problems

- Keyword searches, especially in claims and abstract fields using all terms for drug target names (use synonyms databases)
- Use all pertinent sequence databases
- Use sequence analysis tools that may work more appropriately for sequences claimed in patents

Summary

- FTO searches for drug target sequences may be performed in target validation stage
- Scope of sequence analysis; criteria of selecting answers
- Challenges in FTO searches for patented sequences:
  - -- Nature of claims made on sequences
  - -- Sequence database coverage
  - -- Sequence analysis tools
- Suggested approaches to alleviate problems

Questions?

- For more information, see Yoo et al., Intellectual property management of biosequence information from a patent searching perspective, World Patent Information **27** (2005) 203-211.
- For future correspondence, contact heahyun.yoo@bms.com
- Disclaimer: The views expressed herein are those of authors and should not be attributed to Bristol-Myers Squibb Company