

GENESEQ™

THOROUGH COVERAGE OF BIOLOGICAL SEQUENCES
FROM WORLDWIDE PATENTS



WHAT YOU CAN DO

- Find biological sequences not available in other databases
- Understand the global patent landscape
- Reveal potential patent infringements
- Uncover the activities of your competitors
- Identify prior art
- Establish freedom to operate and patentability
- Quickly assess the context of the sequence in the patent

WHO CAN BENEFIT

- Intellectual property (IP) professionals
- Bioinformatics specialists
- Molecular biologists
- Licensing and business development professionals

UNRAVEL THE COMPLEX STRANDS OF BIOLOGICAL SEQUENCES

GENESEQ ensures you're seeing the entire patent landscape surrounding the biological sequence under investigation.

Manually curated and professionally annotated, it covers all biological sequences patented since 1981 from 41 worldwide patent-issuing authorities, including WO, US, EP, JP, DE, IN and CN.

With over 27 million biological sequences from more than 150,000 unique patents, it contains a wealth of essential sequence information not available anywhere else.

The database is continually growing with an average of 600 unique patents containing biological sequences added every two weeks. This means biological scientists can spend less time searching for sequence data and more time on research and development, and IP professionals can be sure they don't miss any potential infringements or competitor activities.

.....
GENESEQ contains genetic sequences from worldwide patents, manually annotated to highlight IP context and biological significance—an essential tool for any organization that needs to know what's happening in genomic-based life science discovery.
.....

COMPLETE, TIMELY, ANNOTATED ANALYSIS WITH GLOBAL COVERAGE

Our dedicated editorial team provides thorough coverage of all biological sequence information from patents including:

- Nucleic acid sequences 10 or more bases in length
- Amino acid sequences 4 or more residues in length
- All PCR primers and probes of any length
- Sequence and contextual information, irrespective of its location within a patent, including organism name, gene/protein name, sequence modifications, associated disease information
- Non-English patents (in total, these comprise more than a third of the database), including Japanese, Chinese, Korean and European languages, with English language indexing
- Links to identical records from NCBI, SWISSPROT, Gene Ontology (GO) annotation, and sequence aliases

The team adds a unique commentary, clarifying sequence novelty and the actual utility underlying each invention. Each record is fully indexed using standardized terminology, enabling accurate search and retrieval.

This annotation includes:

- Enhanced patent titles
- English language abstract from non-English sources
- Sequence location within the patent document
- Original source organism
- Highlighted biologically significant regions of the sequence
- Standardized, full bibliographic data



THOMSON REUTERS

HOW GENESEQ AIDS IP PROFESSIONALS

GENESEQ provides manually-captured sequence information not available in any other sequence database, nor in patent office databases.

Many biological sequence patents do not have electronically available or machine-readable sequence information—the only way to capture this information is manually. *GENESEQ* captures up to three million characters of this sequence data every week.

Manual indexing standardizes terminology, enabling you to narrow your results quickly and accurately to the relevant disease, technology, or sequence itself. You can search all sequences in a patent, irrespective of their location (in figures, disclosure, claims, and so on).

GENESEQ helps you to identify key sequences by putting each sequence into context. It is used by major patent issuing authorities, so you can be sure you're checking the same information that your examiners are.

HOW GENESEQ AIDS BIOLOGICAL SCIENTISTS

Currently only 10% of the sequences in *GENESEQ* are identical in length and identity to those in NCBI—which means that approximately 24 million sequences can be found only in *GENESEQ*. Much of this unique content is scientifically important, and hence vital to your research.

GENESEQ enables you to understand the context of each sequence in the patent, quickly identifying new disease, gene, or target associations, and sequence modifications. Its standardized terminology makes it easy to map *GENESEQ* into your in-house bioinformatics systems.

GENESEQ contains raw sequence information for:

- Gene/protein disease association
- Target identification
- Biotherapeutic creation
- Biomedical technology creation
- Gene/protein families enrichment

COMPREHENSIVE ACCESS AND RELEASE OPTIONS

GENESEQ is available through leading sequence search and data management solutions including:

- GenomeQuest — integrated with the extended patent family coverage of GQ-Pat, so that you can find biological sequence information more accurately than ever
- STN International — suitable for advanced information searchers
- Thomson Pharma® — provides an at-a-glance view of the patent portfolios and activities of your competitors

GENESEQ is also available via an EMBL-like FTP-delivered flat file, delivered every two weeks for inclusion in your in-house information system.

An alternative option is the companion database *GENESEQ*™ FASTAlert that enables rapid searching of new sequences. It is released as a rolling FASTA database on Tuesday and Thursday every week.

Scientific Head Offices**Americas**

Philadelphia +1 800 336 4474
+1 215 386 0100

Europe, Middle East and Africa

London +44 20 7433 4000

Asia Pacific

Singapore +65 6411 6888
Tokyo +81 3 5218 6500

For a complete office list visit:

scientific.thomsonreuters.com/contact

GET A QUOTATION

Contact us today to request a no obligation quotation or visit scientific.thomsonreuters.com

