



# Patent Sequence Search: Why You're Missing Crucial Sequences



At Aptean GenomeQuest, we do things differently. From our user-friendly search interface to our customer support, there's a lot that sets us apart, but it's our content - content that can only be found in our GQ-Pat database - that truly makes us the market leader in patent sequence search. Think all patent sequence databases are the same? Read on to find out what makes our content different - and why you might be missing crucial sequences in your searches today.

When an inventor files a life science patent application at the USPTO, they are asked to put all sequences into a specifically formatted list. This so-called "ST.25 listing" helps the examiners with their workflow and makes it straightforward to collect all sequences submitted to the office over time. In an ideal world, every inventor and every patent office would list sequences like this and that would be the end of it.

Unfortunately, very few patent offices in the world actually have official sequence-filing rules, and even when they do have them, they're frequently ignored. As a result, sequences can be found anywhere in a patent: inside the text, in tables, or even as part of the figures. If your search only spans the ST.25 sequence listings, you're going to miss out on a lot of them!

You might be wondering, "But how big is this problem, really? How many sequences can be found outside of the official ST.25 listings?"

Unfortunately, quite a few. In fact, you may easily be missing more than 40% of US, WO, EP, JP, KR patent documents with sequences in them. This number is based on a massive amount of internal sequence data that we confirmed by comparing the number of patent numbers (PNs) with competitors' products based solely on ST.25 listings.

Sure, anyone can download some ST.25 sequence listings and index them for BLAST search, but this approach will cause you to miss out on real business-critical sequences. Why? Because all those sequences that are located in text, tables, and figures aren't indexed!

Are you still not convinced that you're missing documents relevant to you? Here's a list with some of the largest patent assignees in our patent sequence database, and the percentage of patent applications they filed with sequences hidden in the text, tables, and figures. These are all documents that would never be found in ST-25 listings.

Patent Assignee	Percentage patent applications with sequences outside ST25 listings
Shanghai Biowindow	53%
Abbott	41%
Sanofi Aventis	39%
Astrazeneca	31%
Novo Nordisk	29%
Pfizer	28%
Novartis	27%
Johnson & Johnson	26%
Merck Sharp Dome	24%
Bayer	20%
Bristol Myers Squibb	20%

How do we find these documents? Over the years, we have invested countless hours and large sums of money to find and index every last sequence that is out there. We use proprietary algorithms to flag documents with even a minuscule chance of containing sequences. Flagged documents are manually curated to confirm that there are sequences in there, after which they are captured. Our human curator team also verifies additional information like the SEQ ID NO and whether a sequence is mentioned in the claims or not. We have examined the entire backfile of over 120 million historical patents for sequences.

Of course, we haven't stopped there – we continuously process the new patents that are being published to ensure our patent database contains the most up-to-date sequence information available. In addition to 40% more US, WO, EP, JP, KR patents, we have also indexed over 200,000 documents from authorities in China, Canada, India, and Brazil that you will not easily find anywhere else.

Your IP strategy depends on having access to complete and accurate information. If sequences are not even in the database, how are you ever going to find them?

To learn more about Apteian GenomeQuest or to schedule a demo, email [info@aptean.com](mailto:info@aptean.com).



Apteian is a global provider of mission-critical, industry-specific software solutions. Apteian's purpose-built ERP and supply chain management solutions help address the unique challenges facing process and discrete manufacturers, distributors, and other focused organizations. Apteian's compliance solutions are built for companies serving specific markets such as finance, healthcare, biotech and pharmaceuticals. Over 2,500 organizations in more than 20 industries across 54 countries trust Apteian's solutions at their core to assist with running their operations. To learn more about Apteian and the markets we serve, visit [www.aptean.com](http://www.aptean.com).