aptean

Ebola Vaccine Produced in Plants, a Patent Searcher's Story

www.aptean.com

Plant-based vaccine production has become a viable alternative to other manufacturing methods, as it often involves fewer purification steps, is less expensive than traditional production, and yields ambient temperature-stable molecules[1]. This is particularly important for less developed regions of the world, such as Africa, where viruses like Ebola have devastated populations[2]. Because they have such high potential for providing low-cost immediate production, plant-produced vaccines hold immense promise to help the areas that need them the most[3]. Recent studies have demonstrated the feasibility of expressing an Ebola Immune Complex (EIC), coupling the Ebola glycoprotein GP1 to a humanized monoclonal antibody in tobacco plants[4]. Preliminary clinical trials for virus protection are showing promise in animal model systems[5].

An in-house R&D team at a leading pharmaceutical company found genome sequencing of one Ebola strain identified a potential protein epitope that could be coupled with an immunoglobulin and expressed in plants for purification and usage as a vaccine to protect against this devastating filovirus. Freedom to operate (FTO) for each step of the process had to be ensured, and patentability for potential inventive aspects needed to be evaluated.

Definition of the query parameters

For assessing nucleotide and protein sequence similarities for the EIC protein the R&D group identified, Aptean GenomeQuest's GenePAST algorithm was the first algorithm used, as it allows percent identity cutoffs to be specified in the query step. The results were easy to analyze versus typical claim language in the technology space. If particular motifs or fragments of sequences were identified that were commonly referenced in patent claims, refinement searches using the fragment algorithm were run. The fragment algorithm is particularly suited to smaller sequence lengths that have higher percent identity across local regions.

Both the nucleotide and corresponding polypeptide sequences were searched against all relevant patent and non-patent databases. GenomeQuest allowed the selection of any one or multiple subject databases as well as the option to create a custom database, such as in-house genomic sequencing collections, against which to search. Coverage of every important patent authority ensured comprehensive results. Both the GQ Gold+ collection (ST.25 listings from the US, EPO, WIPO, JP, KR) and the Platinum collection (non-ST.25 listings as well as extended legal status, normalized patent assignee, unique family sequence IDs for US, EPO, WIPO, JP, KR, AT, AU, BE, CA, CH, DE, ES, FR, GB, LU, NL, NO, TW, with in-country documents for CN, BR, IN, RU) were selected.

In this case, as the prior art potentially overlapped many different data sources such as plants (vaccine expression), humans (disease target), animals (model studies), viruses (Ebola), as well as likely hits in high throughput sequence collections, the first query run was against every database that GenomeQuest offers – encompassing more sequences than any other search platform in the world.

Identification of different patentability and FTO considerations with metadata

Before assessing sequence hits, GenomeQuest made it possible to view a snapshot of patent vs. non-patent results, earliest priority dates, and other key details. Metadata, such as earliest instance of the sequence in a patent document as well as in literature, informed further investigations. GenomeQuest revealed in a summary table that the EIC sequence had a non-patent disclosure nearly a decade prior to a hit against a US patent, indicating that the composition and methods of use would have different patentability and FTO considerations. From a single query, it was straightforward to then assess the patent and non-patent literature, either simultaneously or separately.

Organization of results with user-defined requirements

GenomeQuest provided results that were sortable and filterable and offered customizable views, with many options for displaying and ordering fields. Key information was also displayed in graphics for quick visual assessment. Hyperlinks to legal authority original documents were available in all views for easy retrieval and reference. This was in stark contrast with previous platforms the team worked with. Before using GenomeQuest, searchers were frustrated with the limitations around viewing and using information, which caused unnecessary delays and complicated the search process. Custom filters, sorting, and grouping options in GenomeQuest meant that different types of views were created to facilitate different analysis paths. For example, once Percent ID matches were filtered to narrow the hit pool to only the most relevant sequences, a customized view was implemented to group the hits by normalized patent assignee in order to more easily track the filing histories as well as identify which patent families may be related.

Elimination of multiple platforms

Individual hits were quickly and easily evaluated in the same window, maintaining continuity and focus during the analysis. Every single piece of information relevant to decision making was accessible within each record, and could be opened/closed with a single click, without losing place in the overall results list: SEQ ID match parameters (%ID, Subject, Query, Coverage), patent claims, priority history, family members, and much more. Different tabs within each expandable view allowed detailed evaluation of exact sequence match parameters, patent status and claim evaluation, as well as a graphical representation of priority dates. Additional windows provided the ability to scroll through family members, evaluate at a glance which claims comprised sequences, and more. Scrutiny of the earliest priority date document that relates to the EIC query revealed a granted sequence-independent claim – in a lapsed US patent. This piece of information was of significance in both FTO and patentability decisions, and the ability to evaluate all aspects of that information efficiently within the GenomeQuest platform saved time and resources.

Sharing and exporting results

As with many projects, the EIC sequence evaluation was performed by several different members of the group who were each evaluating different aspects. Results were easily exported for further analysis and inclusion in reports – for example to Orbit, PatBase, and even GenomeQuest's own full-text search platform, LifeQuest – and shared among GenomeQuest users in the group. Colleagues assessing freedom to operate were able to utilize the exact same dataset with the same query run timestamp as colleagues assessing patentability. Errors were minimized and risks were reduced. By sharing real-time results with others involved in both analysis and decision making, the team eliminated the need to execute redundant searches or guess which search parameters another searcher might have used. Attorneys rendering opinions were assured that the dataset on which they were basing their decisions included complete and accurate information.

Evaluation of non-patent literature

The same approach that was used for patent document evaluation was also leveraged for non-patent sequence hits, which were then saved to a workfile for assessment of disclosure dates relative to the patent hits. As with the patent search, each record was expandable to access all details, and original entries in NCBI and other resources were easily accessed via direct links.

GenomeQuest was the key for every stage of the project

Attorneys and business managers trusted that the IP search group's information analysis was exhaustive and thorough. GenomeQuest allowed an initial evaluation of patentability and FTO of the Ebola Immune Complex composition and method of use for the R&D group's activities within hours, not days or weeks.

Although some outside counsel and search firms offer GenomeQuest searching services, it was important to the company to have GenomeQuest in-house for accessibility of queries and results, as well as continued re-analysis at different time points and as the underlying databases are updated. Audit trails were kept of every search parameter, execution, and result set. Instant alerts were also set up to identify newly published sequences that would impact existing queries. Sharing between GenomeQuest users, with customizable privileges, meant that group projects were streamlined and seamless. Everyone was working with the same information, in one place.

GenomeQuest was the only solution that could provide all of the content required for a comprehensive analysis and all of the features needed for reporting and decision making. No other IP sequence platform approached the scope, accuracy, and flexibility of GenomeQuest, and making it the only sequence search tool that was needed to conduct FTO and patentability analyses.

Summary Table

Hits to Patient and Non-Patients		Nucleotides	Proteins
Patents	Sequences	366	114
	Patents	124	82
	Patent Families	40	33
	100% Identity	0	2
Non-Patents	Sequences	134	386
	100% Identity	1	1

To learn more or to schedule a demo, email **info@aptean.com**.

References

1 Hefferon K. "Plant-derived pharmaceuticals for the developing world". Biotechnology Journal Special Issue: Plant Biotechnology, Volume 8, Issue 10, pages 1193–1202, October 2013.2.

2 Feldmann H and Geisbert TW. "Ebola haemorrhagic fever". The Lancet, Volume 377, Issue 9768, 5–11 March 2011, Pages 849–862.3

3 Cohen J. "Ebola vaccine: Little and late". Science 19 September 2014. Vol. 345 no. 6203 pp. 1441-14424.

4 Phoolcharoen W, Bhoo SH, Lai H, Ma J, Arntzen CJ, Chen Q, Mason HS. "Expression of an immunogenic Ebola immune complex in Nicotiana benthamiana". Plant Biotechnology Journal. Volume 9, Issue 7, pages 807–816, September 2011.

5 Olinger Jr GG et al., "Delayed treatment of Ebola virus infection with plant-derived monoclonal antibodies provides protection in rhesus macaques". PNAS October 30, 2012 vol. 109 no. 44 pp. 18030-18035.Ebola Glycoprotein