Software for
Business Intelligence

BizInt Smart Charts

Patents & IP Sequences | Clinical Trials | Drug Pipelines

Workshop - Reporting Multiple Sequences in BizInt Smart Charts for Patents

13 January 2021 - incorporating feedback from the workshop

John Willmore, VP Product Development

Diane Webb, President

www.bizint.com

Background

- BizInt Smart Charts is used by many companies to report IP sequence search results.
- The ability to summarize all sequence hits in a family is useful, but limited in a complex multi-query search such as for an antibody.

© 2021 BizInt Solutions, Inc

Purpose of today's workshop

- To explore possible improvements for display of multiple alignments in a report.
- Review several possible changes that we have collected.
- Learn your better ways to report these results.
- Make sure we are meeting your reporting needs.

© 2021 BizInt Solutions, Inc

Today's Topics

- IP Sequence Sources
- Presentation of Multiple Alignments in Charts
- Summary Records
- Query Coverage
- Index of Alignments
- Deduplication of Alignments
- Your feedback

© 2021 BizInt Solutions, Inc

IP Sequence Databases

Provide data on sequences filed in patents

- GenomeQuest (Geneseq, GQ-PAT)
- STN (USGENE, DGENE, PCTGEN)
- CAS Biosequences on GenomeQuest (new in 2020)

© 2021 BizInt Solutions, Inc

Future IP Sequence Databases

Provide data on sequences filed in patents

- Clarivate SequenceBase [Winter 2021]
  Encompasses DGENE, USGENE, PCTGEN
- Orbit BioSequences [Winter 2021]

Both expected to be released soon.

Future IP Sequence Databases

Provide data on sequences filed in patents

- BLAST on REGISTRY (classic STN)
- HITSEQ on STN (?)
- “Annotate” results with data from full REGISTRY records (?)

Basically we are looking at how we can best support STN… let us know what the right approach is

Future IP Sequence Databases

Provide data on sequences filed in patents

- Virtual databases on GenomeQuest
- Bug in GQ Antibody Module export???

- GQ Variations module?
  We aren't sure there is an export available here for this to use

Presenting Multiple Alignments in Charts

- Recipe for creating a subtable from selected columns
- Limitation: no text control (alignments lose structure)
- Feature: include text attributes in subtables
- Feature: save subtables in chart templates

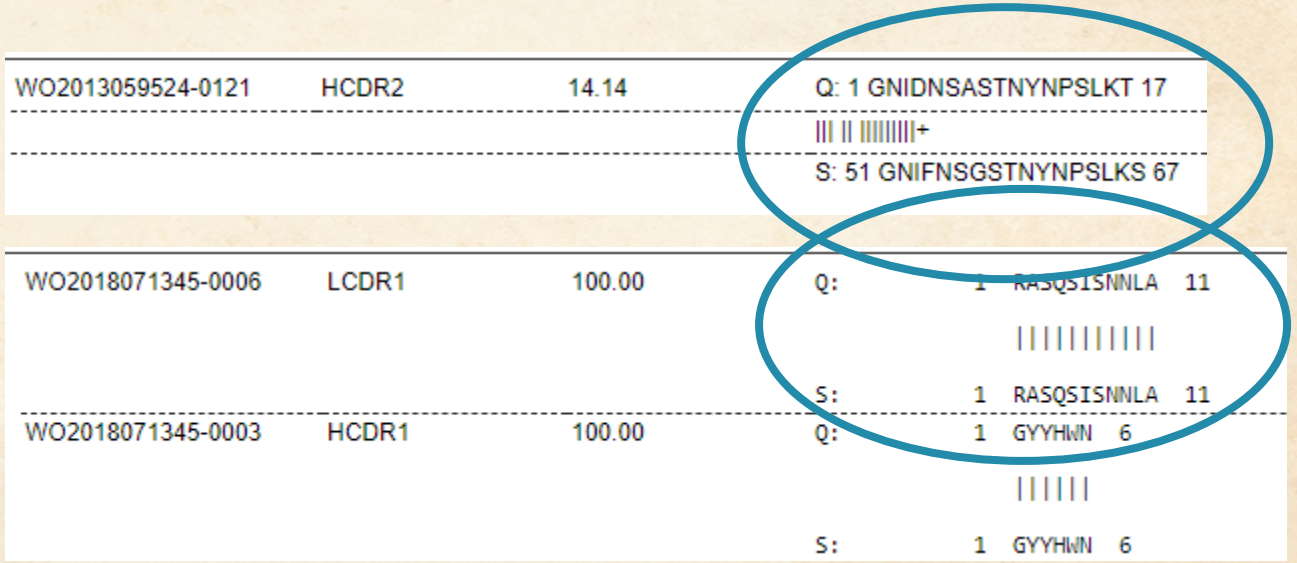
WO2013059524-0121	HCDR2	14.14	Q: 1 GNIDNSASTNYNPSLKT 17
			||| || |||||||+
			S: 51 GNIFNSGSTNYNPSLKS 67

WO2018071345-0006	LCDR1	100.00	Q:	1	RASQSISNNLA 11
					|||||||||||
			S:	1	RASQSISNNLA 11
WO2018071345-0003	HCDR1	100.00	Q:	1	GYYHWN 6
					||||||
			S:	1	GYYHWN 6

© 2021 BizInt Solutions, Inc | www.bizint.com

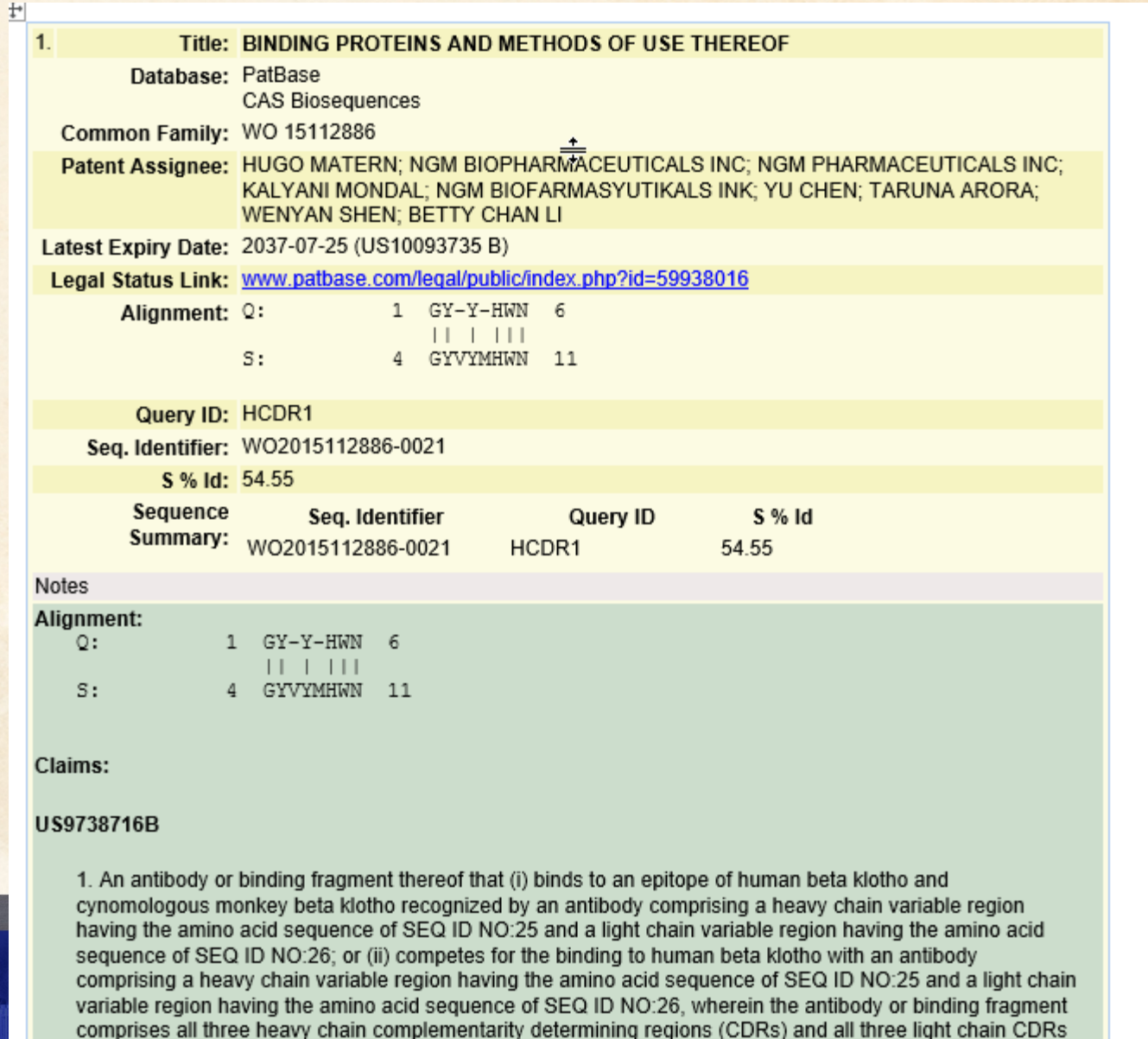
Presenting Multiple Alignments in Charts

- While we're at it…
- Clean up display for tables with multi-line cells
- Better vertical spacing for alignments
- **Consistent** sorting of the rows within a subtable (e.g. by query ID, SEQIDNO, etc)



Multiple Alignments in Summary Records

- Summary Record export has an option to include an alignment
- Fix: Identify the source
- Feature:
  A table of alignments, identifying, other data?
- Note: could include the summarized table from last item



1. **Title:** BINDING PROTEINS AND METHODS OF USE THEREOF

Database: PatBase
CAS Biosequences

Common Family: WO 15112886

Patent Assignee: HUGO MATERN; NGM BIOPHARMACEUTICALS INC; NGM PHARMACEUTICALS INC; KALYANI MONDAL; NGM BIOFARMASYUTIKALS INK; YU CHEN; TARUNA ARORA; WENYAN SHEN; BETTY CHAN LI

Latest Expiry Date: 2037-07-25 (US10093735 B)

Legal Status Link: www.patbase.com/legal/public/index.php?id=59938016

Alignment:
```
Q:          1  GY-Y-HWN  6
               ||  |  |||
S:          4  GYVYMHWN  11
```

Query ID: HCDR1

Seq. Identifier: WO2015112886-0021

S % Id: 54.55

Sequence Summary:

Seq. Identifier	Query ID	S % Id
WO2015112886-0021	HCDR1	54.55

Notes

Alignment:
```
Q:          1  GY-Y-HWN  6
               ||  |  |||
S:          4  GYVYMHWN  11
```

Claims:

US9738716B

1. An antibody or binding fragment thereof that (i) binds to an epitope of human beta klotho and cynomologous monkey beta klotho recognized by an antibody comprising a heavy chain variable region having the amino acid sequence of SEQ ID NO:25 and a light chain variable region having the amino acid sequence of SEQ ID NO:26; or (ii) competes for the binding to human beta klotho with an antibody comprising a heavy chain variable region having the amino acid sequence of SEQ ID NO:25 and a light chain variable region having the amino acid sequence of SEQ ID NO:26, wherein the antibody or binding fragment comprises all three heavy chain complementarity determining regions (CDRs) and all three light chain CDRs

© 2021 BizInt Solutions, Inc | www.bizint.com

Idea: A Table of Unique Alignments

- Suggestion: Include a list of alignments for each family in the green section, hyperlinks from Seq Identifier in the summary/table to the alignments table
- Table: Sequence ID, Query, Alignment, list of equivalent publications (grouped by UFS)
- Might have several templates depending on what fields are needed for a particular search.

© 2021 BizInt Solutions, Inc

Can alignments sit in a separate, linked file?

- An interesting question centered around whether large alignments could be stored somewhere linked.
- A chart table and summary records could be linked, perhaps we could do something similar between exported files.

Workflow idea: a fast way to review alignments to qualify families in the chart

- One idea that came up was having a quick way to quickly scan alignments for a family and use this to annotate the family (either include as a good match or exclude)
- Sort and filter…
- The key question is can the user go from the alignments to the family in the chart.

© 2021 BizInt Solutions, Inc

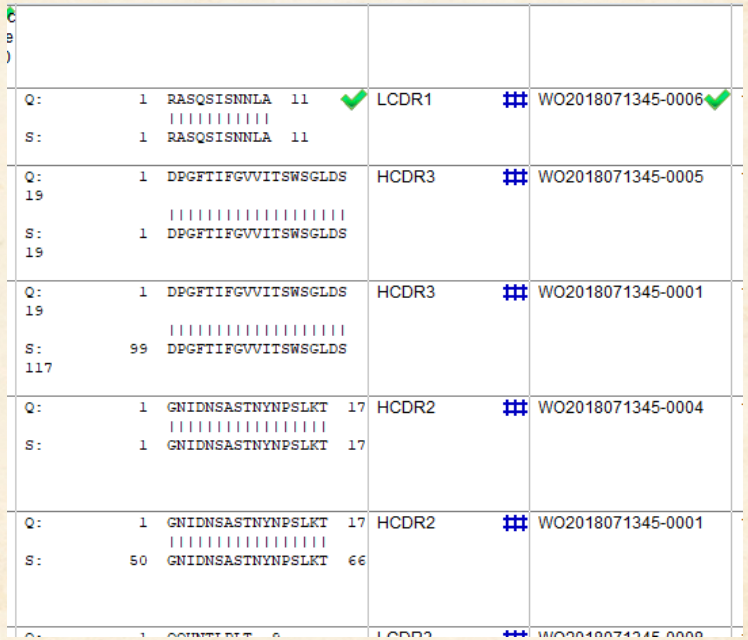
Alignments Embedded in Claims - Summary Records

- We are occasionally asked if we can mix alignments into the flow of claims
- Does this make sense in a family context (placing an alignment from an EP sequence hit into a US claims set)?
- How to handle unclaimed sequences?
- How to handle sequences appearing in multiple claims?
- How to handle multiple alignments in a claim?
- Styling the display
- Suggestion: link to the table of alignments

© 2021 BizInt Solutions, Inc

Query Coverage

- Indicate which queries hit for a family
- Use Summarize Unique rule
- Feature: Sort entries
- Graphical representations? Other representations?
- Idea: show a little graphic of query coverage for each family



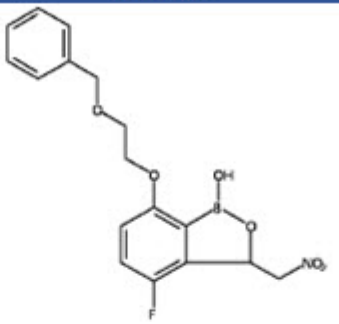
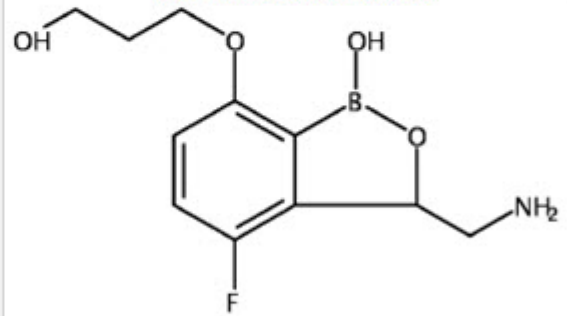
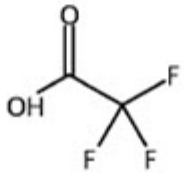
LCDR1
HCDR3
HCDR2
LCDR3
HCDR1
LCDR2
LC-Ebola
HC-Ebola

© 2021 BizInt Solutions, Inc | www.bizint.com

Index of Alignments

- Index of Hit Structures can link to multiple families with the same chemical structure
- What about an Index of Sequences or Index of Alignments?
- Reference to indicate query coverage (for example)
- What determines when two sequences are the same?

Index of Hit Structures

	Substance	Structure	Reference
1	1655492-02-6 2,1-Benzoxaborole, 4-fluoro-1,3-dihydro-1-hydroxy-3-(nitromethyl)-7-[2-(phenylmethoxy)ethoxy]-		prepn. and anti-mycobacterial activity of benzoxaborole compds. *Reference 1* prepn. and biol. applications of tricyclic benzoxaborole compds. *Reference 2*
2	1364682-96-1 1-Propanol, 3-[[3-(aminomethyl)-4-fluoro-1,3-dihydro-1-hydroxy-2,1-benzoxaborol-7-yl]oxy]-, 2,2,2-trifluoroacetate (1:2)	CM1 CRN 1364682-95-0  CM2 CRN 76-05-1 	prepn. of benzoxaborole derivs. useful for treating bacterial infections *Reference 3*

Do we need to remove duplicates?

- Any thinning of hits in a group needs to be optional
- One hit per query? “Best” hit per query?
- Grouping by family (current behavior), publication, sequence?
- Unique Family Sequence (GQ)

- Need a way to manually select sequences to include/exclude from a report.

© 2021 BizInt Solutions, Inc

Manually selecting which sequence to display?

- There was a discussion during the workshop of thinning hits based on authority or grant vs application (like we do for selecting claims – the rules should be the same for claims)
- This ties in with the discussion in summary records, the representative publication in the table is based on rules
- Other publications show up in the equivalents list

© 2021 BizInt Solutions, Inc

On a related topic

- Filtering patent families
- Could the presence of a sequence hit be used? Basically only show members with hits

Software for
Business Intelligence

BizInt Smart Charts

support@bizint.com

Feedback?