Software for
Business Intelligence

BizInt Smart Charts

Patents & IP Sequences | Clinical Trials | Drug Pipelines

Creating IP Sequence Reports from Multiple Sources

Based on our workshop at the PIUG Biotechnology Conference

17 March 2020

John Willmore, VP Product Development

www.bizint.com

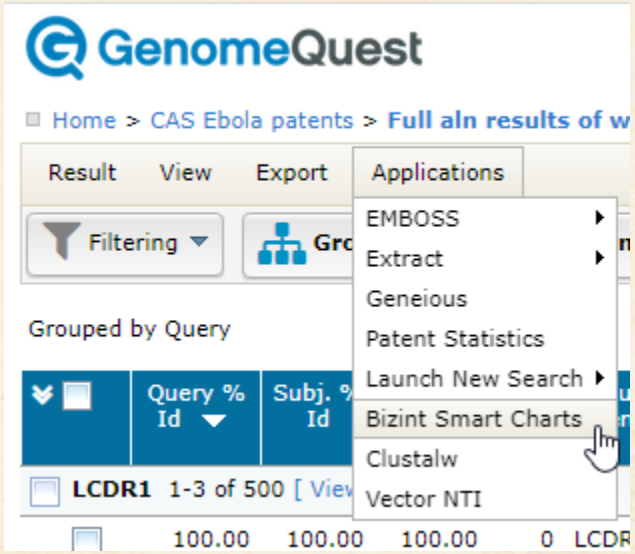
Today's Agenda

- CAS Biosequences on GenomeQuest
- Transferring publication numbers to PatBase
- Combining reports
- Identify Common Patent Family
- Reference Rows
- Creating a Summary Table of Sequence Hits
- Exports, including Summary Records

© 2020 BizInt Solutions, Inc

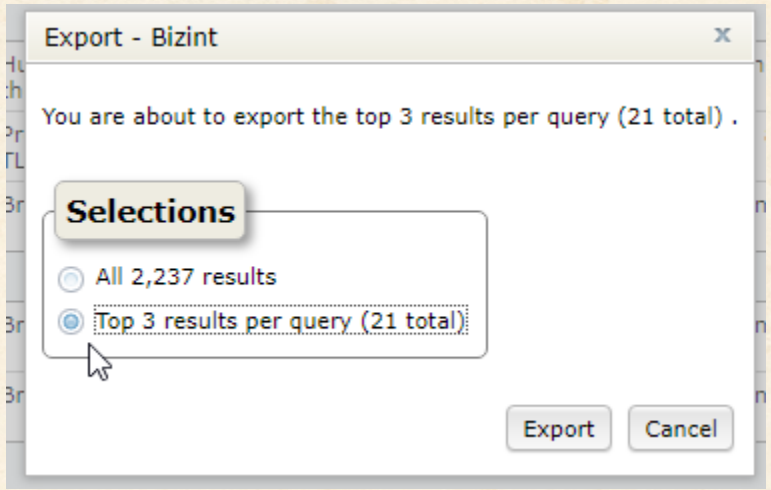
CAS Biosequences on GenomeQuest

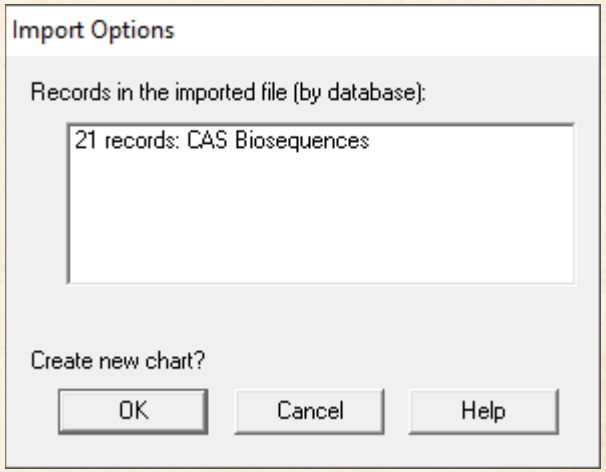
- New support in Version 5.3.2 (patent and non-patent)
- Use the BizInt application to export



CAS Biosequences on GenomeQuest

- Filters and selections now carry over to exports

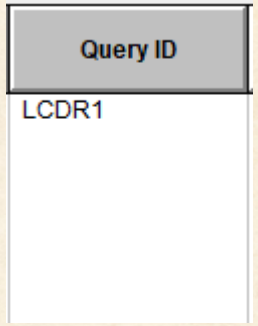




© 2020 BizInt Solutions, Inc | www.bizint.com

Multi-query searches on GenomeQuest

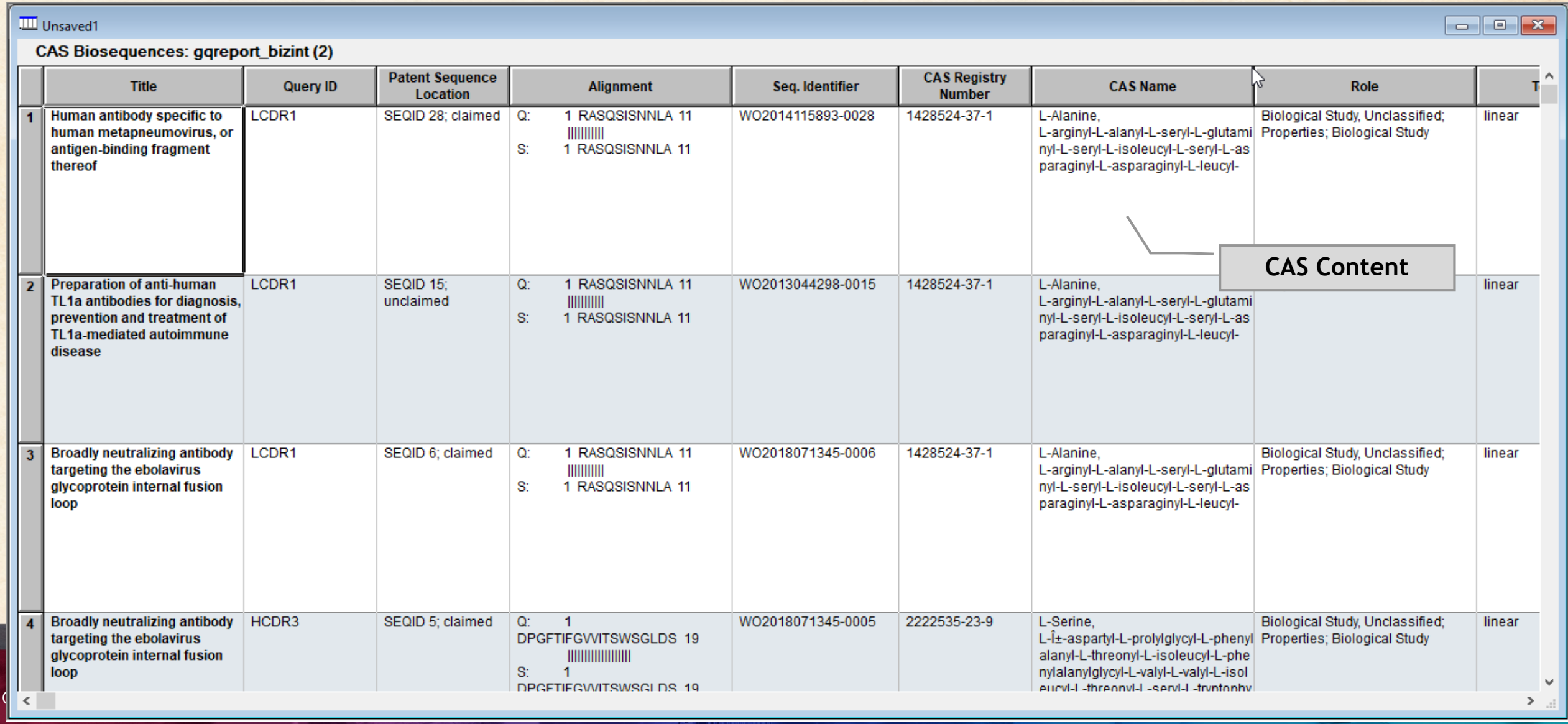
- Query labels are available in the QueryID column e.g. LC, LCDR1/2/3…



- If you run each query as a separate search, fill the Query ID in each chart by selecting the column and pasting in the new value.
- Then combine.

© 2020 BizInt Solutions, Inc

CAS Biosequences


Unsaved1

CAS Biosequences: gqreport_bizint (2)

	Title	Query ID	Patent Sequence Location	Alignment	Seq. Identifier	CAS Registry Number	CAS Name	Role	T
1	Human antibody specific to human metapneumovirus, or antigen-binding fragment thereof	LCDR1	SEQID 28; claimed	Q: 1 RASQSISNNLA 11 |||||||||| S: 1 RASQSISNNLA 11	WO2014115893-0028	1428524-37-1	L-Alanine, L-arginyl-L-alanyl-L-seryl-L-glutaminyl-L-seryl-L-isoleucyl-L-seryl-L-asparaginyl-L-asparaginyl-L-leucyl-	Biological Study, Unclassified; Properties; Biological Study	linear
2	Preparation of anti-human TL1a antibodies for diagnosis, prevention and treatment of TL1a-mediated autoimmune disease	LCDR1	SEQID 15; unclaimed	Q: 1 RASQSISNNLA 11 |||||||||| S: 1 RASQSISNNLA 11	WO2013044298-0015	1428524-37-1	L-Alanine, L-arginyl-L-alanyl-L-seryl-L-glutaminyl-L-seryl-L-isoleucyl-L-seryl-L-asparaginyl-L-asparaginyl-L-leucyl-		linear
3	Broadly neutralizing antibody targeting the ebolavirus glycoprotein internal fusion loop	LCDR1	SEQID 6; claimed	Q: 1 RASQSISNNLA 11 |||||||||| S: 1 RASQSISNNLA 11	WO2018071345-0006	1428524-37-1	L-Alanine, L-arginyl-L-alanyl-L-seryl-L-glutaminyl-L-seryl-L-isoleucyl-L-seryl-L-asparaginyl-L-asparaginyl-L-leucyl-	Biological Study, Unclassified; Properties; Biological Study	linear
4	Broadly neutralizing antibody targeting the ebolavirus glycoprotein internal fusion loop	HCDR3	SEQID 5; claimed	Q: 1 DPGFTIFGVVITSWSGLDS 19 ||||||||||||||||||| S: 1 DPGFTIFGVVITSWSGLDS 19	WO2018071345-0005	2222535-23-9	L-Serine, L-Î±-aspartyl-L-prolylglycyl-L-phenylalanyl-L-threonyl-L-isoleucyl-L-phenylalanylglycyl-L-valyl-L-valyl-L-isoleucyl-L-threonyl-L-seryl-L-tryptophy	Biological Study, Unclassified; Properties; Biological Study	linear

CAS Content

CAS Biosequences (non-patent content)

Unsaved2

gq_cas_journals

	Title	Source	Biosequence Modifications		Seq. Identifier	Subject Pct Identity
			Position	Modification		
7	**Nucleotide sequence of the tcmII-tcmIV region of the tetracenomycin C biosynthetic gene cluster of Streptomyces glaucescens and evidence that the tcmN gene encodes a multifunctional cyclase-dehydratase-O-methyl transferase**	Summers, Richard G.; Wendt-Pienkowski, Evelyn; Motamedi, Haideh; Hutchinson, C. R.. 1992. Nucleotide sequence of the tcmII-tcmIV region of the tetracenomycin C biosynthetic gene cluster of Streptomyces glaucescens and evidence that the tcmN gene encodes a multifunctional cyclase-dehydratase-O-methyl transferase. Journal of Bacteriology 174(6):1810-20	1	terminal mod., methionine-1, N-formyl	1992:485843-142845-47-4	2.06
8	**Analysis of expressed sequence tags derived from developing seed and pollen cones of Cryptomeria japonica**	Ujino-Ihara, T.; Taguchi, Y.; Yoshimura, K.; Tsumura, Y.. 2003. Analysis of expressed sequence tags derived from developing seed and pollen cones of Cryptomeria japonica. Plant Biology (Stuttgart, Germany) 5(6):600-607			2004:186071-576510-15-1	5.70
9	**Large-scale analysis of the barley transcriptome based on expressed sequence tags**	Zhang, Hangning; Sreenivasulu, Nese; Weschke, Winfriede; Stein, Nils; Rudd, Stephen; Radchuk, Volodymyr; Potokina, Elena; Scholz, Uwe; Schweizer, Patrick; Zierold, Uwe; Langridge, Peter; Varshney, Rajeev K.; Wobus, Ulrich; Graner, Andreas. 2004.			2004:1000626-427324-98-9	5.62

Sequence Modifications

© 2020 BizInt

CAS Biosequences

Unsaved1

CAS Biosequences: gqreport_bizint (2)

	Title	Query ID	Patent Sequence Location	Alignment	Seq. Identifier	CAS Registry Number	CAS Name	Role	T
1	**Human antibody specific to human metapneumovirus, or antigen-binding fragment thereof**	LCDR1	SEQID 28; claimed	Q: 1 RASQSISNNLA 11 |||||||||| S: 1 RASQSISNNLA 11	WO2014115893-0028	1428524-37-1	L-Alanine, L-arginyl-L-alanyl-L-seryl-L-glutaminyl-L-seryl-L-isoleucyl-L-seryl-L-asparaginyl-L-asparaginyl-L-leucyl-	Biological Study, Unclassified; Properties; Biological Study	linear
2	**Preparation of anti-human TL1a antibodies for diagnosis, prevention and treatment of TL1a-mediated autoimmune disease**	LCDR1	SEQID 15; unclaimed	Q: 1 RASQSISNNLA 11 |||||||||| S: 1 RASQSISNNLA 11	WO2013044298-0015	1428524-37-1	L-Alanine, L-arginyl-L-alanyl-L-seryl-L-glutaminyl-L-seryl-L-isoleucyl-L-seryl-L-asparaginyl-L-asparaginyl-L-leucyl-	Properties	linear
3	**Broadly neutralizing antibody targeting the ebolavirus glycoprotein internal fusion loop**	LCDR1	SEQID 6; claimed	Q: 1 RASQSISNNLA 11 |||||||||| S: 1 RASQSISNNLA 11	WO2018071345-0006	1428524-37-1	L-Alanine, L-arginyl-L-alanyl-L-seryl-L-glutaminyl-L-seryl-L-isoleucyl-L-seryl-L-asparaginyl-L-asparaginyl-L-leucyl-	Biological Study, Unclassified; Properties; Biological Study	linear
4	**Broadly neutralizing antibody targeting the ebolavirus glycoprotein internal fusion loop**	HCDR3	SEQID 5; claimed	Q: 1 DPGFTIFGVVITSWSGLDS 19 ||||||||||||||||||| S: 1 DPGFTIFGVVITSWSGLDS 19	WO2018071345-0005	2222535-23-9	L-Serine, L-Î±-aspartyl-L-prolylglycyl-L-phenylalanyl-L-threonyl-L-isoleucyl-L-phenylalanylglycyl-L-valyl-L-valyl-L-isoleucyl-L-threonyl-L-seryl-L-tryptophy	Biological Study, Unclassified; Properties; Biological Study	linear

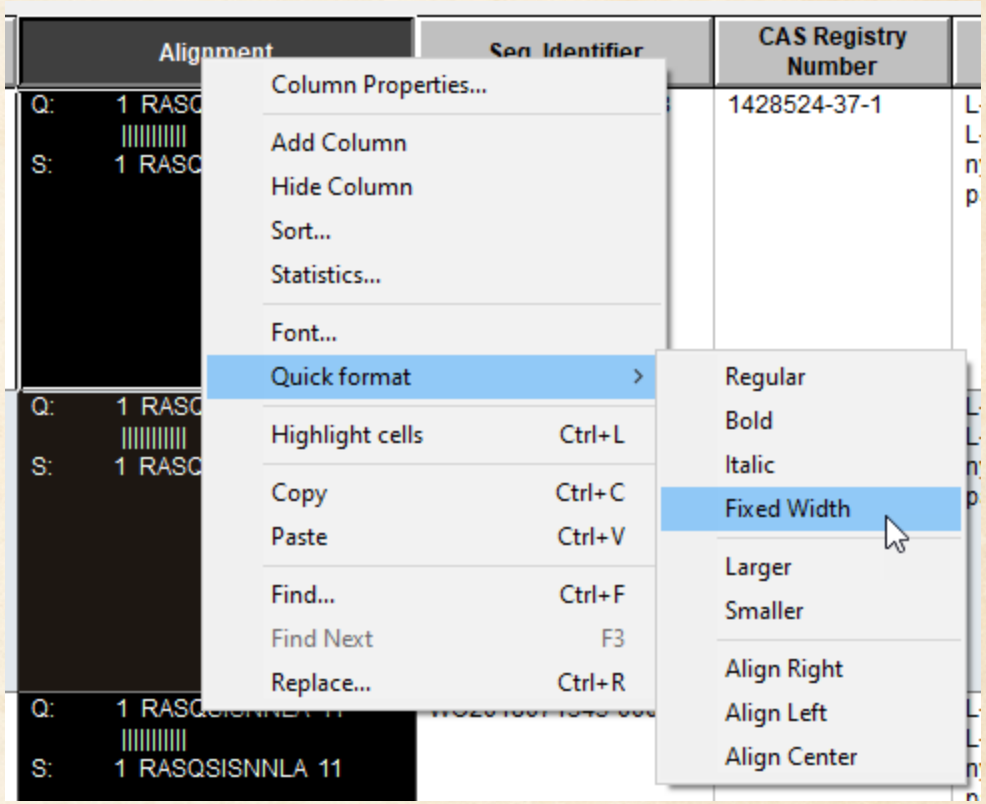
GQ Content

Alignments

Formatting for Alignments

- Choose the Fixed Width option for Quick Format
- Don’t just change the font for the column … that looks ok in BizInt Smart Charts, but loses spaces when exported!

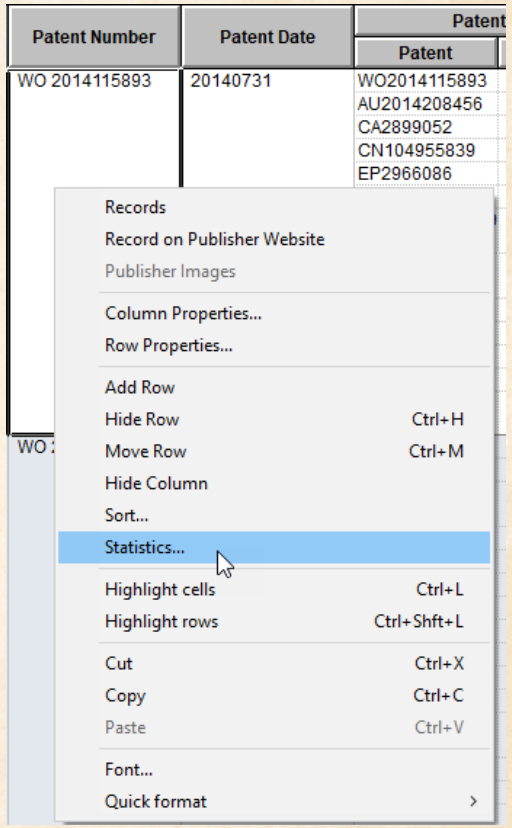
Alignment
Q: 1 RASQSISNNLA 11 ||||||||||| S: 1 RASQSISNNLA 11



Transferring publication numbers to PatBase

- In a small set like this, collect all publication numbers
- Tools | Statistics creates an Excel sheet

- Copy and Paste numbers into PatBase “Upload publication numbers” panel



© 2020 BizInt Solutions, Inc | www.bizint.com

PatBase – export results, enhanced legal status support

- Creating Reports from Databases and Hosts for instructions on each platform

Latest Expiry Date	Family Status Expiry			Latest Legal Status Table			Legal Status Link
	Pub No.	State	Exp Date	Pub	Date	Event	
2037-07-25 (US10093735 B)	US 2015210764 A	ALIVE		EP3097122 A2	2017-05-03	(DAX) REQUEST FOR EXTENSION OF THE EUROPEAN PATENT (TO ANY COUNTRY) (DELETED)	http://www.patbase.com/legal/public/index.php?id=59938016
	US 9738716 B	ALIVE	2035-03-01				
	US 2018100014 A	ALIVE					
	US 10093735 B	ALIVE	2037-07-25				
	US 2019106490 A	ALIVE		EP3097122 A4	2019-12-18	+ (LSGT / INTG) INTENTION TO GRANT ANNOUNCED	
	W						
	W						
	CA 2937898 AA	ALIVE		US10093735 B	2018-11-13	(LSRE / CC) CERTIFICATE OF CORRECTION	
	AU 2015209131 AA	ALIVE					
	SG 11201606018U A1	ALIVE			8-02	(LSGT / STCF) INFORMATION ON STATUS: PATENT GRANT	
	SG 10201806108T A1	ALIVE		US2018100014 A	2018-09-19	(LSGT / STCF) INFORMATION ON STATUS: PATENT GRANT	
	IL 246921 A0	ALIVE					
	IL 246921 A1	ALIVE					
	IN 201647028679 A	ALIVE					
	KR 20160125381 A	ALIVE		US2019106490 A	2019-12-10	(LSES / N Action Ma	
	PH 2016501644 A	ALIVE		US9738716 B	2018-05-23	(LSFE / F PAYMENT PROCEDURE	
	CO 20160000995 A2	ALIVE					
	EP 3097122 A2	ALIVE		WO15112886 A2	2015-09-16	(121) EP: THE EPO HAS BEEN INFORMED BY WIPO THAT EP WAS DESIGNATED IN THIS APPLICATION	
	EP 3097122 A4	ALIVE					
	MX 2016009555 A1	ALIVE					
	JP 2017507652 T2	ALIVE					
	PE 20170256 A1	ALIVE		WO15112886 A3	2017-10-03	(LSNP / ENP) ENTRY	

Latest Expiry Date

Family Status with Expiry Dates

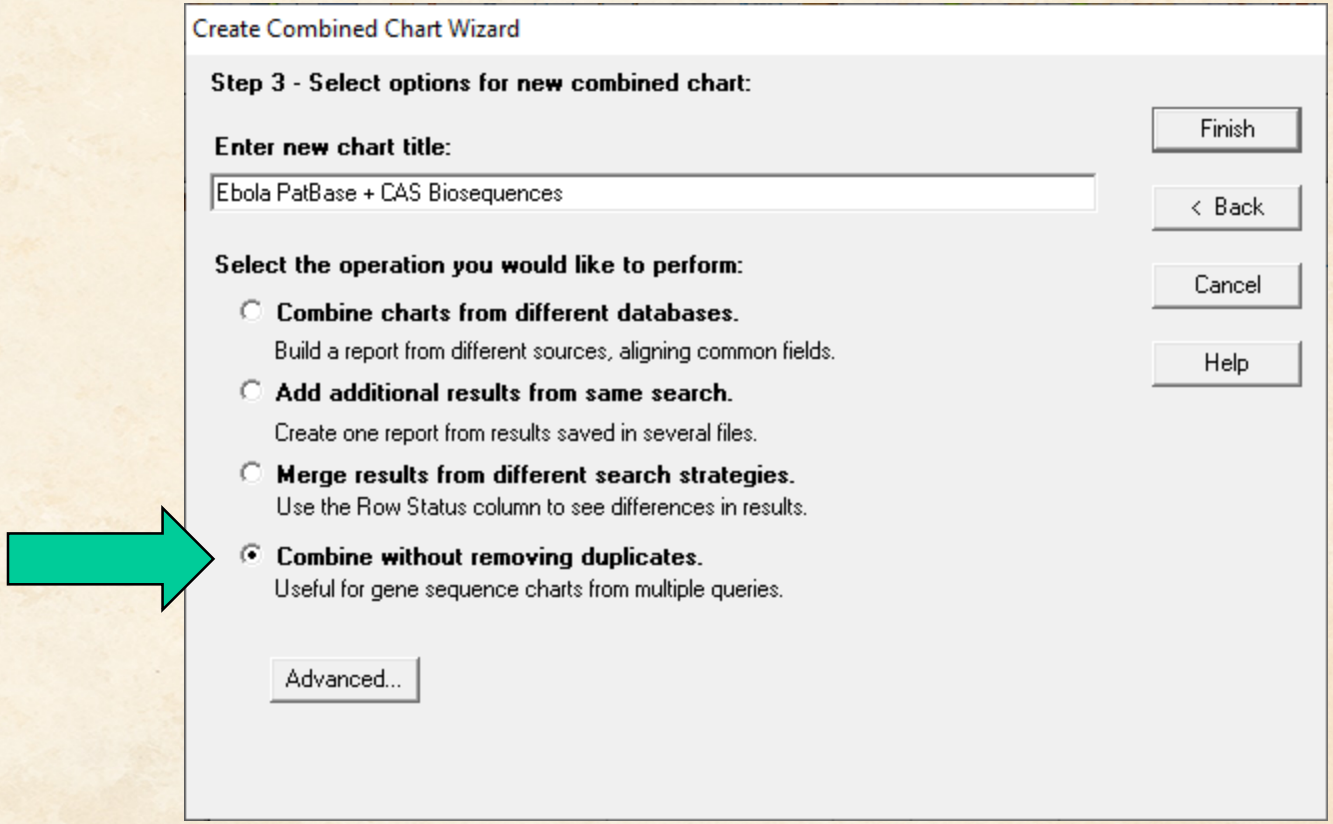
Latest Status per publication

Legal status browser

© 2020 BizIn

Combine Charts

- Both charts open
- File | Combine



Match records by Common Patent Family

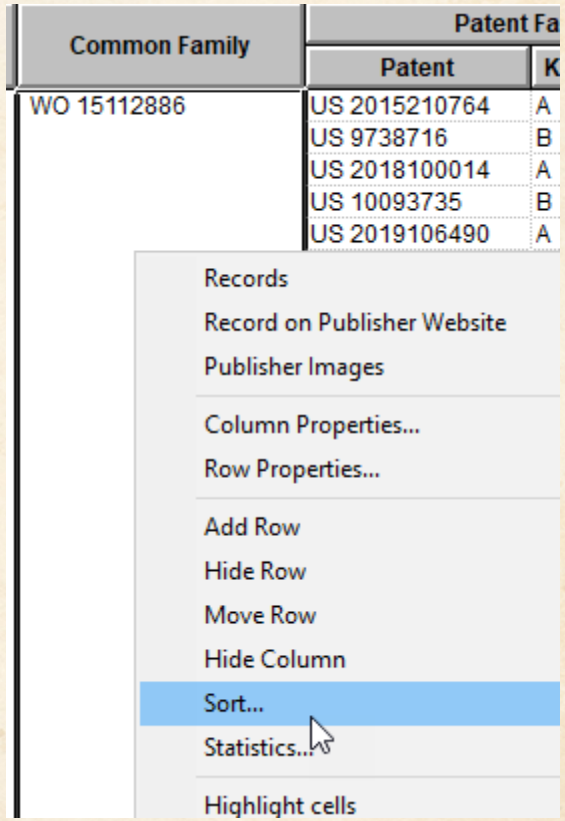
- Links rows in the chart based on publication numbers in families
- Remember:
  Common Family is a sort key

Database	Common Family	Patent Family		
		Patent	Kind	Date
Derwent World Patents Index	US 2014356956	US20140356959	A1	20141204
Derwent World Patents Index	US 2014356956	US20140356956	A1	20141204
		[illegible]	A2	20141211
		WO2014197568	A3	20150312
		CA2914638	A1	20141211
FAMPAT	US 2014356956	~~US 2014356956~~	A1	2014-12-04
		~~US 2014356959~~	A1	2014-12-04
		[illegible]	B2	2016-02-23
GQPAT Gold+ Proteins	US 2014356956	US20140356959		20141204
GQPAT Gold+ Proteins	US 2014356956	US20140356956		20141204
PatBase	US 2014356956	US 2014356959	A	2014-12-04
		US 2014356956	A	2014-12-04
		AU 2014274939	AA	2014-12-11
		WO 14197568	A2	2014-12-11
		WO 14197568	A3	2015-03-12
		CA 2914638	AA	2015-12-04
		KR 20160014036	A	2016-02-05

© 2020 BizInt Solutions, Inc

Common Patent Family

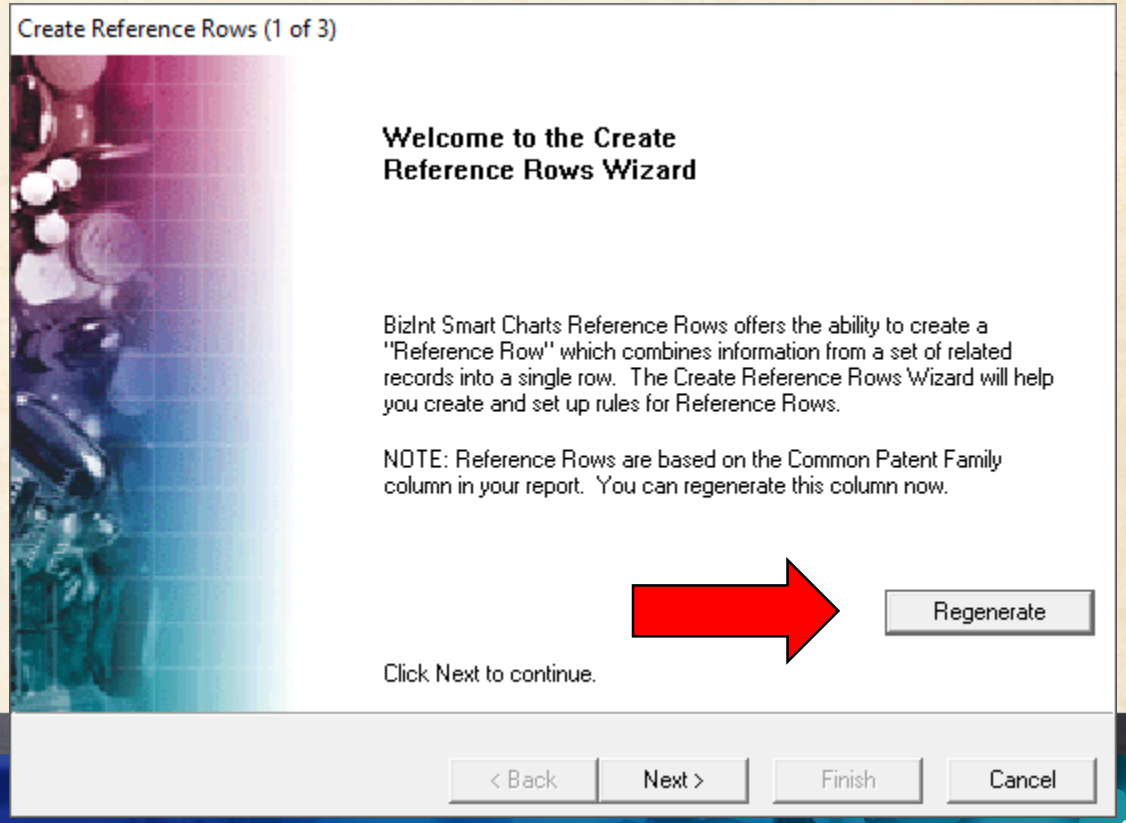
- Remember that Common Family is simply a sort key
- You can edit the assigned values
- You can paste another field into Common Family
- Patent Number – group by publication
- Sequence ID – group by sequence



© 2020 BizInt Solutions, Inc | www.bizint.com

Send to Reference Rows

- Save the combined chart first
- File | Send to Reference Rows
- On the first step of the wizard, do NOT select “Regenerate” if you have modified Common Family



© 2020 BizInt Solutions, Inc | www.bizint.com

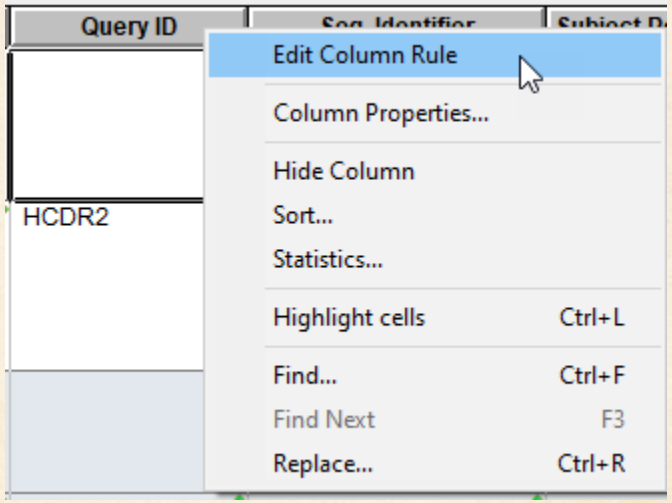
Standard behavior - fill in the blanks

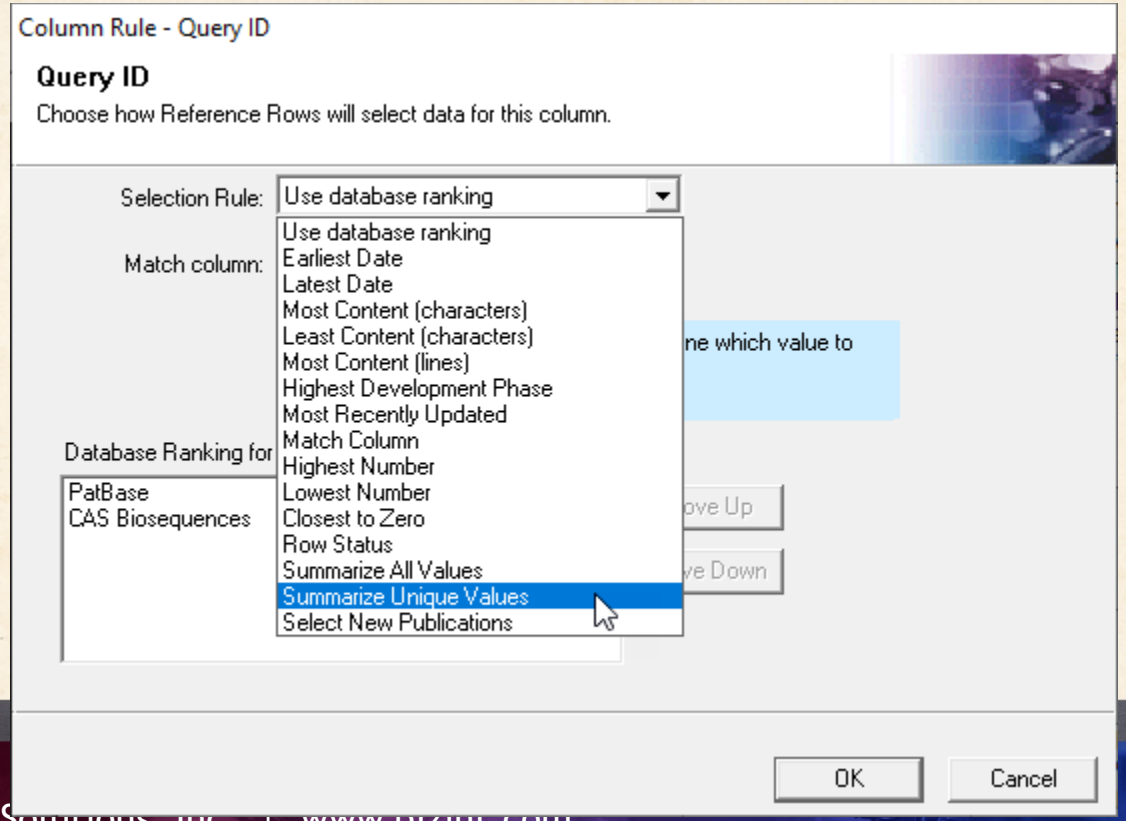
- Check marks show representative values for each group

5.1	BROADLY NEUTRALIZING ANTIBODY TARGETING THE EBOLAVIRUS GLYCOPROTEIN INTERNAL FUSION LOOP	PatBase	WO 18071345	INTEGRATED BIOTHERAPEUTICS INC UNIV OF MARYLAND		http://www.patbase.com/legal/public/index.php?id=69341340					
5.2	Broadly neutralizing antibody targeting the ebolavirus glycoprotein internal fusion loop	CAS Biosequences	WO 18071345	Integrated BioTherapeutics, Inc.; The University of Maryland			Q: 1 RASQSISNNLA 11 ||||||||||| S: 1 RASQSISNNLA 11	LCDR1	WO2018071345-0006	100.00	
5.3	Broadly neutralizing antibody targeting the ebolavirus glycoprotein internal fusion loop	CAS Biosequences	WO 18071345	Integrated BioTherapeutics, Inc.; The University of Maryland			Q: 1 DPGFTIFGVVITSW GLDS 19 |||||||||||||| |||| S: 1 DPGFTIFGVVITSW GLDS 19	HCDR3	WO2018071345-0005	100.00	
5.4	Broadly neutralizing antibody targeting the ebolavirus glycoprotein internal fusion loop	CAS Biosequences	WO 18071345	Integrated BioTherapeutics, Inc.; The University of Maryland			Q: 1 DPGFTIFGVVITSW GLDS 19 |||||||||||||| |||| S: 99 DPGFTIFGVVITSW GLDS 117	HCDR3	WO2018071345-0001	14.84	
5.5	Broadly neutralizing antibody targeting the ebolavirus glycoprotein internal fusion loop	CAS Biosequences	WO 18071345	Integrated BioTherapeutics, Inc.; The University of Maryland			Q: 1 GNIDNSASTNYNPS KT 17 |||||||||||||| || S: 1 GNIDNSASTNYNPS KT 17	HCDR2	WO2018071345-0004	100.00	
5.6	Broadly neutralizing antibody targeting the ebolavirus glycoprotein internal fusion loop	CAS Biosequences	WO 18071345	Integrated BioTherapeutics, Inc.; The University of Maryland			Q: 1 GNIDNSASTNYNPS KT 17 |||||||||||||| || S: 50 GNIDNSASTNYNPS KT 66	HCDR2	WO2018071345-0001	13.28	

Show all Query IDs for a Family

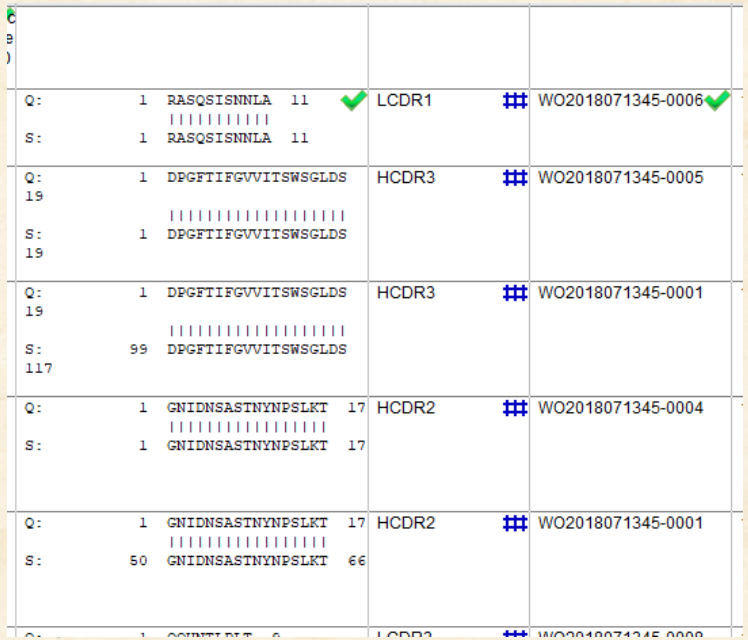
- Edit Column Rule for Query ID
- Choose Summarize Unique Values



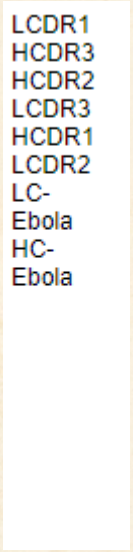


Summarize Unique Values

- Cell Glyph Changes
- Export or Statistics to see value



Q: 1 RASQSISNNLA 11 S: 1 RASQSISNNLA 11	LCDR1	WO2018071345-0006
Q: 1 DPGFTIFGVVITSWSGLDS 19 S: 1 DPGFTIFGVVITSWSGLDS 19	HCDR3	WO2018071345-0005
Q: 1 DPGFTIFGVVITSWSGLDS 19 S: 99 DPGFTIFGVVITSWSGLDS 117	HCDR3	WO2018071345-0001
Q: 1 GNIDNSASTNYNPSLKT 17 S: 1 GNIDNSASTNYNPSLKT 17	HCDR2	WO2018071345-0004
Q: 1 GNIDNSASTNYNPSLKT 17 S: 50 GNIDNSASTNYNPSLKT 66	HCDR2	WO2018071345-0001



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

Create Subtable to Summarize Sequence Results

- But that solution only gives a list of Query IDs associated with the family – it doesn't tell you which sequence was returned with each query.

LCDR1	WO2018071345-0006	100.00
HCDR3	WO2018071345-0005	100.00
HCDR3	WO2018071345-0001	14.84
HCDR2	WO2018071345-0004	100.00
HCDR2	WO2018071345-0001	13.28
LCDR3	WO2018071345-0008	100.00
LCDR3	WO2018071345-0002	8.49

© 2020 BizInt Solutions, Inc

Create Subtable to Summarize Sequence Results

- The answer is to create a table showing data from different sequence hits.
- Complete steps are available in a recipe at bizint.com/piugbio :
- Create subtable from columns
- Change rule to Summarize All Values
- Export

WO2018071345-0006	LCDR1	100.00	5.2
WO2018071345-0005	HCDR3	100.00	5.3
WO2018071345-0001	HCDR3	14.84	5.4
WO2018071345-0004	HCDR2	100.00	5.5
WO2018071345-0001	HCDR2	13.28	5.6
WO2018071345-0008	LCDR3	100.00	5.7
WO2018071345-0002	LCDR3	8.49	5.8
WO2018071345-0003	HCDR1	100.00	5.9
WO2018071345-0007	LCDR2	100.00	5.10
WO2018071345-0002	LC-Ebola	100.00	5.11
WO2018071345-0001	HC-Ebola	100.00	5.12

© 2020 BizInt Solutions, Inc

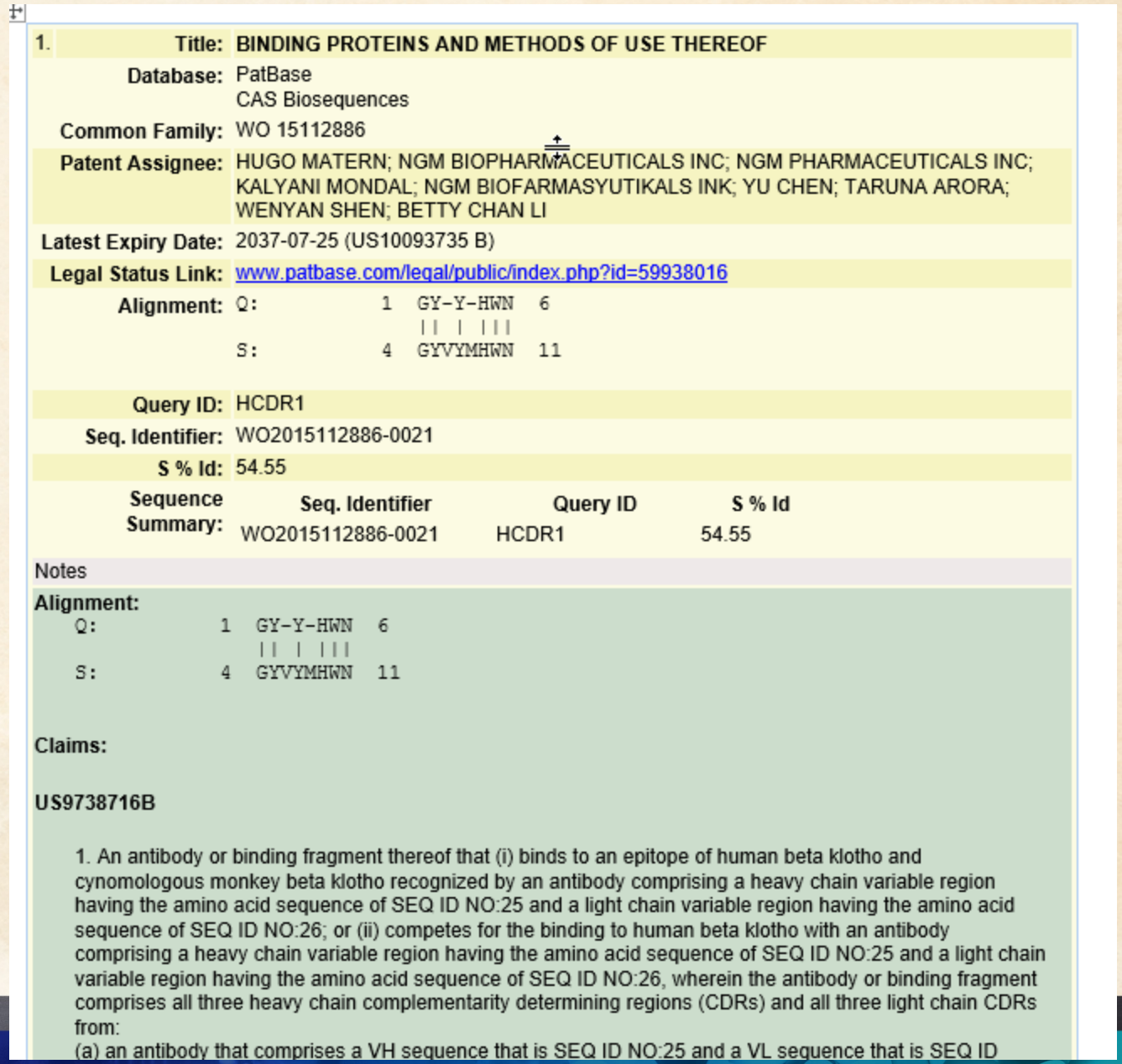
Key Tips of Summarized Subtable

- Rename columns before creating subtable
  you can’t change them after the fact
- It is best to include a column which helps identify the data in each row (Sequence ID in the example)
- Fixed Width property of the Alignment column doesn’t work in a subtable

© 2020 BizInt Solutions, Inc

Summary Records export

- A Word export containing content of the chart, full claims set, full alignment from selected records
- Yellow section contains columns of the chart
- Green section can contain one alignment, one full set of claims



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 from:
(a) an antibody that comprises a VH sequence that is SEQ ID NO:25 and a VL sequence that is SEQ ID

© 2020 BizInt Solutions, Inc | www.bizint.com

Resources

- bizint.com/tips for links to key documentation
- BizInt Smart Charts for Patents Mini Guide
- Bizint.com/piugbio for the recipe handout

© 2020 BizInt Solutions, Inc

BizInt Smart Charts

for Patents

VERSION 5

THE JOURNEY CONTINUES...

Thank you!

www.bizint.com