Software for Business Intelligence

BizInt Smart Charts

Patents & IP Sequences | Clinical Trials | Drug Pipelines

PIUG®

Creating IP Reports Integrating Sequence, Family, and Hit Structure data with BizInt Smart Charts

PIUG 2019 Biotechnology Conference, Boston MA

John Willmore

19 February 2019

THE JOURNEY BEGINS...

doxifun.com/puppies

Agenda

- IP Sequence reports
- Hit Structure reports
- Summary Record exports
- Reference Rows

BizInt.com/slides for other presentations

Agenda

- IP Sequence reports
- Hit Structure reports
- Summary Record exports
- Reference Rows

BizInt Smart Charts

VERSION 5

for Patents

IP Sequence Databases

Provide data on sequences filed in patents

- GQ Life Sciences GenomeQuest (Geneseq, GQ-PAT)
- STN (USGENE, DGENE, PCTGEN)

Sequence Databases on STN

- DGENE, USGENE, PCTGEN
- Jim Brown's workshop this morning provided step-by-step instructions for searching
- D BIB SCORE ALIGN
- Import your transcripts into BizInt Smart Charts
- Each row is a sequence from a patent pub.
- Don't skip Validate Sequences step!

Sequence Databases on STN

Combined: sequences2019

	Title	Database	Sequence ID	Patent Sequence Location	Score
29	New vector comprising promoter and polynucleotide useful for producing fusion protein and target protein e.g. enzyme, blood protein, binding protein, hormone, synthetic protein and peptide.	Derwent GeneSeq	US20110151514-0005	Example; SEQ ID NO 5	44 2% of query self score 2022
30	Vectors, Methods, Systems and Kits for Protein Purification (PublishedApplication)	USGENE	US20110151514-0005	SEQ ID NO 5	44 2% of query self score 2022
31	New vector comprising promoter and polynucleotide useful for producing fusion protein and target protein e.g. enzyme, blood protein, binding protein, hormone, synthetic protein and peptide.	Derwent GeneSeq	US20110151514-0010	Example; SEQ ID NO 10	44 2% of query self score 2022
32	Vectors, Methods, Systems and Kits for Protein Purification (PublishedApplication)	USGENE	US20110151514-0010	SEQ ID NO 10	44 2% of query self score 2022
33	Genetic resistance prediction against antimicrobial drugs in microorganism using structural changes in the genome	PCTGEN	WO20170021529-0144		109 5% of query self score 2022
34	Determining structural variations in genome of microorganism particularly bacterial microorganism, comprising change in genome, by obtaining first data set gene sequences of clinical isolates, and correlating data with second data set.	Derwent GeneSeq	WO20170021529-0144	Disclosure; SEQ ID NO 144	109 5% of query self score 2022

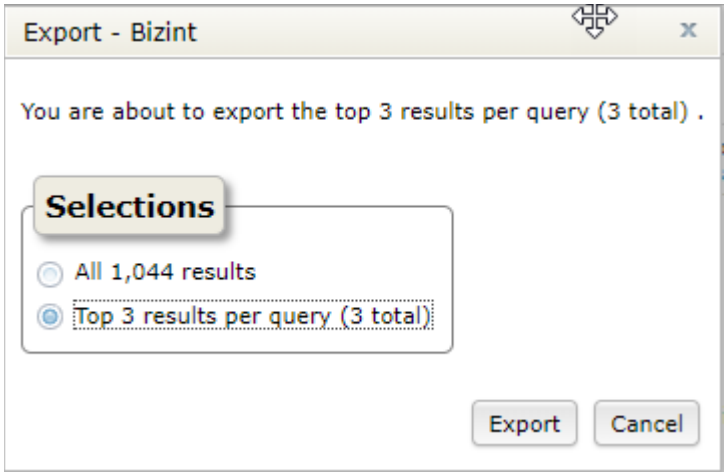
Sequence Databases on STN

- CAplus

- Not supported … yet
- Save both your transcript…
- D BIB HITRN
- …and your Alignment report (.xss)

Sequence Databases on GenomeQuest

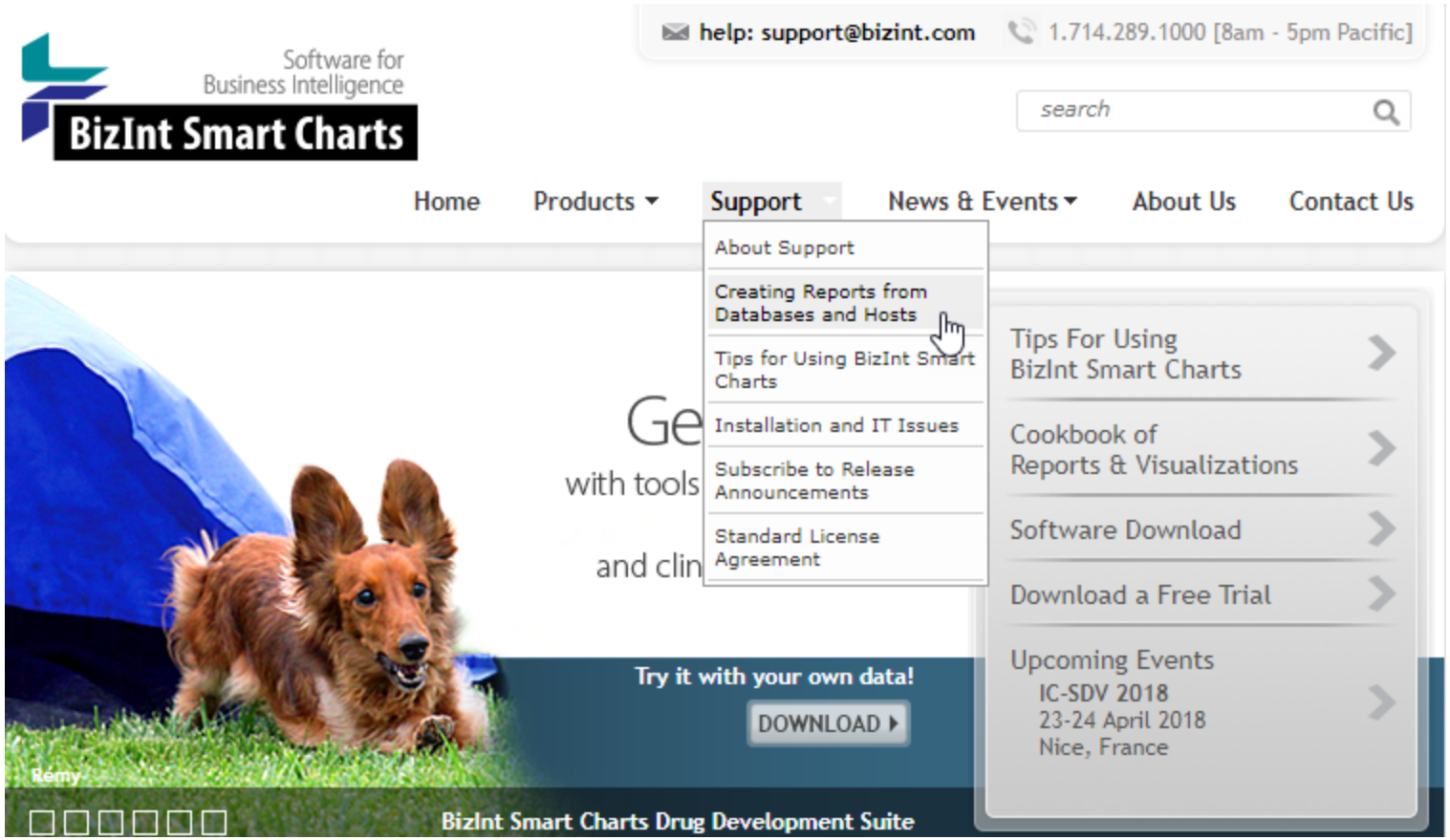
- GQPat, GENESEQ
- Results filters now apply to BizInt exports (e.g. top results per query)



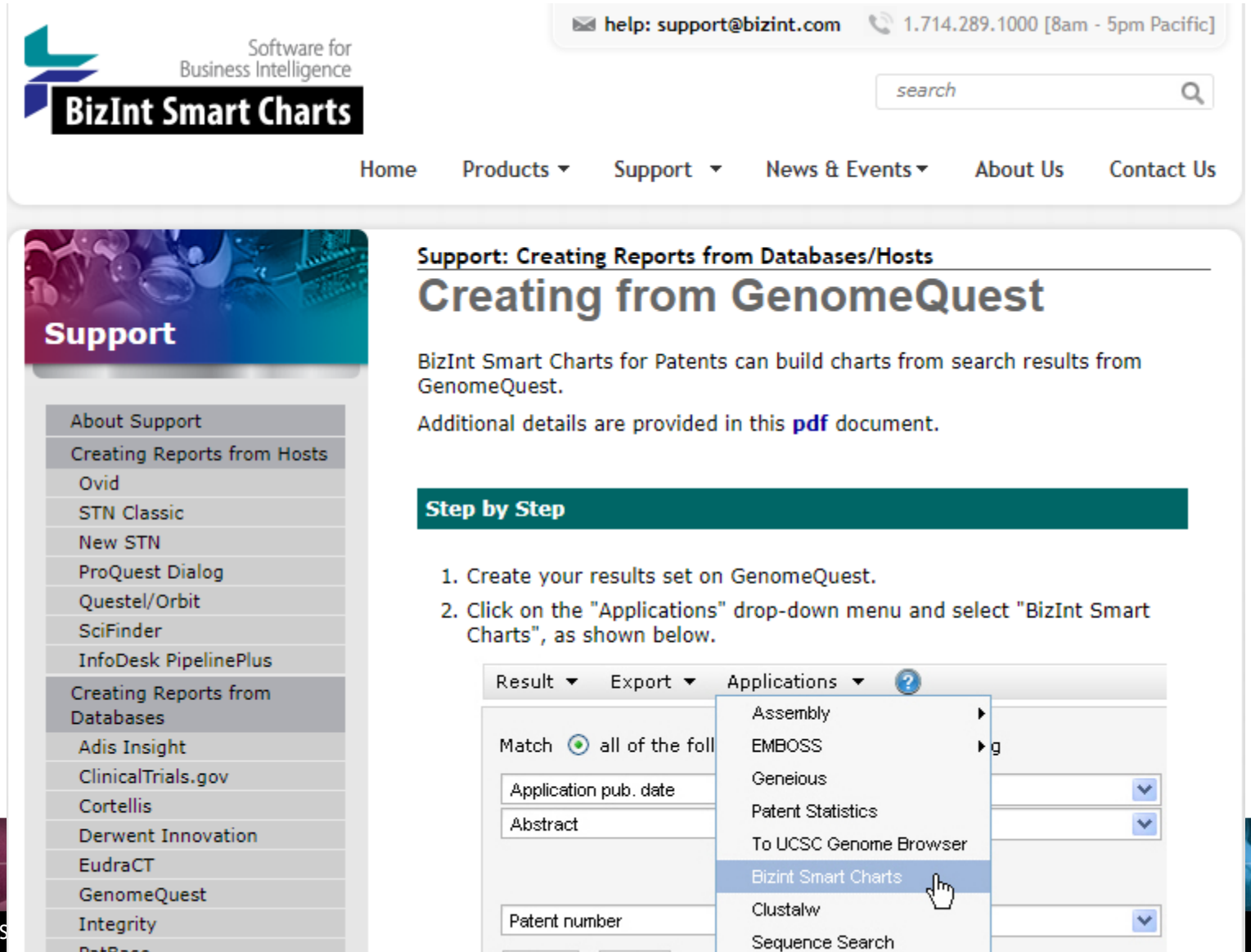
Sequence Databases on GenomeQuest

- CAS Biosequences
- Watch this space!

Export your search results



Export from GenomeQuest



GenomeQuest

- Added Unique Family Sequence ID
- Possible development: each unique sequence only appears once per family
- Added Sequence Listing Equivalents
- Useful in publication-level reports

Features of IP Sequences in BizInt reports

- Each row corresponds to a sequence in the context of a query.
- Columns contain bibliographic data, sequence details, and query results.
- You may see the same sequence more than once in a report.
- Family equivalents are not removed.

Sample sequence data report

GQPAT Proteins: Antibodies_GenomeQuest

	Title	Patent Assignee	Seq. ID Number	Organism Species	Alignment	Percenta
1	**Nucleic acid sequences relating to Bacteroides fragilis for diagnostics and therapeutics**	OSCIENT PHARMACEUTICALS CORPORATION WALTHAM, MA	US7090973-6862	Bacteroides fragilis	Q: 1 KV--SNR-LY 7 || ||| || S: 340 KVDMSNRILY 349	70.00
2	**Expression of microbial proteins in plants for production of plants with improved properties**	MONSANTO TECHNOLOGY, LLC ST. LOUIS, MO	US7314974-14121	Pseudomonas fluorescens	Q: 1 K-VS--NRLY 7 | || |||| S: 597 KLVSDLNRLY 606	70.00
3	**Production and use of novel peptide-based agents for use with bi-specific antibodies**	IMMUNOMEDICS INC. MORRIS PLAINS, NJ	US6962702-0008	Artificial Sequence	Q: 1 RSSQSIVHSNGNTYLQ 16 |||||||||||||||+ S: 24 RSSQSIVHSNGNTYLE 39	93.75
4	**Chimeric, human and humanized anti-CSAP monoclonal antibodies**	IMMUNIMEDICS, INC. MORRIS PLAINS, NJ	US7387772-0032	Murine sp.	Q: 1 RSSQSIVHSNGNTYLQ 16 |||||||||||||||+ S: 24 RSSQSIVHSNGNTYLE 39	93.75
5	**Chimeric, human and humanized anti-CSAp monoclonal antibodies**	IMMUNOMEDICS, INC. MORRIS PLAINS, NJ	US7414121-0032	Murine sp.	Q: 1 RSSQSIVHSNGNTYLQ 16 |||||||||||||||+ S: 24 RSSQSIVHSNGNTYLE 39	93.75
6	**Production and use of novel peptide-based agents for use with bi-specific antibodies**	IMMUNOMEDICS, INC. MORRIS PLAINS, NJ	US7429381-0008	Artificial Sequence	Q: 1 RSSQSIVHSNGNTYLQ 16 |||||||||||||||+ S: 24 RSSQSIVHSNGNTYLE 39	93.75
7	**Production and use of novel peptide-based agents for use with bi-specific antibodies**	IMMUNOMEDICS INC. MORRIS PLAINS, NJ	US6962702-0012	Artificial Sequence	Q: 1 RSSQSIVHSNGNTYLQ 16 |||||||||||||||+ S: 24 RSSQSIVHSNGNTYLE 39	93.75
8	**Production and use of novel peptide-based agents for use with bi-specific antibodies**	IMMUNOMEDICS, INC. MORRIS PLAINS, NJ	US7429381-0012	Artificial Sequence	Q: 1 RSSQSIVHSNGNTYLQ 16 |||||||||||||||+ S: 24 RSSQSIVHSNGNTYLE 39	93.75
9	**Covalently reactive transition state analogs and methods of use thereof**	BOARD OF REGENTS, THE UNIVERSITY OF TEXAS SYSTEM AUSTIN, TX	US6855804-0042	Mus musculus domesticus	Q: 1 RSSQSIVHSNGNTYLQ 16 |||||||||||||||+ S: 24 RSSQSIVHSNGNTYLE 39	93.75
	Covalently reactive transition	ADLER; BENJAMIN	US7524663-0042	Mus musculus	Q: 1 RSSQSIVHSNGNTYLQ 16	93.75

Sample sequence data report - features

GQPAT Proteins: Antibodies_GenomeQuest

	Title	Patent Assignee	Organism Species	Alignment	Percenta
1	Nucleic acid sequences relating to Bacteroides fragilis for diagnostics and therapeutics	OSCIENT PHARMACEUTICALS CORPORATION WALTHAM, MA	Bacteroides fragilis	Q: 1 KV--SNR-LY 7 || ||| || S: 340 KVDMSNRILY 349	70.00
2	Expression of microbial proteins in plants for production of plants with improved properties	MONSANTO TECHNOLOGY, LLC ST. LOUIS, MO	1 Pseudomonas fluorescens	Q: 1 K-VS--NRLY 7 | || |||| S: 597 KLVSDLNRLY 606	70.00
3			Artificial Sequence	Q: 1 RSSQSIVHSNGNTYLQ 16 |||||||||||||||+ S: 24 RSSQSIVHSNGNTYLE 39	93.75
4	Production and use of novel peptide-based agents for use with bi-specific antibodies	IMMUNOMEDICS INC. MORRIS PLAINS, NJ	Murine sp.	Q: 1 RSSQSIVHSNGNTYLQ 16 |||||||||||||||+ S: 24 RSSQSIVHSNGNTYLE 39	93.75
5	Chimeric, human and hu… mo…	IMMUNIMED…CS, INC. …AINS, NJ	Murine sp.	Q: 1 RSSQSIVHSNGNTYLQ 16 |||||||||||||||+ S: 24 RSSQSIVHSNGNTYLE 39	93.75
6	Chimeric, human and humanized anti-CSAp monoclonal antibodies	IMMUNOMEDICS, INC. MORRIS PLAINS, NJ	Artificial Sequence	Q: 1 RSSQSIVHSNGNTYLQ 16 |||||||||||||||+ S: 24 RSSQSIVHSNGNTYLE 39	93.75
7			Artificial Sequence	Q: 1 RSSQSIVHSNGNTYLQ 16 |||||||||||||||+ S: 24 RSSQSIVHSNGNTYLE 39	93.75
8	Production and use of novel peptide-based agents for use with bi-specific antibodies	IMMUNOMEDICS, INC. MORRIS PLAINS, NJ	Artificial Sequence	Q: 1 RSSQSIVHSNGNTYLQ 16 |||||||||||||||+ S: 24 RSSQSIVHSNGNTYLE 39	93.75
9	Production and use of novel peptide-based agents for use with bi-specific antibodies	IMMUNOMEDICS INC. MORRIS PLAINS, NJ	Mus musculus domesticus	Q: 1 RSSQSIVHSNGNTYLQ 16 |||||||||||||||+ S: 24 RSSQSIVHSNGNTYLE 39	93.75
			Mus musculus	Q: 1 RSSQSIVHSNGNTYLQ 16	93.75

Bibliographic Data

Sample sequence data report - features

GQPAT Proteins: Antibodies_GenomeQuest

	Title	Patent Assignee	Seq. ID Number	Organism Species	Alignment	Percenta
1	Nucleic acid sequences relating to Bacteroides fragilis for diagnostics and therapeutics	OSCIENT PHARMACEUTICALS CORPORATION WALTHAM, MA	US7090973-6862	Bacteroides fragilis	1 KV--SNR-LY 7 || ||| || 340 KVDMSNRILY 349	70.00
2	Expression of microbial proteins in plants for production of plants with improved properties	MONSANTO TECHNOLOGY, LLC ST. LOUIS, MO	US7314974-14121	Pseudomonas fluorescens	1 K-VS--NRLY 7 | || |||| 597 KLVSDLNRLY 606	70.00
3	Production and use of novel peptide-based agents for use with bi-specific antibodies	IMMUNOMEDICS IN MORRIS PLAINS, N.	US6962702-0008	Artificial Sequence	1 RSSQSIVHSNGNTYLQ 16 |||||||||||||||+ 24 RSSQSIVHSNGNTYLE 39	93.75
4	Chimeric, human and humanized anti-CSAP monoclonal antibodies	IMMUNIMEDICS, INC MORRIS PLAINS, N.				
5	Chimeric, human and humanized anti-CSAp monoclonal antibodies	IMMUNOMEDICS, IN MORRIS PLAINS, N.	US7387772-0032	Murine sp.		
6	Production and use of novel peptide-based agents for use with bi-specific antibodies	IMMUNOMEDI MORRIS PLAI	US7414121-0032	Murine sp.		
7	Production and use of novel peptide-based agents for use with bi-specific antibodies	IMMUNOMEDICS IN MORRIS PLAINS, N.				
8	Production and use of novel peptide-based agents for use with bi-specific antibodies	IMMUNOMEDICS, IN MORRIS PLAINS, N.	US7429381-0008	Artificial Sequence		
9	Covalently reactive transition state analogs and methods of use thereof	BOARD OF REGENT THE UNIVERSITY O TEXAS SYSTEM AUSTIN, TX	US6962702-0012	Artificial Sequence		
	Covalently reactive transition	ADLER; BENJAMIN				

Sequence Data

Sample sequence data report - features

GQPAT Proteins: Antibodies_GenomeQuest

	Title	Patent Assignee	Seq. ID Num	Alignment	Percentage
1	Nucleic acid sequences relating to Bacteroides fragilis for diagnostics and therapeutics	OSCIENT PHARMACEUTICALS CORPORATION WALTHAM, MA	US7090973-6	Q: 1 KV--SNR-LY 7 S: 340 KVDMSNRILY 349	70.00
2	Expression of microbial proteins in plants for production of plants with improved properties	MONSANTO TECHNOLOGY, LLC ST. LOUIS, MO	US7314974-1	Q: 1 K-VS--NRLY 7 S: 597 KLVSDLNRLY 606	70.00
3	Production and use of novel peptide-based agents for use with bi-specific antibodies	IMMUNOMEDICS INC. MORRIS PLAINS, NJ	US6962702-0	Q: 1 RSSQSIVHSNGNTYLQ 16 S: 24 RSSQSIVHSNGNTYLE 39	93.75
4	Chimeric, human and humanized anti-CSAP monoclonal antibodies	IMMUNIMEDICS, INC. MORRIS PLAINS, NJ	US7387772-0		
5	Chimeric, human and humanized anti-CSAp monoclonal antibodies	IMMUNOMEDICS, INC. MORRIS PLAINS, NJ	US7414121-0	Q: 1 RSSQSIVHSNGNTYLQ 16 S: 24 RSS	93.75
6	Production and use of novel peptide-based agents for use with bi-specific antibodies	IMMUNOMEDICS, INC. MORRIS PLAINS, NJ	US7429381-0		
7	Production and use of novel peptide-based agents for use with bi-specific antibodies	IMMUNOMEDICS INC. MORRIS PLAINS, NJ	US6962702-0	Q: 1 RSSQSIVHSNGNTYLQ 16 S: 24 RSSQSIVHSNGNTYLE 39	93.75
8	Production and use of novel peptide-based agents for use with bi-specific antibodies	IMMUNOMEDICS, INC. MORRIS PLAINS, NJ	US7429381-0	Q: 1 RSSQSIVHSNGNTYLQ 16 S: 24 RSSQSIVHSNGNTYLE 39	93.75
9	Covalently reactive transition state analogs and methods of use thereof	BOARD OF REGENTS, THE UNIVERSITY OF TEXAS SYSTEM AUSTIN, TX	US6855804-0	Q: 1 RSSQSIVHSNGNTYLQ 16 S: 24 RSSQSIVHSNGNTYLE 39	93.75
	Covalently reactive transition	ADLER; BENJAMIN	US7524663-0		

Query Results

About the Alignment Column

- The Alignment column appears in the same default font as all columns
- You might be tempted to change the font to Courier New – DON'T!
- **Text menu | Fixed Width (e.g. Alignment)** is the correct technique
- Preserves runs of whitespace when exporting to HTML, Word, or Excel

Features of IP Sequences in BizInt reports

	Title	Alignment	Alignment Style
1	HEPATITIS C VIRUS GENE	Q: 1 MSTNPKPQRKTKRNTNRRPQDVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRG 60 S: 1 MSTNPKPQRKTKRNTNRRPQDVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRG 60 Q: 61 RRQPIPKDRRSTGKSWGKPGYPWPLYGNEGCGWAGWLLSPRGSRPTWGPTDPRHRSRNLG 120 [CONT.]	Default
2	HCV GENE	Q: 1 MSTNPKPQRKTKRNTNRRPQDVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRG 60 S: 1 MSTNPKPQRKTKRNTNRRPQDVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRE 60 Q: 61 RRQPIPKDRRSTGKSWGKPGYPWPLYGNEGCGWAGWLLSPRGSRPTWGPTDPRHRSRNLG 120 [CONT.]	Changed Font
3	HCV GENE	Q: 1 MSTNPKPQRKTKRNTNRRPQDVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRG 60 S: 1 MSTNPKPQRKTKRNTNRRPQDVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRE 60 Q: 61 RRQPIPKDRRSTGKSWGKPGYPWPLYGNEGCGWAGWLLSPRGSRPTWGPTDPRHRSRNLG 120 [CONT.]	Text | Fixed Width

Features of IP Sequences in BizInt reports

GQPAT Gold+ Proteins: hepc_claimed_gqprt

	Title	Alignment	Alignment Style
1 Link	**HEPATITIS C VIRUS GENE**	Q: 1 MSTNPKPQRKTKRNTNRRPQDVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRG 60 |||||||||||||||||||||||||||||||||||||||||||||||||||||||| S: 1 MSTNPKPQRKTKRNTNRRPQDVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRG 60 Q: 61 RRQPIPKDRRSTGKSWGKPGYPWPLYGNEGCGWAGWLLSPRGSRPTWGPTDPRHRSRNLG 120 ||||||| |+ |++| +|||||||||||| |||||||||||||+||||||| |||||| [CONT.]	Default
2 Link	**HCV GENE**	Q: 1 MSTNPKPQRKTKRNTNRRPQDVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRG 60 |||||||||||||||||||||||||||||||||||||||||||||||||||||||| S: 1 MSTNPKPQRKTKRNTNRRPQDVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRE 60 Q: 61 RRQPIPKDRRSTGKSWGKPGYPWPLYGNEGCGWAGWLLSPRGSRPTWGPTDPRHRSRNLG 120 ||||||| || |++| +|||||||||||| |||||||||||||| ||||||| |||||| [CONT.]	Changed Font
3 Link	**HCV GENE**	Q: 1 MSTNPKPQRKTKRNTNRRPQDVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRG 60 |||||||||||||||||||||||||||||||||||||||||||||||||||||||| S: 1 MSTNPKPQRKTKRNTNRRPQDVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRE 60 Q: 61 RRQPIPKDRRSTGKSWGKPGYPWPLYGNEGCGWAGWLLSPRGSRPTWGPTDPRHRSRNLG 120 ||||||| || |++| +||||||||||||| ||||||||||||||| ||||||| |||||| [CONT.]	Text | Fixed Width

Non-Patent Sequence Databases

- Can support as needed
- e.g. SWISSPROT, DrugBank

- Can be combined with IP sequences
- Cannot be grouped with patent families
- No tools for deduplication

Sequences in Literature

- Most collections are structed around sequences with multiple citations per seq
- CAplus is organized by document, with multiple sequences per doc

Improved Literature Support -- Project Goals

- Bring literature support up to the same level as drug pipelines, clinical trials, patents, and IP sequences
- Collecting user priorities to help set our development plans
- Focus of development for the next year.

Improved literature support - summary

- “Identify Common Citation” to allow Reference Rows operations and deduplication
- Literature support in all products
- Combine literature with patents OR trials
- Options for Citation formats, exports
- Support for additional sources
- Additional fields for analysis
- Improved hyperlink generation

Agenda

- IP Sequence reports
- Hit Structure reports
- Summary Record exports
- Reference Rows

Deliver reports in HTML, Word, Excel…

LifeQuest: Sample

	Title	Patent Family: Patent	Kind	Date	Inventor(s)	CPC	Abstract	Claims	Pub. Status	
1 Link	Drug delivery compositions	US7329638	B2	2008-02-12	Victor C. Yang Yoon Jeong Park Junfeng Liang	A61K2039/505 A61K47/48315 A61K47/48884 A61K47/4823 A61K47/48561	The present invention relates to multicomponent compositions and methods of administering these compositions, which specifically translocate	1. A composition, comprising: a. a first targeting component, wherein said first targeting component comprises i. a molecular recognition element, wherein said molecular	Grant	0
2 Link	Drug delivery compos									0
3 Link	POLYMORPHISMS OF REGION OF THE HUMA HT1A GENE, ASSOCIA PROTEINS OF THE 5' AND A DIAGNOSTIC T FOR MAJOR DEPRESSI RELATED MENTAL ILL									0
4 Link	Oncology drug innova									0

Link in the

Link to web resource (like USPTO)

Link to record on source platform

United States Patent: 732

patft.uspto.

ml&r=1&f=G&l=50&d=PALL&RefSrch=ye...

LL-TEXT AND IMAGE DATABASE

Advanced | Pat Num | Help

Bottom

View Cart | Add to Cart

Images

(1 of 1)

United States Patent **7,329,638**
Yang , et al. **February 12, 2008**

Drug delivery compositions

Abstract

The present invention relates to multicomponent compositions and methods of administering these compositions, which specifically translocate therapeutic molecules (e.g., drugs or prodrugs) across biological membranes thus reducing potential toxic side effects on nontargeted cells and tissues.

Inventors: **Yang; Victor C.** (Ann Arbor, MI), **Park; Yoon Jeong** (Seoul, **KR**), **Liang; Junfeng** (Westfield, NJ)
Assignee: **The Regents of the University of Michigan** (Ann Arbor, MI)
Family ID: **34198915**
Appl. No.: **10/835,151**
Filed: **April 29, 2004**

Prior Publication Data

Document Identifier	Publication Date
US 20050042753 A1	Feb 24, 2005

Related U.S. Patent Documents

Summary Record export in Word

3.	**Basic Patent Number:**	**WO2012033858A2**
	Title:	Boron-containing small molecules
	Inventor(s):	Hernandez, Vincent S.; Ding, Charles; Plattner, Jacob J.; Alley, Michael Richard Kevin; Rock, Fernando; Zhang, Suoming; Easom, Eric; Li, Xianfeng; Zhou, Ding
	Patent Assignee:	Anacor Pharmaceuticals, Inc., USA
	Hyperlinks:	Source | WO2012033858A2 |

- Fields are the columns in chart
- Content, like hyperlinks, is included in the Summary Record

THE JOURNEY BEGINS...

Summary Record export with Hit Structures

3. **Basic Patent Number:** WO2012033858A2

Title: Boron-containing small molecules

Inventor(s): Hernandez, Vincent S.; Ding, Charles; Plattner, Jacob J.; Alley, Michael Richard Kevin; Rock, Fernando; Zhang, Suoming; Easom, Eric; Li, Xianfeng; Zhou, Ding

Patent Assignee: Anacor Pharmaceuticals, Inc., USA

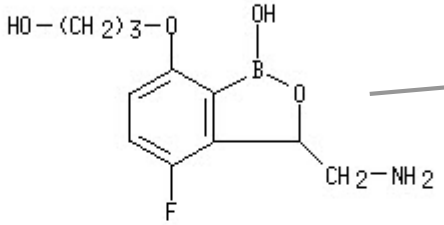
Hyperlinks: Source | WO2012033858A2

Hit Structures:

1364682-96-1 *(Cmpd. 2)*

1-Propanol, 3-[[3-(aminomethyl)-4-fluoro-1,3-dihydro-1-hydroxy-2,1-benzoxaborol-7-yl]oxy]-, 2,2,2-trifluoroacetate (1:2) (CA INDEX NAME)

CM 1 CRN 1364682-95-0



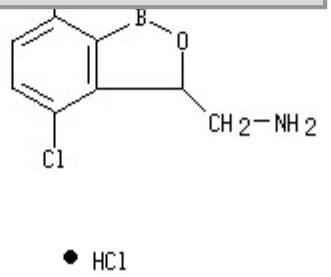
CM 2 CRN 76-05-1

RL: PAC (Pharmacological activity); SPN (Synthetic preparation); TH (Therapeutic us (Biological stud (Preparation); U (Uses)

prepn. of benzoxaborole useful for treati bacterial infections

1364683-03-3 *(Cmpd. 3)*

1-Propanol, 3-[[3-(aminomethyl)-4-chloro-1,3-dihydro-1-hydroxy-2,1-benzoxaborol-7-yl]oxy]-, hydrochloride (1:1) (CA INDEX NAME)



RL: PAC (Pharmacological activity); SPN (Synthetic preparation); THU (Therapeutic use); BIOL (Biological stud (Preparation); U (Uses)

prepn. of benzoxaborole derivs. useful for treating bacterial infections

Text from CAplus record

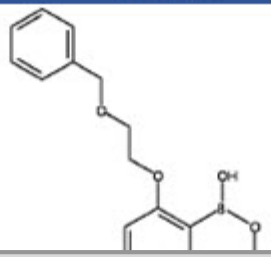
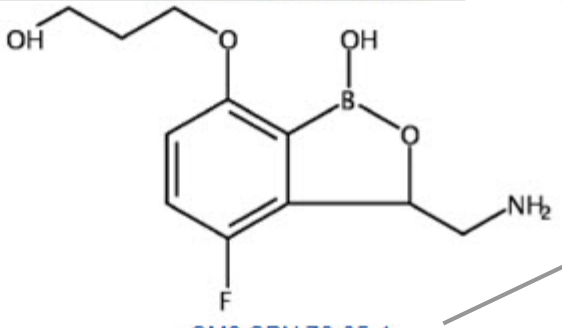
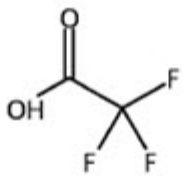
Structures with annotations including multi page images

Each hit substance identified

Index Terms

A structure oriented “Index of Hit Structures”

Index of Hit Structures

	Substance	Structure	
1	1655492-02-6 2,1-Benzoxaborole, 4-fluoro-1,3-dihydro-1-hydroxy-3-(nitromethyl)-7-[2-(phenylmethoxy)ethoxy]-		pr… 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)	 CM2 CRN 76-05-1 	prepn. of benzoxaborole derivs. useful for treating bacterial infections *Reference 3*

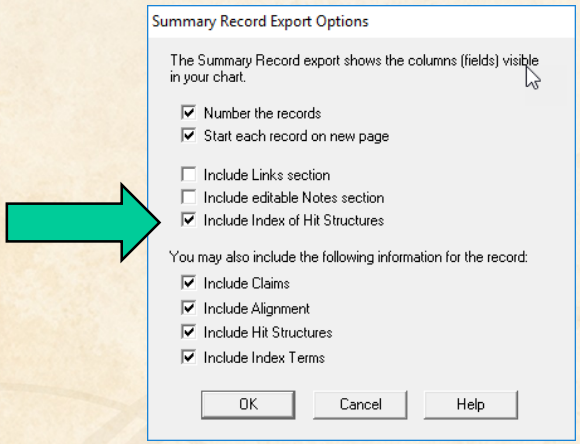
Structures with annotations

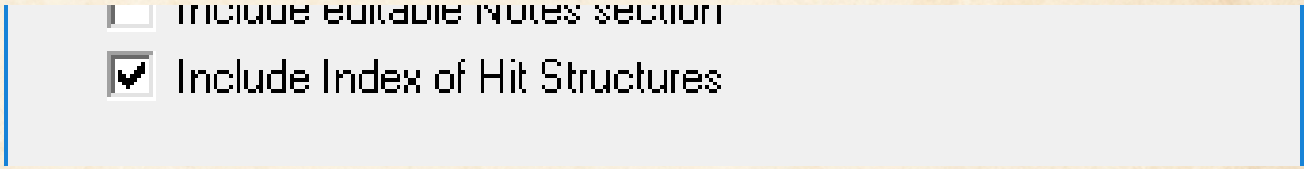
Each hit substance identified

Links to references for each structure

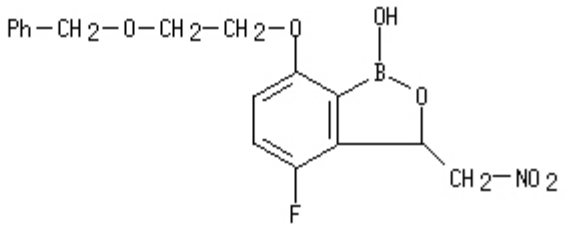
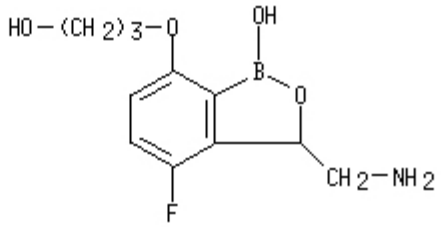
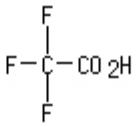
Multiple images (continuation) or mixtures

Option: Index of Hit Structures

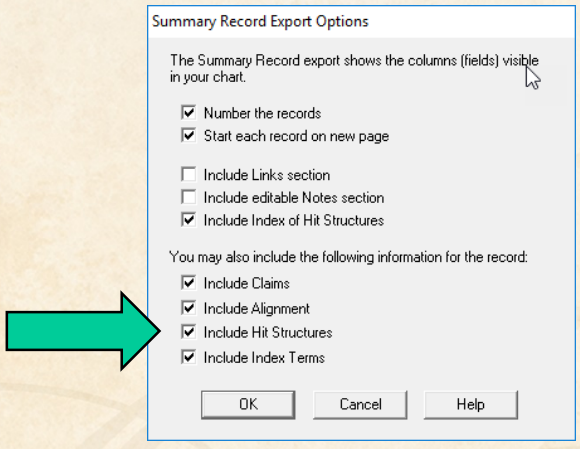


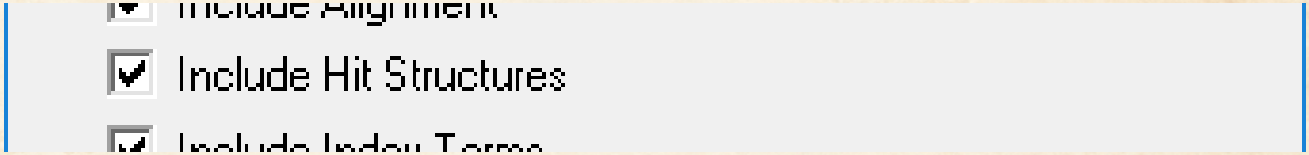


Index of Hit Structures

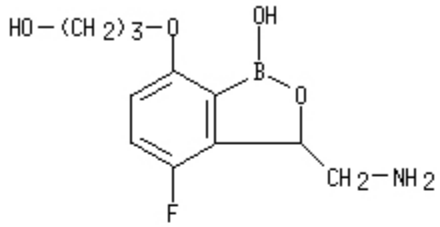
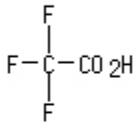
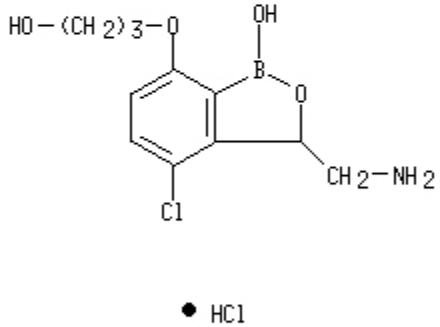
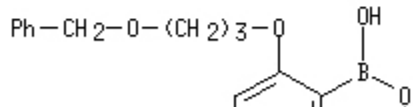
	Substance	Structure	Reference
1	1655492-02-6 2,1-Benzoxaborole, 4-fluoro-1,3-dihydro-1-hydroxy-3-(nitromethyl)-7-[2- (phenylmethoxy)ethoxy]- (CA INDEX NAME)		prepn. and anti-mycobacterial activity of benzoxaborole compds. Reference 1 prepn. and biol. applications of tricyclic benzoxaborole compds. Reference 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) (CA INDEX NAME)	CM 1 CRN 1364682-95-0  CM 2 CRN 76-05-1 	prepn. of benzoxaborole derivs. useful for treating bacterial infections Reference 3

Option: Hit Structures

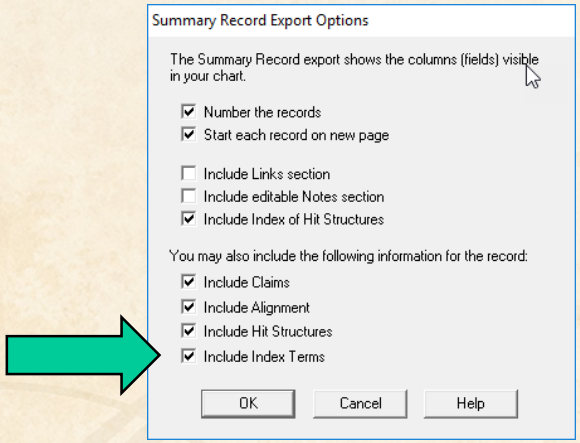


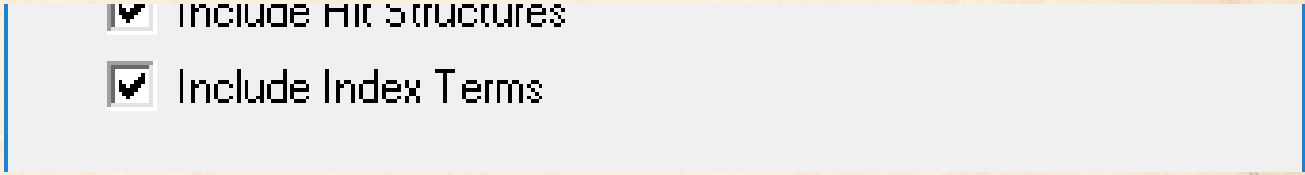


Hit Structures:

1364682-96-1 *(Cmpd. 2)* 1-Propanol, 3-[[3-(aminomethyl)-4-fluoro-1,3-dihydro-1-hydroxy-2,1-benzoxaborol-7-yl]oxy]-, 2,2,2-trifluoroacetate (1:2) (CA INDEX NAME)	CM 1 CRN 1364682-95-0  CM 2 CRN 76-05-1 	RL: PAC (Pharmacological activity); SPN (Synthetic preparation); THU (Therapeutic use); BIOL (Biological study); PREP (Preparation); USES (Uses) prepn. of benzoxaborole derivs. useful for treating bacterial infections
1364683-03-3 *(Cmpd. 3)* 1-Propanol, 3-[[3-(aminomethyl)-4-chloro-1,3-dihydro-1-hydroxy-2,1-benzoxaborol-7-yl]oxy]-, hydrochloride (1:1) (CA INDEX NAME)		RL: PAC (Pharmacological activity); SPN (Synthetic preparation); THU (Therapeutic use); BIOL (Biological study); PREP (Preparation); USES (Uses) prepn. of benzoxaborole derivs. useful for treating bacterial infections
1364684-69-4 *(Cmpd. 4)* 2,1-Benzoxaborole, 4-fluoro-1,3-dihydro-1-hydroxy-3-		RL: RCT (Reactant); SPN (Synthetic preparation); PREP (Preparation); RACT

Option: Index Terms





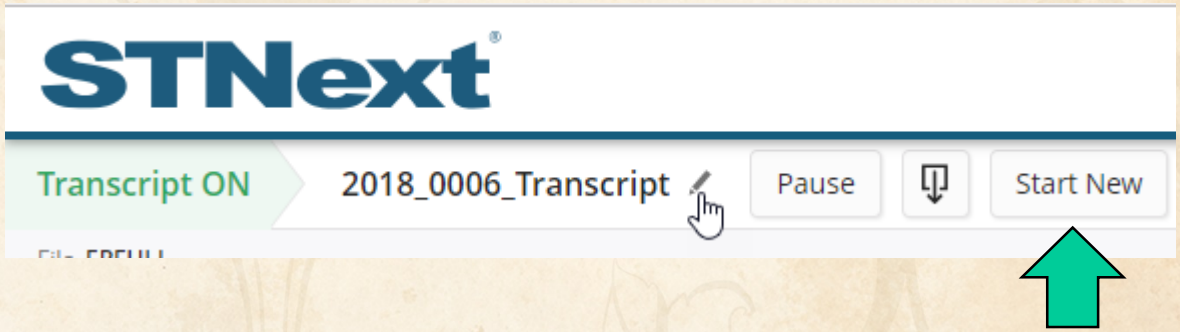
Index Terms:

1364682-96-1P (Cmpd. 2) 1364683-03-3P (Cmpd. 3) PAC (Pharmacological activity); SPN (Synthetic preparation); THU (Therapeutic use); BIOL (Biological study); PREP (Preparation); USES (Uses) (prepn. of benzoxaborole derivs. useful for treating bacterial infections)

1364684-69-4P (Cmpd. 4) 1364684-75-2P (Cmpd. 5) RCT (Reactant); SPN (Synthetic preparation); PREP (Preparation); RACT (Reactant or reagent) (prepn. of benzoxaborole derivs. useful for treating bacterial infections)

Importing transcripts with hit structures (STNext)

- Make sure that Classic Display is on
- Start new transcript before displaying.



BIB vs. IBIB

- We recommend using tagged (BIB), rather than indented (IBIB), display formats
- Some field contents (table headings) appear before the label in IBIB
- Indent levels in RTF are more reliably detected in BIB

Agenda

- IP Sequence reports
- Hit Structure reports
- Summary Record exports
- Reference Rows

Summary Record export in Word

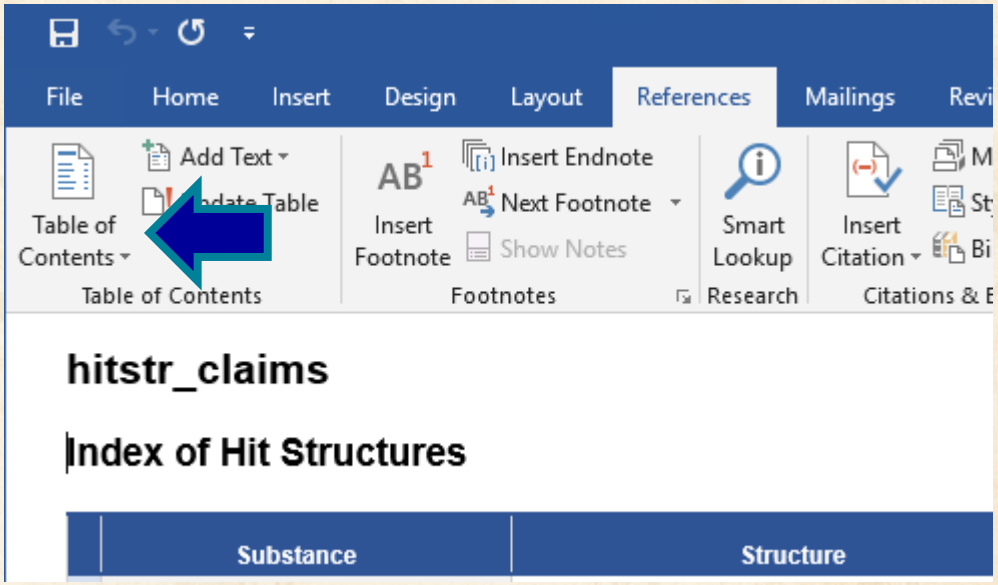
3.	**Basic Patent Number:**	**WO2012033858A2**
	Title:	Boron-containing small molecules
	Inventor(s):	Hernandez, Vincent S.; Ding, Charles; Plattner, Jacob J.; Alley, Michael Richard Kevin; Rock, Fernando; Zhang, Suoming; Easom, Eric; Li, Xianfeng; Zhou, Ding
	Patent Assignee:	Anacor Pharmaceuticals, Inc., USA
	Hyperlinks:	Source | WO2012033858A2

- Fields are the columns in chart
- Content, like hyperlinks, is included in the Summary Record

THE JOURNEY BEGINS...

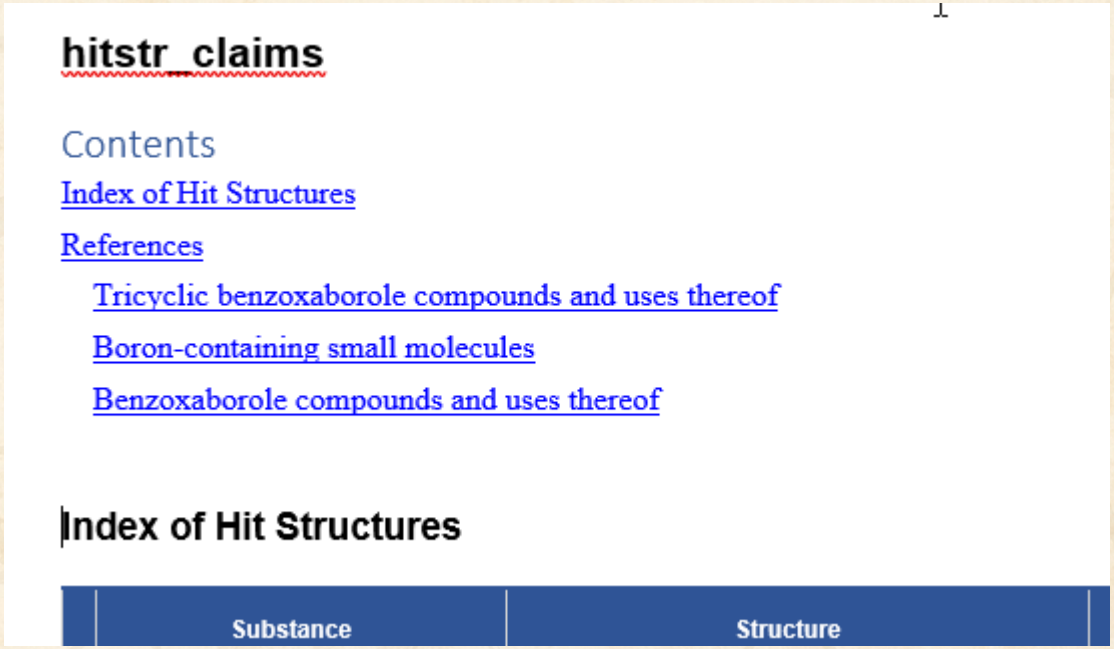
Summary Records – Table of Contents

- Add a table of contents in Word



Summary Records - Table of Contents

- Sections are marked up for headings
- First column in chart is used as TOC entry



hitstr_claims

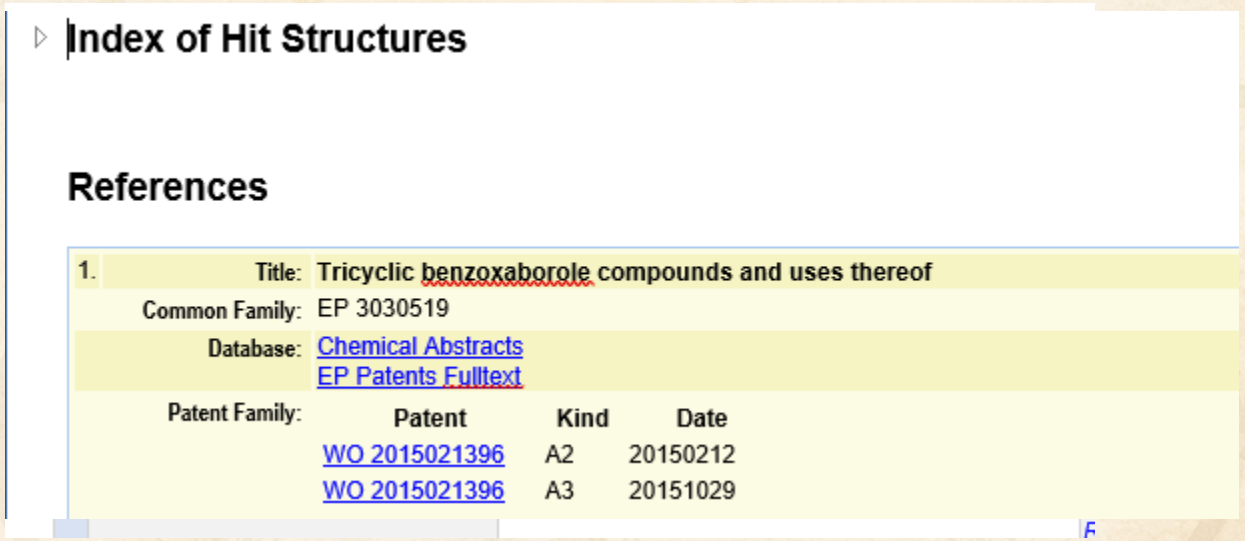
Contents

Index of Hit Structures

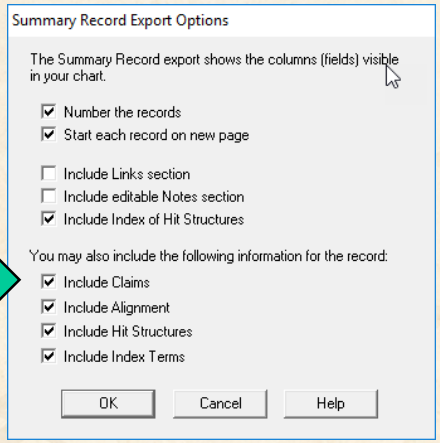
Substance	Structure

Summary Records - Table of Contents

- Can collapse the Index of Hit Structures (in recent versions of Word)



Option: Claims + Hit Structures



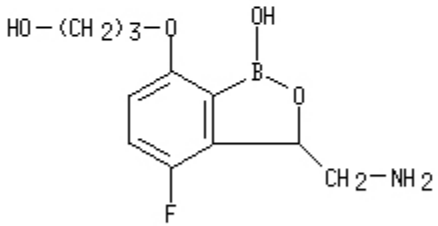
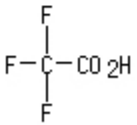
10. An in vitro method of:

(A) inhibiting an enzyme, comprising: contacting the enzyme with the compound of any of claims 1 to 5, thereby inhibiting the enzyme;

(B) killing and/or preventing the growth of a microorganism, comprising: contacting the microorganism with an effective amount of the compound of any of claims 1 to 5, thereby killing and/or preventing the growth of the microorganism; or

(C) inhibiting the editing domain of a t-RNA synthetase, comprising: contacting the synthetase with an effective amount of a compound of any of claims 1 to 5, or a pharmaceutically-acceptable salt thereof, thereby inhibiting the synthetase.

Hit Structures:

1364682-96-1 *(Cmpd. 2)* 1-Propanol, 3-[[3-(aminomethyl)-4-fluoro-1,3-dihydro-1-hydroxy-2,1-benzoxaborol-7-yl]oxy]-, 2,2,2-trifluoroacetate (1:2) (CA INDEX NAME)	CM 1 CRN 1364682-95-0  CM 2 CRN 76-05-1 	RL: PAC (Pharmacological activity); SPN (Synthetic preparation); THU (Therapeutic use); BIOL (Biological study); PREP (Preparation); USES (Uses) prepn. of benzoxaborole derivs. useful for treating bacterial infections
1364683-03-3 *(Cmpd. 3)*		RL: PAC

Link sequence results with hit structures

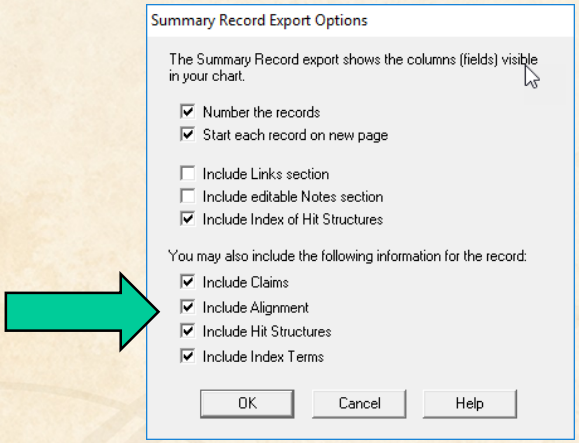
Follow Cookbook recipe to create a summary of sequence hits for each

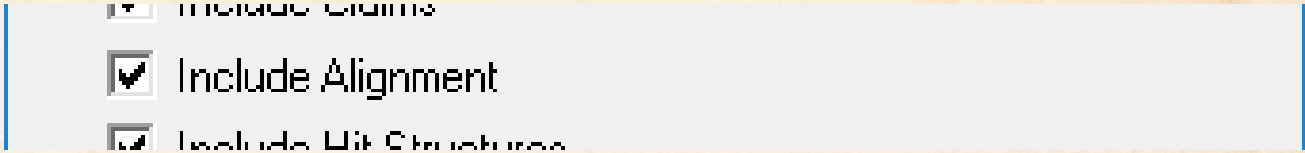
1.	Title:	**Selective high-affinity polydentate ligands and methods of making such**
	Database:	GQPAT Gold+ Proteins GQPAT Gold+ Proteins GQPAT Gold+ Proteins GQPAT Gold+ Proteins GQPAT Gold+ Proteins GQPAT Gold+ Proteins GQPAT Gold+ Proteins Chemical Abstracts Chemical Abstracts
	Organism Species:	Homo sapiens (human)

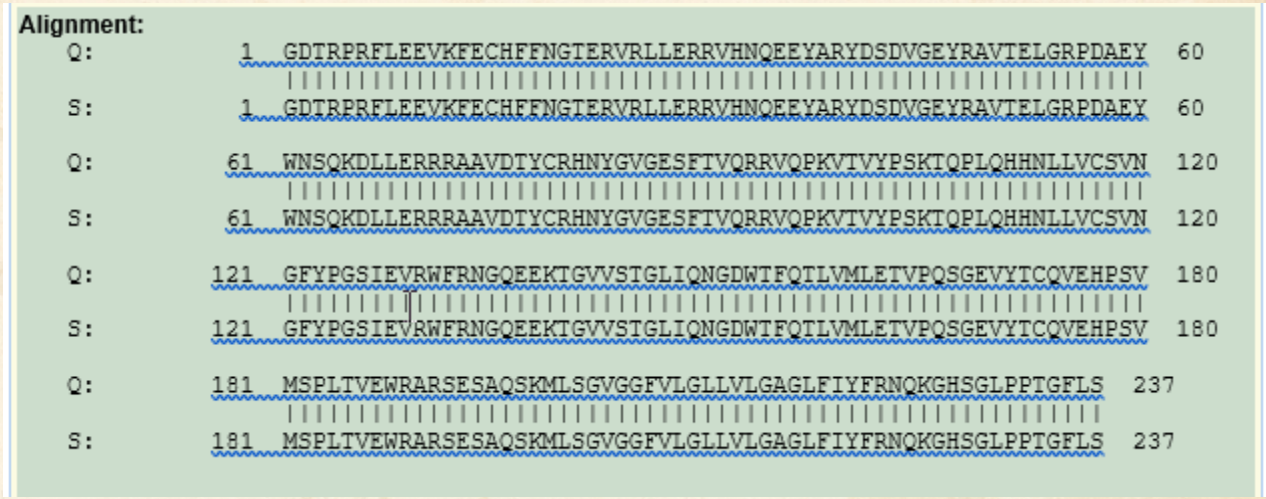
Sequence Summary:

Seq. ID Number	Length	% Identity	Location
US20180008621-0003	237	100.00	probable disclosure (not found by automated parsing)
US9884070-0003	237	100.00	probable disclosure (not found by automated parsing)
US20180008622-0003	237	100.00	probable disclosure (not found by automated parsing)
JP5623384-0003	237	100.00	probable disclosure (not found by automated parsing)
JP2014122234-0003	237	100.00	probable disclosure (not found by automated parsing)
US20110144065-0003	237	100.00	probable disclosure (not found by automated parsing)
CA2721980-0003	237	100.00	probable disclosure (not found by automated parsing)

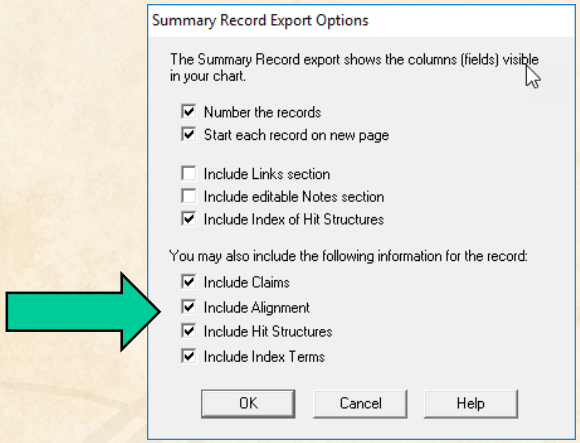
Option: Alignments + Hit Structures

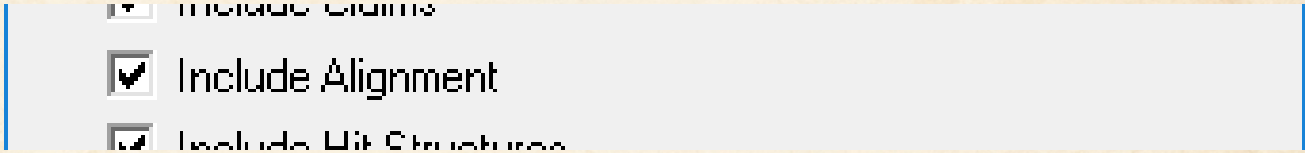






Option: Alignments + Hit Structures



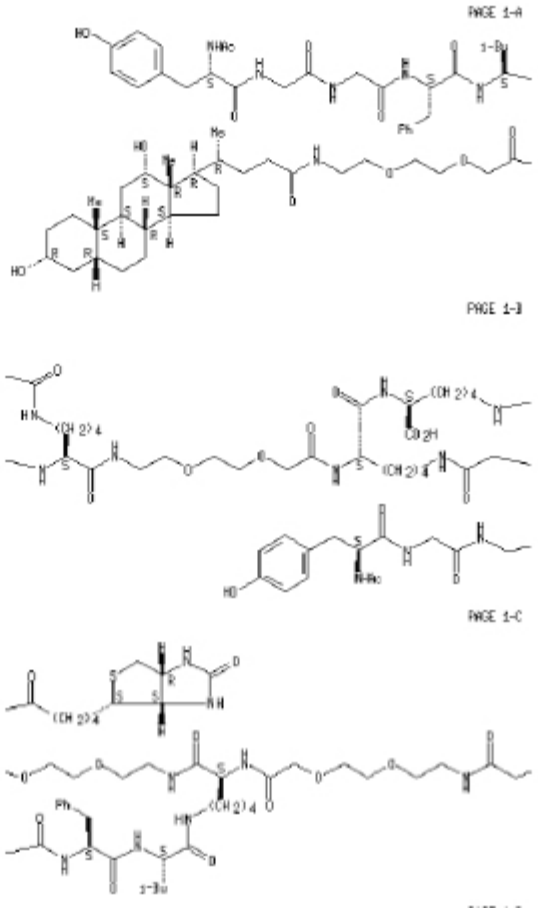


Hit Structures:

1045703-23-8 *(Cmpd. 1)*

L-Lysine, N2,N6-bis[N6-(N-acetyl-L-tyrosylglycylglycyl-L-phenylalanyl-L-leucyl)-N2-[2-[2-[2-[[(3 oxocholan-24-yl]amino]ethoxy]ethoxy]acetyl]-L-lysyl-2-[2-(2-aminoethoxy)ethoxy]acetyl]-L-lysyl-N6-[5-[(3aS,4S,6aR)-hexahydro-2-oxo-1H- thieno[3,4-d]imidazol-4-yl]-1-oxopentyl]- (CA INDEX NAME)

Absolute stereochemistry.



RL: BSU (Biological study, unclassified); DGN (Diagnostic use); PAC (Pharmacological activity); THU (Therapeutic use); BIOL (Biological study); USES (Uses)

HLA-DR10 selective high-affinity polydentate ligand; selective high-affinity polydentate ligands of target mols. and methods of making such and uses for diagnosis and therapeutics in relation to delivery of effectors

Agenda

- IP Sequence reports
- Hit Structure reports
- Summary Record exports
- Reference Rows

BizInt Smart Charts

VERSION 5

for Patents

Patent Databases

Provide data on patents filed worldwide

- STN – Classic, STNext, & New STN
- Questel Orbit.com
- Minesoft PatBase
- Innovation, Cortellis IP, Integrity Patents
- LexisNexis TotalPatent
- GQ LifeSciences LifeQuest

Tools for integrating patent data

- **Combine charts** using File|Combine command
- **Identify related records** using the **"Identify Common Patent Family"** tool.

BizInt Smart Charts
for Patents

Tools for integrating patent data

- **Combine charts** using File|Combine command
- **Identify related records** using the **"Identify Common Patent Family"** tool.
- Use **BizInt Smart Charts Reference Rows** to summarize related records in a single row.

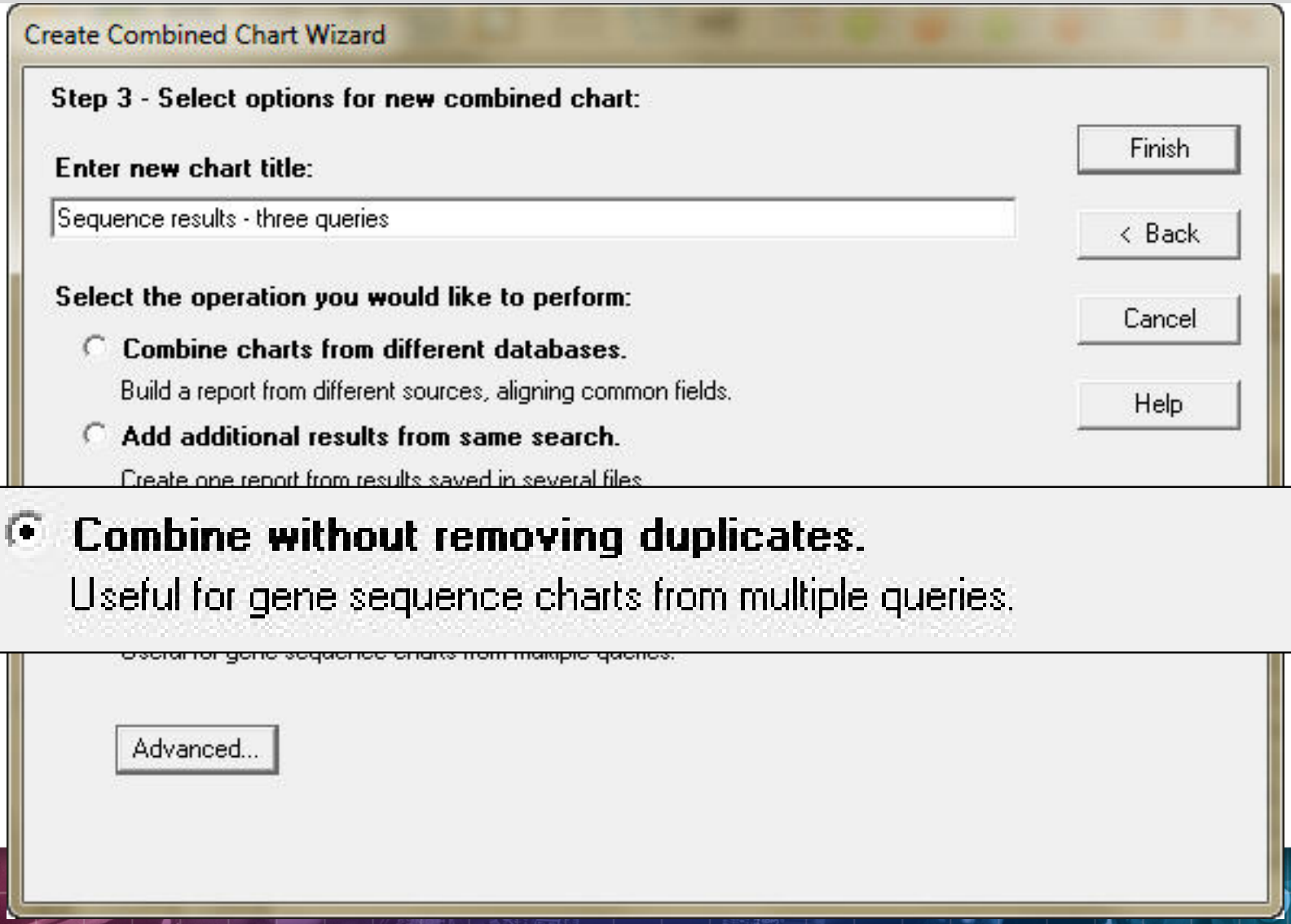
BizInt Smart Charts
for Patents

BizInt Smart Charts
Reference Rows™

Combining Charts

- You can combine two or more charts into a single report
- BizInt Smart Charts usually removes duplicate rows when combining
- Combining different queries? Use the “Combine without removing duplicates” option in the Combine wizard
- See 2013 PIUG Biotech presentation for more details (@ bizint.com/slides)

Combining multiple queries



Preserving Multiple Queries

- First combine non-sequence charts with standard options (remove duplicates)
- Then combine the resulting non-sequence chart with the sequence results (without removing duplicates)
- Only use “without removing duplicates” option when you want to see variations on a row “side by side”

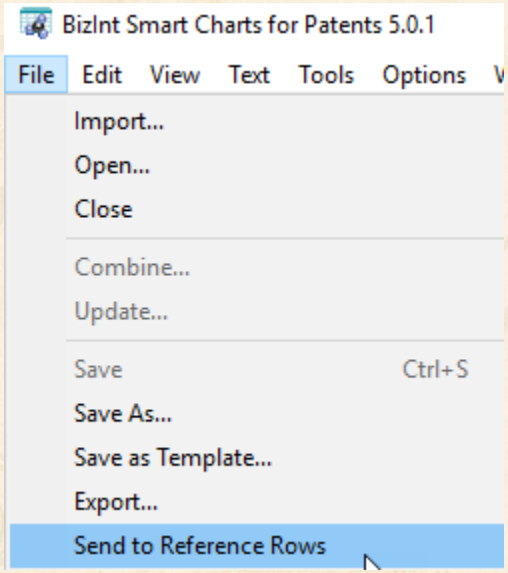
Common Patent Family

"Identify Common Patent Families" tool

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
		WO2014197568	A2	20141211
		WO2014197568	A3	20150312
		CA2914638	A1	20141211
FAMPAT	US 2014356956	US 2014356956	A1	2014-12-04
		US 2014356959	A1	2014-12-04
		US 9267135	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

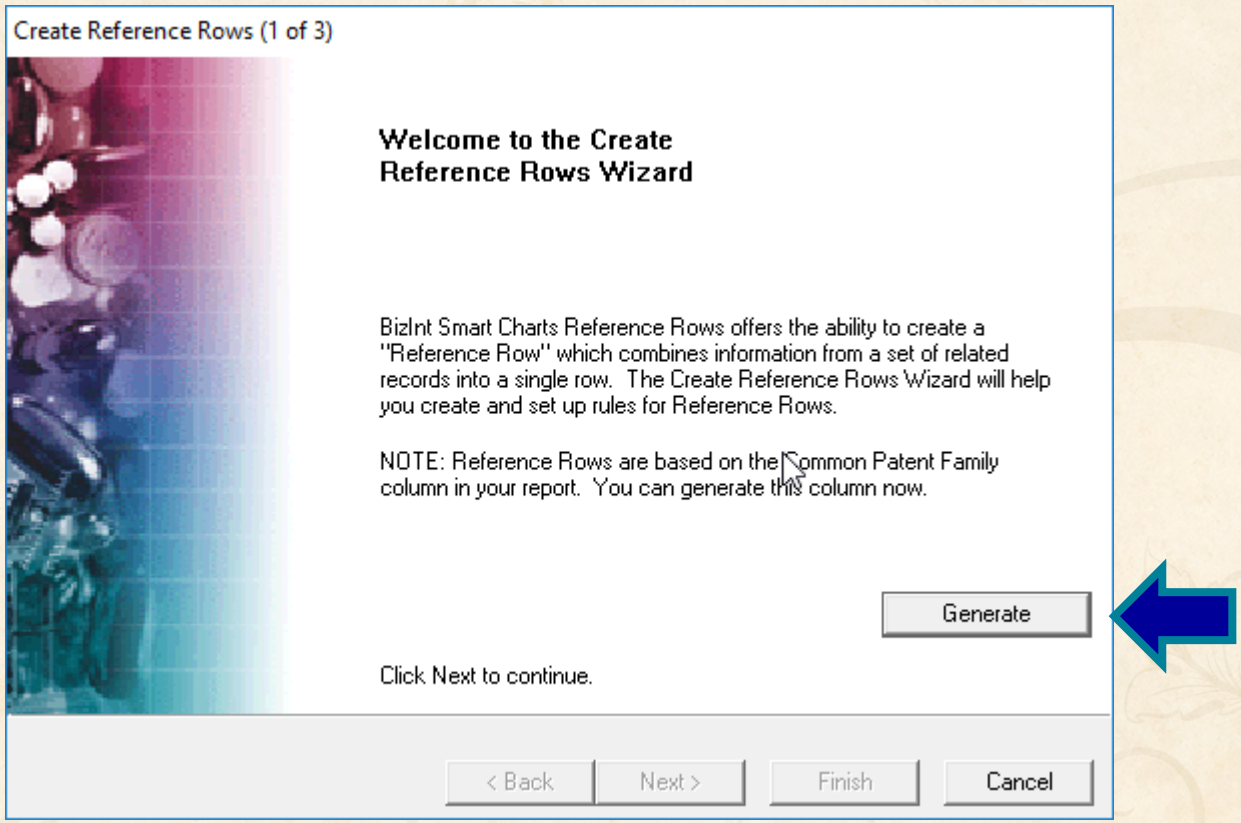
Minimum effort with Reference Rows

- Save chart in BizInt Smart Charts for Patents
- Send to Reference Rows



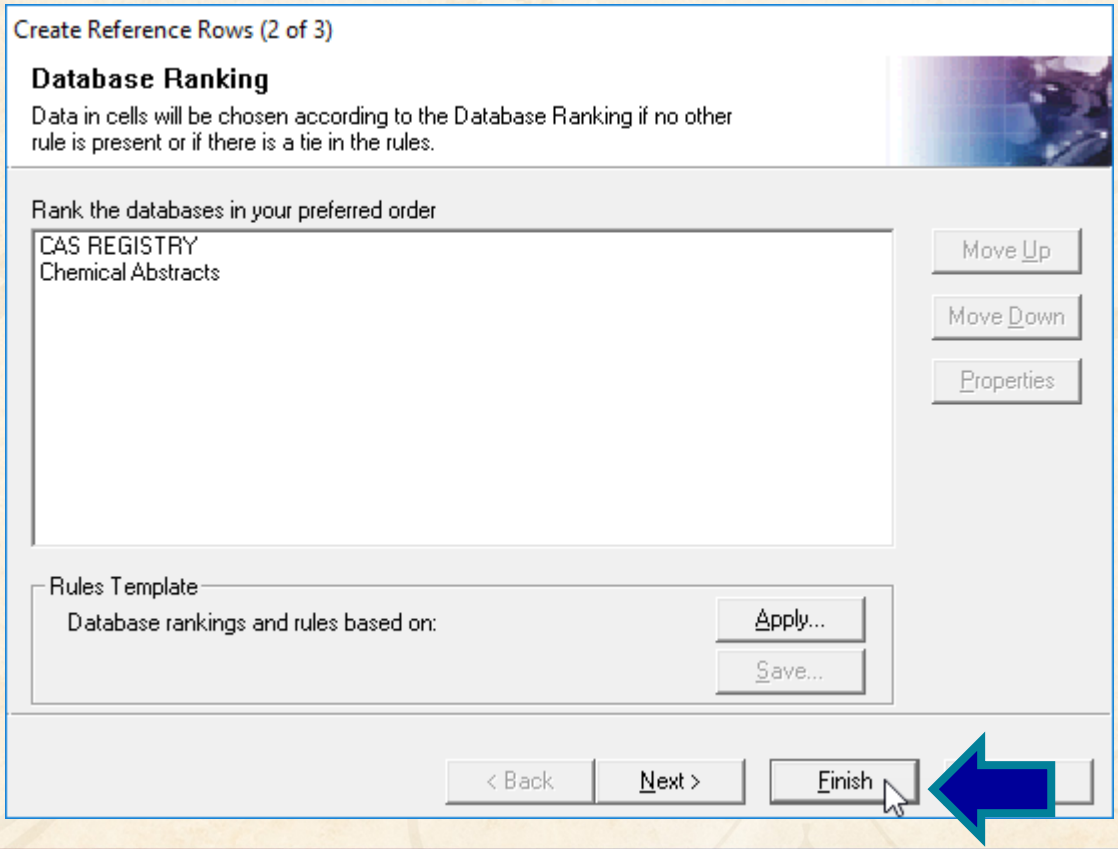
Minimum effort with Reference Rows

- Generate Common Patent Family (if needed)



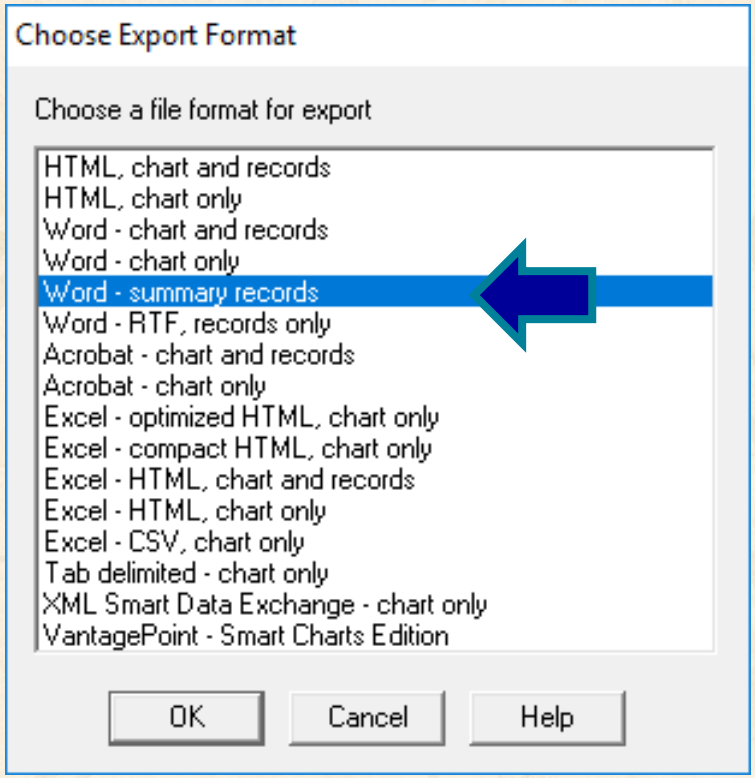
Minimum effort with Reference Rows

- Simply “Finish” on step two



Minimum effort with Reference Rows

- File | Export … Word - summary records



Reference Rows: Selection View

Unique fields are easily integrated in BizInt Smart Charts Reference Rows

Enhanced Title	Indications	Patent Type	Classifications	Family Status			
				Pub No.	State	Status	Expiry
Monoclonal antibodies and vaccines against epitopes on the Ebola virus glycoprotein	Ebola virus infection	Product	Anti-Infectives Biologicals and Immunologicals				
				WO200116183	DEAD	LAPSED	2006-03-26
				AU7089600	DEAD	LAPSED	2006-03-26
				US6630144	ALIVE	GRANTED	2020-08-29
Monoclonal antibodies against glycoprotein of Ebola Sudan Boniface (ESB) virus - useful in the diagnosis and treatment of ESB virus infection.	Ebola virus infection	Diagnostic, Analysis and Assay Product (Macromolecule)	Anti-Infectives Biologicals and Immunologicals Diagnostics				
				WO2011071574	ALIVE	PENDING	2030-09-01
				EP2473525	DEAD	LAPSED	2014-08-27
				US2012164153	ALIVE	PENDING	2030-09-01
Ebola virus liposome vaccines - useful in eliciting immune responses against Ebola virus infection.	Ebola virus infection	Formulation	Anti-Infectives Biologicals and Immunologicals Pharmaceutics				
				WO2012050193	DEAD	LAPSED	2013-12-03
				JP2014005205	ALIVE	PENDING	2030-10-14

Reference Rows: HTML exports

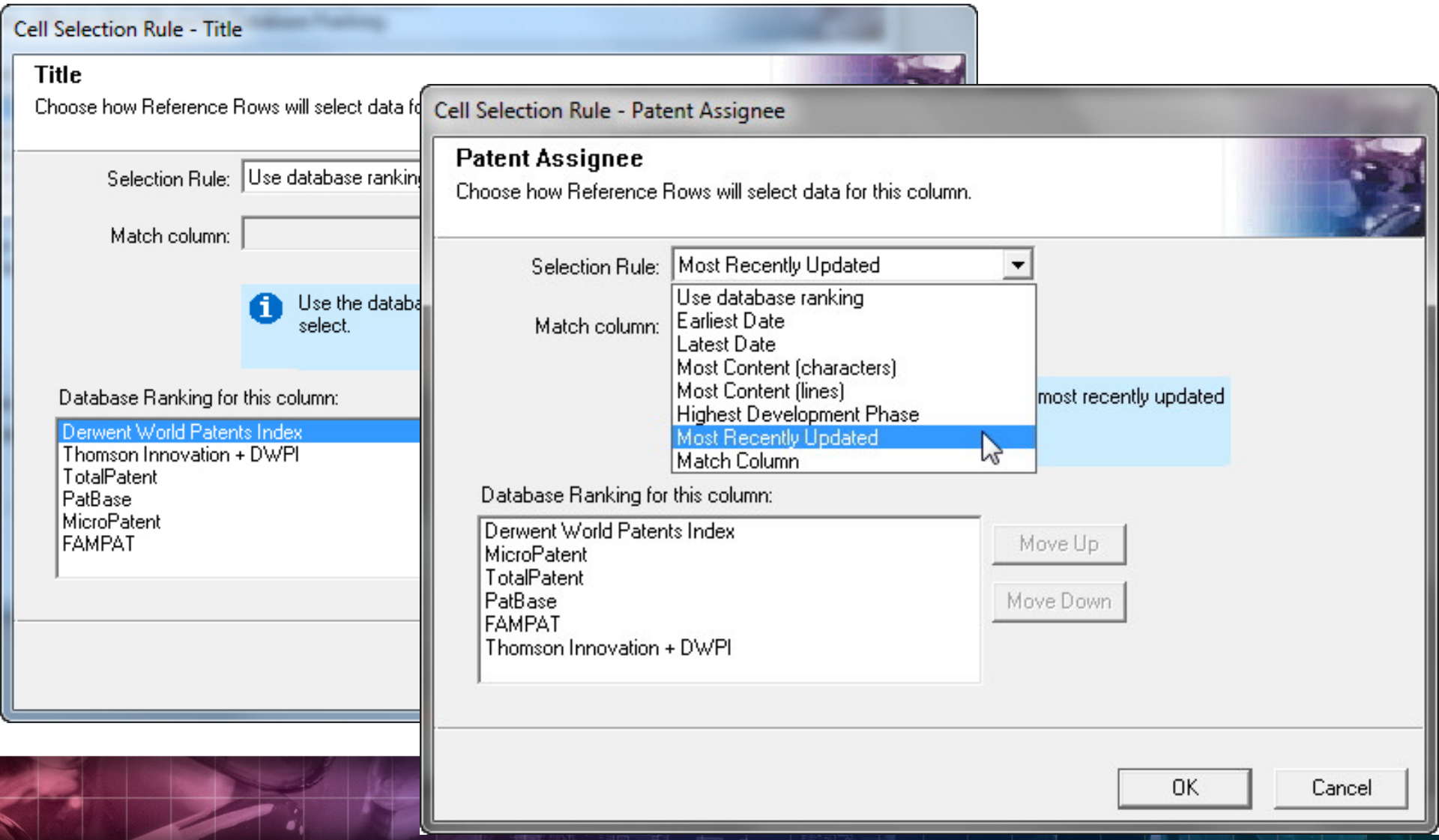
As seen in the fully integrated view

	Enhanced Title	Indications	Patent Type	Classifications	Family Status				Database
					Pub No.	State	Status	Expiry	
2.	**Monoclonal antibodies and vaccines against epitopes on the Ebola virus glycoprotein** 2.1 CORTP	Ebola virus infection 2.1 CORTP	Product 2.1 CORTP	Anti-Infectives Biologicals and Immunologicals 2.1 CORTP	WO200116183 AU7089600 US6630144	DEAD DEAD ALIVE	LAPSED LAPSED GRANTED	2006-03-26 2006-03-26 2020-08-29 2.2 FAMPAT	2.1 CORTP | *link* 2.2 FAMPAT | *link*
3.	**Monoclonal antibodies against glycoprotein of Ebola Sudan Boniface (ESB) virus - useful in the diagnosis and treatment of ESB virus infection.** 3.1 CORTP	Ebola virus infection 3.1 CORTP	Diagnostic, Analysis and Assay Product (Macromolecule) 3.1 CORTP	Anti-Infectives Biologicals and Immunologicals Diagnostics 3.1 CORTP	WO2011071574 EP2473525 US2012164153	ALIVE DEAD ALIVE	PENDING LAPSED PENDING	2030-09-01 2014-08-27 2030-09-01 3.2 FAMPAT	3.1 CORTP | *link* 3.2 FAMPAT | *link*
4.	**Ebola virus liposome vaccines - useful in eliciting immune responses against Ebola virus infection.** 4.1 CORTP	Ebola virus infection 4.1 CORTP	Formulation 4.1 CORTP	Anti-Infectives Biologicals and Immunologicals Pharmaceutics 4.1 CORTP	WO2012050193 JP2014005205	DEAD ALIVE	LAPSED PENDING	2013-12-03 2030-10-14 4.2 FAMPAT	4.1 CORTP | *link* 4.2 FAMPAT | *link*
5.	**Chimeric filovirus glycoproteins useful in vaccines against Ebola and Marburg virus infections** 5.1 CORTP	Marburg virus infection Ebola virus infection 5.1 CORTP	Product 5.1 CORTP	Anti-Infectives Biologicals and Immunologicals 5.1 CORTP	WO02079239 US7731975	DEAD DEAD	LAPSED LAPSED	2006-03-29 2014-06-08 5.2 FAMPAT	5.1 CORTP | *link* 5.2 FAMPAT | *link*

Integrate data from related records

	Enhanced Title	Database	Patent Family: Patent	Kind	Date	Family Status: Pub No.	State	Status	Expiry	Alignment	% Identity
5.	Methods for detecting the presence of isolated attenuated hEbola virus - useful as vaccines.	5.1 FAMPAT | link	WO 201048615	A2	2010-04-29	WO2010048615	ALIVE	PENDING	2029-10-26	Q: 1 SFKAALSSL 9	100.00
		5.2 CORTP | link	CA 2741523	A1	2010-04-29	AU2009308422	ALIVE	PENDING	2029-10-26	|||||||||	
		5.3 GPATPRT | link	AU 2009308422	A1	2010-04-29	CA2741523	ALIVE	PENDING	2029-10-26	S: 279 SFKAALSSL 287	
		5.4 GPATPRT | link	WO 201048615	A3	2010-11-25	EP2350270	ALIVE	PENDING	2029-10-26		
		5.5 GPATNUC | link	EP 2350270	A2	2011-08-03	IN3817/DELNP/2011	ALIVE	PENDING	2029-10-26		
		5.6 GPATNUC | link	EP 2350270	A4	2012-04-11	US2012251502	ALIVE	PENDING	2029-10-26		
		5.7 GENESEQ | link	US 20120251502	A1	2012-10-04						
	5.2 CORTP		IN 2011DN03817	A	2013-09-27	5.1 FAMPAT				5.3 GPATPRT	5.3 GPATPRT
6.	Recombinant biol... con... filov...	6.1 FAMPAT | link	WO 2009128867	A2	2009-10-22	WO2009128867	DEAD	LAPSED	2010...08	Q: 1 SFKAALSSL 9	100.00
		6.2 GENESEQ | link	WO 2009128867	A3	2010-03-25					|||||||||	
	6.1 FAMPAT					6.1 FAMPAT				S: 1 SF...AALSSL 9	
										6.2 GENESE	6.2 GENESE
7.	Nucleic acid comprising a polynucleotide encoding a modified filovirus glycoprotein - useful as vaccines against filovirus infections, specifically Ebola virus.	7.1 FAMPAT | link	WO 200637038	A1	2006-04-06	WO2006037038	ALIVE	PENDING	2025-09-27	Q: 1 HNTPVYKLDISEATQVE 17	100.00
		7.2 CORTP | link	CA 2581840	A1	2006-04-06	AU2005289439	ALIVE	GRANTED	2025-09-27	|||||||||||||||||	
		7.3 GPATPRT | link	AU 2005289439	A1	2006-04-06	CA2581840	ALIVE	GRANTED	2025-09-27	S: 389 HNTPVYKLDISEATQVE 405	
		7.4 GPATPRT | link	WO 200637038	A9	2006-05-26	EP1797113	ALIVE	GRANTED	2025-09-27		
		7.5 GPATPRT | link	WO 200637038	B1	2006-08-03	IL182225	DEAD	LAPSED	2012-09-20		
		7.6 GPATPRT | link	EP 1797113	A1	2007-06-20	IN2674/DELNP/2007	ALIVE	GRANTED	2025-09-27		
		7.7 GPATPRT | link	IN 2007DN02674	A	2007-08-03	JP2008514203	ALIVE	GRANTED	2025-09-27		
		7.8 GENESEQ | link	IL 182225	D0	2007-09-20	US2009232841	ALIVE	GRANTED	2027-06-07		
		7.9 GENESEQ | link	JP 2008514203	A	2008-05-08	US8101739	ALIVE	GRANTED	2027-06-07		
		7.10 GENESEQ | link	US 20090232841	A1	2009-09-17	US2012156239	ALIVE	PENDING	2025-09-27		
			AU 2005289439	B2	2011-12-01						
			US 8101739	B2	2012-01-24						
			US 20120156239	A1	2012-06-21						
			JP 5046941	B2	2012-10-10						
			IN 259912	B	2014-04-04						
			CA 2581840	C	2014-08-05						
	7.2 CORTP		EP 1797113	B1	2014-11-26	7.1 FAMPAT				7.3 GPATPRT	7.3 GPATPRT

Reference Rows: user-defined rules



Summarize data from related records

	Title	Database	Patent Assignee	Query ID	Sequence Locations				
					Seq. ID Number	% Identity	Length	Location	
1.	PRODUCTION OF PEPTIDES IN PLANTS AS VIRAL COAT PROTEIN FUSION 1.1 Patbase	1.1 Patbase | link 1.2 GENESEQ | link	LARGE SCALE BIOLOGY CORP. 1.2 GENESE	query2	WO20050108564-0101	100.00	17	Example 6; SEQ ID NO 101;	1.2
2.	Chimeric ebola virus envelopes and uses therefor 2.1 Patbase	2.1 Patbase | link 2.2 GPATPRT | link 2.3 GPATPRT | link 2.4 GPATPRT | link 2.5 GPATPRT | link 2.6 GENESEQ | link 2.7 GENESEQ | link	UNIV PENNSYLVANIA. 2.6 GENESE	query2 query	US20050255123-0001	100.00	17	claim: 17	2.2
					WO03092582-0009	100.00	498	claim: 17	2.3
					WO03092582-0001	100.00	17	claim: 17	2.4
					US20050255123-0009	100.00	498	claim: 17	2.5
					WO20030092582-0001	100.00	17	Claim 17; SEQ ID NO 1; 107pp; English.	2.6
					WO20030092582-0009	100.00	498	Claim 17; SEQ ID NO 9; 107pp; English.	2.7
3.	ANTIGEN FRAGMENT AND TRUNCATION BASED ON EBOLA VIRUS ENVELOPE PROTEIN AS WELL AS APPLICATION 3.1 Patbase	3.1 Patbase | link 3.2 GENESEQ | link 3.3 GENESEQ | link	BIOENGINEERING RES INST ACAD MEDICAL SCI. 3.2 GENESE	query2	CN103864904-0008	100.00		SEQ ID NO 8; 28pp; Chinese.	3.2
					CN103864904-0002	100.00	17	Example 1; SEQ ID NO 2; 28pp; Chinese.	3.3
4.	HUMAN EBOLA VIRUS SPECIES AND COMPOSITIONS AND METHODS THEREOF 4.1 Patbase	4.1 Patbase | link 4.2 GPATPRT | link 4.3 GPATPRT | link 4.4 GPATNUC | link 4.5 GPATNUC | link 4.6 GENESEQ | link	US DEPT HEALTH & HUMAN SERVICES. 4.6 GENESE	query7 query5	US20120251502-0011	100.00	9	claim: 8; 11; 12	4.2
					EP2350270-0011	100.00	9	TBD (information not in GQ-Pat)	4.3
					US20120251502-0027	100.00	20	probable disclosure (not found by automated parsing)	4.4
					EP2350270-0027	100.00	20	TBD (information not in GQ-Pat)	4.5
					WO20100048615-0027	100.00	20	Claim 30; SEQ ID NO 27; 98pp; English.	4.6

Sequence Summary Recipe

- Recipe for creating the sequence summary table at bizint.com/cookbook
- Create Subtable from any columns you want
  NOTE: alignment loses fixed width formatting
- In Reference Rows, choose Summarize All Values column rule
- Export chart

Summary Record export

1.	Title:	**Modulating expression of a target nucleic acid comprises providing to the cell a guide RNA including a transcriptional activator or repressor domain as a fusion protein, and providing to the cell a nuclease null Cas9 protein**
	Database:	Derwent World Patents Index Derwent World Patents Index GQPAT Gold+ Proteins GQPAT Gold+ Proteins PatBase FAMPAT

Patent Family:

Patent	Kind	Date
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

Family Status:

Pub No.	State	Status	Expiry
US 20140356956 A1	ALIVE	PENDING	2034-06-04
US 9267135 B2	ALIVE	GRANTED	2034-06-04

Probable Assignee: PRESIDENT AND FELLOWS OF HARVARD COLLEGE

Sequence Locations:

Seq. ID Number	% Identity	Length	Location
US20140356959-0001	100.00	1368	probable disclosure (not found by automated parsing)
US20140356956-0001	100.00	1368	probable disclosure (not found by automated parsing)

Notes

Alignment:

```
Q:          1  MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAE  60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
S:          1  MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAE  60
```

Summarize at Patent or Sequence Level

- Always start with Tools | Identify Common Patent Family to create the Common Family column (this is a “magic” column)
- Replace contents of the Common Family column with the data you want to group by
  Select column, copy
  Select Common Family, paste
- Patent level: Patent Number
- Sequence level: Sequence ID (pub+seqidno)

Software for
Business Intelligence

BizInt Smart Charts

Questions?

www.bizint.com