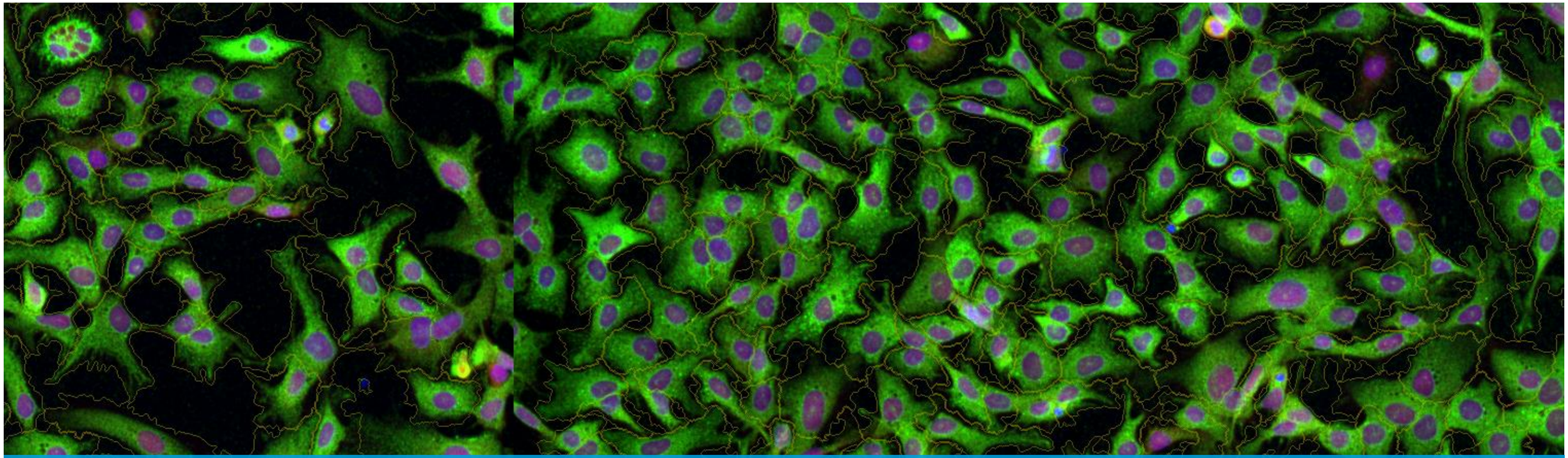


Build & explore virtual libraries for drug discovery projects in Janssen

Zhijie Liu - representing Janssen US CompChem Group
Senior Scientist
Discovery Chemistry, Janssen R&D

NIH Virtual Workshop on Ultra-Large Chemistry Databases, Dec 1-3, 2020

Discovery Sciences

HepG2 hepatocellular carcinoma cells treated with tunicamycin. Nuclei in magenta, and PDI in green.

janssen
PHARMACEUTICAL COMPANIES OF Johnson & Johnson

Outline

1. **Construction of unique virtual libraries within Janssen CompChem group**

   - **Janssen virtual library (JVL)**
   - **Janssen fragment spaces (JFS)**

2. **Applications of ultra-large libraries for an ongoing kinase project**

   - **Identify new hits/leads from Enamine REAL database**
   - **Discover the second new series from JVL and JFS**

Virtual screening – from commercial to ultra-large virtual

Commercially available		Commercially synthesizable	
eMolecules®	ZINC DATABASE	Enamine REAL DATABASE	Bio Solve IT 药明康德 WuXi AppTec
19M	ZINC20 >750M	1.4B	GalaXi™ 1.7B
Virtually synthesizable		**Virtually enumerated**	
Synthetically Accessible Virtual Inventory (SAVI) Database First Full Release | First Complete Beta Release	SAVI-2020 1.8B	GDB-13 970M	
BioSolveIT expect actives!	KnowledgeSpace 290T	GDB-17 170B	

Construct Janssen unique & synthesizable virtual libraries!

Construct unique Janssen virtual libraries

Skin cells at 20x magnification

janssen
PHARMACEUTICAL COMPANIES OF Johnson & Johnson

Process for construction of Janssen Virtual Library (JVL)

Building blocks (BB): Internal & commercial

2M readily available BBs from Janssen & 21 vendors

→

Create filters for BB collections

900K BBs after filtering with physicochemical properties and organic filters*

→

Check for proprietary BBs

300K BBs after excluding vendor BBs and Zinc in-stock BBs

↓

Define functional groups (FG)

Classify BBs via 90 functional groups

←

Reaction registration

50 reactions with chemistry intelligence including reaction scope and limitations

←

JVL – utilizes basis product approach

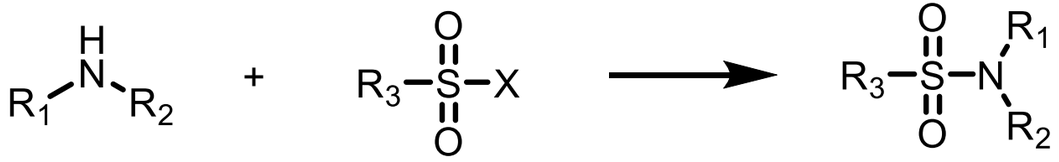
Registration of 10^8 molecules from 50 reactions

* Organic filter defined in Pipeline Pilot

All compounds in the JVL will have at least one Janssen proprietary building block!

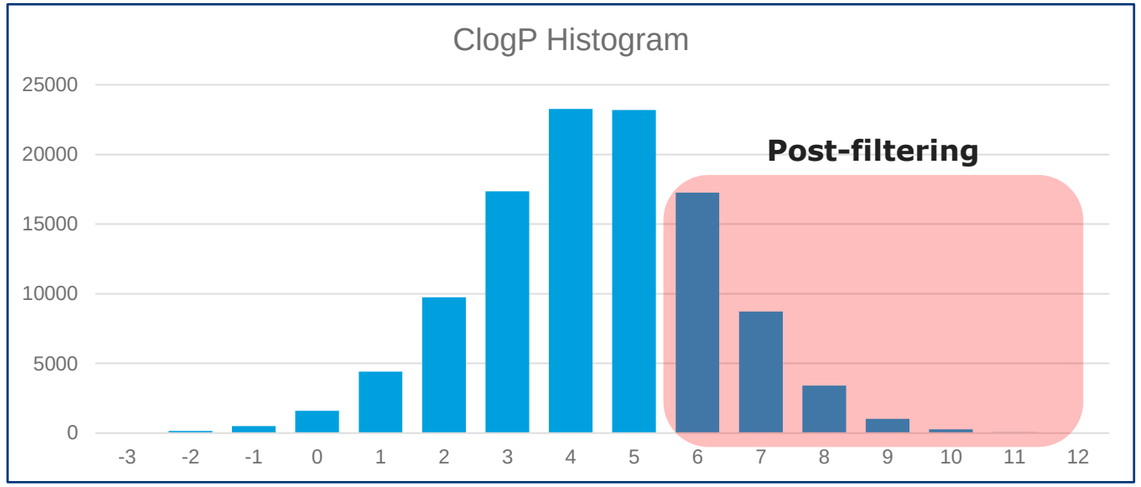
JVL reaction example: sulfonamide formation

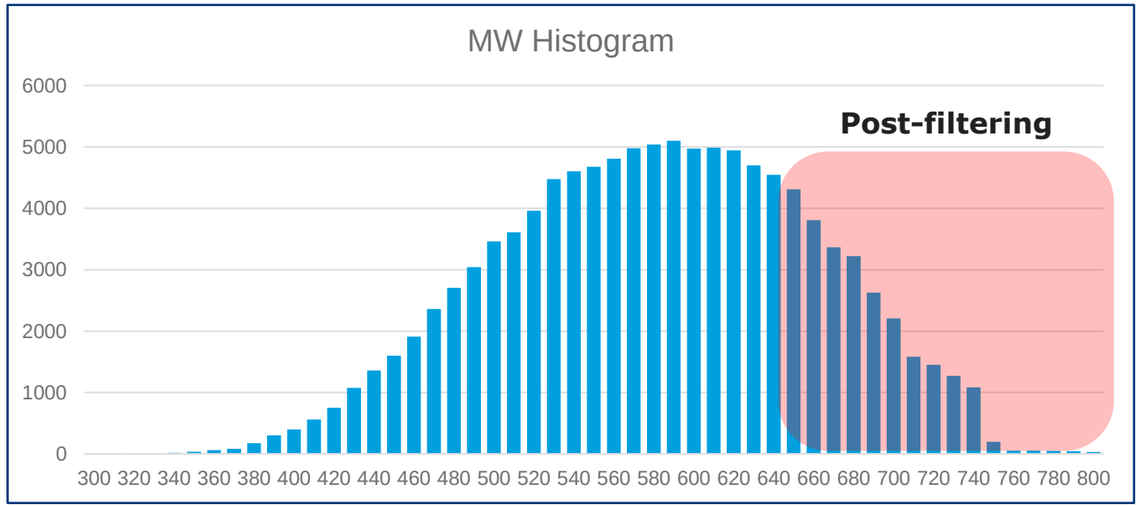
Sulfonamide Formation from Amines and Sulfonyl Halides



X = Cl,Br

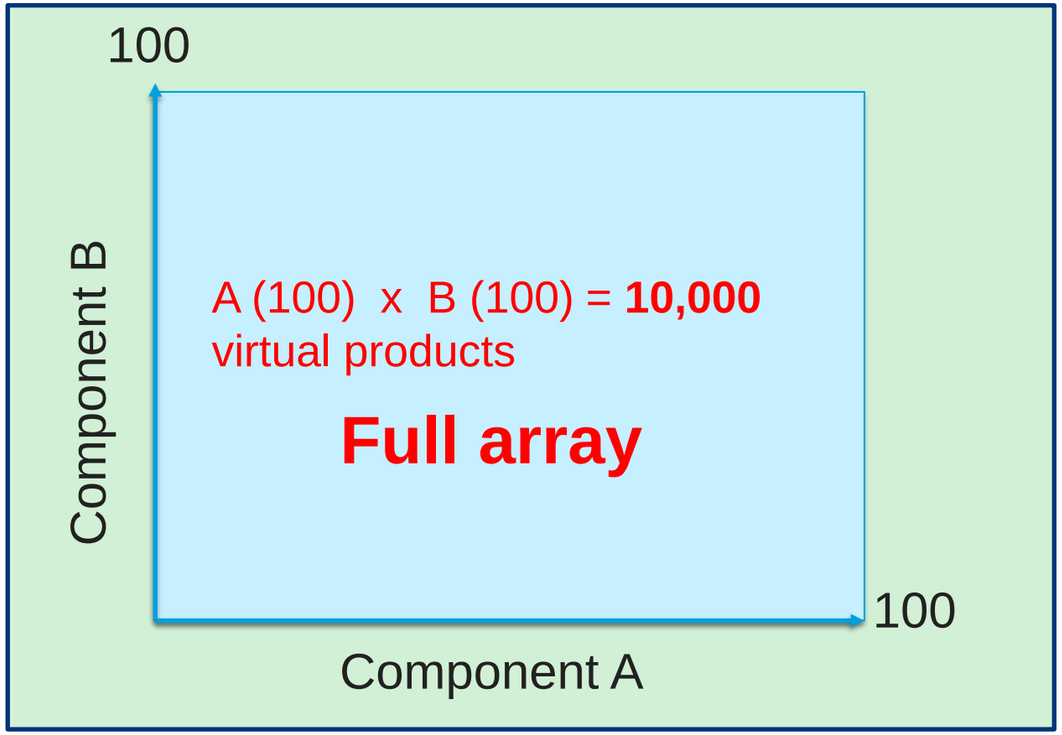
	Proprietary	Non-Proprietary
Amines	~25,000	~44,000
Sulfonyl Halides	12	~3,000
VL Size = ~76,000,000		



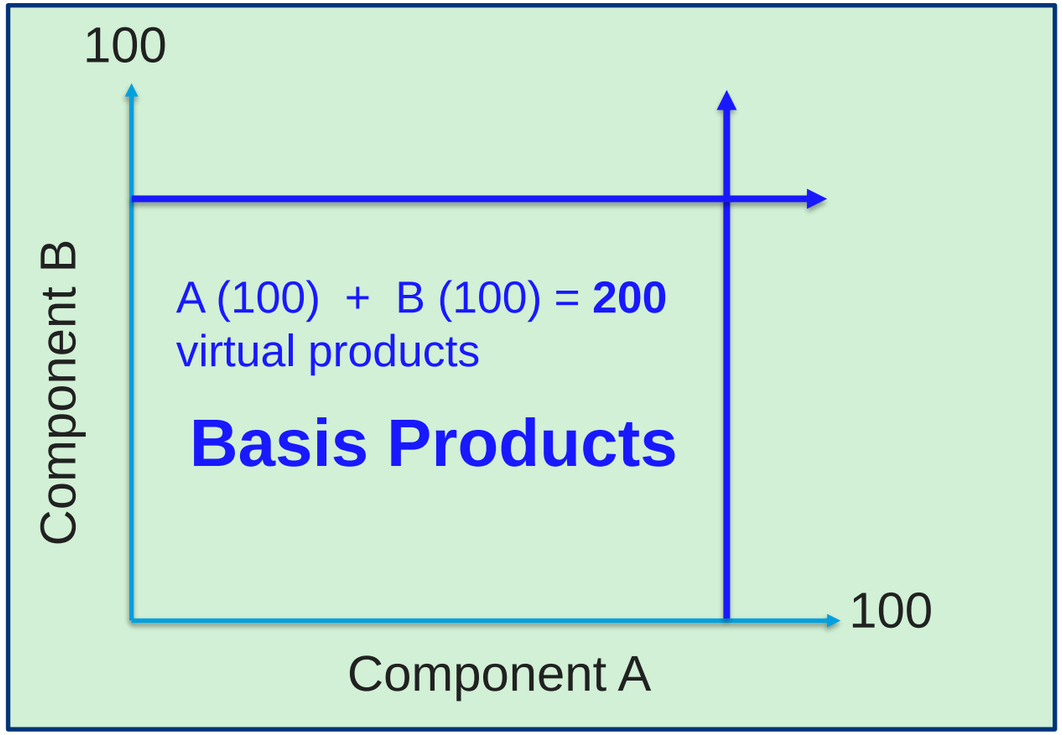


Strategy of using basis products to represent huge VLs

In a typical 2 component reaction, a **full combinatorial array** would be A x B = #Products



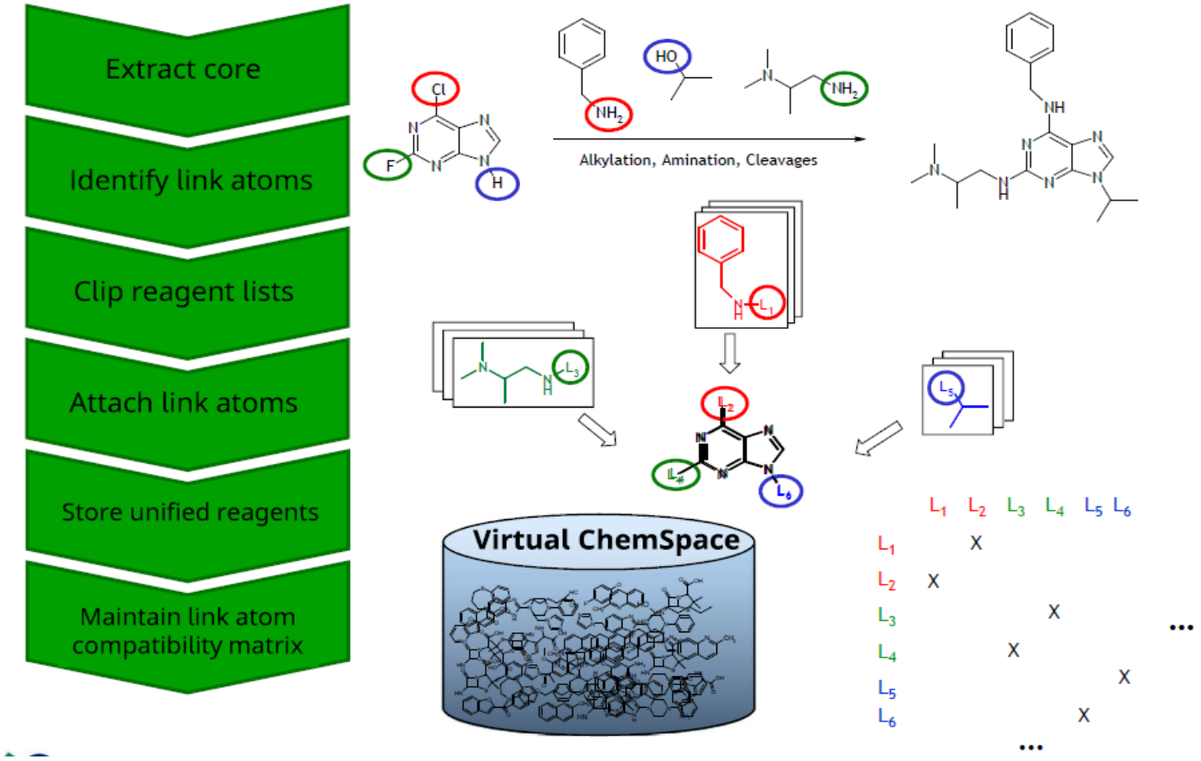
To fully sample the R-group space covered by each component, form the **simplest basis products**. Now you have A + B = #Products



Construct ultra-large Janssen fragment space (JFS) using CoLibri and apply for FTrees similarity search

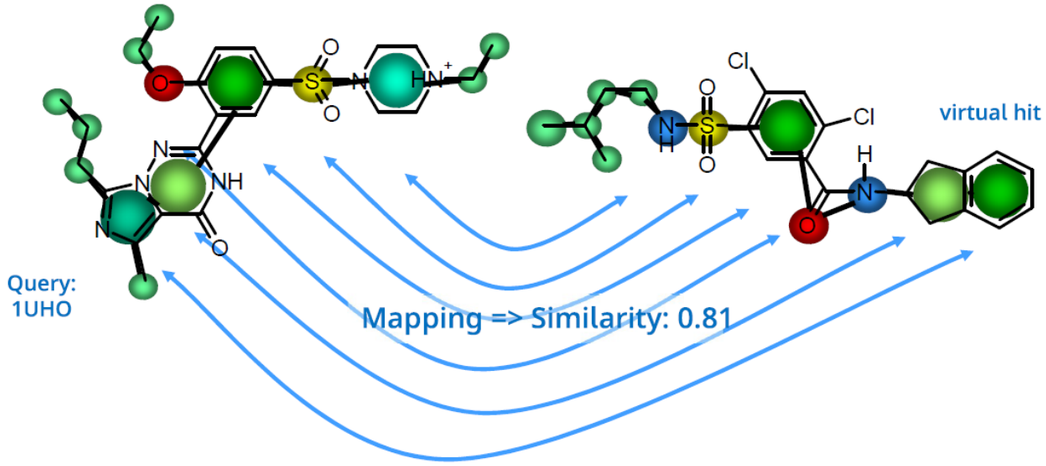
CoLibri fragment space

Reagent matrix linked by reactions



FTrees similarity search

Topological matching molecular pharmacophore feature trees



https://www.biosolveit.de/wp-content/uploads/2020/10/FTrees.pdf

https://www.biosolveit.de/wp-content/uploads/2020/11/CoLibri.pdf

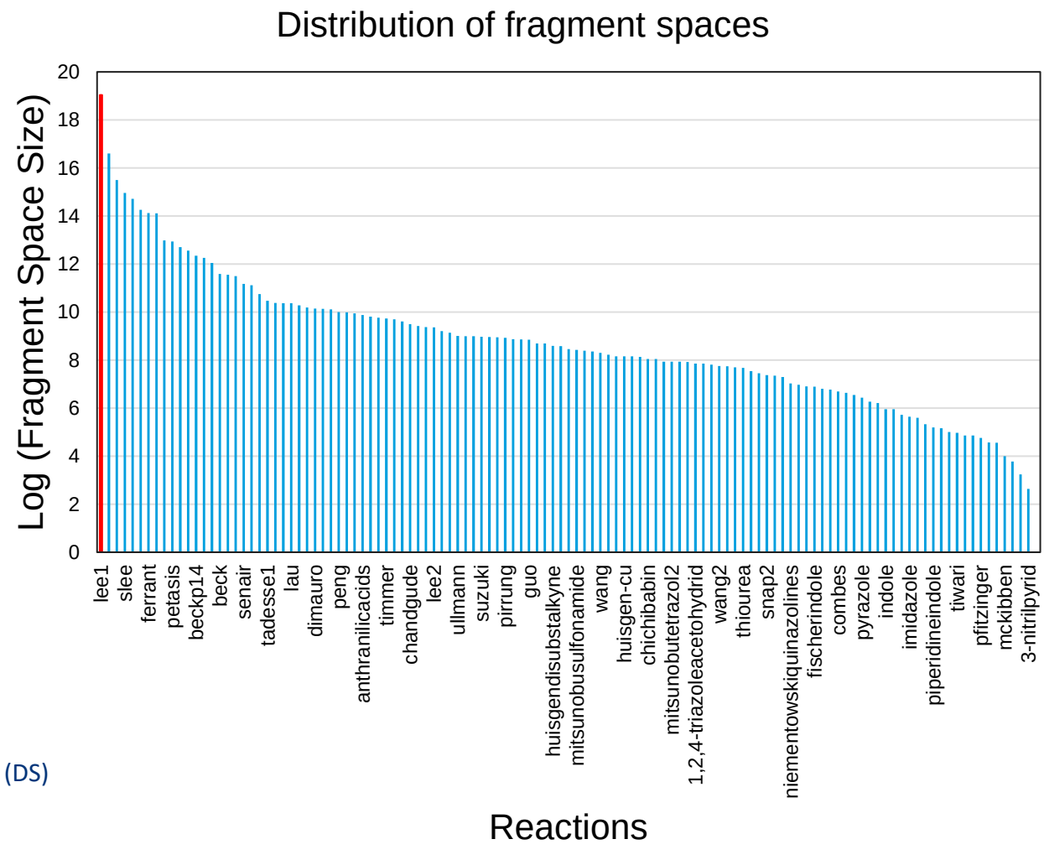
Process for construction of ultra-large Janssen fragment space

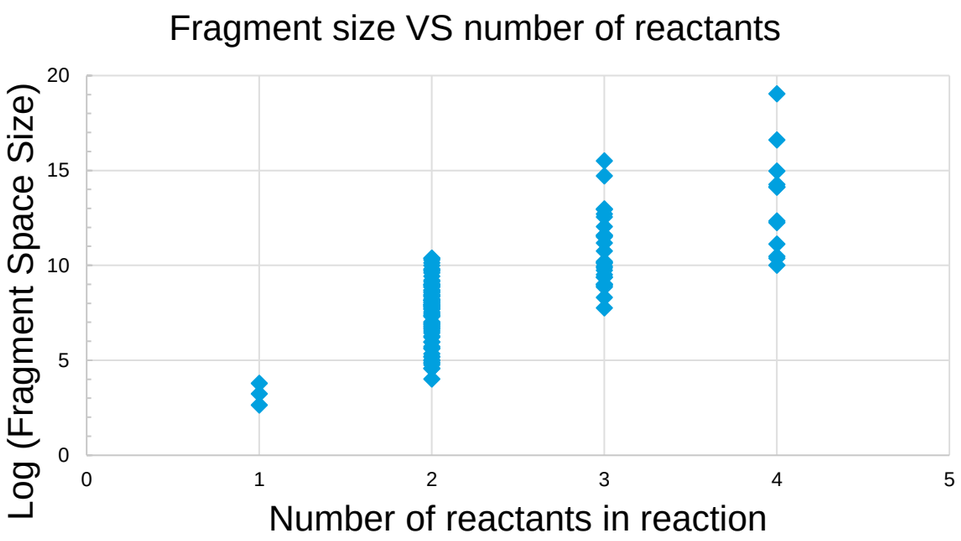
Collect BBs → Filter & clean BBs → Janssen Proprietary → Janssen Fragment Spaces → Similarity Search

- CoLibri 120 reactions
- 2 fragment spaces

- FTrees
- Kinase project

- Fragment spaces dominated by molecules from reactions with multiple reactants





JFS includes more than 10^{19} molecules

Janssen spaces	Space size
Janssen Fragment Space (JFS)	1.1×10^{19}
All Fragment Space (AFS)	3.0×10^{19}

CRO spaces	Space size
Enamine REAL Space	1.3×10^{10}
Wuxi GalaXi fragment space	1.7×10^{9}
BioSolveIt Knowledge Space	2.9×10^{14}

Pharma spaces	Space size*
Lilly LPC 2016	10^{10}
BI BICLAIM 2012	10^{11}
Pfizer PGVL 2008	10^{14}
Evotec EvoSpace 2016	10^{16}
AstraZeneca 2018	10^{17}
Merck MASSIV 2018	10^{20}

All compounds in the JFS have at least one Janssen proprietary building block!

JFS has comparable size to VLs from CRO & pharma !

* Hoffmann T, Gastreich M, Drug Discov Today, 2019, 24(5), 1148-1156

Applications of ultra-large libraries for an ongoing kinase project

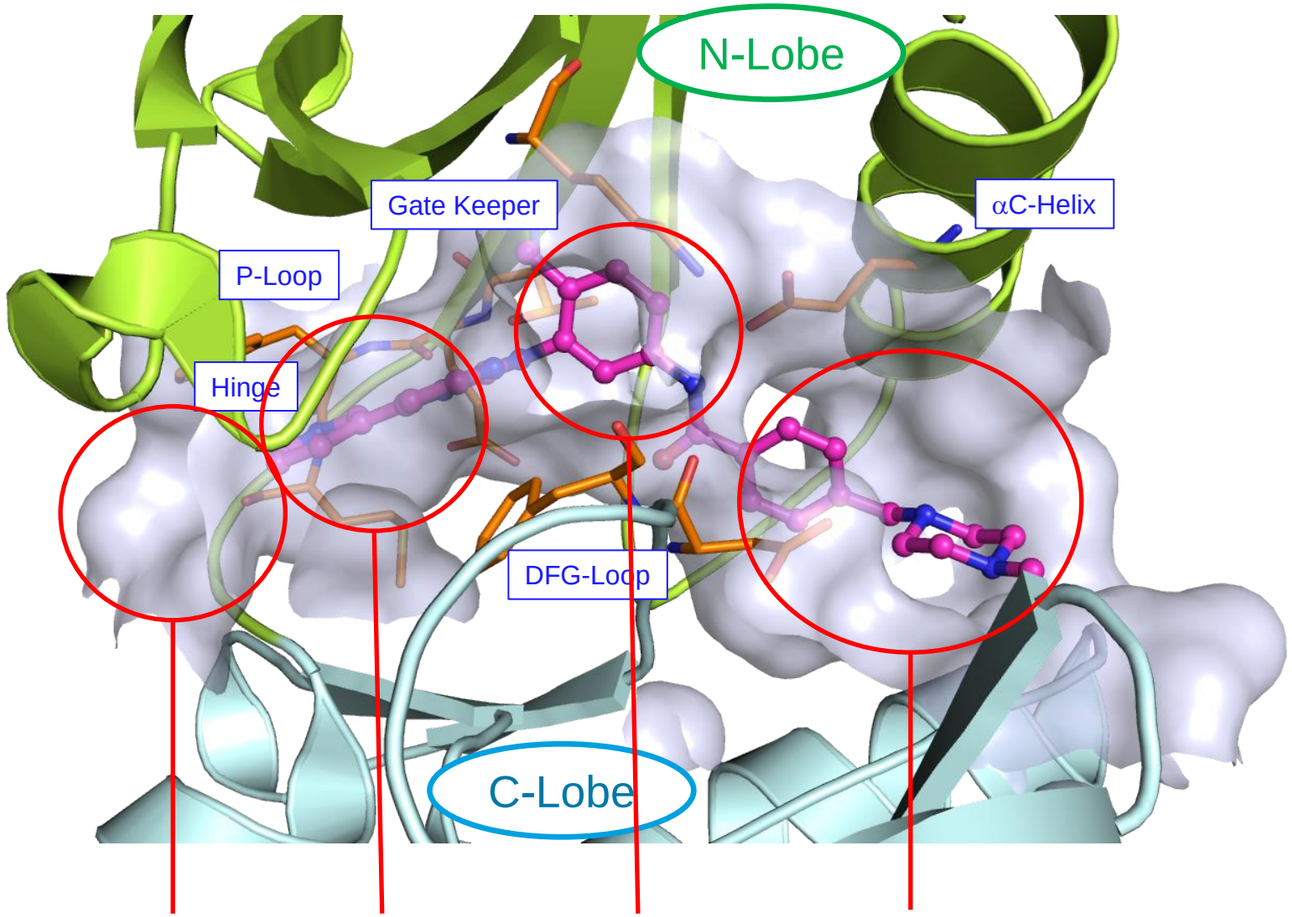
Skin cells at 20x magnification

janssen
PHARMACEUTICAL COMPANIES OF Johnson & Johnson

Develop type-2 inhibitors for an ongoing kinase project

Target inactive state (0P) of kinase to achieve:

- High potency
- Slow off-rate
- High kinase selectivity
- High solubility
- Multiple scaffolds



Front Pocket | ATP Pocket | Hydrophobic Pocket | Back Pocket

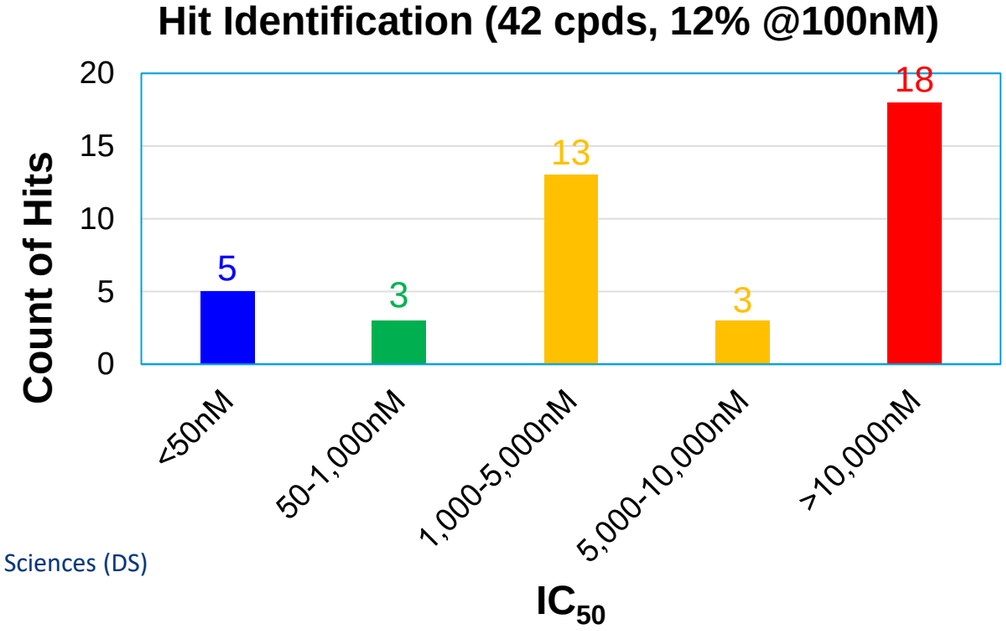
Example: Imatinib with c-Abl (PDB: 1IEP)

Virtual screening discovered potent series from commercial VL

740M Enamine REAL database

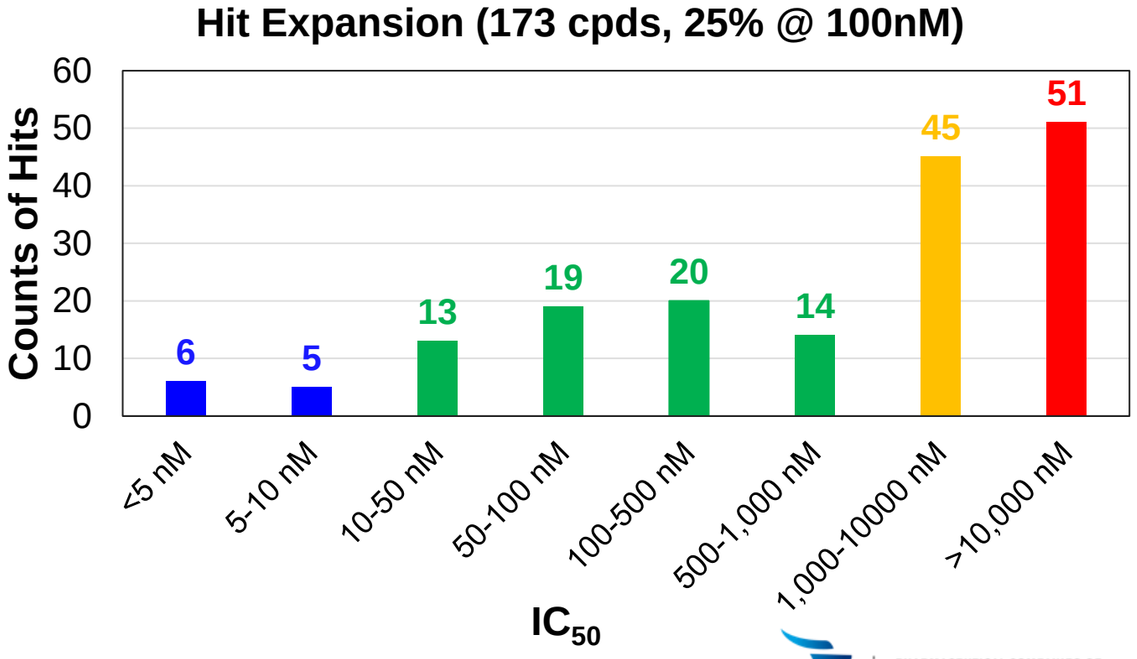
Hit identification via virtual screening (VS)

1. Ligand-based similarity search
   - FastROCS: shape & color
   - FTrees: topology & pharmacophore
2. Structure-based: Glide docking
3. Interaction & property filtering
4. Clustering & selection

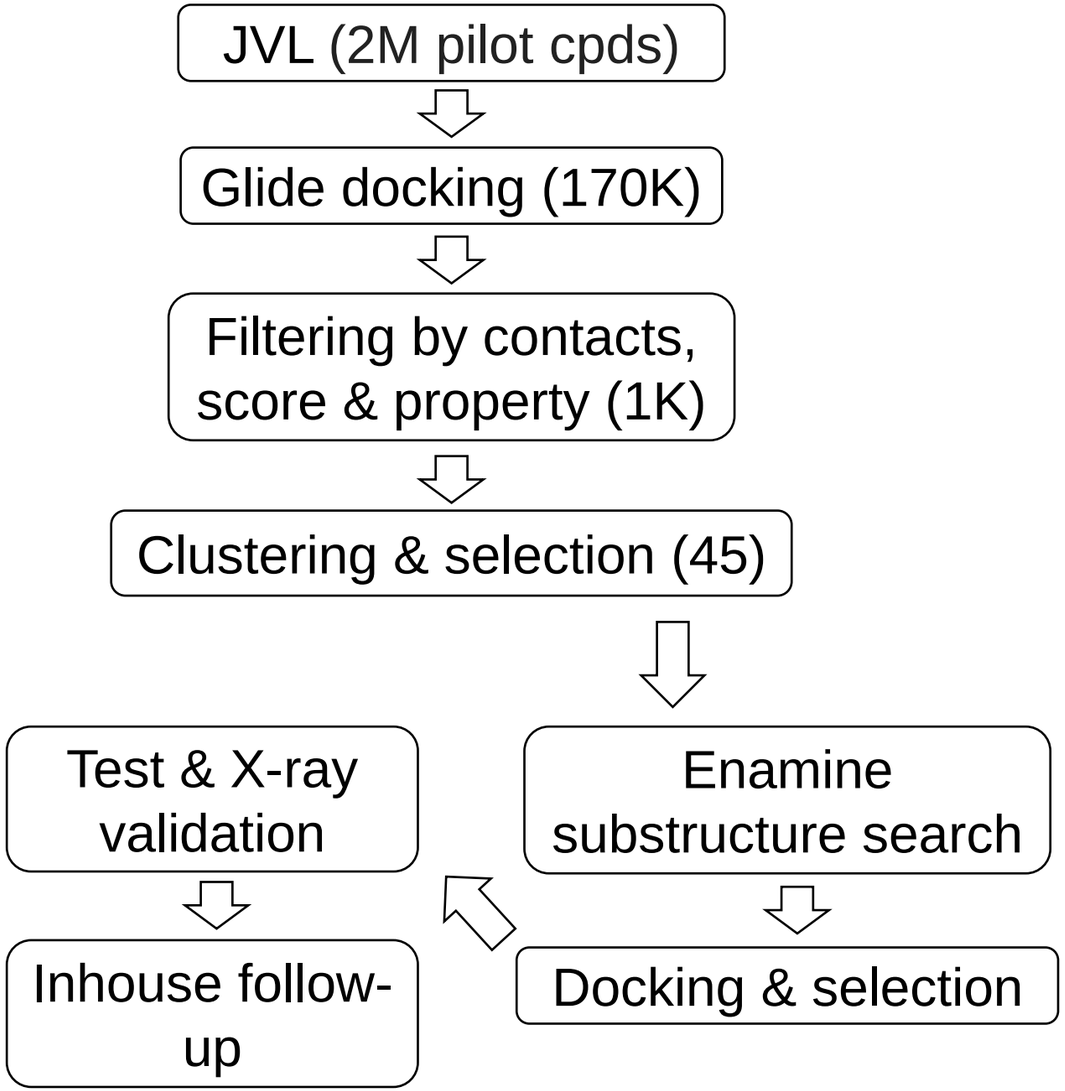


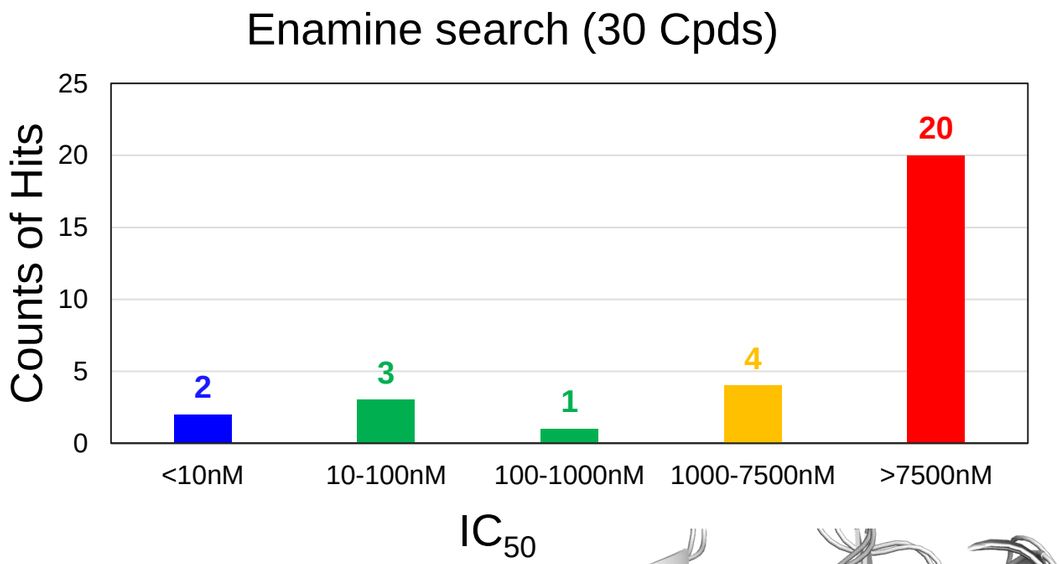
Hit expansion

1. Substructure search
2. Structure-based: Glide docking
3. Property filtering & selection

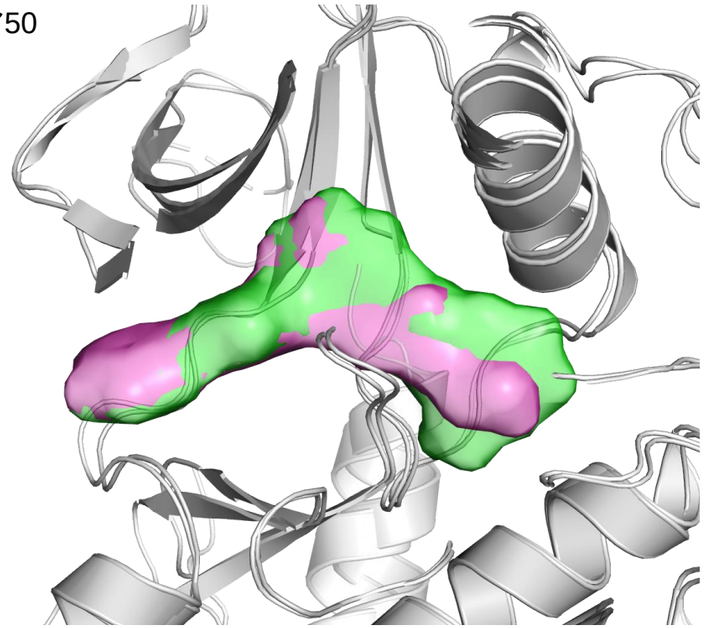


Structure-based VS identified new scaffold from JVL





- Single-digit nanomolar hit
- Docking model is validated by X-ray structure



JFS contains unique chemical space

Query kinase inhibitor

⇩

FTrees similarity search

⇩

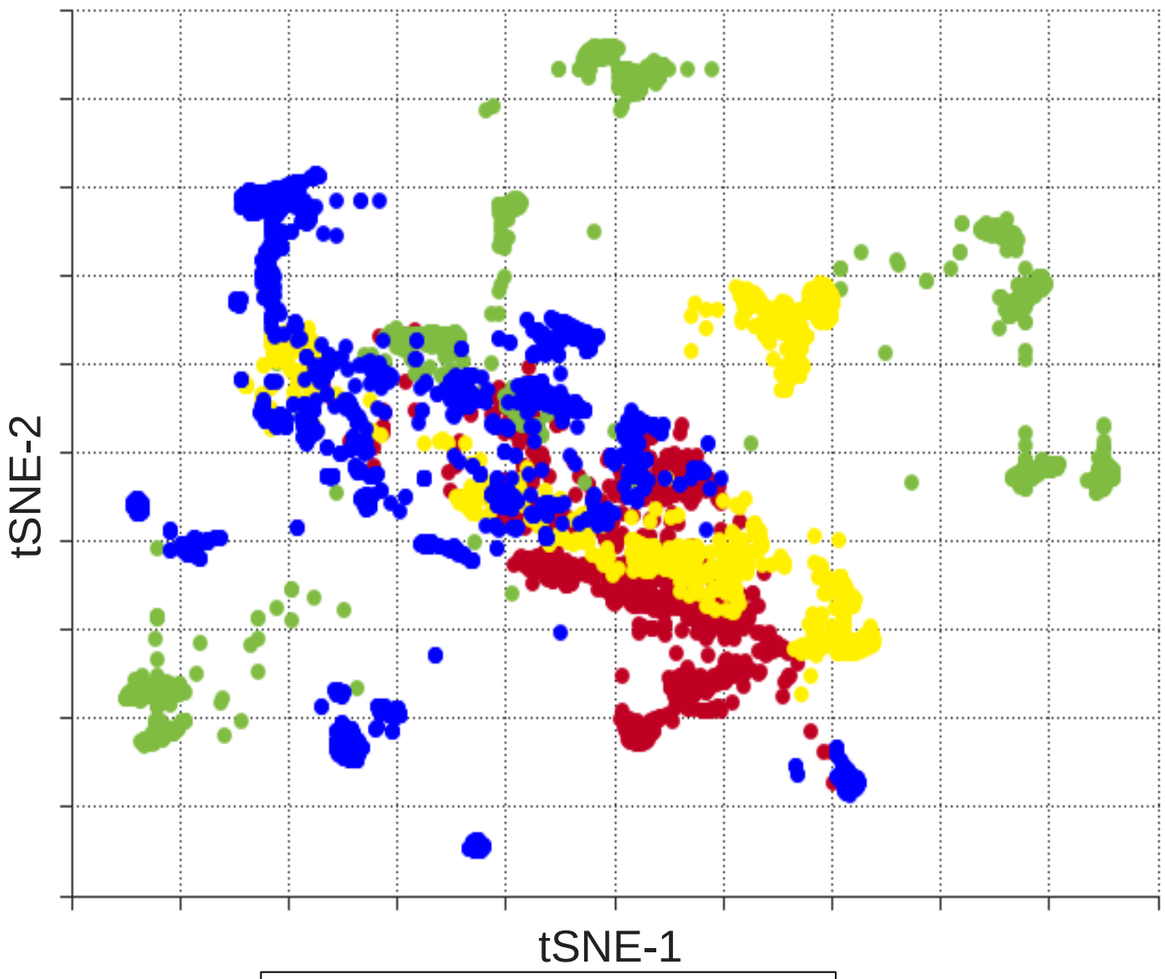
JFS | Enamine REAL space | GalaXi | BioSolveIt Knowledge space

⇩

Top 1K selected each

⇩

Chemical space comparison



Analysis of chemical space

- JNJ_FG_FragmentSpace
- GalaXi_2019-10
- EnamineRealSpace_2019-12
- KnowledgeSpace_2019-05

Discovered additional 608 new molecules

Summary

To support drug discovery projects with much large and unique Janssen chemical space:

- Applied the strategy of basis products to cover all Janssen BBs and constructed JVL to explore unique chemical space with all applicable virtual screening methods
- Constructed ultra-large Janssen fragment space (JFS) to enhance the sampling of Janssen chemical space
- Successfully identified two new potent series for an ongoing kinase project using JVL & JFS and commercial virtual library

Acknowledgement

Janssen CompChem

- Jim Na
- Zhengwei Peng
- Nadeem Vellore
- Mark Seierstad
- Michael Hack
- Berthold Wroblowski
- Herman Van Vlijmen
- Renee DesJarlais
- Tara Mirzadegan
- James Edwards

Janssen Kinase Project Team

- Tianbao Lu
- David Bauman
- Vladimir Chupakhin
- Daniel Krosky
- Rodrigo Ortiz-Meoz
- David Duda

BioSolveIt

Christian Lemmen

Franca Klingler

Markus Lilienthal