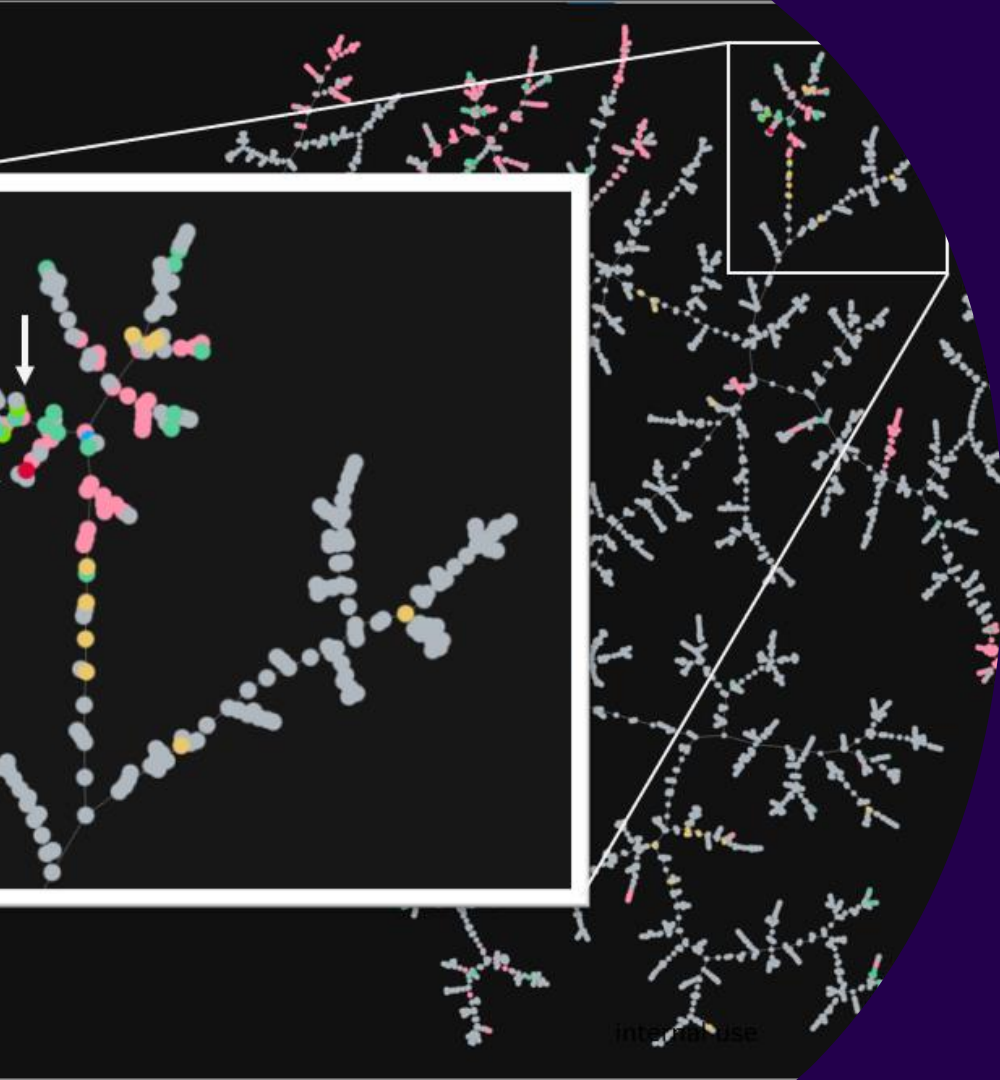


sanofi



sanofi

•
Practice of
Generative AI in
lead optimization

Marc Bianciotto

March 14th 2024

AiChemist Spring School - Berlin

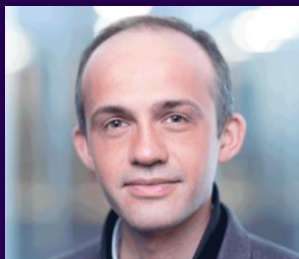
•

Dr. Maximilien Levesque
(now Aqemia)



Pr. Rodolphe Vuilleumier

Dr. Maxime Langevin



And in Sanofi

CADD: B. Filoche, C. Terrier, C. Grebner, Y. Li, S. Guesreguen, H. Matter, S. Sauer

Chemistry: S. Desprets and the medicinal chemists involved in the Molecular Turing Test

"Marc Bianciotto is a Sanofi employee and may hold shares and/or stock options in the company"



We operate in
90
countries



There are
91,000
Sanofians
worldwide

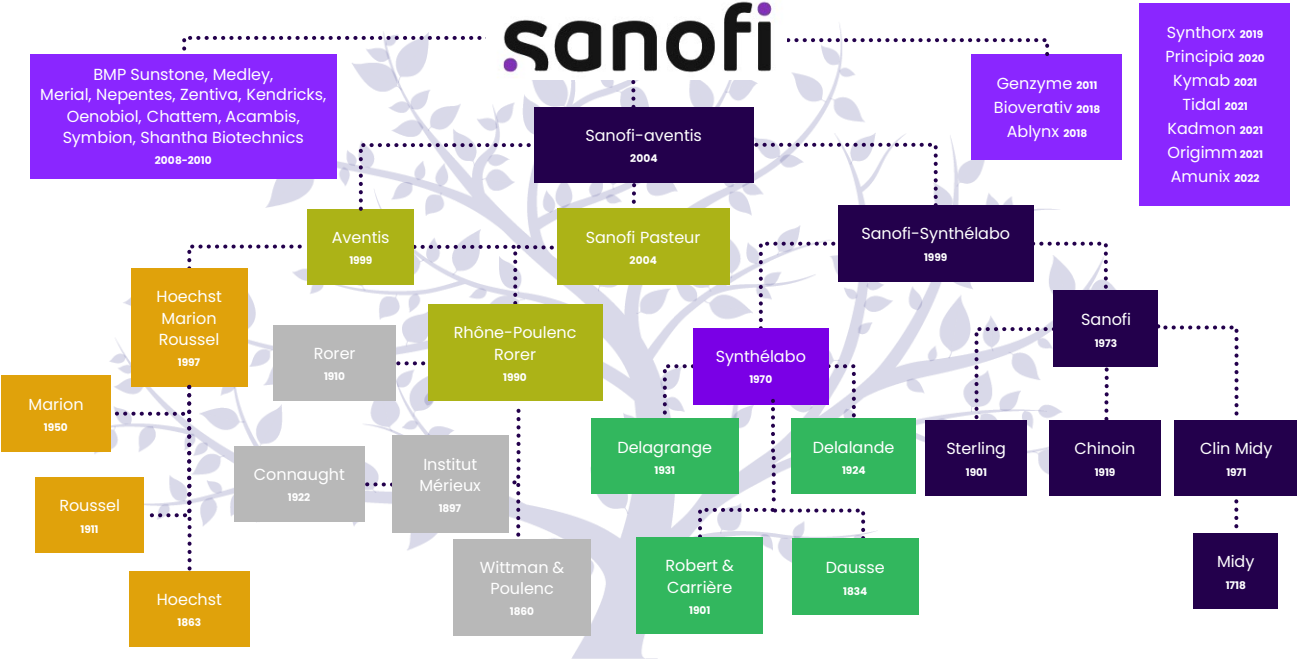


We work in
59
manufacturing
sites



We have
20
R&D sites

The history of Sanofi

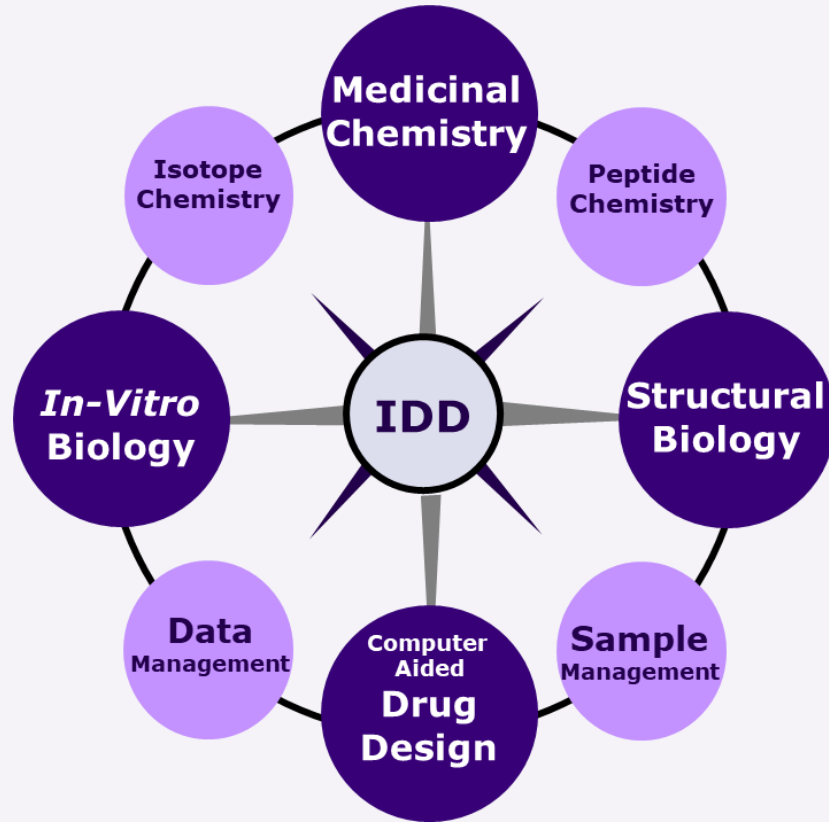


Integrated Drug Discovery

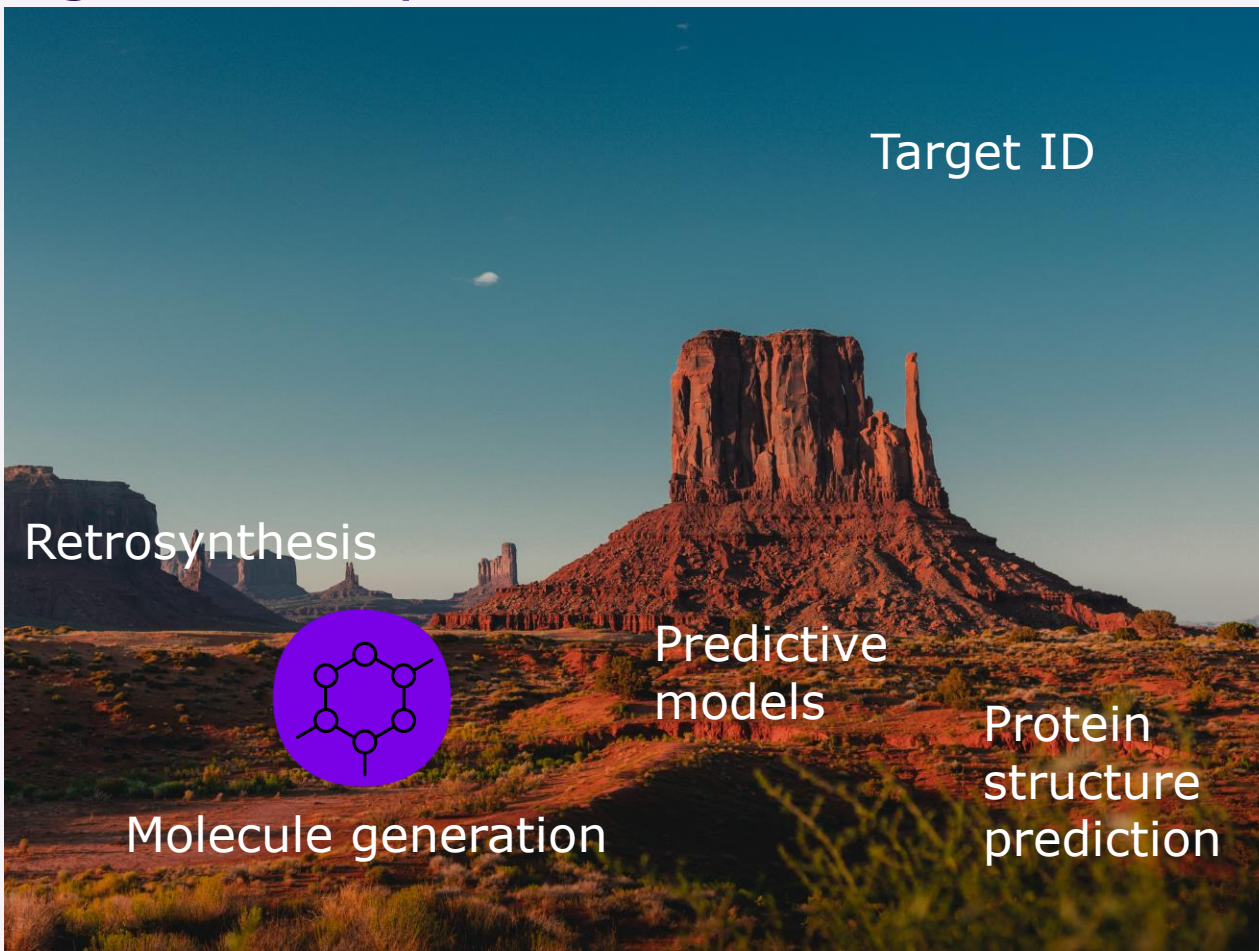


sanofi

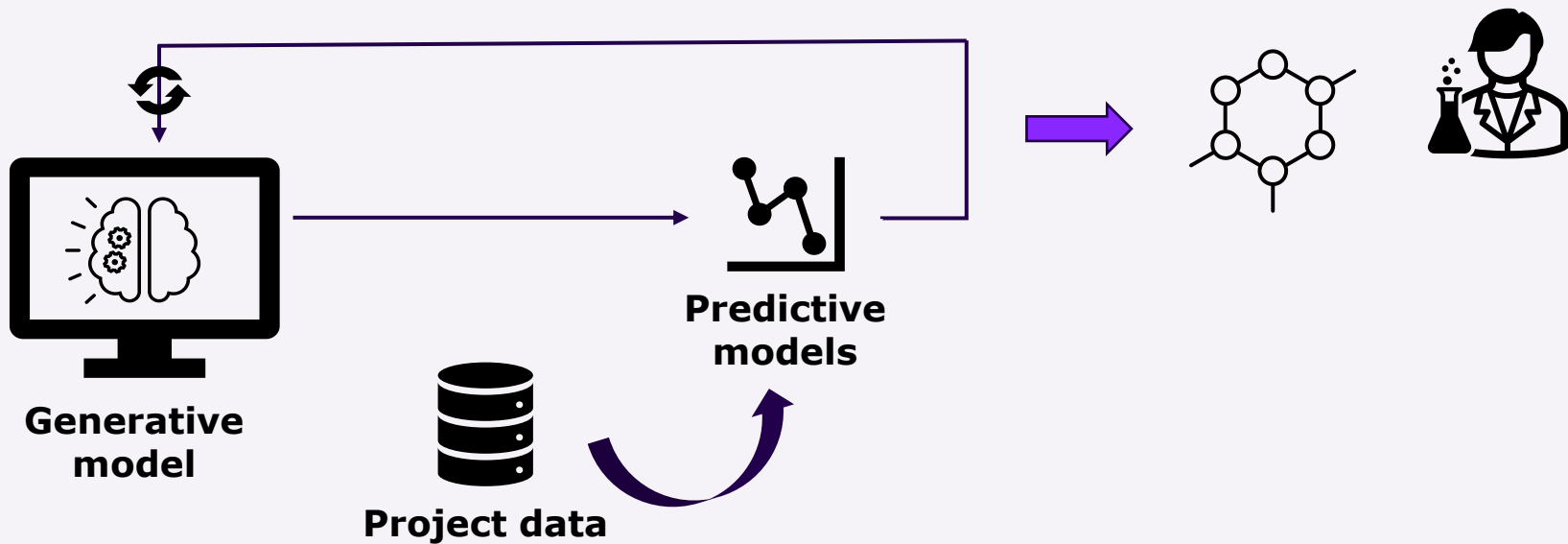
Integrated Drug Discovery France



AI for Drug Discovery: the Gold Rush

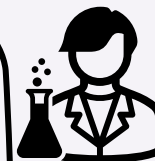


Generate molecules with AI?

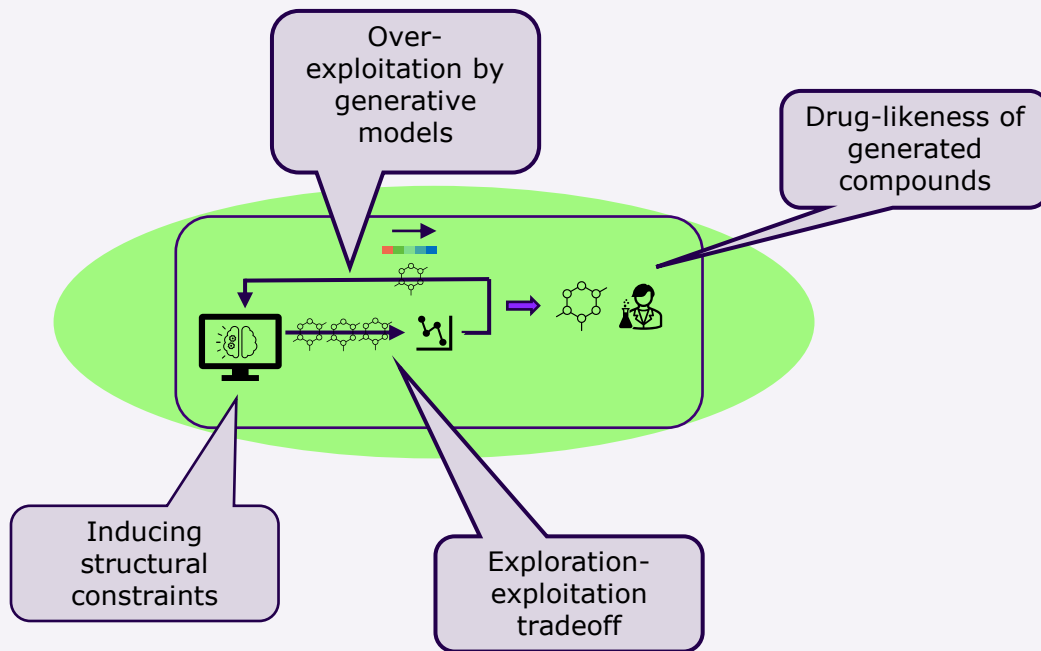


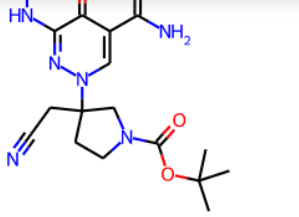
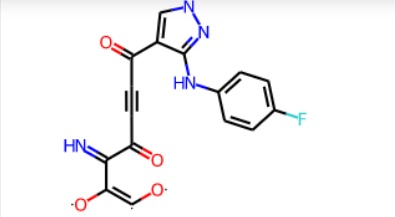
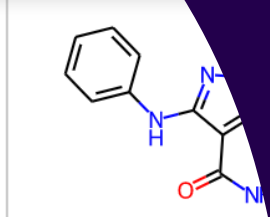
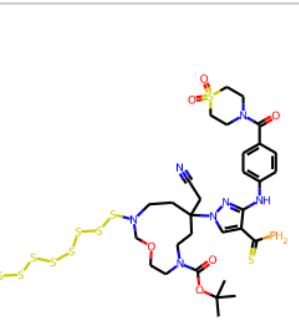
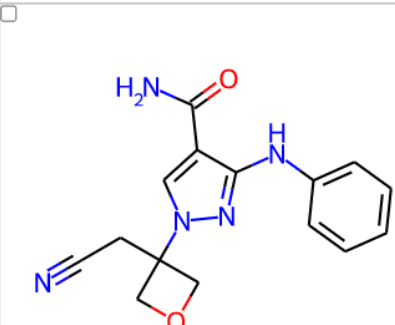
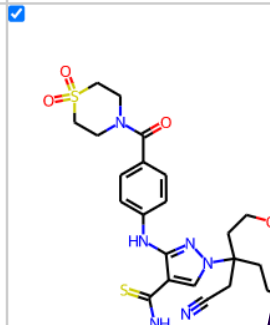
Generate molecules with AI?

	Lead Gen.	Lead Opt.
Project data	Less	More
Predictive models performance	Worse	Better
Chemical space to explore	Larger	More focused
Synthesis likelihood	Lower	Higher
Integration in the legacy workflow	More challenging	Less challenging



Topics of concern in small molecule GenAI for lead optimization



 <p>3</p>	 <p>4</p>	 <p>5</p>
 <p>6</p>	 <p>7</p>	 <p>8</p>

sanofi

Translation to practice

number to scroll through the different panels, from 1 to

Generative AI for protein kinase X: last effort

Objectives:

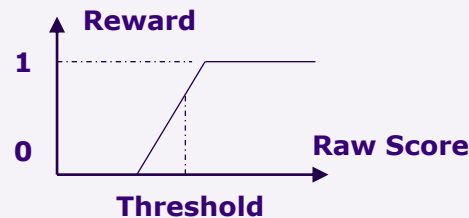
- Evaluate Med Chemists proposals
- Generate a last round of compounds
 - Close to the pre-candidate (same scaffold, 2 diversification points)
 - **Optimized on 8 properties**
 - Synthetically readily accessible



Protein kinase X: reward function

```
▼ root:  
  ▶ GLOBAL:  
  ▶ NNGP_APDP_regression-bioch_pIC50:  
  ▶ NNGP_RDKit_regression-cell_pIC50:  
  ▶ NNGP_RDKit_regression-FUM_Brain:  
  ▶ NNGP_APDP_regression-CLint:  
  ▶ NNGP_APDP_regression-  
  ▶ NNGP_RDKit_regression-  
  ▶ deepchem-  
  ▶ deepchem-  
  ▶ deepchem-  
  ▶ deepchem-logD74_all-logD-1std:  
  ▶ deepchem-Solubility-uM_pH74-1std:  
  ▶ has_not_substructure:
```

- Composite scoring function
 - Biochemical, cell pIC50s, counter-targets, unbound fraction in Brain, solubility, logD
 - Several models per readout if the models are correct
- AD for generation
- Score thresholds from the Med. Chemists' propositions



Generation outcome

Human intelligence + ML score:

- **73 close analogs of known cpds**
- 12 « flat cpds » selected
- => **23 synthesized with stereo**

AI design:

- 857 cpds generated in-criteria
 - **108 selected for Med Chemists**
- 10 « flat » cpds selected for synthesis
- => **17 synthesized with stereo.**
- **Non-trivial R-groups**

Origin	Ki vT (nM)	IC50 (nM)	IC50 vT (nM)	cell vT (nM)	Caco-2/efflux	INH Eurofins (Midazolam) (µM)	Kinact (min-1) Ki (µM)	IEM (CYP1A2 / 2B6 / 3A4)	Clearance Results human HES	Clearance Results mouse HES	Clearance Results rat HES	Cav1.2 (µM); Cav3.1 (µM); Kv4.3	Equilibrated solubility at pH2 - pH6.5; Fassif (µM)	solubility pH 7.4 (nephelometric) (µM)	Fu plasma (b/male) in %	Fu brain (female) (female)
generative design	1	1	1	1	2000.0	1.00	1.00010 / 30	See full vT33 CYP1A2, 2B6 and CYP3A4	0.47	2.97	2.97	1.00 - 1.00; 1.00 - 1.00	900-999-023	94	1.010.0	0.4
Med Chem then scoring (ML)	4	7	4	5	2000.0	1.00	Not used	See full vT33 CYP1A2, 2B6 and CYP3A4	0.30	1.00	1.1	1.00 - 1.00; 1.00 - 1.00	220-229-001	1907	0.110.0	0.3
Med Chem then scoring (ML)	3	3	2	20	2000.0	1.00	0.04710.0	See full vT33 CYP1A2, 2B6 and CYP3A4	0.70, 1.07	0.00-0.40	40, 1.07	1.00 - 1.00; 1.00 - 1.00	4407-000-001	1980	0.110.0	0.1
Med Chem then scoring (ML)	2	1	2	5	2000.0	1.00	1.00010 / 30	See full vT33 CYP1A2 and 2B6, moderate vT of CYP3A4	0.40	2.07	2.07	1.00 - 1.00; 1.00 - 1.00	1400-999-000	1.000	0.110.0	0.1
Med Chem then scoring (ML)	100+0	200+0	200+0	27	2000.0	1.00	Not used	See full vT33 CYP1A2, 2B6 and CYP3A4	10.1	1.21, 2.00	40, 0.4	1.00 - 1.00; 1.00 - 1.00	210-000-010	1907	0.110.0	0.3
generative design	3	4	5	5	200.0	1.00	1.00010 / 30	See full vT33 CYP1A2, 2B6 and CYP3A4	0.46	0.72, 2.02	40, 0.30	1.00 - 1.00; 1.00 - 1.00	1.000-000-000	1902	0.110.0	0.3
generative design	4	10	7	20	2000.0	1.00	1.00010 / 30	See full vT33 CYP1A2, 2B6 and CYP3A4	0.02	1.00, 2.00	40, 0.02	1.00 - 1.00; 1.00 - 1.00	700-000-000	1.010	0.110.0	0.3
Med Chem then scoring (ML)	4	10	5	40	2000.0	1.00	1.00010 / 30	See full vT33 CYP1A2, 2B6 and CYP3A4	10.1	0.77, 1.00	40, 0.07	1.00 - 1.00; 1.00 - 1.00	4200-000-020	1907	0.110.0	0.1

A shortlist of 5+3 cpds that fulfill the vast majority of 15 criteria



Chemical space

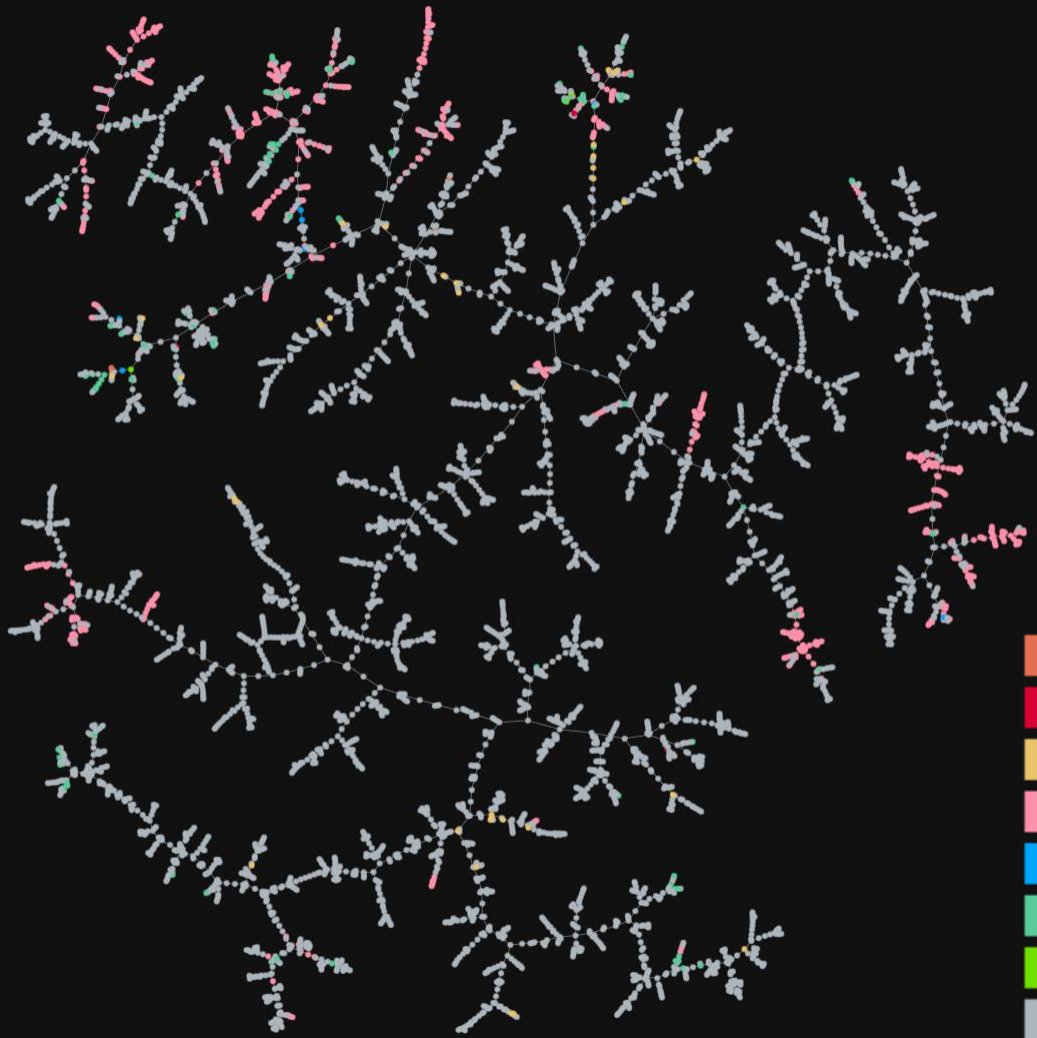
Colored by
molecule origin and fate



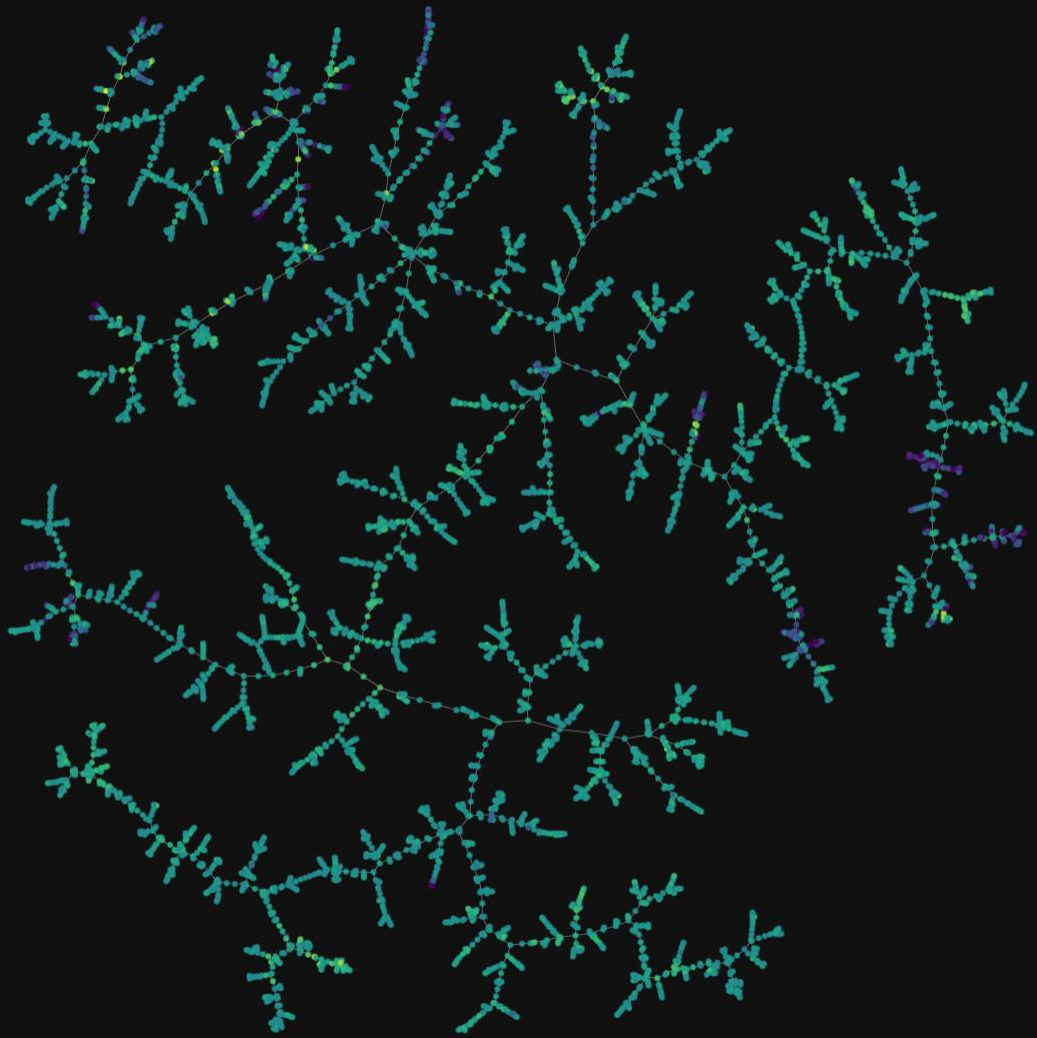
Legend

Assay
Set

- MedChem, Synthesized
- MedChem, Best
- MedChem
- Known
- Generated, Synthesized
- Generated, Selected
- Generated, Best
- Generated

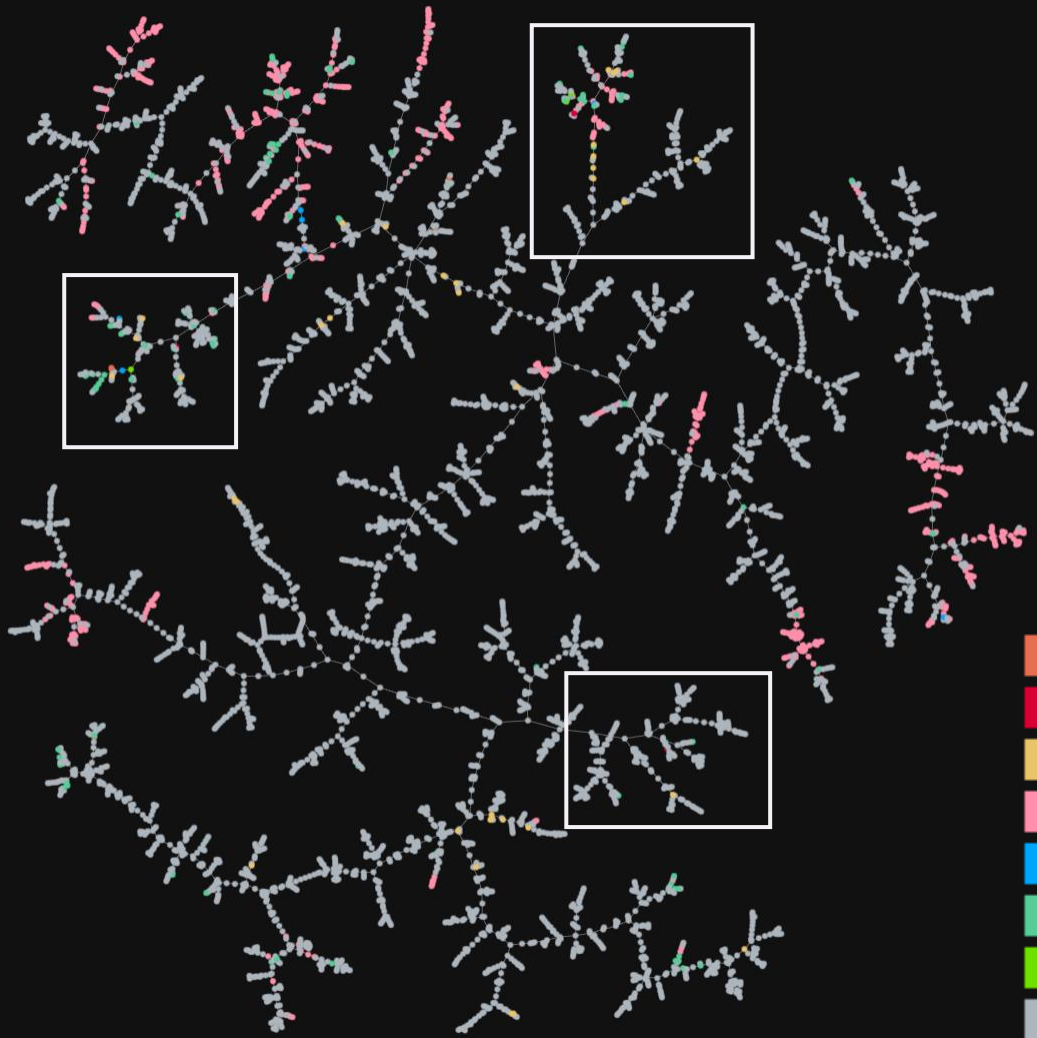


- MedChem, Synthesized
- MedChem, Best
- MedChem
- Known
- Generated, Synthesized
- Generated, Selected
- Generated, Best
- Generated

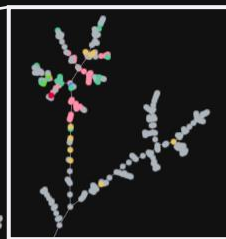
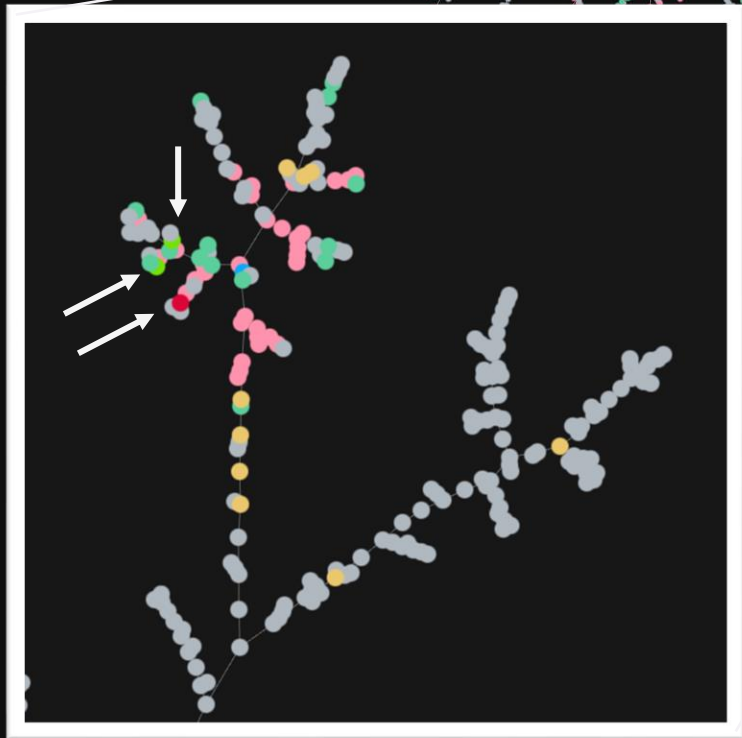


Colored by
Multi-parametric score

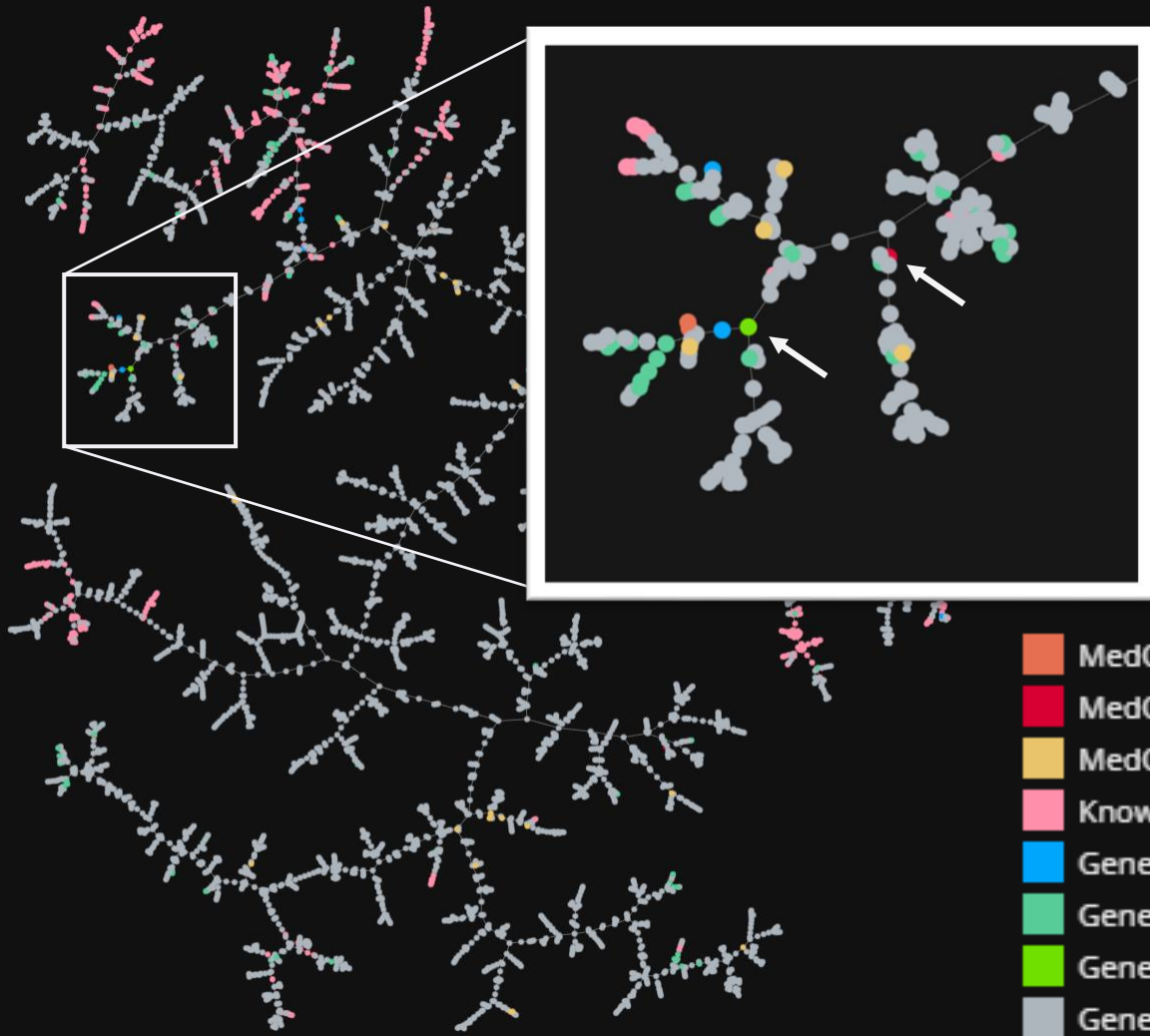


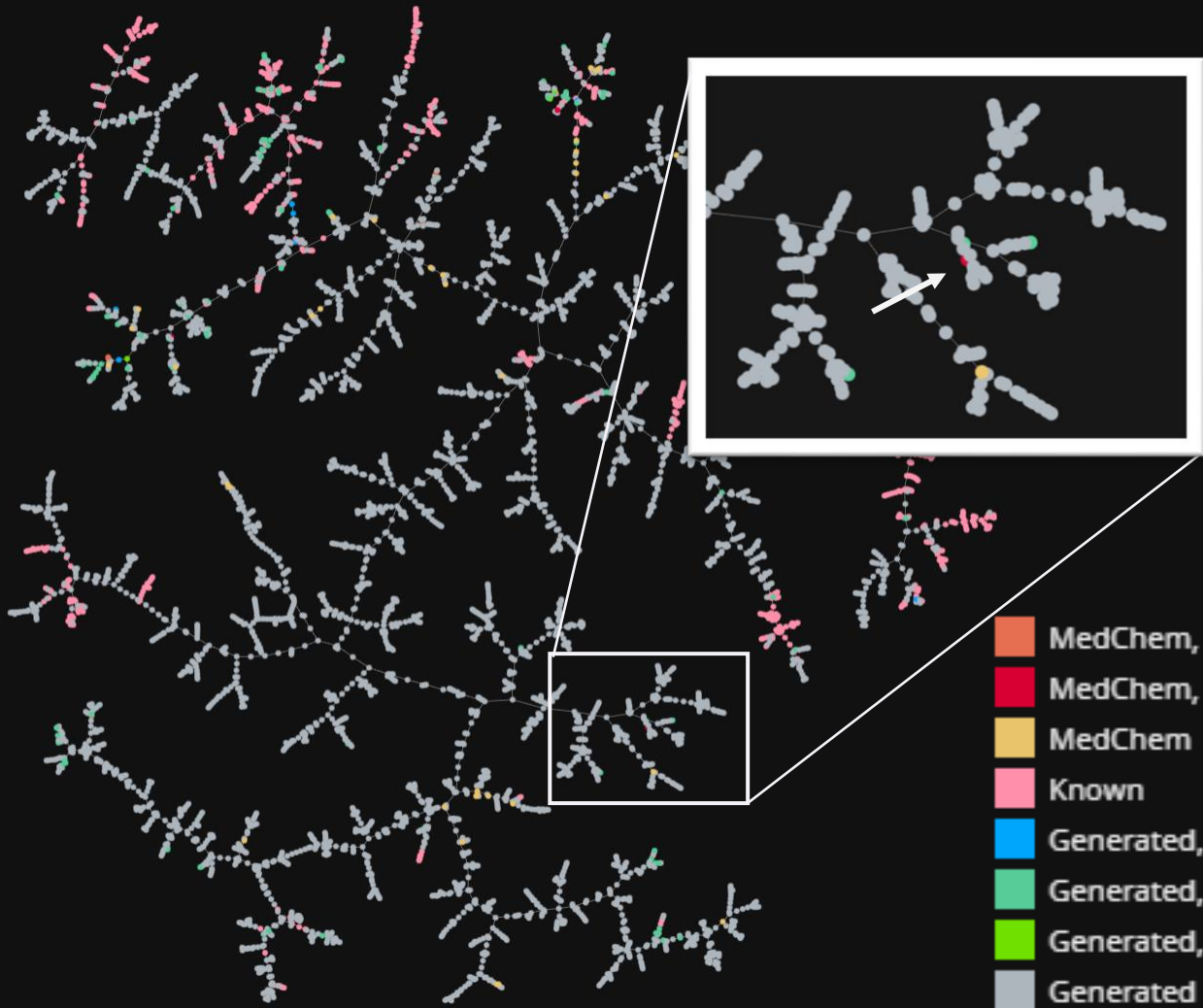


- MedChem, Synthesized
- MedChem, Best
- MedChem
- Known
- Generated, Synthesized
- Generated, Selected
- Generated, Best
- Generated

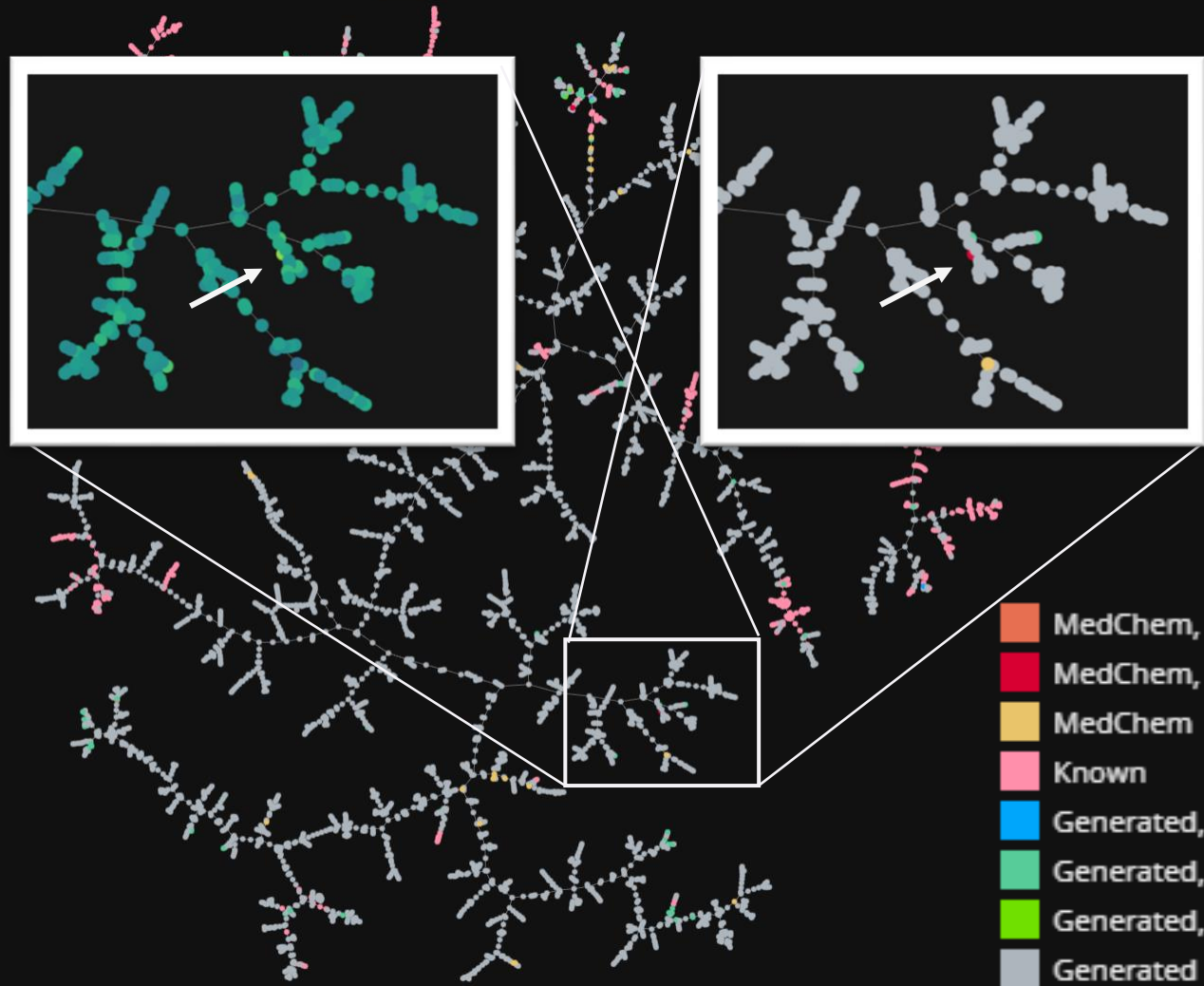


- MedChem, Synthesized
- MedChem, Best
- MedChem
- Known
- Generated, Synthesized
- Generated, Selected
- Generated, Best
- Generated



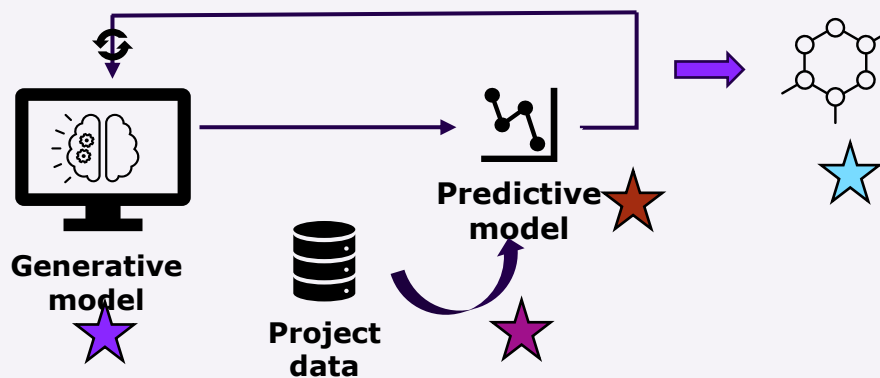


- MedChem, Synthesized
- MedChem, Best
- MedChem
- Known
- Generated, Synthesized
- Generated, Selected
- Generated, Best
- Generated



Using Chemical Language Models: Practical considerations on execution

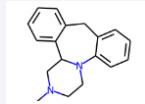
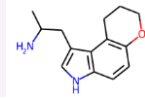
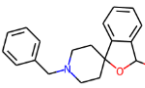
- Timing
 - The lead optimization GenAI window
 - Enough data for good models
 - Still useful to explore (vs exploit)
- "Data Science is teamwork"
 - Even with chemical lead opt data
 - Curate curate curate: Speed and quality



Using Chemical Language Models: Practical considerations on integration

- Define clear and shared objectives
- From screen to compound:
 - Quality & synthesisability
 - Or BB availability?
- From screen to data: cycle time
- Acculturate on statistics

- Explicit, transparent, shared design choices
- Generative approaches as part of a strategy
 - Explainable decision making

	pIC50	Caco2	logS	FUM	IEN																																																																											
	<table border="1"><thead><tr><th colspan="2"></th><th colspan="2">Prediction</th></tr><tr><th colspan="2"></th><th>0</th><th>1</th></tr></thead><tbody><tr><th rowspan="2">Actual</th><th>0</th><td>TN</td><td>FP</td></tr><tr><th>1</th><td>FN</td><td>TP</td></tr></tbody></table> ✓			Prediction				0	1	Actual	0	TN	FP	1	FN	TP	<table border="1"><thead><tr><th colspan="2"></th><th colspan="2">Prediction</th></tr><tr><th colspan="2"></th><th>0</th><th>1</th></tr></thead><tbody><tr><th rowspan="2">Actual</th><th>0</th><td>TN</td><td>FP</td></tr><tr><th>1</th><td>FN</td><td>TP</td></tr></tbody></table> ✓			Prediction				0	1	Actual	0	TN	FP	1	FN	TP	<table border="1"><thead><tr><th colspan="2"></th><th colspan="2">Prediction</th></tr><tr><th colspan="2"></th><th>0</th><th>1</th></tr></thead><tbody><tr><th rowspan="2">Actual</th><th>0</th><td>TN</td><td>FP</td></tr><tr><th>1</th><td>FN</td><td>TP</td></tr></tbody></table> ✗			Prediction				0	1	Actual	0	TN	FP	1	FN	TP	<table border="1"><thead><tr><th colspan="2"></th><th colspan="2">Prediction</th></tr><tr><th colspan="2"></th><th>0</th><th>1</th></tr></thead><tbody><tr><th rowspan="2">Actual</th><th>0</th><td>TN</td><td>FP</td></tr><tr><th>1</th><td>FN</td><td>TP</td></tr></tbody></table> ✓			Prediction				0	1	Actual	0	TN	FP	1	FN	TP	<table border="1"><thead><tr><th colspan="2"></th><th colspan="2">Prediction</th></tr><tr><th colspan="2"></th><th>0</th><th>1</th></tr></thead><tbody><tr><th rowspan="2">Actual</th><th>0</th><td>TN</td><td>FP</td></tr><tr><th>1</th><td>FN</td><td>TP</td></tr></tbody></table> ✓			Prediction				0	1	Actual	0	TN	FP	1	FN	TP
		Prediction																																																																														
		0	1																																																																													
Actual	0	TN	FP																																																																													
	1	FN	TP																																																																													
		Prediction																																																																														
		0	1																																																																													
Actual	0	TN	FP																																																																													
	1	FN	TP																																																																													
		Prediction																																																																														
		0	1																																																																													
Actual	0	TN	FP																																																																													
	1	FN	TP																																																																													
		Prediction																																																																														
		0	1																																																																													
Actual	0	TN	FP																																																																													
	1	FN	TP																																																																													
		Prediction																																																																														
		0	1																																																																													
Actual	0	TN	FP																																																																													
	1	FN	TP																																																																													
	<table border="1"><thead><tr><th colspan="2"></th><th colspan="2">Prediction</th></tr><tr><th colspan="2"></th><th>0</th><th>1</th></tr></thead><tbody><tr><th rowspan="2">Actual</th><th>0</th><td>TN</td><td>FP</td></tr><tr><th>1</th><td>FN</td><td>TP</td></tr></tbody></table> ✓			Prediction				0	1	Actual	0	TN	FP	1	FN	TP	<table border="1"><thead><tr><th colspan="2"></th><th colspan="2">Prediction</th></tr><tr><th colspan="2"></th><th>0</th><th>1</th></tr></thead><tbody><tr><th rowspan="2">Actual</th><th>0</th><td>TN</td><td>FP</td></tr><tr><th>1</th><td>FN</td><td>TP</td></tr></tbody></table> ✗			Prediction				0	1	Actual	0	TN	FP	1	FN	TP	<table border="1"><thead><tr><th colspan="2"></th><th colspan="2">Prediction</th></tr><tr><th colspan="2"></th><th>0</th><th>1</th></tr></thead><tbody><tr><th rowspan="2">Actual</th><th>0</th><td>TN</td><td>FP</td></tr><tr><th>1</th><td>FN</td><td>TP</td></tr></tbody></table> ✓			Prediction				0	1	Actual	0	TN	FP	1	FN	TP	<table border="1"><thead><tr><th colspan="2"></th><th colspan="2">Prediction</th></tr><tr><th colspan="2"></th><th>0</th><th>1</th></tr></thead><tbody><tr><th rowspan="2">Actual</th><th>0</th><td>TN</td><td>FP</td></tr><tr><th>1</th><td>FN</td><td>TP</td></tr></tbody></table> ✗			Prediction				0	1	Actual	0	TN	FP	1	FN	TP	<table border="1"><thead><tr><th colspan="2"></th><th colspan="2">Prediction</th></tr><tr><th colspan="2"></th><th>0</th><th>1</th></tr></thead><tbody><tr><th rowspan="2">Actual</th><th>0</th><td>TN</td><td>FP</td></tr><tr><th>1</th><td>FN</td><td>TP</td></tr></tbody></table> ✓			Prediction				0	1	Actual	0	TN	FP	1	FN	TP
		Prediction																																																																														
		0	1																																																																													
Actual	0	TN	FP																																																																													
	1	FN	TP																																																																													
		Prediction																																																																														
		0	1																																																																													
Actual	0	TN	FP																																																																													
	1	FN	TP																																																																													
		Prediction																																																																														
		0	1																																																																													
Actual	0	TN	FP																																																																													
	1	FN	TP																																																																													
		Prediction																																																																														
		0	1																																																																													
Actual	0	TN	FP																																																																													
	1	FN	TP																																																																													
		Prediction																																																																														
		0	1																																																																													
Actual	0	TN	FP																																																																													
	1	FN	TP																																																																													
	<table border="1"><thead><tr><th colspan="2"></th><th colspan="2">Prediction</th></tr><tr><th colspan="2"></th><th>0</th><th>1</th></tr></thead><tbody><tr><th rowspan="2">Actual</th><th>0</th><td>TN</td><td>FP</td></tr><tr><th>1</th><td>FN</td><td>TP</td></tr></tbody></table> ✓			Prediction				0	1	Actual	0	TN	FP	1	FN	TP	<table border="1"><thead><tr><th colspan="2"></th><th colspan="2">Prediction</th></tr><tr><th colspan="2"></th><th>0</th><th>1</th></tr></thead><tbody><tr><th rowspan="2">Actual</th><th>0</th><td>TN</td><td>FP</td></tr><tr><th>1</th><td>FN</td><td>TP</td></tr></tbody></table> ✓			Prediction				0	1	Actual	0	TN	FP	1	FN	TP	<table border="1"><thead><tr><th colspan="2"></th><th colspan="2">Prediction</th></tr><tr><th colspan="2"></th><th>0</th><th>1</th></tr></thead><tbody><tr><th rowspan="2">Actual</th><th>0</th><td>TN</td><td>FP</td></tr><tr><th>1</th><td>FN</td><td>TP</td></tr></tbody></table> ✓			Prediction				0	1	Actual	0	TN	FP	1	FN	TP	<table border="1"><thead><tr><th colspan="2"></th><th colspan="2">Prediction</th></tr><tr><th colspan="2"></th><th>0</th><th>1</th></tr></thead><tbody><tr><th rowspan="2">Actual</th><th>0</th><td>TN</td><td>FP</td></tr><tr><th>1</th><td>FN</td><td>TP</td></tr></tbody></table> ✓			Prediction				0	1	Actual	0	TN	FP	1	FN	TP	<table border="1"><thead><tr><th colspan="2"></th><th colspan="2">Prediction</th></tr><tr><th colspan="2"></th><th>0</th><th>1</th></tr></thead><tbody><tr><th rowspan="2">Actual</th><th>0</th><td>TN</td><td>FP</td></tr><tr><th>1</th><td>FN</td><td>TP</td></tr></tbody></table> ✓			Prediction				0	1	Actual	0	TN	FP	1	FN	TP
		Prediction																																																																														
		0	1																																																																													
Actual	0	TN	FP																																																																													
	1	FN	TP																																																																													
		Prediction																																																																														
		0	1																																																																													
Actual	0	TN	FP																																																																													
	1	FN	TP																																																																													
		Prediction																																																																														
		0	1																																																																													
Actual	0	TN	FP																																																																													
	1	FN	TP																																																																													
		Prediction																																																																														
		0	1																																																																													
Actual	0	TN	FP																																																																													
	1	FN	TP																																																																													
		Prediction																																																																														
		0	1																																																																													
Actual	0	TN	FP																																																																													
	1	FN	TP																																																																													

5 binary models,
success rate 0.8

$$0,8^5 = 0.33$$

•

Thank *you!*

•

sanofi