Scaling *De novo* Design of Bispecific Ligands

Jérémy Besnard

UK-QSAR & Chemoinformatics Group
19 October 2016

exscientia

exscientia

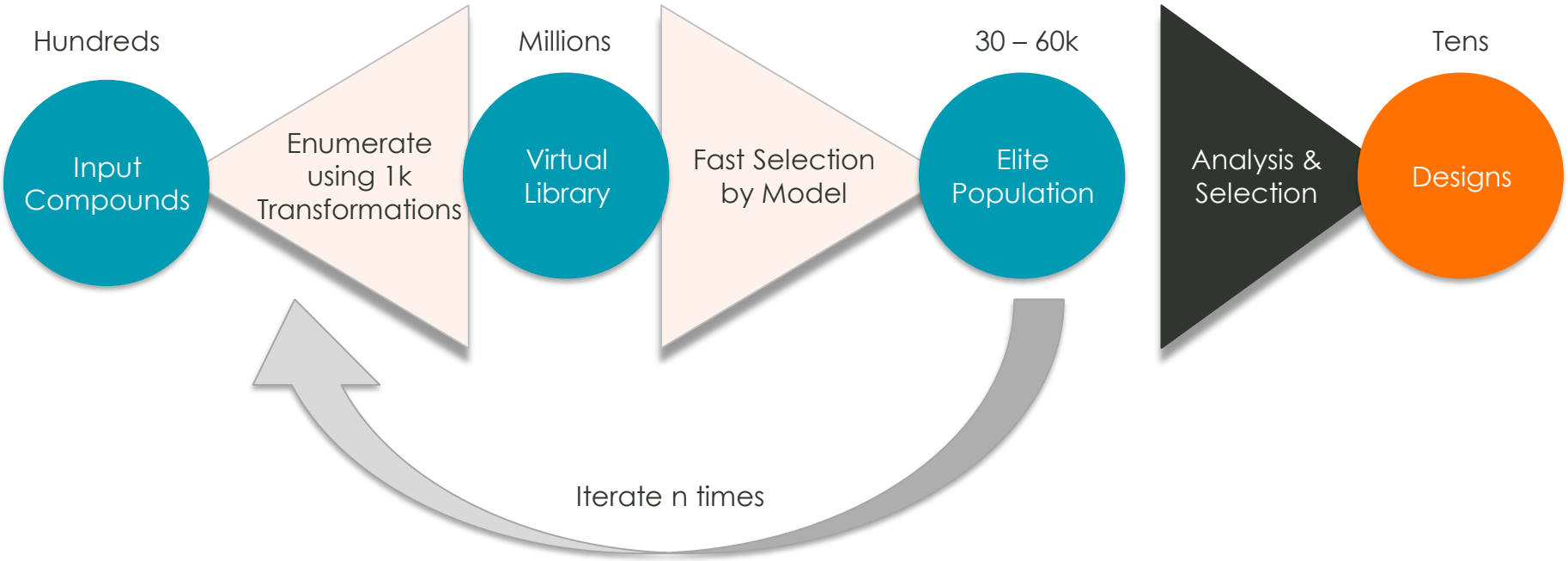
- We are a young spin-out company from the University of Dundee (2012)
- Employees working in multiple locations
- Platform built from revenues through partnerships with pharmaceutical companies – organic yet exponential growth
- Medicinal chemistry design as a routine process

e^x scientia

Scaling *de novo* design

1. Validate on single target
   - Show that the method can generate patentable compounds
2. Expand to dual combinations
   - Generate clinical candidate
3. Expand again to disease portfolio
   - Design against a matrix of disease-related targets

GA *de novo* design algorithm



Besnard *et al.* Automated design of ligands to polypharmacological profiles, Nature 492, 215–220
http://doi.org/10.1038/nature11691

exscientia

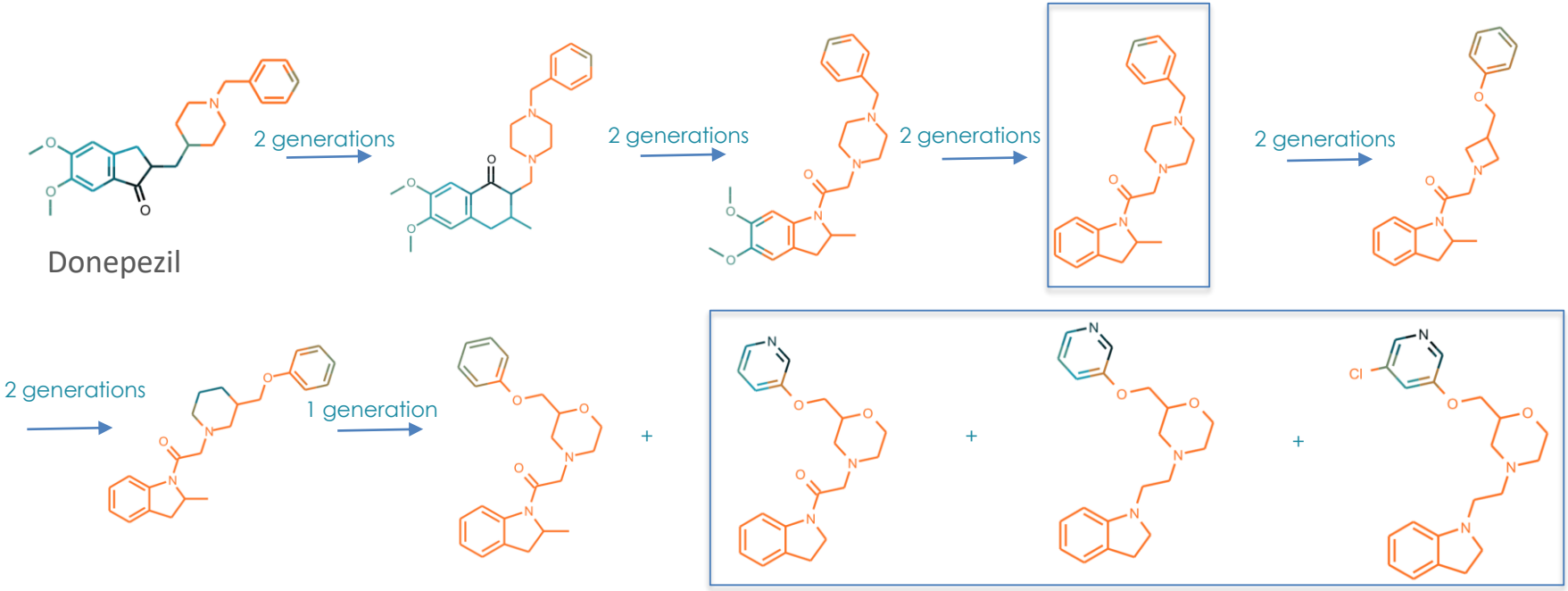
1. Validating the method

Besnard *et al.* Automated design of ligands to polypharmacological profiles, Nature 492, 215–220
http://doi.org/10.1038/nature11691

exscientia

Optimise D4 activity and selectivity

<30 compounds required to discover, synthesize and patent compound 27s, a selective D4 compound with early lead properties



Compound 27s
D4 K_i = 90 nM, selective
Patent: PCT/GB2012/051194 / WO2012160392

> O(6) compounds evolved & scored for D4 and off-targets, but only most promising compounds were synthesized and screened

exscientia

Metabolic Disorder Example

Dual inhibitor for two unrelated targets

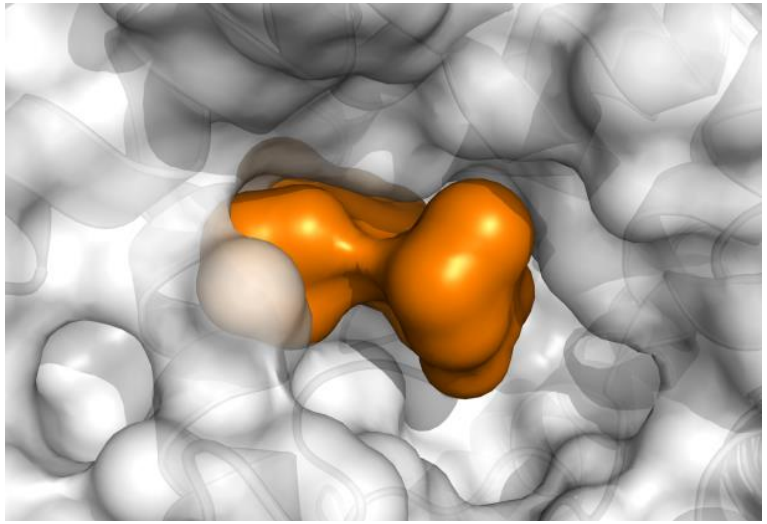
- Design against polypharmacology profiles
- Confirmatory X-ray structures of both complexes

Bispecific Compounds

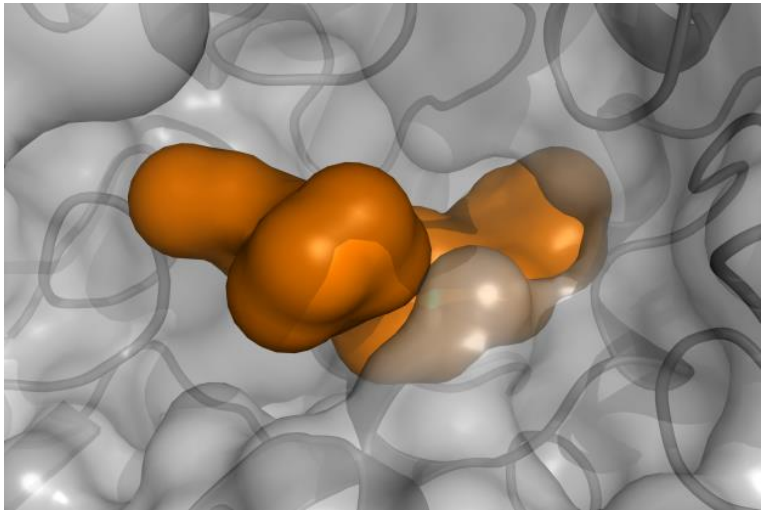
- Goal is to find first-in-class bispecific small molecule inhibiting two enzymes of unrelated families
- Process:
  - Gather public and patent data to build models
  - *De novo* design with evolutionary algorithm
  - Docking of top ranked compounds to assess if the compounds could fit the two binding sites
  - Synthesize top-ranking
  - In-vitro assay followed by crystallography

Structural validation

Protein-ligand **X-ray structures** of both complexes with the top prioritized compound



Enzyme A
IC_{50}= 350 nM



Enzyme B
IC_{50}= 10 nM

exscientia

2. Delivering a candidate

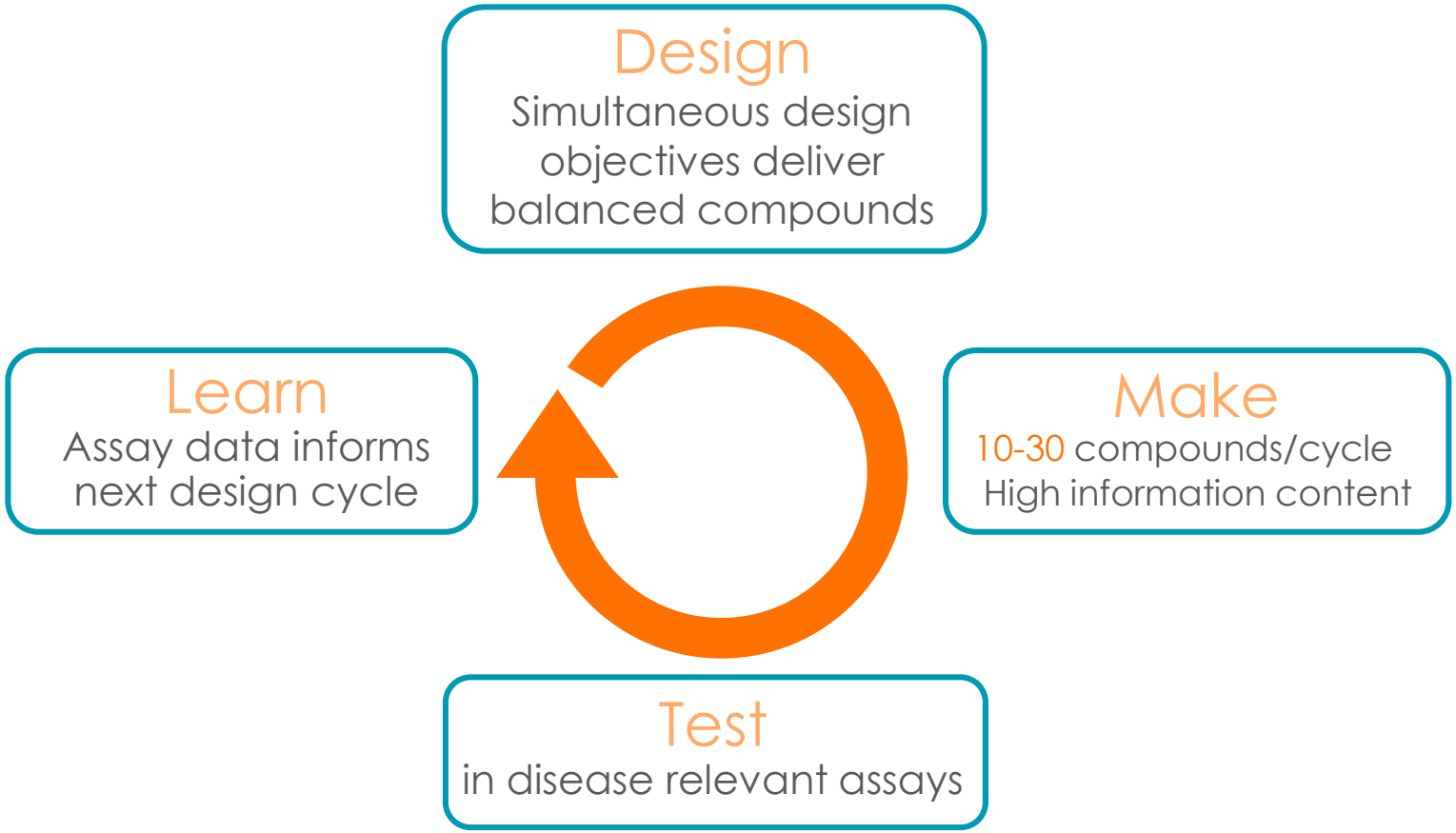
Dual agonist for two distinct GPCRs

- Collaboration with Sumitomo Dainippon Pharma
- Design against polypharmacology profile
- *in vitro* assessment
- Rapid delivery of candidate to *in vivo* safety study

exscientia

Technology in practice

Automated lead generation with rapid design cycle and efficient evolution to drug candidate profile



exscientia

Lead Identification

Design, synthesis, assay: 5-15 compounds per 2-week cycle

Design of 5 chemotypes	Compounds synthesized	Ease of synthesis	Dual agonist activity❖	Best Affinity	GPCR selectivity*	
	25	✓	✓	80 nM 70 nM	✓	 to lead optimization
	5	X	X		X	
	30	✓	X		X	
	45	✓	✓	70 nM 100 nM	✓	 to lead optimization
	5	X	X		X	

❖ multiple compounds <150 nM at both targets.

* <50% activity at 1μM over 20 GPCR receptors

exscientia

Lead Optimization

80 further compounds for each prioritized scaffold

	Dual agonist	<20nM target 1	<20nM target 2	
▲	✓	✓	✗	scaffold designated as backup
●	✓	✓	✓	scaffold prioritized additional assays progressed further compounds made on this scaffold

Candidate Seeking

40 compounds for prioritized scaffold

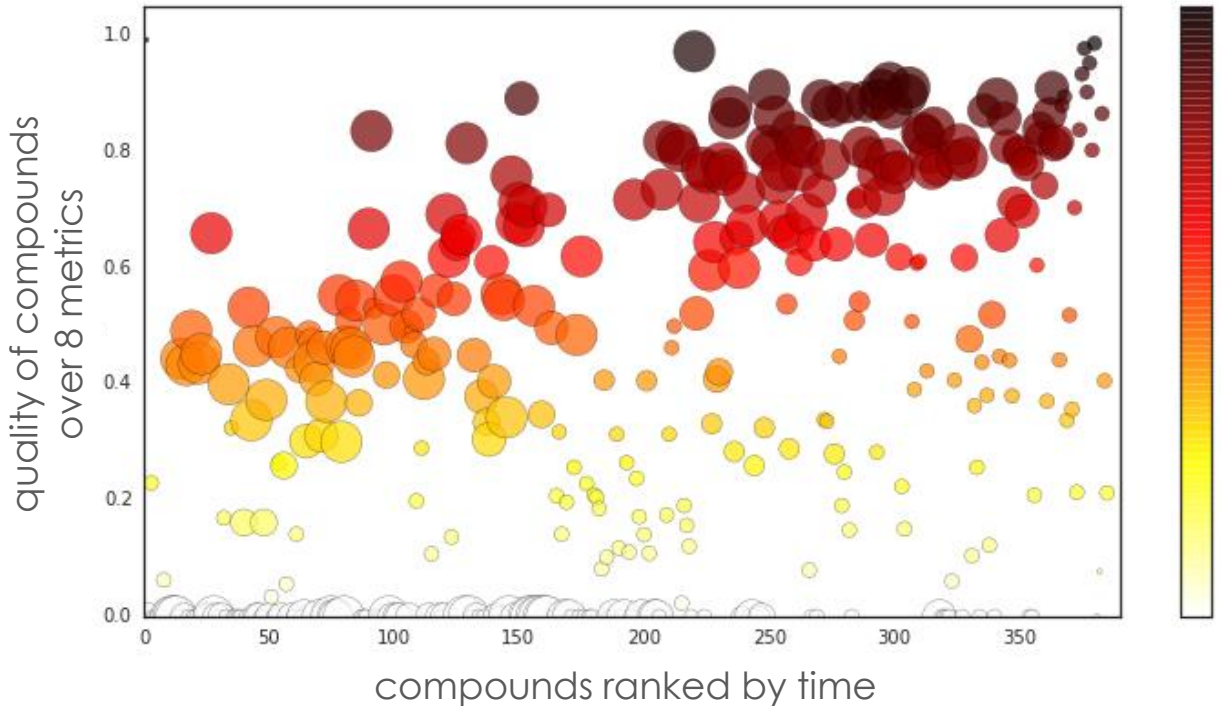
	Dual agonist	<20nM target 1	<20nM target 2	solubility	HERG >10µM	GPCR selectivity	DMPK
●	✓	✓	✓	✓	✓	✓	✓

exscientia

Towards candidate nomination

Successful bispecific project for CNS disease

- 2 chemotypes progressing to candidate selection (ongoing)
- <400 compounds synthesized and assayed
- 12 month project



circle size
number of assays performed for each compound

circle colour
compound quality/proximity to objectives

exscientia

3. Scaling to disease level

From one pair to a matrix of pairs

Big Data

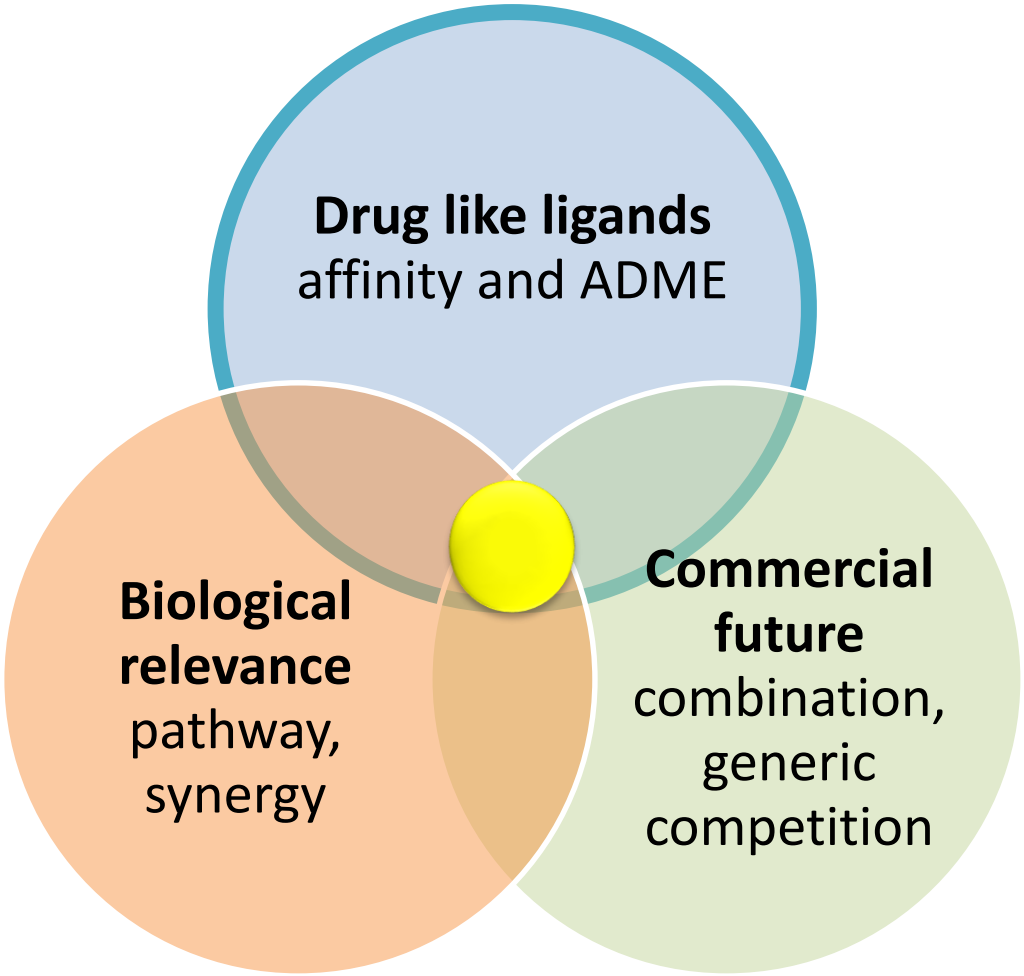
- Big is relative, big compared to what?
- Philosophy is more important than size
- Collect and use all data rather than a small sample
- Accept messiness of data - using more data of variable quality outweighs using small, very exact data
- All models are wrong but some are useful (link)

exscientia

Design against a disease

- Find targets and select potential pairs
- Generate designs for all pairs
- Analyse the results

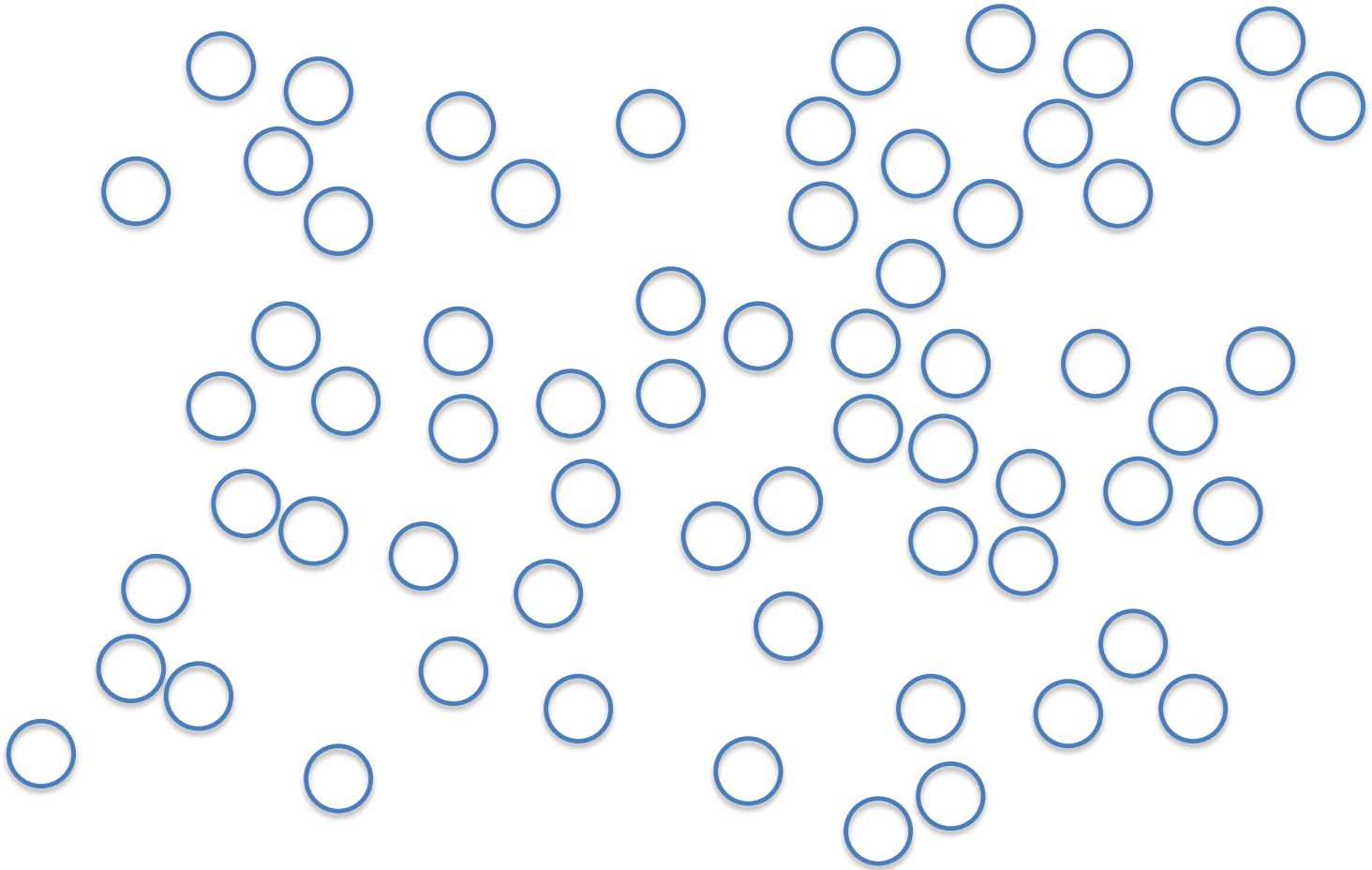
Selecting target combinations



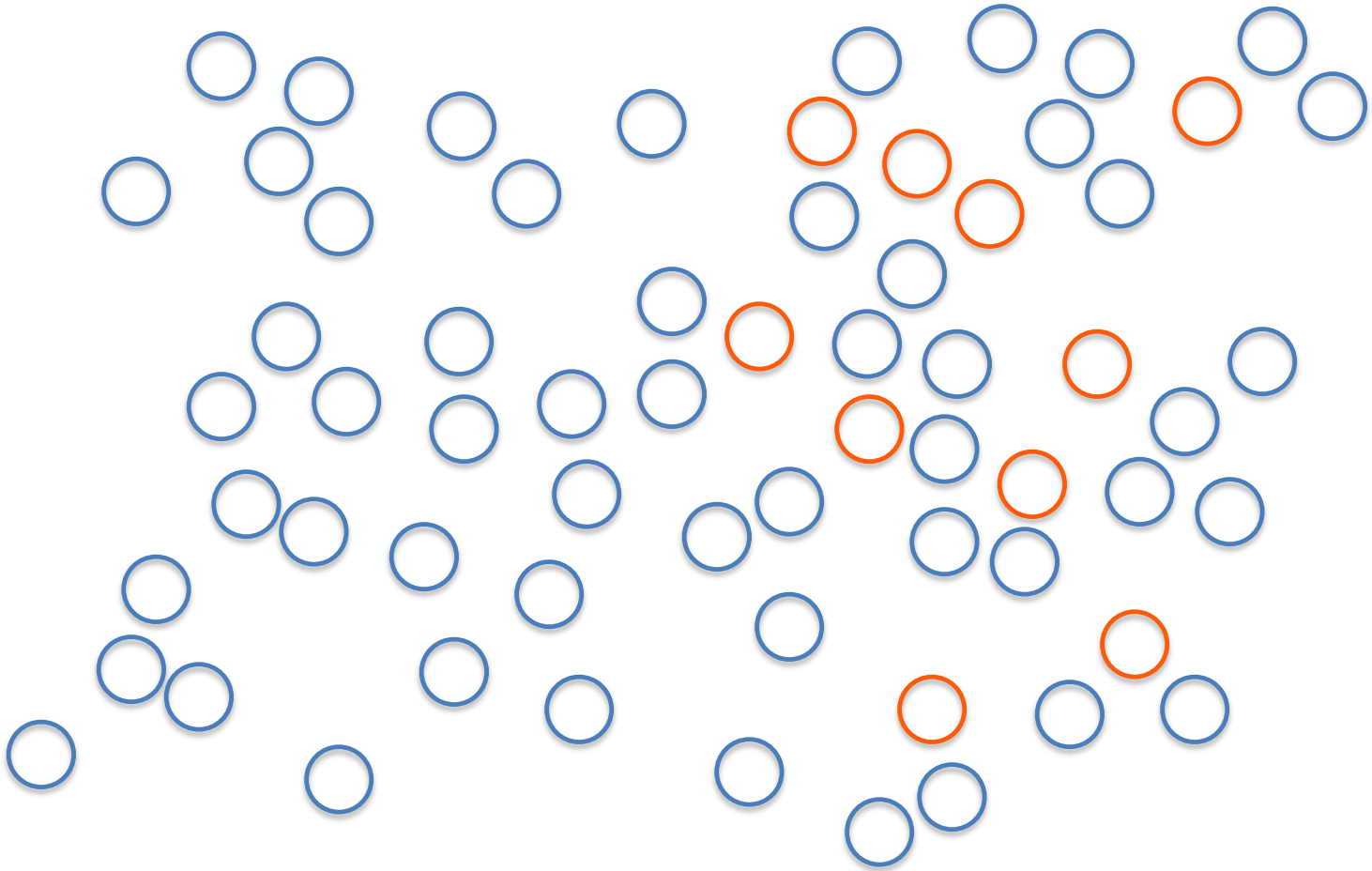
Commonality between two targets

- Disease link
  - Text mining for co-occurrence of target-disease and target-target
- Chemical compatibility
  - Known compounds that bind both targets
    - Take into account selectivity, "easy" to hit everything
  - Halfway compounds: known active against one, predicted active against the other target
  - Similar chemical property space of known actives

Druggable genome

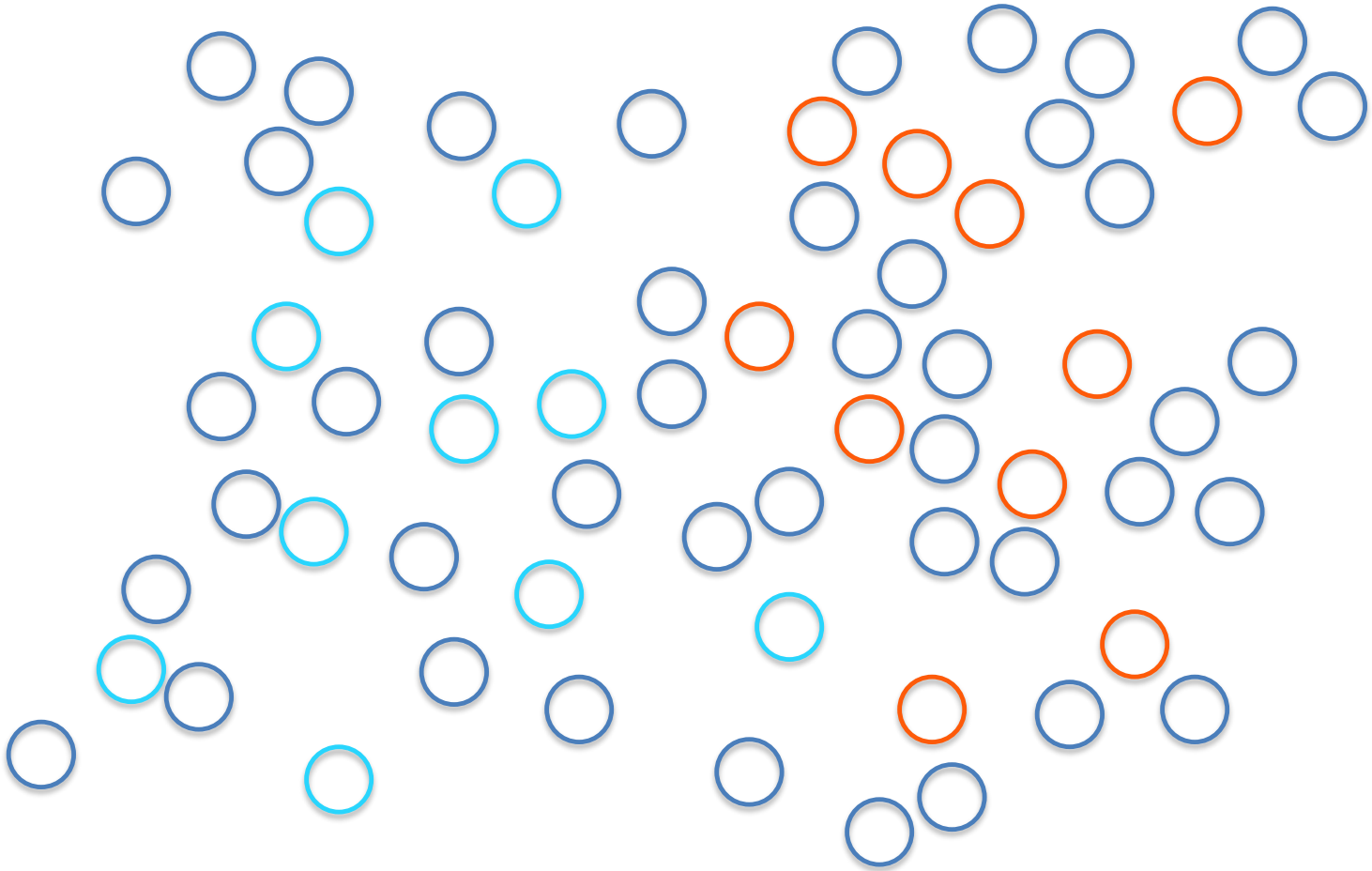


Clinically validated

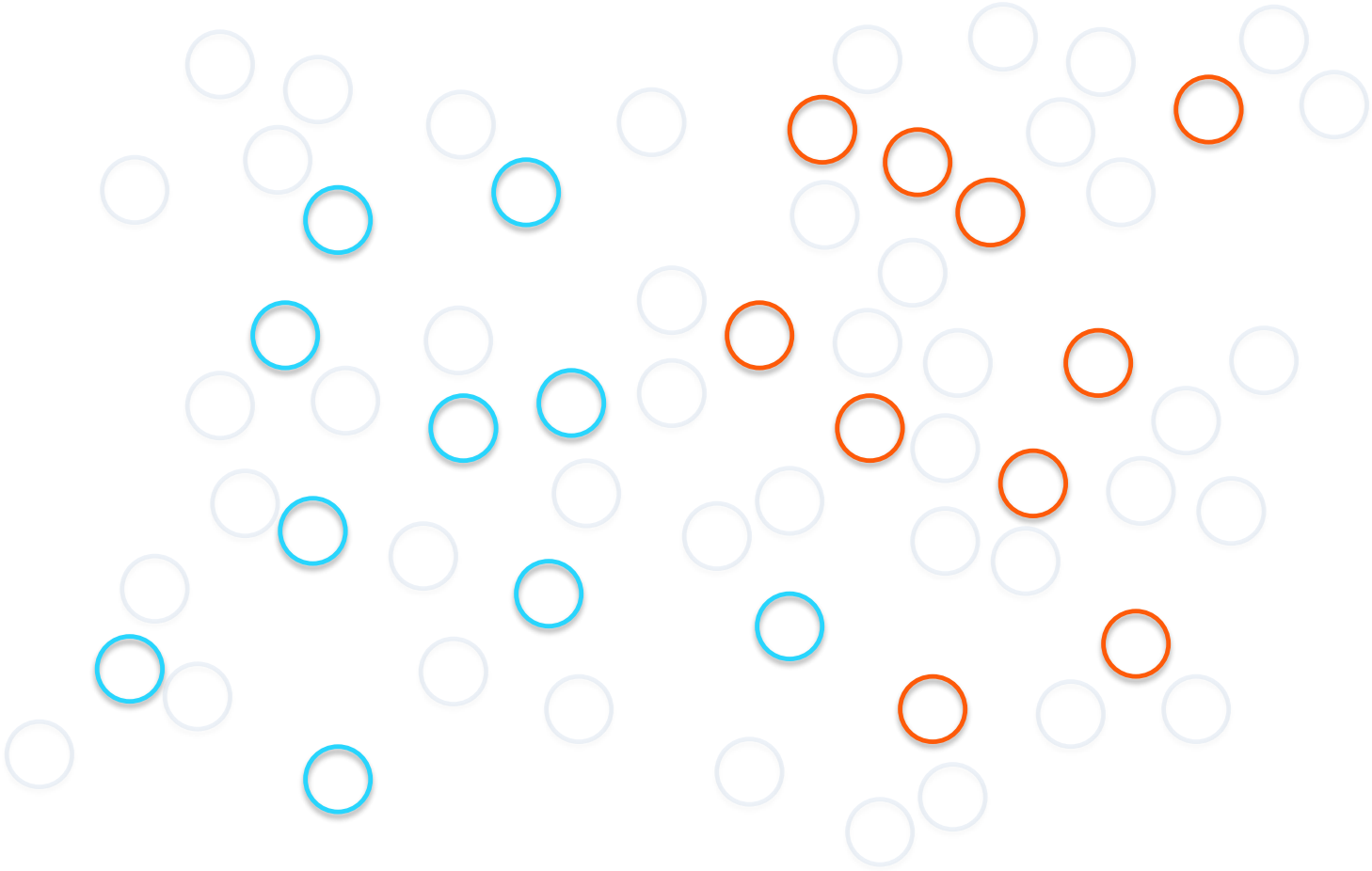


exscientia

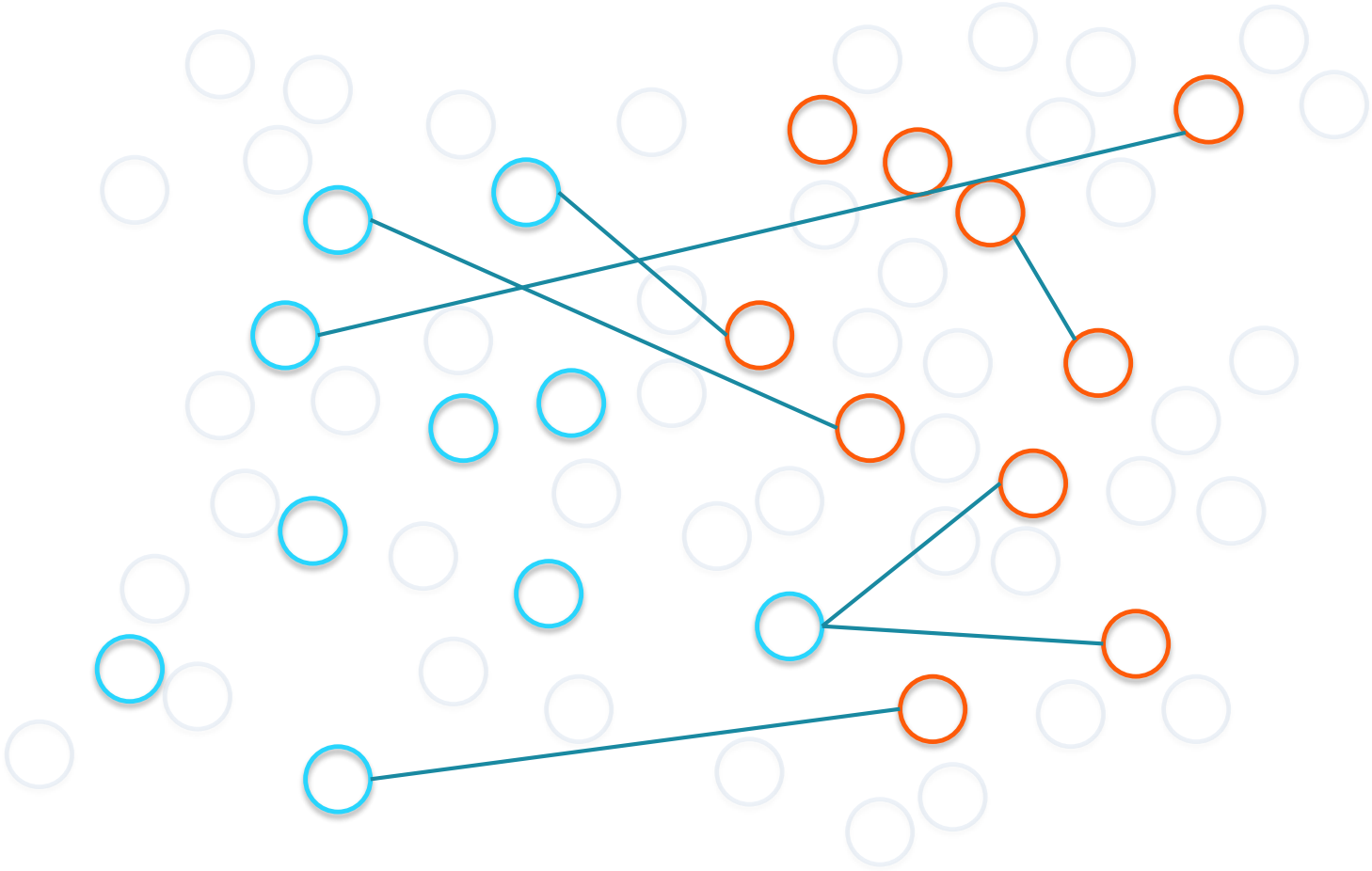
Emerging biology



Target pool

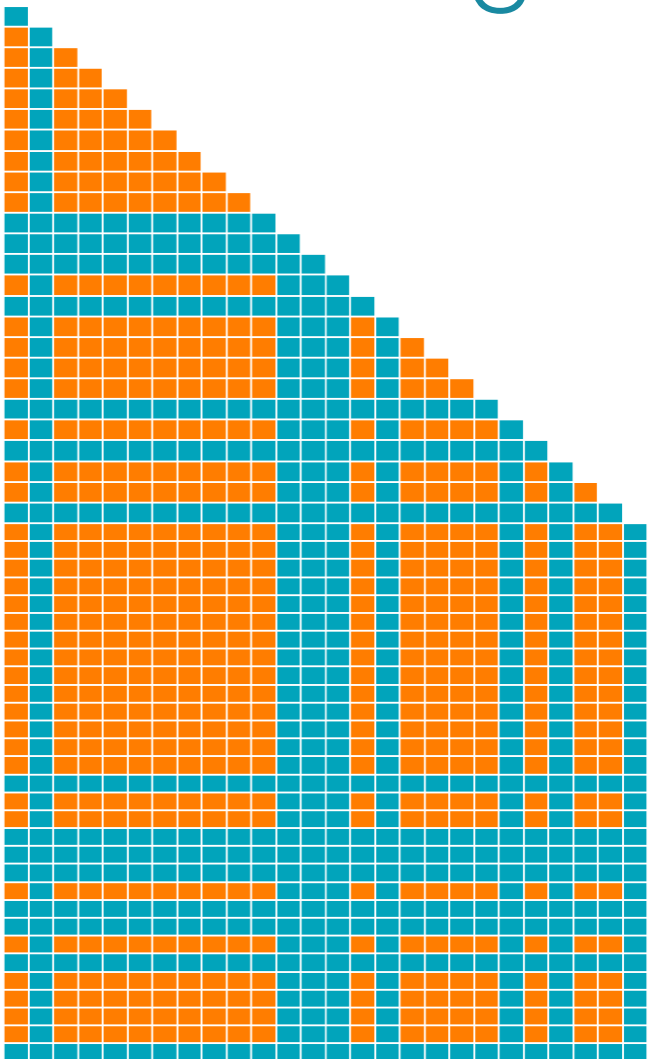


Chemical compatibility



exscientia

Design against a disease



- Matrix of targets linked to disease
- Automated design for all pairs
- ~1200 designs
- Scale: all results within ~ 2 days
- < 1% technical failures
- 100% result capture

exscientia

Calculation details

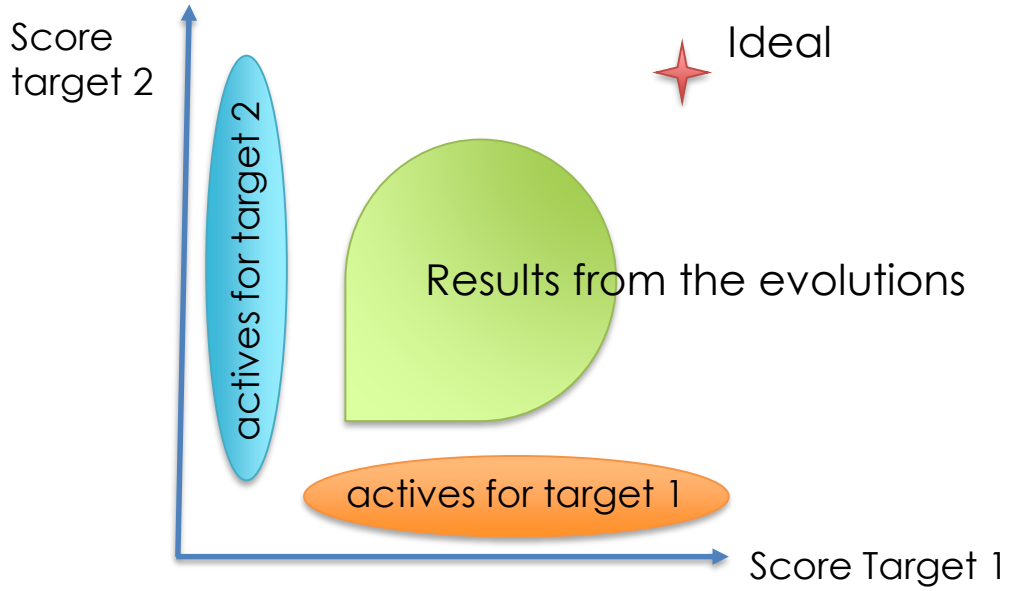
Target matrix pairs	1235
No model for target	60
No inputs pass filters	1
No evolution	26
Technical failures	0

- 1150 successful runs
- 13 billion enumerated compounds
- 60 million compound designs captured
- A weekend on AWS

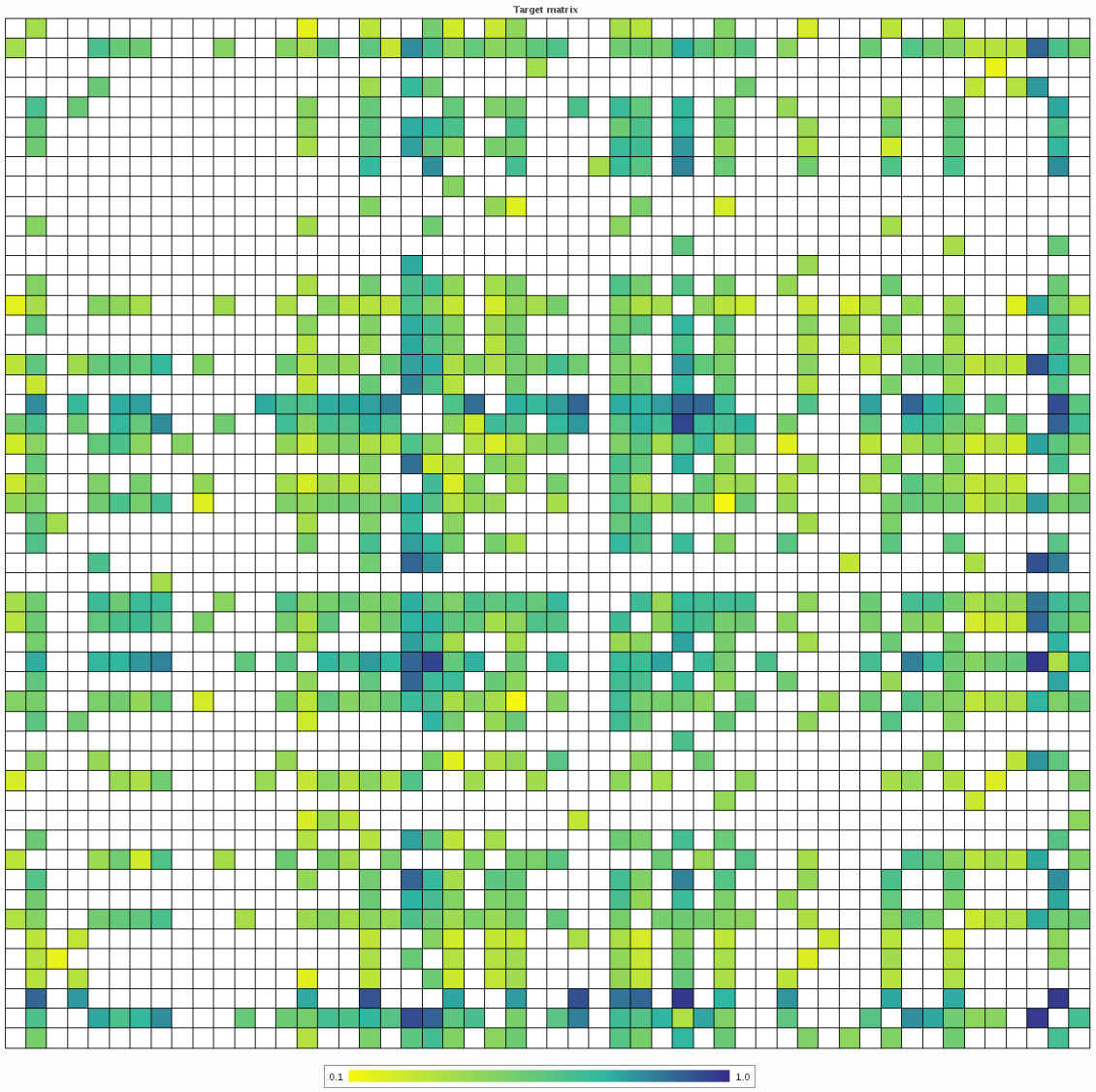
exscientia

What we are looking for per pair

- Combinations with good model scores for both targets and drug-like molecules



“Distance to ideal” per target pair



Ideal pair:

- Score Target 1 = 1
- Score Target 2 = 1

Want: small distance (yellow)

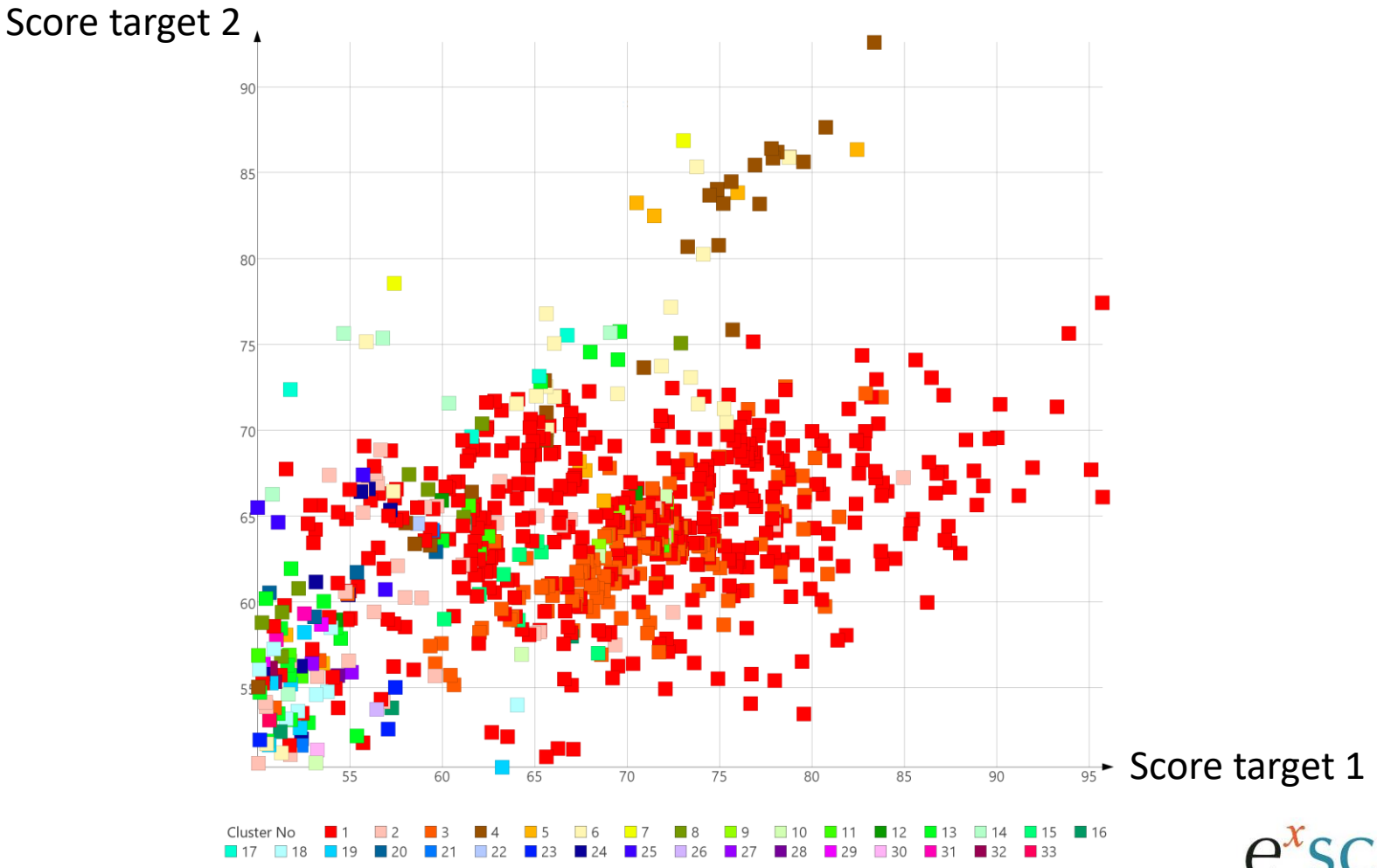
Enough to chose from!

White = no result (not run or no desirable compound)

exscientia

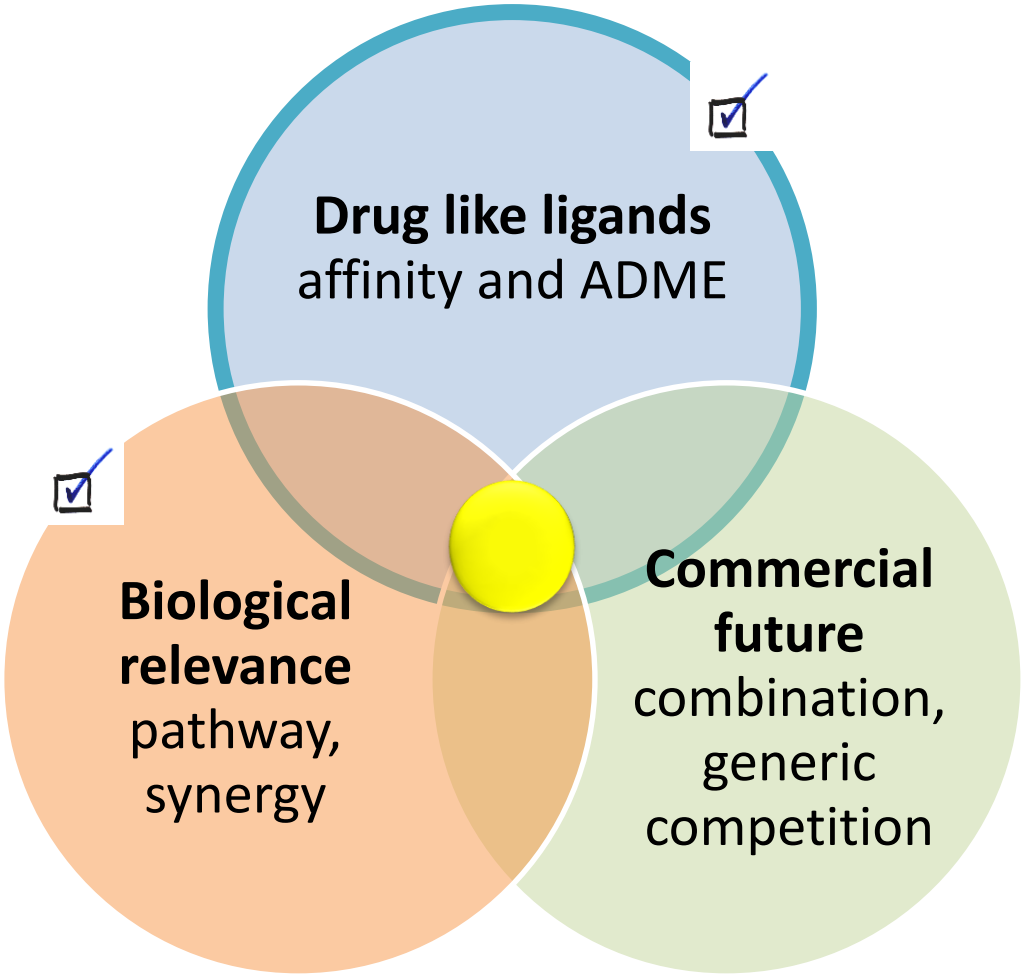
Multiple series for a target pair

Data clustered and visualised in DataWarrior



exscientia

Selecting target combinations



Scaling *de novo* design

1. Validate on single target
   - Show that the method can generate patentable compounds
2. Expand to dual combinations
   - Generate clinical candidate
3. Expand again to disease portfolio
   - Design against a matrix of disease-related targets

exscientia team

CEO
Andrew Hopkins
Founder

Chief Chemist
Andy Bell
2 FDA drugs:
sildenafil (Viagra)
voriconazole
(Vfend).

CIO
Richard Bickerton
Co-founder
.

CTO
Jérémy Besnard
Co-founder

Chemo-informatics
Willem van Hoorn

Molecular Informatics
Adrian Schreyer

COO
Mark Swindells

Systems Architect
Richard Cox

Executive Officer
Senga Oxenham

THANK YOU