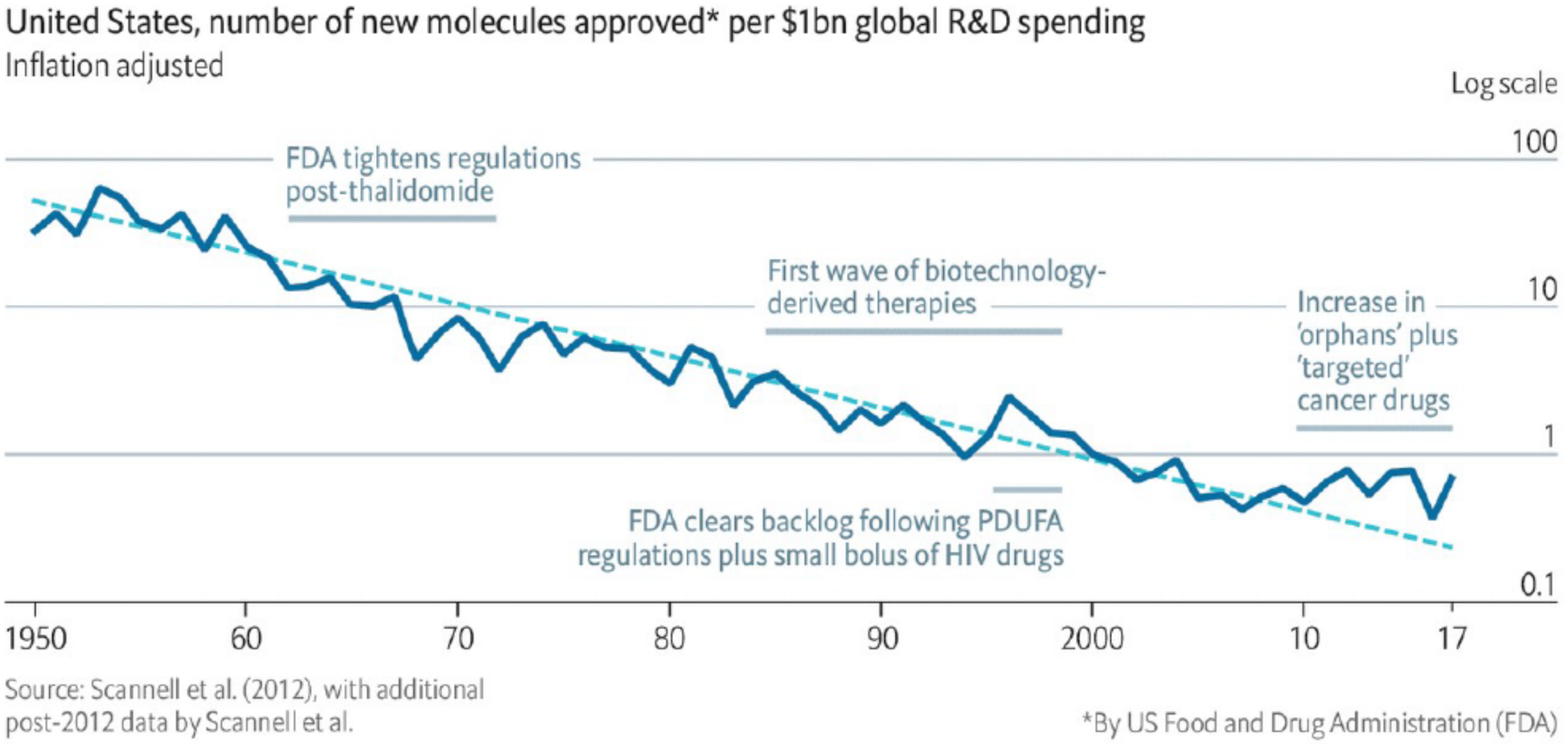
anagenex®

Platform Overview for Drug Discovery Workflows Symposium

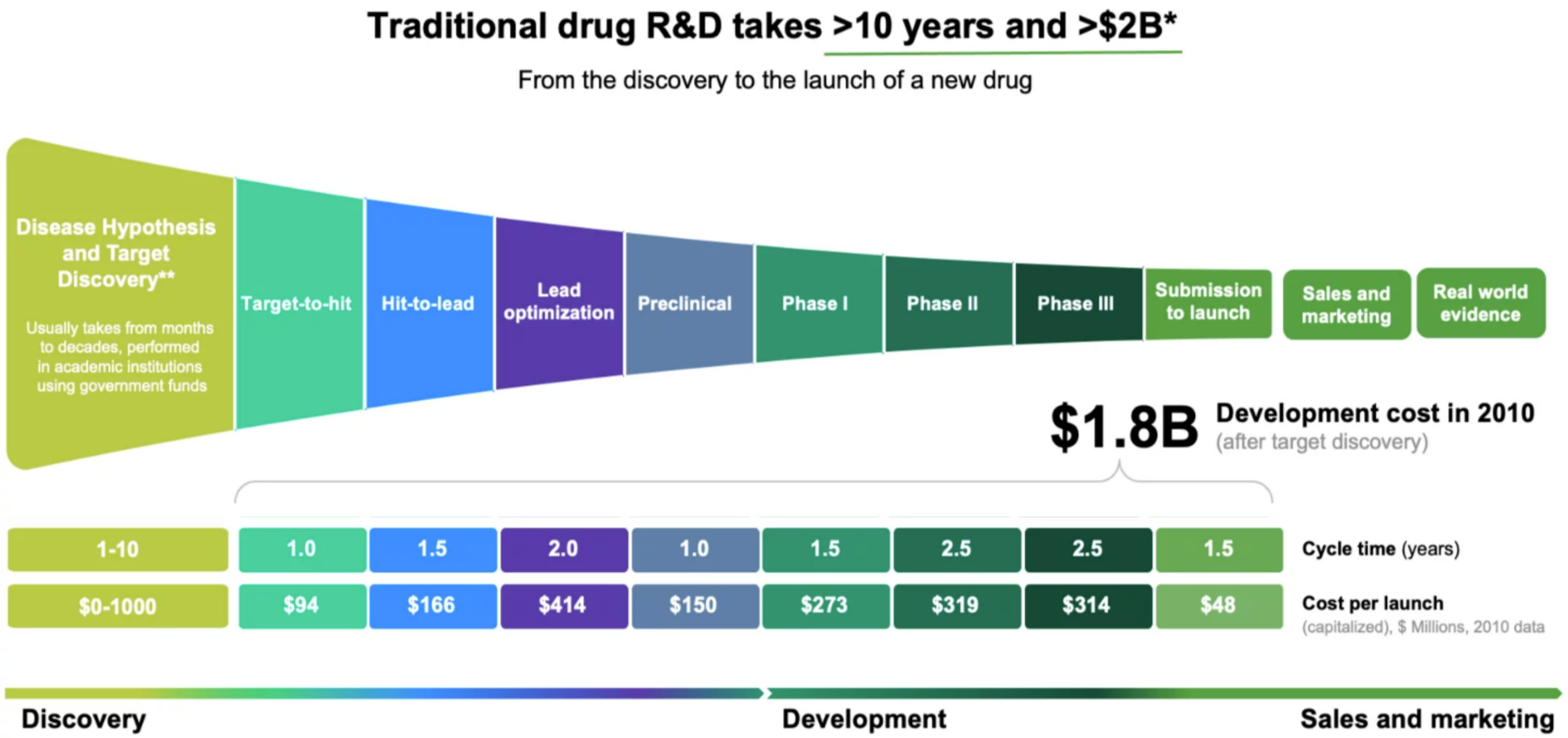
February 10, 2023

Current Drug Discovery Trends are Unsustainable



anagenex

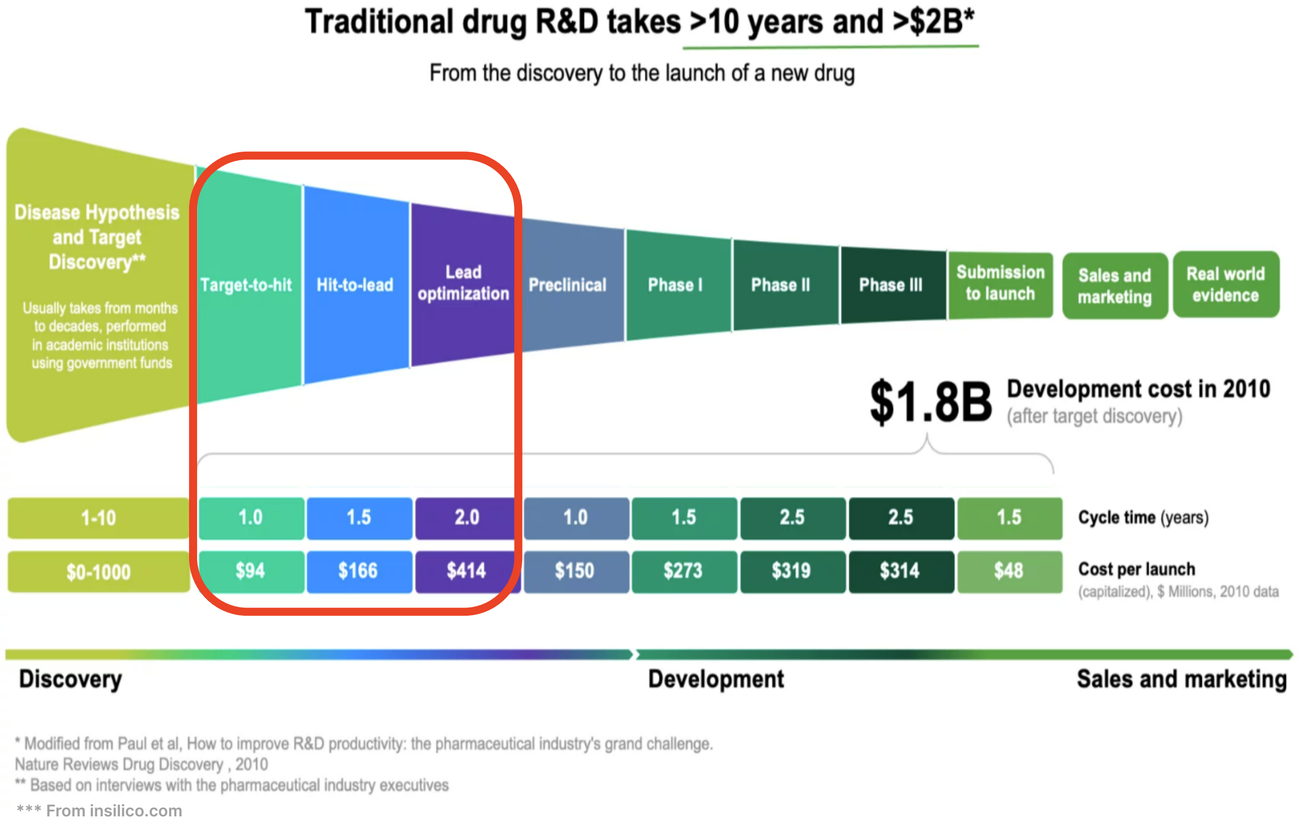
Current Drug Discovery Trends are Unsustainable



* Modified from Paul et al, How to improve R&D productivity: the pharmaceutical industry's grand challenge. Nature Reviews Drug Discovery , 2010
** Based on interviews with the pharmaceutical industry executives
*** From insilico.com

anagenex

How Can We Improve Efficiency?



* Modified from Paul et al, How to improve R&D productivity: the pharmaceutical industry's grand challenge. Nature Reviews Drug Discovery , 2010
** Based on interviews with the pharmaceutical industry executives
*** From insilico.com

Opportunities for improvement throughout the process:

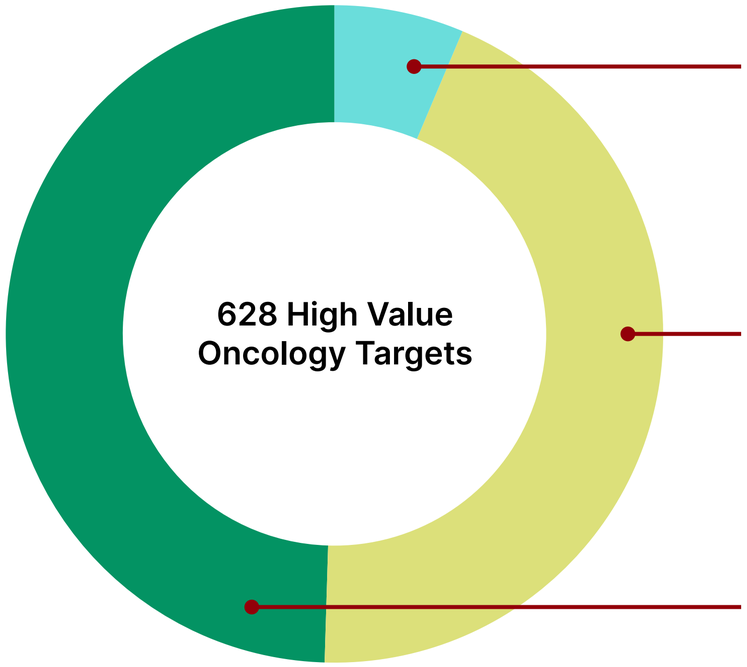
- Better Target Validation
- Faster Hit Identification
- Drugging “Undruggable” Targets
- Better clinical biomarkers

How can we:

- Identify chemical starting points for difficult targets?
- Speed up the process of Hit ID and Lead Optimization?

anagenex

Problem: Existing Technologies Miss Valuable Targets



Behan et al, Nature, 2019

Why can’t we drug these high value targets?

1. Novel high value targets need new chemical matter

- Screening collections biased by what used to work
- Pure computational methods biased by static structures and incomplete physical models

Misses:

Cryptic pockets | **Protein protein interactions** | **Nucleic acid binders**

2. Hit & lead optimization is slow & low throughput

- Synthesizing 10 compounds per chemist per month delays in-vitro and in-vivo compound validation

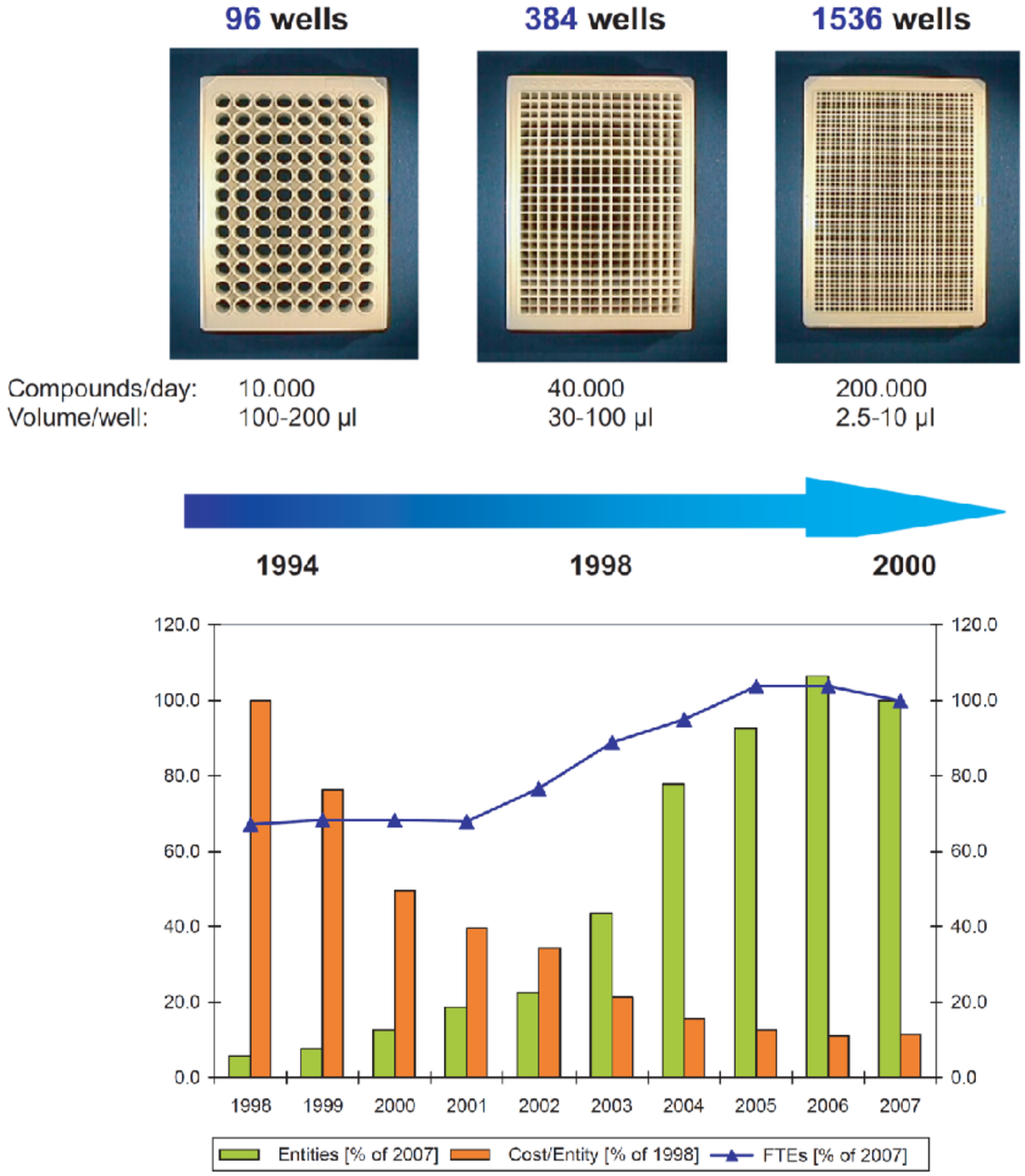
anagenex

Small Molecule Discovery Tools

High Throughput Screening and DNA Encoded Libraries

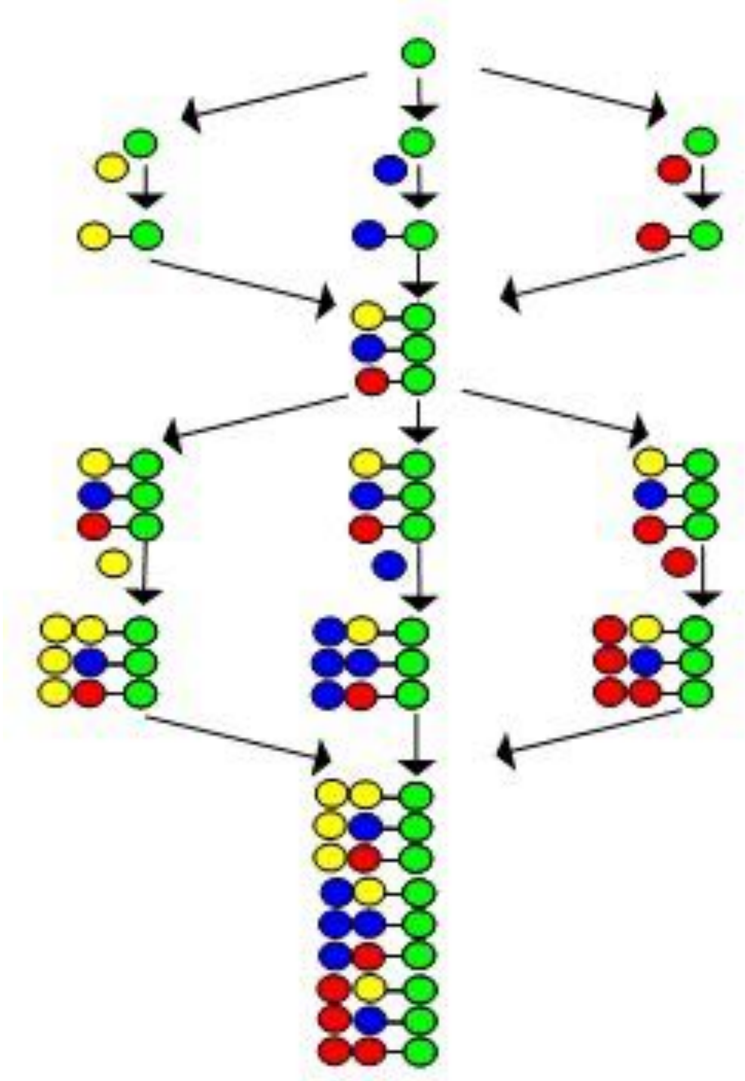
High Throughput Screening

- Hit identification in Big Pharma has been driven through High Throughput Screening over the past several decades
- Increased miniaturization has led to:
  - Screening ~1-2 Million highly purified, discreet compounds
  - 1536 well plates, ~ 1 week to screen
- Screening gains have plateaued:
  - Months to optimize assay conditions
  - ~$1 Million per screen
  - Compound collections comprised of chemical matter from legacy drug discovery efforts
  - Key limitation: 1 compound per well



anagenex

How Can We Test More?

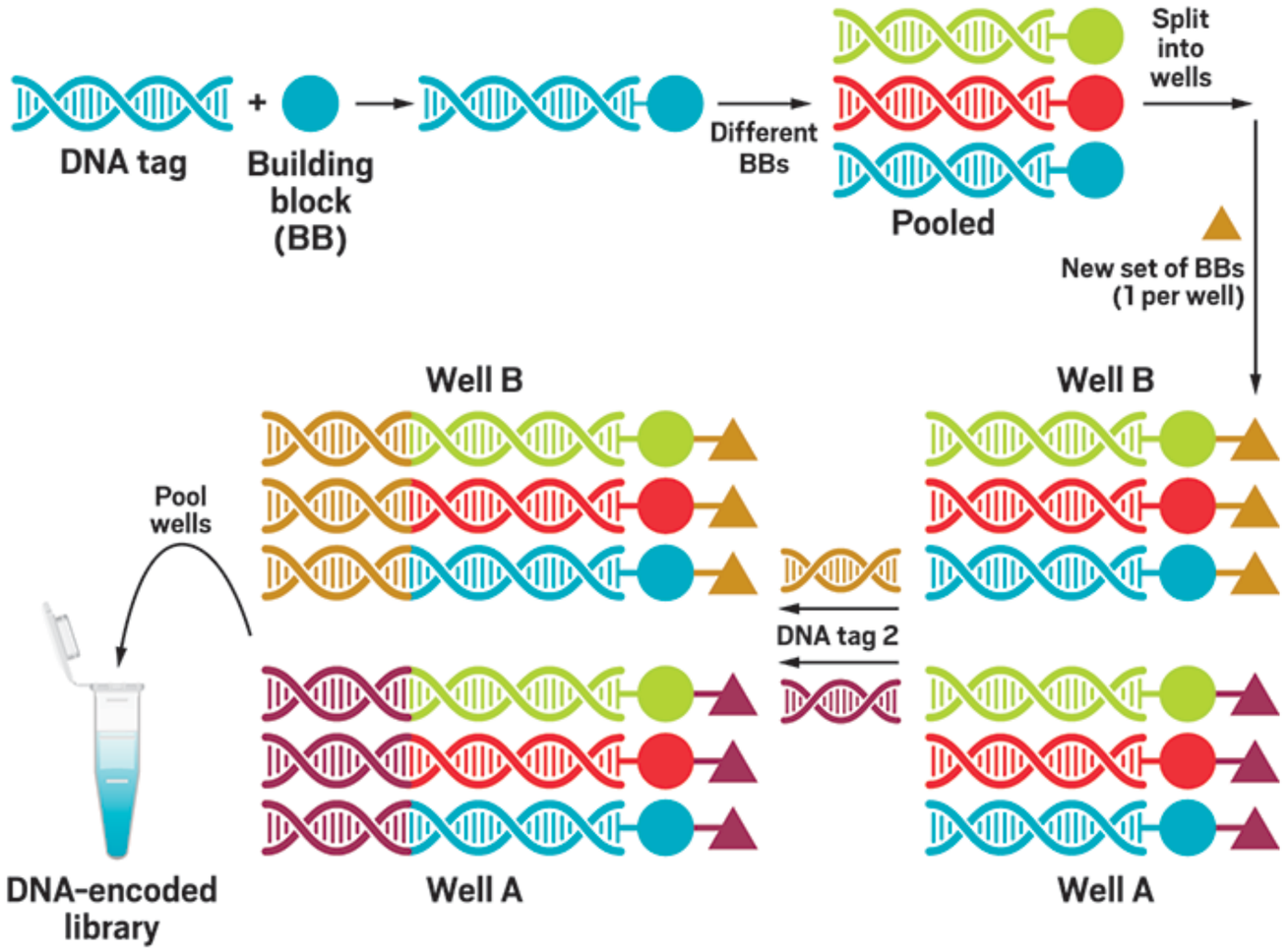


Combinatorial Chemistry

- “Split and Pool” method developed by Bruce Merrifield in the 1960s
- In just a few cycles of chemistry, allows the synthesis of Millions to Billions of compounds
- Molecules exist as a mixture and each molecule represented in tiny quantities
- How to screen a mixture of a Billion compounds and know which one inhibited your target?????

Advent of DNA Encoded Libraries (DEL)

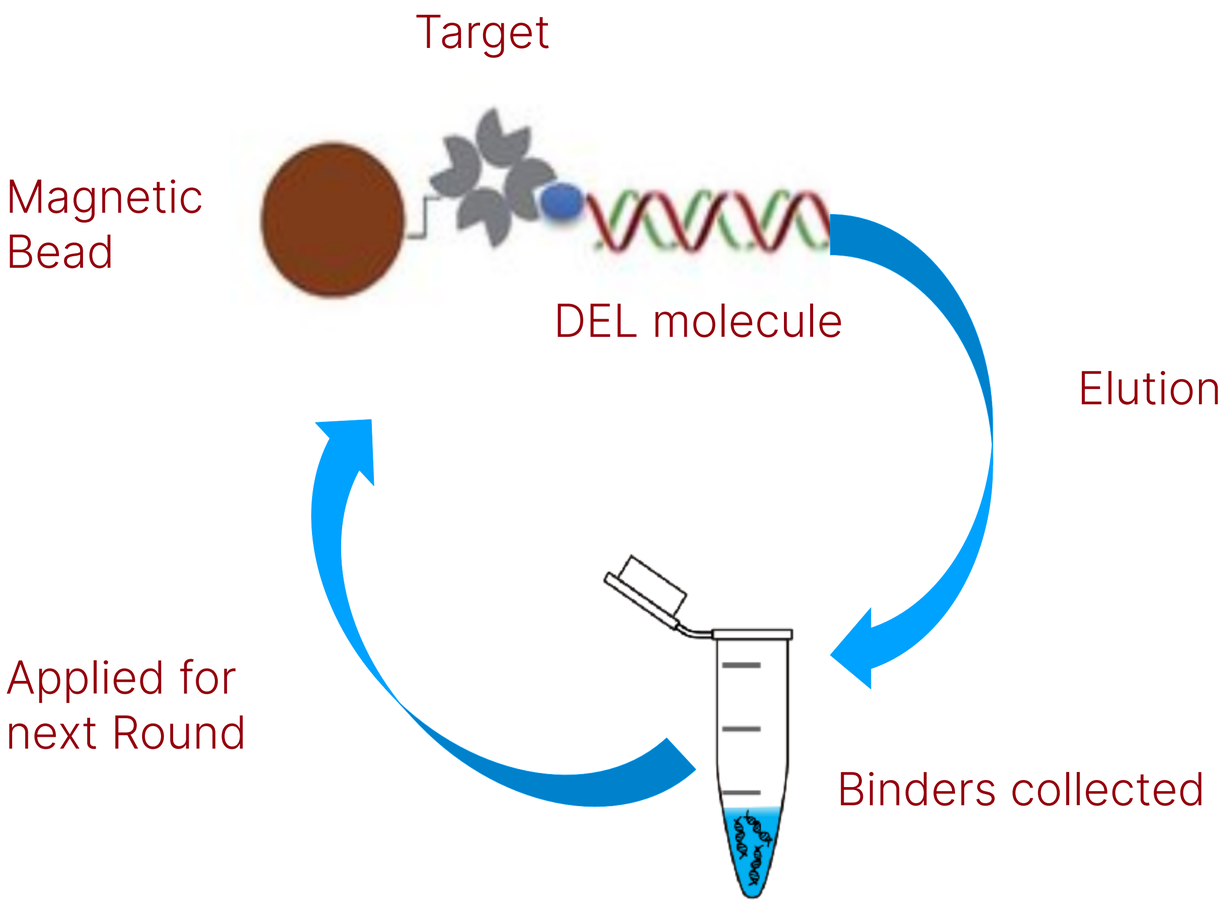
- Combinatorial chemistry in two directions
- DNA tag added at each step that encodes for the building block added
- Through use of PCR and Next gen DNA sequencing, we can now identify each compound from the mixture



HTTPS://CEN.ACS.ORG/ARTICLES/95/I25/DNA-ENCODED-LIBRARIES-REVOLUTIONIZING-DRUG.HTML

Affinity Selection to Identify Binders

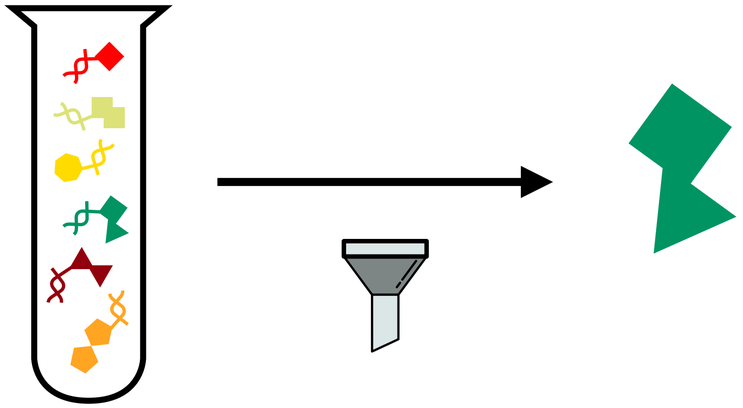
ONE ROUND:



- Typically 2-3 rounds
- More rounds = less noise, but also less signal
- Need to calibrate amount of input material to # of rounds
- Read out via PCR then DNA Sequencing

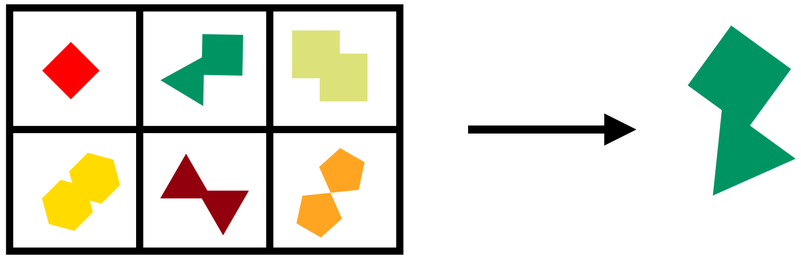
A Better Approach: Parallel Data Generation

DNA ENCODED LIBRARIES:
BILLIONS AT A TIME



- Billions of Compounds per tube
- 100s of Parallel Experiments
- Similar to Antibody Development

TRADITIONAL SCREENING:
ONE AT A TIME



- Single Compound Tested per well
- One Experiment at a Time
- 1-2M Compound Practical limit

Anagenex Platform

Completely integrated parallel biochemistry platform to generate the best data on the right compounds to enable ML and Iteration

Solution: Iterative Lab + ML Drug Discovery Funnel

Anagenex Platform

Unbiased, proprietary billion drug-like compound collection hits hard targets

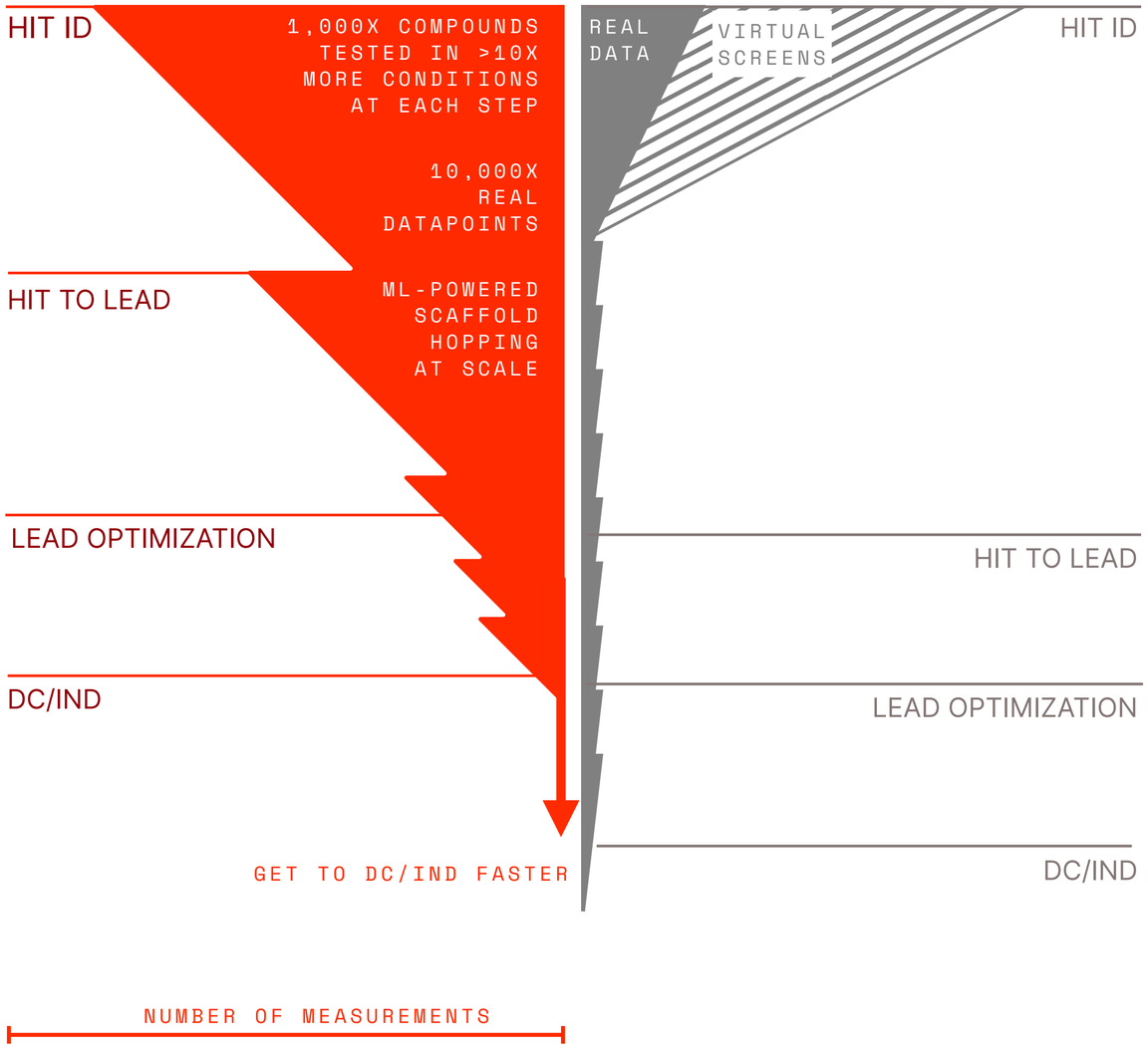
Proprietary ML identifies even more developable chemical series

Refresh the funnel with one million new, ML designed, never tested compounds

Many ML designed and optimized series with better initial properties

Fewer total iterations, faster proof of concept, faster time to clinic

Many backup options if any problems



Traditional Drug Discovery

Start with one million (or fewer) compounds biased by previous work misses new target classes

Or virtual screens with simulated data

Filtering leads to fewer developable compounds

Use limited SAR to design followup, build 10 compounds per month per chemist.

Fewer backup series

Need more optimization cycles, slowing time to proof of concept and clinic

Few options if any problems

anagenex

Our ML-Lab Discovery Platform Delivers:

Quick Tests In-Vitro & Beyond

- Rapid ID of drug-like compounds to get to validated chemical matter faster
- Case study 1, 2, 3

On Undruggable Targets

- ML and iteration find compounds despite ambiguous, non-selective results from initial screen
- Case study 2

Optimized Selectivity

- Get selective compounds early in the process
- Case study 3

anagenex

Iterative, Parallel Biochemistry Designed by ML Transforms Small Molecule Discovery

Parallel Biochemistry

- ✓ Proven on challenging targets
- ✓ Billions of datapoints per tube
- ✓ REAL not virtual Measurements
- - Simple chemistry
- - Too much data for humans to interpret

Hard to iteratively build & test libraries

Big, Iterative Experiments + Purpose Built ML

- ✓ ML-optimized experimental measurements
- ✓ Proprietary high-signal, low noise multibillion point datasets
- ✓ Proprietary algorithms optimized for our lab processes
- ✓ Rapid, complete exploration of chemical space
- ✓ Build+Test 1M new ML-designed compounds in 1.5 months

Best of both, enhanced by large scale iteration

Machine Learning

- - Needs real measurements not simulation
- - Requires large, quality datasets
- ✓ Extract meaning from massive datasets
- ✓ Extrapolates to new chemical space

Huge benefits from iterative refinement

anagenex

Experienced Blend of Drug Discovery and Computation

FOUNDER & CEO

Nicolas Tilmans

16 yr. ML & DEL. DiCE spinout from PhD | Stanford

CSO

Ryan Kruger

20+ yr. Drug Discovery, 6 INDs | Foghorn, GSK

SVP EARLY DISCOVERY

Joe Franklin

20 yr. Pioneer Parallel Chemistry, 3 INDs | FORMA

VP BIOLOGY

Svetlana Belyanskaya

20 yr. Pioneer Parallel Biochemistry | GSK

VP ML

Henri Palacci

11+ yr. ML | DE Shaw Research

LUX

AIR STREET CAPITAL.

khosla ventures

CATALIO CAPITAL MANAGEMENT

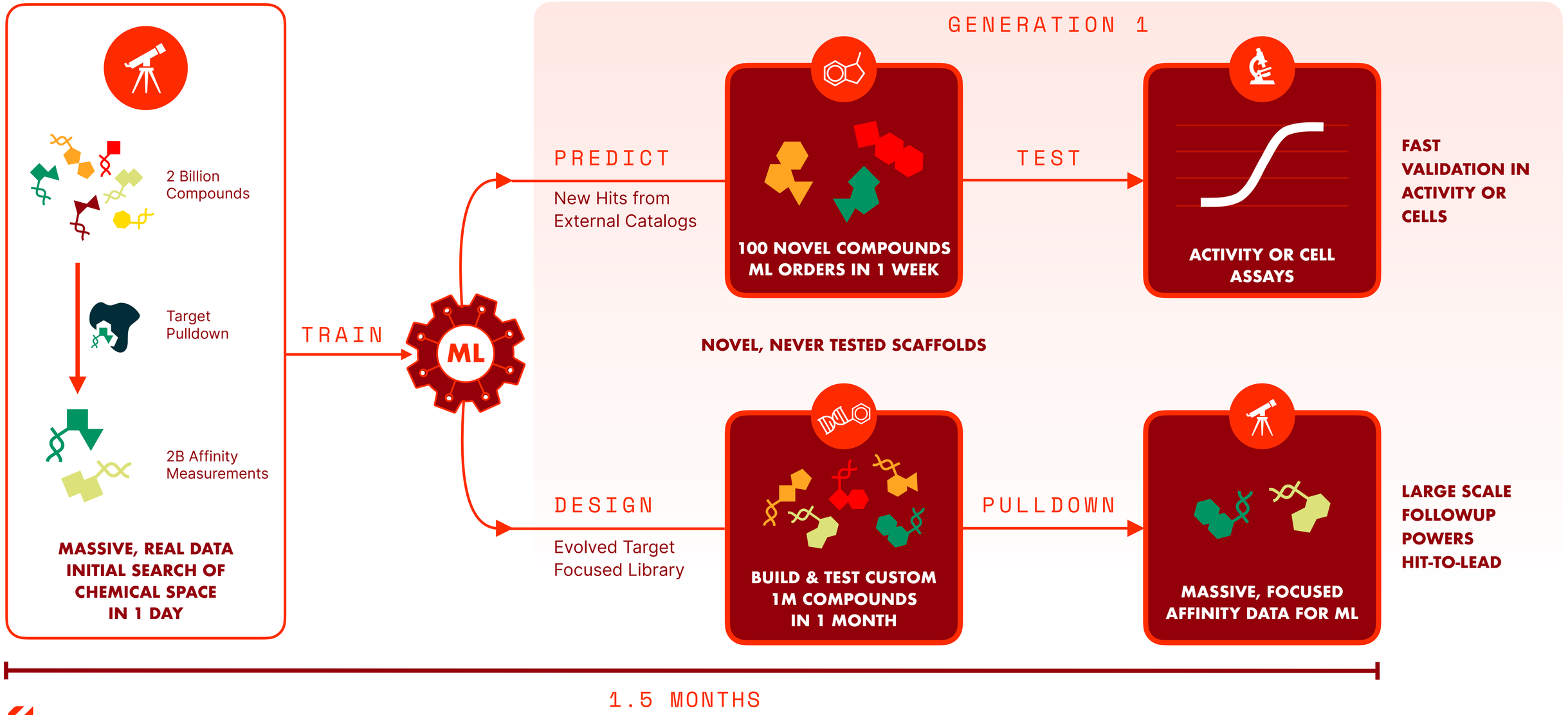
OBVIOUS

MENLO VENTURES

anagenex

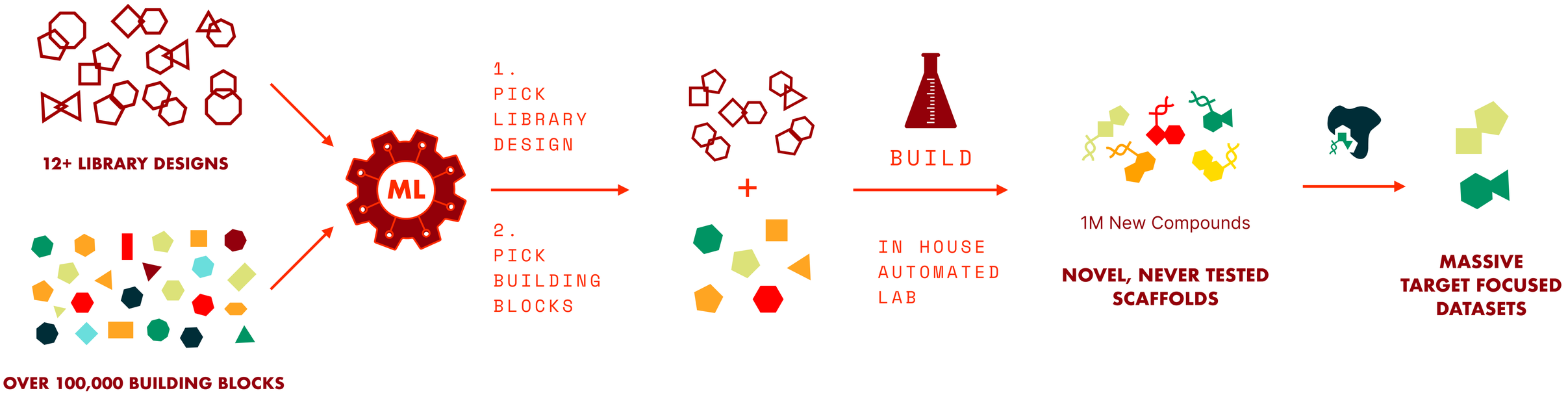
+16 others from diverse backgrounds form our uniquely talented team.

Step 1: Data+ML Powered Rapid and Large-Scale Followup



anagenex

ML-Designed Evolved Libraries: The Anagenex Difference



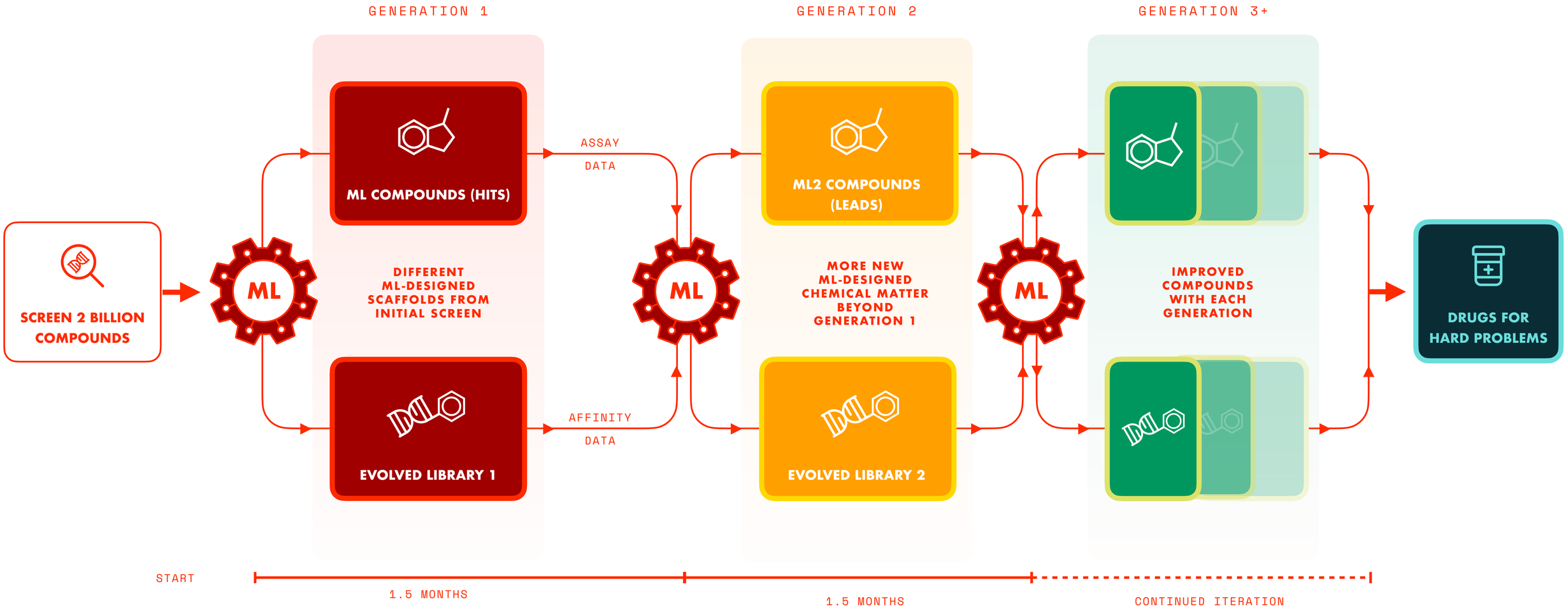
1.5 MONTHS

ML-Designed, Target Specific “Evolved” libraries Built and Tested in <6 weeks

Allows us to explore chemical space 8,000x more efficiently than traditional approaches

anagenex

Step 2+: ML-Driven Iteration to Find New Scaffolds



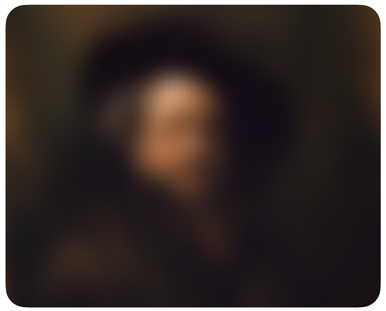
Our platform dynamically designs and builds novel previously unseen compounds with better properties at each step

These are experimentally tested to drive the next directed evolution loop with massive, real datasets

anagenex

Our Ability to Iterate is Unique and Unlocks New Opportunities

STANDARD DEL

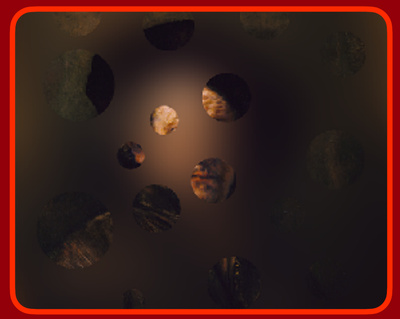


Blurry snapshot, once

Finds hits, cannot confidently explore beyond

Poor ML performance

anagenex



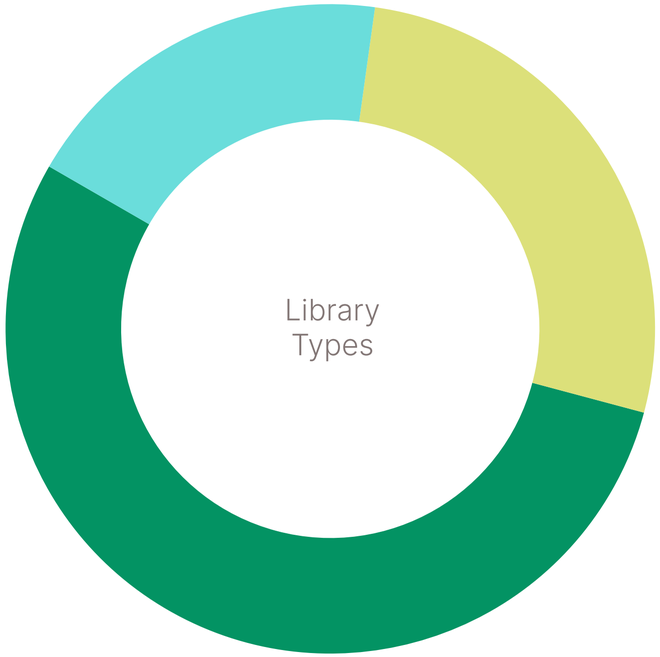
Clear snapshots

Better hits & understanding of SAR to optimize drugs

Excellent ML performance

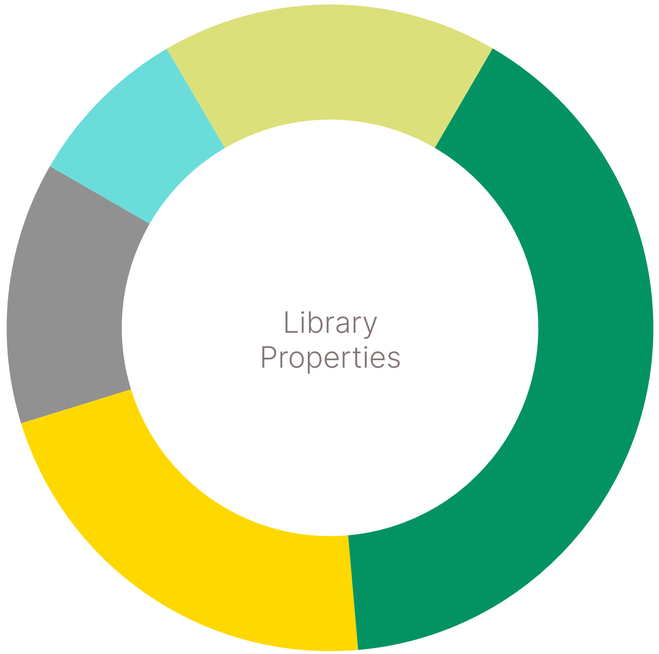
Expanded Chemical Space and Useful ML Predictions

30+ VARIED CONSTRUCTION LIBRARIES



- Scaffold
- Building Block
- Ring Forming

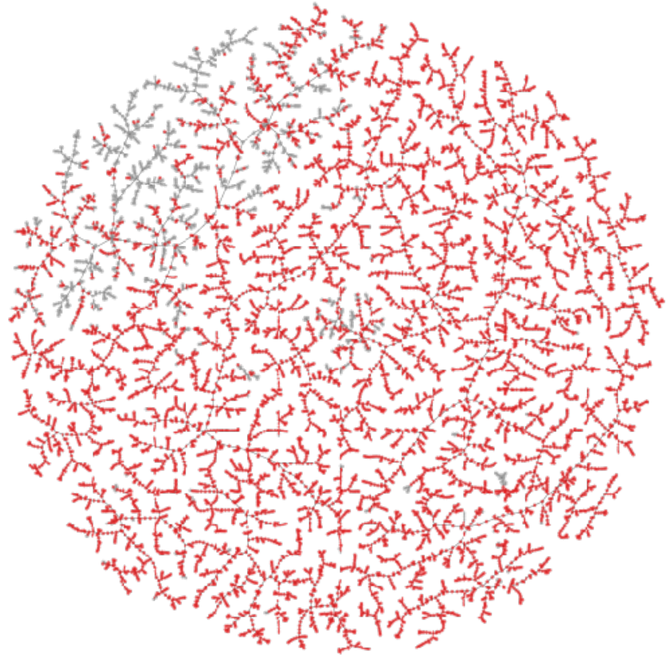
MEASURE DRUG-RELEVANT CHEMICAL SPACE



87% DRUG LIKE

- Ro4
- Ro4 Miss 1
- Ro5
- Ro5 Miss 1
- Other

WITH MUCH EXPANDED CHEMICAL MATTER



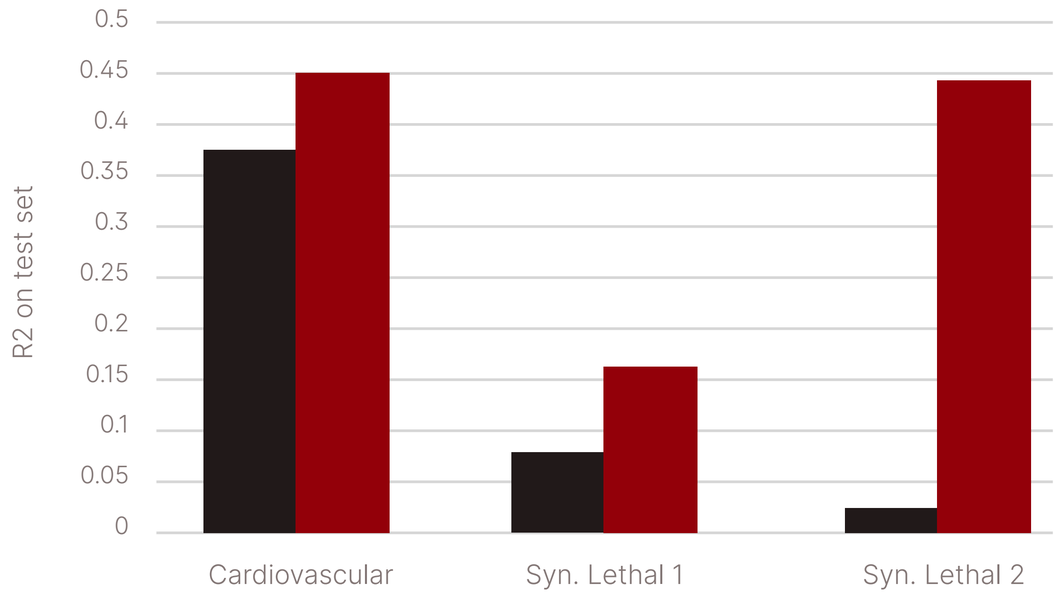
- Known Chemical Space
- Anagenex Addressable Space

anagenex

Superior ML by Collecting + Modeling Data Others Don't

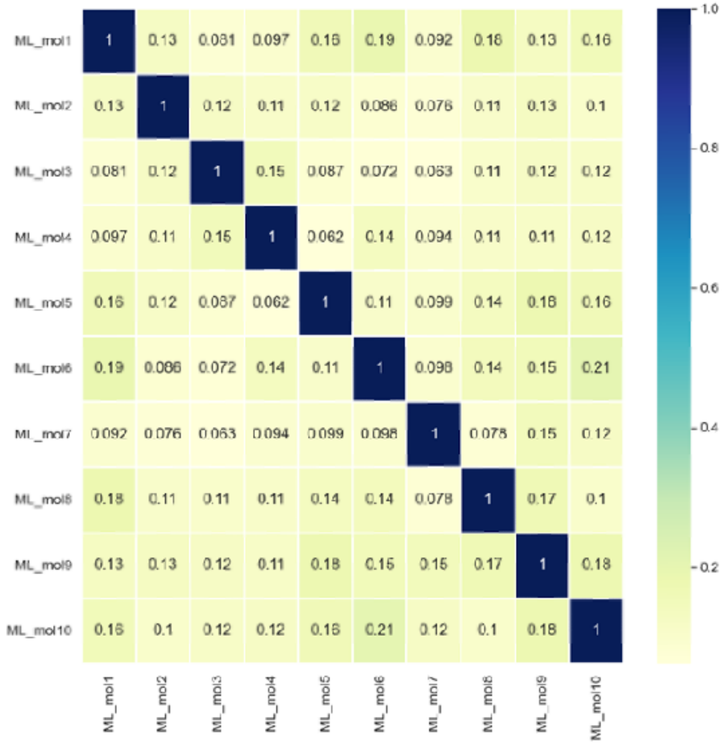
Our DEL Process Enables Better Model Performance

PROPRIETARY LAB PROCESS ENHANCES MODEL PERFORMANCE ON VARIOUS TARGETS



External Library Anagenex

USE ML TO FIND DISTINCT HIT FAMILIES



ML finds Novel, Chemically Distinct Compounds
(0.2 Tanimoto Similarity to original data)

anagenex

Differentiated To Identify New Chemical Matter

Large Scale, Real View of Chemical Space	Billion datapoint lab experiments on top of a 100+ billion datapoint database enables success on challenging targets
Fast Validation	ML powered predictions and fast iterative refinement identifies compounds ready for cells and beyond in six months
Large Scale Iteration	Build and test up to one million compounds in six weeks, allowing the ML to check its work, learn faster and solve harder targets
Accurate Predictions	By training our ML on real measurements rather than simulations, and updating them with fresh data, our models power high fidelity decision making

Validated to find lead compounds not hits or tools to hard, even “undruggable” target classes

So... Does it Deliver?

Three test cases on challenging targets to prove Parallel Biochemistry + ML
can solve more than just easy problems

Case Study 1: ML-Lab Iteration Improves Quality

Target: DNA Binder, PPI

- **Protein-Protein Interaction disruption**
  Hard for small molecules
- **No known compound co-crystal structure**
  Challenging for structure based simulations

Results: Lead Series in Nine Months

- **Out of 20M purchasable compounds our ML model picked 135 to test**
- **21% confirmed as specific target binders**
  (at least 10x better than HTS or virtual screening)
- Synthesized & Tested **300k ML designed compounds** to update ML model
- **Second generation predictions have more, higher quality compounds**
- **Beat nM tool compound, all lead-like properties**

Hit = experimentally confirmed binding of an ML predicted compound

anagenex

Case Study 2: ML-Lab Iteration Unlocks “Undruggables”

Target: Undruggable

- **Globular target with no known binding pockets**
  Hard for small molecules
- **Structure based approaches fail completely**
- **Antibody therapy, but no commercialized small molecules**

Results: Iteration Finds a Way

- **Generation 1 ML model predictions:**
  2 weak chemotypes from ML1
  1% hit rate
- **Generation 2 ML model predictions:**
  >50 chemotypes from Evolved ML2
  16% hit rate
- Best series <1uM in AS/MS dose response

Hit = experimentally confirmed binding of an ML predicted compound

Our iterative process found a novel cryptic pocket with several promising series in months

anagenex

Case Study 3: Clinically Valuable Synthetic Lethal Oncology

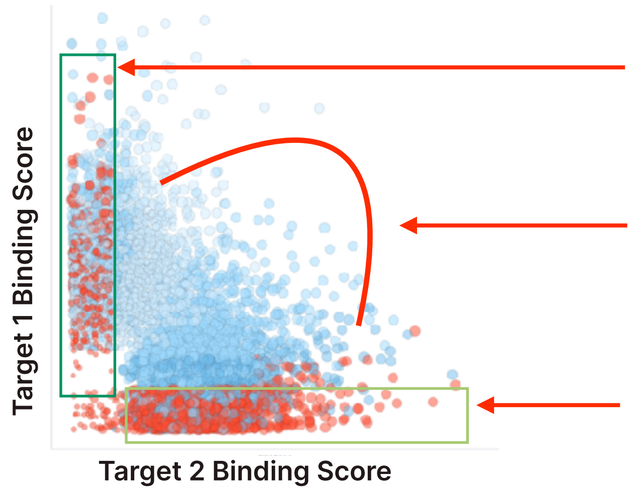
Target: Selectivity Challenge

- High industry interest target pair
- >80% sequence similarity
- No known selective compounds

Results: ML Finds Selectivity

- Anagenex ML identified promising compounds with initial selectivity
- Rapid medicinal chemistry improved selectivity to 14x
- Active in cellular assays (6 months to in-vitro PoC)

Anagenex Measures Billions of Binders to Each Target to Learn Selectivity

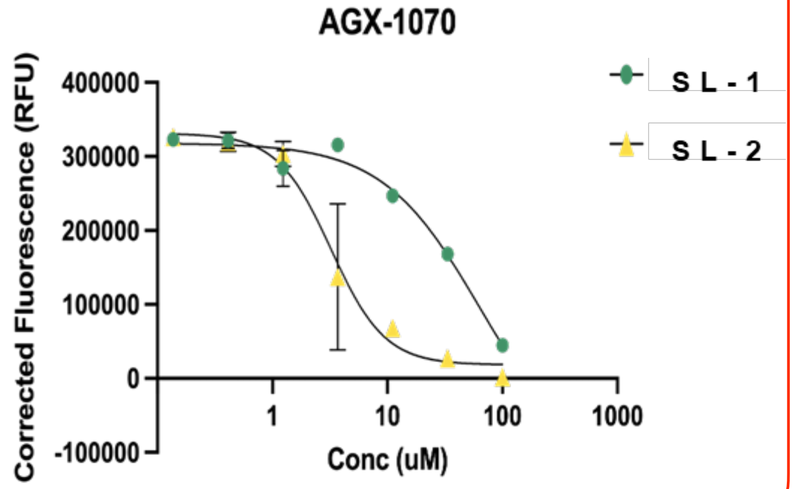


Compounds that hit only Target 1

Teaches us how molecules change to hit Target 1 or 2

Compounds that hit only Target 2

ML Identified 14x Selective Hits



Generation 2 iteration to identify improved compounds in progress

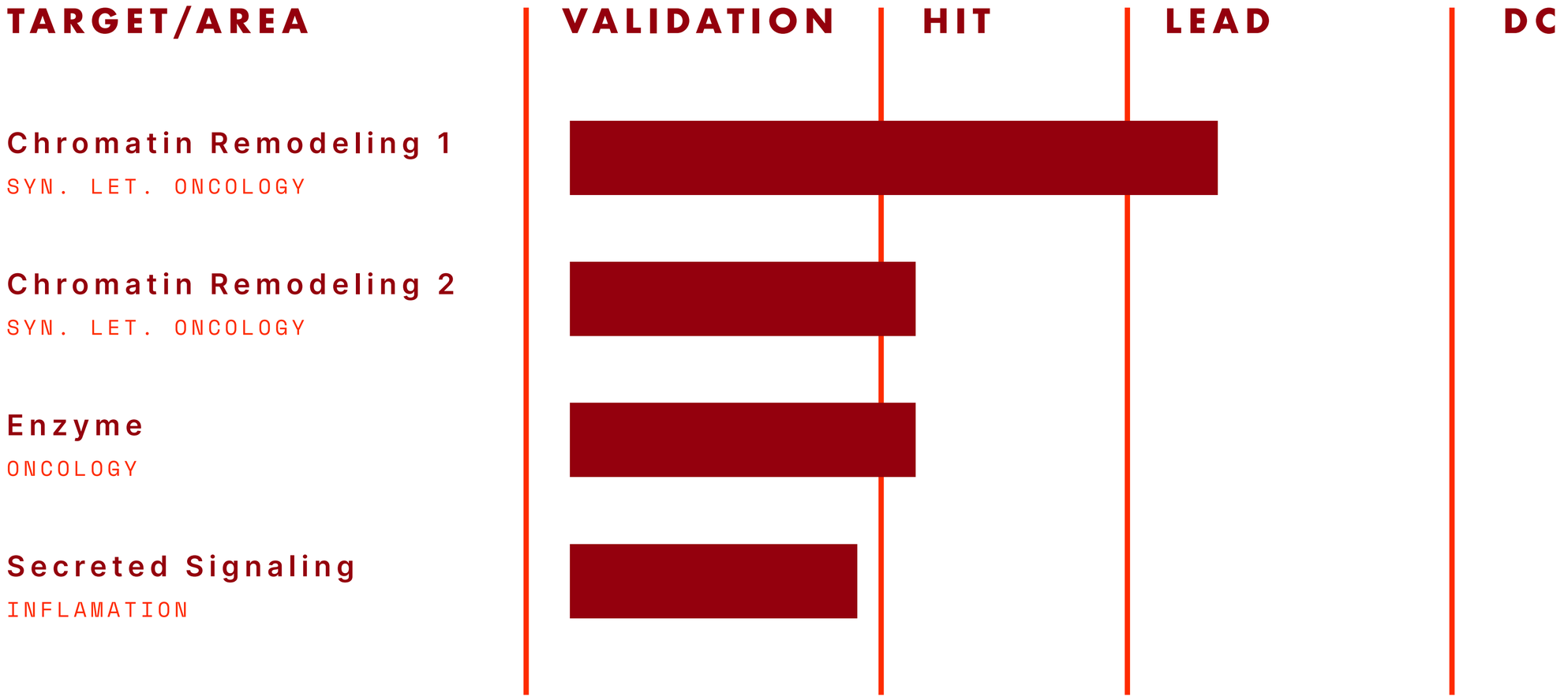
anagenex

A Proven Iterative ML-Lab Platform for Important Targets

2B Diverse Compounds Tested in 1 day	87% drug-like, selected for diversity and historical success
25 Targets Processed, over 100 Billion Experimentally Measured Datapoints Collected	Proprietary multiplexed selection process reduces noise & increases signal
Identified Hits With an Iterative ML-Driven Design Loop	Custom model learns better from rich internal datasets, improves >2x using a single ML-lab iteration
10 Targets with ML Identified Hits Including Known Challenging Target Classes	By training our ML on real measurements rather than simulations, our models power high fidelity decision making
3 Active Programs	Mix across synthetic lethal oncology and cardiovascular

anagenex

Our Pipeline: Focus on Synthetic Lethal Oncology



anagenex

Thank you!