Algorithms, Evolution and Network-Based Approaches in Molecular Discovery
Drug Discovery

The State of the Art
Antipsychotic drugs and their Poly-Pharmacology

The chart shows the known affinity (Ki) values of antipsychotic drugs for a panel of receptors.

How can we go about discovering a novel antipsychotic?
What do I make next?

Medicinal Chemistry Optimisation

- 3-Ph, -EtNH₂, CO₂H, CH₃CO₂H, CH₂CO₂Et, Imidazole, all inactive
- Cl – tolerated
- OBN, CO₂H - inactive
- Rapid drop off in activity
  - H > F > OH > Cl > Me
- Inactive Br, OCF₃, CN, OAlkyl, OBN, Amidine
- 2-3 Fused Systems/Cyclisation of carboxyl inactive
- Carboxyl homologation inactive
- N-Alkyl, Heteroatom substitution, N-Aryl all inactive
- 5-6 Bis-diol weakly active
- Cl – weakly active
- MeO - inactive
- 6-7 Fused Pyridyl inactive
- 7-Aryl substitution, CO₂H inactive

Diagram:

- Hit
- Potency
- Safety
- Absorption
- Solubility
- Metabolic stability
ChEMBL

ChEMBLSpace

[Image of ChEMBLSpace]

10.6k Targets
1.4m Compounds
12.8m Activities

ChEMBLSpace – a graphical explorer of the chemogenomic space covered by ChEMBL

Bioinformatics (2013) 29 (4): 523-524

https://www.ebi.ac.uk/chembldb/
ChEMBLSpace search: D2 & α1BA

Dopamine D2

α1B-Adrenergic
ChEMBLSpace: D2, α1BA, H1

Available for download search: ChEMBLSpace@sourceforge.net
Similarity Ensemble Approach (SEA)
Keiser et al.
Predictive Pharmacology

- Combination of:
  - Target Prediction
  - BioSAR - Laplacian-modified Naïve Bayes algorithm
  - Information gain prefiltering
  - Decision Trees (C4.5) for Classification

- Yields 70% accurate,
- Interpretable model for sleep outcome.

IF
ACTIVITY_ON D(2) Dopamine receptor
AND
ACTIVITY_ON Histamine H1 receptor
AND
ACTIVITY_ON 5-hydroxytryptamine receptor 2A
THEN
“Good Sleep”
A virtual enumeration of chemical space up to 17 heavy atoms generated 166,443,860,262 molecules.

A Pharma screening collection up to 17 heavy atoms is typically 100–500K molecules.

$\text{== 0.000003\% of accessible space}$
Strategies in Automated Molecule Design

Grow

Core X

A — Core X — B — C

Constraints

2D or 3D Generation

Protein Cavity

Volume Based

Feature Based

Synthetic Feasibility

Reaction Based

Reagent Enumeration

Transformation Based

Complexity Score

User Assessment

Replace Scaffold

Merge Molecules
Motivation

Using reaction databases

- Adding say ~250,000 reactions per year, strong medicinal chemistry bias
- Wealth of reaction data
  - Extract the knowledge hidden in these data
  - Use this knowledge to assist the medicinal chemist
  - Suggest new, synthetically feasible molecules with desired bio profile
Reaction Vectors

![Chemical reaction diagram with reactant and product structures and reaction vectors]

**Reactant vector,** $R = (R1 + R2)$

<table>
<thead>
<tr>
<th>Bond</th>
<th>R1</th>
<th>R2</th>
</tr>
</thead>
<tbody>
<tr>
<td>C=C</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>C=O</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>C-OH</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>C-OR</td>
<td>0</td>
<td>2</td>
</tr>
</tbody>
</table>

**Product vector,** $P$

<table>
<thead>
<tr>
<th>Bond</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>C=C</td>
<td>1</td>
</tr>
<tr>
<td>C=O</td>
<td>2</td>
</tr>
<tr>
<td>C-OH</td>
<td>3</td>
</tr>
<tr>
<td>C-OR</td>
<td>4</td>
</tr>
</tbody>
</table>

**Reaction vector,** $D = P - R$

<table>
<thead>
<tr>
<th>Bond</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>C=C</td>
<td>1</td>
</tr>
<tr>
<td>C=O</td>
<td>2</td>
</tr>
<tr>
<td>C-OH</td>
<td>1</td>
</tr>
<tr>
<td>C-OR</td>
<td>-2</td>
</tr>
</tbody>
</table>

Reaction Vectors in Structure Generation

• The reaction vector, \( D \), equals the difference between the product vector, \( P \), and the reactant vector, \( R \)

\[ D = P - R \]

Given a reaction vector, \( D \), and a reactant vector, \( R \), the product vector, \( P \), can be obtained

\[ P = D + R \]

Given a product vector, \( P \), can we reconstruct the product molecule(s)?

<table>
<thead>
<tr>
<th>I</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bond</td>
<td>C-C</td>
<td>C=O</td>
<td>C-OH</td>
<td>C-OR</td>
</tr>
<tr>
<td>#</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>2</td>
</tr>
</tbody>
</table>

better descriptor is required
Modified Atom Pairs

Atom Pairs 2 (AP2): $X_1(n, p, r)$-2(BO)-$X_2(n, p, r)$

- $X$: element type
- $n$: number of bonds to heavy atoms
- $p$: number of $\pi$ bonds
- $r$: number of ring memberships
- $BO$: bond order

Extending the bond distance in atom pairs encodes more of the environment of the reaction centre.
**Beckmann Rearrangement**

![Image of Beckmann Rearrangement](image)

**Reaction vector**

<table>
<thead>
<tr>
<th>Negative APs</th>
<th>Positive APs</th>
</tr>
</thead>
<tbody>
<tr>
<td>1- C(3,2,1)-2(1)-C(3,1,0)</td>
<td>1+ C(3,2,1)-2(1)-N(2,0,0)</td>
</tr>
<tr>
<td>2- C(3,1,0)-2(2)-N(2,1,0)</td>
<td>2+ C(3,1,0)-2(1)-N(2,0,0)</td>
</tr>
<tr>
<td>3- N(2,1,0)-2(1)-O(1,0,0)</td>
<td>3+ C(3,1,0)-2(2)-O(1,1,0)</td>
</tr>
<tr>
<td>a- C(3,2,1)-3-N(2,1,0)</td>
<td>a+ C(2,2,1)-3-N(2,0,0)</td>
</tr>
<tr>
<td>b- C(3,2,1)-3-C(1,0,0)</td>
<td>b+ C(2,2,1)-3-N(2,0,0)</td>
</tr>
<tr>
<td>c- C(3,1,0)-3-C(2,2,1)</td>
<td>c+ C(2,2,1)-3-N(2,0,0)</td>
</tr>
<tr>
<td>d- C(3,1,0)-3-C(2,2,1)</td>
<td>d+ N(2,0,0)-3-C(1,0,0)</td>
</tr>
<tr>
<td>e- C(3,1,0)-3-O(1,0,0)</td>
<td>e+ N(2,0,0)-3-O(1,1,0)</td>
</tr>
<tr>
<td>f- N(2,1,0)-3-C(1,0,0)</td>
<td>f+ O(1,1,0)-3-C(1,0,0)</td>
</tr>
</tbody>
</table>

**Atom Pairs 2 (AP2):** X1(n, p, r)-2(BO)-X2(n, p, r)

**Atom Pairs 3 (AP3):** X1(n, p, r)-3-X2(n, p, r)

**Legend:**
- **X**: element type
- **n**: number of bonds to heavy atoms
- **p**: number of π bonds
- **r**: number of ring memberships
- **BO**: bond order
Applying a RV to a reactant to generate a Product

1. Removing the negative atom pairs from the reactant

<table>
<thead>
<tr>
<th>Negative APs</th>
<th>Positive APs</th>
</tr>
</thead>
<tbody>
<tr>
<td>1- C(3,2,1)-2(1)-C(3,1,0)</td>
<td>1+ C(3,2,1)-2(1)-N(2,0,0)</td>
</tr>
<tr>
<td>2- C(3,1,0)-2(2)-N(2,1,0)</td>
<td>2+ C(3,1,0)-2(1)-N(2,0,0)</td>
</tr>
<tr>
<td>3- N(2,1,0)-2(1)-O(1,0,0)</td>
<td>3+ C(3,1,0)-2(2)-O(1,1,0)</td>
</tr>
<tr>
<td>a- C(3,2,1)-3-N(2,1,0)</td>
<td>a+ C(2,2,1)-3-N(2,0,0)</td>
</tr>
<tr>
<td>b- C(3,2,1)-3-C(1,0,0)</td>
<td>b+ C(2,2,1)-3-N(2,0,0)</td>
</tr>
<tr>
<td>c- C(3,1,0)-3-C(2,2,1)</td>
<td>c+ C(2,2,1)-3-N(2,0,0)</td>
</tr>
<tr>
<td>d- C(3,1,0)-3-C(2,2,1)</td>
<td>d+ N(2,0,0)-3-C(1,0,0)</td>
</tr>
<tr>
<td>e- C(3,1,0)-3-O(1,0,0)</td>
<td>e+ N(2,0,0)-3-O(1,1,0)</td>
</tr>
<tr>
<td>f- N(2,1,0)-3-C(1,0,0)</td>
<td>f+ O(1,1,0)-3-C(1,0,0)</td>
</tr>
</tbody>
</table>
Applying a RV to a reactant to generate a Product

2. Adding positive atom pairs to the fragment

<table>
<thead>
<tr>
<th>Negative APs</th>
<th>Positive APs</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-</td>
<td>1+</td>
</tr>
<tr>
<td>C(3,2,1)-2(1)-C(3,1,0)</td>
<td>C(3,2,1)-2(1)-N(2,0,0)</td>
</tr>
<tr>
<td>2-</td>
<td>2+</td>
</tr>
<tr>
<td>C(3,1,0)-2(2)-N(2,1,0)</td>
<td>C(3,1,0)-2(1)-N(2,0,0)</td>
</tr>
<tr>
<td>3-</td>
<td>3+</td>
</tr>
<tr>
<td>N(2,1,0)-2(1)-O(1,0,0)</td>
<td>C(3,1,0)-2(2)-O(1,1,0)</td>
</tr>
<tr>
<td>a-</td>
<td>a+</td>
</tr>
<tr>
<td>C(3,2,1)-3-N(2,1,0)</td>
<td>C(2,2,1)-3-N(2,0,0)</td>
</tr>
<tr>
<td>b-</td>
<td>b+</td>
</tr>
<tr>
<td>C(3,2,1)-3-C(1,0,0)</td>
<td>C(2,2,1)-3-N(2,0,0)</td>
</tr>
<tr>
<td>c-</td>
<td>c+</td>
</tr>
<tr>
<td>C(3,1,0)-3-C(2,2,1)</td>
<td>C(2,2,1)-3-N(2,0,0)</td>
</tr>
<tr>
<td>d-</td>
<td>d+</td>
</tr>
<tr>
<td>C(3,1,0)-3-C(2,2,1)</td>
<td>N(2,0,0)-3-C(1,0,0)</td>
</tr>
<tr>
<td>e-</td>
<td>e+</td>
</tr>
<tr>
<td>C(3,1,0)-3-O(1,0,0)</td>
<td>N(2,0,0)-3-O(1,1,0)</td>
</tr>
<tr>
<td>f-</td>
<td>f+</td>
</tr>
<tr>
<td>N(2,1,0)-3-C(1,0,0)</td>
<td>O(1,1,0)-3-C(1,0,0)</td>
</tr>
</tbody>
</table>

Atom Pairs 2 (AP2) : X1(n, p, r)-2(BO)-X2(n, p, r)

- X: element type
- n: number of bonds to heavy atoms
- p: number of π bonds
- r: number of ring memberships
- BO: bond order

No AP2s left in the reaction vector that match atom 11

Final Solution

Duplicate solution
How well does it work?

Organic Chemistry Database

<table>
<thead>
<tr>
<th>Reaction Type</th>
<th>Number of Reactions</th>
<th>Correctly Reproduced</th>
<th>Number</th>
<th>Per cent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Epoxide reduction</td>
<td>450</td>
<td>449</td>
<td>99.8</td>
<td></td>
</tr>
<tr>
<td>Epoxide formation</td>
<td>450</td>
<td>444</td>
<td>98.7</td>
<td></td>
</tr>
<tr>
<td>Ester to amide</td>
<td>172</td>
<td>172</td>
<td>100.0</td>
<td></td>
</tr>
<tr>
<td>Alcohol dehydration</td>
<td>171</td>
<td>169</td>
<td>98.8</td>
<td></td>
</tr>
<tr>
<td>Claisen rearrangement</td>
<td>61</td>
<td>54</td>
<td>88.5</td>
<td></td>
</tr>
<tr>
<td>Beckmann rearrangement</td>
<td>123</td>
<td>123</td>
<td>100.0</td>
<td></td>
</tr>
<tr>
<td>Friedel Crafts acylation</td>
<td>113</td>
<td>113</td>
<td>100.0</td>
<td></td>
</tr>
<tr>
<td>Olefin metathesis</td>
<td>9</td>
<td>7</td>
<td>77.8</td>
<td></td>
</tr>
<tr>
<td>Dieckmann condensation</td>
<td>98</td>
<td>91</td>
<td>92.9</td>
<td></td>
</tr>
<tr>
<td>Nitro reduction</td>
<td>231</td>
<td>230</td>
<td>99.6</td>
<td></td>
</tr>
<tr>
<td>Alkene oxidation</td>
<td>272</td>
<td>272</td>
<td>100.0</td>
<td></td>
</tr>
<tr>
<td>Cope rearrangement</td>
<td>453</td>
<td>306</td>
<td>67.5</td>
<td></td>
</tr>
<tr>
<td>Aldol condensation</td>
<td>134</td>
<td>134</td>
<td>100.0</td>
<td></td>
</tr>
<tr>
<td>Alcohol amination</td>
<td>97</td>
<td>97</td>
<td>100.0</td>
<td></td>
</tr>
<tr>
<td>Amide reduction</td>
<td>51</td>
<td>51</td>
<td>100.0</td>
<td></td>
</tr>
<tr>
<td>Diels-Alder hetero</td>
<td>441</td>
<td>320</td>
<td>72.6</td>
<td></td>
</tr>
<tr>
<td>Ether halogenation</td>
<td>58</td>
<td>58</td>
<td>100.0</td>
<td></td>
</tr>
<tr>
<td>Ozonolysis</td>
<td>132</td>
<td>125</td>
<td>94.7</td>
<td></td>
</tr>
<tr>
<td>Claisen condensation</td>
<td>98</td>
<td>77</td>
<td>78.6</td>
<td></td>
</tr>
<tr>
<td>Carboxylic acids to aldehydes</td>
<td>194</td>
<td>194</td>
<td>100.0</td>
<td></td>
</tr>
<tr>
<td>Nitrile reduction</td>
<td>102</td>
<td>102</td>
<td>100.0</td>
<td></td>
</tr>
<tr>
<td>Diels-Alder cycloaddition</td>
<td>106</td>
<td>65</td>
<td>61.3</td>
<td></td>
</tr>
<tr>
<td>Fischer indole</td>
<td>230</td>
<td>94</td>
<td>40.9</td>
<td></td>
</tr>
<tr>
<td>Alkene halogenation</td>
<td>310</td>
<td>281</td>
<td>90.6</td>
<td></td>
</tr>
<tr>
<td>Nitrile hydrolysis</td>
<td>460</td>
<td>460</td>
<td>100.0</td>
<td></td>
</tr>
<tr>
<td>Olefination</td>
<td>455</td>
<td>427</td>
<td>93.8</td>
<td></td>
</tr>
<tr>
<td>Wittig-Horner</td>
<td>211</td>
<td>190</td>
<td>90.0</td>
<td></td>
</tr>
<tr>
<td>Robinson annihilation</td>
<td>13</td>
<td>10</td>
<td>76.9</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>5,695</td>
<td>5,115</td>
<td>89.8</td>
<td></td>
</tr>
</tbody>
</table>

- Products generated for 5,115 reactions (~90% of the 5,695)
- ~3 seconds per reaction average, 0.015 seconds median run time
Evolutionary Design

From fragment to drug-like molecules with designed polypharmacology

1. Using known chemistry for in-silico design and optimisation

2. Evolutionary Algorithm

- Algorithm is well suited to fragment based design (FBDD) approach wherein small fragments are grown into a protein active site.

- Only the best scoring molecules survive and are resubmitted to the design algorithm to genetically evolve into better scoring solutions.

- Methods to score the designs include: QSARs, pharmacophore fit, docking scores, similarity to known actives, polypharmacology prediction, physical property predictions etc.

Starting material → Evotec Reaction Transformation Database → Pool of possible products → Reaction Vectors → In-silico reaction → Score → Select

Q² = 0.76

Activity X_QSAR

Local Models

Multi-objective Pareto optimisation algorithm

Q² = 0.68

ADMET_QSAR

Global Models

Cream-off top scoring molecules for evolution

Methods to score the designs include: QSARs, pharmacophore fit, docking scores, similarity to known actives, polypharmacology prediction, physical property predictions etc.
Multi-Objective Optimisation

“Score” = Similarity + D2_{predAct} + \alpha 1BA_{predAct} + H1_{predAct}

Haloperidol + Ziprasidone

Union of Descriptor

Q^2 = 0.40

Q^2 = 0.68

Q^2 = 0.76
Results: Piperidone

26K Reactions, 93K Reagents

Starting Material

Chlorpromazine

Fluphenazine
It looks great but ...

1. The algorithm only knows about transformation types that are in the Db!
2. The AP2/3’s cover 1 and 2 bonds. Remote functionality isn’t considered.
3. A reaction path is not a drug “optimisation”!

But how do we get from here to here?
Reaction Sequence Vectors

Tools for molecular design

Sequence Vectors

a → b
a → c
a → d
a → e
b → x
**SAR Exploration**

**Succinyl Hydroxamates**

Succinyl hydroxamates as potent and selective non-peptidic inhibitors of procollagen C-proteinase: Design, synthesis, and evaluation as topically applied, dermal anti-scarring agents

Simon Bailey**, a,b,** Paul V. Fish a,** Stephane Billotte a,** Jon Bordner a,** Doris Greiling b,** Kim James a,** Andrew McElroy a,** James E. Mills a,** Charlotte Reed a,** Robert Webster a,d

a Department of Discovery Chemistry, Pfizer Global Research and Development, Sandwich Laboratories, Ramsgate Road, Sandwich, Kent CT13 9NJ, UK
b Department of Discovery Biology, Pfizer Global Research and Development, Sandwich Laboratories, Ramsgate Road, Sandwich, Kent CT13 9NJ, UK
c Department of Pharmaceutical Sciences, Pfizer Global Research and Development, Sandwich Laboratories, Ramsgate Road, Sandwich, Kent CT13 9NJ, UK
d Department of Pharmaceutics, Dynamic & Metabolism, Pfizer Global Research and Development, Sandwich Laboratories, Ramsgate Road, Sandwich, Kent CT13 9NJ, UK

**Abstract**

Succinyl hydroxamates 1 and 2 are disclosed as novel series of potent and selective inhibitors of procollagen C-proteinase (PCP) which may have potential as anti-fibrotic agents. Carboxamide 7 demonstrated good PCP inhibition and had excellent selectivity over MMPs involved in wound healing. In addition, 7 was effective in a cell-based model of collagen deposition (fibroplasia model) and was very effective at penetrating human skin in vitro. Compound 7 (UK-383,367) was selected as a candidate for evaluation in clinical studies as a topically applied, dermal anti-scarring agent.

**Table:**

<table>
<thead>
<tr>
<th>Compound</th>
<th>NR ( R^2 )</th>
<th>PCP ( IC_{50} ) (nM)</th>
<th>MMP-2 IC( 50 ) (nM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>NH(_2)</td>
<td>44</td>
<td>&gt;50,000</td>
</tr>
<tr>
<td>(S)-7</td>
<td>NH(_2)</td>
<td>&gt;2000</td>
<td>NT</td>
</tr>
<tr>
<td>8</td>
<td>NHMe</td>
<td>28</td>
<td>&gt;30,000</td>
</tr>
<tr>
<td>9</td>
<td>NH(_2)O</td>
<td>10</td>
<td>&gt;10,000</td>
</tr>
<tr>
<td>10</td>
<td>NH-i-Pr</td>
<td>32</td>
<td>21,700</td>
</tr>
<tr>
<td>11</td>
<td>NH-CH(_2)-i-Pr</td>
<td>21</td>
<td>&gt;100,000</td>
</tr>
<tr>
<td>12</td>
<td>NHCH(_2)-i-Pr</td>
<td>33</td>
<td>&gt;30,000</td>
</tr>
<tr>
<td>13</td>
<td>NHCH(_2)-OH</td>
<td>57</td>
<td>NT</td>
</tr>
<tr>
<td>14</td>
<td>NHCH(_2)-Py</td>
<td>37</td>
<td>&gt;30,000</td>
</tr>
<tr>
<td>15</td>
<td>NHCH(_2)-CO(_2)H</td>
<td>21</td>
<td>&gt;100,000</td>
</tr>
<tr>
<td>16</td>
<td>Pyrrolidine</td>
<td>43</td>
<td>NT</td>
</tr>
<tr>
<td>17</td>
<td>Piperidine</td>
<td>26</td>
<td>74,000</td>
</tr>
<tr>
<td>18</td>
<td>Morpholine</td>
<td>25</td>
<td>NT</td>
</tr>
<tr>
<td>19</td>
<td>1,4-Dihydropyridine</td>
<td>54</td>
<td>NT</td>
</tr>
<tr>
<td>20</td>
<td>NMeCH(_2)-Ph</td>
<td>65</td>
<td>81,800</td>
</tr>
<tr>
<td>21</td>
<td>NMeCH(_2)-2-Py</td>
<td>40</td>
<td>50,500</td>
</tr>
</tbody>
</table>
Novel SAR

Succinyl Hydroxamates
Principle Components Analysis of Property Space

Succinyl Hydroxamates

Legend

- **Known products**
- **Near neighbours** (Tanimoto 1.0-0.8)

Legend

- **Known products**
- **Near neighbours** (Tanimoto 1.0-0.8)
- **All other products**
Mapping Discovery Space

Sequence vector network

600K+ reactions from US Patent Database: Nextmove
The PGC GWAS

Genome Wide Association Studies

Psychiatric Genomics Consortium

RESULTS
Results to date
ADHD
Bipolar disorder
Cross-disorder
MDD
Schizophrenia
Background papers
GPCRs associated with Schizophrenia GWAS genes

2 step shortest paths network

- Schizophrenia GWAS genes
- GPCR receptor subtypes

Eg, From: Dopamine receptor subtypes
To: Proteins defined by PGC2
schizophrenia GWAS genes

Histamine receptor network
Dopamine receptor network
Adrenergic receptor network

Remove proteins degree <10

Hubs involved in GPCR signalling in schizophrenia
In-Silico Network Based Design

Connecting Gene Expression Data from Connectivity Map and In Silico Target Predictions For Small Molecule Mechanism-of-Action Analysis
Bender et al., Molecular BioSystems 2015, 11, 86-96
PhenoTarget - Phenotypic screening to target de-convolution

Networks in Molecular Discovery

Compound selection → Phenotypic Screen → Build SAR
Evolve targets and pathways → Hit val.-H2L-LO
Target identification and validation → In vivo studies

Focused set Chemically diverse compounds
Structurally diverse hits
Hits with target and pathway annotations
Biologically diverse (annotated) compounds

Structural clustering
Orthogonal assays
Mechanism-informed phenotypic assay

Hit expansion by structure
Secondary assays
 Knockdown studies

Predictive Pharmacology
 Selectivity assays (Cerep, Kinaffinity)
Cellular Target Profiling

Proteome / Transcriptome analysis
Proteome
Target specific assays
Gene targeting mouse model

Hit expansion by target/pathway
Biomarker development

Disease models in vivo efficacy
Biomarker analysis PK/PD and MoA

legend
Target-agnostic route and App’s
Target-agnostic route and App’s
App for both routes
An informatics workflow

When more is not always more

<table>
<thead>
<tr>
<th></th>
<th>Target count</th>
<th>Pathway count</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>Significant</td>
</tr>
<tr>
<td>(1) Exact match</td>
<td>895</td>
<td>30</td>
</tr>
<tr>
<td>(2) + similarity search</td>
<td>940</td>
<td>41</td>
</tr>
<tr>
<td>(3) + predictive models</td>
<td>1380</td>
<td>134</td>
</tr>
</tbody>
</table>
Given a disease signature how do we best sample appropriate chemistry space?
Happiness

What makes you happy?

A man doesn’t know what happiness is until he’s married, by then it’s too late!  *(Frank Skinner)*
Acknowledgments

Val Gillet
Beining Chen
Hina Patel
Ben Allen
James Wallace

Andreas Bender
Georgios Drakakis

Tim James
Inaki Morai
Jon Hollick
York Rudhard
Alex Böcker

Dimitar Hristozov
David Thorner
David Evans
Nik Fechner
George Papadatos
Suzanne Brewerton
Lewis Vidler

John Liebeschuetz
Jason Cole
Colin Groom
Your contact:
Mike Bodkin
VP Research Informatics
114 Innovation Drive,
Milton Park, Abingdon
Oxfordshire OX14 4RZ, UK
T: +44 (0)1235 44 1207
Mike.Bodkin@evotec.com
## Track record

**Evotec Phenotypic DD projects and applied methods for TI/TV**

<table>
<thead>
<tr>
<th>Assay Dev</th>
<th>Hit ID, Hit validation</th>
<th>H2L</th>
<th>LO</th>
<th>in vivo studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kidney fibrosis</td>
<td>diverse / bioann. cpds</td>
<td>SimSearch</td>
<td></td>
<td>in vivo val</td>
</tr>
<tr>
<td>Melanocytes</td>
<td>diverse cpds</td>
<td></td>
<td>Pred.Pharm</td>
<td></td>
</tr>
<tr>
<td>Orphan transporter</td>
<td>diverse cpds</td>
<td></td>
<td>CTP</td>
<td></td>
</tr>
<tr>
<td>Cancer immunotherapy</td>
<td>diverse / bioann. cpds</td>
<td>CTP</td>
<td>Pred.Pharm</td>
<td></td>
</tr>
<tr>
<td>Diabetic nephropathy</td>
<td>diverse / bioann. cpds</td>
<td>Tomics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B cell proliferate</td>
<td>diverse / bioann. cpds</td>
<td>Tomics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neuroprotection</td>
<td>bioann. cpds</td>
<td>SimSearch</td>
<td></td>
<td>in vivo val</td>
</tr>
<tr>
<td>Neuroinflammation</td>
<td>diverse / bioann. cpds</td>
<td>Tomics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lung Fibrosis</td>
<td>diverse / bioannotated cpds</td>
<td>Pred.Pharm</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cancer double strand break repair</td>
<td>diverse / bioannotated cpds</td>
<td>Pred.Pharm</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anti-viral</td>
<td>anti-viral focussed library from Novira</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Muscular dystrophy</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Motor Neuron Disease</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Others…</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1. Target ID achieved by known/annotated drug application
2. Annotated Cpd X prevents pericyte to myofibroblast transition and attenuates kidney fibrosis in vivo

- **Target ID**
- **Omic**
- **In vivo val.**
- **planned**
Informing the Discovery Pipeline

Cheminformatics and Bioinformatics toolkit

1. Target Known
   - Different operational paradigms apply depending on whether the target is known
   - Chemically diverse
   - Biologically diverse

   - Target Identification
   - Ligand Identification
   - Ligand Optimisation

2. Target Unknown

   - Phenotypic Screen
   - Build SAR
   - Hit Val-H2L-LO
   - Target Identification & Validation

   Evolve Target & Pathway Hypothesis