Pharmacogenomics and bioinformatics in drug discovery

Anil Jegga, DVM, MRes

Division of Biomedical Informatics
Cincinnati Children's Hospital Medical Center
Department of Pediatrics and Biomedical Informatics
University of Cincinnati College of Medicine
Department of Electrical Engineering and Computing Systems
University of Cincinnati College of Engineering
Outline

- **Drug discovery and development – Overview**
- **Pharmacogenomics**
  - Resources
  - Role in drug discovery and development
- **Bioinformatics – Leveraging omics big-data**
  - Resources
  - Role in drug discovery and development
  - Drug repositioning/repurposing
- **Integrating Pharmacogenomics & Bioinformatics – Case studies**
  - Cystic fibrosis
  - Drug-induced adverse events
Drug discovery and development – Overview
New Drug Development - Problems

• Expensive
• Time consuming
• **High attrition** - Only five in 5,000 compounds that enter preclinical testing make it to human testing. Only 1 of these 5 tested in humans is approved.

• **Post-marketing drug failure**

• **Decreased return on investment** by 50% in 10 years – patent expiry (generics, etc.)
Expensive and time consuming

**US$2.6 billion** - Tufts Center for the Study of Drug Development, 2015

$800 million and $1.2 billion

- **Discovery**: Identify the molecular basis of disease; find an “ideal” molecular target and compound (specificity, off-target effects, etc.)
- **Preclinical testing**: Model systems – Compounds successfully passing these move to clinical trial
  - **Phase-1 clinical trial**: Safety in humans
  - **Phase-2 clinical trial**: Determine optimal dose – Assess efficacy and safety
  - **Phase-3 clinical trial**: Look for statistically significant benefit - 1000 to 2000 patients

10-12 years
High attrition

Root-cause analysis for 359 phase 3 and 95 NDA/BLA suspended programs

Hay et al., 2014, Nature Biotechnology
Post-marketing drug failure

**Examples of some withdrawn drugs post-approval for safety reasons (FDA Orange Book)**

<table>
<thead>
<tr>
<th>Generic Name (Brand Name)</th>
<th>Approval Date (US FDA)</th>
<th>Withdrawal Date (US FDA)</th>
<th>Years on Market</th>
</tr>
</thead>
<tbody>
<tr>
<td>sibutramine (MERIDIA)</td>
<td>11/22/1997</td>
<td>10/8/2010</td>
<td>12.9</td>
</tr>
<tr>
<td>trasylol (APRÖTININ)</td>
<td>12/29/1993</td>
<td>11/5/2007</td>
<td>3.9</td>
</tr>
<tr>
<td>tegaserod (ZELNORM)</td>
<td>7/24/2002</td>
<td>3/30/2007</td>
<td>4.7</td>
</tr>
<tr>
<td>gatifloxacin (TEQUIN)</td>
<td>12/17/1999</td>
<td>5/1/2006</td>
<td>6.4</td>
</tr>
<tr>
<td>technetium (99m TC) fanolesomab (NEUTROSPEC)</td>
<td>7/2/2004</td>
<td>12/19/2005</td>
<td>1.5</td>
</tr>
<tr>
<td>hydromorphone (PALLADONE)</td>
<td>9/24/2004</td>
<td>7/13/2005</td>
<td>0.8</td>
</tr>
<tr>
<td>valdecoxib (BEXTRA)</td>
<td>11/16/2001</td>
<td>4/7/2005</td>
<td>3.4</td>
</tr>
<tr>
<td>natalizumab (TYSABRI)</td>
<td>11/23/2004</td>
<td>2/28/2005</td>
<td>0.3</td>
</tr>
<tr>
<td>rofecoxib (VIOXX)</td>
<td>5/20/1999</td>
<td>9/29/2004</td>
<td>5.4</td>
</tr>
<tr>
<td>levomethadyl (ORLAAM)</td>
<td>7/9/1993</td>
<td>9/2/2003</td>
<td>10.2</td>
</tr>
<tr>
<td>cerivastatin (BAYCOL)</td>
<td>6/26/1997</td>
<td>8/8/2001</td>
<td>7.3</td>
</tr>
<tr>
<td>rapacuronium (RAPLON)</td>
<td>8/18/1999</td>
<td>3/30/2001</td>
<td>1.6</td>
</tr>
<tr>
<td>alosetron (LOTRONEX)</td>
<td>2/9/2000</td>
<td>11/28/2000</td>
<td>0.8</td>
</tr>
<tr>
<td>troglitazone (REZULIN)</td>
<td>1/29/1997</td>
<td>3/21/2000</td>
<td>3.1</td>
</tr>
<tr>
<td>grepafloxacin (RAXAR)</td>
<td>11/6/1997</td>
<td>8/11/1999</td>
<td>1.8</td>
</tr>
<tr>
<td>bromfenac (DURACT)</td>
<td>7/15/1997</td>
<td>6/22/1998</td>
<td>0.9</td>
</tr>
<tr>
<td>mibebradil (POSICOR)</td>
<td>6/20/1997</td>
<td>6/8/1998</td>
<td>1</td>
</tr>
<tr>
<td>dexfenfluramine (REDUX)</td>
<td>6/1/1996</td>
<td>9/15/1997</td>
<td>1.3</td>
</tr>
</tbody>
</table>

**Strengthening the warnings to withdrawal for safety reasons**

**Propoxyphene** (Darvocet/Darvon): narcotic pain-killer withdrawn in 2010 because of increased risk of heart attacks and stroke.

**NSAIDs**: Change the labels from “may cause” to “cause an increased risk” of serious heart failure.
## Patent Expiry & Loss of Exclusivity

**Top 10 patent losses of 2018** (worth nearly $26 billion in U.S. sales last year.)

<table>
<thead>
<tr>
<th>Drug</th>
<th>Company</th>
<th>2017 U.S. sales (billion)</th>
<th>Disease/Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rituxan</td>
<td>Roche</td>
<td>$4.41</td>
<td>Blood cancers, rheumatoid arthritis</td>
</tr>
<tr>
<td>Neulasta</td>
<td>Amgen</td>
<td>$3.93</td>
<td>White blood cell booster</td>
</tr>
<tr>
<td>Lyrica</td>
<td>Pfizer</td>
<td>$3.46</td>
<td>Nerve and muscle pain</td>
</tr>
<tr>
<td>Advair</td>
<td>GlaxoSmithKline</td>
<td>$2.23</td>
<td>Asthma and COPD</td>
</tr>
<tr>
<td>Xolair</td>
<td>Roche/Novartis</td>
<td>$1.83</td>
<td>Allergic asthma and chronic idiopathic urticaria</td>
</tr>
<tr>
<td>Epogen/Procrit</td>
<td>Amgen and Johnson &amp; Johnson</td>
<td>$1.77</td>
<td>Anemia</td>
</tr>
<tr>
<td>Restasis</td>
<td>Allergan</td>
<td>$1.41</td>
<td>Dry eye</td>
</tr>
<tr>
<td>Cialis</td>
<td>Eli Lilly</td>
<td>$1.36</td>
<td>Erectile dysfunction</td>
</tr>
<tr>
<td>Sensipar</td>
<td>Amgen</td>
<td>$1.35</td>
<td>Calcium reducer</td>
</tr>
<tr>
<td>Zytiga</td>
<td>Johnson &amp; Johnson</td>
<td>$1.23</td>
<td>Prostate cancer</td>
</tr>
</tbody>
</table>

Pharmacogenomics
Pharmacogenomics

- Refers to the effects of genetic variants drug response – efficacy, side-effects.
- Pharmacogenomics can help in selection of the optimal drug, dose, treatment regimens, & and avoid potential drug-induced adverse drug reactions.
  - Responders, non-responders, or responding adversely to medications
- FDA drug labeling may contain information on genomic biomarkers:
  - Drug exposure and clinical response variability
  - Risk for adverse events
  - Genotype-specific dosing
  - Mechanisms of drug action
  - Polymorphic drug target and disposition genes
  - Trial design features
### Pharmacogenomic Biomarkers in Drug Labeling

<table>
<thead>
<tr>
<th>Drug</th>
<th>Therapeutic Area</th>
<th>Biomarker</th>
<th>Labeling Sections</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abacavir</td>
<td>Infectious Diseases</td>
<td>HLA-B</td>
<td>Boxed Warning, Dosage and Administration, Contraindications, Warnings and Precautions</td>
</tr>
<tr>
<td>Abemaciclib (1)</td>
<td>Oncology</td>
<td>ESR</td>
<td>Indications and Usage, Adverse Reactions, Clinical Studies</td>
</tr>
<tr>
<td>Abemaciclib (2)</td>
<td>Oncology</td>
<td>EREB2 (HER2)</td>
<td>Indications and Usage, Adverse Reactions, Clinical Studies</td>
</tr>
<tr>
<td>Ado-Trastuzumab</td>
<td>Oncology</td>
<td>EREB2 (HER2)</td>
<td>Indications and Usage, Adverse Reactions, Clinical Studies</td>
</tr>
<tr>
<td>Emtrastine</td>
<td>Oncology</td>
<td>EREB2 (HER2)</td>
<td>Indications and Usage, Adverse Reactions, Clinical Studies</td>
</tr>
<tr>
<td>Afatinib</td>
<td>Oncology</td>
<td>EGFR</td>
<td>Indications and Usage, Adverse Reactions, Clinical Studies</td>
</tr>
<tr>
<td>Alectinib</td>
<td>Oncology</td>
<td>ALK</td>
<td>Indications and Usage, Adverse Reactions, Clinical Studies</td>
</tr>
<tr>
<td>Amtriptyline</td>
<td>Psychiatry</td>
<td>CYP2D6</td>
<td>Indications and Usage, Adverse Reactions, Clinical Studies</td>
</tr>
<tr>
<td>Anastrozole</td>
<td>Oncology</td>
<td>ESR, PGR</td>
<td>Indications and Usage, Adverse Reactions, Clinical Studies</td>
</tr>
</tbody>
</table>

**Boxed Warning**

**WARNING HYPERSENSITIVITY REACTIONS, and LACTIC ACIDOSIS, AND SEVERE HEPATOMEGALY**

Serious and sometimes fatal hypersensitivity reactions, with multiple organ involvement, have occurred with ZIAGEN® (abacavir). Patients who carry the HLA-B*5701 allele are at a higher risk of a hypersensitivity reaction to abacavir; although, hypersensitivity reactions have occurred in patients who do not carry the HLA-B*5701 allele [see Warnings and Precautions (5.1)]. ZIAGEN is contraindicated in patients with a prior hypersensitivity reaction to abacavir and in HLA-B*5701-positive patients [see Contraindications (4), Warnings and Precautions (5.1)]. All patients should be screened for the HLA-B*5701 allele prior to initiating therapy with ZIAGEN or remittance of therapy with ZIAGEN, unless patients have a previously documented HLA-B*5701 allele assessment. Discontinue ZIAGEN immediately if a hypersensitivity reaction is suspected, regardless of HLA-B*5701 status and even when other diagnoses are possible [see Contraindications (4), Warnings and Precautions (5.1)]. Following a hypersensitivity reaction to ZIAGEN, NEVER restart ZIAGEN or any other abacavir-containing product because more severe symptoms, including death can occur within hours. Similar severe reactions have also occurred rarely following the reintroduction of abacavir-containing products in patients who have no history of abacavir hypersensitivity [see Warnings and Precautions (5.1)].

**2 DOSAGE AND ADMINISTRATION**

2.1 Screening for HLA-B*5701 Allele Prior to Starting ZIAGEN

Screen for the HLA-B*5701 allele prior to initiating therapy with ZIAGEN [see Boxed Warning, Warnings and Precautions (5.1)].

**4 CONTRAINDICATIONS**

ZIAGEN is contraindicated in patients:
- who have the HLA-B*5701 allele [see Warnings and Precautions (5.1)].
Pharmacogenomics Database

What is PharmGKB?

- NIH-funded resource – Established in 2000 – one of the first “post-genomic” database for the storage and curation of genotype-phenotype data from pharmacogenomics studies.

  - Provides information about how human genetic variation affects response to medications.

  - Collects, curates, and disseminates clinically actionable gene-drug associations and genotype-phenotype relationships.
PharmGKB Clinical Annotations: Curated variant-drug pairs based on variant annotations in the database.
Atorvastatin/Lovastatin/Simvastatin Pathway, Pharmacokinetics

Summary
Drug-specific representation of the candidate genes involved in transport, metabolism and clearance.

Annotations Data
Downloads contain information from PharmGKB annotations.

- Variant and Clinical Annotations Data
  Clinical & Variant annotations summary.
  - annotations.zip 3.81 MB

- Dosing Guidelines
  Detailed dosing guidelines in JSON format:
  - dosingGuidelines.json.zip 1.2 MB

- Pathways
  Pathways data in BioPax XML and TSV formats:
  - pathways-biopax.zip 524.76 KB
  - pathways.tsv.zip 73.57 KB

- Literature Occurrence
  A list of objects that occur in PharmGKB literature annotations and pathways.
  - occurrences.zip 1.62 MB
PharmGenEd - Pharmacogenomics education program
http://pharmacogenomics.ucsd.edu

Objectives:
• Bridge the gap between pharmacogenomics and practice - program designed for healthcare professionals and trainees.
• Increase awareness of validity & usage of pharmacogenomics tests and their implications on their therapeutic use.
  ➢ Web-based presentations
  ➢ A shared curriculum platform to "Train-the-Trainer"
  ➢ Live presentations
  ➢ Written articles
Pharmacogenomics – Benefits

• Disease risk stratification: Knowing the genetic risk factors, disease susceptibility – lifestyle and environmental changes, more informed therapeutic regimens

• Drug discovery – Improvements: “lead” identification and optimization; Resurrecting previously failed drugs (unselected disease population vs niche population)

• Customized therapeutic regimens – personalized medicine: Modified paradigm of dosing taking into account not just age/weight but also genetic composition

• Economics – cost of healthcare: Potential lowering of healthcare costs
Bioinformatics – Leveraging omics big-data
Role of bioinformatics – Drug discovery

• Traditional approach - pharmacology and chemistry-based drug discovery approaches – laborious, expensive, several challenges in finding new drugs
• Highly competitive - “winner takes all”
• Need for accelerating any phase of the drug development
• Increasing pressure to generate more and more drugs in a short period of time with low risk has resulted in remarkable interest in bioinformatics
• Computer-aided drug design (CADD)
Role of bioinformatics – Drug discovery

• **Drug Target Identification** - Bioinformatics allows the identification and analysis of more and more biological drug targets - increases the number of potential drugs in the discovery pipelines

• **Drug Target Validation** - Helps to moderate the potential for failure in the clinical testing and approval phases

• **Reducing Costs** - Accelerate drug discovery process
  - Drug target identification and validation (e.g., Docking)
  - Assay development
  - Virtual-high-throughput screening (v-HTS)

• **Enable Novel Drug Discovery & Development** – Data-driven approaches, hypotheses generation.

**Drug discovery cycle → Data-intensive**
Drugbank - https://www.drugbank.ca/

- Bioinformatics and cheminformatics resource
- Detailed drug data + comprehensive drug target information.
- Current release (version 5.1.1, 2018-07-03):
  11,885 drug entries:
  - 2,529 approved small molecule drugs,
  - 1,184 approved biotech (protein/peptide) drugs,
  - 129 nutraceuticals and
  - >5,755 experimental drugs.
  - 5,131 non-redundant protein (i.e. drug target or enzyme or transporter or carrier) sequences are linked to these drug entries.
- Each DrugCard entry contains more than 200 data fields with half of the information being devoted to drug/chemical data and the other half devoted to drug target or protein data.
<table>
<thead>
<tr>
<th>DRUG</th>
<th>INTERACTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abemaciclib</td>
<td>The metabolism of Simvastatin can be decreased when combined with Abemaciclib.</td>
</tr>
<tr>
<td>Acetaminophen</td>
<td>The metabolism of Simvastatin can be increased when combined with Acetaminophen.</td>
</tr>
<tr>
<td>Acetylcysteine</td>
<td>The excretion of Simvastatin can be decreased when combined with Acetylcysteine.</td>
</tr>
<tr>
<td>Acipimox</td>
<td>Acipimox may increase the myopathic rhabdomyolysis activities of Simvastatin.</td>
</tr>
<tr>
<td>Adenine</td>
<td>The metabolism of Simvastatin can be decreased when combined with Adenine.</td>
</tr>
<tr>
<td>Afatinib</td>
<td>The serum concentration of Simvastatin can be increased when it is combined with Afatinib.</td>
</tr>
<tr>
<td>Albendazole</td>
<td>The metabolism of Simvastatin can be decreased when combined with Albendazole.</td>
</tr>
<tr>
<td>Aldosterone</td>
<td>The metabolism of Simvastatin can be decreased when combined with Aldosterone.</td>
</tr>
<tr>
<td>Alectinib</td>
<td>The serum concentration of Simvastatin can be increased when it is combined with Alectinib.</td>
</tr>
<tr>
<td>Alendronic acid</td>
<td>The risk or severity of myopathy, rhabdomyolysis, and myoglobinuria can be increased when Alendronic acid is combined with Simvastatin.</td>
</tr>
</tbody>
</table>
### 3-hydroxy-3-methylglutaryl-coenzyme A reductase

**Kind**: Protein  
**General Function**: Nadph binding

**Organism**: Human  
**Specific Function**: Transmembrane glycoprotein that is the rate-limiting enzyme in cholesterol biosynthesis as well as in the biosynthesis of nonsterol isoprenoids that are essential for normal cell function including...

**Pharmacological action**: Yes

**Actions**: Inhibitor

**Gene Name**: HMGCR

**Uniprot ID**: P04035

**Uniprot Name**: 3-hydroxy-3-methylglutaryl-coenzyme A reductase

**Molecular Weight**: 97475.155 Da

### References

DrugCentral - http://drugcentral.org

Online drug information resource - provides information on active ingredients chemical entities, pharmaceutical products, drug mode of action, indications, pharmacologic action.

<table>
<thead>
<tr>
<th>Entity</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Active Ingredients</td>
<td>4,531</td>
</tr>
<tr>
<td>Small molecule</td>
<td>3,807</td>
</tr>
<tr>
<td>Biologic</td>
<td>279</td>
</tr>
<tr>
<td>Other</td>
<td>445</td>
</tr>
<tr>
<td>FDA drug labels</td>
<td>77,484</td>
</tr>
<tr>
<td>Rx drug labels</td>
<td>34,192</td>
</tr>
<tr>
<td>OTC drug labels</td>
<td>43,292</td>
</tr>
<tr>
<td>Pharmaceutical formulations in FDA drug labels</td>
<td>97,271</td>
</tr>
</tbody>
</table>
## FDA Adverse Event Reporting System

<table>
<thead>
<tr>
<th>MedDRA adverse event term</th>
<th>Likelihood ratio</th>
<th>Likelihood ratio threshold</th>
<th>Patients taking drug having adverse event</th>
<th>Patients taking drug not having adverse event</th>
<th>Patients not taking drug having adverse event</th>
<th>Patients not taking drug not having adverse event</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rhabdomyolysis</td>
<td>13758.02</td>
<td>28.60</td>
<td>2994</td>
<td>14654</td>
<td>12735</td>
<td>3355476</td>
</tr>
<tr>
<td>Myalgia</td>
<td>7892.50</td>
<td>28.60</td>
<td>2215</td>
<td>15433</td>
<td>29860</td>
<td>3338351</td>
</tr>
<tr>
<td>Blood creatine phosphokinase increased</td>
<td>4915.20</td>
<td>28.60</td>
<td>1268</td>
<td>16380</td>
<td>11745</td>
<td>3356466</td>
</tr>
<tr>
<td>Disease</td>
<td>Relation</td>
<td>SNOMED_ID</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------------------------------------------</td>
<td>----------</td>
<td>-----------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hypercholesterolemia</td>
<td>Indication</td>
<td>13644009</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arteriosclerotic vascular disease</td>
<td>Indication</td>
<td>72092001</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hypocalalphoproteinemias</td>
<td>Indication</td>
<td>190786000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cerebrovascular accident</td>
<td>Indication</td>
<td>230690007</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Familial hypercholesterolemia - homozygous</td>
<td>Indication</td>
<td>238078005</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Familial hypercholesterolemia - heterozygous</td>
<td>Indication</td>
<td>238079002</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mixed hyperlipidemia</td>
<td>Indication</td>
<td>267434003</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disease of liver</td>
<td>Contraindication</td>
<td>235858003</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rhabdomyolysis</td>
<td>Contraindication</td>
<td>240818008</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pregnancy, function</td>
<td>Contraindication</td>
<td>289008002</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thrombocytopenic disorder</td>
<td>Contraindication</td>
<td>302215000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Surgical procedure</td>
<td>Contraindication</td>
<td>387713003</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acute coronary syndrome</td>
<td>Contraindication</td>
<td>394859003</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breastfeeding (mother)</td>
<td>Contraindication</td>
<td>413712001</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disorder of coronary artery</td>
<td>Contraindication</td>
<td>414024009</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Slco1b1 gene variant for CYP1b1 transporter affecting hepatic clearance</td>
<td>Contraindication</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Myopathy related to Slco1b1 gene variant</td>
<td>Contraindication</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Myocardial infarction prevention</td>
<td>Off-label Use</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary Prevention of Coronary Heart Disease</td>
<td>Off-label Use</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
DGIdb - http://www.dgidb.org
ChEMBL - https://www.ebi.ac.uk/chembl

- Manually curated database of bioactive molecules with drug-like properties.
- Chemical, bioactivity and genomic data
Inxight: Drugs
https://drugs.ncats.io/ginas/app

• Incorporates and unifies a wealth of data, including manually curated data supplied by the FDA and private companies
• Provides marketing and regulatory status
• Rigorous drug ingredient definitions, biological activity, clinical use, etc.
Helping an old dog perform new tricks

Drug Repositioning/Repurposing
New Drug Development - Problems

- **Expensive**
- **Time consuming**
- **High attrition** - Only five in 5,000 compounds that enter preclinical testing make it to human testing. Only 1 of these 5 tested in humans is approved.
- **Post-marketing drug failure**
- **Decreased return on investment** by 50% in 10 yrs – patent expiry (generics, etc.)
Drug Repositioning or Drug Repurposing

• Also referred to as drug reprofiling or drug retasking

• Search for potential new therapeutic applications of existing (approved) compounds

• Reinvestigation of drug candidates that have not succeeded in previous advanced clinical trials, for reasons other than safety
Drug repositioning – Benefits & Examples

- Reduction of time and costs - Since the drug is already approved – initial timeline can be bypassed (1.5 to 2 years of preclinical and Phase I development time)
- Better/smart resource utilization
- De-Risking - lower development risk for investor
- Relatively lower patient risk – “known” drug-related adverse effects

**Sildenafil** (Viagra): From failed antihypertensive to erectile dysfunction and to orphan disease

**Thalidomide**: From a dangerous drug to a promising start

**Azidothymidine**: Anti-cancer to AIDSs

**Ropinirole** and **Pramipexole**: Parkinson’s Drugs for Restless Legs Syndrome

**Clioquinol**: Antiprotozoal as a lead compound for neuroprotection

**Finasteride**: Prostate cancer to baldness/hair loss
Computational Drug repositioning

• Can we make drug discovery (drug repositioning) systematically serendipitous?
• Formulate testable hypothesis from serendipitous discoveries
Computational drug repositioning

Two principles:

“similar drugs” → similar therapeutic effects

“similar diseases” → treated with the same drugs

Computational approach: Measuring similarity with diverse data
Connecting small molecules to diseases

- Animal Model
- Differentially expressed genes
- Perturbagen signature
- NCBI GEO
- Human Patients
- LINCS

Signature-based

- Small Molecule Perturbagen
- Signature-based

Target- & Knowledge-based

- OMIM
- GAD
- GWAS
- PPI
- Gene Ontology
- Pathways
- Phenotype
- Literature
- Approved Drugs-Targets
- Druggable Genome

Disease

- Reverse signature
to disease
Potential therapeutic
- Similar signature
to disease
Potential indicator

10/24/2018
L1000 Signatures (http://www.lincscloud.org)

Library of Integrated Cellular Signatures (LINCS)

- Bead-based, high-throughput gene expression assay
- Cultured cells treated with various chemical and genetic perturbations
- Measure the corresponding transcriptional responses

**Compute connections between user-submitted signatures (e.g., disease signatures) and LINCS data**
L1000 Signatures (http://www.lincscloud.org)

**20,413 Small-molecule compounds**
- Selected from multiple sources
- Include known drugs
- Pathway-specific tool compounds
- Compounds in NIH-sponsored small-molecule screening efforts
- Nominated by members of the research community
  - ~1,300 FDA-approved drugs
  - ~5,585 bioactive tool compounds

**22,119 Genetic constructs for knocking-down genes (shRNA) or over-expressing genes (cDNA)**
- Known targets of FDA-approved drugs
- Drug-target pathway members
- Associated with disease
- Nominated by the research community
  - ~900 targets/pathways of FDA-approved drugs
  - ~600 candidate disease genes
  - 500+ community nominations
Compound signatures
- ~20k compounds
- 72 cell lines
- 2 time points – 6 hr and 24 hr
- Dose: different doses
- 1.4 million gene expression experiments

~200,000 signatures
~7000 genes
7000 X 200,000 matrix
Top 100 and bottom 100 probe sets/genes
Original Connectivity Map Concept

BIOLOGICAL STATE OF INTEREST (SIGNATURE)

REFERENCE DATABASE (PROFILES)

CONNECTIONS

up

query

down

strong positive

weak positive

null

strong negative

output

positive

negative

Lamb et al., 2006, Science
# Transcriptomics data – NCBI GEO

## Total holdings

<table>
<thead>
<tr>
<th></th>
<th>Public</th>
<th>Unreleased</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Series</td>
<td>97,837</td>
<td>11,712</td>
<td>109,549</td>
</tr>
<tr>
<td>Platforms</td>
<td>18,473</td>
<td>246</td>
<td>18,719</td>
</tr>
<tr>
<td>Samples</td>
<td>2,481,196</td>
<td>369,433</td>
<td>2,850,629</td>
</tr>
</tbody>
</table>

## Series type

| Series type                                                      | Count   |
|                                                               |         |
| Expression profiling by array                                   | 53,956  |
| Expression profiling by genome tiling array                     | 730     |
| Expression profiling by high throughput sequencing               | 18,639  |

## Organism

<table>
<thead>
<tr>
<th>Organism</th>
<th>Series</th>
<th>Platforms</th>
<th>Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homo sapiens</td>
<td>41,055</td>
<td>5,266</td>
<td>1,360,137</td>
</tr>
<tr>
<td>Mus musculus</td>
<td>28,389</td>
<td>2,300</td>
<td>547,601</td>
</tr>
</tbody>
</table>

10/24/2018
Integrating Pharmacogenomics & Bioinformatics

Case studies
Use Case 1: Cystic Fibrosis (CF) - Candidate therapeutic discovery
>2000 known CFTR mutations: Categorized into different groups

<table>
<thead>
<tr>
<th>Class</th>
<th>Description</th>
<th>CFTR Protein Status</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>CFTR is created, reaches cell surface and functions properly, allowing transfer of chloride and water.</td>
<td>Functional</td>
<td>G542X, W1282X, R553X</td>
</tr>
<tr>
<td>Class I</td>
<td>No functional CFTR created.</td>
<td>Misfolded</td>
<td>F508del, E1103K, I507del</td>
</tr>
<tr>
<td>Class II</td>
<td>CFTR protein is created, but misfolded, keeping it from reaching the cell surface.</td>
<td>Partially functional</td>
<td>G551D, S549N, V520F</td>
</tr>
<tr>
<td>Class III</td>
<td>CFTR protein is created and reaches cell surface, but does not function properly.</td>
<td>Non-functional</td>
<td>R117H, D1152H, R347P</td>
</tr>
<tr>
<td>Class IV</td>
<td>The opening in the CFTR protein ion channel is faulty.</td>
<td>Insufficient</td>
<td>3849+10kbc-&gt;T, 2789+5g-&gt;a, A455E</td>
</tr>
</tbody>
</table>

‘severe’ mutations
pancreatic insufficiency
decreased survival

‘mild’ mutations
pancreatic sufficiency

Three approved drugs

KALYDECO (Ivacaftor) – Potentiator – Restore gating defects
• CF patients age 2 years and older who have one of the following mutations in their CFTR gene: G551D, G1244E, G1349D, G178R, G551S, S1251N, S1255P, S549N, R117H, or S549R.
• Not for use in people with CF due to other mutations in the CFTR gene.
• Not effective in patients with CF with two copies of the F508del mutation (F508del/F508del) in the CF gene.

ORKAMBI (Lumacaftor+Ivacaftor) - Corrector (restores protein folding) + Potentiator
• CF patients age 2 years and older
• Have two copies of the F508del mutation (F508del/F508del) in their CFTR gene.
• Should not be used in patients other than those who have two copies of the F508del mutation in their CFTR gene.

SYMDEKO (Tezacaftor+Ivacaftor) - Corrector (restores protein folding) + Potentiator
• CF patients age 12 years and older
• Have two copies of the F508del mutation (F508del/F508del) in their CFTR gene or with one mutation that responds to tezacaftor/ivacaftor.

10/24/2018
Three approved drugs - Limitations

• In vitro studies have shown that the approved combinatorial has limited clinical efficacy

• Potential interference of potentiators with corrector actions and destabilizing of corrected ΔF508-CFTR

• Pseudomonas infections – interference

• Will not work for all patients with CF

• Rare mutations

• Additional/New Side-effects? – Patient use

• Costs $250-300K per year’s supply
### Published CF-related gene expression data sets

<table>
<thead>
<tr>
<th>GEO/Array Express Accession No.</th>
<th>Samples</th>
<th>Year</th>
<th>Publication</th>
</tr>
</thead>
<tbody>
<tr>
<td>GSE15568</td>
<td>16 CF and 13 nonCF&lt;br&gt;Rectal suction specimens</td>
<td>2013</td>
<td>Stanke et al.</td>
</tr>
<tr>
<td>GSE39843</td>
<td>3 Control (NuLi - no-Cystic Fibrosis deltaF508 mutation) vs 3 CF (CuFi cell line)</td>
<td>2012</td>
<td>Voisin et al.</td>
</tr>
<tr>
<td>GSE26482</td>
<td>Gene expression of immortalized cystic fibrosis and non-cystic fibrosis airway epithelial cells</td>
<td>2011</td>
<td>GSE26482</td>
</tr>
<tr>
<td>GSE10406</td>
<td>Sinus mucosa; 9 control vs 6 CF</td>
<td>2008</td>
<td>GSE10406</td>
</tr>
<tr>
<td>E-MEXP-980</td>
<td>comparison of expression profiles between a CF cell line and a control cell line; 3 control and 3 CF (tracheal epithelia)</td>
<td>2007</td>
<td>Verhaeghe et al.</td>
</tr>
<tr>
<td>GSE2395</td>
<td>Control vs Severe&lt;br&gt;Respiratory epithelial gene expression in patients with severe CF lung and control 11 control and 3 severe CF</td>
<td>2005</td>
<td>Wright et al.</td>
</tr>
</tbody>
</table>

**Approaches that incorporate multiple disease-related available transcriptional knowledge should produce improved and robust drug predictions**
In vitro assays  
(in collaboration with Naren Lab, CCHMC)

- Forskolin induces rapid swelling of organoids derived from healthy controls or wild-type mice
- Swelling strongly reduced in organoids of CF patients or in mice carrying the Cftr F508del mutation and is absent in Cftr-deficient organoids

Forskolin is commonly used to increase the amount of intracellular cyclic AMP (cAMP) and cAMP activates CFTR

Dekkers et al., 2013, Nature Medicine
Monitoring CFTR-mediated fluid secretion using Phase-contrast time lapse microscopy

* Calculation of fluid secretion

Volume of a sphere

\[ V = \frac{4}{3} \pi r^3 \]

Fluid secretion (\%) = \( \frac{\text{Luminal}}{\text{Enterosphere}} \times 100 \)

<table>
<thead>
<tr>
<th>Experimental conditions</th>
<th>FSK</th>
<th>Compound</th>
<th>VX809</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-FSK (DMSO): Compound-VX809</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Post-FSK: Compound-VX809</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Pre-FSK (DMSO): Compound+VX809</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Post-FSK: Compound+VX809</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

Fluid secretion

- 9.145 %
- 32.395 %
- 63.902 %

*Proposed experiments for drug repositioning candidates*
CF patients differentially expressed genes

NCBI GEO

Small molecules-specific gene signatures

LINCS

Connectivity Mapping

Small molecule hits

Hit prioritization

CF Knowledgebase

CF Candidate Therapeutics

Pre-clinical Validation

- ΔF508-CFTR homozygous CF patients
  - FIS Assay – Intestinal organoids
  - Primary bronchial epithelial cells - Short-circuit current ($I_{sc}$) measurements
- HEK 293 cells overexpressing ΔF508 CFTR
  - Western blot
- RNA-Seq – Differentially expressed genes – pre- and post- treatment

~20K Compound Library

Signature reversal

184 Compounds significantly “reversing” CF gene signature

24 Compounds enriched for CF-relevant pathways &/or CFTR interactome (WT/ΔF508)

2 Compounds selected for combinatorial tests (with known correctors and potentiators and CFTR-INH)
New Results

PP-2, a src-kinase inhibitor, is a potential corrector for F508del-CFTR in cystic fibrosis

Yunguan Wang, Kavisha Arora, Fanmuyi Yang, Woong-Hee Shin, Jing Chen, Daisuke Kihara, Anjaparavanda P Naren, Anil G Jegga

doi: https://doi.org/10.1101/288324
Use Case 2: Drug-induced Adverse Events
AERSMine: a phenome+pharmacome clinical datamine

AERSMine is a web application for mining FDA’s Adverse Event Reporting System data. AERSMine generates analyzable data matrices that can be filtered, clustered, and scored by a variety of approaches including the WHO signal reporting algorithms to identify unexpectedly high-risk subgroups. Our long-term hypothesis is that by correlation of adverse reactions with known drug-phenotype-gene relationships, we will improve our ability to modify therapeutic strategies and improve therapeutic efficacy.


- >2 million cases of prescription drug-related adverse events (AE) occur annually, including ~100,000 deaths.
- 4.2 - 30% of hospital admissions in the USA and Canada
- 11.4 - 35.5% of emergency department visits in older adults are due to drug-related causes
- The impact and the management of ADRs is complex - ~30 billion dollars annually.

Sultana et al., 2013, J Pharmacol Pharmacother.
Rich Segmentation ⊢ High Resolution analyses

1. Between Class effects
2. Within Class effects
3. Identify Signatures

10.6 million patient reports representing clinical drug effects across a wide variety of population subgroups (2004-2018q1)

~300,000 new reports added each year.
A. Search or browse for indications, classes of drugs, outcomes

User-driven choice of any or all from Drugs (sets or classes), Clinical Indications (specific, general, class-based), Demographics, Outcomes

B. Query or select from ontology classes of drugs and/or indication groups

C. Save/name each cohort for comparison; Rx(s), +/- indication(s), demographic(s)

Hypothesis: Differential occurrences and correlations of outcomes reflects mixed effects of intrinsic cohort risks and differential drug actions

Cohort quantitation metric: absolute counts, per 1000 patients, relative risk, safety signal
## Differential response to drugs

<table>
<thead>
<tr>
<th>Rank</th>
<th>Known Drug Reactions</th>
<th>Adverse Events</th>
<th>Total Adverse Events Reports</th>
<th>Relative Risk ARBs non-cancer aggregated</th>
<th>Relative Risk ACE non-cancer aggregated</th>
<th>Relative Risk thiazides non-cancer aggregated</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>renal failure acute</td>
<td>60,084</td>
<td>3.176</td>
<td>4.732</td>
<td>1.937</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>hyperkalaemia</td>
<td>14,895</td>
<td>5.669</td>
<td>9.493</td>
<td>1.594</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>hypertension</td>
<td>87,361</td>
<td>2.473</td>
<td>2.354</td>
<td>2.756</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>blood pressure increased</td>
<td>52,843</td>
<td>3.848</td>
<td>2.295</td>
<td>2.716</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>atrial fibrillation</td>
<td>39,484</td>
<td>2.664</td>
<td>3.823</td>
<td>2.313</td>
</tr>
<tr>
<td>6</td>
<td>3</td>
<td>hypotension</td>
<td>73,593</td>
<td>2.297</td>
<td>3.005</td>
<td>1.765</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
<td>blood creatinine increased</td>
<td>33,018</td>
<td>3.723</td>
<td>3.081</td>
<td>2.108</td>
</tr>
<tr>
<td>8</td>
<td>3</td>
<td>hyponatraemia</td>
<td>24,150</td>
<td>3.500</td>
<td>3.168</td>
<td>6.625</td>
</tr>
<tr>
<td>9</td>
<td>2</td>
<td>drug interaction</td>
<td>62,052</td>
<td>1.842</td>
<td>2.922</td>
<td>2.058</td>
</tr>
<tr>
<td>10</td>
<td>2</td>
<td>coronary artery disease</td>
<td>26,759</td>
<td>1.621</td>
<td>4.283</td>
<td>2.639</td>
</tr>
<tr>
<td>11</td>
<td>2</td>
<td>blood urea increased</td>
<td>13,518</td>
<td>4.546</td>
<td>3.513</td>
<td>2.795</td>
</tr>
<tr>
<td>12</td>
<td>3</td>
<td>bradycardia increased</td>
<td>22,513</td>
<td>2.469</td>
<td>3.590</td>
<td>1.871</td>
</tr>
<tr>
<td>12</td>
<td>2</td>
<td>renal failure chronic</td>
<td>11,778</td>
<td>3.440</td>
<td>4.069</td>
<td>2.603</td>
</tr>
<tr>
<td>14</td>
<td>2</td>
<td>cardiomegaly</td>
<td>12,124</td>
<td>2.409</td>
<td>4.097</td>
<td>2.941</td>
</tr>
<tr>
<td>15</td>
<td>2</td>
<td>renal impairment</td>
<td>29,644</td>
<td>3.104</td>
<td>2.651</td>
<td>1.573</td>
</tr>
</tbody>
</table>

*Indicates on-label event for corresponding drugs/class*
Drug-induced GI Disorders (DIGID):
- Patient without previous GI disease develops GI symptoms, deterioration of GI function, etc. in association with drug therapy
- Underdiagnosed major safety concern
  - ~20-40% of GI tract disorders are drug-induced.
- **Current treatment**: stopping the drug

**Unpublished 10/24/2018**
Concomitant use of roflumilast/apremilast (PDE4 inhibitors) with tiotropium bromide significantly mitigated the DIGID risk ($p \leq 0.05$, two-tailed Mann-Whitney-Wilcoxon test) in comparison with baseline roflumilast/apremilast monotherapy. Similar trend is seen for irinotecan and solifenacin combinatorial.
Intersection between drug targets, genes over expressed in GI tissue, & pharmacogenes associated with drug-induced AEs.

ABC1, ABCC2, ABCG2, ACE, CPS1, CYP2C19, CYP2D6, CYP3A4, CYP3A5, CYP4F2, DDC, FOLH1, GIPR, HNF4A, HTR1B, MME, NAGS, NAT2, NOS2, NQO1, NR1I2, SLC10A2, SLC28A1, SLC28A2, SLC6A4, SULT1A2, TNFRSF11A, UGT1A1, XD

<table>
<thead>
<tr>
<th>Variant</th>
<th>Gene</th>
<th>Type</th>
<th>Chemicals</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs6785049</td>
<td>NR1I2</td>
<td>Toxicity/ADR</td>
<td>sunitinib</td>
</tr>
<tr>
<td>rs3732360</td>
<td>NR1I2</td>
<td>Toxicity/ADR</td>
<td>docetaxel</td>
</tr>
<tr>
<td>rs3732359</td>
<td>NR1I2</td>
<td>Toxicity/ADR</td>
<td>docetaxel</td>
</tr>
<tr>
<td>rs3814058</td>
<td>NR1I2</td>
<td>Toxicity/ADR</td>
<td>docetaxel</td>
</tr>
<tr>
<td>rs3814055</td>
<td>NR1I2</td>
<td>Toxicity/ADR</td>
<td>flucloxacillin</td>
</tr>
<tr>
<td>rs3814055</td>
<td>NR1I2</td>
<td>Toxicity/ADR</td>
<td>sirolimus, temsirolimus</td>
</tr>
<tr>
<td>rs6785049</td>
<td>NR1I2</td>
<td>Toxicity/ADR</td>
<td>sirolimus, temsirolimus</td>
</tr>
<tr>
<td>rs8175347</td>
<td>UGT1A1</td>
<td>Toxicity/ADR</td>
<td>atazanavir, ritonavir</td>
</tr>
<tr>
<td>rs1042640</td>
<td>UGT1A1</td>
<td>Toxicity/ADR</td>
<td>atazanavir, ritonavir</td>
</tr>
<tr>
<td>rs10929303</td>
<td>UGT1A1</td>
<td>Toxicity/ADR</td>
<td>atazanavir, ritonavir</td>
</tr>
<tr>
<td>rs8330</td>
<td>UGT1A1</td>
<td>Toxicity/ADR</td>
<td>atazanavir, ritonavir</td>
</tr>
<tr>
<td>rs8175347</td>
<td>UGT1A1</td>
<td>Toxicity/ADR</td>
<td>deferasirox</td>
</tr>
<tr>
<td>rs4148323</td>
<td>UGT1A1</td>
<td>Toxicity/ADR</td>
<td>deferasirox</td>
</tr>
<tr>
<td>rs4124874</td>
<td>UGT1A1</td>
<td>Toxicity/ADR</td>
<td>deferasirox</td>
</tr>
<tr>
<td>rs4148323</td>
<td>UGT1A1</td>
<td>Toxicity/ADR</td>
<td>indinavir</td>
</tr>
<tr>
<td>rs8175347</td>
<td>UGT1A1</td>
<td>Toxicity/ADR</td>
<td>indinavir</td>
</tr>
<tr>
<td>rs8175347</td>
<td>UGT1A1</td>
<td>Toxicity/ADR</td>
<td>irinotecan</td>
</tr>
<tr>
<td>rs4124874</td>
<td>UGT1A1</td>
<td>Toxicity/ADR</td>
<td>irinotecan</td>
</tr>
<tr>
<td>rs8175347</td>
<td>UGT1A1</td>
<td>Toxicity/ADR</td>
<td>nilotinib</td>
</tr>
<tr>
<td>rs8175347</td>
<td>UGT1A1</td>
<td>Toxicity/ADR</td>
<td>sorafenib</td>
</tr>
<tr>
<td>rs8175347</td>
<td>UGT1A1</td>
<td>Toxicity/ADR</td>
<td>tranilast</td>
</tr>
</tbody>
</table>

Genes NR1I2 & UGT1A1 are overexpressed in GI tissue.

Known variants in NR1I2 and UGT1A1 (along with drugs) associated with drug toxicity or drug-induced AE.
Summary

- **Pharmacogenomics**
  - Resources – PharmGKB and FDA Drug Labels

- **Bioinformatics – Leveraging omics big-data**
  - Resources: DrugBank, DrugCentral, DGIdb, ChEMBL, Inxight
  - Drug repositioning/repurposing

- **Integrating Pharmacogenomics & Bioinformatics – Case studies**
  - Cystic fibrosis – Repurposing existing CF patient gene expression data for novel candidate CF corrector discovery
  - Drug-induced GI disorders – Integrating heterogeneous biomedical and genomic big data – Hypothesis generation
Acknowledgements

**Jegga Lab**
Yunguan Wang, Ph.D.
Surya Yaddanapudi
Jaswanth Yella
Sudhir Ghandikota

**Cystic Fibrosis**
Anjaparavanda P Naren, Ph.D.
John P. Clancy, M.D.
Kavisha Arora, Ph.D.
*CF Clinical Translational Research Division of Pulmonary Medicine*

**Drug-induced Adverse Events**
Mayur Sarangdhar, Ph.D.

---

Cincinnati Children’s

National Center for Advancing Translational Sciences

University of Cincinnati

National Heart, Lung, and Blood Institute
Connect the dots/events:

Creatively, Comprehensively, Computationally

Thank you!