Cheminformatics in drug discovery and public health
Progress and challenges ahead

José L Medina-Franco
DIFACQUIM research group
Department of Pharmacy
School of Chemistry, UNAM, Mexico

XXVIII Symposium on Bioinformatics and Computer-Aided Drug Discovery
May 26th, 2022
Outline

• Introduction
  – UNAM & research group
  – Chemoinformatics

• Progress in drug discovery and public health
  – Case study: Epigenetic drug discovery; inhibitors of DNA methyltransferases

• Cheminformatics: Challenges ahead

• Summary
National Autonomous University of Mexico

UNAM

- 350,000 students
  - 6,836 students: School of Chemistry
- 12,400 full-time professors
- 2,200 buildings in Mexico and abroad

Main campus in Mexico City

Extension schools abroad

School of Medicine
Main Library
Concert Hall
Soccer Stadium

UNAM - Tucson
Centro de Estudios Mexicanos

www.english.unam.mx
Research group DIFACQUIM
Computer-Aided Drug Design at School of Chemistry

Bachelors
Jocelyn Salazar
Alexis Flores
Claudia D. López
Alejandro Ochoa
Hassan Villegas

MSc
Liliam Martínez (UNAM)
Camila Garibay (UMSNH)
Alberto Zavala (UMSNH)
Eva Palma (UMSNH)
Co-tutor: Dr. Víctor M. Baizabal

Ph.D.
Fernanda Saldivar
Diana Prado
Bárbara Díaz
Ana Chávez
Emma Andrade
Alejandro Gonzalez
Edgar López (CINVESTAV)
Co-tutor: Dr. Carlos C. Rojas

@difacquim
www.difacquim.com
medinajl@unam.mx
Chemoinformatics

“All concepts and methods that are designed to interface theoretical and experimental efforts involving small molecules”

Drug Discovery Today 2004 9:13–14
Chemoinformatics and bioinformatics

Main focus

Gene $\leftrightarrow$ Protein $\leftrightarrow$ Drug $\leftrightarrow$ Lead

Computational representation

- Sequences
- 3D coordinates

- Molecular fingerprints
- Numerical descriptors.
Timeline and impact of chemoinformatics

Today

Applications in different areas

- Toxicity
- Natural products informatics
- Biomedical informatics
- Food science
- Epigenetic drug discovery
Progress in drug discovery and public health

Epigenetic drug discovery

Inhibitors of DNA methyltransferases
Epigenetics

Modifications that happen on chromosomes without alteration of the DNA sequence and that lead to a stable phenotype.

www.celgene.com/the-potential-of-epigenetics
Epigenetic targets

Histone methylation and acetylation are frequently dysregulated in cancer cells and other diseases.

Epigenetic protein families

- **Writers**
  - Histone acetyltransferases, Histone methyltransferases

- **Erasers**
  - Histone deacetylases, Lysine demethylases

- **Readers**
  - Bromodomains, chromodomains, PHD fingers, malignant brain tumour domains, Tudor domains, PWWP domains

3 active DNMTs in eukaryotes:
- DNMT3A
- DNMT3B
- DNMT1

DNA methyltransferases (DNMTs)

Methylation: first epigenetic modification linked to cancer

Nat Rev Drug Discov 2012 11:384
Increased interest of epigenetic drug discovery

Epigenetics and therapeutic potential

Epigenetic drug discovery
Epi-drugs in clinical use and clinical development

A. Drugs approved for clinical use.

B. Examples of small molecules in clinical development.
DNA Methylation

- Epigenetic change
- Addition of a CH$_3$ at C5 of cytosine
- Mediated by DNMTs

In cancer cells
- Overexpression of DNMTs
- Hypermethylation of tumor suppressor genes

*FEBS Letters* 2011 585:2078
DNA metiltransferases (DNMTs)

1. De novo DNA methylation
2. Maintenance DNA methylation

Cytosine → 5-methylcytosine

DNMT1 regulatory domain: NLS, CXXC, BAH1, BAH2
Catalytic domain: MTase

DNMT3A regulatory domain: PWWP, ADD, MTase

DNMT3A2 regulatory domain: PWWP, ADD, MTase

DNMT3B regulatory domain: PWWP, ADD, MTase

DNMT3L regulatory domain: ADD
Selected crystal structures of DNMTs

<table>
<thead>
<tr>
<th>PDB code</th>
<th>Type</th>
<th>Region</th>
<th>Res (Å)</th>
<th>Cofactor</th>
<th>Ligand</th>
</tr>
</thead>
<tbody>
<tr>
<td>3SWR</td>
<td>hDNMT1</td>
<td>Autoinhibitory linker, CXXC, BAH1/2, methyltransferase domain</td>
<td>2.49</td>
<td>SFG</td>
<td>5-methyl-2'-deoxycytidine</td>
</tr>
<tr>
<td>4DA4</td>
<td>mDNMT1</td>
<td>BAH1/2, methyltransferase domain</td>
<td>2.60</td>
<td>SAH</td>
<td>5-methyl-2'-deoxycytidine</td>
</tr>
<tr>
<td>3PTA</td>
<td>hDNMT1</td>
<td>Autoinhibitory linker, CXXC, BAH1/2, methyltransferase domain</td>
<td>3.60</td>
<td>SAH</td>
<td>5-methyl-2'-deoxycytidine</td>
</tr>
<tr>
<td>3PT6</td>
<td>mDNMT1</td>
<td>Autoinhibitory linker, CXXC, BAH1/2, methyltransferase domain</td>
<td>3.00</td>
<td>SAH</td>
<td>5-methyl-2'-deoxycytidine</td>
</tr>
<tr>
<td>3PT9</td>
<td>mDNMT1</td>
<td>BAH1/2, methyltransferase domain</td>
<td>2.50</td>
<td>SAH</td>
<td>5-methyl-2'-deoxycytidine</td>
</tr>
</tbody>
</table>
Demethylation compounds

Cytosine analogues (citotoxic effects)

- 5-azacytidine (R = OH)
- 5-aza-2'-deoxycytidine (R = H)
- EGCG
- Nanaomycin A
- Hydralazine

Virtual screening hits

- RG108
- NSC 14778
- NSC 319745
- Olsalazine

Synthesis

- SGI-1027
- 16
- 17
- 5
- 49

HTS

- SID 49645275
- SW155246
- Dichloro
- Laocacic acid A
- 2,3-Dimercaptosuccinic acid

Approved drugs

- GSK3685032
- DNMT1 IC₅₀ 36 = nM
- DNMT3B/3L IC₅₀ > 100000 nM
- DNMT3A/3L IC₅₀ > 100000 nM
Goal of the research program

Discovery and development of DNMT inhibitors and other epigenetic targets

Approach

Computational methods integrated with experimental validation.

Specific aims

• Development of predictive models.
• Virtual screening of compound libraries.
Databases of compounds with experimental activity

- 103,284,373 compounds.
- 271,135,693 bioactivity data.
- 9,643,220 publications.
- 3,173,654 patents.
## Epigenetic data in ChEMBL

Structure-activity data vs. 55 epigenetic targets

<table>
<thead>
<tr>
<th>Target</th>
<th>Function</th>
<th>Families (HGNC)</th>
<th>Cluster (manually annotated)</th>
<th>Molecules</th>
<th>Scaffolds</th>
<th>% Active</th>
</tr>
</thead>
<tbody>
<tr>
<td>BAZ2B</td>
<td>Acetylated histone reader</td>
<td>PHD finger proteins, methyl-CpG binding domain containing</td>
<td>BRD</td>
<td>53</td>
<td>27</td>
<td>25</td>
</tr>
<tr>
<td>BRD2</td>
<td>Histone PTM reader</td>
<td>NA</td>
<td>BRD</td>
<td>277</td>
<td>91</td>
<td>87</td>
</tr>
<tr>
<td>BRD3</td>
<td>Histone PTM reader</td>
<td>NA</td>
<td>BRD</td>
<td>263</td>
<td>89</td>
<td>95</td>
</tr>
<tr>
<td>BRD4</td>
<td>Histone PTM reader</td>
<td>NA</td>
<td>BRD</td>
<td>643</td>
<td>259</td>
<td>80</td>
</tr>
<tr>
<td>BRD9</td>
<td>Histone PTM reader</td>
<td>NA</td>
<td>BRD</td>
<td>13</td>
<td>9</td>
<td>77</td>
</tr>
<tr>
<td>BRPF1</td>
<td>Histone PTM reader</td>
<td>PHD finger proteins, PWPP domain containing</td>
<td>BRD</td>
<td>27</td>
<td>15</td>
<td>89</td>
</tr>
<tr>
<td>DNMT1</td>
<td>DNA methyltransferase</td>
<td>Zinc fingers CXXC-type, seven-beta-strand methyltransferase motif containing</td>
<td>DNMT</td>
<td>248</td>
<td>194</td>
<td>60</td>
</tr>
<tr>
<td>DNMT3A</td>
<td>DNA methyltransferase</td>
<td>PWPP domain containing</td>
<td>DNMT</td>
<td>47</td>
<td>30</td>
<td>55</td>
</tr>
<tr>
<td>DNMT3B</td>
<td>DNA methyltransferase</td>
<td>PWPP domain containing</td>
<td>DNMT</td>
<td>40</td>
<td>22</td>
<td>50</td>
</tr>
<tr>
<td>CREBBP</td>
<td>Histone acetyltransferase</td>
<td>Zinc fingers ZZ-type, lysine acetyltransferases</td>
<td>HAT</td>
<td>180</td>
<td>65</td>
<td>64</td>
</tr>
<tr>
<td>EP300</td>
<td>Histone acetyltransferase</td>
<td>Zinc fingers ZZ-type, lysine acetyltransferases</td>
<td>HAT</td>
<td>73</td>
<td>52</td>
<td>78</td>
</tr>
<tr>
<td>KAT2A</td>
<td>Histone acetyltransferase</td>
<td>Lysine acetyltransferases, ATAC complex, SAGA complex, GCN5 related N-acetyltransferases</td>
<td>HAT</td>
<td>27</td>
<td>20</td>
<td>41</td>
</tr>
<tr>
<td>KAT2B</td>
<td>Histone acetyltransferase</td>
<td>Lysine acetyltransferases, ATAC complex, SAGA complex, GCN5 related N-acetyltransferases</td>
<td>HAT</td>
<td>121</td>
<td>40</td>
<td>61</td>
</tr>
<tr>
<td>NCOA1</td>
<td>Histone acetyltransferase</td>
<td>Basic helix-loop-helix proteins, lysine acetyltransferases</td>
<td>HAT</td>
<td>634</td>
<td>568</td>
<td>22</td>
</tr>
<tr>
<td>NCOA3</td>
<td>Histone acetyltransferase</td>
<td>Basic helix-loop-helix proteins, lysine acetyltransferases, trinucleotide repeat containing</td>
<td>HAT</td>
<td>564</td>
<td>517</td>
<td>32</td>
</tr>
<tr>
<td>HDAC1</td>
<td>Histone deacetylase</td>
<td>Histone deacetylases class I, EMSY complex, NuRD complex, SIN3 histone deacetylase complex</td>
<td>HDAC</td>
<td>3304</td>
<td>1418</td>
<td>90</td>
</tr>
<tr>
<td>HDAC2</td>
<td>Histone deacetylase</td>
<td>Histone deacetylases class I, EMSY complex, NuRD complex, SIN3 histone deacetylase complex</td>
<td>HDAC</td>
<td>942</td>
<td>427</td>
<td>84</td>
</tr>
<tr>
<td>HDAC3</td>
<td>Histone deacetylase</td>
<td>Histone deacetylases class I</td>
<td>HDAC</td>
<td>854</td>
<td>395</td>
<td>80</td>
</tr>
<tr>
<td>HDAC4</td>
<td>Histone deacetylase</td>
<td>Histone deacetylases class II</td>
<td>HDAC</td>
<td>704</td>
<td>348</td>
<td>69</td>
</tr>
<tr>
<td>HDAC5</td>
<td>Histone deacetylase</td>
<td>Histone deacetylases class II</td>
<td>HDAC</td>
<td>235</td>
<td>150</td>
<td>58</td>
</tr>
</tbody>
</table>

Jesús Naveja
Development of predictive models

Binary classification

- Precision: 0.92 - 0.81
- Sensibility: 0.89 - 0.65

Target prediction

- Precision: 0.95 - 0.77
- Sensibility: 0.89 - 0.56

Norberto Sánchez
Epigenetic target fishing

- Implementation of predictive models.
- Free webserver to predict the activity of small organic molecules with 55 epigenetic targets.

Norberto Sánchez

www.epigenetictargetprofiler.com
Epigenetic target fishing
Inverse virtual screening with epigenetic targets

Norberto Sánchez
Screening of compound databases

1. *In silico* characterization and “purification” (data curation) of compound libraries
   - In silico characterization
   - Purification (data curation)

2. Computational screening
   - In silico screening
   - Selected molecules for testing (virtual screening hits)

3. Experimental screening of selected compounds
   - Experimental screening
   - Active compounds

4. Optimization
   - Chemical synthesis and testing of analogs

Compound libraries
- Natural products
- Food chemicals
- In-house from collaborators
- Commercial synthetic
- Focused libraries

Strategies
- Similarity searching
- Molecular docking & dynamics
- Machine learning: Epigenetic Target Profiler

Approaches
- Enzymatic inhibition assays: collaboration and contract services.
- Additional experiments: Open for collaborations!

Strategy
- Chemical synthesis and testing of analogs.
Compound libraries

Natural products & food chemicals

- COCONUT
  - ~0.5 million compounds
- BIOFACQUIM
  - ~530 natural products
- FOODB
  - ~22K food chemicals

Data sets from collaborators

- Synthetic analogs of caffeic acid
  - Dra. Laura Alvarez (UAEM)
  - Dr. Mayra Antúnez (UAEM)
- Small molecules
  - Dr. Alexander Gagnon (UQAM, Canada)

Commercial general synthetic libraries

Focused libraries

ZINC15
Welcome to ZINC, a free database of commercially available compounds for virtual screening. ZINC contains over 230 million purchasable compounds in ready-to-dock 3D formats. ZINC also contains over 750 million purchasable compounds you can search for analogs in under a minute.

Otava chemicals

Focused Libraries Building Blocks Screening Compounds
Epigenetic focused libraries: characterization

11 focused libraries.
- 56,680 initial compounds.
- **53,443** compounds after data curation.

<table>
<thead>
<tr>
<th>Company</th>
<th>Library size</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Initial</td>
</tr>
<tr>
<td>ApeXBio</td>
<td>328</td>
</tr>
<tr>
<td>Asinex</td>
<td>5,391</td>
</tr>
<tr>
<td>ChemDiv</td>
<td>30,431</td>
</tr>
<tr>
<td>Enamine</td>
<td>9,352</td>
</tr>
<tr>
<td>Life Chemicals</td>
<td>7,019</td>
</tr>
<tr>
<td>MedChemExpress</td>
<td>700</td>
</tr>
<tr>
<td>OTAVA DNMT1</td>
<td>466</td>
</tr>
<tr>
<td>OTAVA DNMT3B</td>
<td>1,261</td>
</tr>
<tr>
<td>Targetmol</td>
<td>932</td>
</tr>
<tr>
<td>Tocris</td>
<td>101</td>
</tr>
<tr>
<td>SelleckChem</td>
<td>699</td>
</tr>
</tbody>
</table>

Alexis Padilla

Visual representation of the chemical space

Principal component analysis of 6 properties of pharmaceutical interest.

- Compound libraries have drug-like properties.
- Have different diversity (MedChemExpress, Targetmol, SelleckChem).
Constellation maps
Epigenetic focused libraries

Representative molecular scaffolds of each “constellation”

Molecular libraries have:
• Different chemical structures.
• Cover different regions of the chemical space.
• Different diversity.
Computational screening of focused libraries

Approaches
1. **Epigenetic Target Profiler (ETP)**
2. **Docking** with MOE
3. Docking AutoDock VINA
4. **Re-scoring**: Extended connectivity interaction features (ECIF). *

53,443 compounds (initial library size)

ETP

119 compounds with best ETP predictions.

Consensus docking and re-scoring

20 compounds selected for enzymatic inhibition assays.

<table>
<thead>
<tr>
<th>Proveedor</th>
<th>ID</th>
<th>AutodockVINA (kcal/mol)</th>
<th>ECIF</th>
<th>MOE (kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enamine</td>
<td>Z991906684</td>
<td>-8.5</td>
<td>5.36</td>
<td>-8.02</td>
</tr>
<tr>
<td>Tocris</td>
<td>TC25</td>
<td>-9.8</td>
<td>5.3</td>
<td>-9.95</td>
</tr>
<tr>
<td>Targetmol</td>
<td>T2354</td>
<td>-9.4</td>
<td>5.36</td>
<td>-9.74</td>
</tr>
<tr>
<td>ChemDiv</td>
<td>C191-0147</td>
<td>-9.0</td>
<td>5.19</td>
<td>-8.57</td>
</tr>
<tr>
<td>ChemDiv</td>
<td>C769-0077</td>
<td>-8.8</td>
<td>5.28</td>
<td>-8.19</td>
</tr>
<tr>
<td>ChemDiv</td>
<td>F477-3331</td>
<td>-9.3</td>
<td>5.26</td>
<td>-8.33</td>
</tr>
<tr>
<td>ChemDiv</td>
<td>G119-0160</td>
<td>-8.7</td>
<td>5.37</td>
<td>-7.29</td>
</tr>
<tr>
<td>ChemDiv</td>
<td>L485-2681</td>
<td>-8.9</td>
<td>5.4</td>
<td>-8.64</td>
</tr>
<tr>
<td>ChemDiv</td>
<td>L485-2718</td>
<td>-9.4</td>
<td>5.48</td>
<td>-8.26</td>
</tr>
<tr>
<td>ChemDiv</td>
<td>L485-2735</td>
<td>-8.9</td>
<td>5.45</td>
<td>-8.12</td>
</tr>
<tr>
<td>ChemDiv</td>
<td>L485-2754</td>
<td>-10</td>
<td>5.56</td>
<td>-7.95</td>
</tr>
<tr>
<td>ChemDiv</td>
<td>L485-2759</td>
<td>-10.1</td>
<td>5.42</td>
<td>-8.16</td>
</tr>
<tr>
<td>MedChemEx HY-10128</td>
<td>-7.7</td>
<td>5.52</td>
<td>-9.83</td>
<td></td>
</tr>
<tr>
<td>ChemDiv</td>
<td>L485-2761</td>
<td>-9.6</td>
<td>5.43</td>
<td>-8.48</td>
</tr>
<tr>
<td>ChemDiv</td>
<td>L485-2767</td>
<td>-9.8</td>
<td>5.23</td>
<td>-8.84</td>
</tr>
</tbody>
</table>

* Alexis Padilla
Focused libraries: experimental screening

% of activity of DNMT1 of virtual screening hits at 100 μM

Quinolines
- Inhibition 10-40 %
- Cases of activation and activity cliffs
Focused libraries: experimental screening

Quinazolines

MolPort-023-277-153

- DNMT1: IC\textsubscript{50} = 30 nM
- DNMT3A: IC\textsubscript{50} = 4870 nM
- DNMT3B: > 100000 nM

MolPort-035-789-726

- DNMT1: IC\textsubscript{50} = 81 nM
- DNMT3A: IC\textsubscript{50} = 14690 nM
- DNMT3B: > 100000 nM

MolPort-006-396-396

- DNMT1: % inhibition = 7.4 +/- 2.4
- DNMT1: IC\textsubscript{50} NA
- DNMT3A, DNMT3B: NA

-- Nanomolar inhibition of DNMT1.
-- ~ 10 times more potent than the control SAH.
-- Selective towards DNMT1.
  - Low or no activity with DNMT3A and DNMT3B

Reference:
SAH: IC\textsubscript{50} = 0.34 \mu M
Survey in the literature...

Quinazolines are also inhibitors of the epigenetic reader G9a

Optimization of Cellular Activity of G9a Inhibitors 7-Aminoalkoxy-quinazolines

Feng Liu,†‡ Dalia Barsyte-Lovejoy,†‡ Abdellah Allali-Hassani,† Yunlong He,§ J. Martin Herold,† Xin Chen,† Christopher M. Yates,† Stephen V. Frye,† Peter J. Brown,§ Jing Huang,§ Masoud Vedadi,‡ Cheryl H. Arrowsmith,‡ and Jian Jin§∗†

†Center for Integrative Chemical Biology and Drug Discovery, Division of Medicinal Chemistry and Natural Products, UNC Eshelman School of Pharmacy, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, United States
‡Structural Genomics Consortium, University of Toronto, Toronto, Ontario, M5G 1L7, Ontario, Canada
§Laboratory of Cancer Biology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892, United States

G9a IC$_{50}$ = 6 nM
Reduction of H3K9me2 in MCF7 cells: IC$_{50}$ = 10 nM
Cell toxicity (MCF7 cells): EC$_{50}$ = 4,700 nM
Progress to elucidate the mechanism of inhibition and selectivity

M. en C. Edgar López

Computational
(docking + dynamics)
Interaction with the CXXX domain.

Experimental
It does not compete with DNA

MolPort-023-277-153
DNMT1: IC₅₀ = 30 nM
DNMT3A: IC₅₀ = 4870 nM
DNMT3B: > 100000 nM
Towards multitarget epigenetic drug discovery

Methods & Protocols

An *in silico* pipeline for the discovery of multitarget ligands: A case study for epi-polypharmacology based on DNMT1/HDAC2 inhibition

Fernando D. Prieto-Martínez, Eli Fernández-de Gortari, José L. Medina-Franco, L. Michel Espinoza-Fonseca

Databases

- Structural information
  - HDAC & DNMT inhibitors activity
    - Experimentally based
    - Public repositories: ChEMBL, PubChem

- Data preprocessing
  - Open Babel: Curation, filtering

- Inactive compounds
  - DUDC decoys
  - DeepCoy
  - Class labeling

- Data Representation
  - RDKit
  - Morgan FP
  - Physicochemical properties: MW, TPSA, RB, HBD(A), Csp3

- Scaling, test/trainsets
  - Scikit learn

Classification Models Selection

- Models Training
  - Scikit learn
  - RFC
  - DTC
  - LRC
  - XGBC

- Parameter tuning and validation
  - Scikit learn
  - Grid search
  - Ten-fold cross-validation

Model selection

- ROC AUC, Accu, STD, implementation, etc.

Generator

- Pretrained ANN-Model
  - Learning continuous and data-driven
  - Molecular descriptors

Scoring Function

- Multi-objective particle swarm latent space optimization

In-house classification models integration

- Prebuilt RDKit functions

Seed Selection

- Literature sources

Virtual Screening

- Virtual Compound
  - Positive: Generated
  - Negative: DeepCoy

Data processing

- Curation, filtering
- Scaling, test/trainsets

Models Training

- XGBoost

Parameter tuning and validation

In-house classification models integration

- ROC AUC

Virtual Screening

- COCONUT

Hit Selection

- Data Processing
  - Probability and PGP filtering

- Hit filtering
  - Classifier probability
  - Cluster selection

- Hit selection
  - Flexible pharmacophoric alignment
  - Molecular docking
  - Protein-ligand interactions
  - Molecular dynamics

Collaboration with Dr. Eli Fernandez
Perspective: De novo design

Design of small molecules from scratch based on anchor points in the binding site

MSc. Diana Prado
Chemoinformatics
Challenges ahead

Grand Challenges of Computer-Aided Drug Design: The Road Ahead

José L. Medina-Franco*

DIFACQUIM Research Group, Department of Pharmacy, School of Chemistry, National Autonomous University of Mexico, Mexico City, Mexico
Chemical space

• Expand the medicinally relevant chemical space.

• Rational design and screen of ultra-large chemical libraries.

• Repurpose existing libraries (drugs and in-house collections).

• Rescue missing hits and lead compounds from screening libraries.

• Explore neglected regions of chemical space.
• Improve multi-target drug design and polypharmacology.

• Explore “dark” targets and identify novel promising regions in the genome.

• Improve targeting protein-protein interactions.

• Continue investigating targets associated with rare and neglected diseases.
Methodological challenges

How to conduct the search for new and better drugs at the intersection of the chemical and biological spaces?

- Computational chemogenomics.
- Automated *de novo* design and computational fragment screening.
- Improve property prediction, including ADME and toxicity.
- Modeling large and complex systems.
- Continue to improve molecular docking and scoring.
- Improve the hit rate of virtual screening and strategies to automatically propose high quality hits.
- Synergize with other methods: consensus approaches.
- Ensure data curation and quality.
Human factor and other challenges

• Communication and human interaction.
  – Improve multidisciplinary research: reach common objectives from different perspectives.
  – Enhance communication across research teams; avoid duplicating efforts.

• Dissemination and data sharing.
  – Rigorous dissemination of information and high-quality data.
  – Transparency and reproducibility.
  – Open science vs. securing intellectual property.

• Education and training.
  – Individuals and teams.
  – Set up realistic expectations of computational methods.
Grand challenge
Use cheminformatics rationally beyond the hype

OPINION ARTICLE
Rationality over fashion and hype in drug design [version 1; peer review: 2 approved]

José L. Medina-Franco¹, Karina Martinez-Mayorga², Eli Fernández-de Gortari³, Johannes Kirchmair⁴, Jürgen Bajorath⁵

“Unrealistic expectations are a quick road to disappointments”
Summary
Take home messages

• Chemoinformatics: an independent discipline that impacts many areas of chemistry.

• Discovery of epi-drug candidates.
  – Development of *Epigenetic Target Profiler*.
  – Virtual screening identifies low micromolar and selective DNMT1 inhibitors.

• Challenges of chemoinformatics and CADD.
  – Revisit and expand chemical and biological spaces.
  – Several methodological challenges: data quality is a must.
  – Effective communication and education/training: beware of the “artificial intelligence extasy”.
Acknowledgments

DIFACQUIM’s students & alumni
- Norberto Sánchez
- Jesús Naveja
- Fernando Prieto
- Edgar López
- Euridice Juárez
- Marisa Santibáñez
- Claudia Oviedo
- Alexis Flores
- Jocelyn Salazar
- Diana Prado

Collaborators

Eli Fernández
INL, Portugal

Paola Arimondo
CNRS, France

Hyang-Min Byun
New Castle, UK

Massimo Bertinaria
Università degli Studi di Torino

Laura Alvarez
UAEM, Mexico

Mark Muller
Topogen, Colorado, USA

Alexandre Gagnon
Université du Québec à Montréal

Mayra Antúnez
UAEM, Mexico

Alfonso Dueñas
INCAN, Mexico

Carlos Velazquez
University of Alberta, Canada

Keith Robertson
Mayo Clinic, USA

Federica Catti
Arkansas State University

Funding

CONACyT 282785
PAPIIT IA203718, IN201321
PAIP 5000-9163
Miztli Supercomputer (UNAM)
Colloquium Chemoinformatics and Artificial Intelligence
Progress and Challenges to Develop Bioactive Compounds

SAVE THE DATE!
JUNE 15-17, 2022

16 speakers.

Progress and challenges to develop peptides, natural products, drug candidates.

Speakers from: Academia, industry and other research institutions.

Organizers:
Secretariat of Graduate Studies and Research
School of Chemistry, National Autonomous University of Mexico (UNAM)
DIFACQUIM research group

Registration
www.difacquim.com @difacquim