

## Systems pharmacology using data deposited in Drugbank

Zastosuj narzędzia biologii systemów do przyjrzenia się bliżej związkom chemicznym i ich celom. Zbuduj różne sieci interakcji lek-cel i oceń czy dotychczasowa wiedza zgromadzona przez lekarzy, chemików medycznych, bioinformatyków i biologów molekularnych pozwala pogłębić nasze rozumienie leczenia chorób oraz efektów ubocznych stosowanych terapii.

### Materialy i Metody:

- dane z [www.drugbank.ca/downloads](http://www.drugbank.ca/downloads)
- Python/Perl <https://github.com/dhimmel/drugbank>
- Cytoscape lub Gephi
- UniprotKB,
- sysPTM lub ELM
- openbabel

Na podstawie danych z DrugBank stwórz ogólne i specyficzne sieci drug-target (lek-cel) w Cytoscape lub Gephi. Na ich podstawie odpowiedz na wybrane pytanie:

1. Jakie są parametry sieci lek-cel zbudowanej na podstawie danych z Drugbank? Czy leki są specyficzne czy wielo-celowe? Które leki mają najwięcej targetów, a które najmniej? Ile mają średnio celów?
2. Jakie typy związków chemicznych występują w lekach (pole <classification>). Czy któreś z nich są preferowane przez dany subtyp targetów (subtypów: "carriers", "transporters", "metabolizing enzymes"). Czy któreś z nich są średnio bardziej specyficzne czy bardziej promiskuitywne?
3. Stwórz sieci interakcji dla różnych subtypów celów: "carriers", "transporters", "metabolizing enzymes". Porównaj właściwości takich sieci. Co jest specyficzne a co wspólne?
4. Jak wyglądają sieci interakcji dla różnych subtypów leków "FDA approved", "withdrawn", "illicit, experimental drugs" i "approved- investigational". Porównaj właściwości takich sieci. Co jest specyficzne a co wspólne?
5. Jak podobne są leki do siebie? Wykorzystaj podobieństwo FP2 (openbabel). Czy podobne leki oddziałują na te same cele biologiczne?
6. Jeśli dwa leki mają wspólny cel (target) to występuje między nimi połączenie (link). Jeśli mają więcej niż jeden wspólny target waga takiego linku jest wyższa. Jakie są parametry

- takiej sieci? Czy na tej podstawie potrafisz znaleźć leki które podane wspólnie mogą współdziałać lub powodować niekorzystne efekty uboczne?
7. Jeśli dwa targety mają wspólny lek to występuje między nimi połączenie (link). Jeśli mają więcej niż jeden wspólny lek waga takiego linku jest wyższa. Jakie są parametry takiej sieci? Czy na tej podstawie potrafisz znaleźć leki, które można wykorzystać w innych chorobach niż pierwotnie zakładano (tzw. drug repurposing)?
  8. Jakie leki oddziałują na różne cele (tzw. multitarget drugs)? Kiedy zostały odkryte/zatwierdzone? Na jakie procesy biologiczne oddziałują (GO). Jakie cechy ma sieć drug-target z i bez tych leków?
  9. Gdzie znajdują się w komórce cele leków, jaka jest ich lokalizacja? Podziel komórkę na komponenty: Plasma membrane, Cytosol, Extracellular space, Mitochondrion, Nucleus, Unknown/other. Dodaj dane dotyczące roku patentu/zatwierdzenia leku. Czy widać jakiś trend?
  10. Zbadaj jak wygląda kwestia modyfikacji posttranslacyjnych (PTM) targetów leków: wykorzystaj sysPTM lub ELM. Typy PTM: glycosylation, phosphorylation, ubiquitination, sulfation, , disulfide bonds, nucleotide binding. Podziel targety na subtypy (transmembrane receptors, ion channels, kinases, metaloenzymes, inne?) i określ czy specyficzne PTMy są nadreprezentowane w różnych subtypach.

Parametry określające cechy sieci to m.in:

- **Average node degree:** For each node, node degree is the number of edges connected to a node. Number of connected nodes/disconnected nodes
- **Centrality metrics:** node degrees of drugs and/or targets
- **Hub node:** Highest degree nodes in a network.
- **Connected component:** A subgraph of a network, in which there is a path between any two nodes. The largest component in a network is called giant component (GC).
- **Average path length:** The average of number of links in shortest paths for all possible pairs of network nodes.
- **Clustering coefficient:** Tendency of nodes for clustering together measured by the ratio of existing edges for a node's neighbors to maximum number of all possible edges
- **Modularity index:** A value between -1 and 1 that quantifies the density of connections inside network subunits compared with the links between subunits.
- **Degree distribution:** The probability distribution of node degrees over the whole network.
- **Power-law degree distribution:** Inhomogeneous distribution of node degrees within a network, i.e. few nodes have high degrees, while most of the nodes have low degree.
- **Network density:** Measurement of how a network is close to complete.  $k$  is the average node degrees, and  $n$  is the total number of nodes.
- **Network entropy:** The 'randomness' in graph is defined as the entropy of its degree distribution.
- Inne poznane na zajęciach z Biologii Systemów.