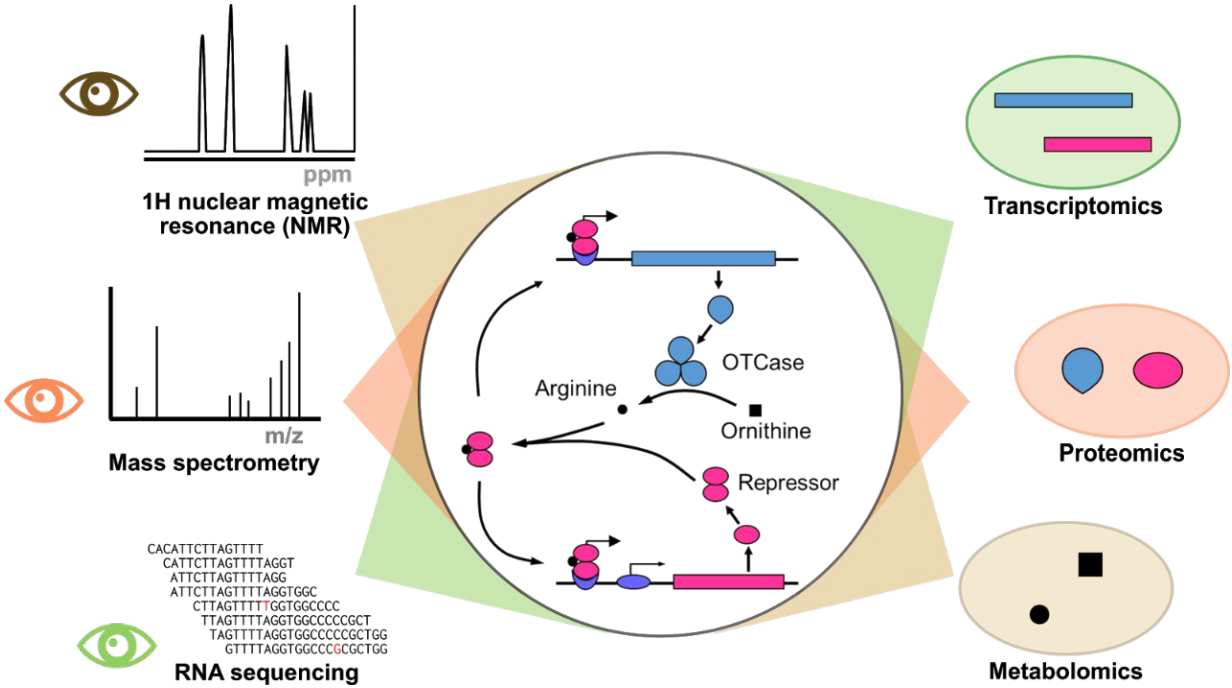
AMIDD 2024 Lecture 9: Biological networks and omics

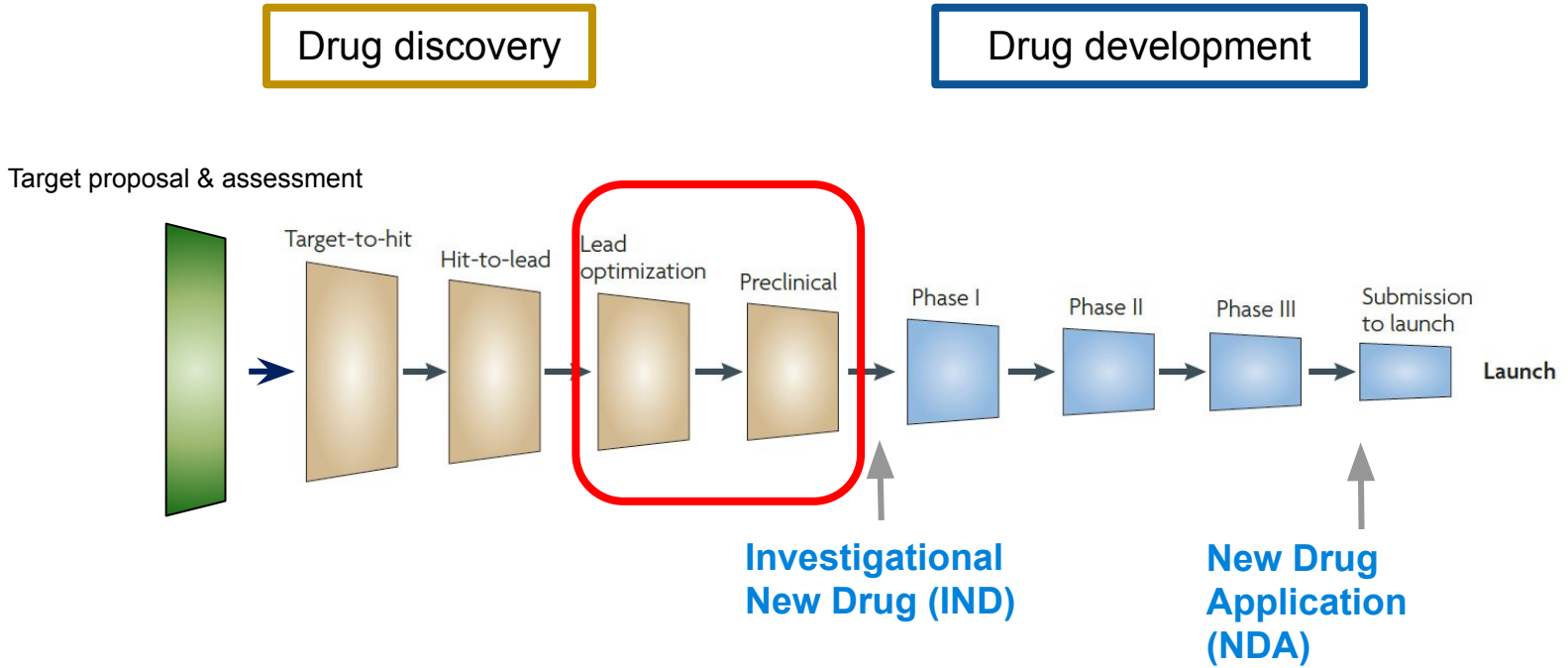


Dr. Jitao David Zhang, Computational Biologist

[1] Pharmaceutical Sciences, Pharma Research and Early Development, Roche Innovation Center Basel, F. Hoffmann-La Roche

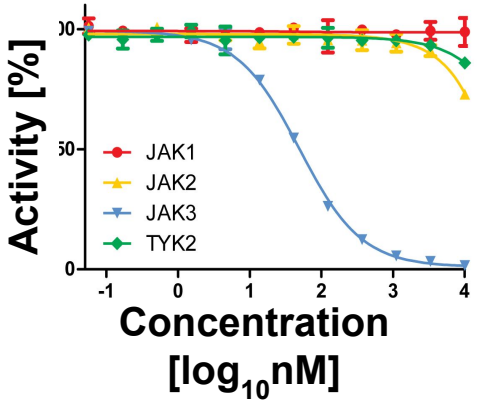
[2] Department of Mathematics and Informatics, University of Basel

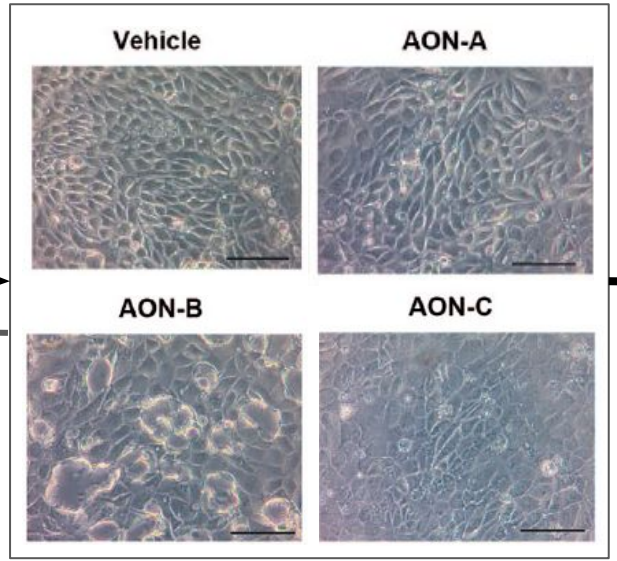
Translational research makes molecules into medicines

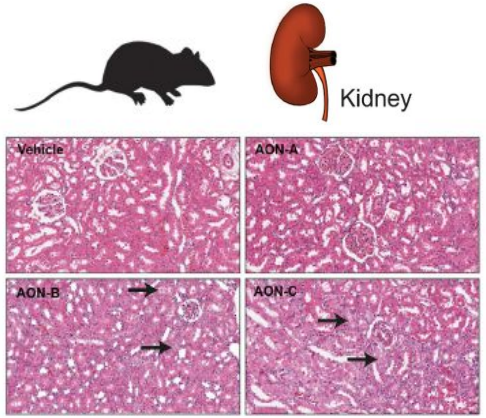


Adapted from Paul *et al.* "How to Improve R&D Productivity: The Pharmaceutical Industry's Grand Challenge." Nature Reviews Drug Discovery, 2010

Classical workflow of efficacy and toxicity assessment







Biochemical & biophysical assays

Cellular assays (*in vitro*)

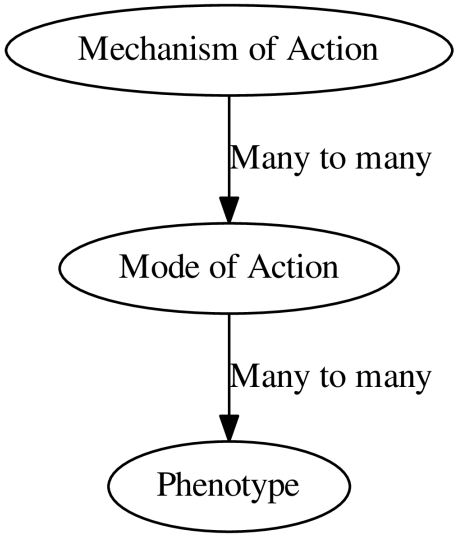
Animal experiments (*in vivo*)

Usual workflow

Assay development

Mechanism of Action and *Mode* of Action

- **Mechanism of Action**: The specific biochemical interaction through which a drug substance produces its pharmacological effect, ***at the molecular level***.
- **Mode of Action**: Functional or anatomical changes, ***at the cellular level*****,** resulting from the exposure of a living organism to a substance.
- For instance, a mechanism of action of a drug can be "*binding to Monoacylglycerol lipase (MAGL)*" while its mode of action would be "*regulating endocannabinoid signaling"* and "*reducing inflammation*".
- In lead optimization (LO) and early development, our goal is to understand both the mechanism of action and the mode of action *in vitro*, *in vivo*, and in human. The term *MoA* is used to refer both.

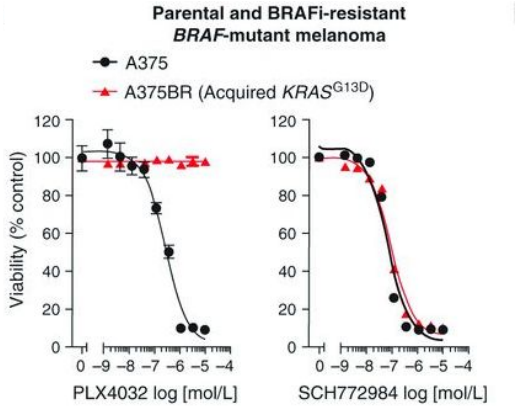


The Hill function is a common model of *in vitro* pharmacology

- The Hill function is one of the mostly useful non-linear functions to model biological systems.
- In its general form, H_{max} indicates the maximal value to which the function is asymptotic, n is the shape parameter (known as the Hill's coefficient), and k is the reflection point, often abbreviated as XC_{50} (X=I, E, C, …), the half-saturation constant.
- The Michaelis-Menten model is a special case of the Hill function ($n=1$).

$$H = H_{max}\frac{x^n}{k^n + x^n}$$

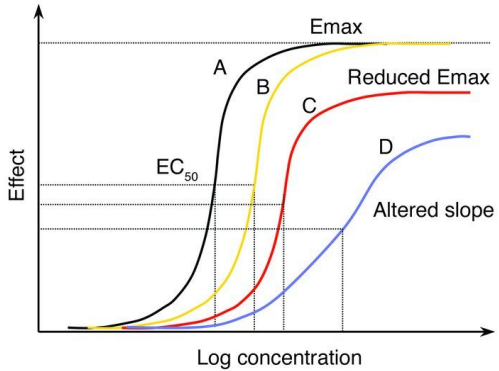
The general form of the Hill function



Morris *et al.* Cancer Discov; 3(7); 742–50. ©2013 AACR.

$$E = E_{max}\frac{[L]^n}{EC_{50}^n + [L]^n} = E_{max}\frac{1}{1 + (\frac{EC_{50}}{[L]})^n}$$

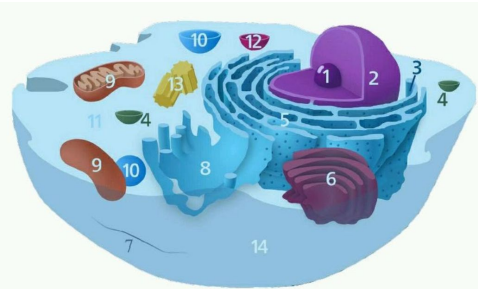
Modelling the dose-dependent effect

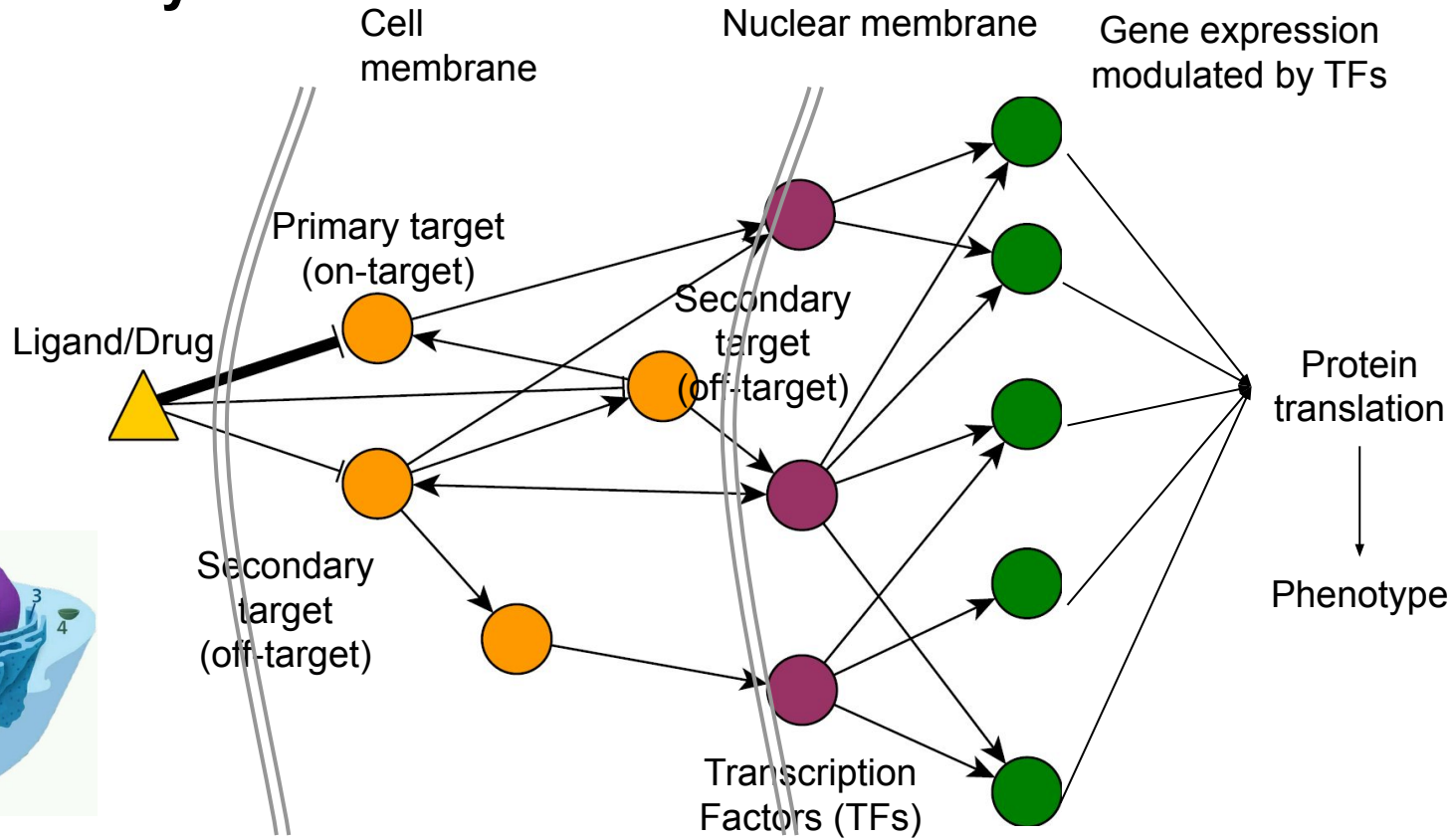


White. *J Clin Invest.* 2004;113(8):1084-1092. https://doi.org/10.1172/JCI21682.

Suppose it is an antiviral drug, compared with curve B, what does curve A, C, and D suggest?

Biological networks interact with drugs and manifest its efficacy and safety





Reaction Rate Equations: a compartment/ODE model of biological chemical reaction

$$\mathrm{E} + \mathrm{S} \underset{k_r}{\overset{k_f}{\rightleftharpoons}} \mathrm{ES} \xrightarrow{k_{cat}} \mathrm{E} + \mathrm{P}$$

$$v = \frac{V_{max}[S]}{K_D + [S]}$$

$$V_{max} \equiv k_{cat}[E]_0$$

$$v = \frac{d[P]}{dt} = k_{cat}[ES] = \frac{k_{cat}[E]_0[S]}{K_D + [S]}$$

$$[ES] = \frac{[E]_0[S]}{K_D + [S]}$$

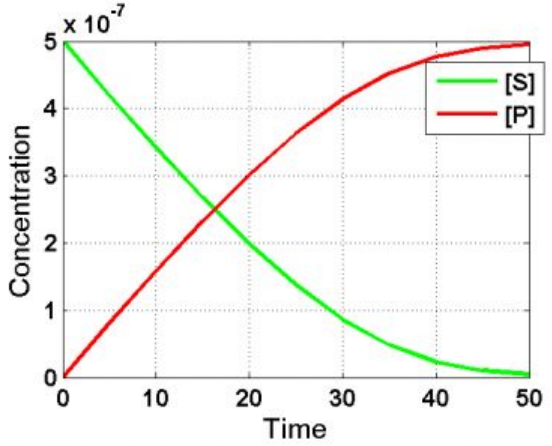
$$K_D \equiv \frac{k_r}{k_f}$$

$$\begin{aligned}
\frac{d[E]}{dt} &= -k_f[E][S] + k_r[ES] + k_{cat}[ES], \\
\frac{d[S]}{dt} &= -k_f[E][S] + k_r[ES], \\
\frac{d[ES]}{dt} &= k_f[E][S] - k_r[ES] - k_{cat}[ES], \\
\frac{d[P]}{dt} &= k_{cat}[ES],
\end{aligned}$$

$$k_f[E][S] = k_r[ES]$$

$$\begin{aligned}
k_f([E]_0 - [ES])[S] &= k_r[ES] \\
k_f[E]_0[S] - k_f[ES][S] &= k_r[ES] \\
k_f[E]_0[S] &= k_r[ES] + k_f[ES][S] \\
k_f[E]_0[S] &= [ES](k_r + k_f[S]) \\
[ES] &= \frac{k_f[E]_0[S]}{k_r + k_f[S]} \\
[ES] &= \frac{k_f[E]_0[S]}{k_f(\frac{k_r}{k_f} + [S])}
\end{aligned}$$

***RRE simulation of the Michaelis-Menten* model**



Source: Systems Engineering Wiki (tue.nl)

RRE is a set of ODEs, with each ODE representing one chemical species. Solution of the *j*th equation at time *t* is a real number representing the concentration of species *j* at time *t*.

Simulation of biological networks with ordinary differential expression

Given the reaction

$$\mathrm{S} + \mathrm{E} \underset{k_2}{\overset{k_1}{\rightleftharpoons}} \mathrm{C} \xrightarrow{k_3} \mathrm{P} + \mathrm{E}$$

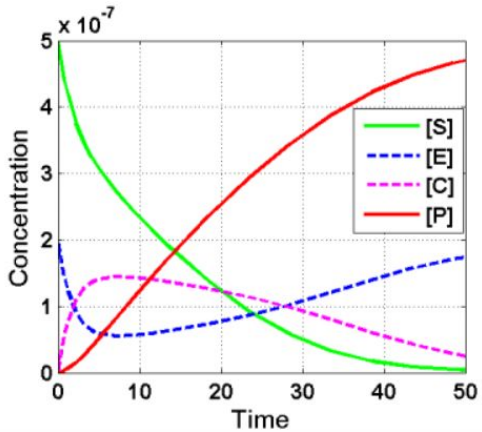
Given the initial values and rate constants

- $S(0) = 5e^{-7}$
- $E(0) = 2e^{-7}$
- $C(0) = P(0) = 0$
- $k_1 = 1e^{6}$
- $k_2 = 1e^{-4}$
- $k_3 = 0.1$

According to the law of mass action

$$\begin{aligned}
\frac{d[S]}{dt} &= -k_1[E][S] + k_2[C], \\
\frac{d[E]}{dt} &= -k_1[E][S] + (k_2 + k_3)[C], \\
\frac{d[C]}{dt} &= k_1[E][S] - (k_2 + k_3)[C], \\
\frac{d[P]}{dt} &= k_3[C],
\end{aligned}$$

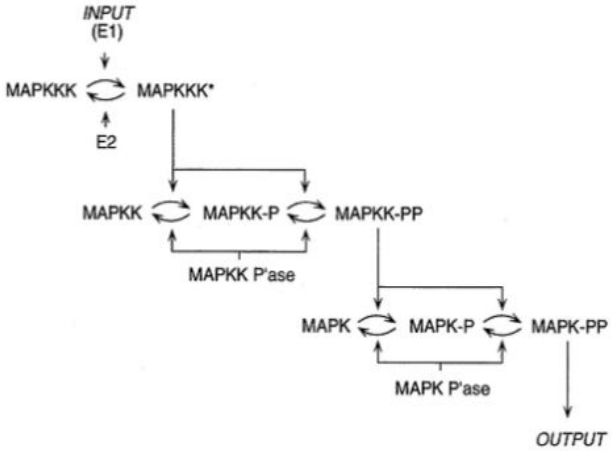
It is possible to simulate the concentration changes by time *deterministically*.

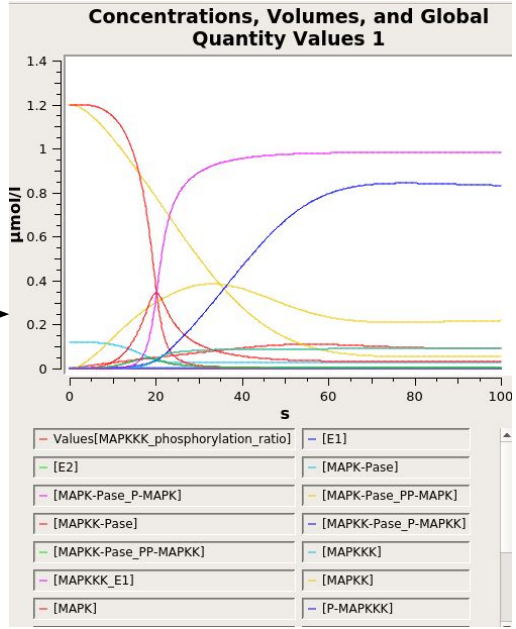


See Systems Engineering Wiki (tue.nl) for MATLAB/COPASI codes and *Stochastic Modelling for Systems Biology* by Darren J. Wilkinson

Simulating behavior of complex ODE systems with COPASI

- COPASI, freely available at http://COPASI.org/, supports both **ordinary differential equation (ODE)** based simulation as well as stochastic kinetic simulation.
- Such tools are important for detailed analysis of enzymatic reactions, for instance in the presence of drugs and/or disease-relevant mutation.

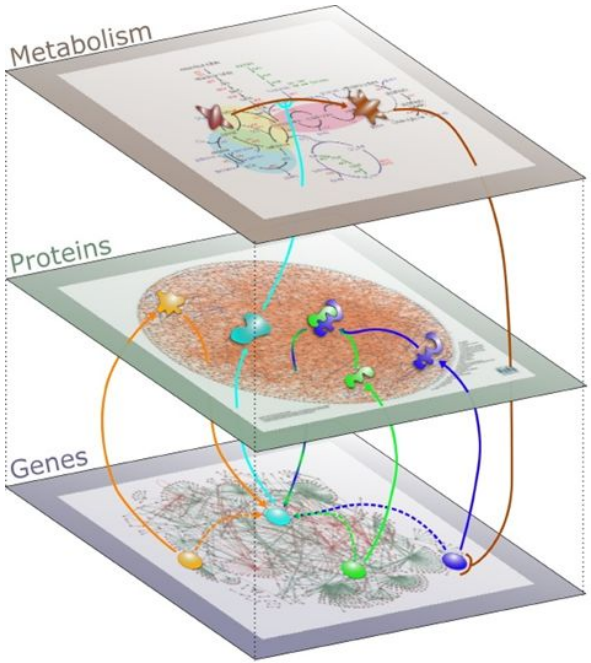




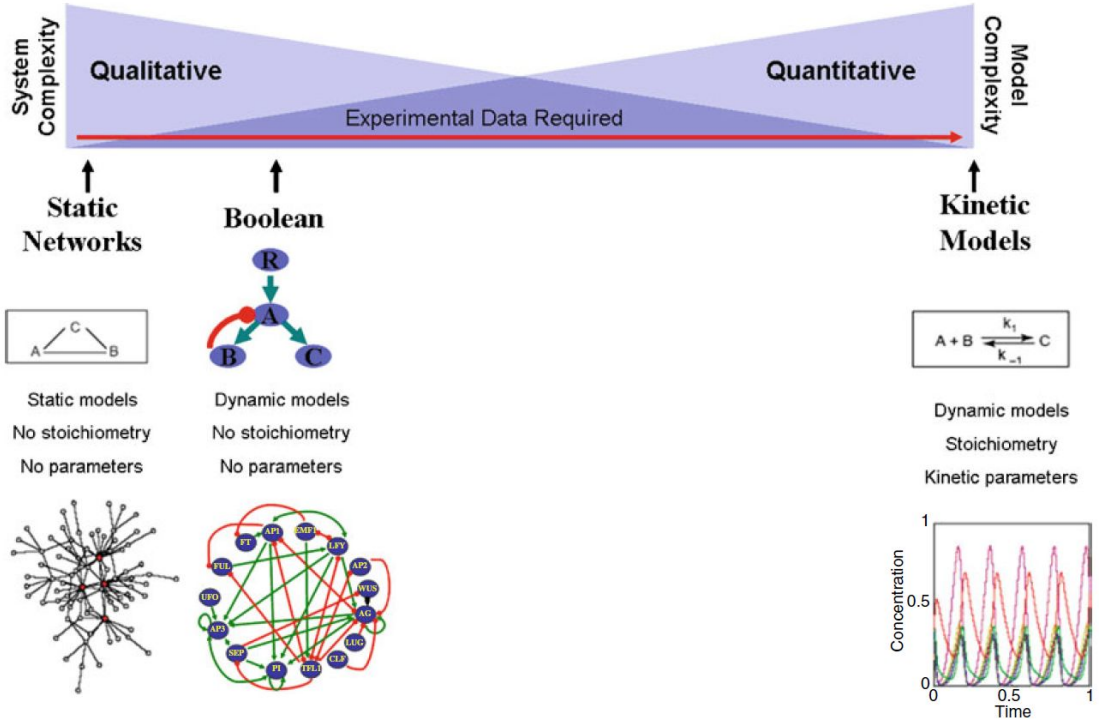
ODE-based simulation of dynamics

Figure: Huang and Ferrell, PNAS, 2006. Resources to learn more about stochastic modelling: MIT OpenCourseWare by Jeff Gore, and Stochastic Processes: An Introduction, Third Edition by Jones and Smith. Tutorials also available on the website of European Bioinformatics Institute (EBI)

Different ways of modelling biological networks

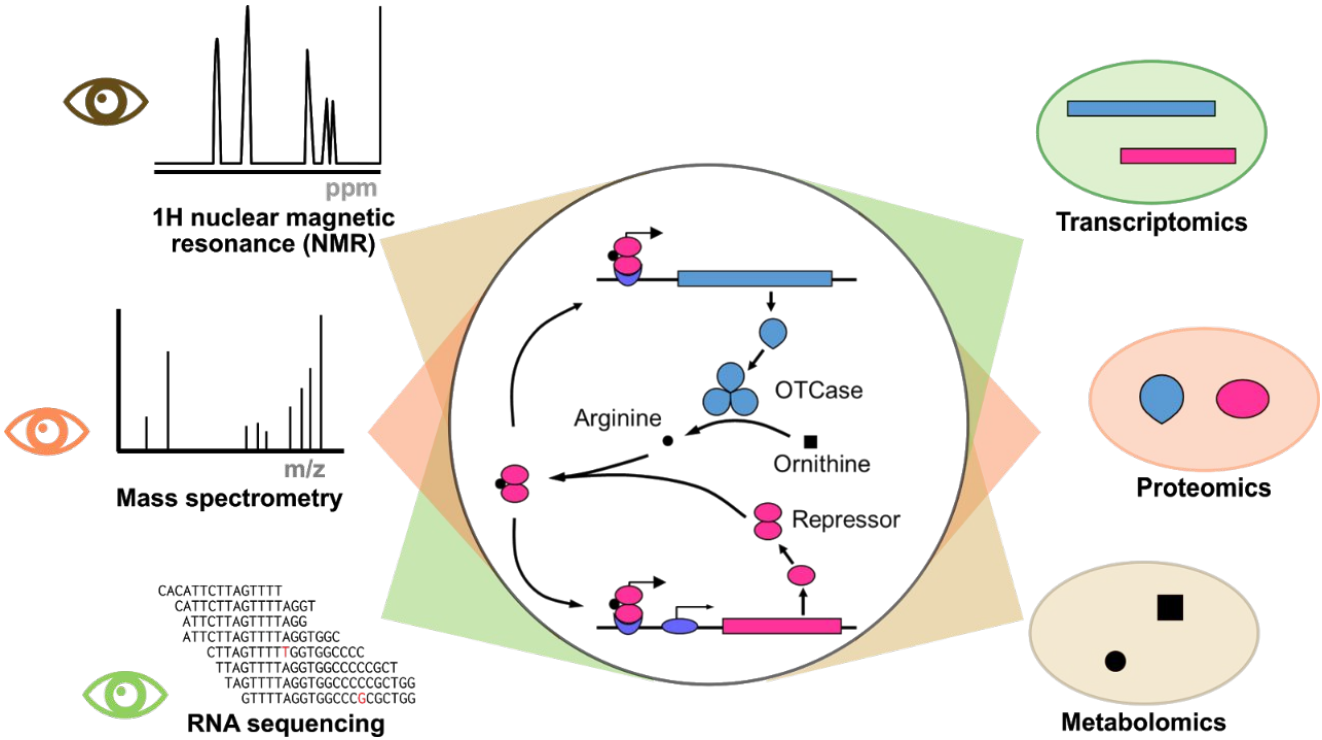


Stéphane CHÉDIN & Jean LABARRE, www-dsv.cea.fr

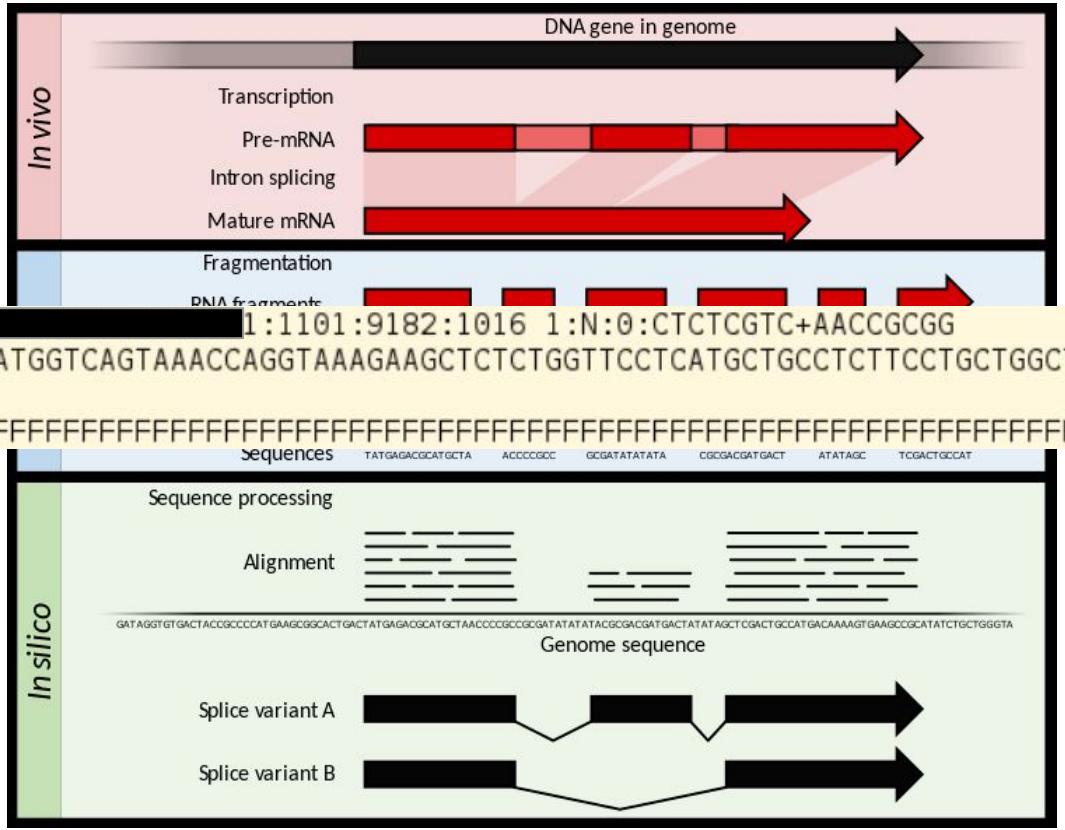


Garg, Abhishek, Kartik Mohanram, Giovanni De Micheli, and Ioannis Xenarios. 2012. "Implicit Methods for Qualitative Modeling of Gene Regulatory Networks." In *Gene Regulatory Networks: Methods and Protocols*, edited by Bart Deplancke and Nele Gheldof, 397–443. Methods in Molecular Biology. Totowa, NJ: Humana Press.

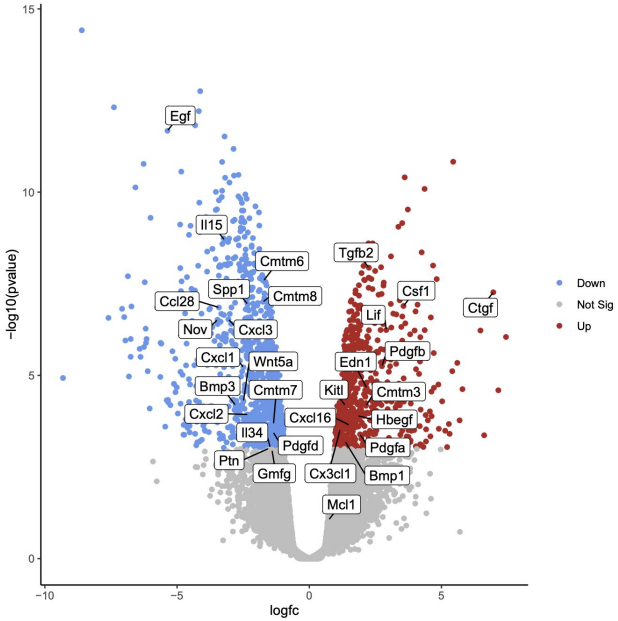
Biological networks can be studied with omics technologies



Principle of next-generation RNA sequencing (NGS)



We can reveal compound's effect on gene expression by performing differential gene expression analysis



Visualization (e.g. volcano plot)

Differential Gene Expression Analysis

Normalization

Count collection

Read Mapping

	sample A1	sample A2	sample B1	sample B2
gene 1	8	10	100	200
gene 2	14	15	15	40
gene 3	33	40	35	70
...	...	...	...	...
gene N	100	120	105	220

	sample A1	sample A2	sample B1	sample B2
gene 1	8	10	100	200
gene 2	14	15	115	40
gene 3	33	40	35	70
...	...	...	...	...
gene N	100	120	105	220

Tot. reads: 5 millions

Tot. reads: 10 millions

	sample A1	sample A2	sample B1	sample B2
gene 1	0.16	0.20	2.00	2.00
gene 2	0.28	0.30	0.30	0.40
gene 3	0.66	0.80	0.70	0.70
...	...	...	...	...
gene N	2.00	2.40	2.10	2.20

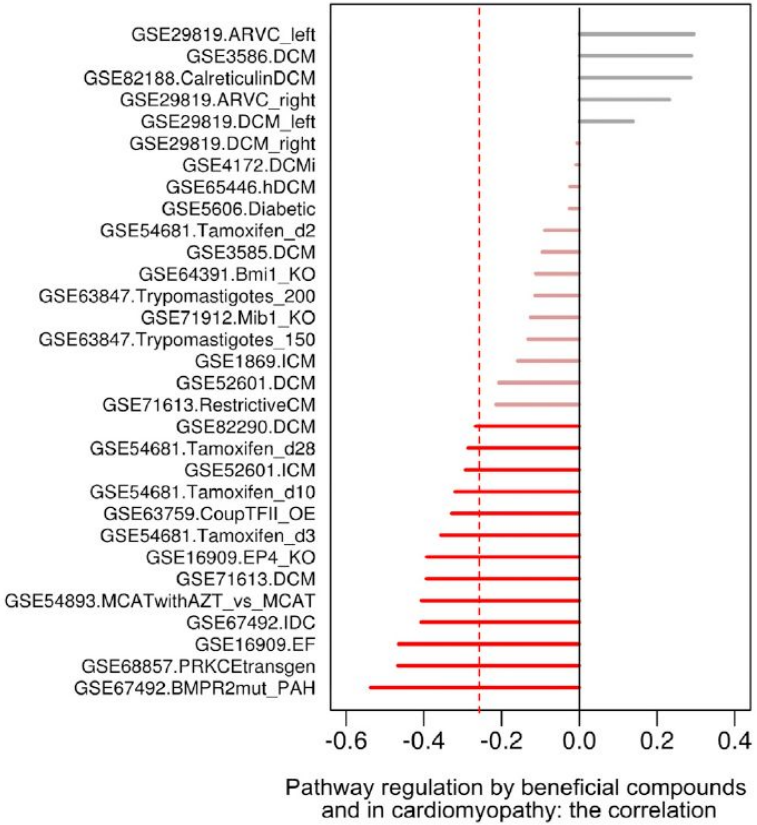
Advantages and challenges of using RNA sequencing to study MoA

Advantages:

1. Since patient samples can be also profiled with omics technologies, it is possible to compare a compound's effect with the changes induced by disease progression (right).
2. Well-designed omics study can reveal strong and subtle effects of the compound (the example with splicing modifier).

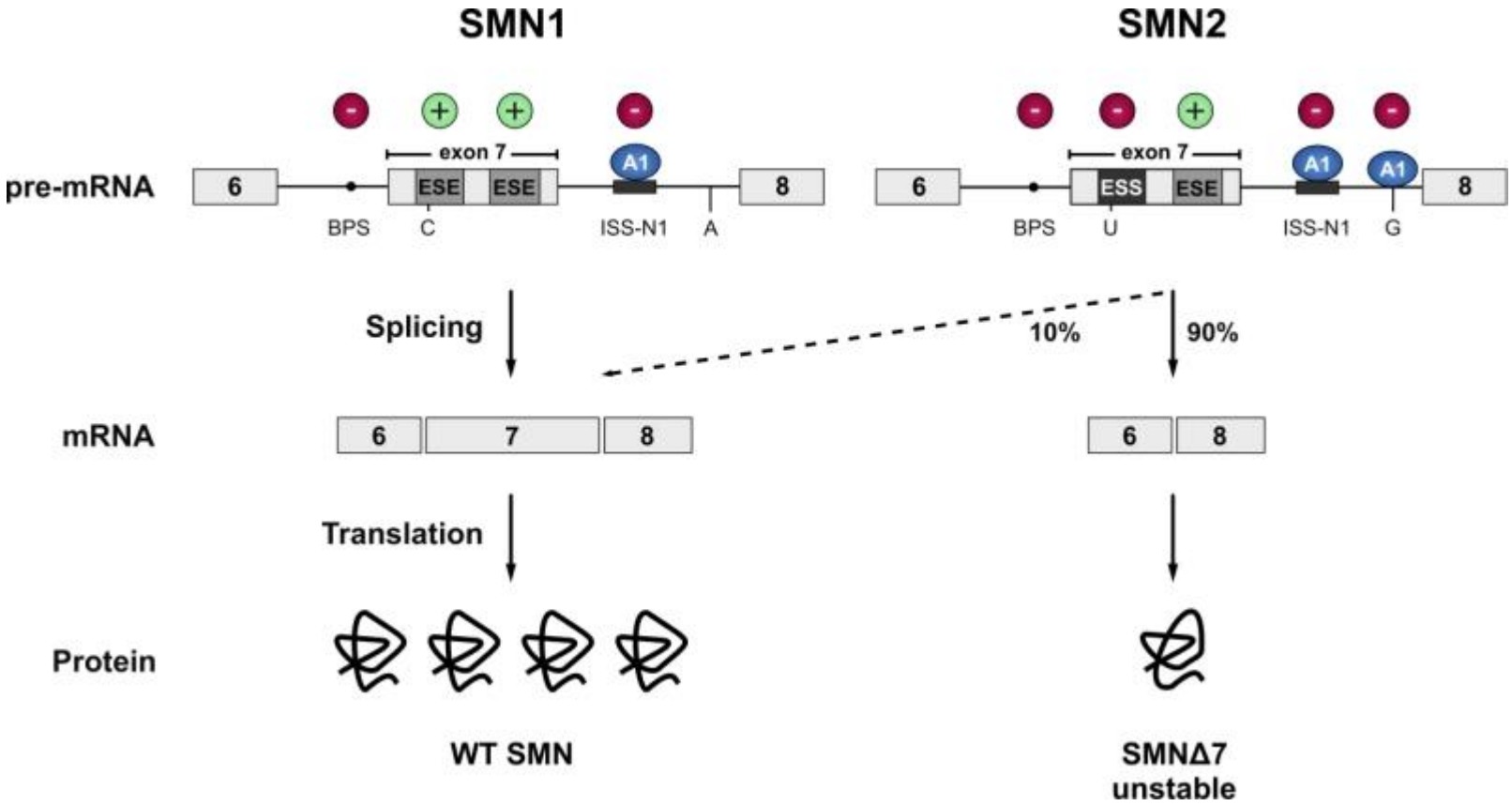
Challenges:

1. Data from biological models that poorly reflect human disease can do more harm than benefits.
2. Curse of dimensionality.



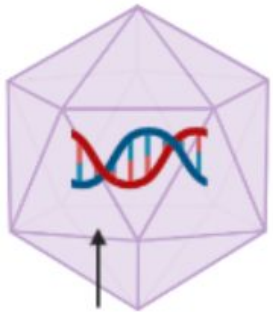
Pathway regulation by beneficial compounds and in cardiomyopathy: the correlation

Splicing of SMN1 and SMN2 genes: patients with mutations in SMN1 gene suffer from Spinal Muscle Atrophy (SMA)

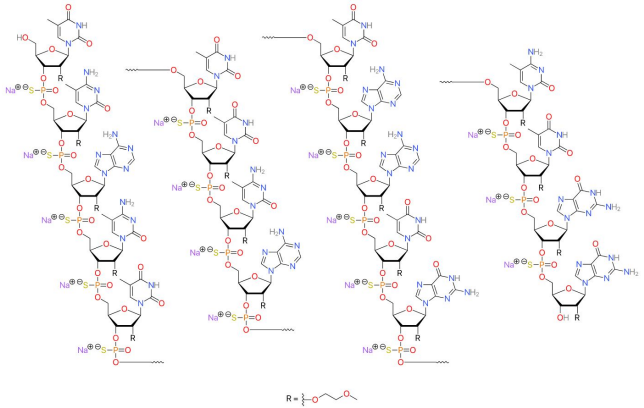


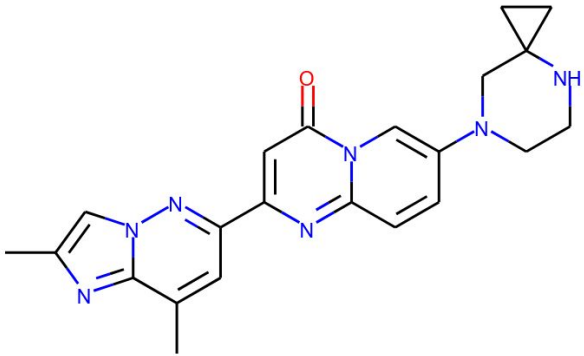
Three drugs of different modalities are approved to treat SMA

AAV9 capsid



SMN1 gene



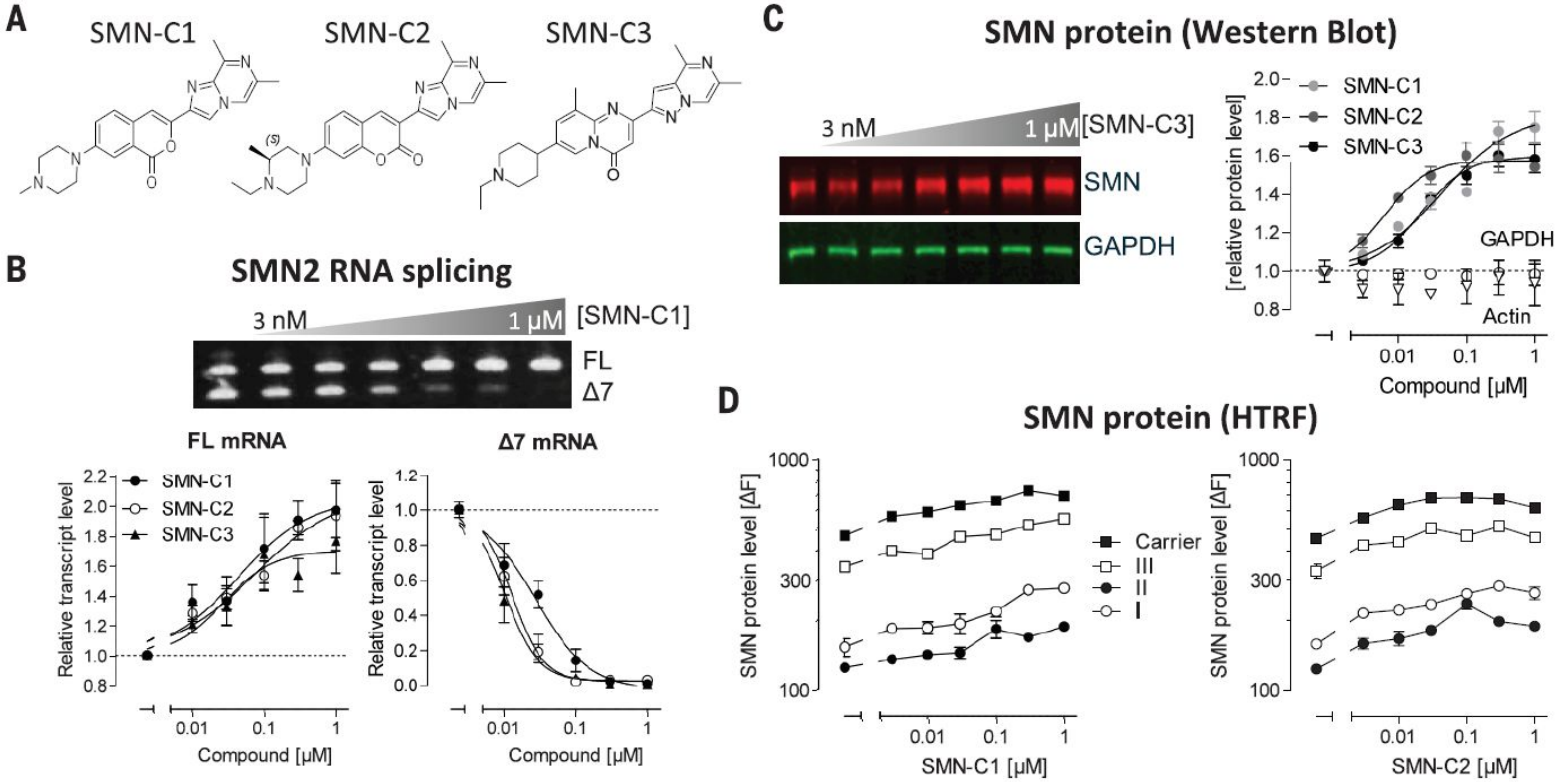


Onasemnogene Abeparvovec/ Zolgensma

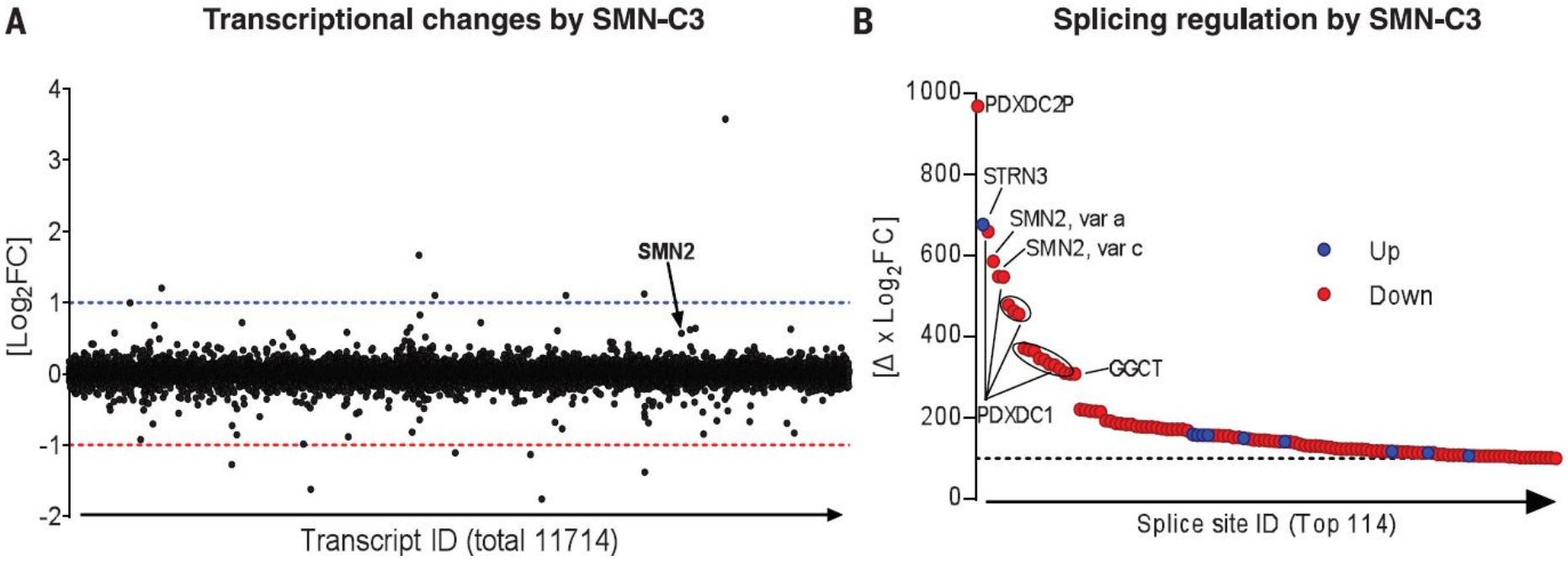
Nusinersen sodium/ Spinraza (CHEMBL3833342)

Risdiplam/ *Evrysdi* (CHEMBL4297528)

Small molecules were identified as RNA splicing modifiers

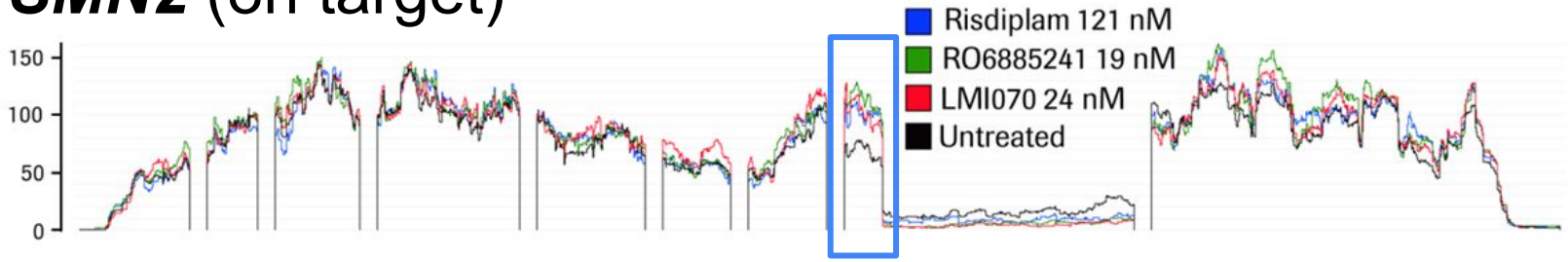


RNA sequencing confirms the specificity of SMN-C3

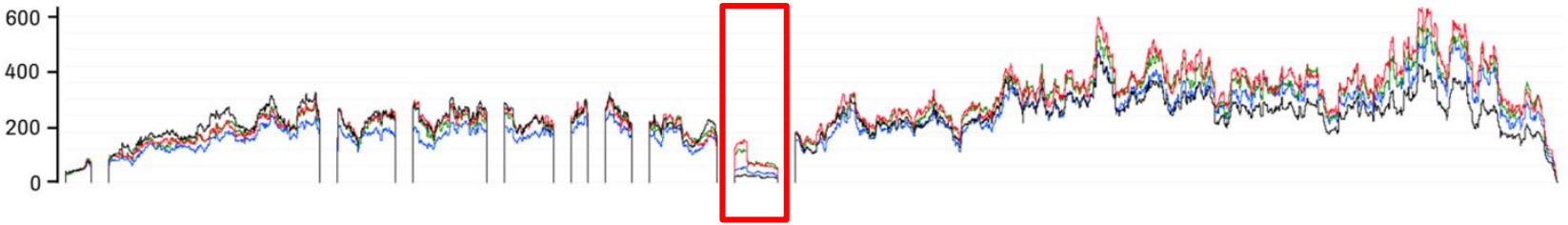


RNA sequencing confirms the superior safety profile of SMN-C3 over other compounds

SMN2 (on target)

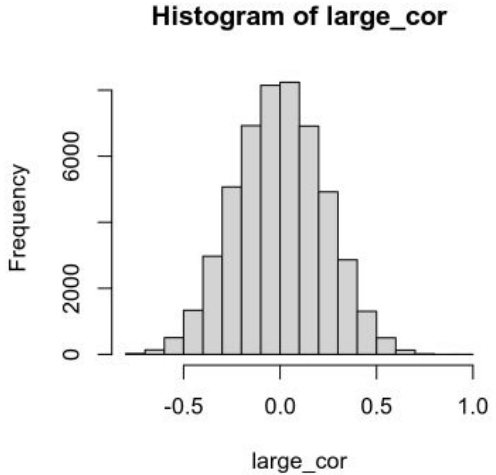


FOXM2 (off target)

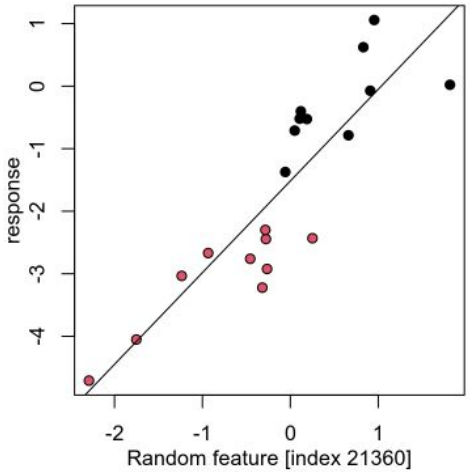


Given enough tests, there will be significant results

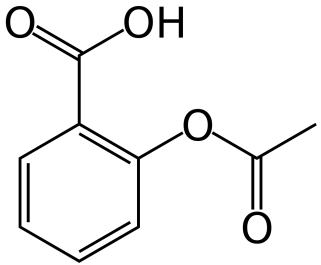
```r
set.seed(1887)
patient_group <- gl(2,10)
response <- c(rnorm(10, 0), rnorm(10, -3))
random_features_large <- matrix(rnorm(20*50000), nrow=20)
large_cor <- cor(response, random_features_large, method="spearman")
hist(large_cor)
```




```r
largest_cor_ind <- which.max(large_cor)
{
  compactPar()
  plot(random_features_large[, largest_cor_ind],
       response,
       bg=patient_group,pch=21,
       xlab=sprintf("Random feature [index %d]", largest_cor_ind))
  abline(lm(response ~ random_features_large[, largest_cor_ind]))
}
```



The road of MoA understanding can be 120 year long



Aspirin trademarked in 1899

Dai *et al*, Cell, 2019

Acetylation blocks cGAS activity and inhibits self-DNA-induced autoimmunity

- Acetylation suppresses cGAS activity
- Aspirin directly acetylates cGAS
- Aspirin inhibits cGAS-mediated interferon production
- Aspirin alleviates DNA-induced autoimmunity in AGS mouse models and patient cells

MoA understanding can be a long process full of surprises

Summary

1. In lead optimization and early development, we are interested in MoA of drug candidates *in vitro*, *in vivo*, and in human.
2. We can study MoA by modeling biological networks, for instance with ODE-based models and its variants.
3. We can also study MoA by performing omics experiments and analysing the data with statistical, machine-learning or AI tools. It is helpful to keep both advantages and challenges in mind.

https://xkcd.com/1217/