Quantifying Perturbed Biological Processes by Analyzing High-Throughput Data Using Causal Networks Models

Alain Sewer[1]; Florian Martin[1]; Ty M. Thomson[2]; Dmitry Vasilyev[2]; Jennifer Park[2]; Brian Frushour[2]; Jurjen W. Westra[2]; Julia Hoeng[1]; Manuel C. Peitsch[1]

[1]Philip Morris International R&D, Philip Morris Products S.A., Neuchâtel, Switzerland
[2]Selventa, One Alewife Center, Cambridge, MA, USA

Abstract

BACKGROUND: The scientific community faces an ongoing challenge in the analysis of high-throughput omics data to accurately characterize the biological processes and molecular mechanisms that are perturbed by diseases, drug treatments and environmental agents. A large number of biological processes taking place in healthy lung and vascular tissues have been recently captured into literature-based causal network models, which include cell proliferation, cellular stress responses, and inflammation.

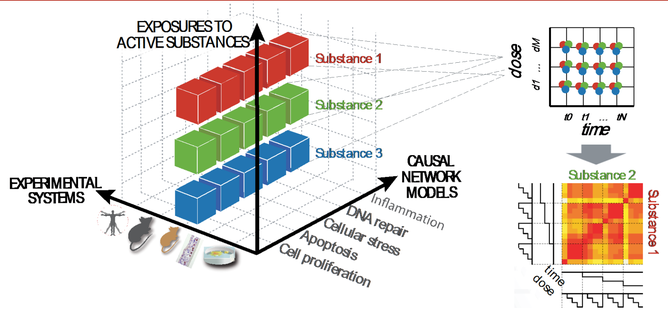
RESULTS: In order to leverage both the measured data and the prior biological knowledge contained in the network models, we developed a novel systems-level scoring approach called Network Perturbation Amplitude (NPA). The NPA algorithm computes the amplitudes of treatment- or disease-induced perturbations in a network model using transcriptomic data as an input, enabling the identification of activated molecular mechanisms. Targeted perturbations on in vitro systems, including normal human bronchial epithelial (NHBE) cells treated with a cyclin-dependent kinase inhibitor, were scored by the NPA approach, which was able to identify relevant mechanisms and to provide a comparative quantitation of the impacted biological processess. When we evaluated perturbations in more complex experimental systems, such as rats exposed to formaldehyde in vivo, the NPA results captured the degree of activation of known mechanistic effects mediated by the exposures.

CONCLUSIONS: The ability of the NPA approach to infer the degree of activation of molecular mechanisms and cellular processes provides a powerful means to broadly assess biological activity using a variety of transcriptomic datasets. We have developed and demonstrated the value of a novel integrative, systems biology approach to evaluate disease progression. We expect the range of applications of the NPA approach to be extended to drug safety and discovery as well as to guide biomarker development.

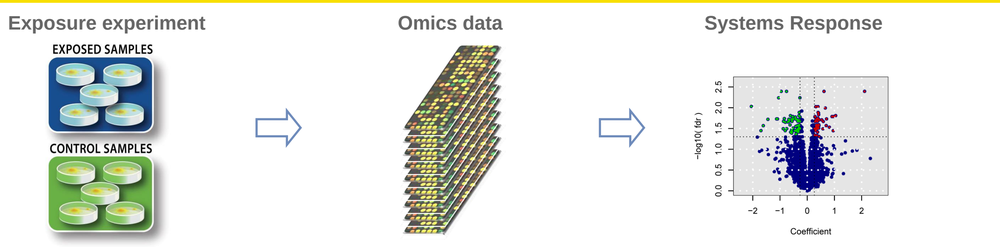
Approach

The quantification of perturbed biological processes is a multistep approach integrating transcriptomic data and prior biological knowledge structured into causal network models [1].

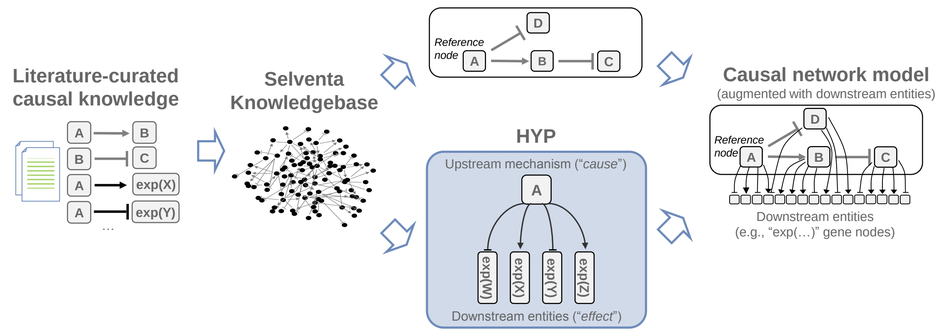
STEP 1: DESIGN OF THE EXPERIMENTS
The biological perturbations induced by exposures to active substances (drug, consumer products, environmental pollutants, etc.) will be evaluated along one axis of the design space {system, active substance, network model}.



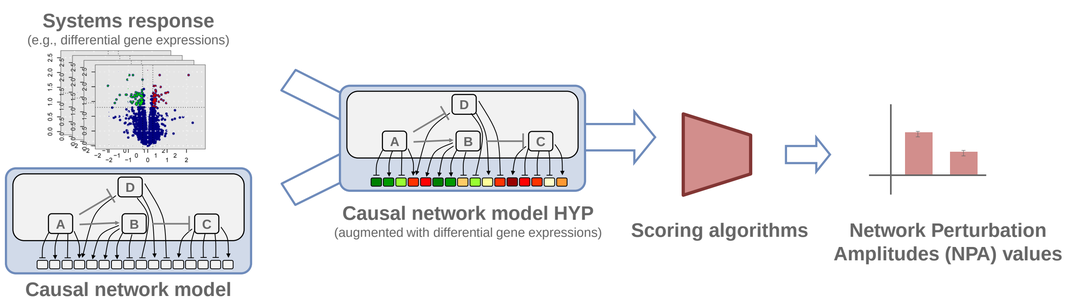
STEP 2: MEASUREMENT OF THE SYSTEMS RESPONSE
High-throughput omics technologies and standard bioinformatic tools are used to obtain the systems-wide response to the exposure to the active substance (e.g., the differential expressions of transcribed genes in a treatment vs. control comparison).



STEP 3: CONSTRUCTION OF CAUSAL NETWORKS MODELS
Public and/or proprietary prior biological knowledge encoded in the BEL language [2] is structured into a collection of biological network models capturing the essential cellular processes and the relevant response mechanisms, e.g., cell proliferation or cellular stress [3].
These network models are complemented by the causal connections between the BEL-encoded biological mechanisms and the downstream entities measured in the systems response (e.g., the genes probed by a transcriptomic platform) [4].



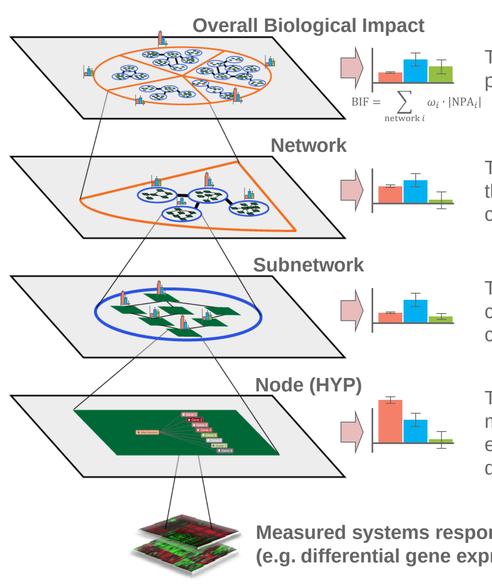
STEP 4: CALCULATION OF THE NETWORK PERTURBATIONS AMPLITUDES (NPA)
The systems response and one causal network model are combined to quantify the treatment-induced perturbations of the biological mechanisms described by the network model [4,5].



Results

STEP 5: CALCULATION OF THE BIOLOGICAL IMPACT FACTOR (BIF)
The NPA approach is combined with the complete collection of causal network models to quantify the overall impact of the applied treatment on the considered biological system [6].

The calculation of the BIF enables a top-down description of the treatment-induced perturbation by leveraging the hierarchical structure of the network models.



$$BIF = \sum_{\text{network } i} \omega_i \cdot |NPA_i|$$

The Biological Impact Factor (BIF) is computed as a combination of the perturbation amplitudes of the network models.

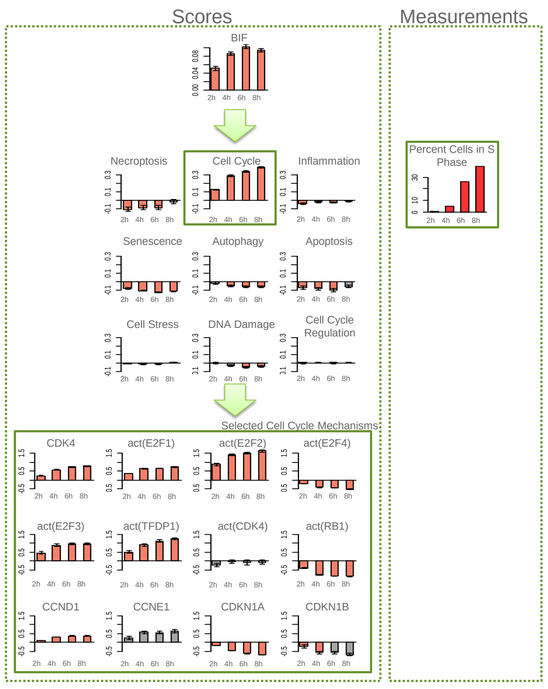
The network perturbation amplitudes are computed as combinations of the perturbation amplitudes of the subnetworks constituting the considered network.

The subnetwork perturbation amplitudes are computed as combinations of the perturbation amplitudes of the nodes (HYPs) constituting the considered subnetwork.

The node (HYP) perturbation amplitudes are computed from the measured systems response, e.g., as a combination of the differential expressions of downstream genes of the biological mechanism described by the HYP.

Proof-of-principle application for a targeted perturbation on an *in vitro* system
(normal human bronchial epithelial cells treated with a cyclin-dependent kinase (CDK) inhibitor [7])

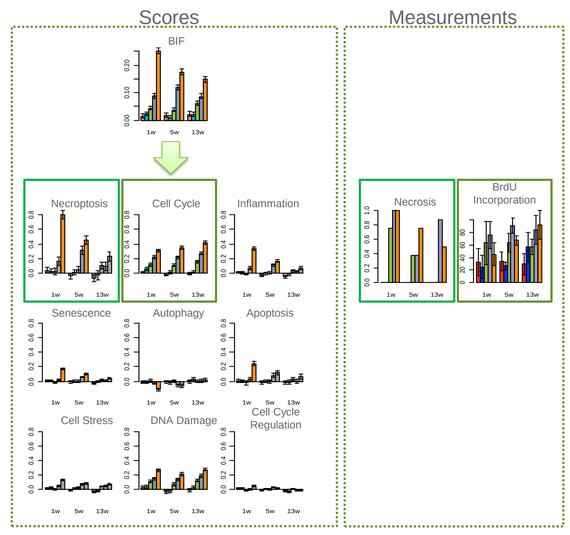
- Exposures: CDK inhibitor washout (=treatment) and continued exposure to the CDK inhibitor (=control).
- Gene differential expressions (treatment vs. control) were measured 2,4,6,8 hours after washout.
- Compared with percentage of cells in S-phase measured by flow cytometry.



→ Quantification of network perturbations identified the *Cell Cycle* network as the primary responder to CDK inhibitors, produced the expected temporal response, and showed that nodes proximal to CDK4/6, which were inhibited, also showed expected temporal response.

Concrete application for a more complex perturbation on an *in vivo* system
(nasal epithelium from rats exposed to several doses of formaldehyde [8])

- Exposures: five doses (0.7, 2, 6, 10, and 15ppm) of formaldehyde compared to time-matched controls.
- Gene differential expressions (non-zero dose vs. control) were measured after 1, 5, 13 weeks of exposure.
- Compared with necrosis (percentage of animals showing nasal necrosis/erosion) and cell cycle (BrdU incorporation) measurements.



→ Quantification of network perturbations identified significant perturbation in *Necrosis* and *Cell Cycle* networks, consistent with explicit experimental measurements. The network perturbations were also consistent with other known effects of formaldehyde, including inflammation and DNA damage.

Conclusions

- The quantification of network perturbation amplitudes (NPA and BIF) is a key component of a multi-step approach for assessing the biological impact of active substances (drug treatments, environmental factors, or toxic substances).
- The NPA and BIF methodologies leverages a collection of hierarchically-structured biological network models and enables a description of the treatment-induced perturbations at multiple resolution levels.
- The applications of the approach to concrete *in vitro* and *in vivo* exposure experiments yielded meaningful results and adequate correspondence with independently measured phenotypic endpoints.

References

[1] Hoeng et *al.*, *A network-based approach to quantifying the impact of biologically active substances*, Drug Discov Today 2012.
[2] Selventa, *The openBEL Portal,* http://www.openbel.org/ 2012.
[3] Westra et *al.*, *Construction of a computable cell proliferation network focused on non-diseased lung cells*, BMC Syst Biol 2011; Schlage et *al.*, *A computable cellular stress network model for non-diseased pulmonary and cardiovascular tissue*, BMC Syst Biol 2011; Gebel et *al.*, *Construction of a computable network model for DNA damage, autophagy, cell death, and senescence*, Bioinform Biol Insights 2013; Westra et *al.*, *A modular cell-type focused inflammatory process network model for non-diseased pulmonary tissue*, Bioinform Biol Insights 2013.
[4] Martin et *al.*, *Assessment of network perturbation amplitudes by applying high-throughput data to causal biological networks*, BMC Syst Biol 2012.
[5] Martin et *al.*, *Quantification of biological network perturbations: Impact assessment and diagnosis using causal biological networks*, submitted to *Bioinformatics*.
[6] Thomson et *al.*, *Quantitative assessment of biological impact using transcriptomic data and mechanistic network models*, Toxicol Appl Pharmacol. 2013.
[7] Belcastro et *al.*, *Systematic verification of upstream regulators of a computable cellular proliferation network model on non-diseased lung cells using a dedicated dataset*, Bioinform Biol Insights 2013.
[8] Andersen et *al.*, *Formaldehyde: integrating dosimetry, cytotoxicity, and genomics to understand dose-dependent transitions for an endogenous compound*, Toxicol. Sci. 2010.

PMI RESEARCH & DEVELOPMENT

The International Conference on Systems Biology (ICSB)
Copenhagen, Denmark
30. August - 3. September 2013

Philip Morris International Research & Development, Quai Jeanrenaud 5, 2000 Neuchâtel, Switzerland
T: +41 58 242 21 11, **F**: +41 58 242 28 11, **W**: www.pmi.com