PMI RESEARCH & DEVELOPMENT

Genomics and Bioinformatics of Lung Diseases

Dr. Nikolai V. Ivanov

Manager, Genomics Research & Technology

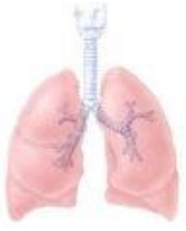
on behalf of

Biological Systems Research Department

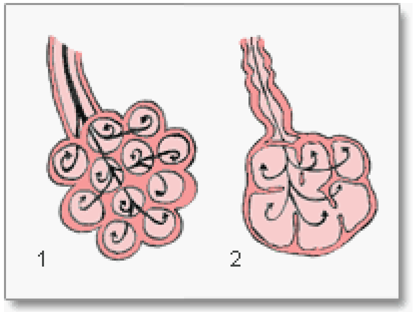
April 25-27th, 2013

Background

- Lung Cancer, CVD and COPD
- Chronic obstructive pulmonary disease (COPD) is the 5th leading cause of death worldwide and cigarette smoking is the main cause of COPD
- Main features of COPD
  - – Lung Inflammation, Chronic Bronchitis and Bronchiolitis
  - – Lung Emphysema



afarewellrescue.com



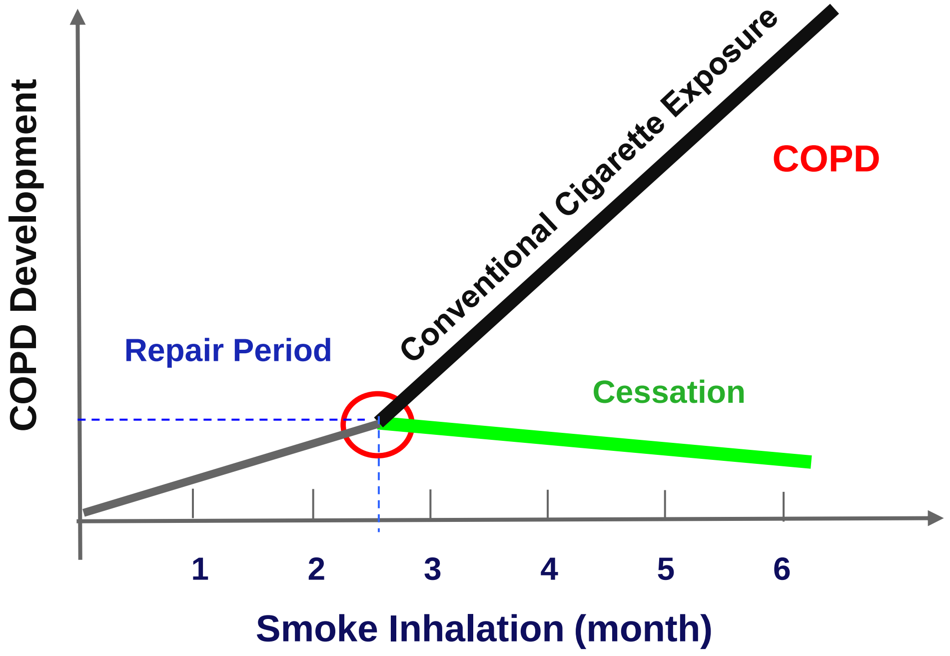
1: Healthy Alveoli

2: Emphysematous Alveoli

www.lung.ca/.../emphysema-emphyseme/index_e.php

Background

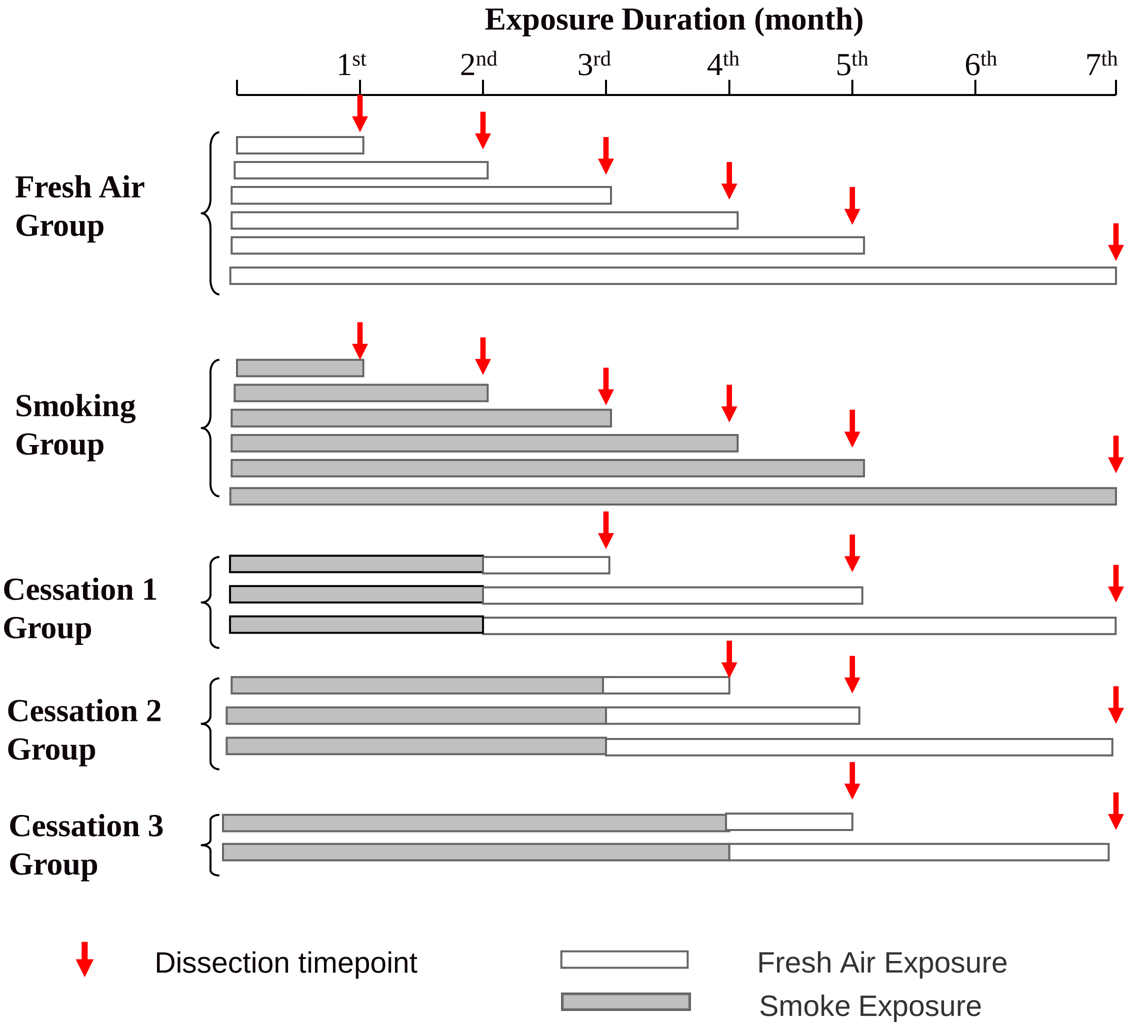
- Human (Kohansal et al 2009) and animal experimental data (Milot J et al 2007; Wright JL et al 1994, 2006) suggest that lung can repair smoke-induced damage for some period before it eventually loses this ability to repair



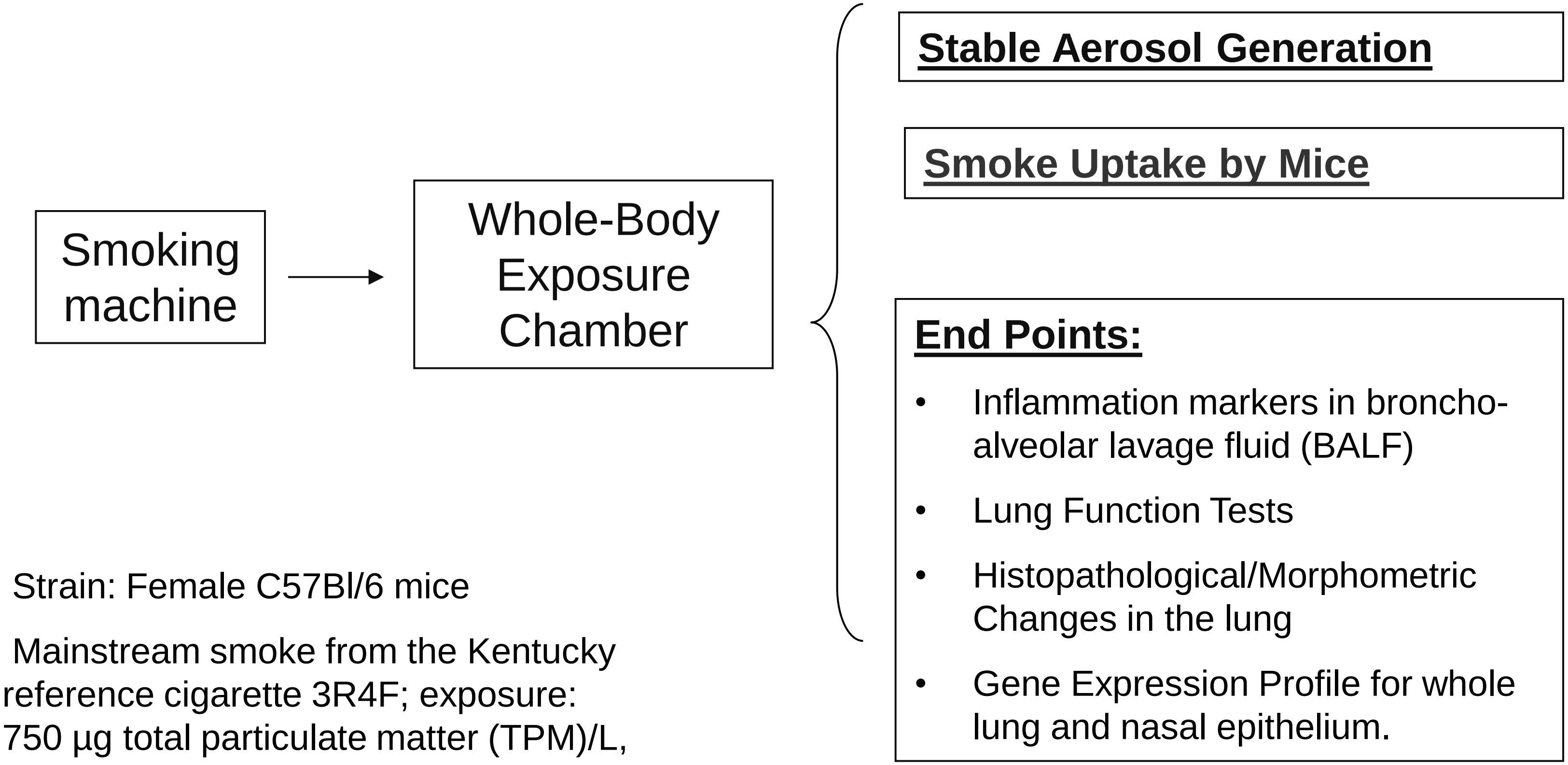
Objectives

- The objectives of COPD studies are
  - To establish C57BL/6 as a mouse model of cigarette smoke-induced COPD
  - To explore the impact of exposure to 3R4F on the onset and the development of COPD in C57BL/6 mice
  - To explore the impact of smoking cessation on the recovery of COPD in C57BL/6 mice
  - To determine the time point that animals start to develop irreversible emphysema
  - To assess the impact of smoking exposure and smoking cessation on the development of COPD based on whole genome gene expression profile.
  - Understand the underlying molecular mechanisms perturbed by smoke and smoking cessation

Study Design - Exposure Groups and Duration



Study Design – Experimental Set up



Smoking machine

Whole-Body Exposure Chamber

Stable Aerosol Generation

Smoke Uptake by Mice

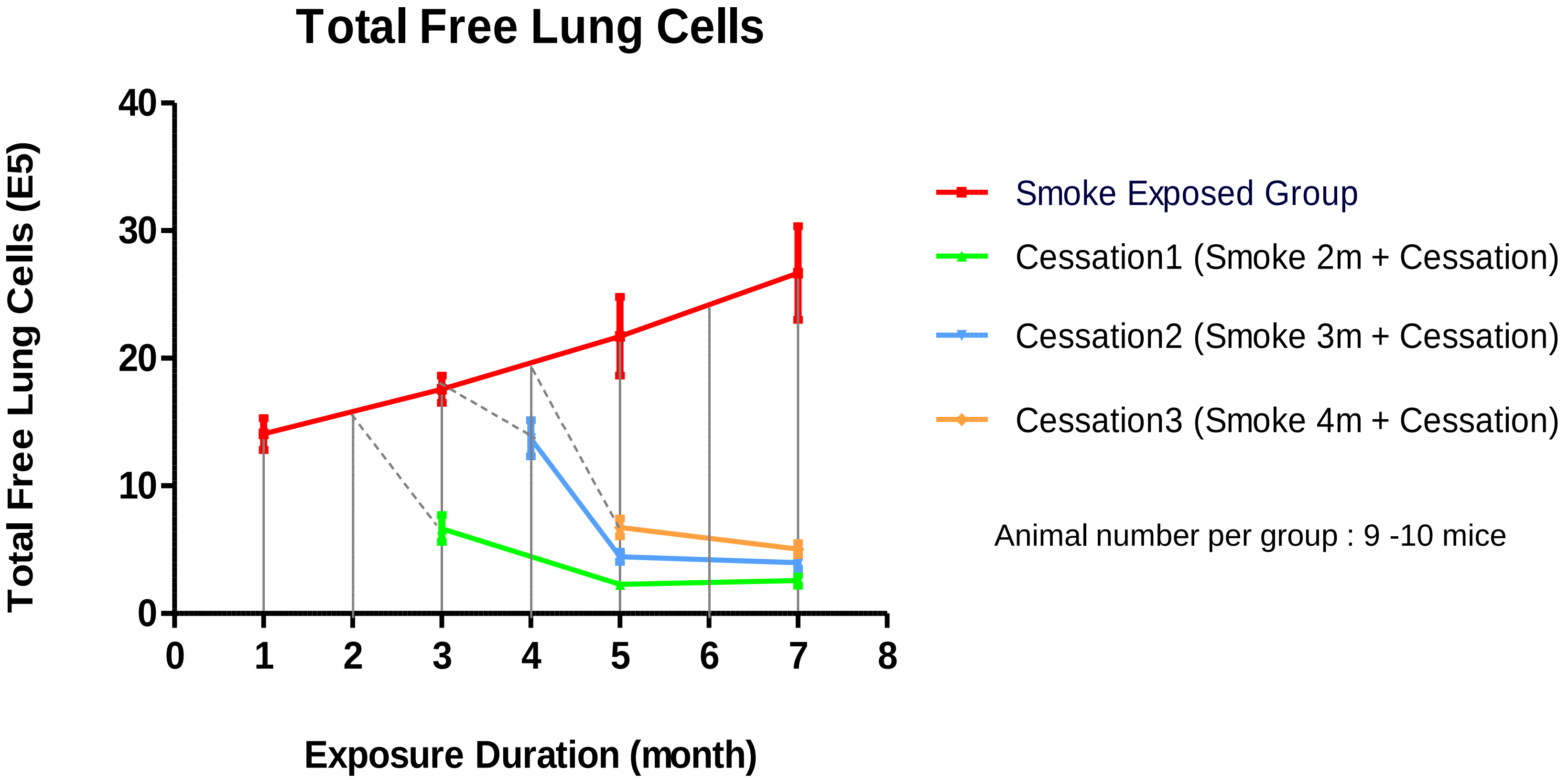
End Points:

- Inflammation markers in broncho-alveolar lavage fluid (BALF)
- Lung Function Tests
- Histopathological/Morphometric Changes in the lung
- Gene Expression Profile for whole lung and nasal epithelium.

- Strain: Female C57Bl/6 mice
- Mainstream smoke from the Kentucky reference cigarette 3R4F; exposure: 750 µg total particulate matter (TPM)/L, 4 hours/day, 5 days/week
- Total duration of the study: 7 months

Lung Inflammation - BALF

Bronchoalveolar Lavage Fluid (BALF) Cells

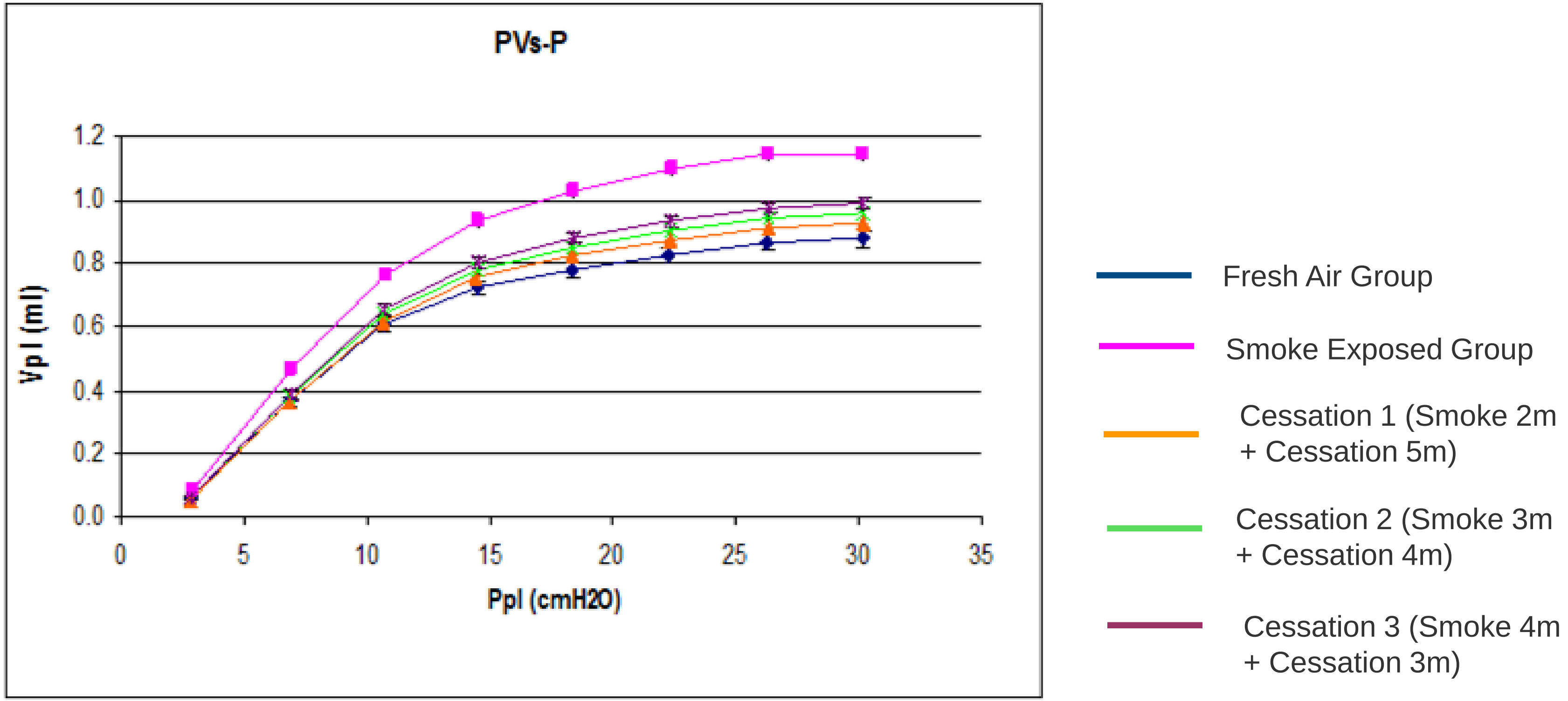


Animal number per group : 9 -10 mice

PMI RESEARCH & DEVELOPMENT

Emphysema: Lung Function

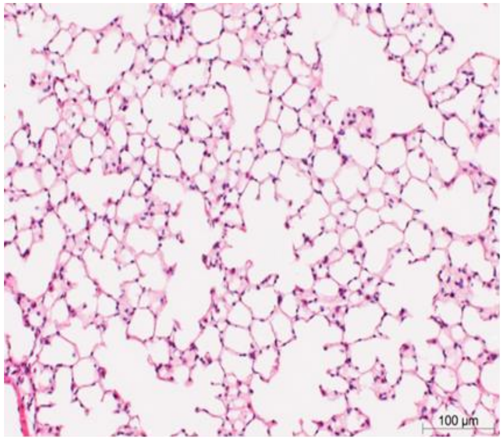
Pressure-Volume Loop – Inhalation Curve



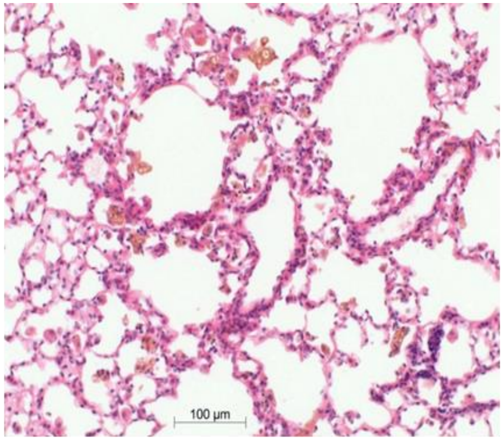
7 Months Exposure

Emphysema: Histopathology/Morphometry

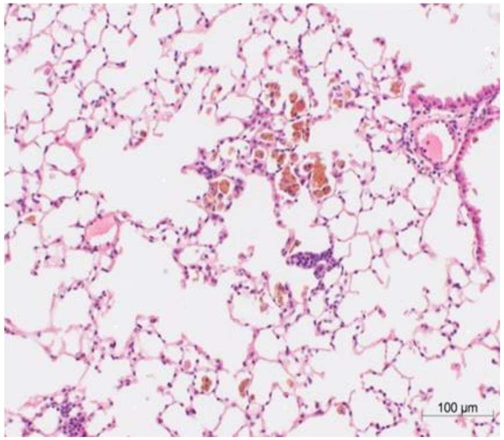
Lung Morphometry after 7 Months of Exposure



Fresh Air Group

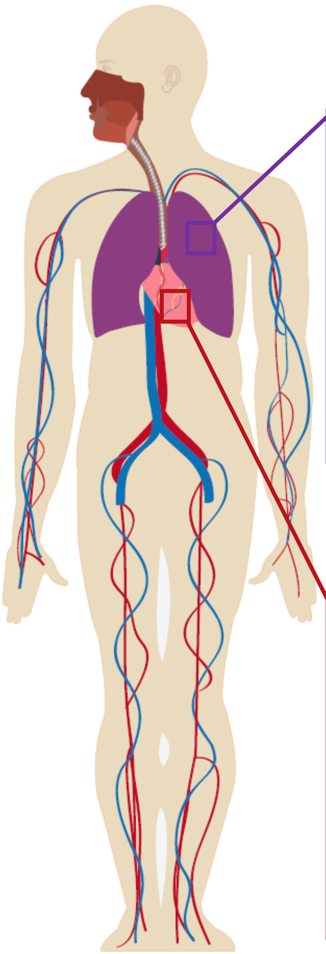


Smoke Exposed Group

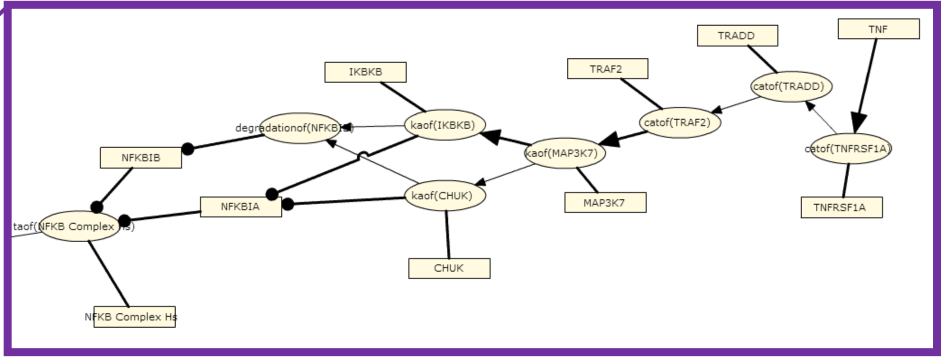


Cessation Group
(Smoke 2m + Cessation 5m)

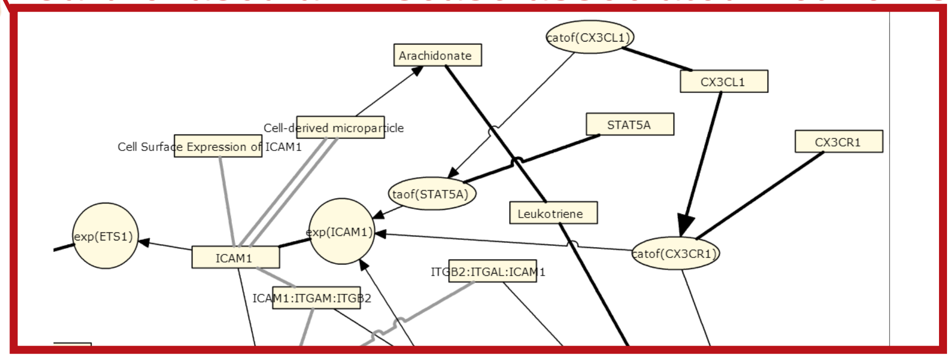
Network Biology for Systems Toxicology and Biomarker Discovery



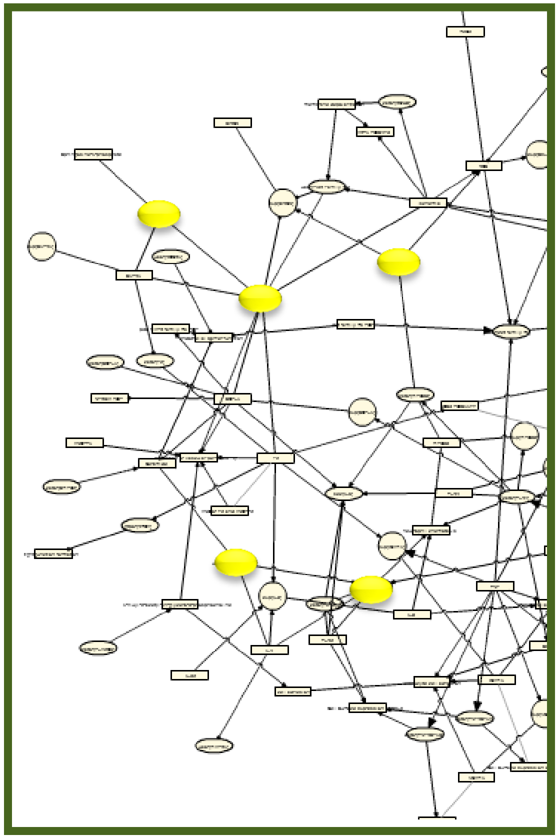
Respiratory Disease-associated Networks



Cardiovascular Disease-associated Networks

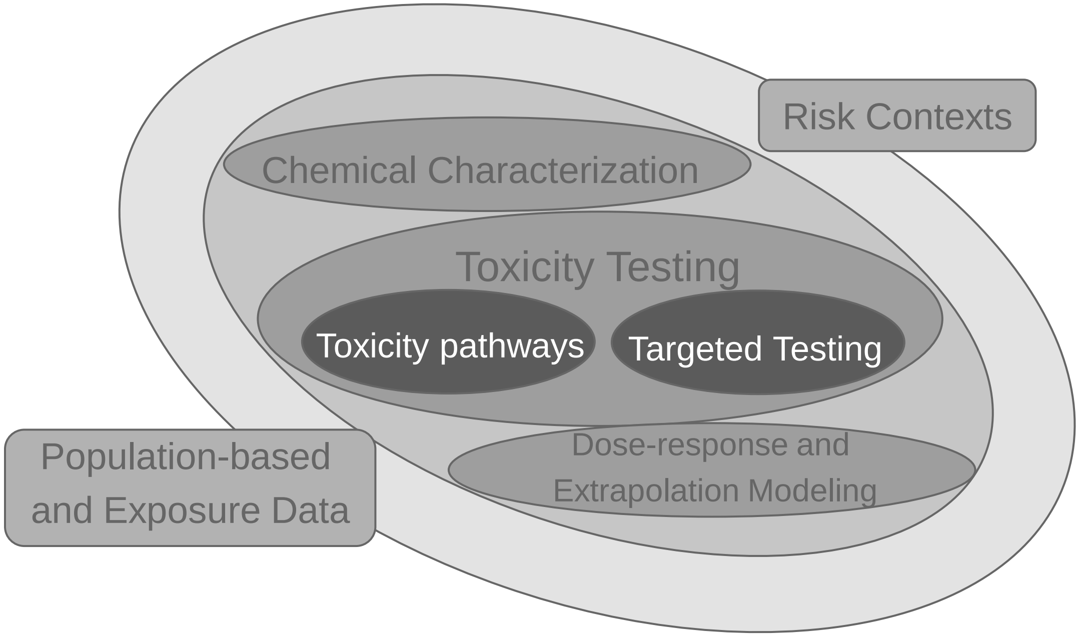


Biomarker



PMI RESEARCH & DEVELOPMENT

New Paradigm in Toxicity Testing in the 21st Century



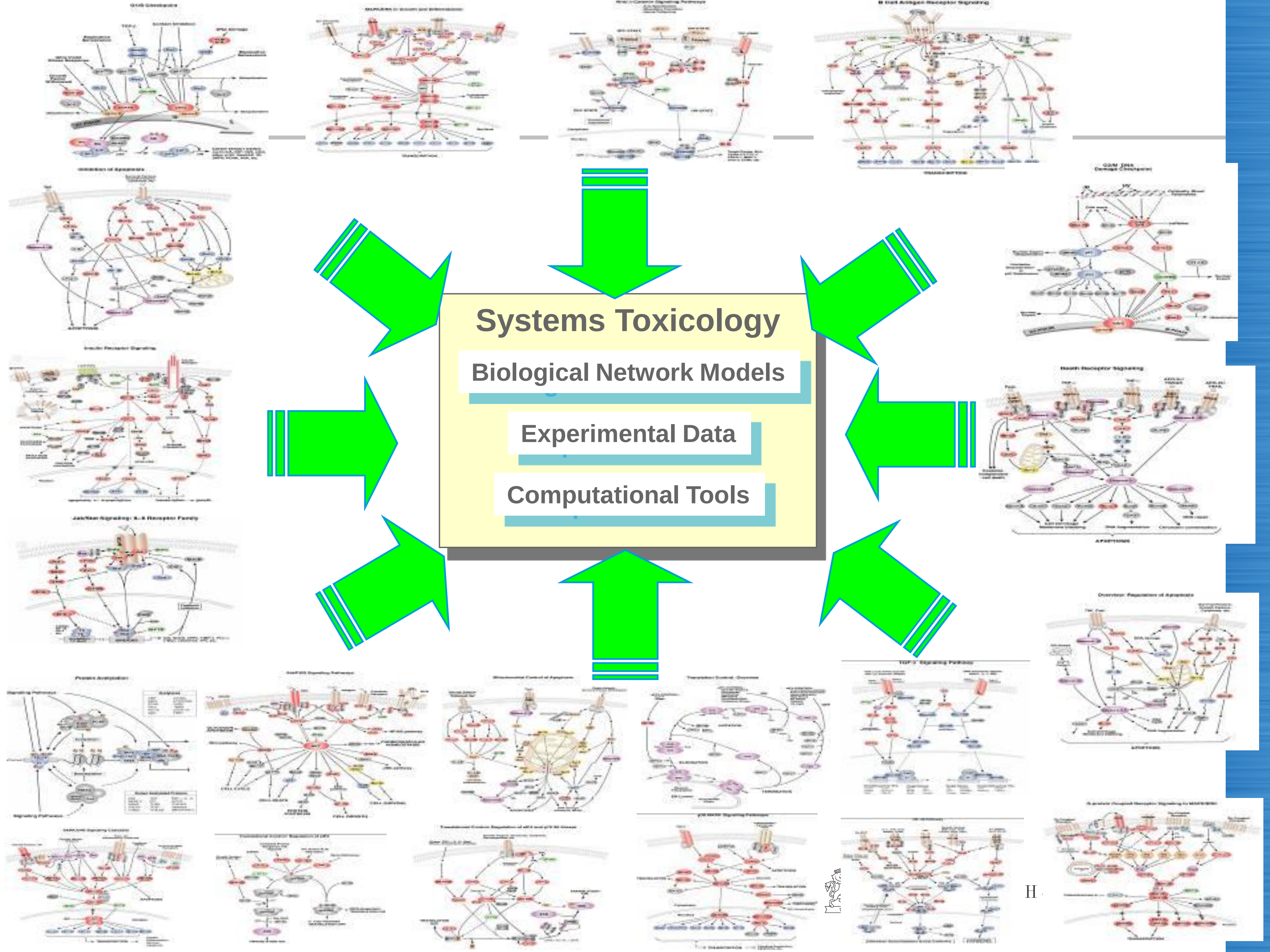
2007 – National Academy of Sciences, issued a landmark report "Toxicity Testing in the 21st Century: A Vision and Strategy"

Vision: Move away from evaluating apical health endpoints in animals **-** turn toward identification of toxicity pathways in *in-vitro* test systems

Transformative change in toxicity testing

- to achieve testing of broad coverage of chemicals, mixtures, outcomes and life stages
- to significantly increase human relevance
- to reduce the cost and time required to conduct chemical safety assessments
- to reduce and potentially eliminate high-dose animal testing

Adapted from: *Toxicity Testing in the Twenty-first Century: A Vision and a Strategy.* The Board on Environmental Studies and Toxicology http://nationalacademies.org/best.



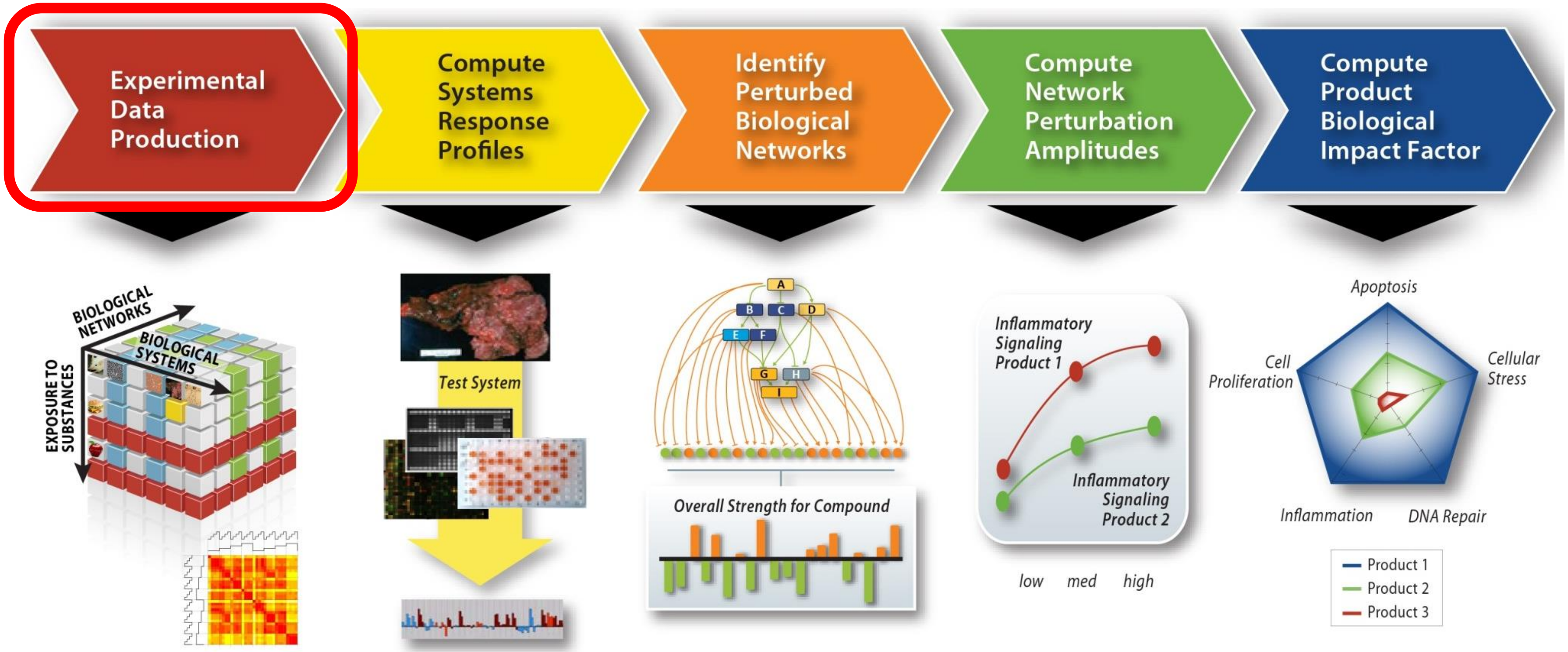
Systems Toxicology

Biological Network Models

Experimental Data

Computational Tools

Quantitative Mechanism-Based Systems Impact Assessment



Experimental Data Production

Compute Systems Response Profiles

Identify Perturbed Biological Networks

Compute Network Perturbation Amplitudes

Compute Product Biological Impact Factor

MEASURE MULTIPLE PERTURBATIONS ACROSS MULTIPLE EXPERIMENTAL SYSTEMS

COMPUTE DIFFERENTIAL RESPONSE PROFILES FROM LARGE NUMBER OF MEASURED BIOLOGICAL VARIABLES

IDENTIFICATION OF BIOLOGICAL MECHANISMS PERTURBED BY A PRODUCT

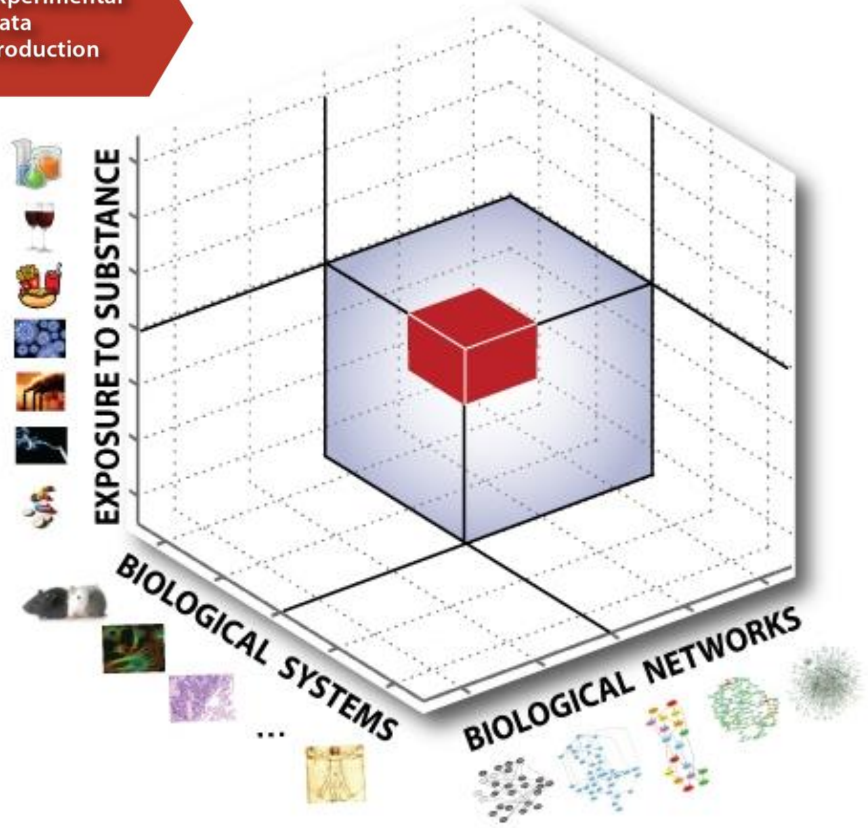
QUANTIFICATION OF PERTURBATIONS OF INDIVIDUAL NETWORKS

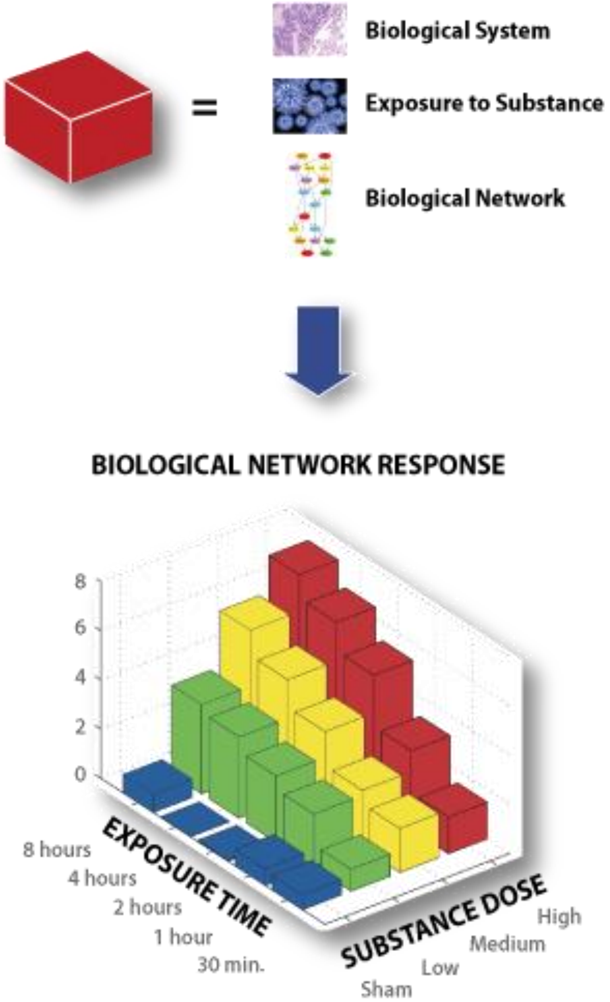
QUANTIFICATION OF OVERALL PERTURBATION INDUCED BY PRODUCTS

Hoeng J, Deehan R, Pratt D, Martin F, Sewer A, Thomson TM, Drubin DA, Waters CA, de Graaf D, and Peitsch MC.
A network-based approach to quantifying the impact of biologically active substances. *Drug Discov Today* 17: 413-418, 2012.

Quantitative Mechanism-Based Systems Impact Assessment

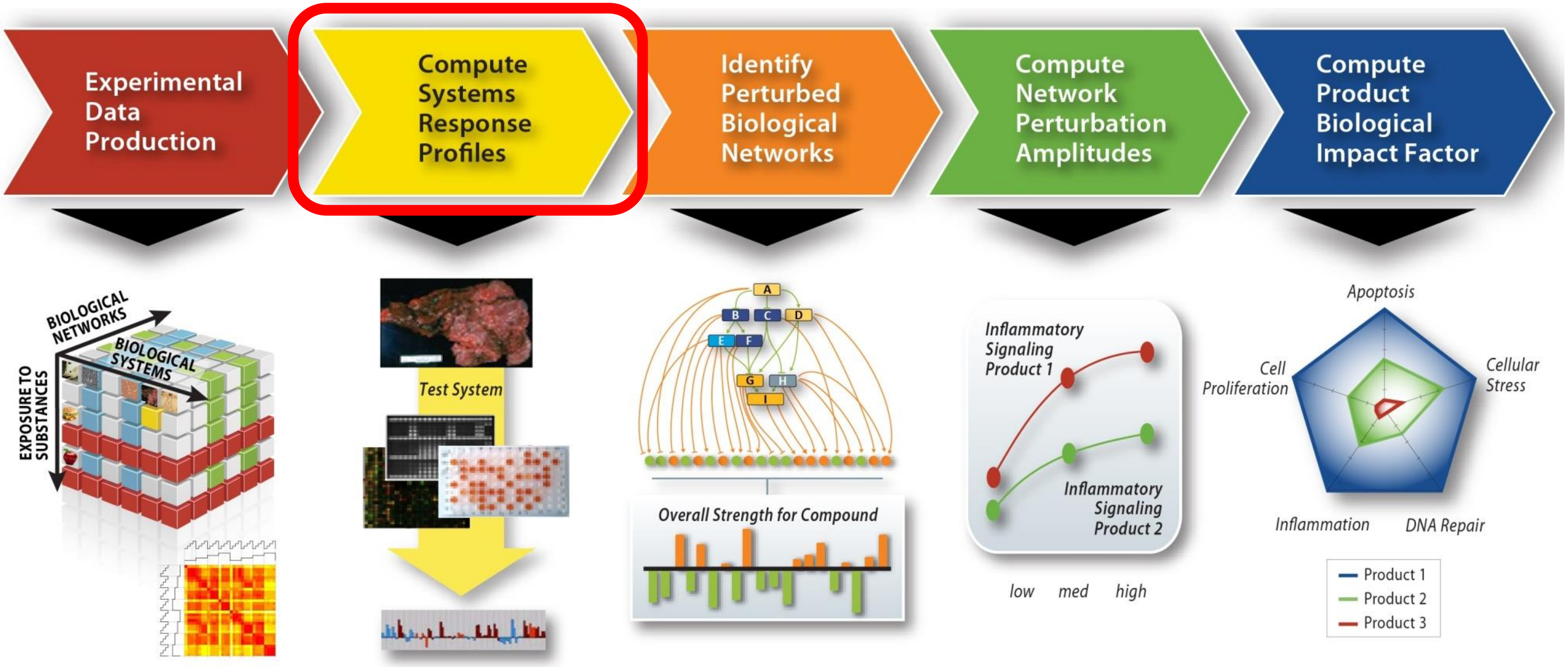
Experimental Data Production





Systematic experiments across numerous test systems

Quantitative Mechanism-Based Systems Impact Assessment



MEASURE MULTIPLE PERTURBATIONS ACROSS MULTIPLE EXPERIMENTAL SYSTEMS

COMPUTE DIFFERENTIAL RESPONSE PROFILES FROM LARGE NUMBER OF MEASURED BIOLOGICAL VARIABLES

IDENTIFICATION OF BIOLOGICAL MECHANISMS PERTURBED BY A PRODUCT

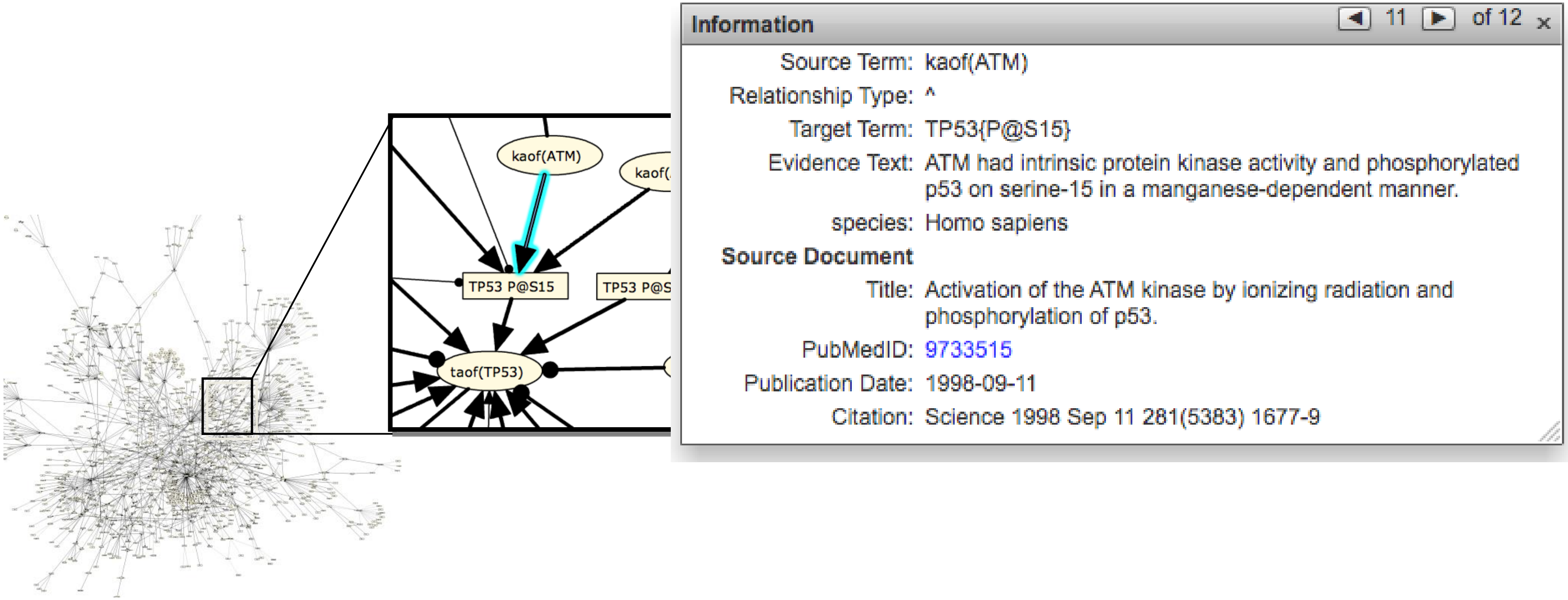
QUANTIFICATION OF PERTURBATIONS OF INDIVIDUAL NETWORKS

QUANTIFICATION OF OVERALL PERTURBATION INDUCED BY PRODUCTS

Hoeng J, Deehan R, Pratt D, Martin F, Sewer A, Thomson TM, Drubin DA, Waters CA, de Graaf D, and Peitsch MC.
A network-based approach to quantifying the impact of biologically active substances. *Drug Discov Today* 17: 413-418, 2012**.**

Construction of BEL Encoded Network Models

kaof(ATM) increases **TP53(P@S15)**



Source Term: kaof(ATM)

Relationship Type: ^

Target Term: TP53{P@S15}

Evidence Text: ATM had intrinsic protein kinase activity and phosphorylated p53 on serine-15 in a manganese-dependent manner.

species: Homo sapiens

Source Document

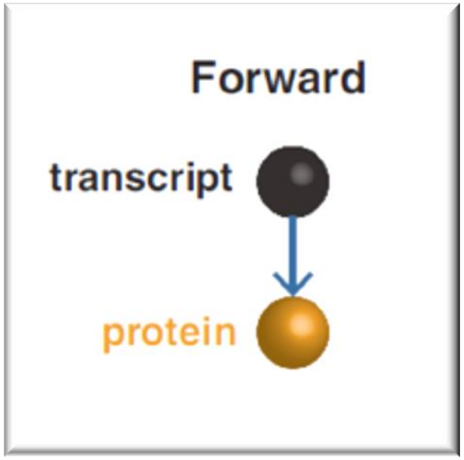
Title: Activation of the ATM kinase by ionizing radiation and phosphorylation of p53.

PubMedID: 9733515

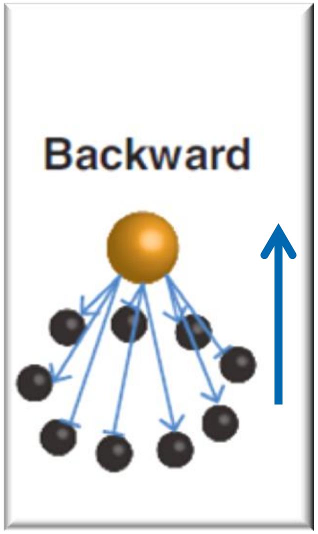
Publication Date: 1998-09-11

Citation: Science 1998 Sep 11 281(5383) 1677-9

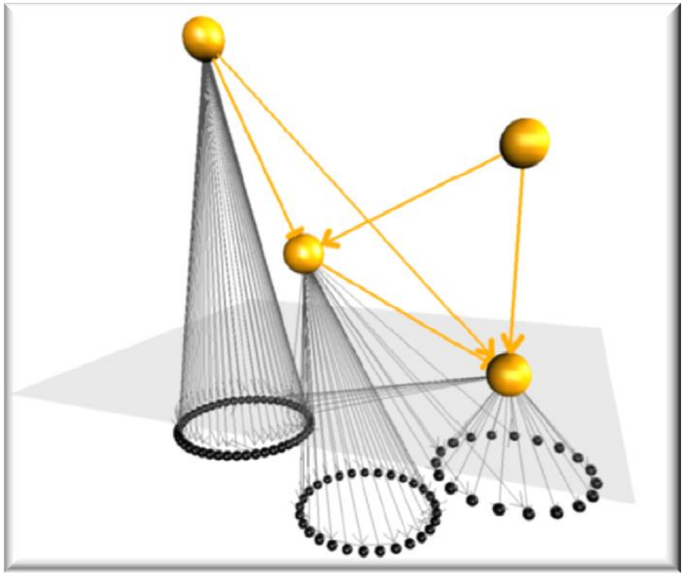
How Do We Use Transcriptomic Data?



Identification of potential downstream consequences of the data observations



Identification of upstream entities that are potential causes of the observed differential elements

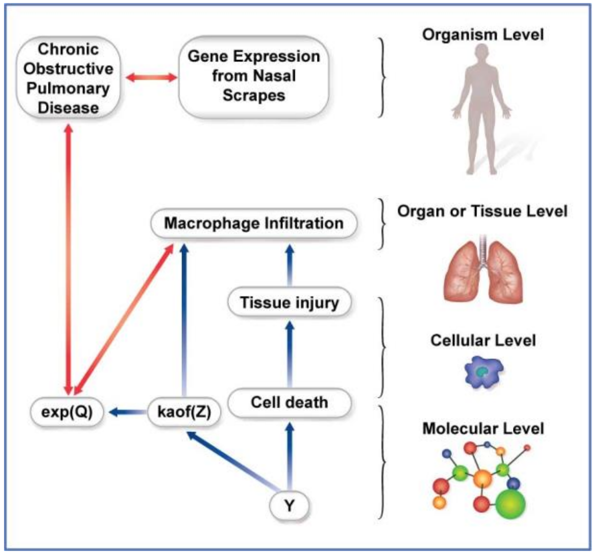


Martin F, Thomson TM, Sewer A, Drubin DA, Mathis C, Weisensee D, Pratt D, Hoeng J, and Peitsch MC.
Assessment of network perturbation amplitude by applying high-throughput data to causal biological networks. *BMC Syst Biol* 6: 54, 2012

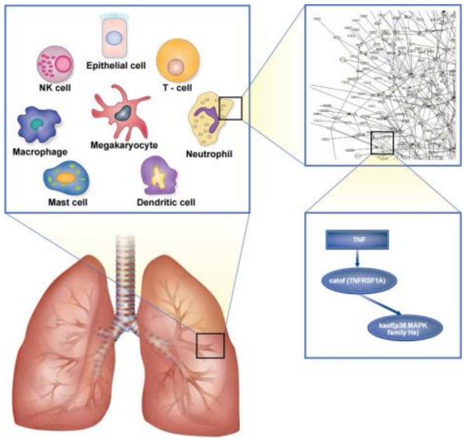
Knowledgebase Model Properties

Open BEL

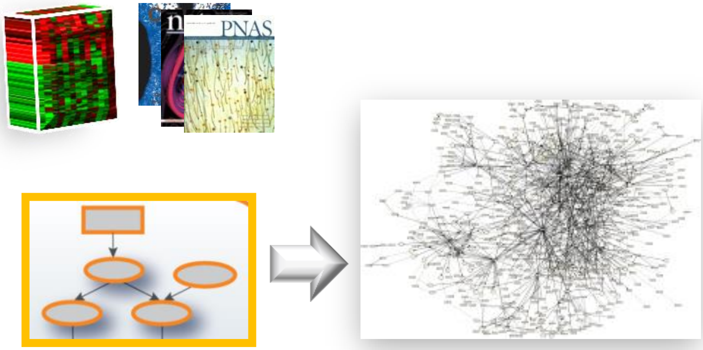
Computable



Capture wide range of biology



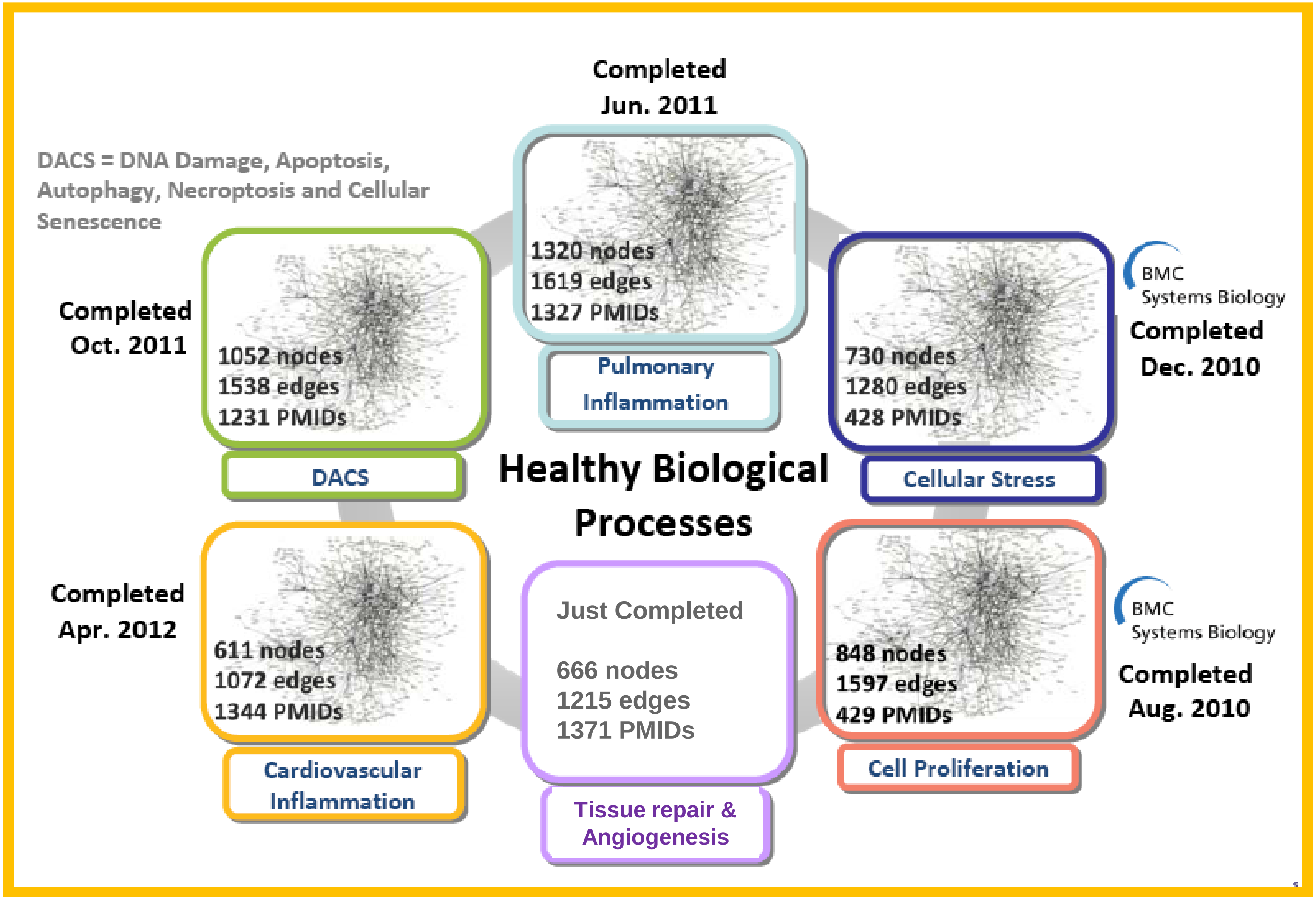
Hierarchal structure



Updatable

Quantitative Mechanism-Based Systems Impact Assessment

Build and Maintain Biological Networks



DACS = DNA Damage, Apoptosis, Autophagy, Necroptosis and Cellular Senescence

Completed Jun. 2011

1320 nodes
1619 edges
1327 PMIDs

Pulmonary Inflammation

Completed Oct. 2011

1052 nodes
1538 edges
1231 PMIDs

DACS

BMC Systems Biology

Completed Dec. 2010

730 nodes
1280 edges
428 PMIDs

Cellular Stress

Healthy Biological Processes

Completed Apr. 2012

611 nodes
1072 edges
1344 PMIDs

Cardiovascular Inflammation

Just Completed

666 nodes
1215 edges
1371 PMIDs

Tissue repair & Angiogenesis

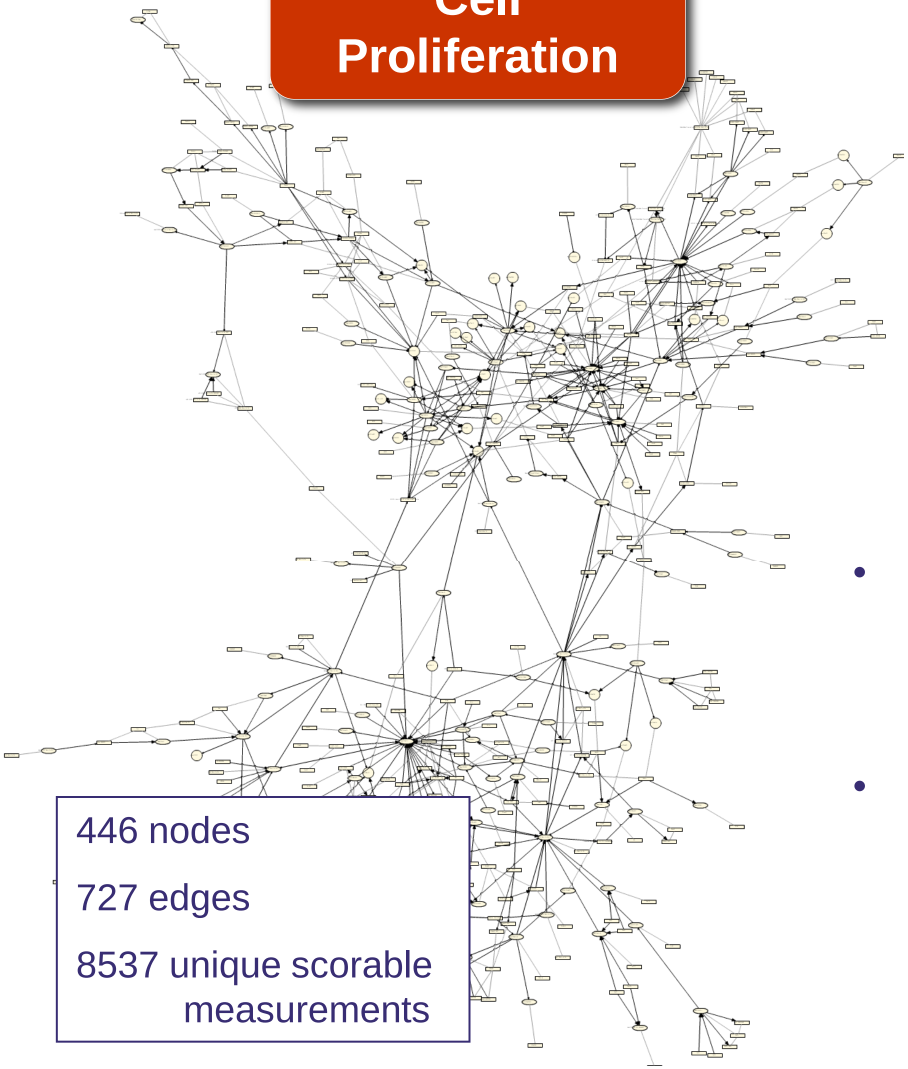
BMC Systems Biology

Completed Aug. 2010

848 nodes
1597 edges
429 PMIDs

Cell Proliferation

Relevant Biological Network Models

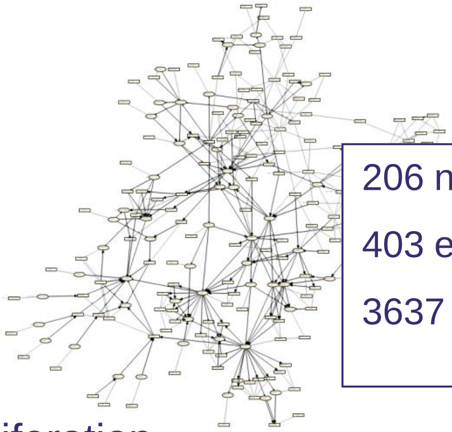


Cell Proliferation

446 nodes

727 edges

8537 unique scorable measurements



Inflammation

206 nodes

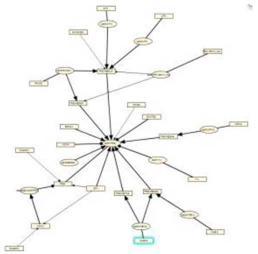
403 edges

3637 unique scorable measurements

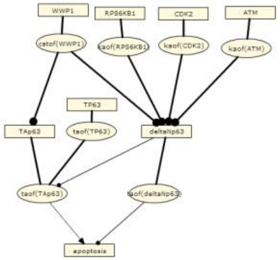
- Cell Proliferation
  - Represents growth factor signaling pathways and cell cycle regulation
  - Data includes measured phenotype
- Inflammation
  - Represents cytokine signaling to NFKB and AP-1 activation
  - No quantified phenotype associated with data

PMI RESEARCH & DEVELOPMENT

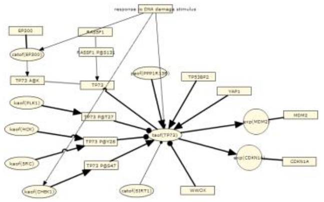
DNA Damage Subnetworks in Selventa Knowledgebase



Components affecting
TP53 activity

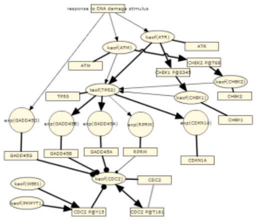


Components affecting
TP63 activity

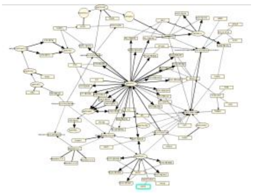


Components affecting
TP73 activity

DNA damage to
G1/S checkpoint



DNA damage to
G2/M checkpoint



Double Strand
Break Response

Inhibition of
DNA repair

NER/XP pathway

Single Strand
Break Response

TP53 Transcriptional
Signature

Using Selventa Knowledgebase®[1], a repository containing cause-and-effect relationships extracted from 35000 PMIDs, one can assemble networks as models of biological mechanism within clearly defined boundaries

1. *Selventa, http://selventa.com/*

PMI RESEARCH & DEVELOPMENT

Quantitative Mechanism-Based Systems Impact Assessment

Experimental Data Production

Compute Systems Response Profiles

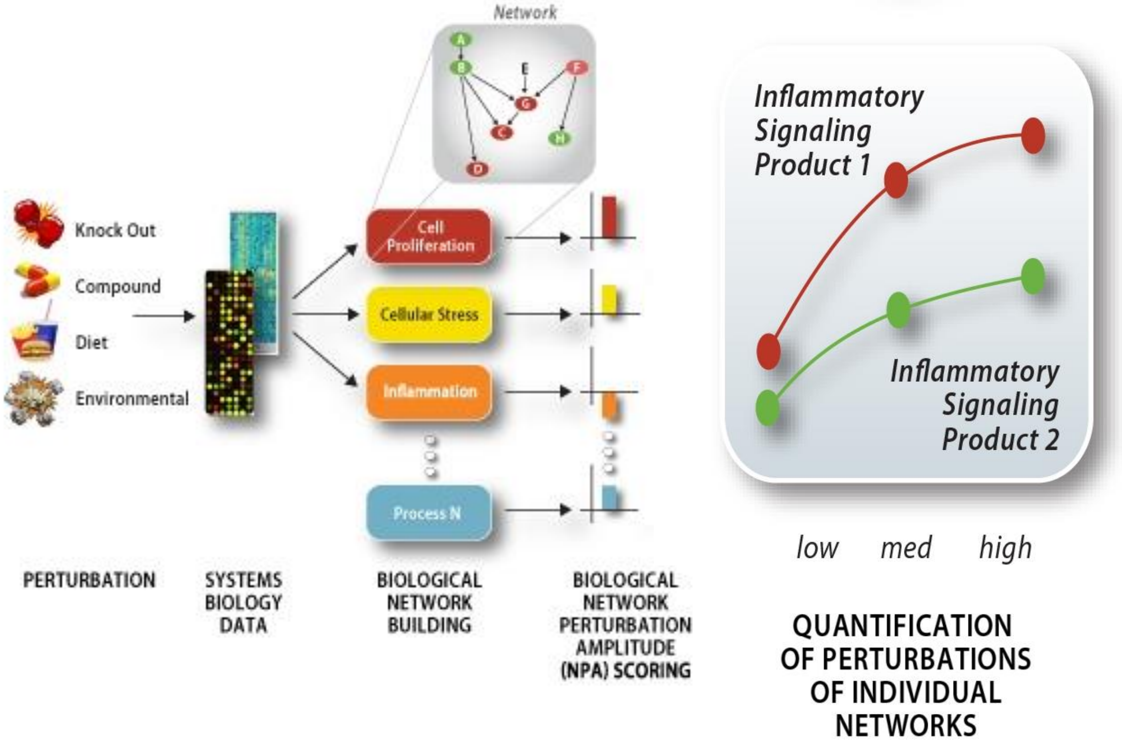
Identify Perturbed Biological Networks

Compute Network Perturbation Amplitudes

- Compute Amplitudes of Perturbation for all identified Biological Networks

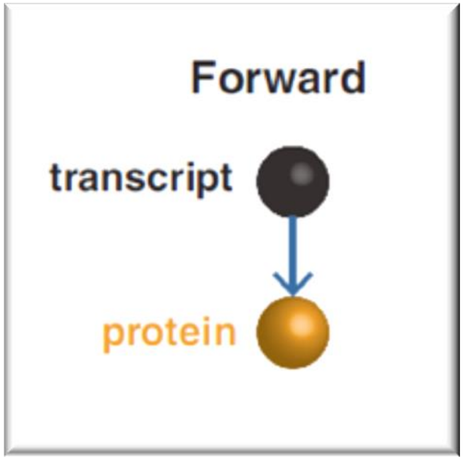
- Compare the Network Perturbation Amplitudes across responses between different perturbations

- Identify potential biomarkers indicative of overall Network Perturbation State

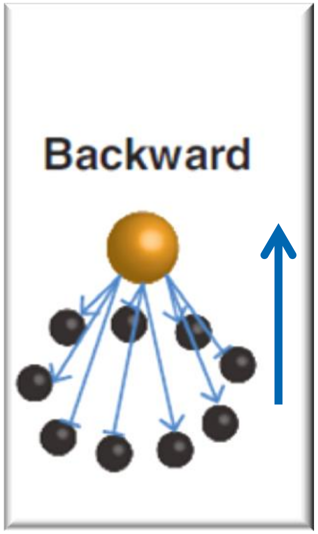


Martin F, Thomson TM, Sewer A, Drubin DA, Mathis C, Weisensee D, Pratt D, Hoeng J, and Peitsch MC.
Assessment of network perturbation amplitude by applying high-throughput data to causal biological networks. *BMC Syst Biol* 6: 54, 2012

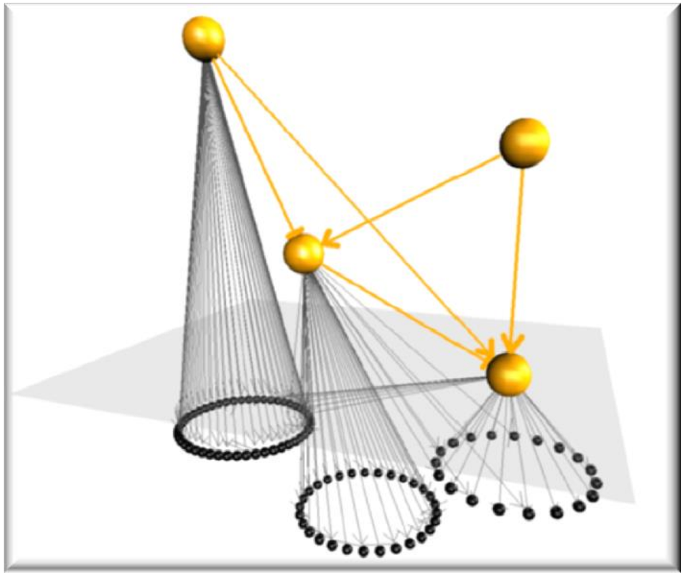
How Do We Use Transcriptomic Data?



Identification of potential downstream consequences of the data observations



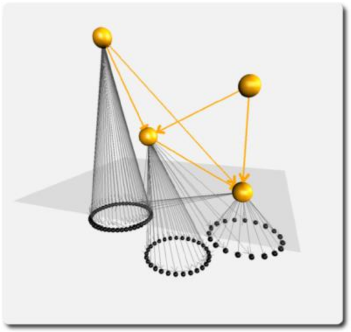
Identification of upstream entities that are potential causes of the observed differential elements



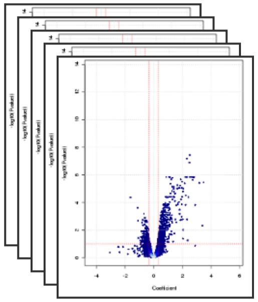
Martin F, Thomson TM, Sewer A, Drubin DA, Mathis C, Weisensee D, Pratt D, Hoeng J, and Peitsch MC.
Assessment of network perturbation amplitude by applying high-throughput data to causal biological networks. *BMC Syst Biol* 6: 54, 2012

NPA for Biological Impact Assessment and Mechanistic Interpretation

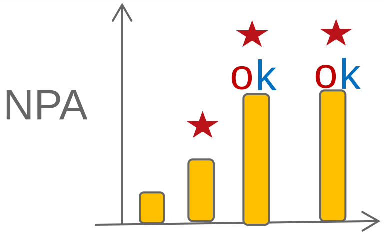
Network X



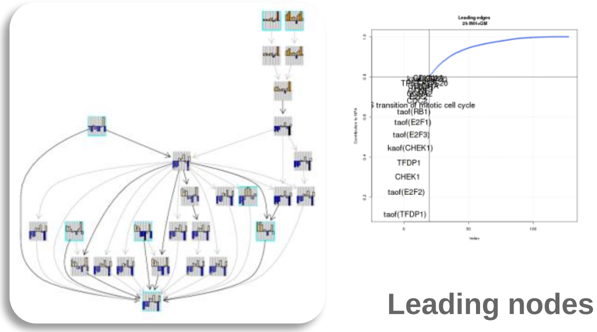
Differential data



Network Perturbation Amplitude



Network Perturbation Shape



Leading nodes

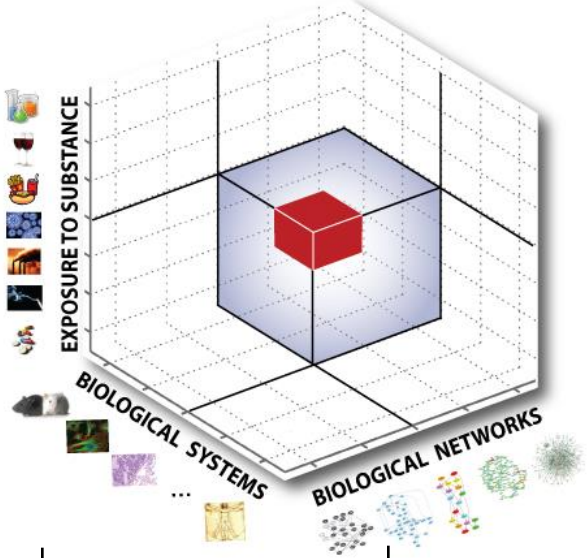
Related publications:

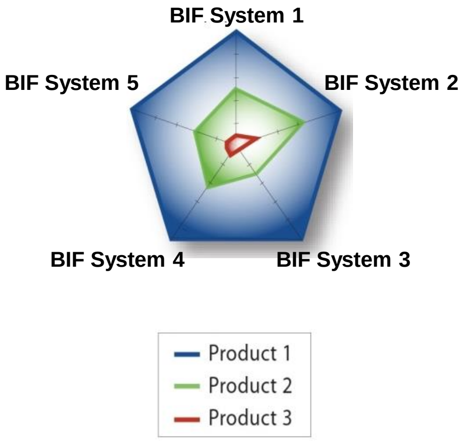
- *Assessment of Network Perturbation Amplitude by Applying High-Throughput Data to Causal Biological Networks,* Martin et al., BMC Syst Biol. 2012 May 31;6:54.
- *Quantification of biological network perturbations: Impact assessment and diagnostic using causal biological networks,* Martin et al., *In preparation.*
- *A network-based approach to quantifying the impact of biologically active substances*, Hoeng et al., Drug Discov Today. 2012 May;17(9-10):413-8.

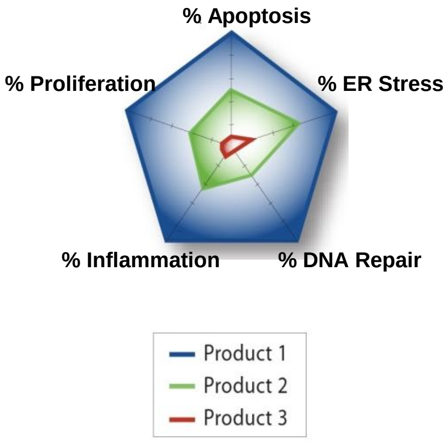
Quantitative Mechanism-Based Systems Impact Assessment

- Experimental Data Production
- Compute Systems Response Profiles
- Identify Perturbed Biological Networks
- Compute Network Perturbation Amplitudes
- Compute Product Biological Impact Factor

Product Biological Impact Factor is a numerical indicator of the impact of a product (a set of substances) on the system

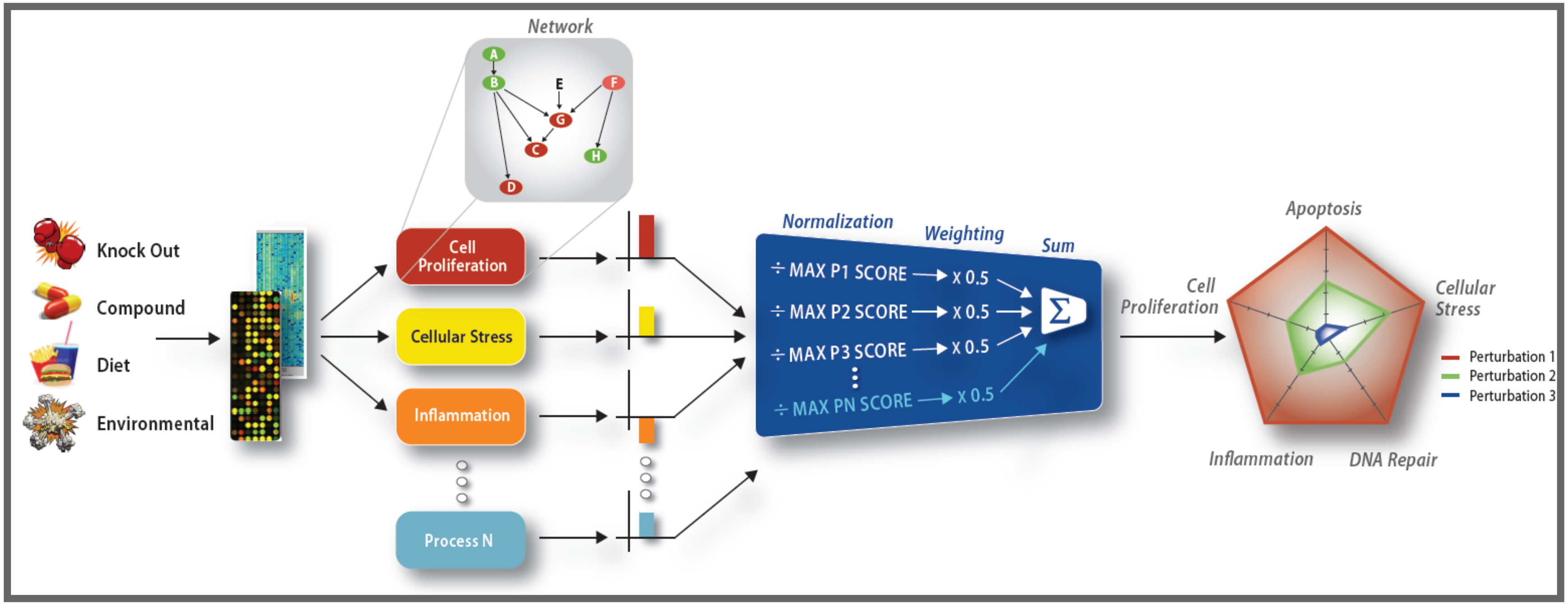






	System 1				System 2			
Exp / BN	Prod 1	Prod 2	Prod 3	Ctrl	Prod 1	Prod 2	Prod 3	Ctrl
Apoptosis	10	3	2	0	5	4	3	1
Oxidative Stress	5	4	4	1	4	2	2	0
Proliferation								
………								
………								

Summary of the Computational Process



Perturbation

Systems Biology Data

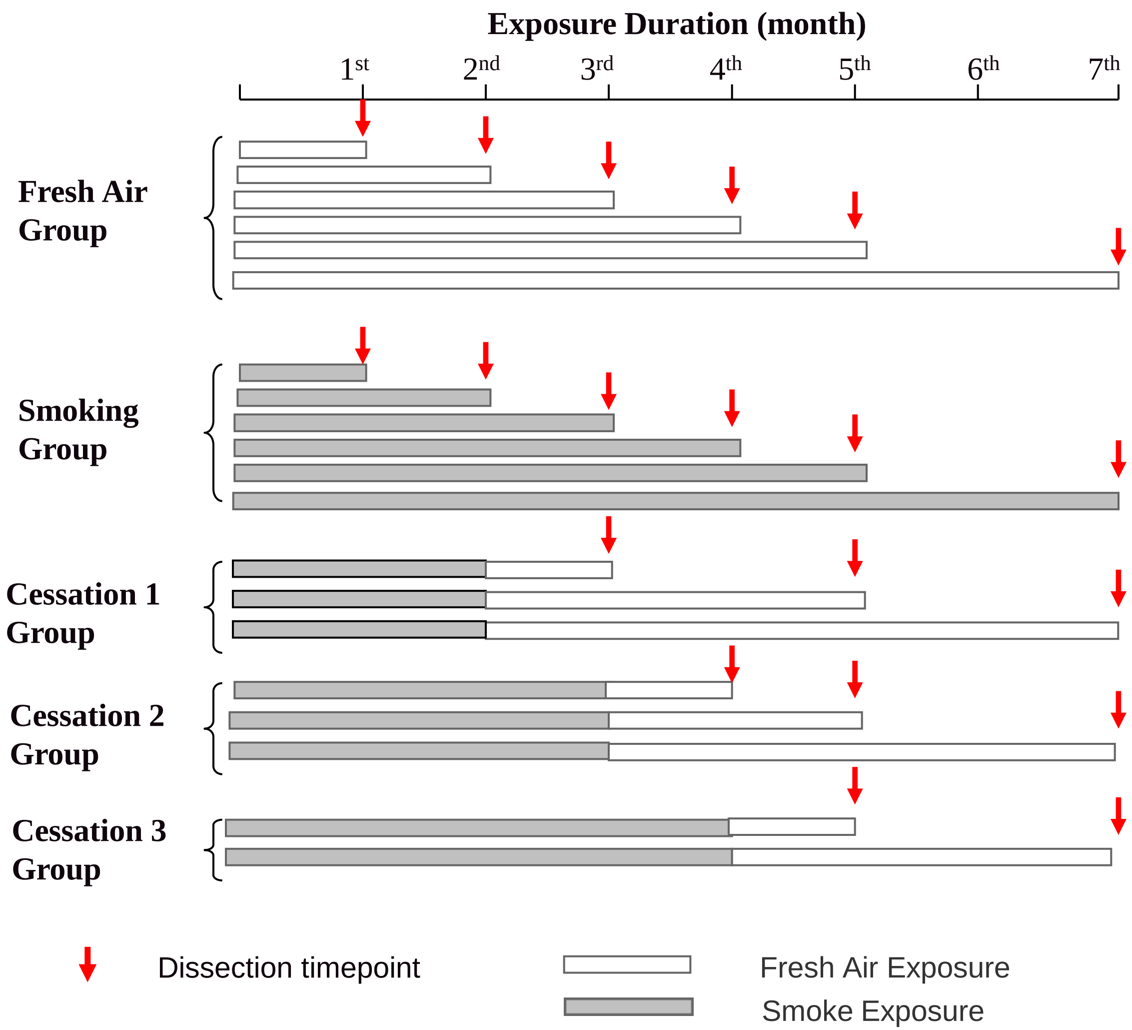
NPA Scores

Biological Impact Factor (BIF)

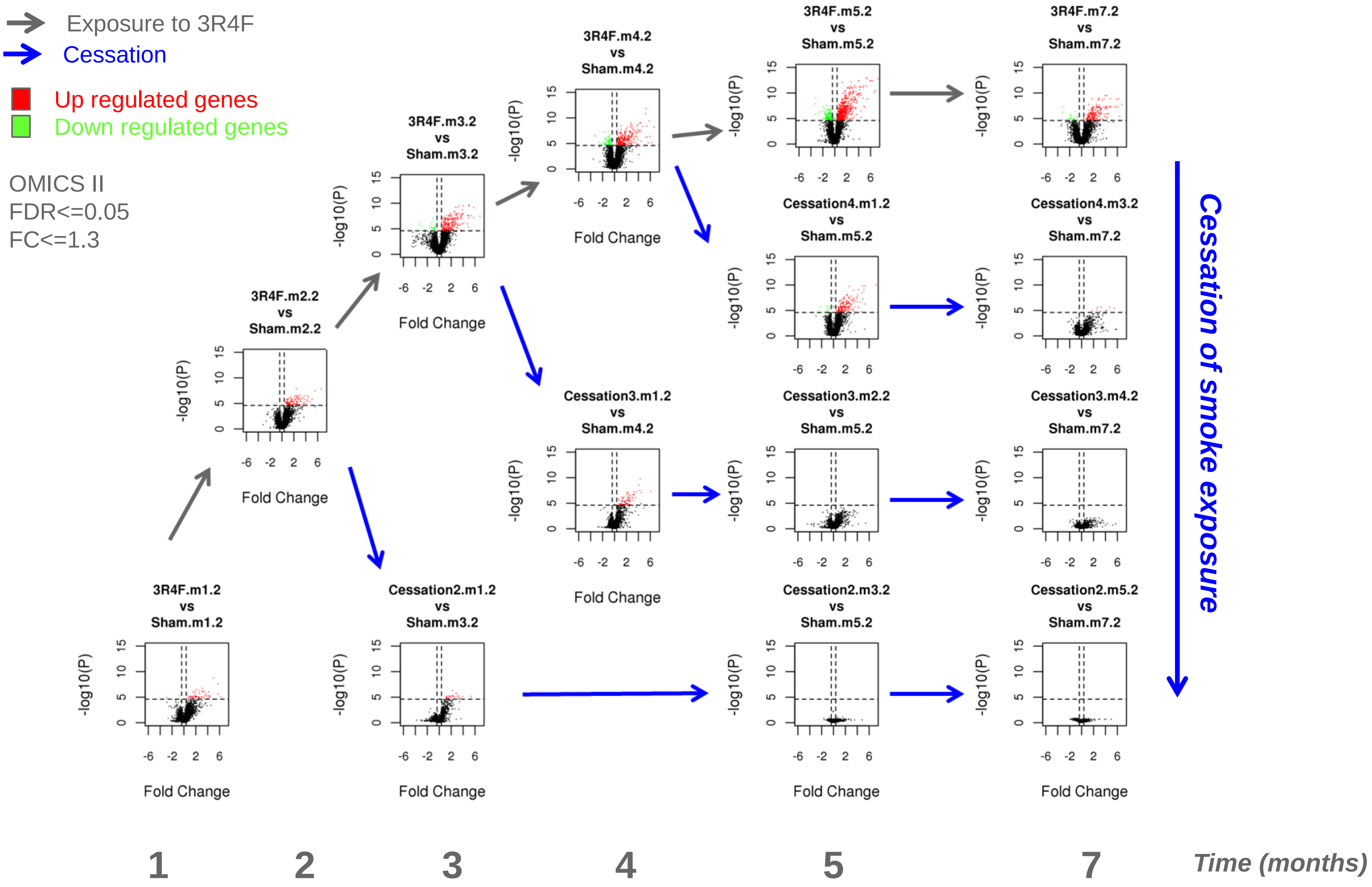
Biological Network Model

Score Aggregation

Study Design - Exposure Groups and Duration



Differential Gene Expression Regulation in Lung



→ Exposure to 3R4F

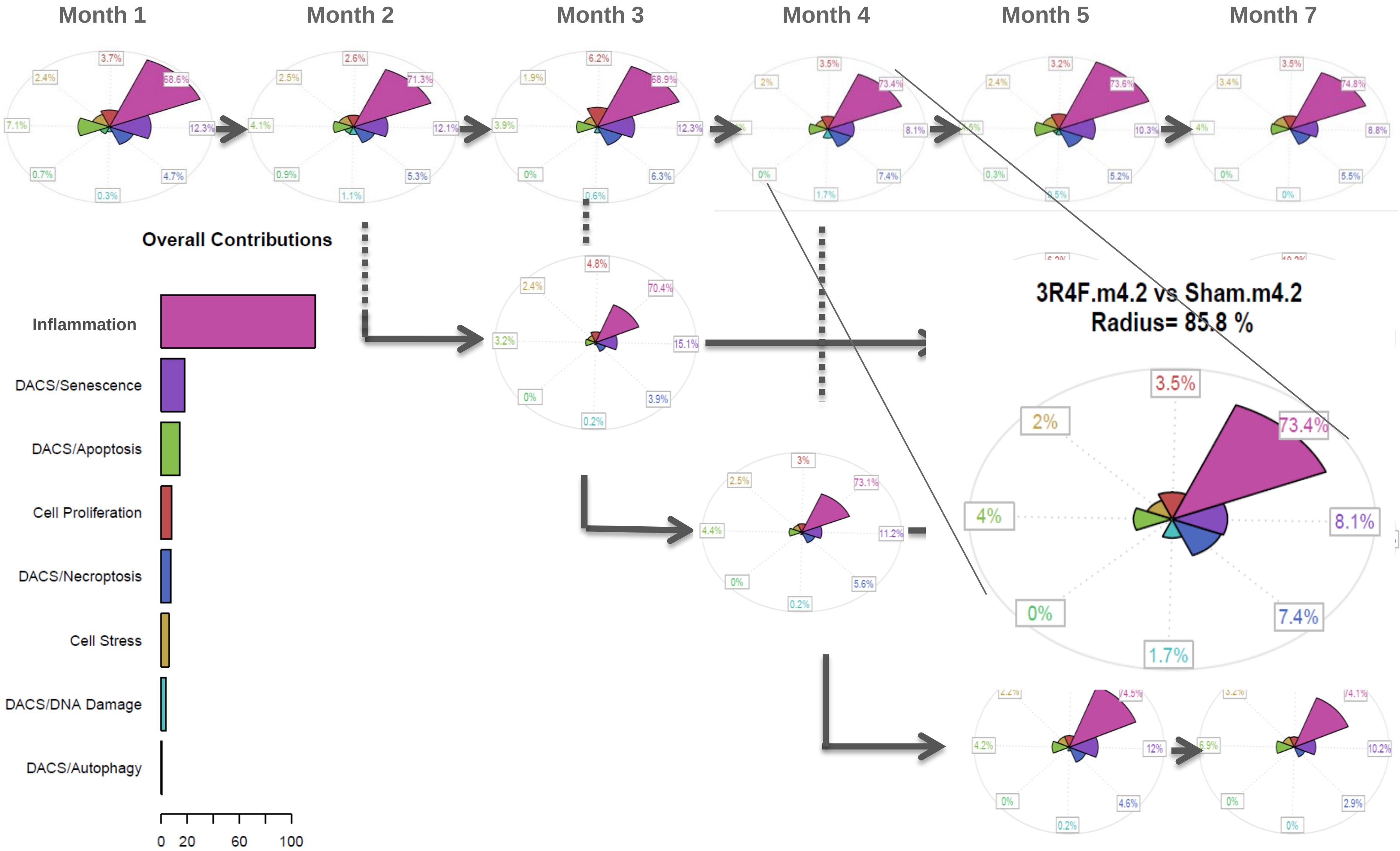
→ Cessation

🟥 Up regulated genes

🟩 Down regulated genes

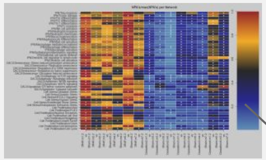
OMICS II
FDR<=0.05
FC<=1.3

Biological Impact Factors (BIF's)

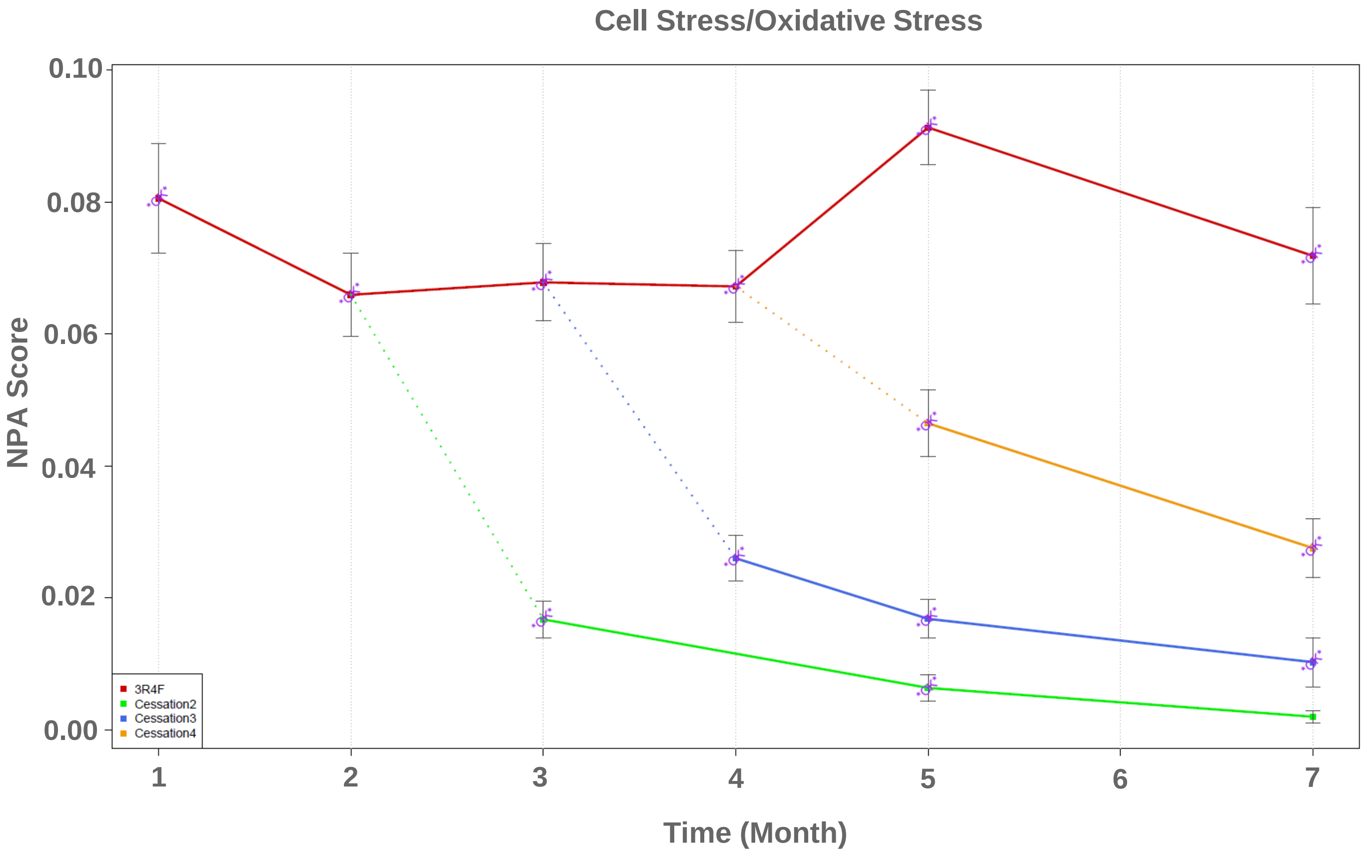


PMI RESEARCH & DEVELOPMENT

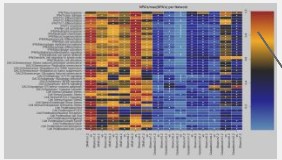
Evolution of NPA in 3R4F and Cessation Exposure Type Regimens over Time - *Cell Oxidative Stress Network*



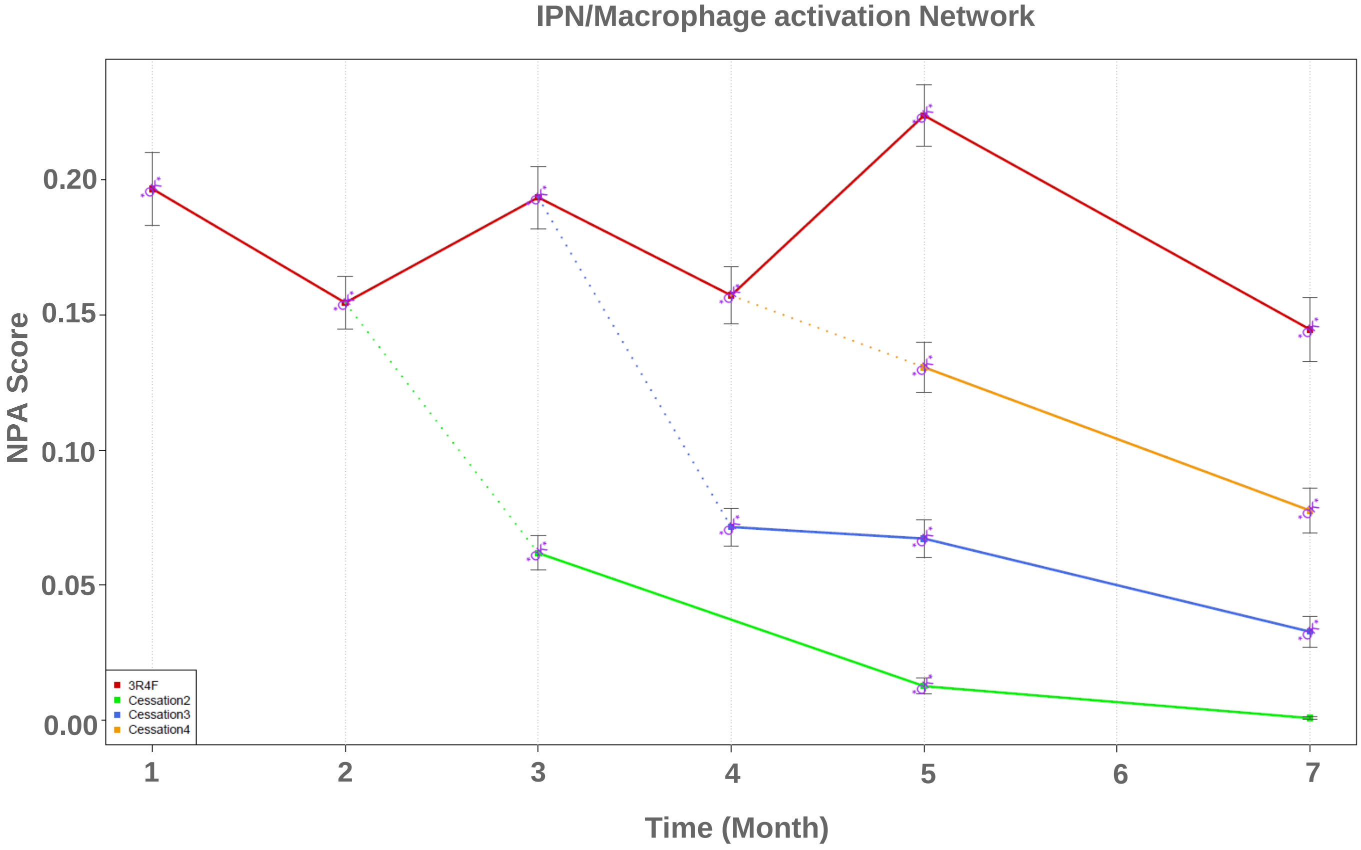
Quantification of the perturbation of 70 cause-and- effect mechanistic networks



Evolution of NPA in 3R4F and Cessation Exposure Type Regimens over Time - *IPN/Macrophage activation Network*



Quantification of the perturbation of 70 cause-and- effect mechanistic networks

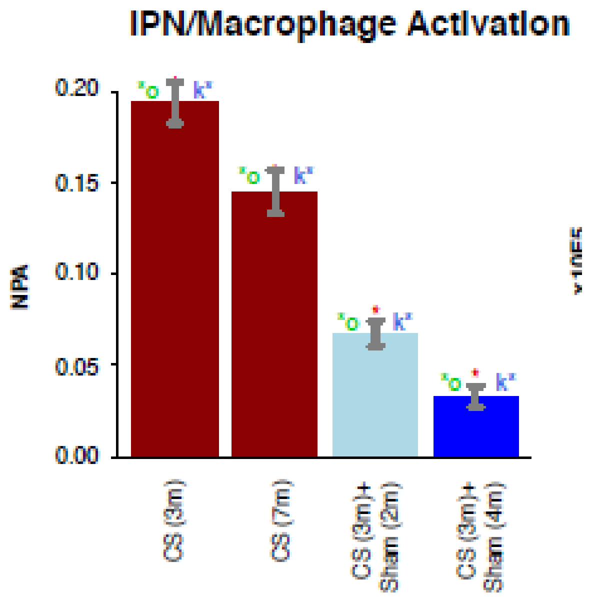


PMI RESEARCH & DEVELOPMENT

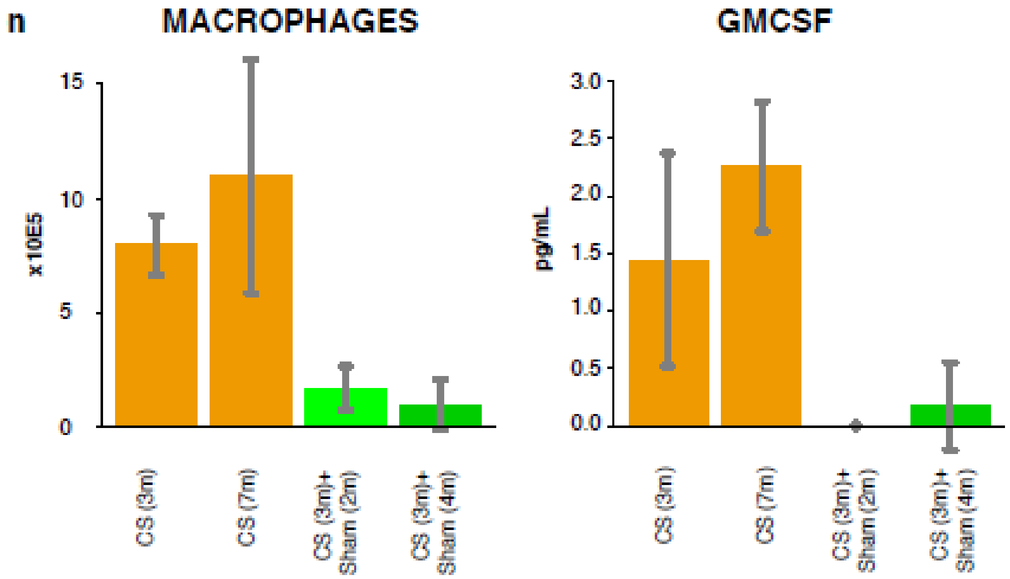
E

Consistent Trends between Network Models (primary tissue) & BALF (surrogate measurements)

Lung NPA Score



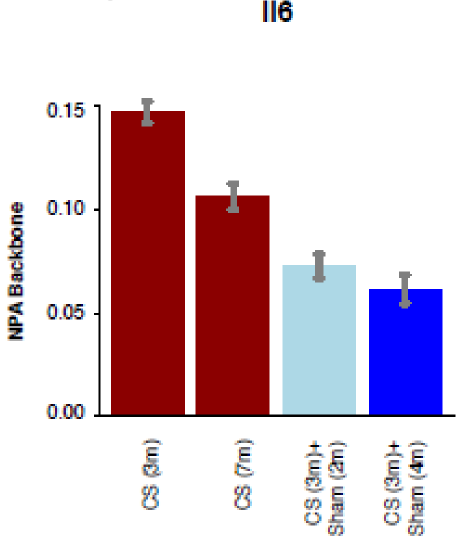
BALF Measurements

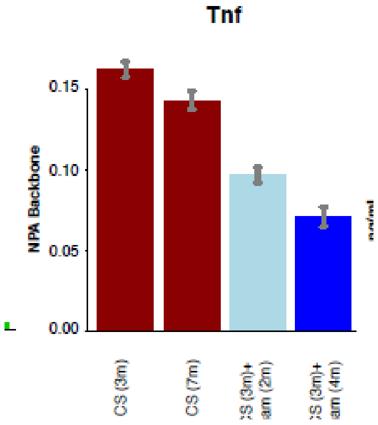


Julia Hoeng, Marja Talikka, Florian Martin; Alain Sewer, Xiang Yang, Anita Iskandar, Walter Schlage , Manuel C. Peitsch
Case Study: The Role of Mechanistic Network Models in Systems Toxicology (manuscript submitted).

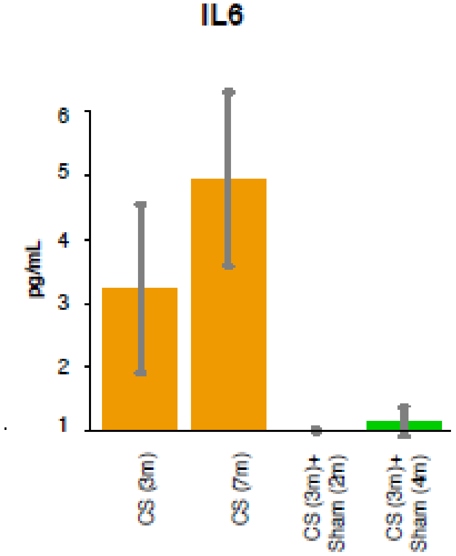
NPA Scores Indicate Residual Signaling in Primary Tissue (Lung)

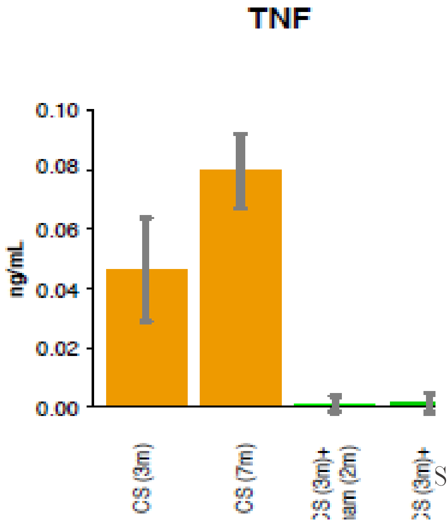
Lung NPA Score



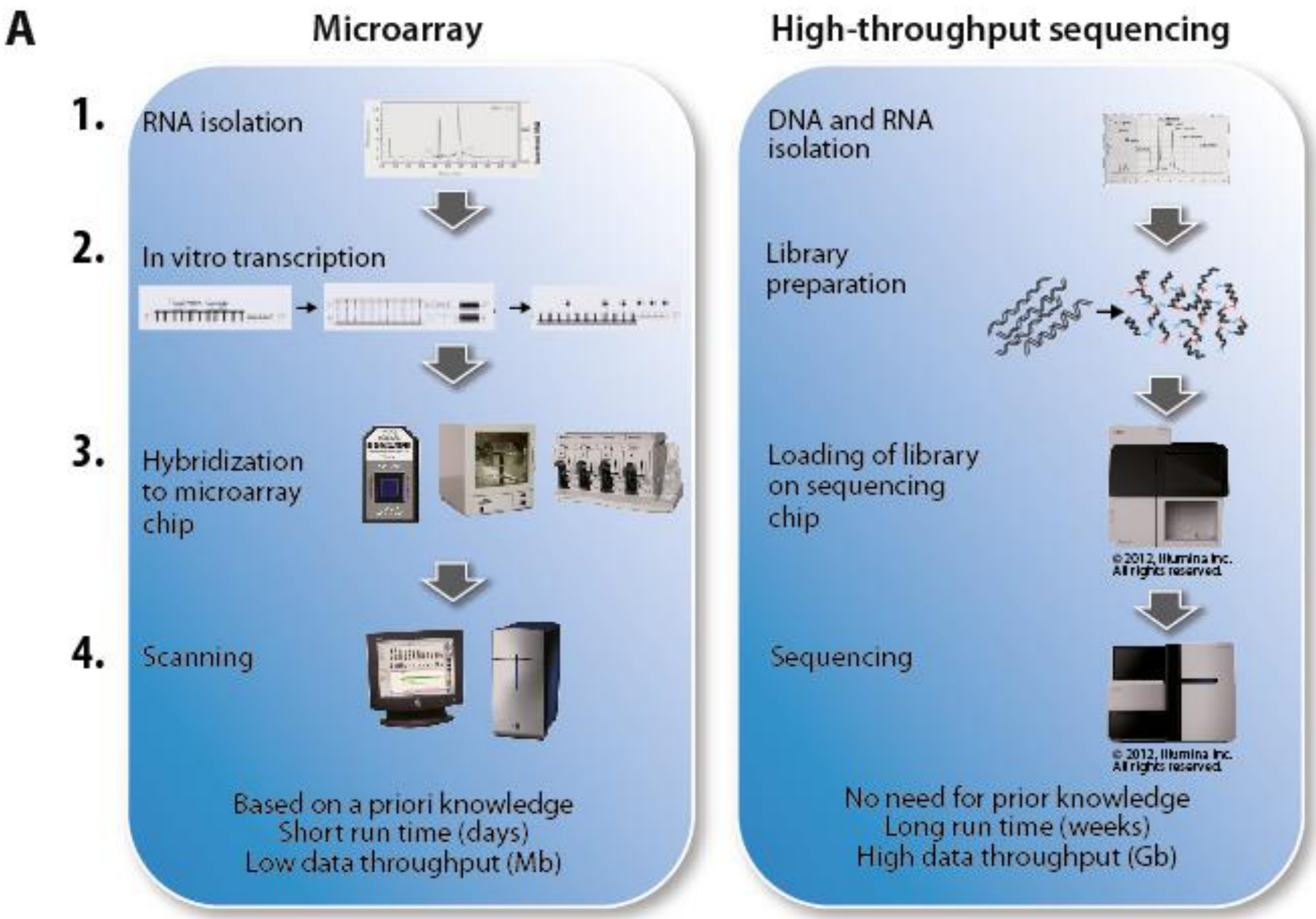


BALF Measurements





New Technologies

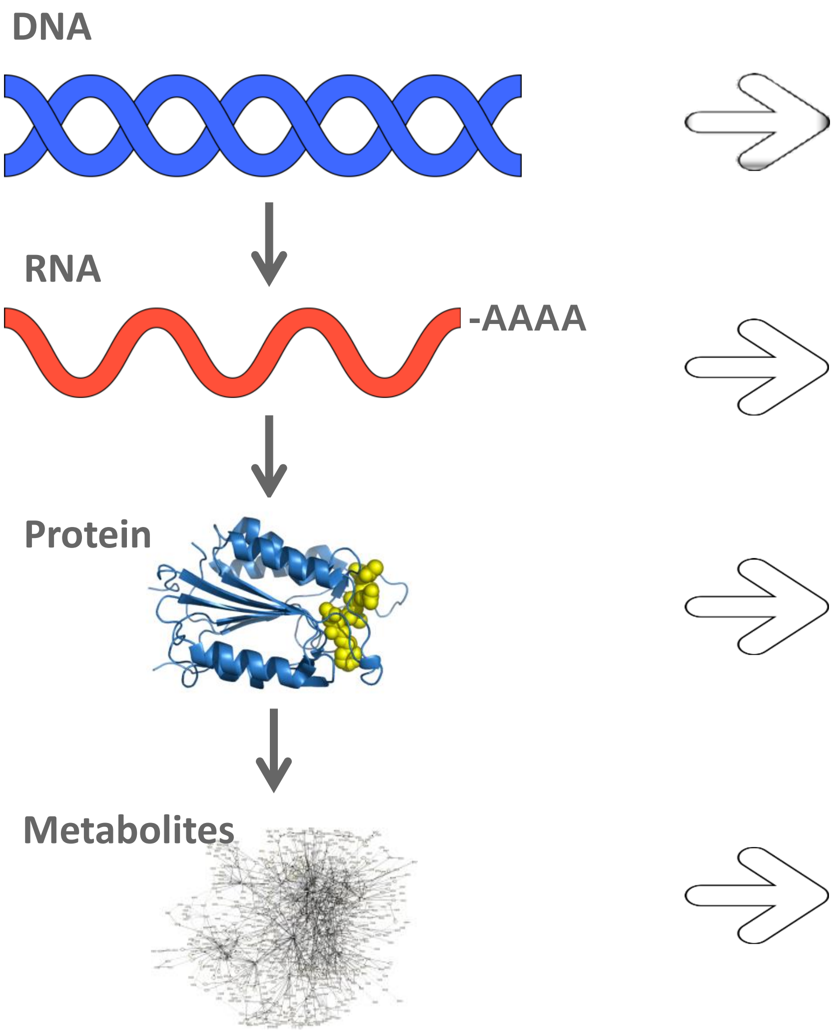


The Affymetrix Gene Expression and Illumina Sequencing Platforms

	Affymetrix Chip	Illumina RNA-seq
Technology	Gene-centric (limited to genes present on the chip)	Sequence-centric (genes, transcript isoforms,…)
Measured datapoints	~30’000	Unlimited
Time to generate data	~3 days	~14 days
Samples per run	96	Typical: 4 samples*16 lanes = 64 Max: 12 samples*16 lanes = 192
Analysis methods	Standardized Well-established	In development
Data storage footprint	Low (Mb)	High (Gb)

PMI RESEARCH & DEVELOPMENT

Future Developments …



- Genomics
- Epigenomics

- Transcriptomics
- miRNA

- Proteomics
- Phosphoproteomics

- Metabolomics
- Lipidomics

Summary

- Obvious lung emphysema was observed by histopathological/morphometric analysis after 2 months of mainstream cigarette smoke exposure.
- Mainstream cigarette smoke exposure cessation reverses the effects of inflammation, lung function changes and histopathological/morphometric changes.
- A strong inflammation component was observed from the gene expression profile followed by senescence and DNA damage. The molecular changes are supported by the biological endpoints.
- PMI computational method based on mechanistic networks enables us to identify the key biological processes involved.
- New technologies (e.g. Next Generation Sequencing and Mass Spectroscopy Proteomics) would allow us to evaluate the effects of cigarette smoke exposure and cessation at the DNA and protein levels.

Acknowledgement

Selventa-USA

Jurjen W. Westra
Michael J. Maria
David de Graaf
Renée Deehan
Dexter Pratt
Ty M. Thomson,
David A. Drubin
Andrea L. Mathews

PMIRL-Singapore

Magali Russell
Shen Yi
Jan Verbeeck
Blaine Phillips
Didier Goedertier
Clement Foong
Teo Hutt Wang
Woon Ching Qing
Eric Yew
Chor Kiat Yeo

PMI R&D Switzerland

Manuel C. Peitsch
Julia Hoeng
Marja Talikka
Patrick Vanscheeuwijck
Nveed Chaudhary
Emilija Veljkovic
Michael Peck
Florian Martin
Yang Xiang
Grégory Vuillaume
Stefan Lebrun
Carine Poussin
Carole Mathis
Sandra Wagner
Ulrike Kogel
Walter Schlage
Vincenzo Belcastro
Alain Sewer
Sam Ansari

PMI RESEARCH & DEVELOPMENT

Thank You

PMI RESEARCH & DEVELOPMENT