Environmental
Engineering
Laboratory

Department of Chemical Engineering
School of Engineering
Aristotle University of Thessaloniki

Multi-scale connectivity
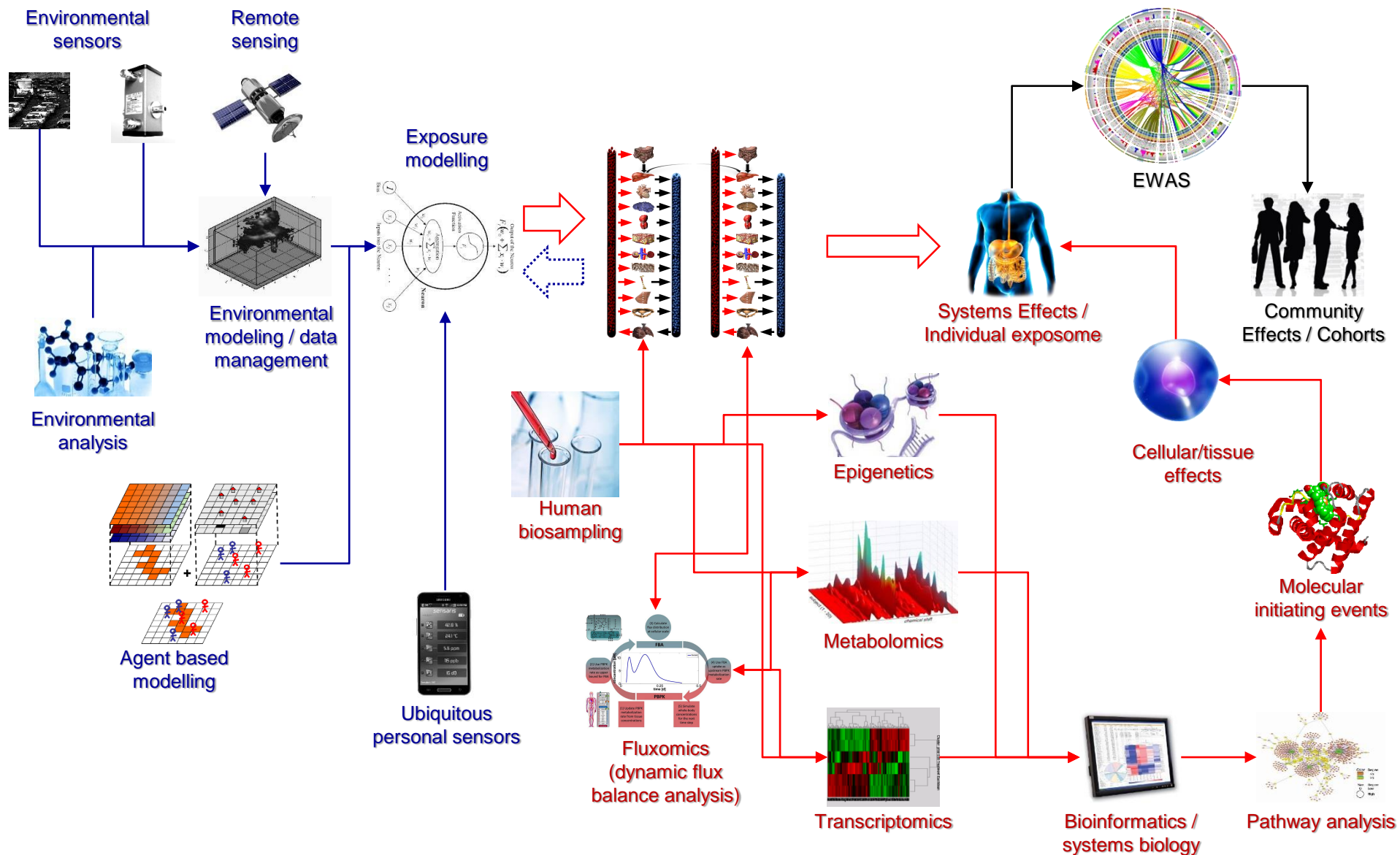# An integrated methodology to unravel the exposome

Denis A. Sarigiannis[1,2]

[1]Department of Chemical Engineering,

Aristotle University of Thessaloniki

GR-54124, Thessaloniki, Greece

[2]Centre for Research and Technology Hellas (CE.R.T.H.),
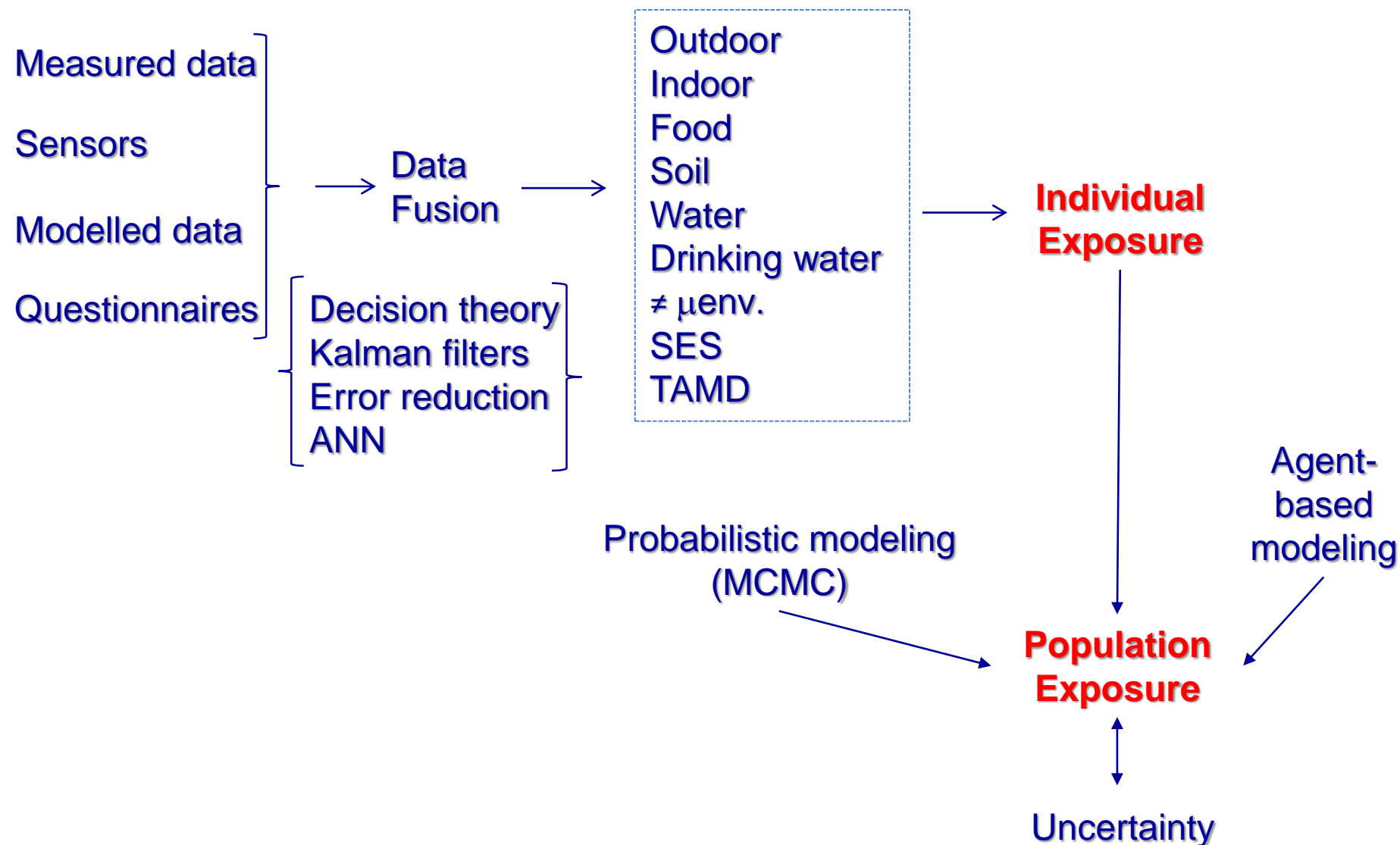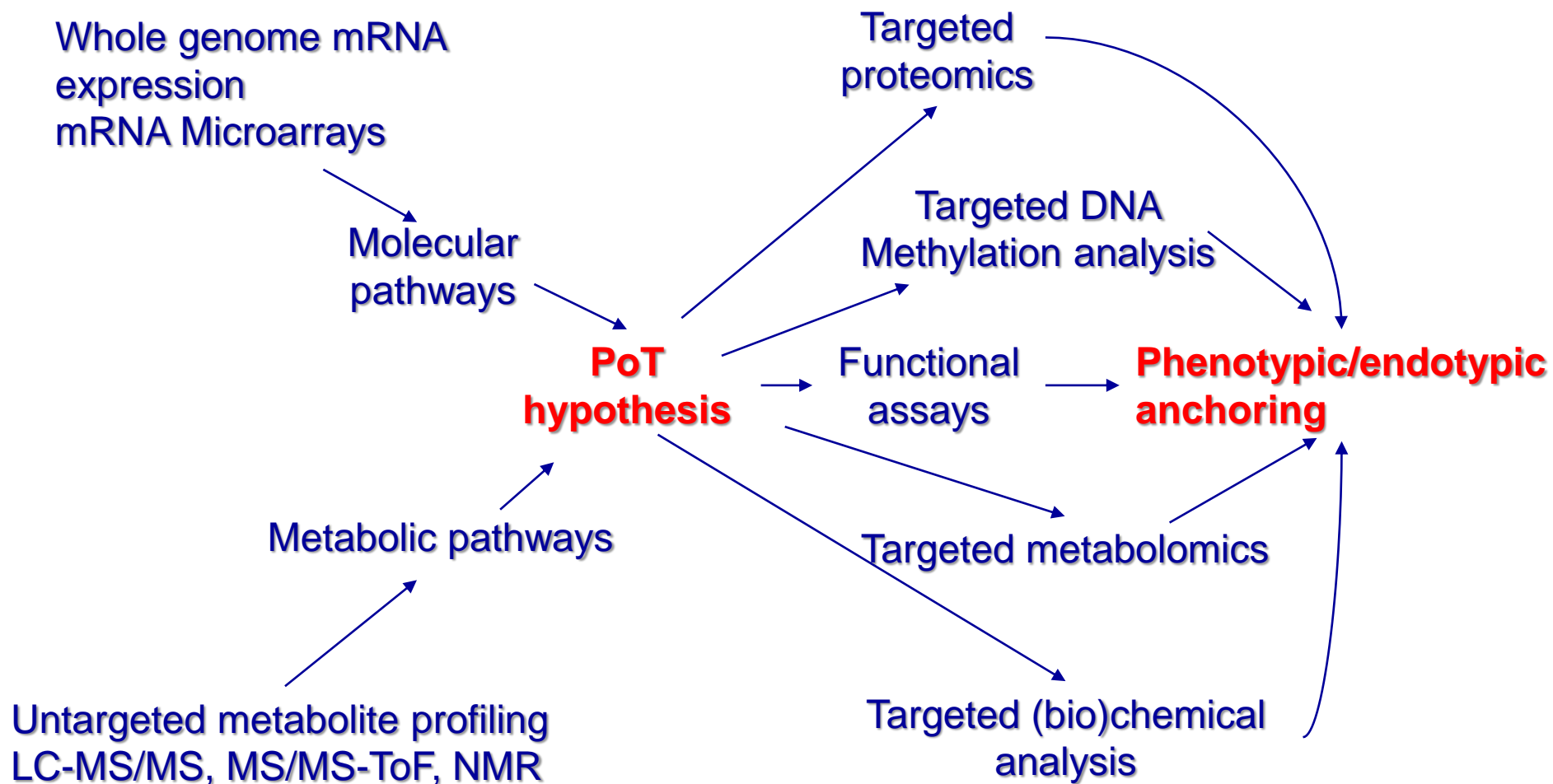
Thessaloniki, 57001, Greece

http://www.enve-lab.eu

The HEALS paradigm

Environmental sensors

Remote sensing

Exposure modelling

Environmental modeling / data management

Environmental analysis

Agent based modelling

Ubiquitous personal sensors

Human biosampling

EWAS

Systems Effects / Individual exposome

Community Effects / Cohorts

Epigenetics

Cellular/tissue effects

Metabolomics

Molecular initiating events

Fluxomics (dynamic flux balance analysis)

Transcriptomics

Bioinformatics / systems biology

Pathway analysis

Environmental
Engineering
Laboratory

Department of Chemical Engineering
School of Engineering
Aristotle University of Thessaloniki

External Exposure workflow

Measured data

Sensors

Modelled data

Questionnaires

→ Data Fusion →

Decision theory
Kalman filters
Error reduction
ANN

Outdoor
Indoor
Food
Soil
Water
Drinking water
$\neq \mu$env.
SES
TAMD

→ **Individual Exposure**

Probabilistic modeling
(MCMC)

Agent-based modeling

Population Exposure

Uncertainty

Environmental
Engineering
Laboratory

Department of Chemical Engineering
School of Engineering
Aristotle University of Thessaloniki

Exposure biology workflow
Rendering high dimension biology operational

Whole genome mRNA expression mRNA Microarrays

Molecular pathways

Metabolic pathways

Untargeted metabolite profiling LC-MS/MS, MS/MS-ToF, NMR

PoT hypothesis

Targeted proteomics

Targeted DNA Methylation analysis

Functional assays

Targeted metabolomics

Targeted (bio)chemical analysis

Phenotypic/endotypic anchoring

Environment Wide Association Studies (EWAS)
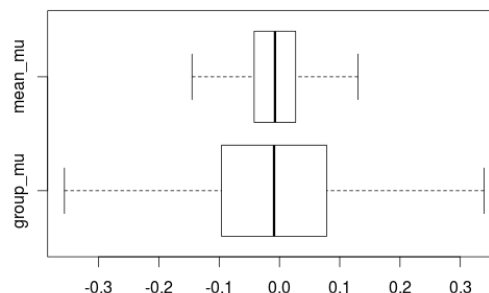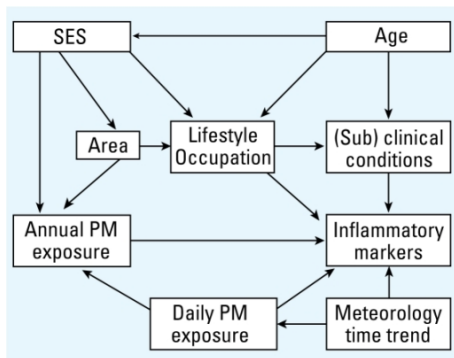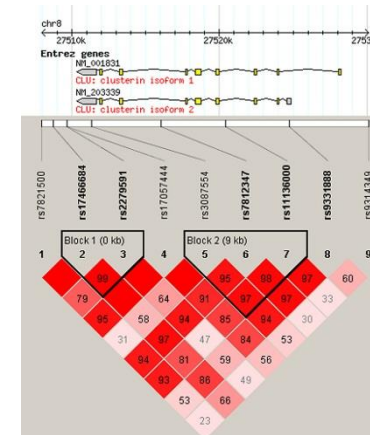
Directed Acyclic graphs (DAGs)
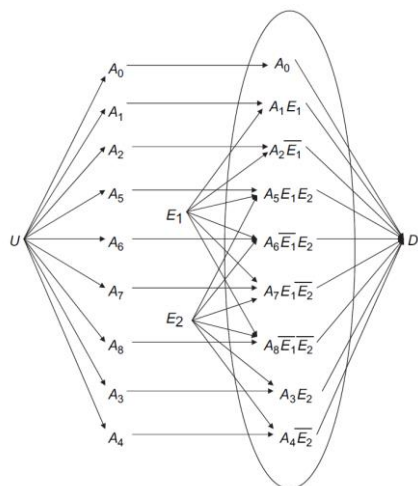
Bayesian Inference
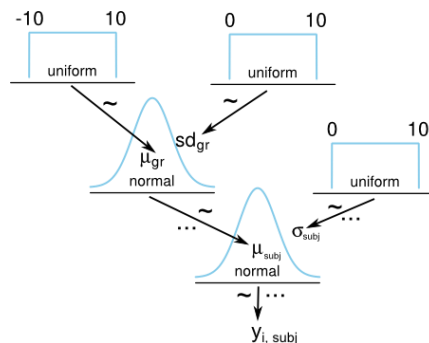
Linkage disequilibrium
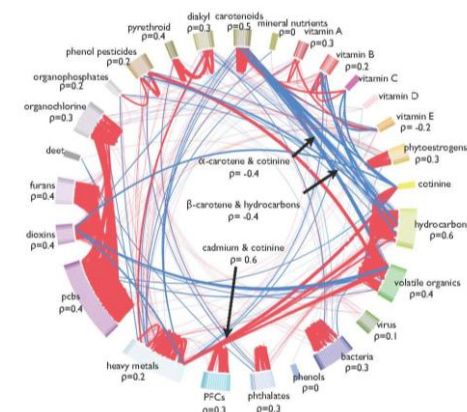
Patel et al. (2011). Canonical Correlation of Biomarkers of Environmental

Environmental
Engineering
Laboratory

Department of Chemical Engineering
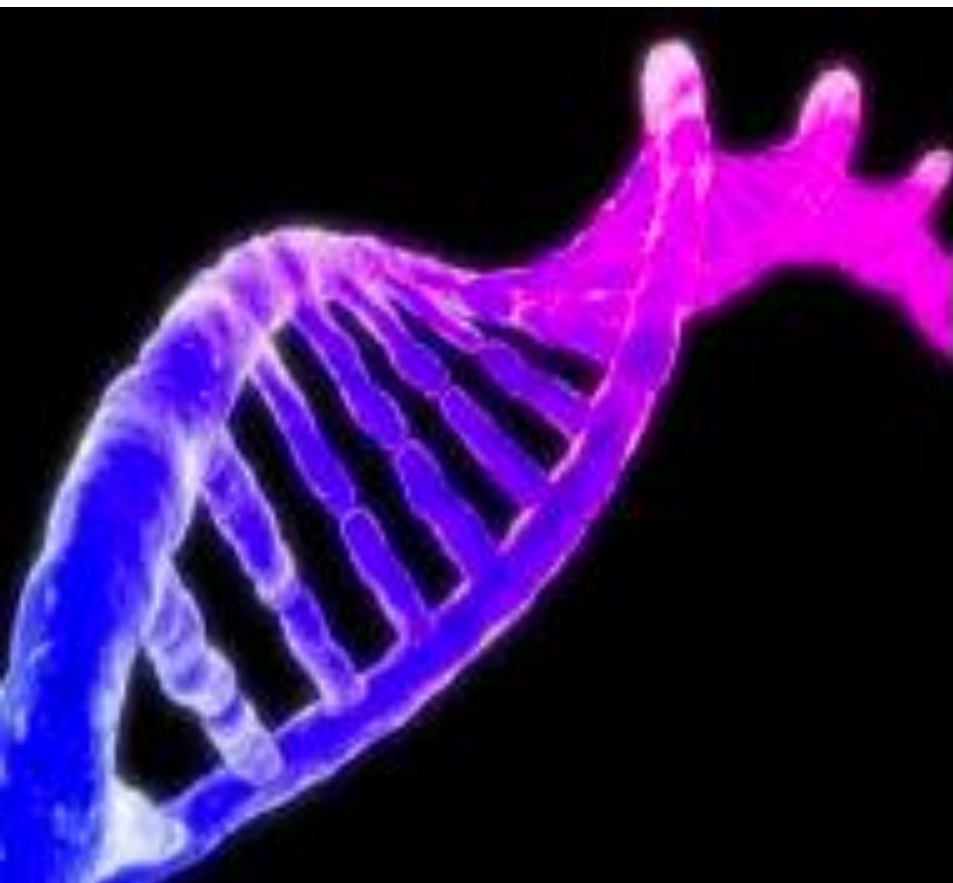School of Engineering
Aristotle University of Thessaloniki

Population studies:
Health outcomes

- Already existing cohort data
  - Allergy and asthma - link with particulate matter (PM) and biologicals
  - Neurodevelopmental and neurodegenerative disorders - link with metals/metalloids and pesticides (incl. endocrine disruptors)
  - Obesity and childhood diabetes - link with endocrine disruptors

- Pilot European Exposure and Health Examination Survey (EXHES)
  - Mother-Father-Children longitudinal and nested case-control study
  - Two phases foreseen:
    - Phase 1:
      - 200 twins, 200 matched singletons, 500 unselected singletons, parents
      - Recruitment within 18 months – follow up for 3 years
      - 10 countries
    - Phase 2:
      - Nested case-control
      - 140 twins, 70 singletons, 140 mothers and fathers
  ✴ Ethnics minorities included and targeted to the extent possible

Environmental
Engineering
Laboratory

Department of Chemical Engineering
School of Engineering
Aristotle University of Thessaloniki

The internal exposome

- **Transcriptomics**
  - Human data
  - In vitro data for mechanistic hypotheses anchoring
- **Metabolomics**
  - Coupled use of NMR, UPLC-MS, MS-MS TOF, GC-MS
  - Key xenobiotics: metals, dioxins, phthalates, PCBs, PAHs
- **Adductomics** of oxidation- and alkylation-induced damage
  - Adducts of electrophiles with DNA, Hb, Alb, Glutathione
  - Oxidative and nitrogen species stress markers (antioxidant capacity, isoprostanes, uric acid, cytokines) in blood
  - UPLC-MS/MS, fixed-step selected reaction monitoring (FS-SRM) for protein adducts
- **SNP profiling** (UK Axiom Biobank array with targeted add-ons) and functional analysis of repair proteins (repair of oxidative DNA damage)
- Genome-wide DNA **methylation profiling**
  - Bisulfite sequencing
  - Characterisation of methylation of selected CpG target sites
  - Aim: to identify differences between epigenetically influenced and independent SNPs

Concept of generic lifetime PBBK model
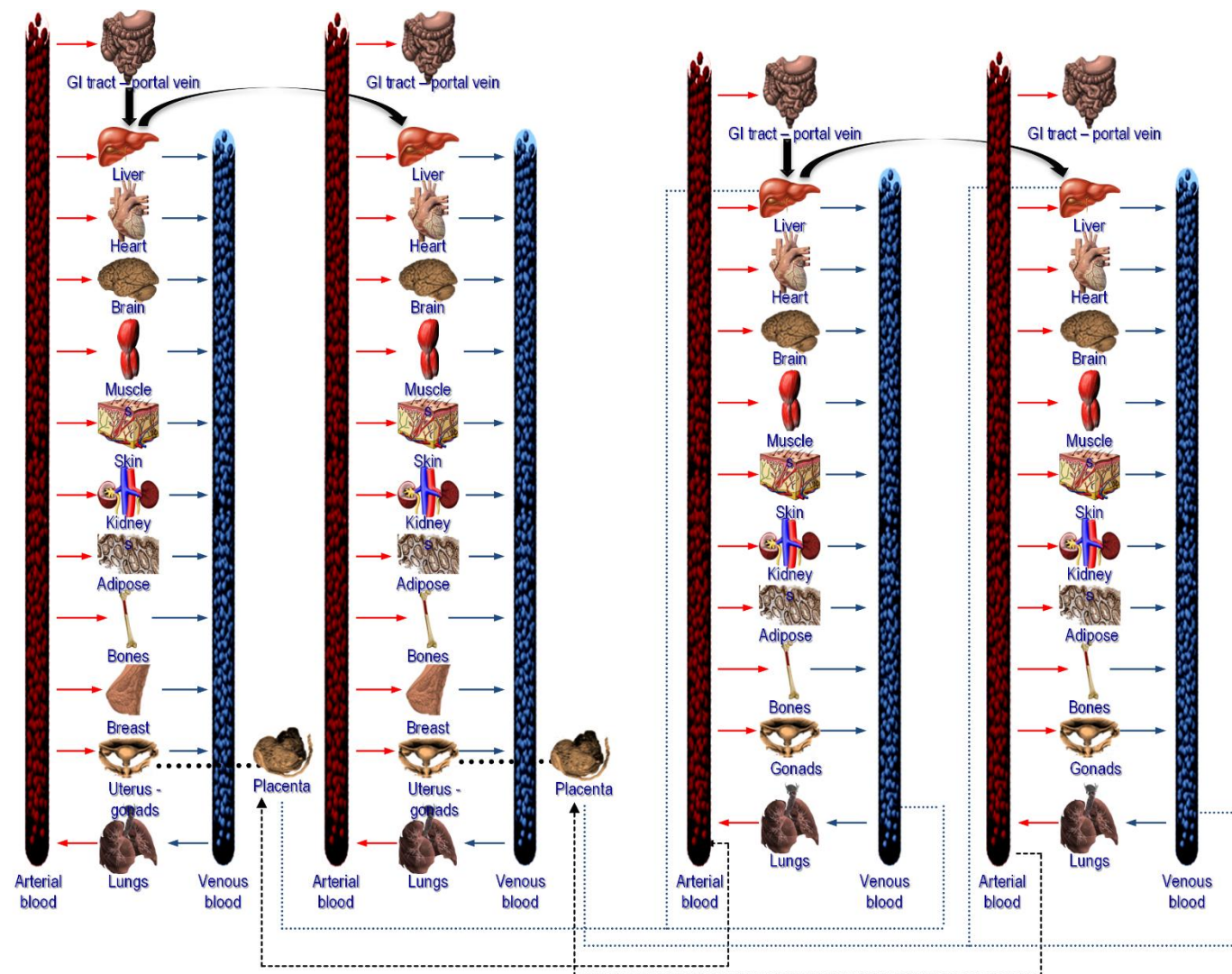
- Detailed description of compartments and tissue composition
- Lifetime evolving parameters
  - Organ volumes
  - Blood flows
  - Age-dependent clearance
- Mother – Fetus interaction
- Breast feeding

Environmental
Engineering
Laboratory

Department of Chemical Engineering
School of Engineering
Aristotle University of Thessaloniki

Expanding the chemical space – use of QSARs

According to Abraham's solvation equation, a biological property *SP is described by the following equation*

$$\log SP = c + r \cdot R_2 + s \cdot \pi_2^{\mathrm{H}} + a \cdot \Sigma\alpha_2^{\mathrm{H}} + b \cdot \Sigma\beta_2^{\mathrm{H}} + v \cdot \log V_x$$
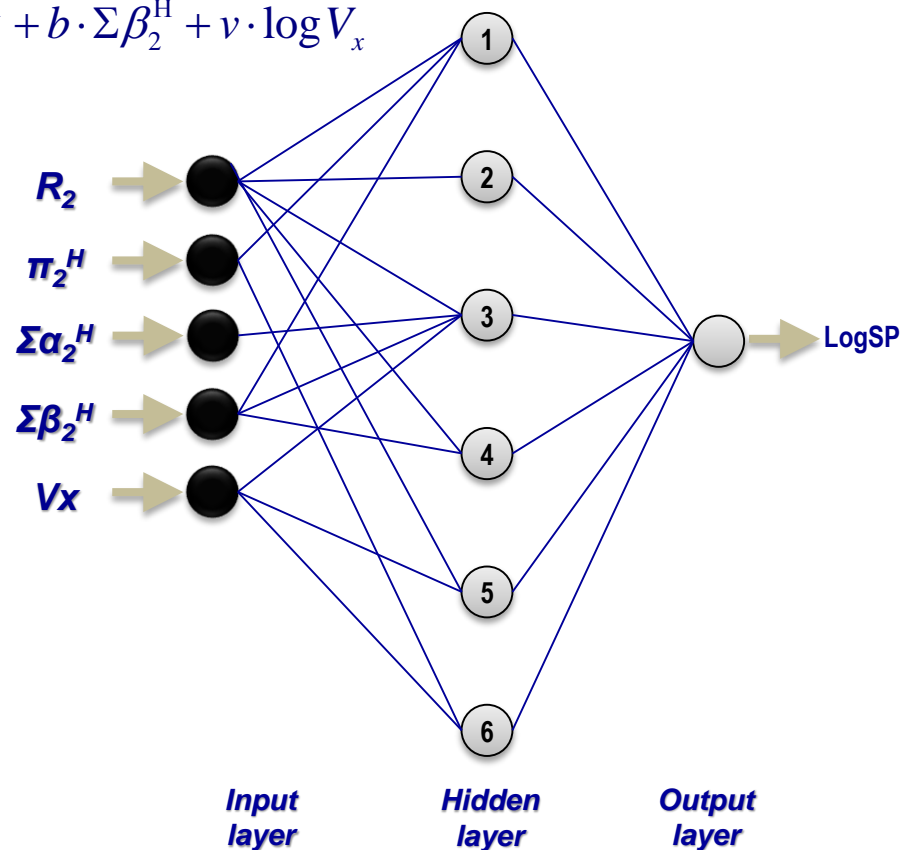
Where:

R_2 is an excess molar refraction that can be determined simply from a knowledge of the compound refractive index

π_2^H is the compound dipolarity/polarizability

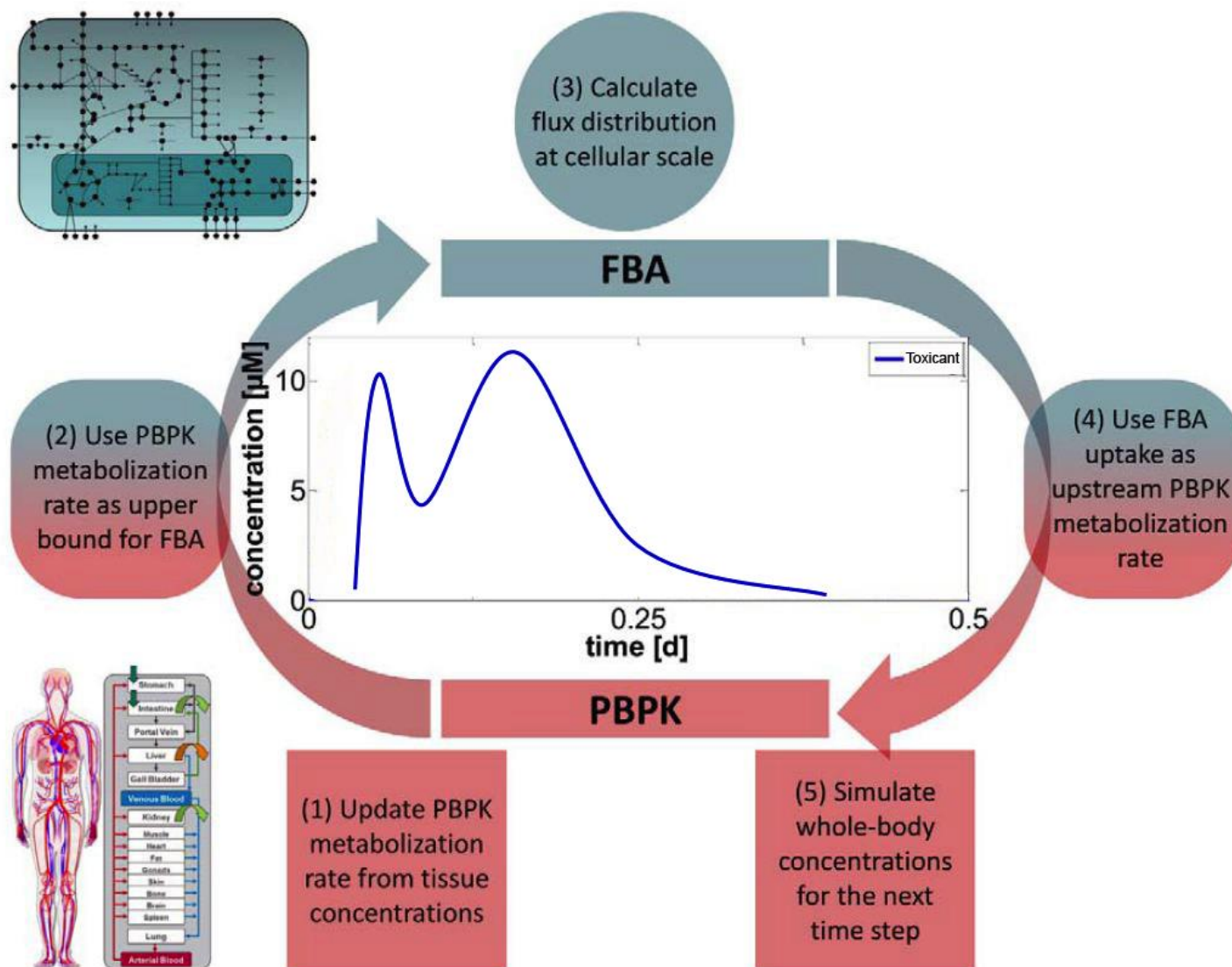$\Sigma\alpha_2^H$ is the solute effective or summation hydrogen-bond acidity

$\Sigma\beta_2^H$ is the solute effective or summation hydrogen-bond basicity
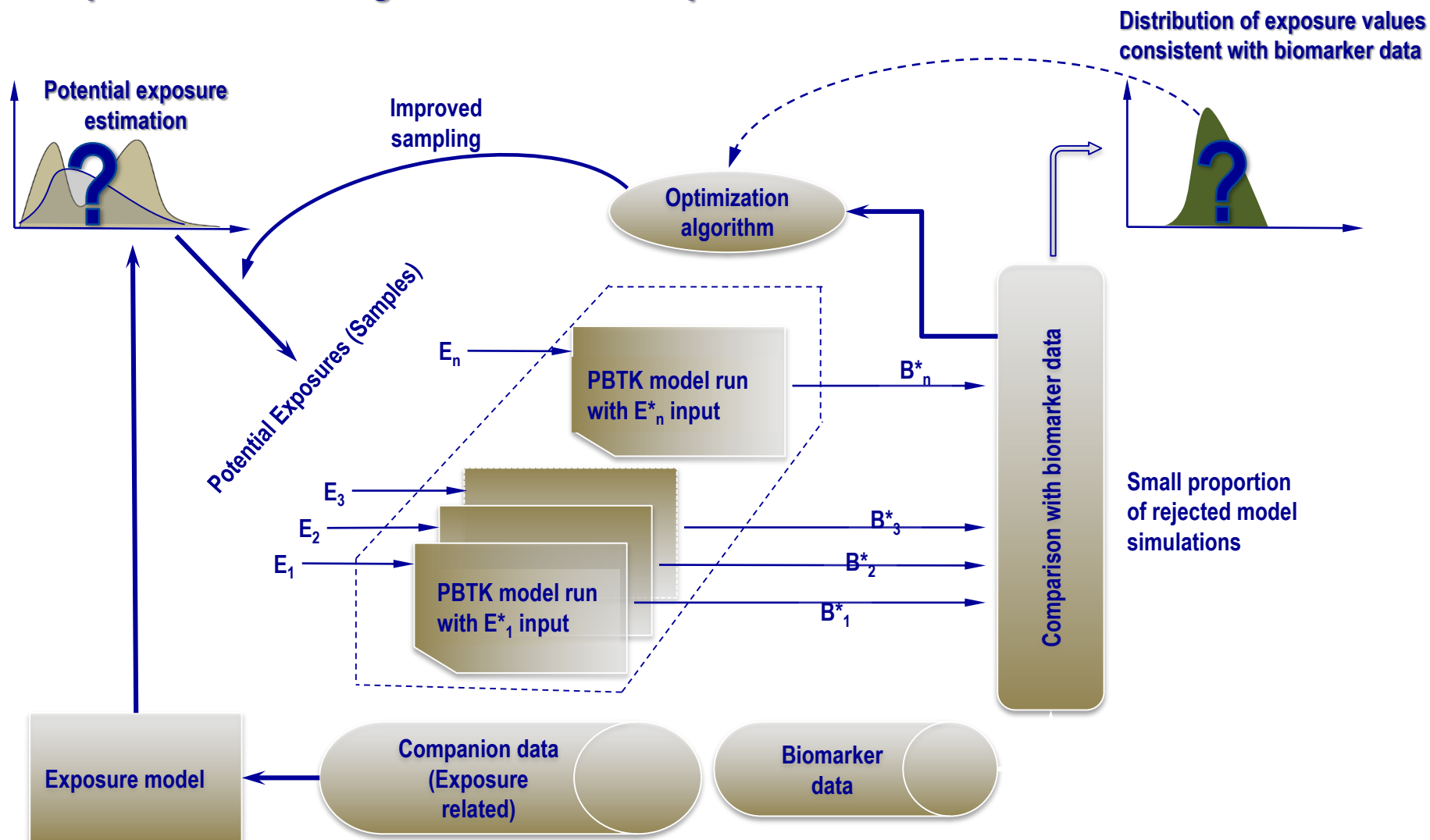
V_x is the McGowan characteristic volume

R_2
π_2^H
$\Sigma\alpha_2^H$
$\Sigma\beta_2^H$
Vx

LogSP

Input layer   **Hidden layer**   **Output layer**

Coupling biokinetics and metabolic regulation

(3) Calculate flux distribution at cellular scale

FBA

(2) Use PBPK metabolization rate as upper bound for FBA

(4) Use FBA uptake as upstream PBPK metabolization rate

PBPK

(1) Update PBPK metabolization rate from tissue concentrations

(5) Simulate whole-body concentrations for the next time step

Environmental
Engineering
Laboratory

Department of Chemical Engineering
School of Engineering
Aristotle University of Thessaloniki

HBM data assimilation

Optimal methodological scheme for exposure reconstruction

Distribution of exposure values consistent with biomarker data

Potential exposure estimation

Improved sampling

Optimization algorithm

Potential Exposures (Samples)

E_n → **PBTK model run with E^*_n input** → B^*_n

E_3 →

E_2 →

E_1 → **PBTK model run with E^*_1 input** → B^*_3 → B^*_2 → B^*_1

Comparison with biomarker data

Small proportion of rejected model simulations

Exposure model ← **Companion data (Exposure related)** **Biomarker data**

Environmental
Engineering
Laboratory

Department of Chemical Engineering
School of Engineering
Aristotle University of Thessaloniki

Exposure Biology Workflow

**LC/MS
GC/MS**
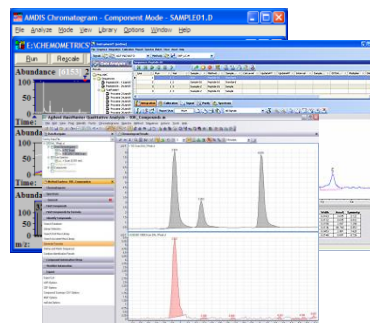
**MassHunter Qual/Quant
ChemStation AMDIS**

Microarrays
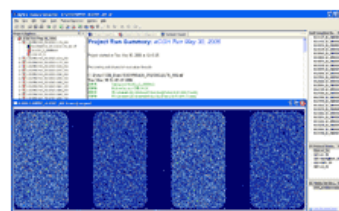
Feature Extraction

GeneSpring Platform
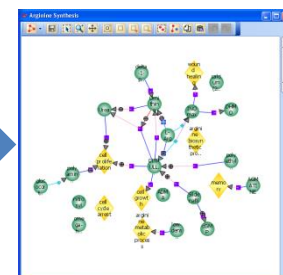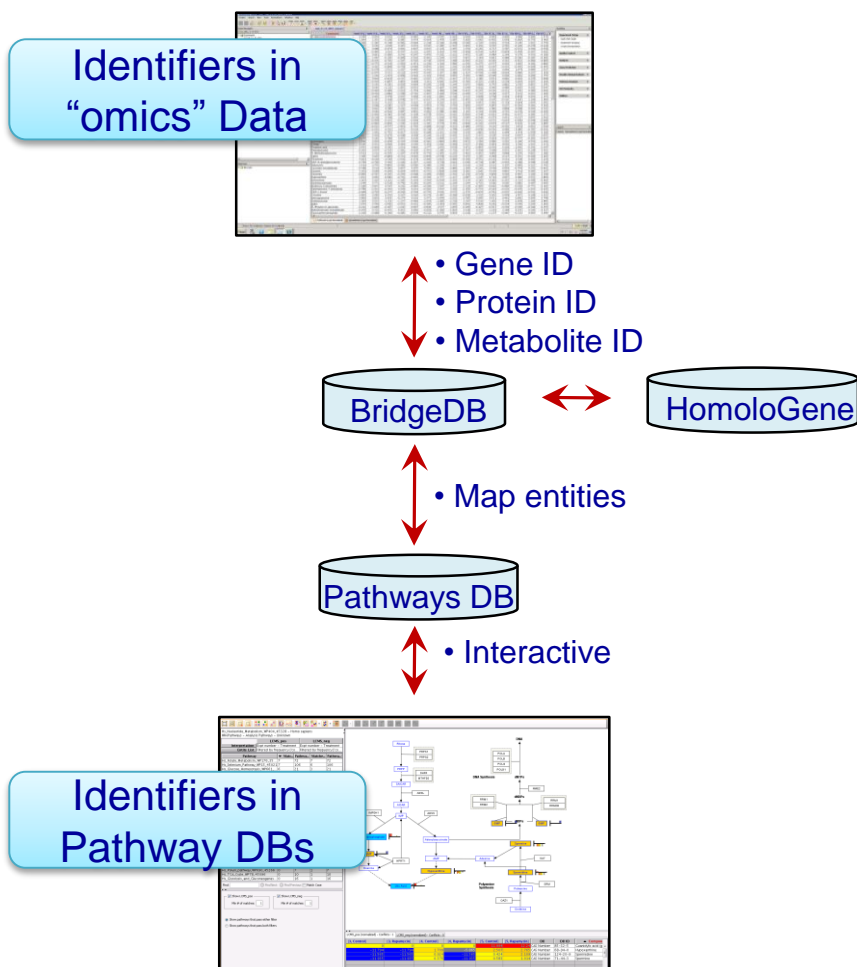
**Biological
Pathways**

NGS

Alignment to Reference Genome

Environmental
Engineering
Laboratory

Department of Chemical Engineering
School of Engineering
Aristotle University of Thessaloniki

BridgeDb for multi-omics pathway mapping

Identifiers in "omics" Data

• Gene ID
• Protein ID
• Metabolite ID

BridgeDB ↔ HomoloGene

• Map entities

Pathways DB

• Interactive

Identifiers in Pathway DBs

- **Metabolite Identifiers**
- KEGG
- HMDB
- ChEBI
- CAS
- **Protein Identifiers**
- Swiss-Prot
- UniProt
- UniProt/TrEMBL
- **Gene Identifiers**
- Entrez Gene, GenBank, Ensembl
- EC #, RefSeq, UniGene, HUGO
- HGNC, EMBL

Disease programming through life

Before conception

Fetal life

Childhood & adult life

Preconceptional exposures

In-utero exposures

Early life exposures

Later life exposures

Exposure of gametes

Epigenome at birth

Epigenome in childhood

Adult/aging epigenome

Epigenome (Parental)

Genome (parental)

Genome (offspring)

Programming of **disease** risks
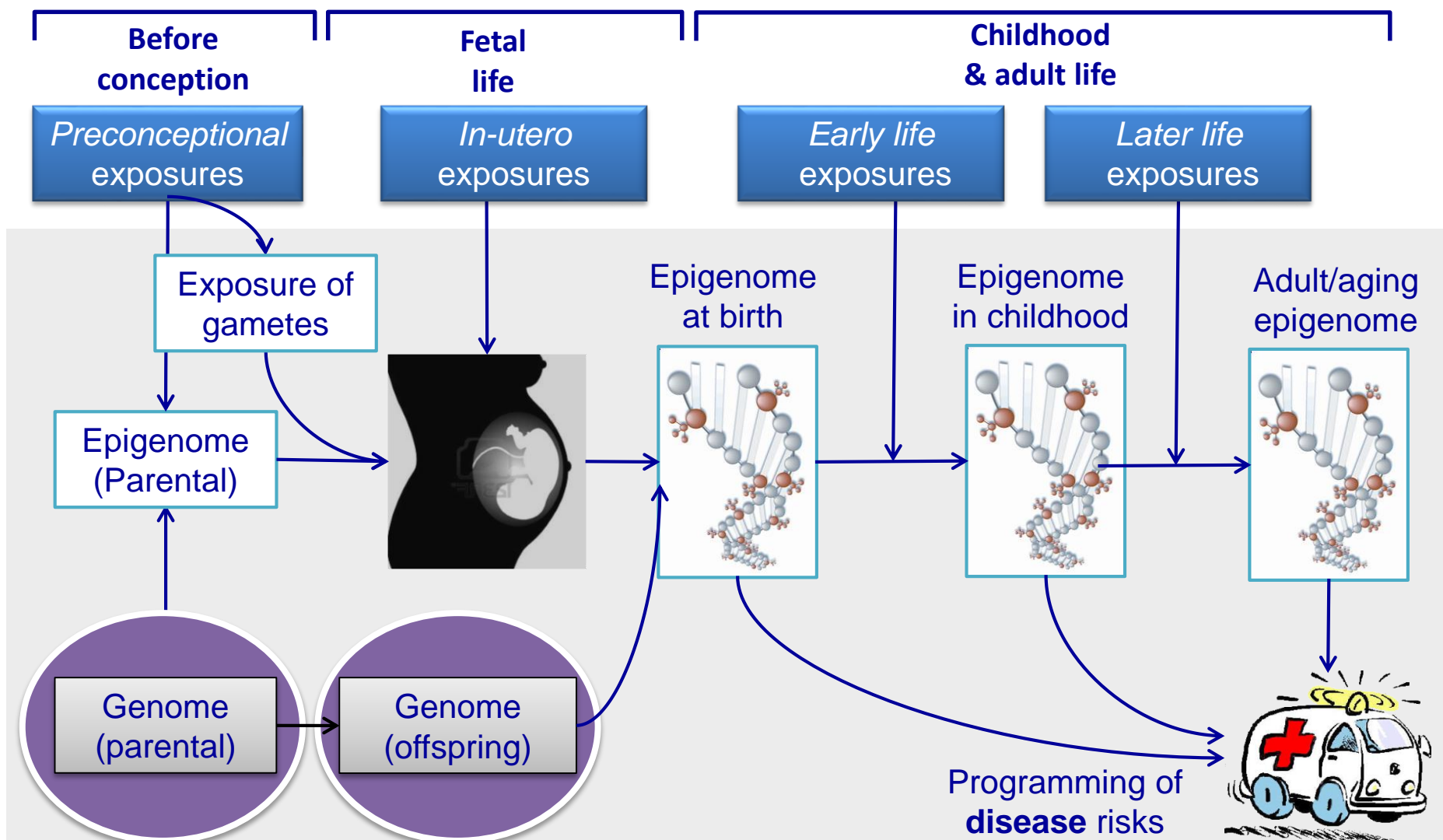
Figure adapted from Fleisch, Wright & Baccarelli, J Mol Endocrinol, 2012

Environmental
Engineering
Laboratory

Department of Chemical Engineering
School of Engineering
Aristotle University of Thessaloniki

The external exposome

Environmental
Engineering
Laboratory

Department of Chemical Engineering
School of Engineering
Aristotle University of Thessaloniki

EDMS – Data retrieval through map representation (1)

Environmental Engineering Laboratory

ENVElab

EDMS – Data retrieval through map representation (2)

Department of Chemical Engineering
School of Engineering
Aristotle University of Thessaloniki

Press the link on a station

External exposure advances

An individual's space-time activity model

Timeline — Hazard prism — Hazard vector
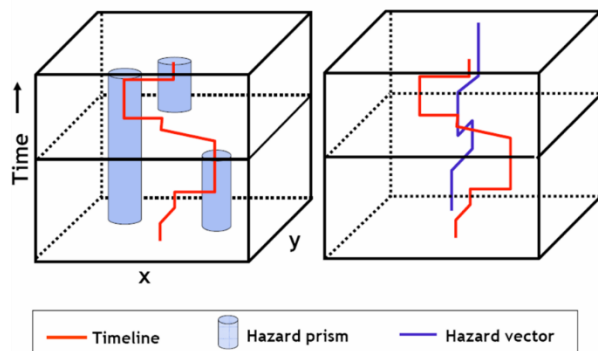
Agent based modelling
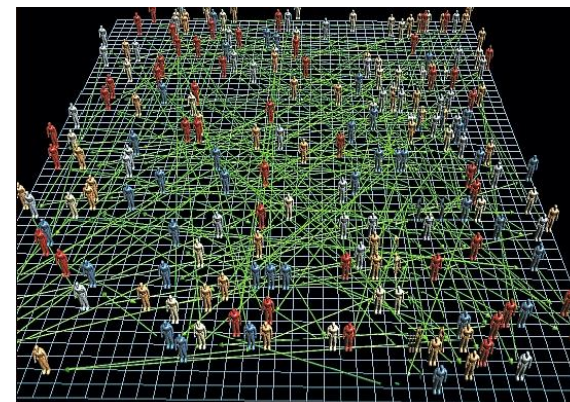
Sensors for exposure assessment

WOCKETS SYSTEM VISION

Multiple, low-cost 3-axis accelerometers stream data in real-time to mobile phone

Sensors miniature, thin, and ergonomic; worn under clothing 24/7

Phone carried in typical fashion (e.g. in pocket)

Wearable sensors (test version 1)

Pattern recognition algorithms running continuously on phone detect physical activities in real-time

Innovative phone apps possible

Environmental
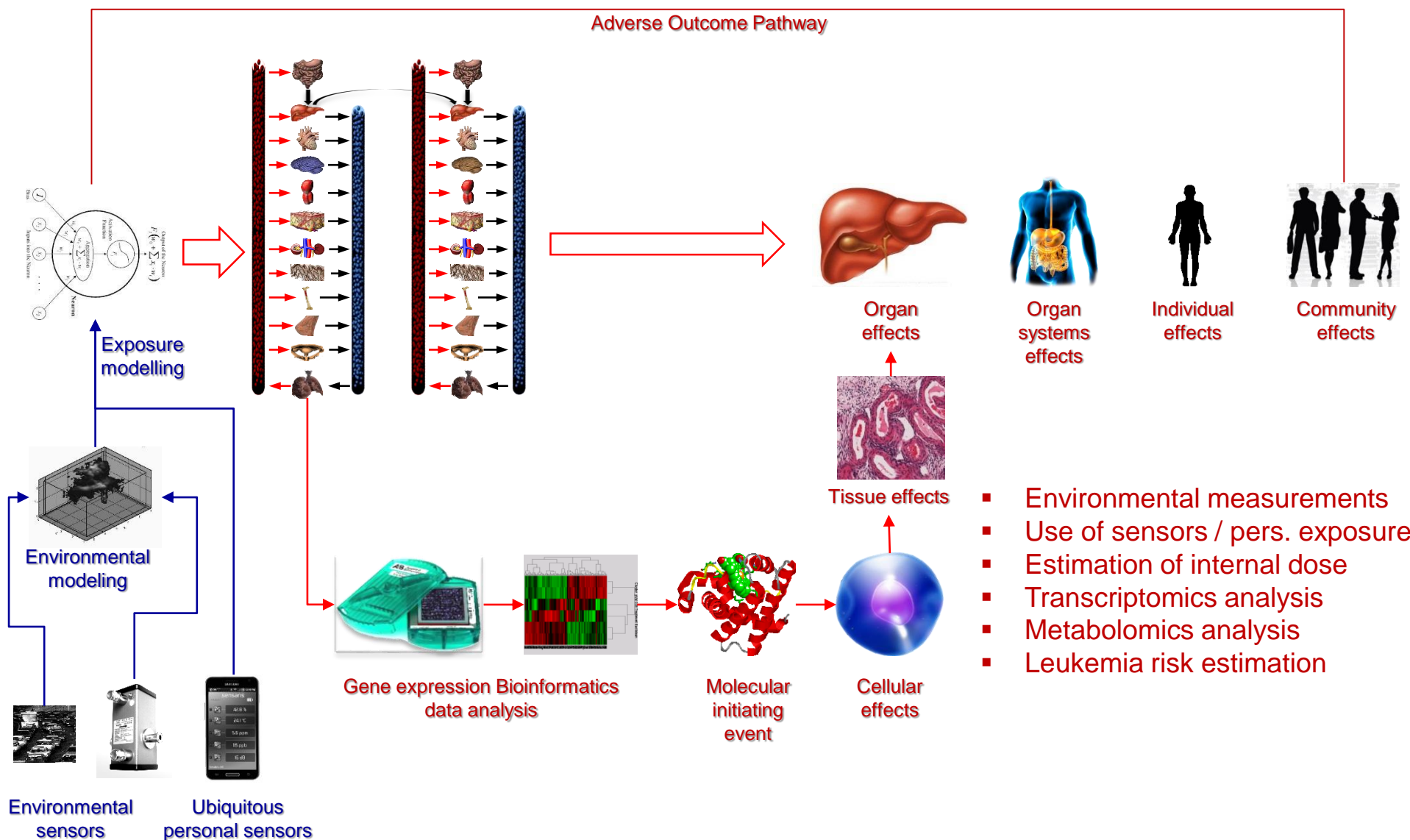Engineering
Laboratory

Department of Chemical Engineering
School of Engineering
Aristotle University of Thessaloniki

A case on the HEALS paradigm: Co-exposure to VOCs

A case on the HEALS paradigm: Co-exposure to VOCs

Adverse Outcome Pathway

Exposure modelling

Environmental modeling

Organ effects

Organ systems effects

Individual effects

Community effects

Tissue effects

Gene expression Bioinformatics data analysis

Molecular initiating event

Cellular effects

Environmental sensors

Ubiquitous personal sensors

- Environmental measurements
- Use of sensors / pers. exposure
- Estimation of internal dose
- Transcriptomics analysis
- Metabolomics analysis
- Leukemia risk estimation

Overall
wear - sync - export
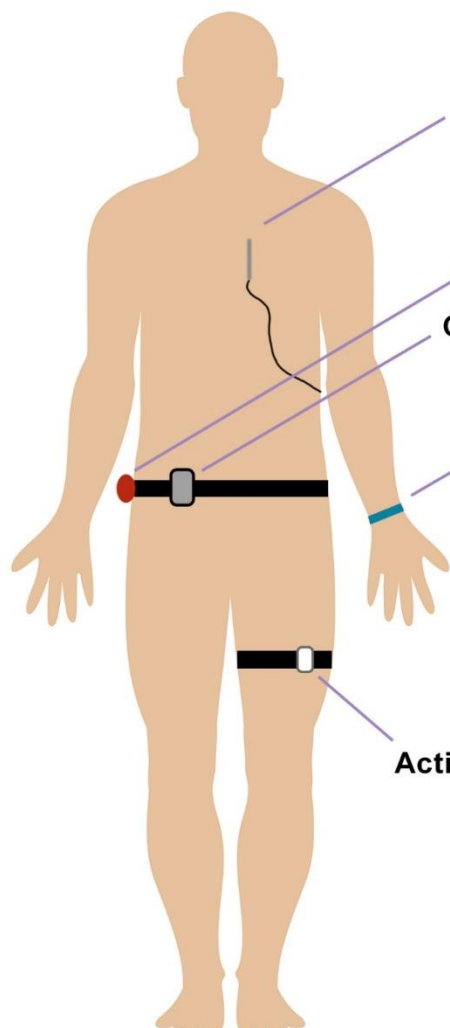
Elitech Temperature Logger
- Worn on the upper torso
- Sensor should not touch skin or other object
- **Export data every day**

Actigraph — Just wear it all the time

GPS
- Keep GPS on at all times
- Logger should be on LOG
- **Charge it each night**

FitBit Flex
- Keep on your hand even in your sleep
- **Daily sync**
- **Charge after 3-4 days**

Activ8
- Worn around the middle of the thigh at all times
- **Sync and charge every 2-3 days**
- **Export data every day**

Moves (iOS / Android)
- Always carry your smartphone (pocket or bag)
- Charge your smartphone every night
- Code "home", "work" etc.

Paper Log — Start time, End time, Location, Comments

Environmental
Engineering
Laboratory

Department of Chemical Engineering
School of Engineering
Aristotle University of Thessaloniki

"Intelligent" location tracking

Statistical Method: Predicting location based on temp and UV logs through an Artificial Neural Network
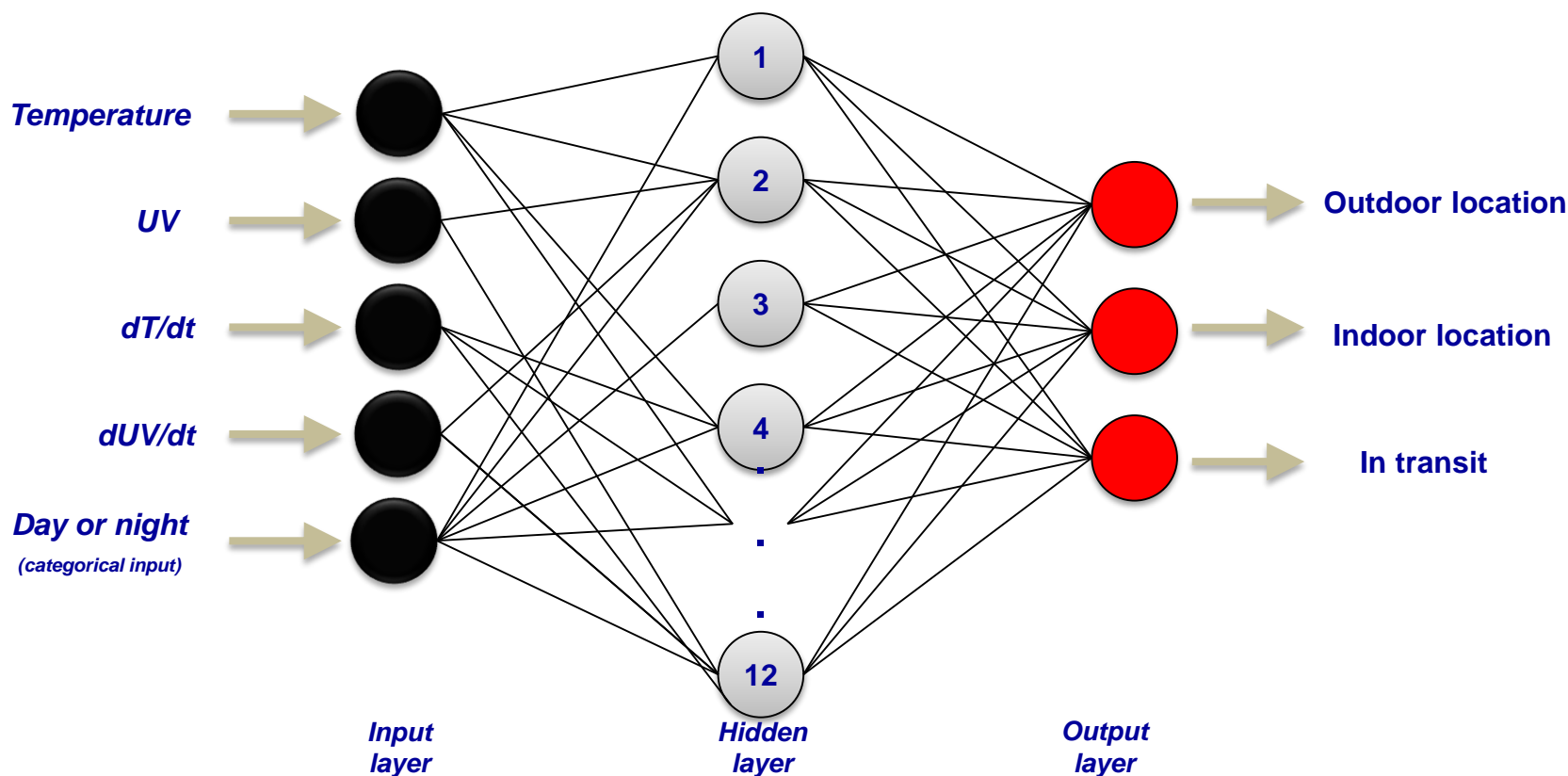
5 input nodes (4 numerical and 1 categorical)
12 hidden nodes (found to yield the best results among several combinations)
3 output nodes (corresponding to the 3 different classes)
Data from 5 days were used for training the ANN model
Data from 2 days were used as an independent dataset to validate the ANN model

Temperature

UV

dT/dt

dUV/dt

Day or night
(categorical input)

1

2

3

4

12

Outdoor location

Indoor location

In transit

Input layer

Hidden layer

Output layer

Environmental
Engineering
Laboratory

Department of Chemical Engineering
School of Engineering
Aristotle University of Thessaloniki

Predicting location
based on temp and UV logs through ANN

Transition from indoors to outdoors is captured

Environmental
Engineering
Laboratory

Department of Chemical Engineering
School of Engineering
Aristotle University of Thessaloniki

Identifying activity patterns Based on Agent Based Modelling
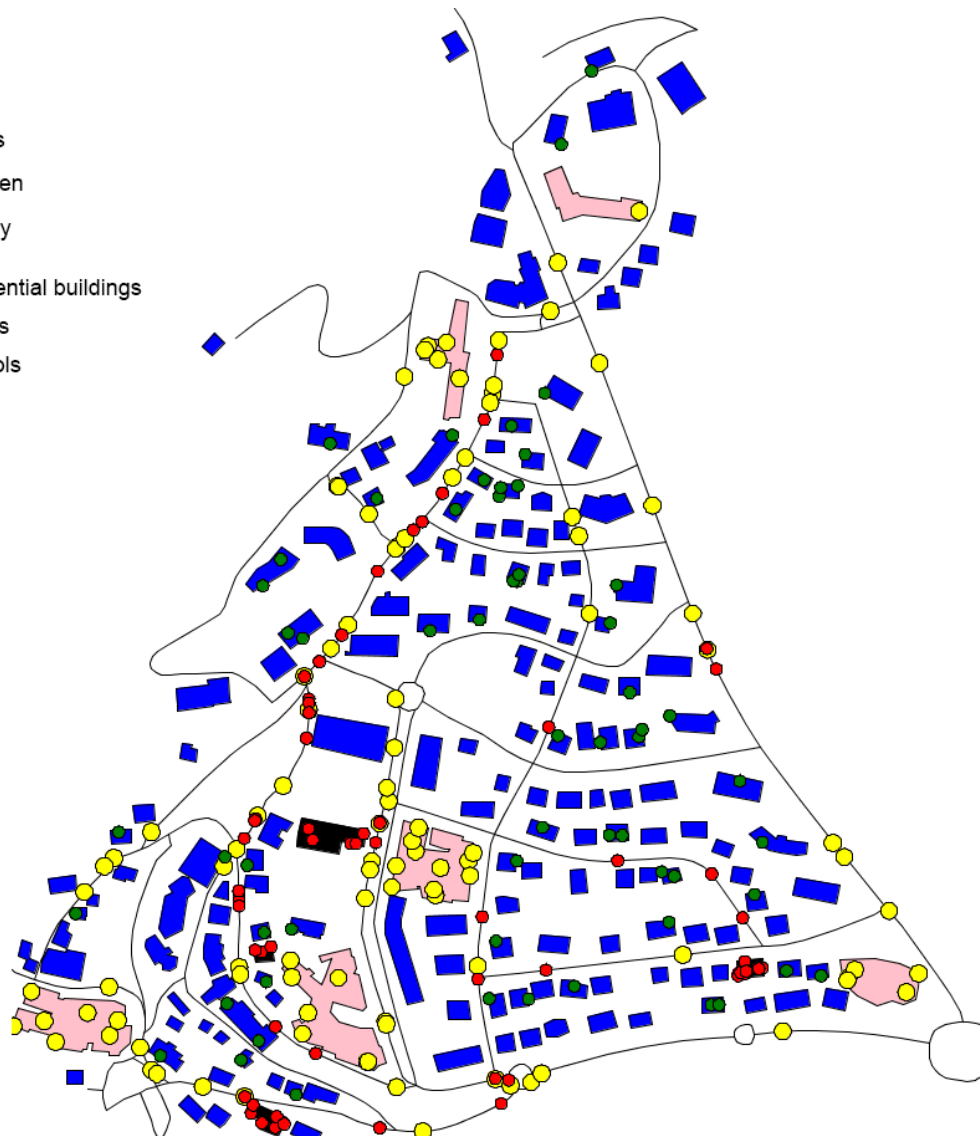
Multi-omics responses and associations

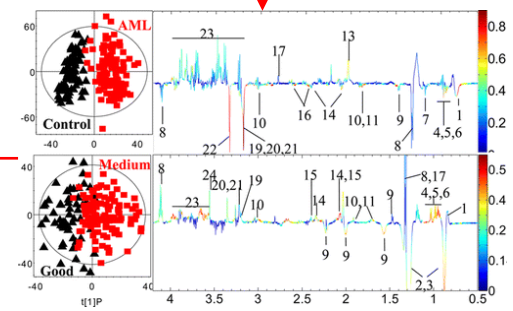## Identification of differentially expressed **proteins**

Apoptosis Signaling Pathway
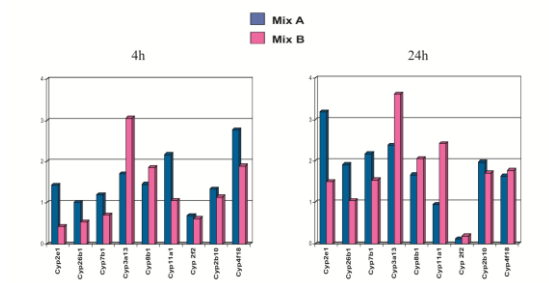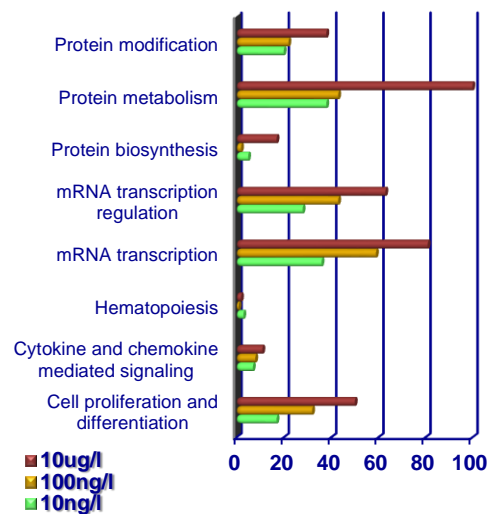
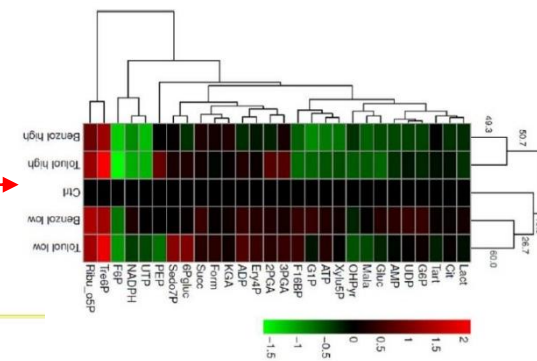**Transcriptomics** responses to chemical BTEX mixtures
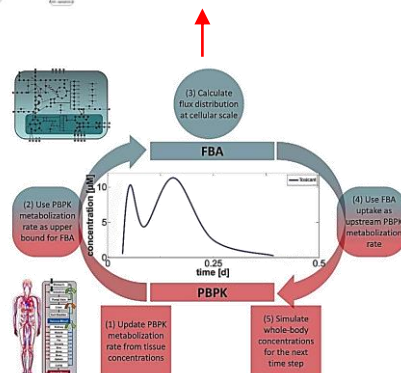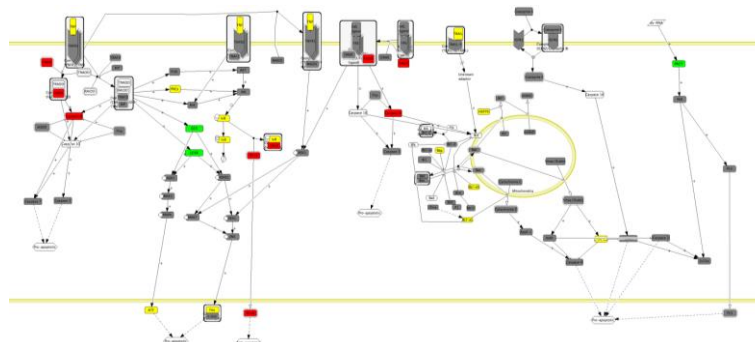
Extracellular perturbations on metabolic states

Identification of AML **metabolomic** fingerprint

Toxic metabolites in bone marrow concentration modulation
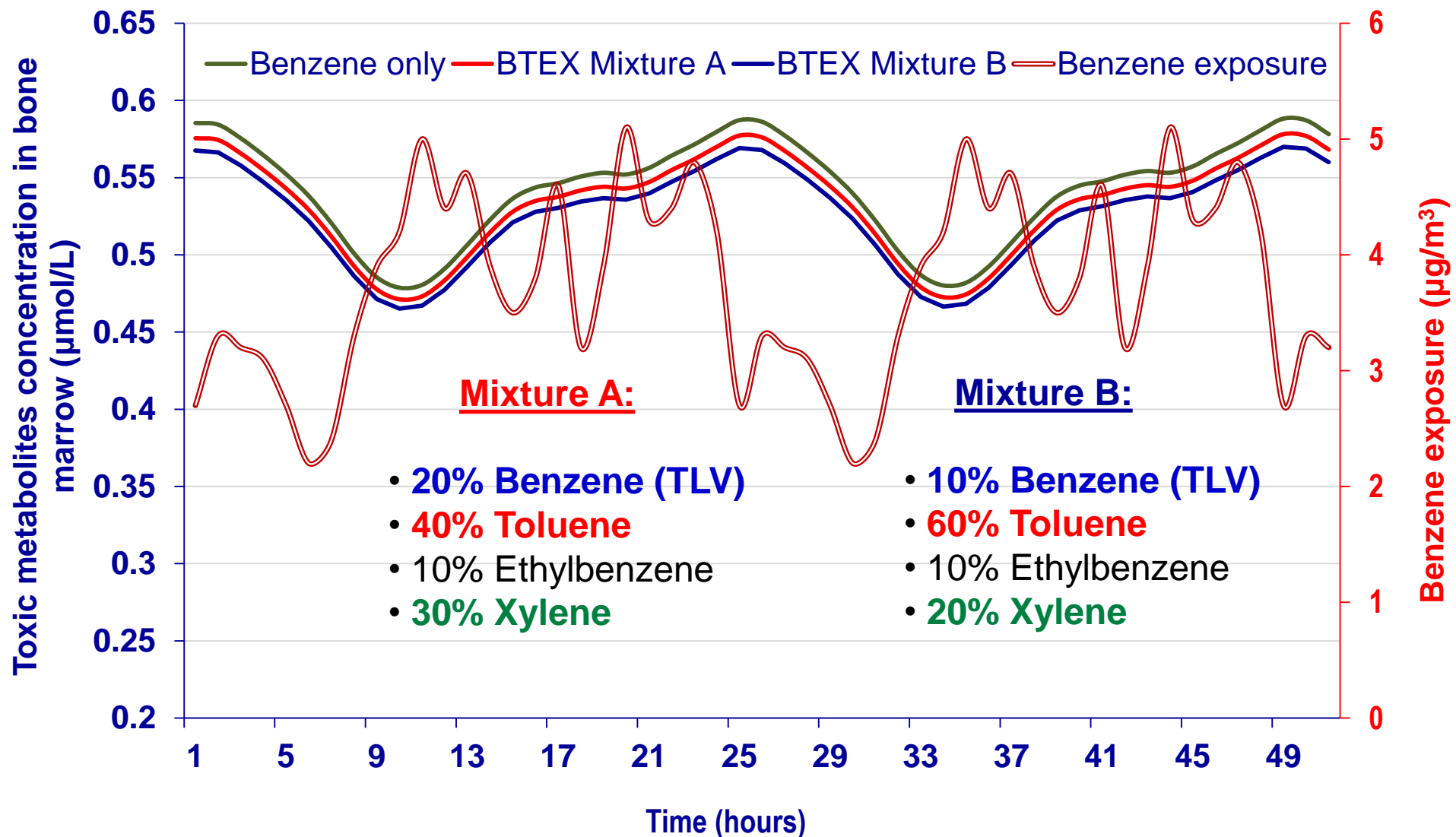
Mixture A:
- **20% Benzene (TLV)**
- **40% Toluene**
- 10% Ethylbenzene
- **30% Xylene**

Mixture B:
- **10% Benzene (TLV)**
- **60% Toluene**
- 10% Ethylbenzene
- **20% Xylene**

Environmental
Engineering
Laboratory

Department of Chemical Engineering
School of Engineering
Aristotle University of Thessaloniki

Conclusions

- Integrated use of existing environmental and biomonitoring data
- Improved assessment of the external exposome
  - Environmental data fusion and Agent-Based Models
  - Mobile phone apps
  - Environmental sensor-webs
  - Micro-sensors
  - Satellite remote sensing
- Linking external and internal exposome
  - Integrated use of –omics and chemical biomarker data
  - Take into account the temporal dimension
- Advanced tools for environmental and biological data analysis
  - PBBK modeling for internal dose estimation and exposure reconstruction
  - Coupling PBBK models with gene regulation models
- Novel bioinformatics strategies for biomarker prediction
  - Meta-modeling for biomarker fusion
- Environment-wide association studies
  - Linkage disequilibrium
  - Use of advanced statistical tools: DAG, Bayesian inference
- Enviromics
  - study of a wide array of environmental factors in relation to health and biology

Environmental
Engineering
Laboratory

Department of Chemical Engineering
School of Engineering
Aristotle University of Thessaloniki

Thank you for your kind attention

www.enve-lab.eu

A connectivity perspective to environmental health