Metabolomic phenotyping in medical systems biology

Matej Orešič
Professor, Systems biology & bioinformatics
VTT Technical Research Centre of Finland
http://sysbio.vtt.fi/

Systems Biology and the Omics Cascade
Stockholm, 13.6.2008
OMICS revolution over the past decade

- Increased ability to measure large number of "parts" of the biological systems and their activities
  - Genes and their expression
  - Proteins and their modifications
  - Small molecules (metabolites) and their reactions
  - Imaging technologies, incl. in vivo
  - Microbial populations
  - "single cell" measurements
  - etc etc
**Systems and levels**

**Hierarchical Mappings**
- **System-wide**

**Compartmental Processes**
- **Organ system**
  - Homeostasis and regulation, metabolite inputs and outputs, endocrine targeting
- **Body**
  - System regulation, Disease staging, treatment monitoring, pharmacodynamics
- **Tissue**
  - Cell Communication, Cytokines, antigen presentation, tissue-type specificity
- **Cell**
  - Gene and Protein Expression and Regulation, Intercellular biochemical activities

**Need for models linking the phenotype with the genetic & environmental factors**

**Systems Biology**
Metabolome

NMR

Mass Spectrometry

CUSTOM

Homeostasis ‘Housekeeping’
Organic Acids
Lipids
Amino Acids
Nucleotides
Steroids
Eicosanoids
Neurotransmitters
Peptides
Trace elements

Biofluid metabolic profile = Phenotype
Metabolomics as a platform for systems biology

SENSITIVITY
As proven via the formalism of Metabolic Control Analysis; small changes in activities of individual enzymes lead to small changes in metabolic fluxes, but can lead to large changes in metabolite concentrations.
Metabolomics platform
Experiment design + Analytical chemistry + Chemometrics + Bioinformatics

Samples (biofluids, tissues, cells) → Analytical methods → Metabolic profiling (LC/MS, GC/MS, NMR) → Statistical analysis → Data processing → Biomarkers, biological insight

Bioinformatics

Analytical methods:
- Screening (hydrophilics)
- Lipidomics
- Eicosanoids
- Amino acids
- Other methods
Why measure lipids?

Membrane Structure & Function; Signaling; Energy; Storage

Fatty Acids and Metabolic Homeostasis

Cell Plasma Membrane

High Density Lipoprotein (HDL)

Storage Lipids

Glycerolipids

Cholesteryl esters

Membrane Lipids

Phospholipids

Sphingolipids

Sterols

TAG 14:0/16:0/16:0: CE 16:0
TAG 16:0/16:0/18:0: CE 18:1
TAG 18:1/18:1/20:4: CE 20:3
TAG 16:0/18:0: CE 20:4
TAG 16:0/18:0: CE 18:1
TAG 18:1/18:1/20:4: CE 20:3
TAG 16:0/18:0: CE 20:4

Cholesterol

PC 16:0/16:0
PC 18:1/18:2
PC 18:0/20:4
PC 16:0/22:6
PC 18:1/20:3
PC 18:1/20:3
PE 16:0/22:6
PE 18:1/20:3
PE 17:0/18:1

LPC 16:0
LPC 18:1
SM 18:0
SM 24:1

VIT
Platforms

UPLC-TOF/MS lipidomics
(major phospholipids, sphingolipids, acylglycerols)
10-15μl serum sample used

GCxGC-TOF/MS
(global metabolome)
20μl serum sample used
MZmine 2.0: data processing for metabolomics

http://mzmine.sourceforge.net
Outline

• 1 genome?
  • Genetic factors affecting the metabolic phenotype

• Metabolic states & development
  • Changes of metabolic phenotypes with age

• Beneficial autoimmunity?
  • Metabolic phenotypes & immune response

• Drug response phenotyping
  • Tissue-specific drug effect on metabolic phenotypes
1 genome?
Gut microbes

We carry 10 times more microbial cells as the host mammalian cells (~100 trillion bacteria).

Human gut microbes are associated with obesity and lipid metabolism

**MICROBIAL ECOLOGY**

Human gut microbes associated with obesity

Serum lipidome is affected by gut microbial composition

Comparison of serum lipidomic profiles of three groups of male Swiss Webster mice of different gut microbial composition:

1. **Germ free**
2. **Conventionalized**: GF mice colonized for two weeks at adulthood
3. **Conventionally raised**, i.e., with normal microbiota
Metabolic states & development
Sample series

- 59 children between 3 months and 4 years of age
  - 27 boys
  - 32 girls
- Serum sample collection every 2-7 months
- Children remained healthy (no chronic disease) throughout the follow-up
- 11 samples per child on average
- Samples from the Type 1 Diabetes Prediction & Prevention study (DIPP)
"Normal" metabolome changes with age

- Describe as progression of metabolic states
- Apply Hidden Markov Model methodology to describe the states and their progression

\[ S_1^x \xrightarrow{p_{11}^x} S_2^x \xrightarrow{p_{22}^x} S_3^x \xrightarrow{p_{33}^x} S_4^x \xrightarrow{p_{44}^x} S_5^x \]

\[ p_{12}^x \quad p_{21}^x \quad p_{23}^x \quad p_{34}^x \quad p_{45}^x \]

\( x \) - Boys/Girls

First five years: progression not the same for each child

Major differences between the states

- Upregulation of proinflammatory lysosphosphatidylcholines and short chain saturated triacylglycerols near 1 year
- Dietary triacylglycerols upregulated near 3 years of age

![Graph showing differences in lipids between girls and boys at different ages.](image-url)
Developmental metabolic differences between girls and boys

Sphingomyelins consistently elevated in girls

Metabolic phenotypes may help detect subtle changes related to early disease development or responses to therapeutic interventions.
Beneficial autoimmunity?
Type 1 diabetes

- T1D is a **chronic autoimmune disease** caused by destruction of the insulin-producing beta cells in the pancreatic islets of Langerhans.
- In most Western countries, the **incidence has increased by 3% per year** during the past 50 years.
- The disease is **multifactorial and polygenic** showing tight linkage with certain HLA-DQ and DR alleles.
- **As only a fraction of those at genetic risk develop T1D**, the impact of environment on disease pathogenesis is obvious.
- A symptom-free prediabetic period is characterized by **T lymphocyte accumulation** to the islets.
Persisting unknowns

• Disease risk and time of onset?

• Triggers of the disease process(es)?

• Mechanisms regulating progression towards T1D?

• Prevention?
• Type 1 Diabetes Prediction and Prevention Project (DIPP) launched Nov 7, 1994 in Turku
• Oulu joined 1 yr and Tampere 3 yrs later
• 20% of newborns screened in Finland

• SYSDIPP – Systems Biology Approach to Biomarker Discovery in Type 1 Diabetes started in 2005 (Tekes FinnWell Program)
In follow-up N=8026

≥1 antibody N=1167

≥2 antibody N=430

Insulin trial N=175

T1D N=51

Cases and controls matched by gender, HLA genotype, city and period of birth.

Metabolomics study design

Progressors N=50

Non-progressors N=67

Progressors N=6

Non-progressors N=6

Metabolomics study design
## Sample series

<table>
<thead>
<tr>
<th>Batch</th>
<th>City of birth</th>
<th>Study</th>
<th>Year of birth</th>
<th>Age at diagnosis</th>
<th>Number of progressors</th>
<th>Number of non-progressors</th>
<th>Number of samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Turku</td>
<td>DIPP</td>
<td>1994-2001</td>
<td>1-11y</td>
<td>13</td>
<td>26</td>
<td>441</td>
</tr>
<tr>
<td>2</td>
<td>Turku</td>
<td>DIPP</td>
<td>1996-1999</td>
<td>1-6y</td>
<td>10</td>
<td>13</td>
<td>185</td>
</tr>
<tr>
<td>3</td>
<td>Oulu</td>
<td>DIPP</td>
<td>1996-2001</td>
<td>1-8y</td>
<td>27</td>
<td>28</td>
<td>483</td>
</tr>
<tr>
<td>4</td>
<td>Turku</td>
<td>STRIP</td>
<td>1990</td>
<td>3-13y</td>
<td>6</td>
<td>6</td>
<td>87</td>
</tr>
</tbody>
</table>

**TOTAL** 56 73 1196

No effect of sample age ($r=-0.05, P=0.94$)
Age-based comparison of molecular profile changes between cases and controls.
Ether-linked phosphocholines decreased in individuals who later developed autoimmunity and Type 1 Diabetes

Plasmalogens are most abundant class of ether linked phospholipids, known as endogenous antioxidants.

**Age 1 year**

<table>
<thead>
<tr>
<th>GPCho(O-18:0/18:2)</th>
<th>Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Progressors</td>
<td>0.001</td>
</tr>
<tr>
<td>Non-progressors</td>
<td></td>
</tr>
</tbody>
</table>

**Age 6 years**

<table>
<thead>
<tr>
<th>GPCho(O-18:0/18:2)</th>
<th>Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Progressors</td>
<td>0.044</td>
</tr>
<tr>
<td>Non-progressors</td>
<td></td>
</tr>
</tbody>
</table>
Ethanolamine plasmalogen also decreased

Plasmalogens are most abundant class of ether linked phospholipids, known as endogenous antioxidants.

**Age 1 year**

GPEtn(O-18:1(1Z)/20:4)

\[ p=0.005 \]

**Age 6 years**

GPEtn(O-18:1(1Z)/20:4)

\[ p=0.019 \]
Can differences be explained by genetic risk?

**High risk**: DR3-DQ2/DR4-DQ8  
**Medium risk**: DR4-DQ8/*x*  
(*x* = any haplotype except DR2-DQ6, DR5-DQ7 or DR3-DQ2)

No significant association of HLA-associated genetic risk and the lipid profiles
Seroconversion for islet autoimmunity

State based comparison of molecular changes near the seroconversion

Time from birth (months)

Progressor 1
Non-progressor 1
Progressor 2
Non-progressor 2
Lipidomic profiles near seroconversion for islet autoimmunity

Elevated lysophosphatidylcholines

GPCho(18:0/0:0)

GPCho(16:0/0:0)

Progressors 3-6 months prior to seroconversion (Ser-) and 3-6 months after seroconversion (Ser+), with matched non-progressors.
Implications

• Our findings clearly favor early immunomodulation, rather than immunosuppression, as a preventive therapy, with the aim to boost the beneficial component of autoimmunity.

• The factors leading to metabolic stress and autoimmune responses clearly need to be investigated in further studies in the context of autoimmune diseases in general.
Drug response phenotyping
Statins

• Lipid lowering drugs
• Reduction in atherosclerotic complications
• Higher doses of statins are being recommended today for lowering of cholesterol
  ➔ Increased risk of myopathy (muscle toxicity)

• Mechanisms or biomarkers of myopathy not known
Statin induced muscle toxicity

• Muscle complaints without creatine kinase elevations occur in 5-10% of patients in clinical trials
  • However, the complaints occur with similar frequency in statin and placebo groups, thus it is commonly believed they are not drug related

• In recent PRIMO study (N=7924, high dose statins) 10.5% patients complained of muscle pain, highest rate of 18% associated with simvastatin treatment
High dose statin trial – strategy to elucidate early mechanisms of statin induced myopathy

Myopathy cases: Muscle damage may be too severe to permit sensible analysis of specific pathophysiological phenomena
**Plasma lipids**

![Bar chart showing the percentage change in cholesterol, LDL cholesterol, and triglycerides over 8 weeks for Placebo, Simva 80 mg/d, and Atorva 40 mg/d groups.]

- **Placebo**
  - Total Cholesterol: -36.0%
  - LDL-Cholesterol: -35.0%
  - Triglycerides: -28.5%

- **Simva 80 mg/d**
  - Total Cholesterol: -54.0%
  - LDL-Cholesterol: -48.5%
  - Triglycerides: -35.6%

- **Atorva 40 mg/d**
  - Total Cholesterol: -28.5%
  - LDL-Cholesterol: -35.6%

* * P < 0.05 Paired t-test -
** P < 0.01 baseline vs. endpoint
*** P < 0.001 within groups

Mitochondrial DNA in skeletal muscle

Figure 1. Changes in mtDNA/nDNA ratio after 8 weeks of high-dose statin treatment.

<table>
<thead>
<tr>
<th></th>
<th>Change in mtDNA/nDNA ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Placebo N=14</td>
<td></td>
</tr>
<tr>
<td>Atorvastatin N=15</td>
<td></td>
</tr>
<tr>
<td>Simvastatin N=14</td>
<td></td>
</tr>
</tbody>
</table>

p = 0.001

BA Schick et al (Clin Pharmacol Ther. 2007)
Pathway analysis; GSEA (muscle)

- No significant changes in placebo or atorvastatin groups
- Several upregulated pathways in simvastatin group (FDR $q<0.1$)

<table>
<thead>
<tr>
<th>NAME</th>
<th>Source</th>
<th>SIZE</th>
<th>ES</th>
<th>NES</th>
<th>NOM p-val</th>
<th>FDR q-val</th>
</tr>
</thead>
<tbody>
<tr>
<td>ST_T_Cell_Signal_Transduction</td>
<td>Signalling Transduction KE</td>
<td>45</td>
<td>0.58</td>
<td>1.68</td>
<td>0.000</td>
<td>0.055</td>
</tr>
<tr>
<td>soddPathway</td>
<td>BioCarta</td>
<td>10</td>
<td>0.83</td>
<td>1.69</td>
<td>0.014</td>
<td>0.060</td>
</tr>
<tr>
<td>Eicosanoid_Synthesis</td>
<td>BioCarta</td>
<td>15</td>
<td>0.79</td>
<td>1.66</td>
<td>0.000</td>
<td>0.071</td>
</tr>
<tr>
<td>plcPathway</td>
<td>BioCarta</td>
<td>8</td>
<td>0.78</td>
<td>1.69</td>
<td>0.000</td>
<td>0.072</td>
</tr>
<tr>
<td>tubbyPathway</td>
<td>BioCarta</td>
<td>7</td>
<td>0.70</td>
<td>1.70</td>
<td>0.000</td>
<td>0.076</td>
</tr>
<tr>
<td>caspasePathway</td>
<td>BioCarta</td>
<td>23</td>
<td>0.68</td>
<td>1.65</td>
<td>0.008</td>
<td>0.077</td>
</tr>
<tr>
<td>CR_DEADTH</td>
<td>Brentani PNAS 2003 (31)</td>
<td>76</td>
<td>0.53</td>
<td>1.64</td>
<td>0.005</td>
<td>0.079</td>
</tr>
<tr>
<td>hivnefPathway</td>
<td>BioCarta</td>
<td>58</td>
<td>0.56</td>
<td>1.64</td>
<td>0.000</td>
<td>0.085</td>
</tr>
<tr>
<td>ephA4Pathway</td>
<td>BioCarta</td>
<td>10</td>
<td>0.86</td>
<td>1.61</td>
<td>0.004</td>
<td>0.085</td>
</tr>
<tr>
<td>deathPathway</td>
<td>BioCarta</td>
<td>33</td>
<td>0.59</td>
<td>1.61</td>
<td>0.005</td>
<td>0.089</td>
</tr>
<tr>
<td>MAP00590_Prostaglandin_and_leukotriene_metabolism</td>
<td>GenMAPP</td>
<td>19</td>
<td>0.72</td>
<td>1.62</td>
<td>0.005</td>
<td>0.089</td>
</tr>
<tr>
<td>nkcellsPathway</td>
<td>BioCarta</td>
<td>20</td>
<td>0.64</td>
<td>1.60</td>
<td>0.000</td>
<td>0.090</td>
</tr>
<tr>
<td>SA_CASPASE_CASCADE</td>
<td>SigmaAldrich</td>
<td>19</td>
<td>0.63</td>
<td>1.59</td>
<td>0.000</td>
<td>0.091</td>
</tr>
<tr>
<td>rac1Pathway</td>
<td>BioCarta</td>
<td>22</td>
<td>0.73</td>
<td>1.62</td>
<td>0.004</td>
<td>0.092</td>
</tr>
<tr>
<td>ST_Dicyostelium_discoideum_CAMP_Chemotaxis_Pathway</td>
<td>Signalling Transduction KE</td>
<td>33</td>
<td>0.64</td>
<td>1.61</td>
<td>0.007</td>
<td>0.093</td>
</tr>
<tr>
<td>nktPathway</td>
<td>BioCarta</td>
<td>29</td>
<td>0.62</td>
<td>1.62</td>
<td>0.018</td>
<td>0.094</td>
</tr>
<tr>
<td>eosinophilsPathway</td>
<td>BioCarta</td>
<td>8</td>
<td>0.77</td>
<td>1.58</td>
<td>0.025</td>
<td>0.094</td>
</tr>
<tr>
<td>tall1Pathway</td>
<td>BioCarta</td>
<td>15</td>
<td>0.56</td>
<td>1.60</td>
<td>0.003</td>
<td>0.095</td>
</tr>
<tr>
<td>il17Pathway</td>
<td>BioCarta</td>
<td>15</td>
<td>0.78</td>
<td>1.57</td>
<td>0.023</td>
<td>0.095</td>
</tr>
<tr>
<td>CBF_LEUKEMIA_DOWNING_AML</td>
<td>Manually Curated</td>
<td>75</td>
<td>0.60</td>
<td>1.58</td>
<td>0.023</td>
<td>0.097</td>
</tr>
<tr>
<td>ureacylePathway</td>
<td>BioCarta</td>
<td>7</td>
<td>0.87</td>
<td>1.57</td>
<td>0.008</td>
<td>0.098</td>
</tr>
<tr>
<td>cell_motility</td>
<td>GO</td>
<td>116</td>
<td>0.61</td>
<td>1.58</td>
<td>0.014</td>
<td>0.098</td>
</tr>
<tr>
<td>MAP00562_Inositol_phosphate_metabolism</td>
<td>GenMAPP</td>
<td>20</td>
<td>0.66</td>
<td>1.57</td>
<td>0.032</td>
<td>0.099</td>
</tr>
</tbody>
</table>
PLS/DA on combined serum lipid and muscle gene expression data (4 pathways: PLC, eicosanoid, sodd, and tubby)

PLS/DA on combined serum lipid and muscle gene expression data (4 pathways: PLC, eicosanoid, sodd, and tubby)

Coregulated muscle transcripts and plasma lipids in simvastatin group
lasso regression of lipid data on ALOX5AP

• Shrinkage regression method, which performs continuous variable selection causing some of the regression coefficients to be exactly zero
• Shrinkage reduces the variance of the regression estimates

Concept: Drug response phenotyping using systems approach

Profiles of dysregulated pathways

Robust regression of biofluid metabolome on tissue pathways

Extend biofluid metabolomics to monitor tissue sensitive biomarkers in larger populations
Summary

- Medical systems biology aims to elucidate complex networks linking phenotypes with genes and environment
- Biofluid metabolome is a quantitative measure of the phenotype
- Metabolic phenotype depends on genetic factors reflected in host and microbial cells
- Changes of metabolic phenotypes can be described in terms of metabolic states
- Autoimmunity may be physiological and beneficial response to metabolic stress
- Tissue-specific drug responses are reflected in changes in metabolic phenotypes
Acknowledgements

Orešič group
- Marko Sysi-Aho
- Tuulikki Seppänen-Laakso
- Mikko Katajamaa
- Laxman Yetukuri
- Tapani Suortti
- Airi Hyrkäs
- Leena Öhrnberg
- Helena Simolin
- Brudy Han Zhao
- Catherine Bounsaythip
- Anna-Liisa Ruskeepää
- Peddinti Gopalacharyulu

Hidden Markov Models
- Sami Kaski, Andrey Ermolov, Janne Nikkilä

High dose statin trial
- Reijo Laaksonen (Zora Biosciences & Tampere University Hospital)

Type 1 diabetes
- Olli Simell, Mikael Knip, Heikki Hyöty, Jorma Ilonen, Riitta Veijola, Riitta Lahesmaa
- Senior research staff and nurses
- DIPP parents & children

Gut microbiota
- Fredrick Bäckhed

Funding: Tekes, EU FP6 & FP7, Academy of Finland

http://sysbio.vtt.fi/