

Applicability of transcriptomics in biomonitoring studies

Jos Kleinjans

Dept. of Health Risk Analysis & Toxicology
Maastricht University

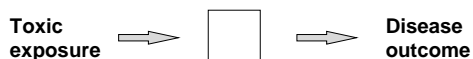


CHILDRENGENONETWORK

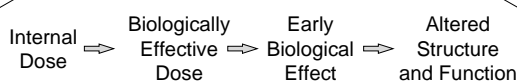
Outline of this presentation

- Methodology of genomics analysis (gene expression profiling) for the purpose of biomarker analysis
- Results from a pilot family study on multiple gene expression induced by air pollution

Traditional Epidemiology



Molecular Epidemiology



Susceptibility

Available biomarkers for monitoring environmental carcinogenesis in humans:

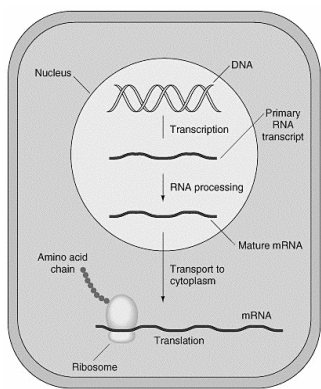
- imply different assays for each precarcinogenic event
- raise questions on assay specificity (biological relevance) and sensitivity

New monitoring methods should thus:

- be more generic, providing mechanistic information on multiple precarcinogenic effects simultaneously
- be suitable for assessing precarcinogenic effects of low dose exposure

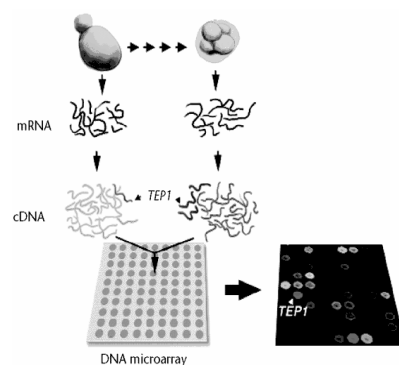
→ GENE EXPRESSION PROFILING AS A NOVEL EFFECT MARKER ?

Central dogma of gene function

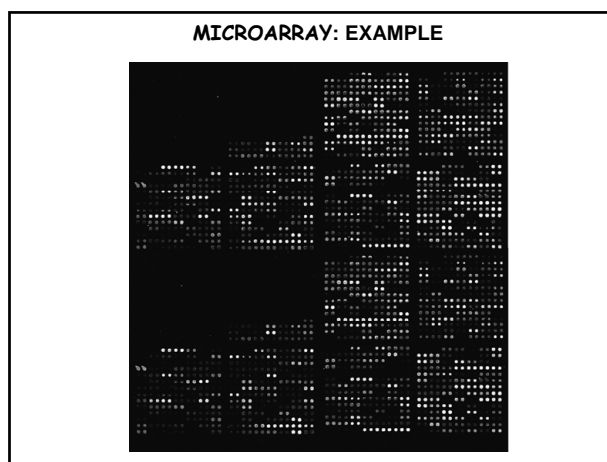
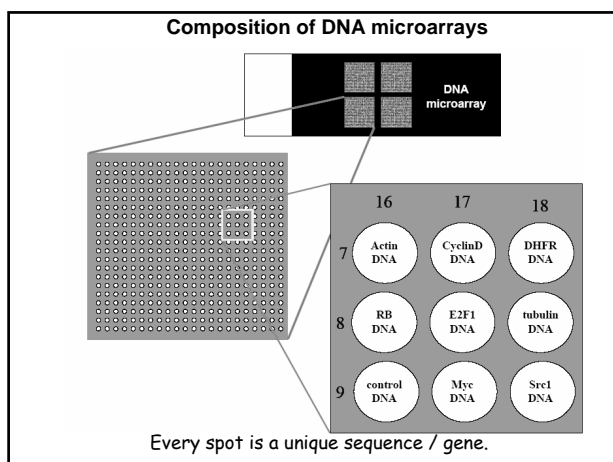


Genome analysis:

dual labeling and microarray hybridisation

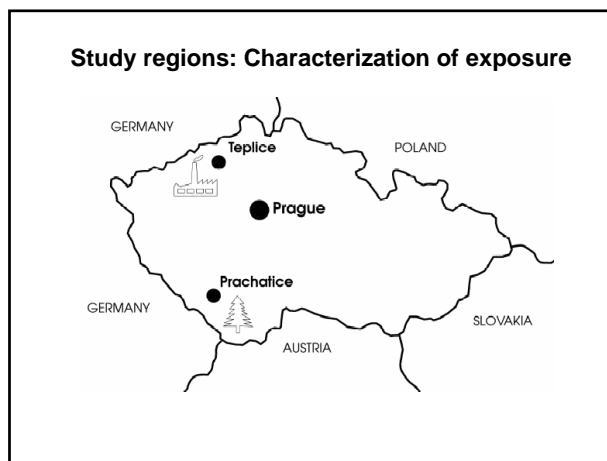


CHILD HEALTH AND THE ENVIRONMENT: RESULTS FROM EU FRAMEWORK 5



Aim of the present study

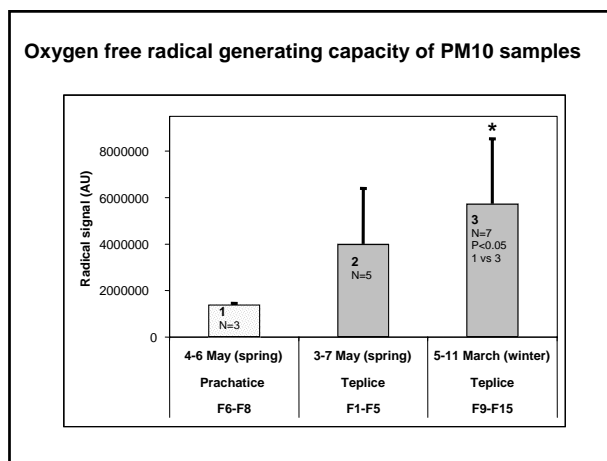
- To generate whole genome gene expression data from peripheral lymphocytes sampled from environmentally exposed subjects and controls, e.g. children from Czech Republic
- To compare obtained transcriptomic profiles with a classical and well validated marker of environmental cancer risk, e.g. lymphocytic micronuclei frequency
- To provide proof of the principle that genetic pathways discriminative for environmental cancer risk, can be identified at rather low exposure levels



Average concentrations of air pollutants measured at the stationary monitoring stations in Prachatice (PR) and Teplice (TP) during the winter months preceding the blood sampling

	SO ₂	NO _x	NO ₂	NO	CO	PM _{2.5} ¹	PM ₁₀ ¹	PAHs ^{2,3}	c- PAHs ^{2,4}
3 months									
PR	10.63	25.48	18.25	4.44	526.9	32.27	38.16	22.37	9.87
TP	17.47	66.05	35.85	19.69	662.3	30.11	37.22	88.90	18.31
6 months									
PR	12.70	25.25	18.08	4.40	607.4	36.15	41.44	19.10	8.32
TP	14.67	74.30	36.63	24.65	525.4	30.19	37.48	75.13	15.79

¹ PM concentrations in µg/m³ air, average based on 24h mean concentrations
² PAH concentrations in ng/m³ air, average based on 24h mean concentrations
³ Total of all PAHs measured
⁴ Total of eight carcinogenic PAHs: benzo[a]pyrene, chrysene, benz[a]anthracene, benzo[b]fluoranthene, benzo[k]fluoranthene, dibenz[a,h]anthracene, indeno[1,2,3-cd]pyrene and benzo[ghi]perylene



CHILD HEALTH AND THE ENVIRONMENT: RESULTS FROM EU FRAMEWORK 5

Characteristics of the study populations

	no. of subjects	gender M/F	no. of smokers	age mean (SD)	age range
Prachatic					
Very young children	12	7/5	-	6,2 (0,5)	5 - 7
Young children	12	6/6	-	9,4 (1,2)	7 - 11
Mothers	9		3	35,0 (4,5)	29 - 44
Teplice					
Very young children	12	6/6	-	6,5 (0,5)	6 - 7
Young children	11	5/6	-	9,3 (1,2)	7 - 11
Mothers	12		1	32,9 (4,1)	28 - 40

Gene expression analysis - Methods

- Blood collected in March 2004 into PAXgene vacutainers (RNA stabilization)
- RNA isolation with PAXgene Blood RNA system
- Purification and DNase treatments
- Quantity (UV absorption) and quality (BioAnalyzer) controls
 → HIGH QUALITY RNA
 ↓
- 1 µg RNA input into Agilent Fluorescent Linear Amplification system
- Generate Cy5-dCTP or Cy3-dCTP labeled cRNA
- Competitive hybridization to Agilent Human 22k oligoarrays

Gene expression – Statistical analyses

Various approaches followed:

- Analysis of differential gene expressions in children from Teplice vs children from Prachatic
 - Hierarchical clustering analysis
 - Principal Component Analysis
- Pathway analysis
- Correlation analysis of gene expression modulations with micronuclei data

Statistics:

- T-test of Teplice children vs pooled reference RNA from Prachatic children
- Pearson correlation analysis

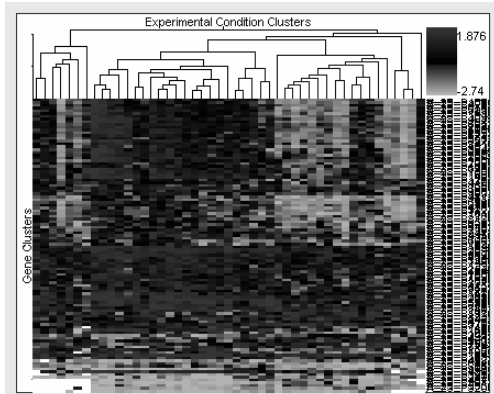
Pathway analyses for biological enrichment:

- Expression Analysis Systematic Explorer (EASE)
- GenMAPP pathway analysis tool

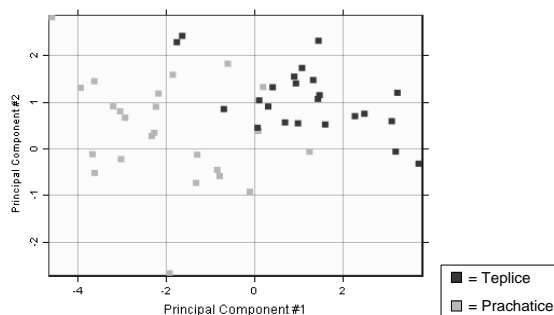
Overview of the numbers of genes significantly differentially expressed after GEPAS permutation t-test

p-value 2-tailed significance	# genes differentially expressed in Teplice versus Prachatic children		
	total	upregulated genes	downregulated genes
<0.0001	27	26	1
<0.001	95	81	14
<0.01	487	315	172
<0.05	1727	1001	726

Hierarchical clustering of genes differentially expressed between exposed (black branches) and non-exposed (blue branches) children at p < 0,001



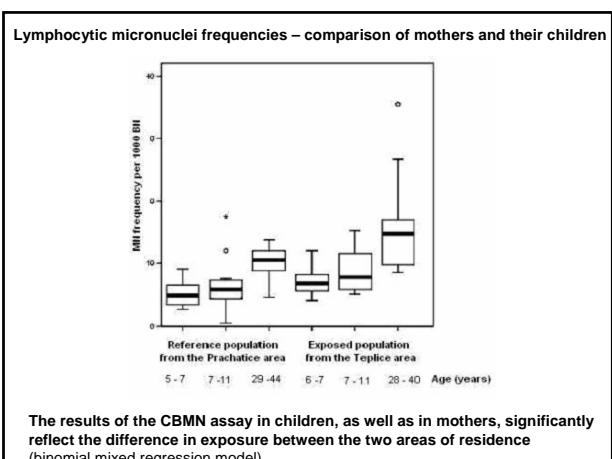
Principal Component Analysis based on the list of genes differentially expressed at p<0.01, n = 489



CHILD HEALTH AND THE ENVIRONMENT: RESULTS FROM EU FRAMEWORK 5

Significance	Regulation	Biological process (EASE score)
p<0.01	Up	Nucleosome assembly (0.0002) Chromatin assembly/disassembly (0.0004) Microtubule based movement (0.006) DNA packaging (0.014) Muscle development (0.046)
	Down	-
All	All	Nucleosome assembly (0.002) Chromatin assembly/disassembly (0.004) Microtubule based movement (0.021) DNA packaging (0.039)
	All	Nucleosome assembly (0.00001) Chromatin assembly/disassembly (0.00002) Cell communication (0.006) Innate immune response (0.015) Inflammatory response (0.016) Circulation (0.031) Muscle development (0.039)
p<0.05	Up	Nucleosome assembly (0.00001) Chromatin assembly/disassembly (0.00002) Cell communication (0.006) Innate immune response (0.015) Inflammatory response (0.016) Circulation (0.031) Muscle development (0.039)
	Down	Nucleobase, nucleoside, nucleotide and nucleic acid metabolism (0.00001) RNA splicing (0.0001) M phase (0.006) Nuclear division (0.001) Transcription (0.006) Response to DNA damage stimulus (0.014)
All	All	Nucleosome assembly (0.001) Chromatin assembly/disassembly (0.002) Circulation (0.012) M phase of mitotic cell cycle (0.022) Vitamin metabolism (0.033) Nuclear division (0.048)

Pathway analysis by EASE (p<0.01, n = 489)



EASE analysis of gene expressions that significantly correlated with the individual micronuclei frequencies

Association with micronuclei frequencies	Biological process (EASE score)
Positive (n=748)	Cell communication (0.00006) G-protein coupled receptor protein signalling pathway (0.002) Signal transduction (0.008) Response to chemical substance (0.015) Cell differentiation (0.023) MAPKKK cascade (0.040) Innate immune response (0.049)
Negative (n=706)	RNA processing ($2 \cdot 10^{-9}$) Protein biosynthesis (0.001) Exogenous antigen processing (0.001) Cell cycle checkpoint (0.034) Mitotic cell cycle (0.043)
All (n=1454)	RNA metabolism (0.001) Exogenous antigen processing (0.015) Cell differentiation (0.019) Vasoconstriction (0.033)

Comparative sensitivity

- Micronuclei frequency exposed/non-exposed is 1,38
- Gene expression level (top 5 modulated genes) exposed/non-exposed is 1,24-1,34
- Expression differences of these genes segregate exposed from non-exposed groups at considerable lower p-values than micronuclei frequencies

Conclusions

- It appears feasible to generate discriminative modulations of gene expression and of genetic pathways at rather low differences in exposure levels (of environmental carcinogens)
- To a certain degree, this correlates with an established phenotypic endpoint of genotoxicity (micronuclei frequency); by this, genetic pathways related to genotoxic risks at low level environmental exposure, may have been identified
- Concomitantly, a wealth of mechanistic information has been gathered, the full biological relevance of which has to be thoroughly explored
- Transcriptomic analysis represents a very promising tool for monitoring environmental carcinogenesis

Many thanks to

- **Maastricht University:**
 - Marcel van Herwijnen
 - Danitsja van Leeuwen
 - Yvonne Staal
 - Joost van Delft
 - Theo de Kok
 - Jacco Briedé
- **Institute of Experimental Medicine, Czech Republic:**
 - Radim Srám
- **Karolinska Institute:**
 - Edyta Bajak
- **Free University Brussels:**
 - Micheline Kirsch-Volders
- **Copenhagen University:**
 - Peter Vinzents
 - Marie Pedersen
 - Lisbeth Knudsen

with the financial support of EU FP5, QLK4-CT-2002-02198, and the Flanders Environmental Health Study Programme