PathWave 2.1: a user-friendly R package for detecting differential regulation and switches in metabolic and signaling pathways

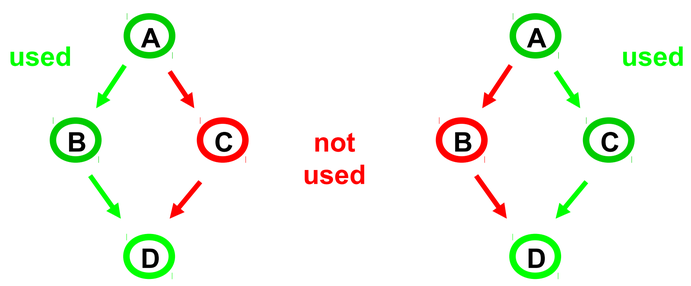
Rosario M. Piro[1,2], Stefan Wiesberg[3,4], Gunnar Schramm[1,2], Nico Rebel[3], Marcus Oswald[1,2], Moritz Aschoff[1,2], Volker Ast[1,2], Marie-Laure Yaspo[5], Gerhard Reinelt[3], Roland Eils[1,2] and Rainer König[1,2]

[1] Department of Bioinformatics and Functional Genomics, Institute of Pharmacy and Molecular Biotechnology, and Bioquant, University of Heidelberg, Im Neuenheimer Feld 267, 69120 Heidelberg,
[2] Department of Theoretical Bioinformatics, German Cancer Research Center (DKFZ), Im Neuenheimer Feld 280, 69120 Heidelberg,
[3] Institute of Computer Science, University of Heidelberg, 69120 Heidelberg, Germany, [4] Interdisciplinary Center for Scientific Computing, University of Heidelberg, 69120 Heidelberg,
[5] Max Panck Institute (MPI) for Molecular Genetics, Ihnestraße 63-73, 14195 Berlin, Germany.

Gene expression profiling by microarrays or transcript sequencing (RNA-Seq) enables observing the pathogenic function of cells on a mesoscopic level. We present a new version of PathWave, a tool that explicitly takes the topology of metabolic networks into account to identify both global and localized (switch-like) regulatory shifts in metabolic pathways. For this purpose, PathWave applies adjusted wavelet transforms on gene expression data after mapping it to optimized 2D grid representations of curated pathways from KEGG or BiGG. Inaddition, the method can also be applied to signaling pathways.
PathWave 2.1 has several improvements including an improved construction of 2D pathway grids for the application of wavelet transforms, a more flexible and user-friendly interface and pre-arranged 2D grid representations for *H. sapiens*, *M. musculus*, *D. melanogaster*, *D. rerio*, *C. elegans*, and *E. coli*. Here, we present PathWave's basic workflow and briefly discuss the improvements and current applications.

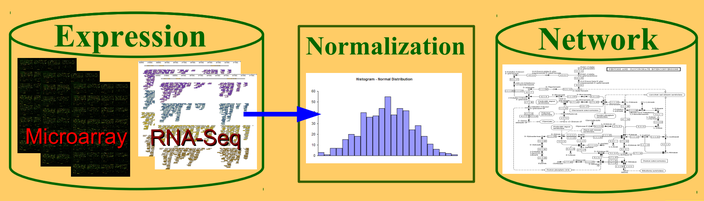
Goal: Finding crucial players in cellular networks for specifically targeting pathogenic cells

normal cell in the human body | pathogenic cell (malignant, virus infected, parasitic)



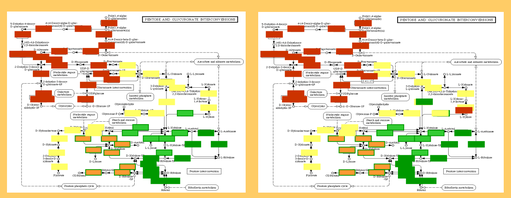
=> targeting C will specifically harm the pathogenic cell

Workflow for defining the relevant pathways

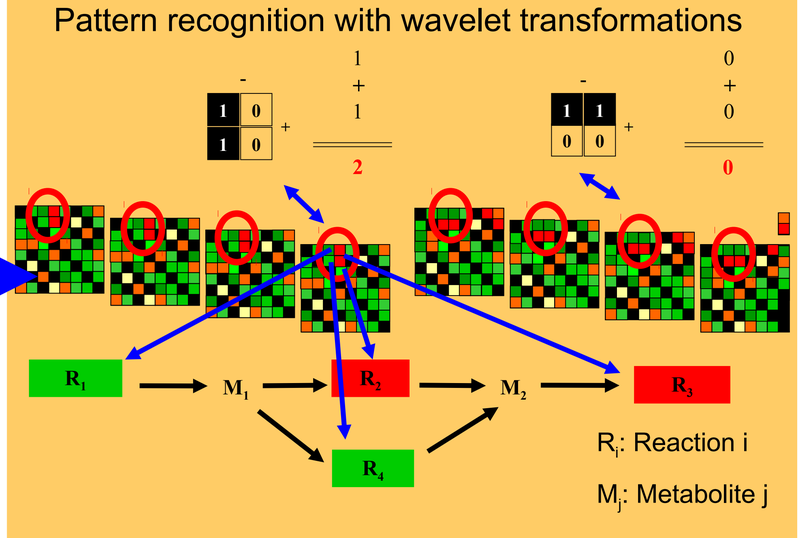


- mapping of expression data on a cellular network

=> extracting coherences and incoherences within the network



Pattern recognition with wavelet transformations



Improvements of PathWave 2.1

R-package with a simple user interface

```
pwres <- pathWave.run(preprocessed.tag="KEGG.hsa",
    input.sampleclasses="sample_classes.tsv",
    input.exprdata="expr_data.tsv", param.numperm=10000)
```

New, improved optimization of 2D pathway maps

up to 10^3 times faster algorithm to map pathways to optimally arranged 2D rappresentations (needed for the application of wavelet transforms)

Extraction of metabolic pathways from SBML (BiGG)

Identification of differentially regulated pathways not only for ***KEGG*** (Kanehisa et al., 2008), but also for pathways extracted from ***BiGG*** (Biochemical Genetics and Genomics knowledgebase of reconstructed metabolic models; Schellenberger et al, 2010). This has already been applied to Alzheimer's Disease (Lewis et al, 2010)



More pre-processed 2D grid arrangements of pathways for immediate application

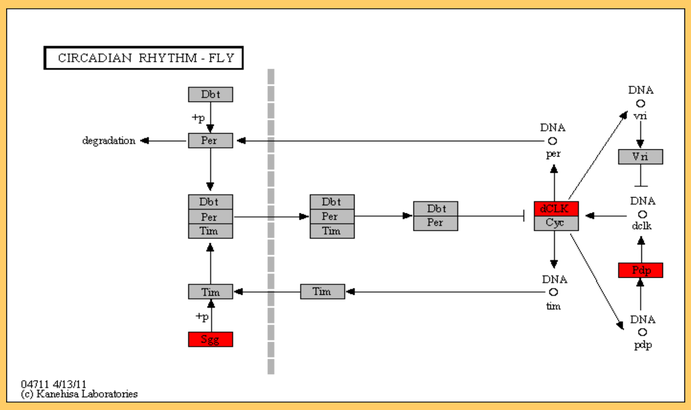
The R-package comes with two sets of arrangements for H. sapiens (KEGG & BiGG pathways) as well as arrangements (KEGG only) for five additional species: M. musculus, D. melanogaster, D. rerio, C. elegans and E. coli. For other species, it provides the necessary tools to produce own 2D grid arrangements.

Examples for the application of PathWave to
- **different species**
- **different types of expression data**
- **metabolic and signaling pathways**

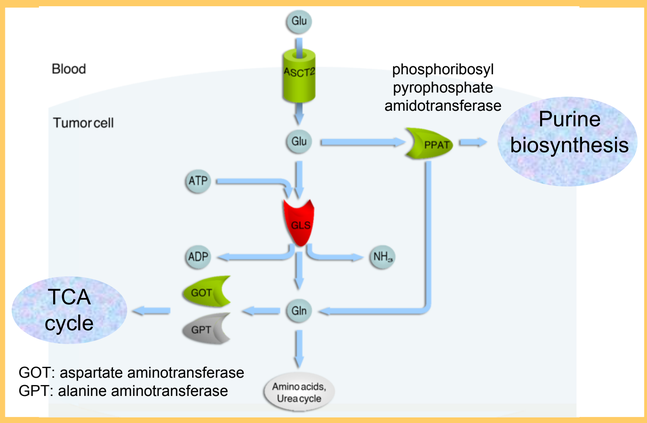
Example: pathways in aging of Drosophila

Microarray data (Doroszuk et al, 2012) from long-lived flies (8 samples) versus normal specimens (8 samples)

KEGG ID	Pathway name	*P*-value	r.up	r.no_ch	r.down
dme00500	Starch and sucrose metabolism	3.35e-03	6	11	1
dme04711	Circadian rhythm - fly	6.62e-03	3	5	0
dme00564	Glycerophospholipid metabolism	1.41e-02	9	25	3
dme00350	Tyrosine metabolism	1.42e-02	4	5	0
dme00760	Nicotinate and nicotinamide metabolism	3.00e-02	3	6	0
dme00130	Ubiquinone and other terpenoid-quinone biosynthesis	3.37e-02	7	4	0
dme04630	Jak-STAT signaling pathway	3.70e-02	1	11	2
dme00260	Glycine, serine and threonine metabolism	4.13e-02	4	14	3



Example: Switch-like regulation pattern in glutamate metabolism in neuroblastoma (Schramm et al., 2010)

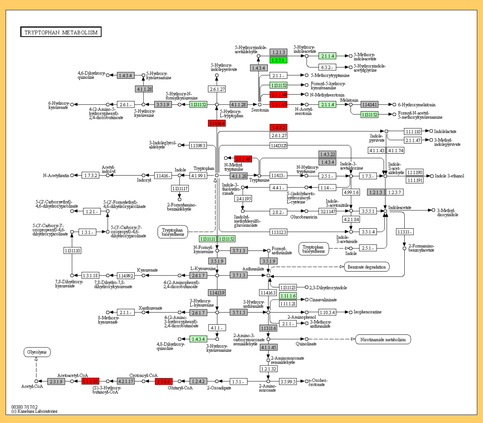


Example: pathways in aggressive medulloblastoma

RNA-Seq data (unpublished) from medulloblastomas (4 WNT; 16 SHH; 11 Grp3; 13 Grp4) and normal control tissue (7 samples)

Two metabolic pathways from KEGG show a significant differential regulatation for **Grp3 vs control** but not for all tumors vs control:

Pathway	*P*	up	-	down
Tryptophan metabolism	3.22E-03	7	24	1
N-Glycan biosynthesis	1.77E-02	6	26	1



Conclusions.

We presented the R-package PathWave 2.1, and discussed its major improvements (user-interface, improved pre-processing of KEGG and BiGG pathway maps, pre-processed pathways for six species).
Examples for the application of PathWave to different biological problems regarding different species and gene expression data types demonstrate the usefulness of our approach to pathway analysis.

References

Doroszuk A, et al. (2012) Transcriptome analysis of a long-lived natural Drosophila variant: a prominent role of stress- and reproduction-genes in lifespan extension. ***BMC Genomics*** 13:167.

Kanehisa M, Araki M, Goto S, Hattori M, Hirakawa M, Itoh M, Katayama T, Kawashima S, Okuda S, Tokimatsu T, Yamanishi Y (2008) KEGG for linking genomes to life and the environment. ***Nucleic Acids Res*** 36:D480-484.

Lewis NE, Schramm G, Bordbar A, Schellenberger J, Andersen MP, Cheng JK, Patel N, Yee A, Lewis RA, Eils R, König R, Palsson BØ (2010) Large-scale in silico modeling of metabolic interactions between cell types in the human brain. ***Nat Biotechnol*** 28:1279-1285.

Schellenberger J, Park JO, Conrad TM, Palsson BØ (2010) BiGG: a Biochemical Genetic and Genomic knowledgebase of large scale metabolic reconstructions. ***BMC Bioinformatics*** 11:213.

Schramm G, Wiesberg S, Diessl N, Kranz AL, Sagulenko V, Oswald M, Reinelt G, Westermann F, Eils R, König R (2010) PathWave: discovering patterns of differentially regulated enzymes in metabolic pathways. ***Bioinformatics*** 26:1225-31.