

Geneset Analysis
**Gene Expression Analysis by staring
at colorfull pictures**
Organizing Maps

Henry Wirth

Interdisciplinary Center for Bioinformatics,
University of Leipzig

Department of Proteomics,
UFZ - Helmholtz-Centre for Environmental Research



Motivation

Gene Expression Analysis:

single gene DE-statistic (t-Test, Anova)

single gene based biological interpretation

	A	B	C	D	E	F
1	Class	delta.tWAD	FC.aIPs.afibr	FC.aIPs.hES	Affy.ID	Gene.Symbols
2	+	14,4379873	1,47987545	-0,18801006	8066925	PTGIS
3	+	12,4303671	1,82965779	-0,11120772	8146921	RDH10
4	+	12,2460534	2,97734209	0,38989647	7967993	FGF9
5	+	11,9033784	2,77767178	-0,29958358	7930482	ADRA2A
6	+	11,2552755	2,86753616	0,4807414	8151369	C8orf84
7	+	10,5323491	2,50643731	0,14669466	7957140	LGR5
8	+	10,1987229	2,39073257	-0,05214042	8022283	FAM38B
9	+	10,186058	2,37575859	0,1277323	8022295	FAM38B2
10	+	10,1587665	0,76191745	0,02819293	7991234	MFG8
11	+	10,1249098	2,39782536	-0,36944771	7923386	LMOD1
12	+	10,0791649	2,42786928	-0,23912324	8140840	STEAP4
13	+	9,92513288	2,22832905	0,17208639	7985786	ACAN
14	+	9,92359538	2,02232157	0,21588585	8140686	SEMA3D
15	+	9,7118542	2,05636983	-0,16539922	7951351	PDGFD
16	+	9,7065094	2,50498518	-0,28916409	7932254	ITGA8
17	+	9,67205438	2,50411111	0,12858748	7907271	FMO2
18	+	9,5343254	2,27045901	-0,49714046	7960919	MFAP5
19	+	9,48728326	2,39916739	-0,31221307	8136200	CPA4
20	+	9,18532943	2,27093189	0,17253479	8091537	IGSF10

UniProtKB

Search in: Protein Knowledgebase (UniProtKB) | Query: [] | Search | Clear | Fields »

★ Reviewed, UniProtKB/Swiss-Prot **P27361** (MK03_HUMAN)
Last modified February 9, 2010. Version 120. History...

Names and origin

Protein names	Recommended name:
	Mitogen-activated protein kinase 3
	Short name=MAP kinase 3
	Short name=MAPK 3
	EC=2.7.11.24
	Alternative name(s):
	Extracellular signal-regulated kinase 1
	Short name=ERK-1
	Insulin-stimulated MAP2 kinase
	Mitogen-activated protein kinase 1
	Short name=MAP kinase 1
	Short name=MAPK 1
	p44-ERK1
	ERT2
	MAP kinase isoform p44
	Short name=p44-MAPK
	Microtubule-associated protein 2 kinase

Gene names

Name:	Synonyms:
MAPK3	ERK1, PRKM3

Motivation

Gene Expression Analysis:

single gene DE-statistic (t-Test, Anova)

→ easy and intuitively way of representation

single gene based biological interpretation

→ statistic for defined genesets with well-known meaning (GO:BP)

Representation and Geneset Analysis

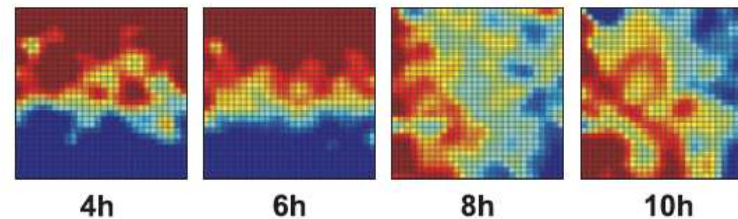
Gene Expression Fingerprinting

Visualisation Tool for Gene Expression Development: “GEDI”

*“Gene Expression Dynamics Inspector (GEDI):
for integrative analysis of expression profiles”*
Gabriel S. Eichler, Sui Huang and Donald E. Ingber

*“Cell Fates as High-Dimensional Attractor
States of a Complex Gene Regulatory Network”*
Sui Huang et al.

Input Data
(2000 Gene
Expressions)



Gene Expression Fingerprinting

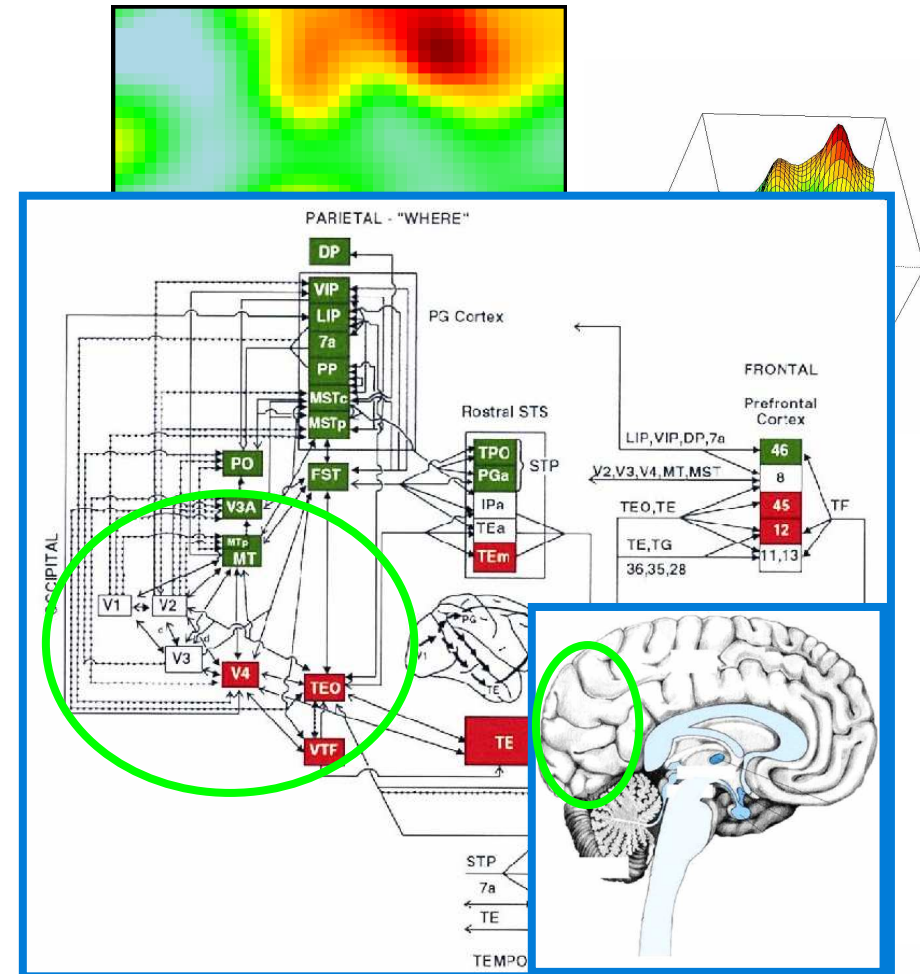
Idea:

Input Data



Self Organizing Map

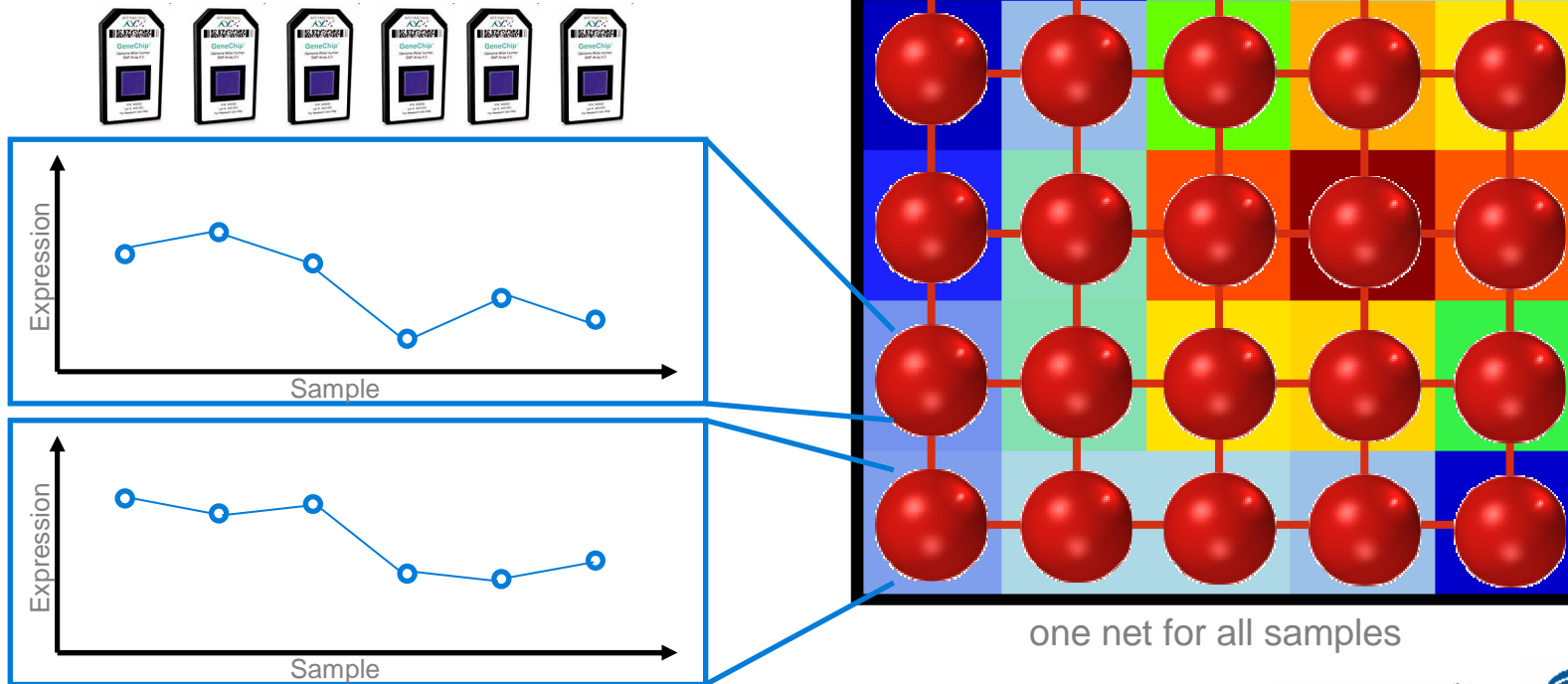
2 dimensional artificial neural net



Gene Expression Fingerprinting

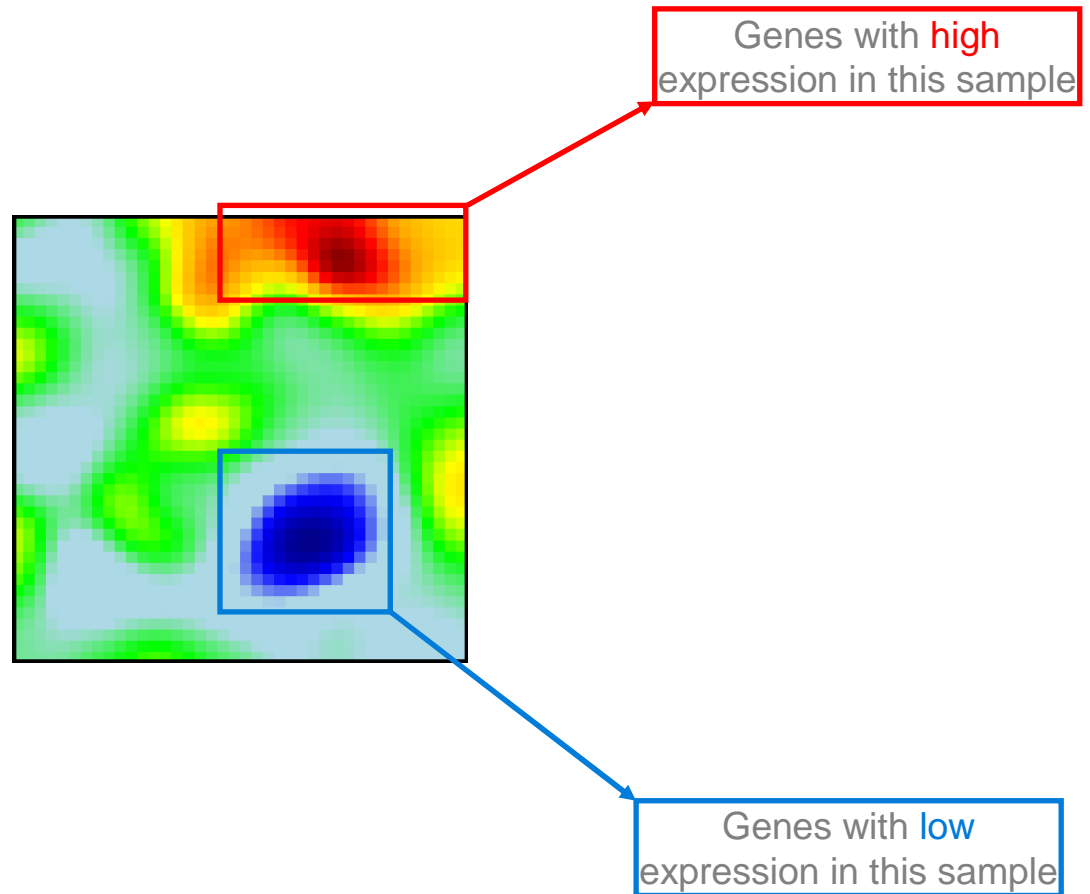
Self Organizing Maps – expression data to mosaic picture

! neighbourhood similarity !



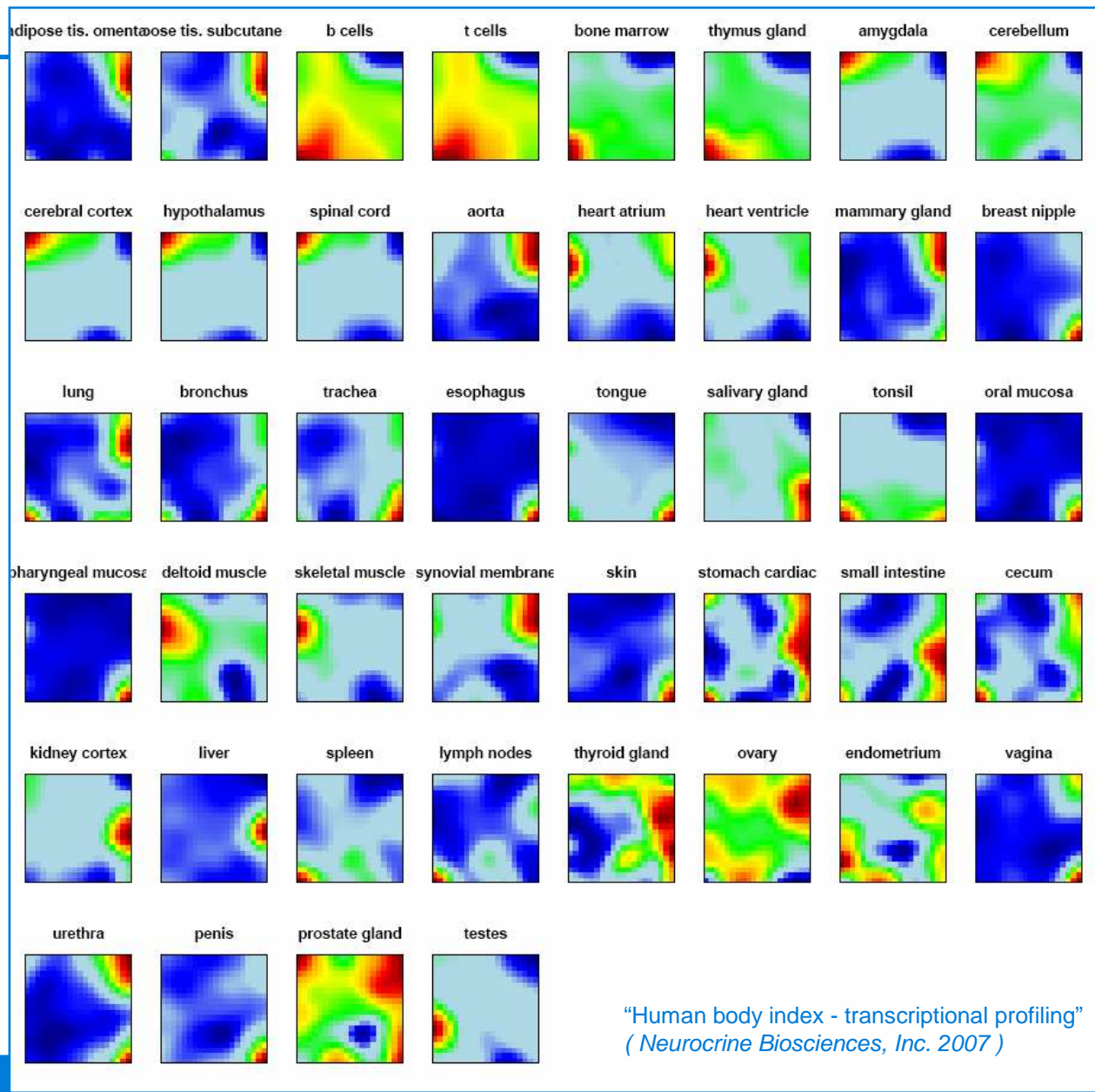
Gene Expression Fingerprinting

Genesets?



Gene Expression Fingerprinting

44 tissues



“Human body index - transcriptional profiling”
 (Neurocrine Biosciences, Inc. 2007)

Geneset Analysis

Geneset Analysis

Geneset Collections (Broad Institute, MIT: "GSEA")

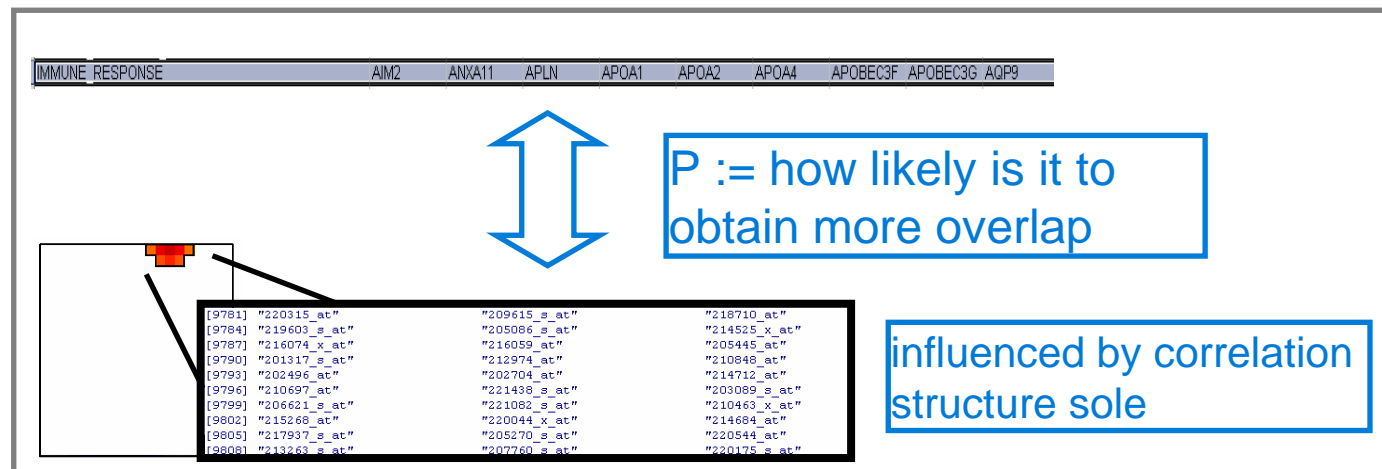
positional geneset lists, curated genesets, GO genesets (biological process, molecular function, cellular component)...

A	B	C	D	E	F	G	H	I	J	K	L
251 GROWTH	Genes annot:ACVR1B	ALOX12	ALOX15B	APBB1	APBB2	APOA5	BBC3	BCAR1	BCL6	BIN3	
252 HEART_DEVELOPMENT	Genes annot:ALPK3	BMP10	BMP2	CASQ2	CBY1	CENTA2	CITED2	DVL1	DVL3	ECE2	
253 HEME_BIOSYNTHETIC_PROCESS	Genes annot:ALAD	ALAS1	ALAS2	COX10	COX15	CPOX	FECH	NFE2L1	PPOX	TSPO	
254 HEME_METABOLIC_PROCESS	Genes annot:ALAD	ALAS1	ALAS2	BLVRA	COX10	COX15	CPOX	FECH	NFE2L1	PPOX	
255 HEMOPOIESIS	Genes annot:ACIN1	ACVR1B	ACVR2A	ALAS2	BLNK	CALCA	CARTPT	CD1D	CD3D	CD4	
256 HEMOPOIETIC_OR_LYMPHOID_ORGAN_DEVELOPMENT	Genes annot:ACIN1	ACVR1B	ACVR2A	ALAS2	BLNK	CALCA	CARTPT	CD1D	CD3D	CD4	
257 HEMOSTASIS	Genes annot:ADORA2A	AVPR2	C4BPB	CD36	CD40LG	CD69	ENTPD1	F10	F12	F13A1	
258 HETEROCYCLE_METABOLIC_PROCESS	Genes annot:ALAD	ALAS1	ALAS2	ALDH1L1	ALDH6A1	ASMTL	BLVRA	CDA	COX10	COX15	
259 HETEROPHILIC_CELL_ADHESION	Genes annot:AMIGO1	AMIGO2	AMIGO3	CADM1	CADM3	CD164	CD209	LGALS7	PVRL1	REG3A	
260 HISTONE_MODIFICATION	Genes annot:CARM1	CREBBP	EHMT1	HDAC10	HDAC11	HDAC6	HTATIP	HUWE1	JMJD2A	MAP3K12	
261 HOMEOSTASIS_OF_NUMBER_OF_CELLS	Genes annot:ACIN1	ACVR1B	ACVR2A	AKT1	ALAS2	BCL10	CDK6	CEBPG	DYRK3	ETS1	
262 HOMEOSTATIC_PROCESS	Genes annot:ABCA1	ABCA2	ABCG1	ACIN1	ACVR1B	ACVR2A	AGTR1	AIFM3	AKR1C1	AKT1	
263 HOMOPHILIC_CELL_ADHESION	Genes annot:AMIGO1	AMIGO2	CADM1	CADM3	CD84	CDH5	ITGB1	MPZL2	NPTN	PKD1	
264 HORMONE_METABOLIC_PROCESS	Genes annot:ADM	AKR1C4	AKR1D1	ALDH8A1	ALDH9A1	ASMTL	BMP6	CHST8	CHST9	CYP11A1	
265 HORMONE_SECRETION	Genes annot:CARTPT	FAM3B	FAM3D	GCK	GHRH	GHRL	GHSR	IL11	INHA	INHBA	
266 HUMORAL_IMMUNE_RESPONSE	Genes annot:BCGF1	BCL2	BLNK	BST1	BST2	C2	CCL2	CCR6	CD28	CD83	
267 I_KAPPA_B_KINASE_NF_KAPPA_B_CASCADE	Genes annot:APOL3	ATP2C1	BCL10	BCL3	BIRC2	BST2	C1orf166	CANT1	CARD10	CARD8	
268 ICOSANOID_METABOLIC_PROCESS	Genes annot:AKR1C2	AKR1C3	CD74	CYP2J2	CYP4F2	CYP4F3	CYP4F8	FADS1	HPGD	LTA4H	
269 IMMUNE_EFFECTOR_PROCESS	Genes annot:APOA1	APOA2	APOBEC3F	APOBEC3G	BCL2	BNIP3	BNIP3L	C2	CADM1	CD40LG	
270 IMMUNE_RESPONSE	Genes annot:AIM2	ANXA11	APLN	APOA1	APOA2	APOA4	APOBEC3F	APOBEC3G	AQP9	ARHGDI3	
271 IMMUNE_SYSTEM_DEVELOPMENT	Genes annot:ACIN1	ACVR1B	ACVR2A	ALAS2	BLNK	CALCA	CARTPT	CD1D	CD3D	CD4	
272 IMMUNE_SYSTEM_PROCESS	Genes annot:ACIN1	ACVR1B	ACVR2A	AIM2	AKT1	ALAS2	ANXA11	APLN	APOA1	APOA2	
273 INACTIVATION_OF_MAPK_ACTIVITY	Genes annot:DUSP16	DUSP2	DUSP22	DUSP6	DUSP8	DUSP9	GPS1	GPS2	LAX1	MBIP	
274 INDUCTION_OF_APOPTOSIS_BY_EXTRACELLULAR_SIGNALS	Genes annot:ADORA1	BAX	BID	BTK	CASP8AP2	CD38	CFLAR	CRADD	DAP	DAP3	
275 INDUCTION_OF_APOPTOSIS_BY_INTRACELLULAR_SIGNALS	Genes annot:ABL1	AIFM1	BAX	BCL3	BRCA1	C16orf5	CDKN1A	CHEK2	CIDEA	CIDEB	
276 INFLAMMATORY_RESPONSE	Genes annot:ABCF1	ADORA1	ADORA2A	ADORA3	AFAP1L2	AGER	AHSG	AIF1	ALOX15	ALOX5AP	
277 INNATE_IMMUNE_RESPONSE	Genes annot:APOA4	APOBEC3F	APOBEC3G	BCL10	CADM1	CD1D	CEBPG	COLEC12	CRTAM	DEFB1	
278 INORGANIC_ANION_TRANSPORT	Genes annot:BEST1	CLCN4	CLIC1	CLIC3	CLIC5	FXYD1	FXYD3	GLRA1	NMUR1	SLC13A4	
279 INSULIN_RECEPTOR_SIGNALING_PATHWAY	Genes annot:AKT1	AP3S1	BAIAP2	BCAR1	FOXC2	FOXO4	GAB1	GRB10	GRB2	IGF1R	
280 INTERACTION_WITH_HOST	Genes annot:ACE2	ALB	C9	CD209	CD81	CLEC4M	DERL1	HS3ST5	HS3ST6	ITCH	
281 INTERCELLULAR_JUNCTION_ASSEMBLY	Genes annot:CD9	GJA1	GJA4	GJA5	GJC1	GJD3	PARDB6	TJP1	TLN1	TLN2	
282 INTERCELLULAR_JUNCTION_ASSEMBLY_AND_MAINTENANCE	Genes annot:CD9	GJA1	GJA4	GJA5	GJC1	GJD3	PARDB6A	PARDB6B	PRKCI	TJP1	
283 INTERFERON_GAMMA_BIOSYNTHETIC_PROCESS	Genes annot:CD276	CEBPG	EBI3	IL12B	IL18	IL27	INHA	INHBA	TLR3	TLR7	
284 INTERFERON_GAMMA_PRODUCTION	Genes annot:CD276	CEBPG	EBI3	FOXP3	IL12A	IL12B	IL18	IL27	INHA	INHBA	

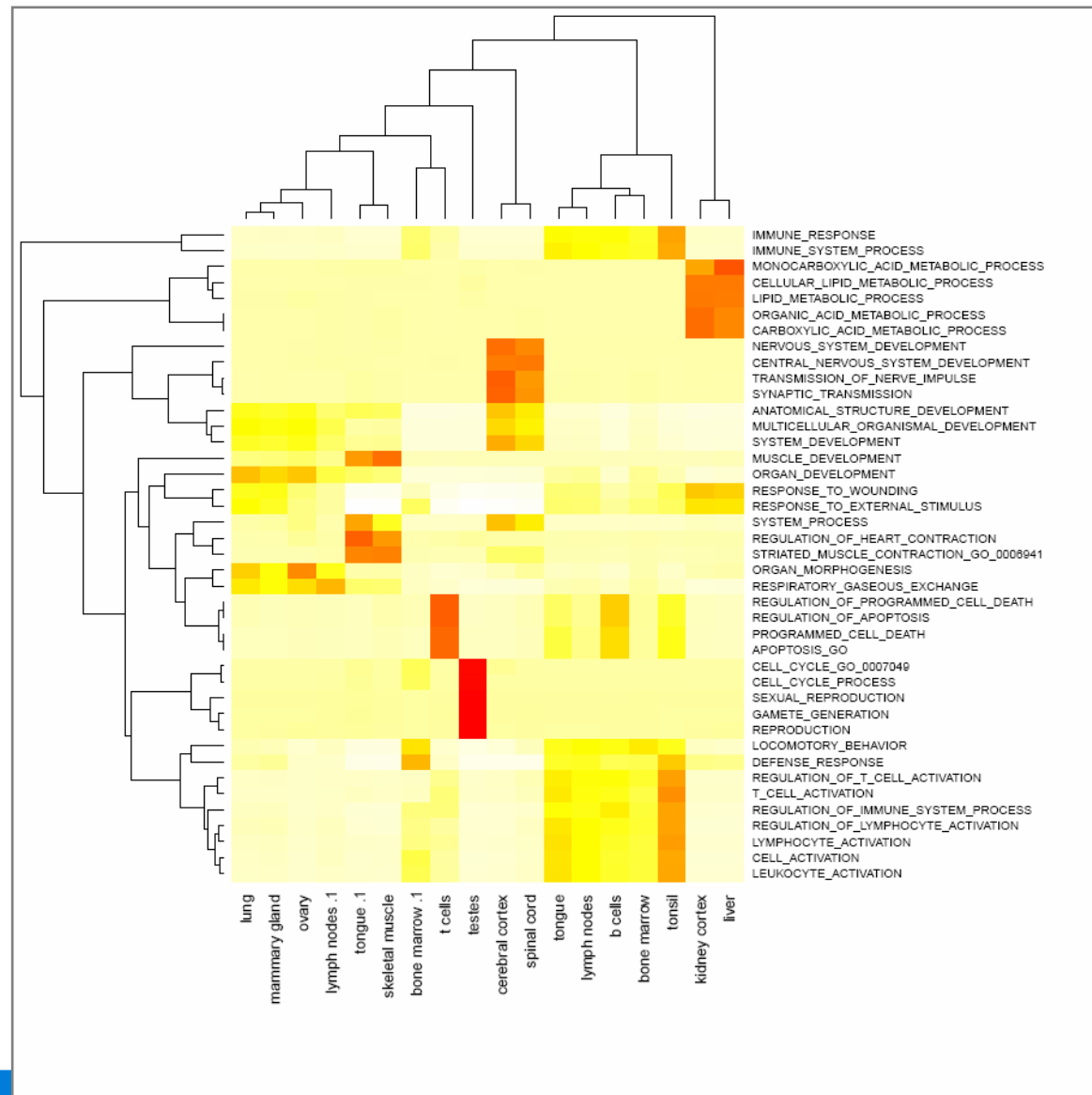
Geneset Analysis

Geneset Analysis

Hypergeometric distribution



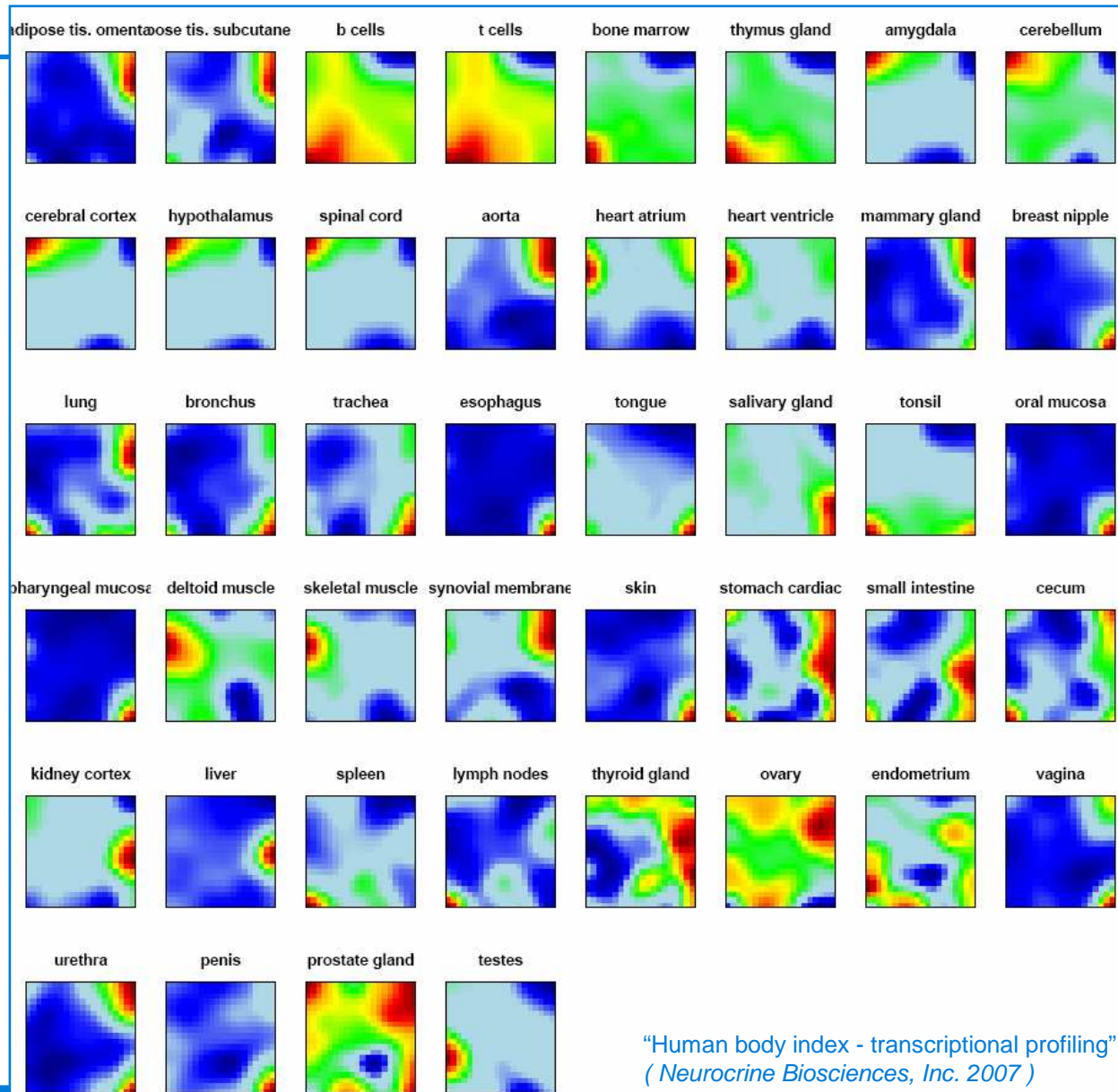
Geneset Analysis



Tissue Cartography

Cartography

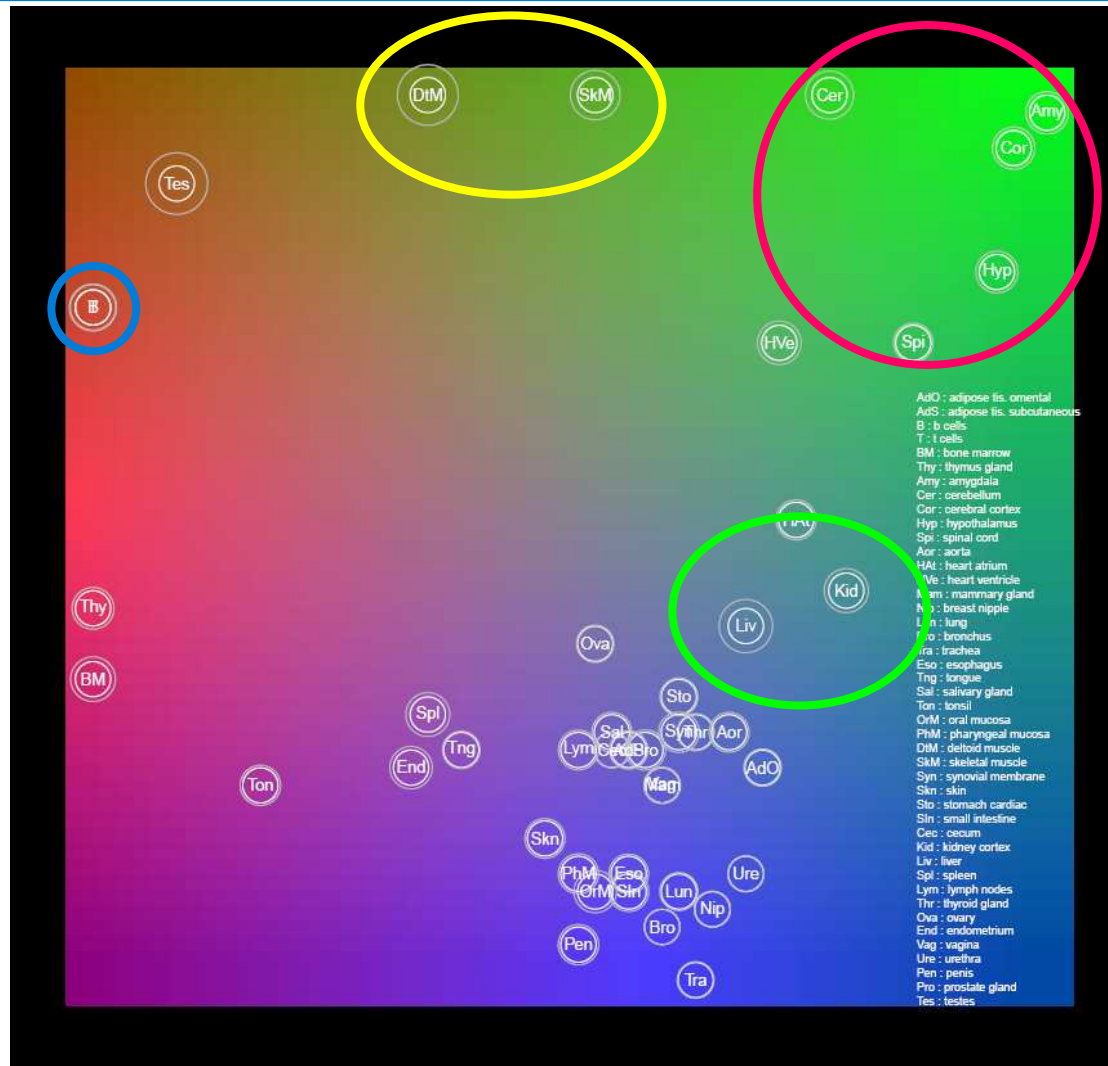
44 tissues



“Human body index - transcriptional profiling”
 (Neurocrine Biosciences, Inc. 2007)

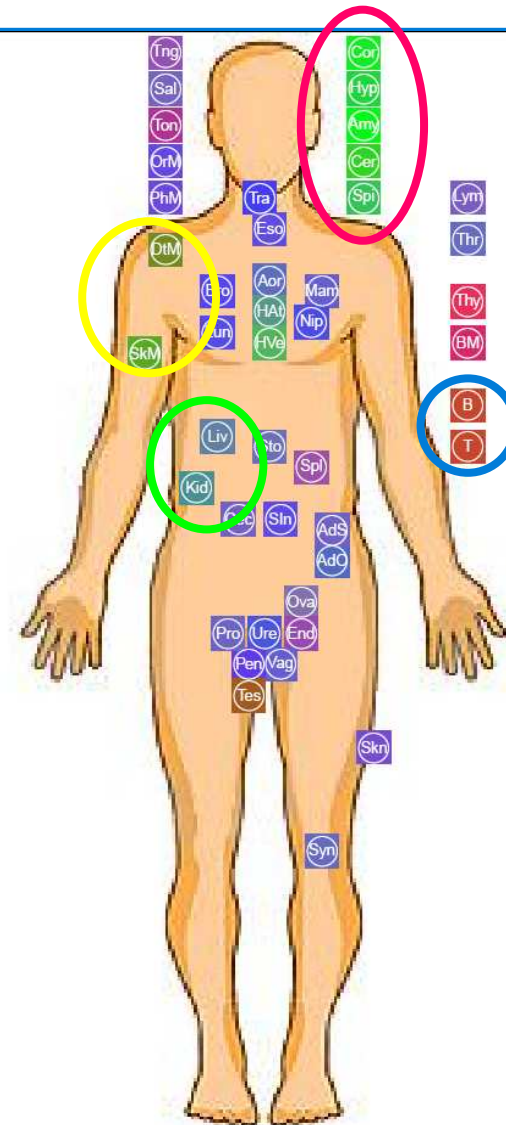
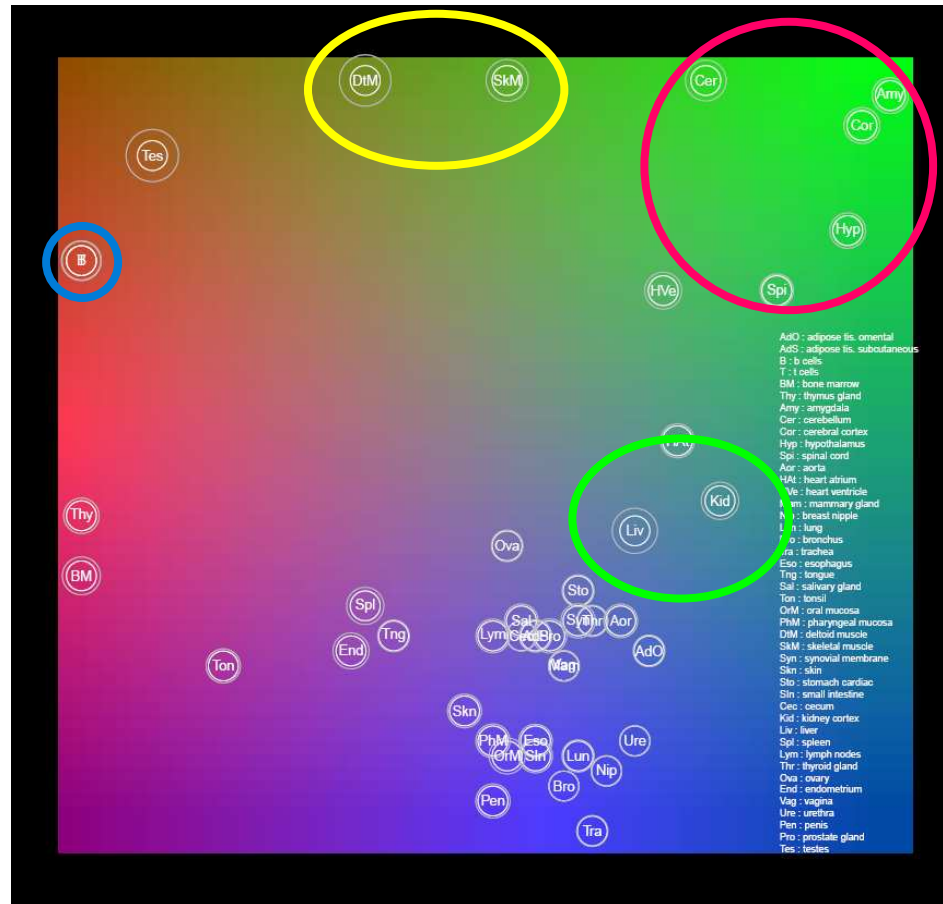
Cartography

44 tissues



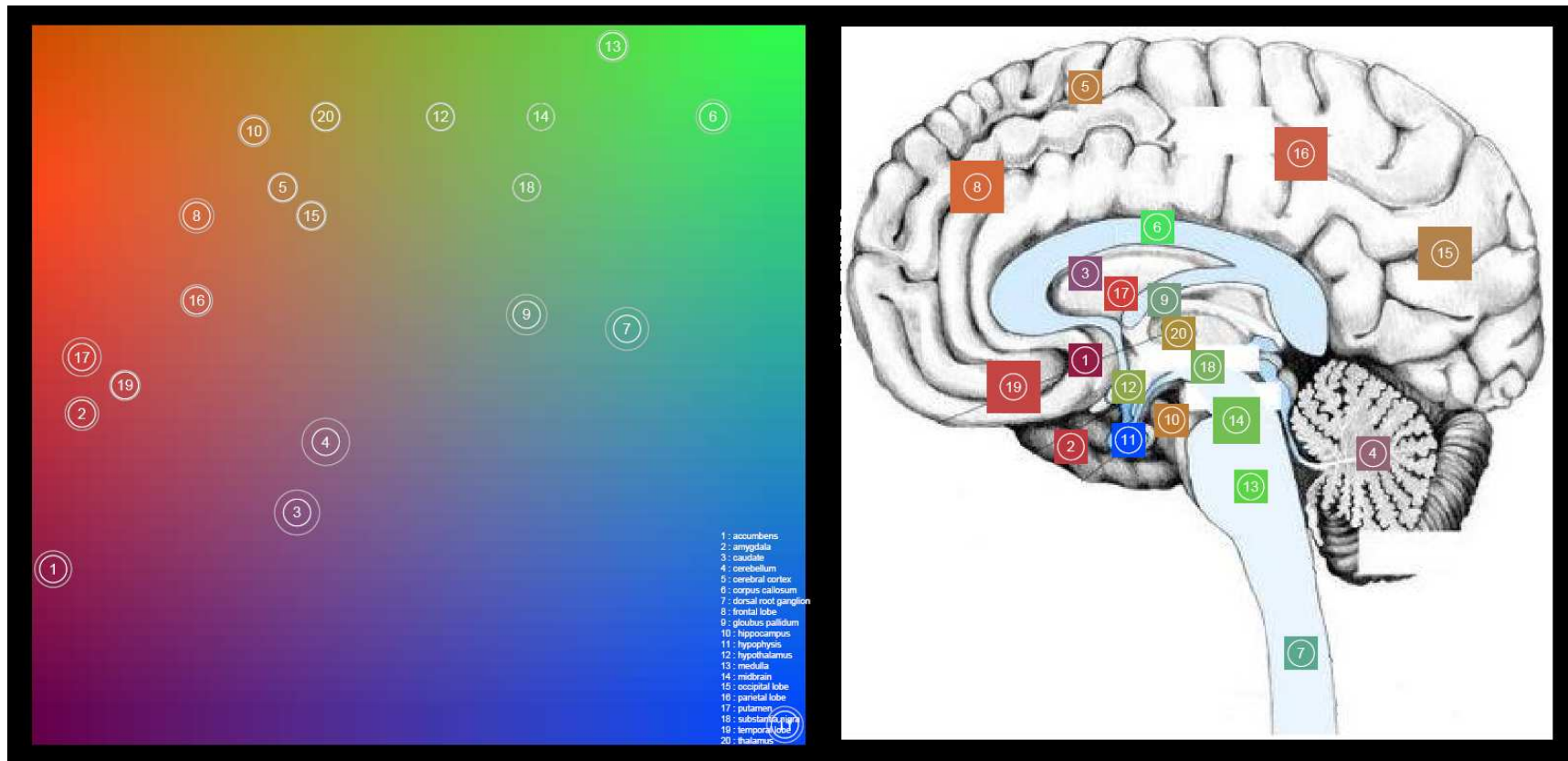
Cartography

44 tissues

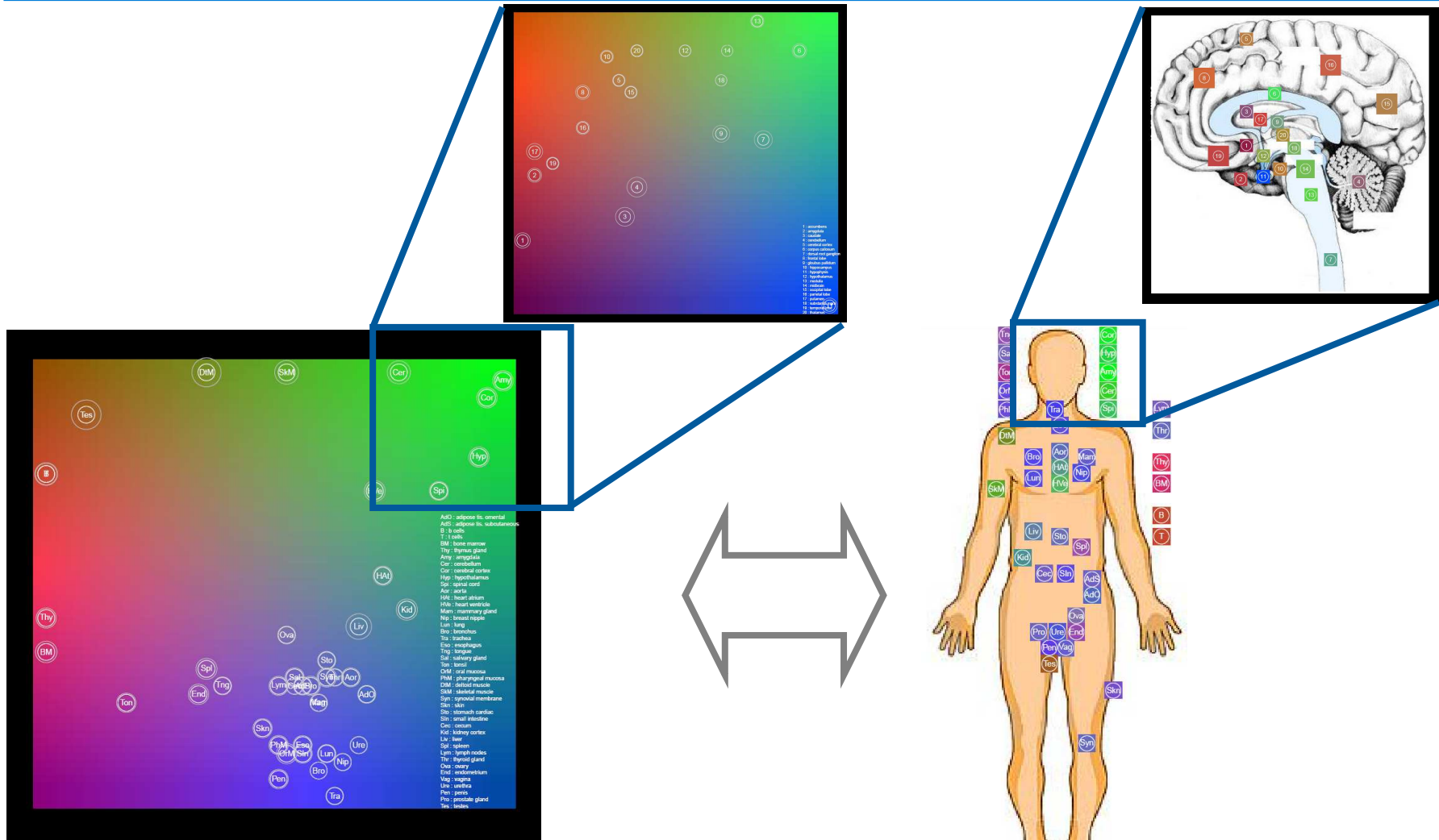


Cartography

20 brain tissues

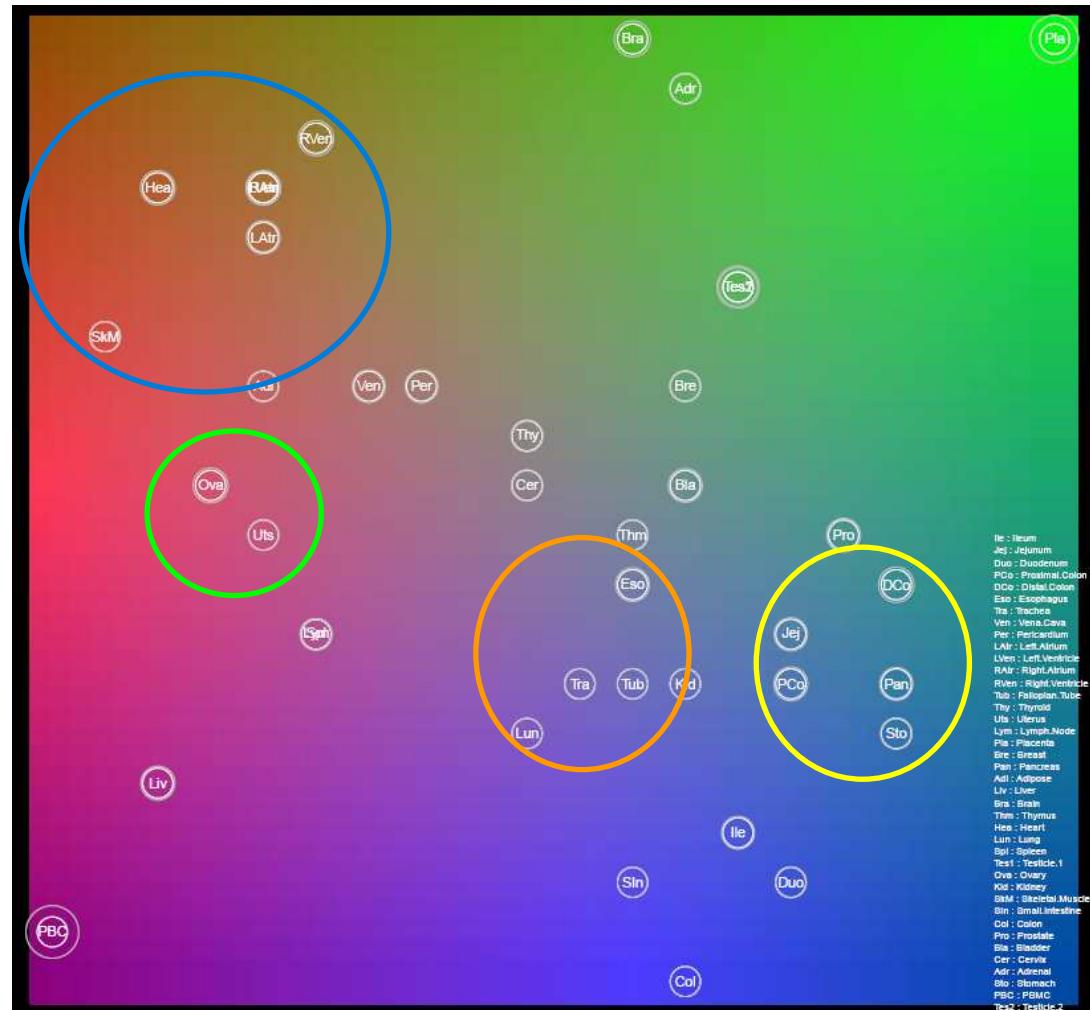
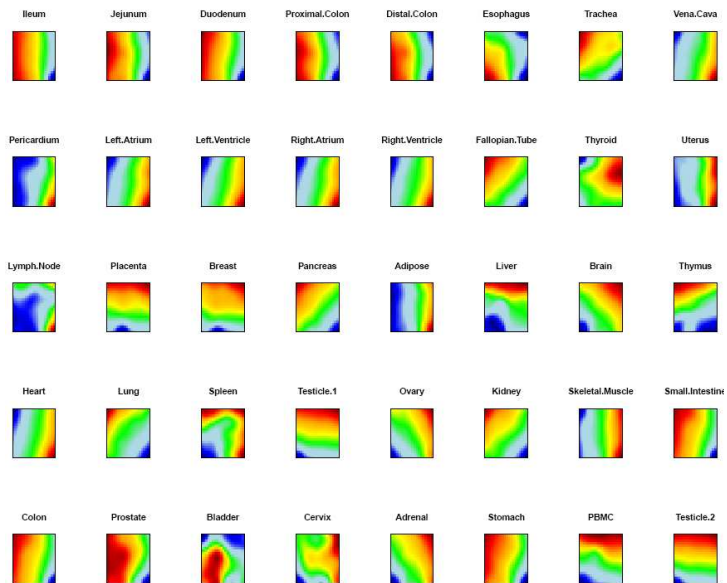


Cartography



Cartography

Adaption to miRNA expression data (Liang et al. 2007)



Thanks to

- PD Dr. Hans Binder



- Prof. Dr. Werner Dilger



- Helmholtz Initiative and Networking Fund

