Toxicology Model Building Based On Omic data sets

Jürgen Borlak
borlak.juergen@mh-hannover.de
Centre for Pharmacology and Toxicology, Hannover Medical School, Germany

Aims of my presentation

- Oncogenomics applied to an EGF transgenic disease models for an improved understanding of non-genotoxic carcino-genicity .
- Model building based on computational biology approaches

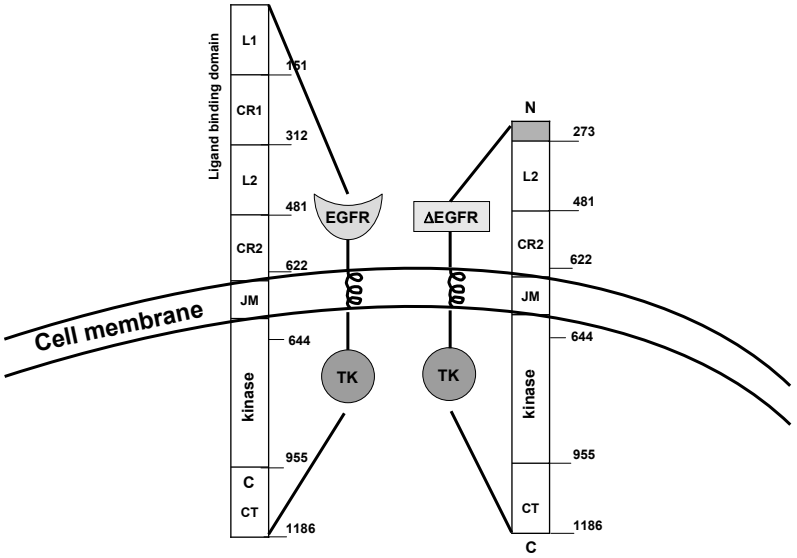
A selection of EGF published studies

- Advanced computational biology methods identify molecular switches for malignancy in an EGF mouse model of liver cancer (PLoS One 2011)
- Combined microPET/CT for imaging of hepatocellular carcinoma in mice (Front Biosci 2009)
- EPS15R, TASP1 and PRPF3 are novel disease candidate genes targeted by HNF4alpha splice variants in hepatocellular carcinoma (Gastroenterology 2008)
- Mapping of the serum proteome of hepatocellular carcinoma induced by targeted overexpression of epidermal growth factor to liver cells of transgenic mice (J Proteome Res 2008)
- Quantitative mass spectrometry to investigate epidermal growth factor receptor phosphorylation dynamics (Mass Spectrom Rev 2008)
- Epidermal growth factor-induced hepatocelluar carcinoma: gene expression profiles in precursor lesions, early stage and solitary tumours (Oncogene 2005)

Rational for investigating the EGF receptor tyrosine kinase

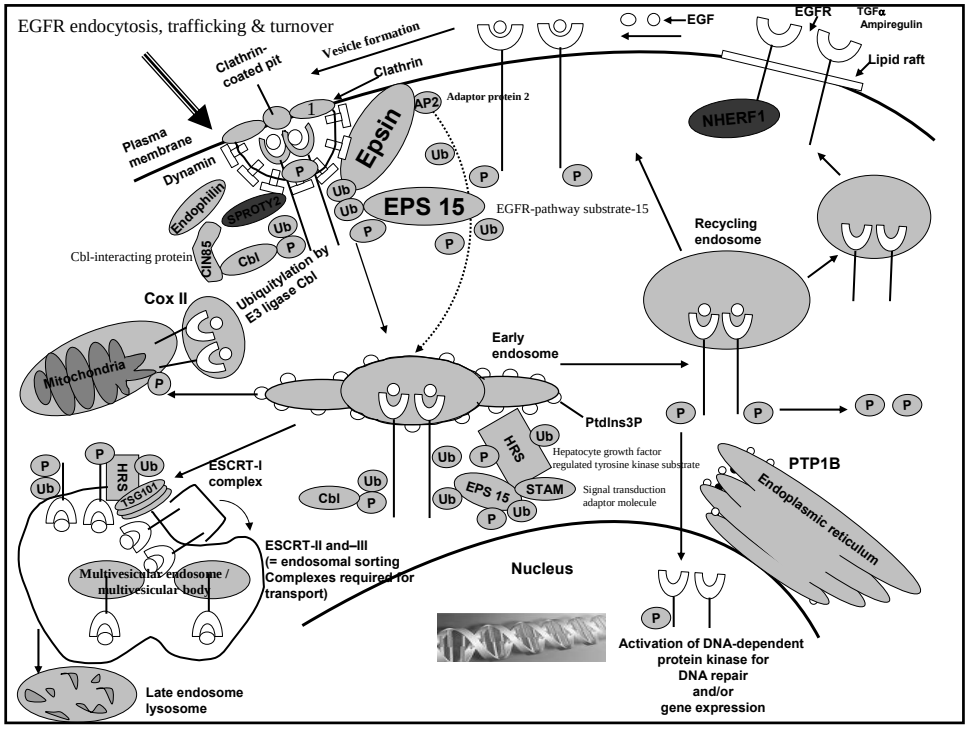
- **Abberant EGFR signaling is common to many cancers**
- **EGFR mutations are also identified in a variety of human cancers**
- **A number of agents targeting EGFR (inhibitory AB or small molecules against kinase activity) are in clinical trails or have been approved**

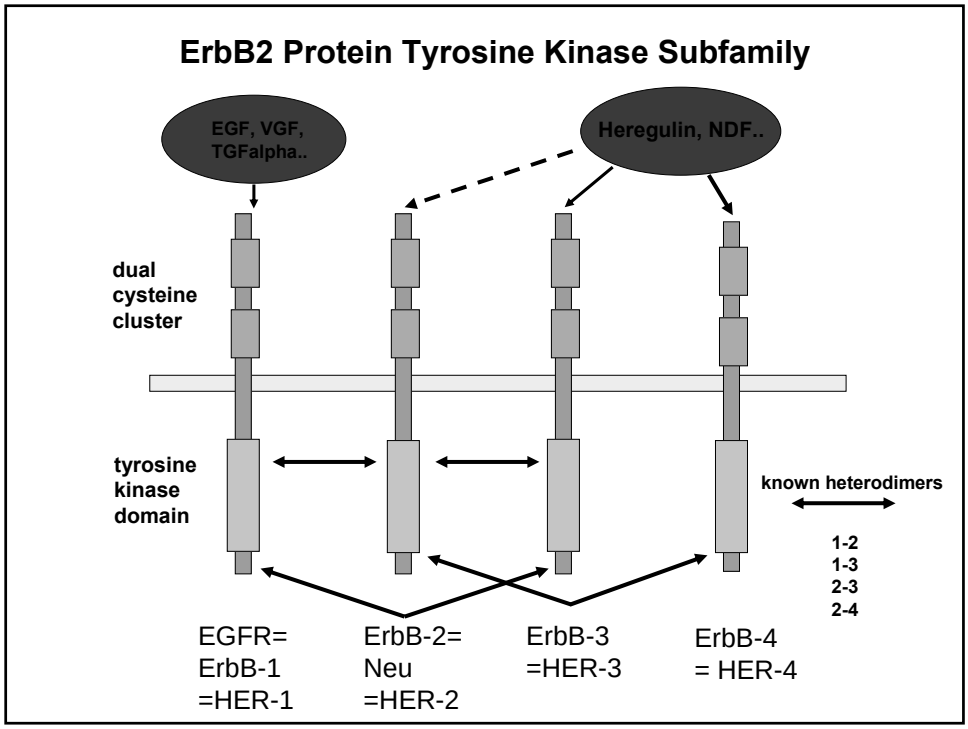
ΔEGFR in glioblastoma; loss of exon 2-7 results in 807 base pair deletion in which amino acids 6-273 are replaced by a single glycin residue. Codes for a 145 kDa EGFR glycoprotein with ligand independent tyrosine kinase activity

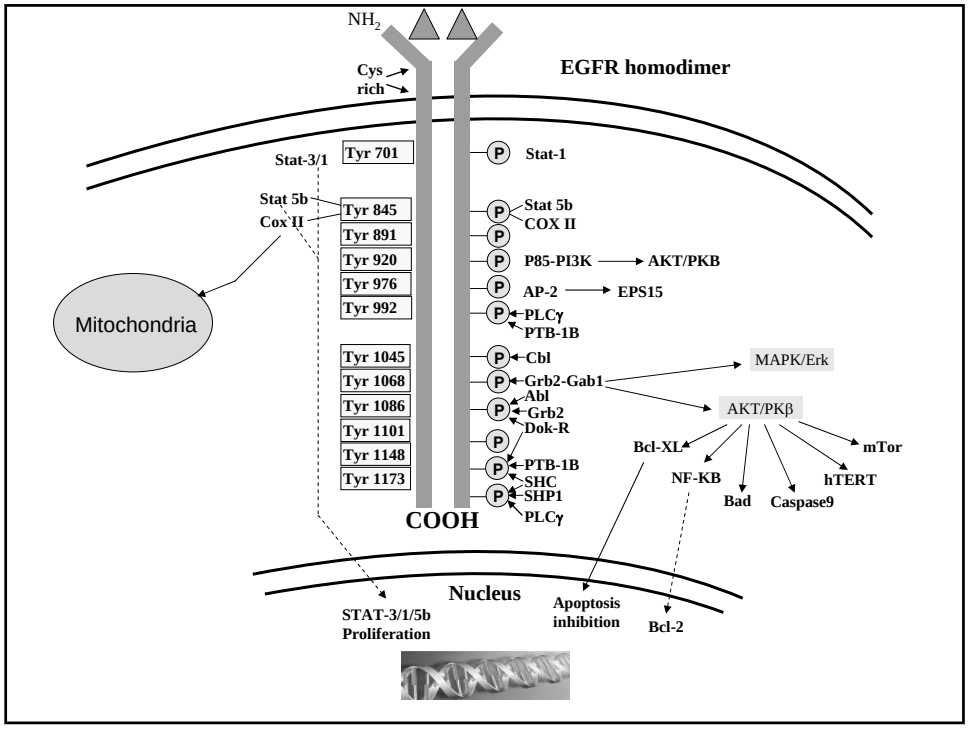


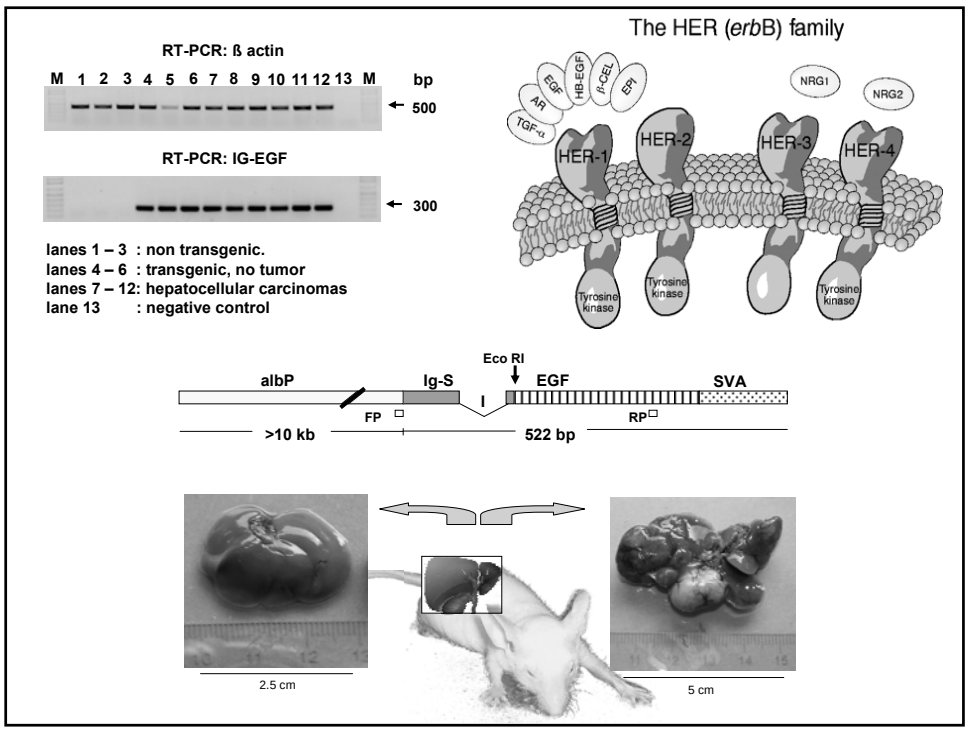
Basic facts about EGFR

- EGFR signaling involves small GTPases of the Rho familiy, whereas EGFR trafficking involves small GTPases of the Rab family.
- Many mutations in the EGFR coding gene identified
- In-frame deletions in exon 19 and point mutations in codon 858 of the EGFR gene are particularly common in NSCLC of never smokers
- Despite the dramatic responses to EGFR inhibitors most patients have a relapse and resistence to drug treatment are frequently associated with secondary mutations

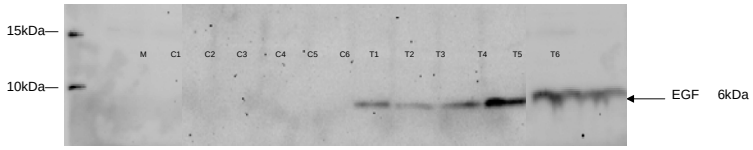

EGFR endocytosis, trafficking & turnover
EGF
EGFR
TGFα
Ampiregulin
Lipid raft
Vesicle formation
Clathrin-coated pit
Clathrin
Plasma membrane
Dynamin
Epsin
AP2
Adaptor protein 2
NHERF1
Endophilin
SPROTY2
EPS 15
EGFR-pathway substrate-15
Recycling endosome
Cbl-interacting protein
CIN85
Cbl
Ubiquitylation by E3 ligase Cbl
Cox II
Mitochondria
Early endosome
PtdIns3P
HRS
Hepatocyte growth factor regulated tyrosine kinase substrate
PTP1B
ESCRT-I complex
TSG101
STAM
Signal transduction adaptor molecule
Endoplasmic reticulum
ESCRT-II and–III (= endosomal sorting Complexes required for transport)
Multivesicular endosome / multivesicular body
Nucleus
Late endosome lysosome
Activation of DNA-dependent protein kinase for DNA repair and/or gene expression

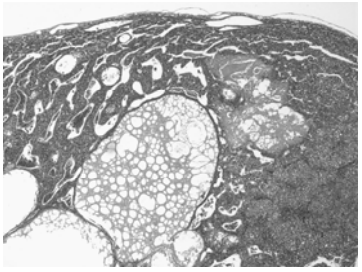

ErbB2 Protein Tyrosine Kinase Subfamily
EGF, VGF, TGFalpha..
Heregulin, NDF..
dual cysteine cluster
tyrosine kinase domain
known heterodimers
1-2
1-3
2-3
2-4
EGFR= ErbB-1 =HER-1
ErbB-2= Neu =HER-2
ErbB-3 =HER-3
ErbB-4 = HER-4

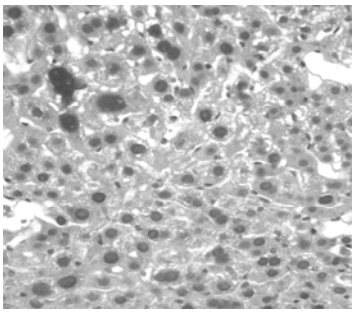

NH2
Cys rich
EGFR homodimer
Stat-3/1
Tyr 701
P
Stat-1
Stat 5b
Cox II
Tyr 845
Tyr 891
Tyr 920
Tyr 976
Tyr 992
Stat 5b
COX II
P85-PI3K
AKT/PKB
AP-2
EPS15
PLCγ
PTB-1B
Mitochondria
Tyr 1045
Tyr 1068
Tyr 1086
Tyr 1101
Tyr 1148
Tyr 1173
Cbl
Grb2-Gab1
Abl
Grb2
Dok-R
PTB-1B
SHC
SHP1
PLCγ
MAPK/Erk
AKT/PKβ
Bcl-XL
NF-KB
Bad
Caspase9
mTor
hTERT
COOH
Nucleus
STAT-3/1/5b
Proliferation
Apoptosis
inhibition
Bcl-2


RT-PCR: ß actin
M 1 2 3 4 5 6 7 8 9 10 11 12 13 M
bp
500
RT-PCR: IG-EGF
300
lanes 1 – 3 : non transgenic.
lanes 4 – 6 : transgenic, no tumor
lanes 7 – 12: hepatocellular carcinomas
lane 13 : negative control
The HER (erbB) family
EGF
AR
TGF-α
HB-EGF
β-CEL
EPR
NRG1
NRG2
HER-1
HER-2
HER-3
HER-4
Tyrosine kinase
Eco RI
albP
Ig-S
I
EGF
SVA
FP
RP
>10 kb
522 bp
2.5 cm
5 cm

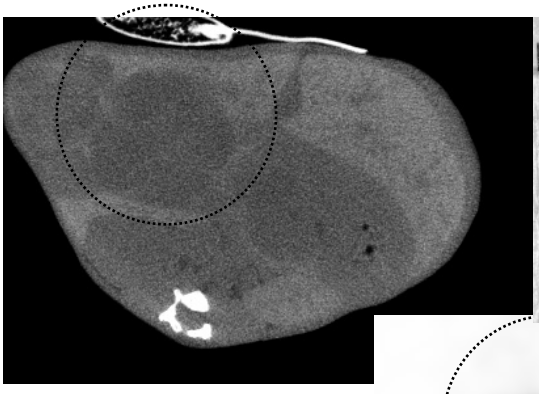
Targeted overexpression of the mitogen EGF induced HCC

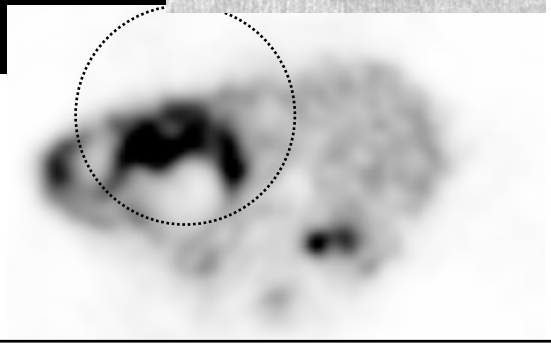






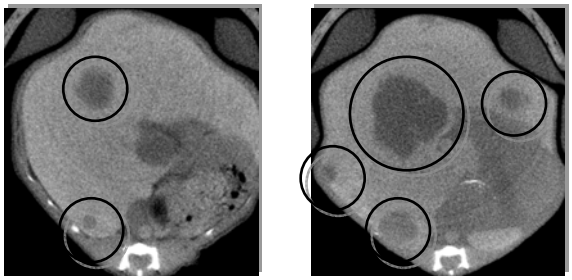
Computed tomography imaging (CT scan) of EGF induced HCC





Transgenic mouse liver tumor
at the age of 9 month
PET: 10MBq [^{18}F]-FDG
CT: 250µl Fenesta LC i.v.

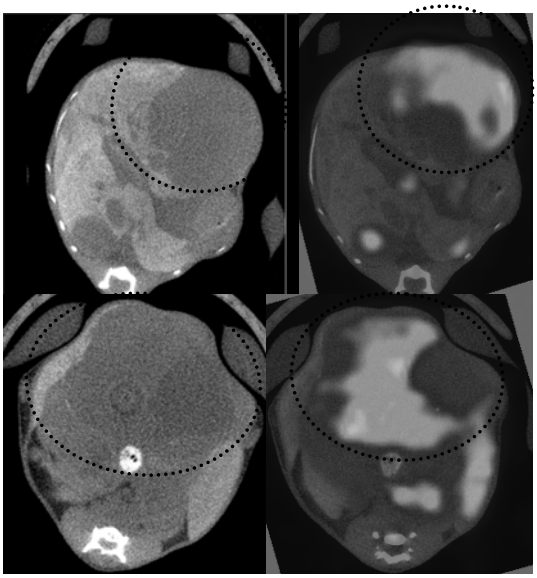
Monitoring tumor growth...



EGF2B
2 weeks interval

Fraunhofer Institut
Toxikologie und
Experimentelle Medizin

PET/CT imaging of EGF liver tumors



CT: 250µl Fenesta LC i.v. PET: 10MBq [^{18}F]-FDG

Genomics enables hypothesis generation to probe for disease causing mechanism(s)

Tox & Disease Phenotype

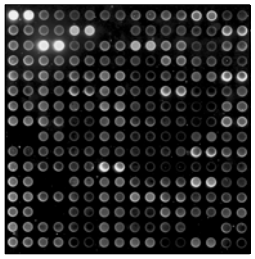
↓ ↑

Predict altered metabolic networks

↓ ↑

Microarrays inform on changes in transcript profiles

↓ ↑



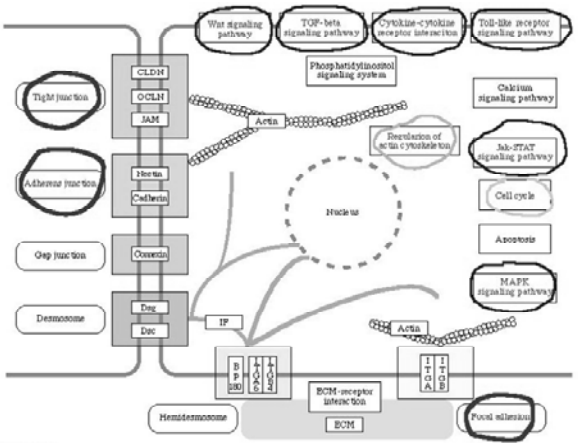
Oncogenomics of EGF induced HCC

· Tumor size dependent transcript signatures of EGF induced hepatocellular carcinomas identified 146 regulated genes which had in at least one group of the tumors a FC > 3 or <-3, a p-value in T-test <0.05, a signal intensity > 70 and were regulated in 100% of pair wise analyses compared with normal liver of control mice

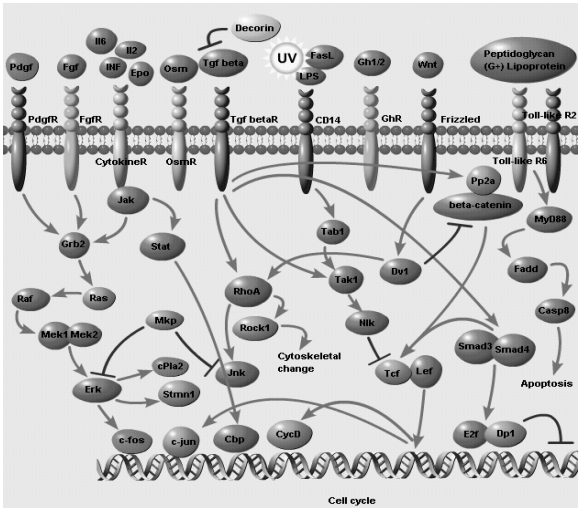
· candidate genes for tumor development could be identified, i.e. molecular switches from transgenicity to low and high grade dysplasia to HCC

Multiple signaling pathways perturbed in EGF induced hepatocellular carcinomas

Cell Communication	
Tight junction	RRAS; PPP2CB; MYH9; CLDN1; ACTN1
Focal adhesion	ZAP70; VASP; RRAS; ROCK2; RELN; Pdgfa; JUN; IGF1; COL4A2; Col4a1; CCND1; CAPNS1; CAPN3; ACTN1
Adherens junction	TCF7; LMO7; IQGAP1; ACTN1
Cell Motility	
Regulation of actin cytoskeleton	RRAS; ROCK2; Pdgfa; MYH9; IQGAP1; FGFR2; CD14; ARPC1B; ACTN1
Signal Transduction	
TGF-beta signaling pathway	TFDP1; ROCK2; PPP2CB; DCN
MAPK signaling pathway	STMN1; RRAS; PLA2G12A; Pdgfa; JUN; FOS; FGFR2; DUSP6; DUSP16; CD14
Toll-like receptor signaling	TLR6; JUN; FOS; CD14
Wnt signaling	TCF7; ROCK2; PPP2CB; JUN; CCND1
Cytokine-cytokine receptor interaction	PDGFA; OSMR; LTBR; IFNGR2; Ifngr1; GHR; CXCL12
Jak-STAT signaling	OSMR; IFNGR2; Ifngr1; GHR; CCND1
Cell Growth and Death	
Cell cycle	TFDP1; PLK1; MAD2L1; CCND1; CCNB2

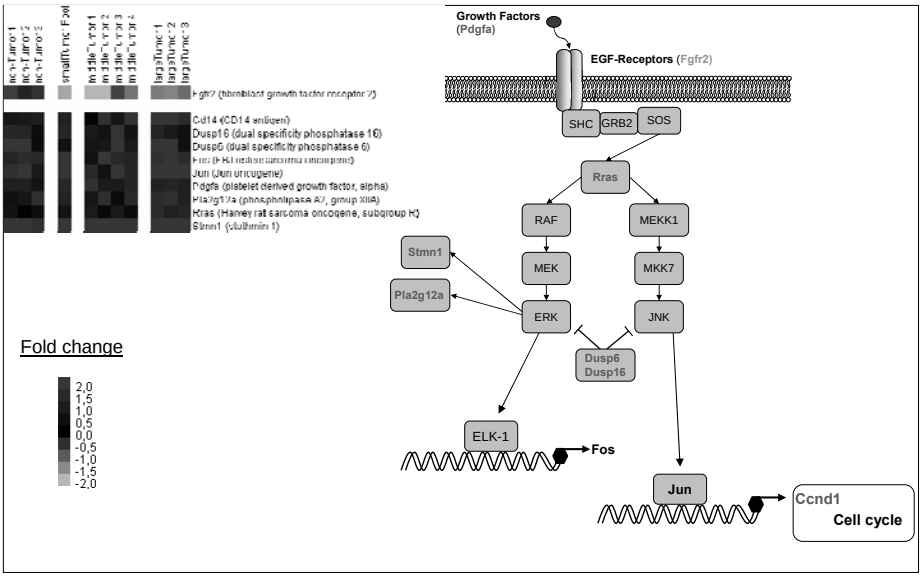


MAPK, TGF beta, Wnt, Jak-STAT, Toll-like receptor signaling pathways and cytokine-cytokine receptor interactions

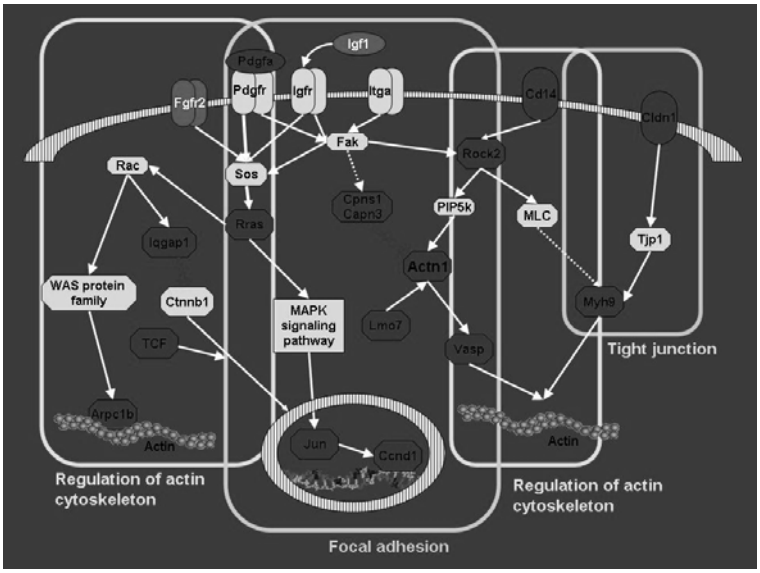


Dcn: decorin; **Fgfr2:** fibroblast growth factor receptor 2; **GhR:** growth hormone receptor; **CD14:** CD14 antigen; **CycD:** Cyclin D1; **CytokineR:** interferon gamma receptor, 2; **Dusp6:** dual specificity phosphatase 6; **Dusp16:** dual specificity phosphatase 16; **Fos:** FBJ osteosarcoma oncogene; **Jun:** Jun oncogene; **Osmr:** oncostatin M receptor; **Pdgfa:** platelet derived growth factor, alpha; **Pla2g12a:** phospholipase A2, group XIIA; **Ppp2cb:** protein phosphatse 2a, catalytic subunit, beta isoform; **Rock2:** Rho-associated coiled-coil forming kinase 2; **Rras:** Harvey rat sarcoma oncogene, subgroup R; **Stmn1:** stathmin1; **Tcf7:** transcription factor, T-cell specific; **Tfdp1:** transcription factor Dp1; **Tlr6:** toll-like receptor 6

Altered MAPK signaling in EGF induced hepatocellular carcinomas

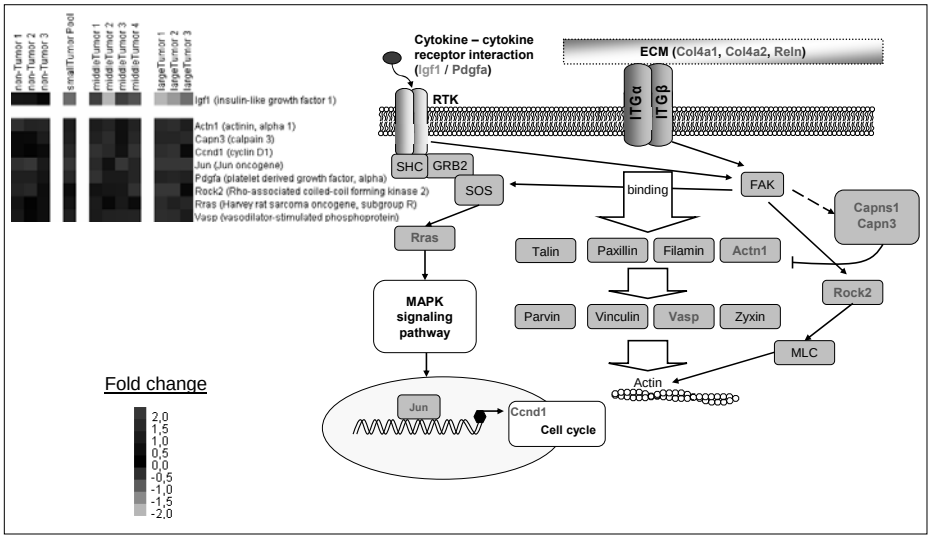


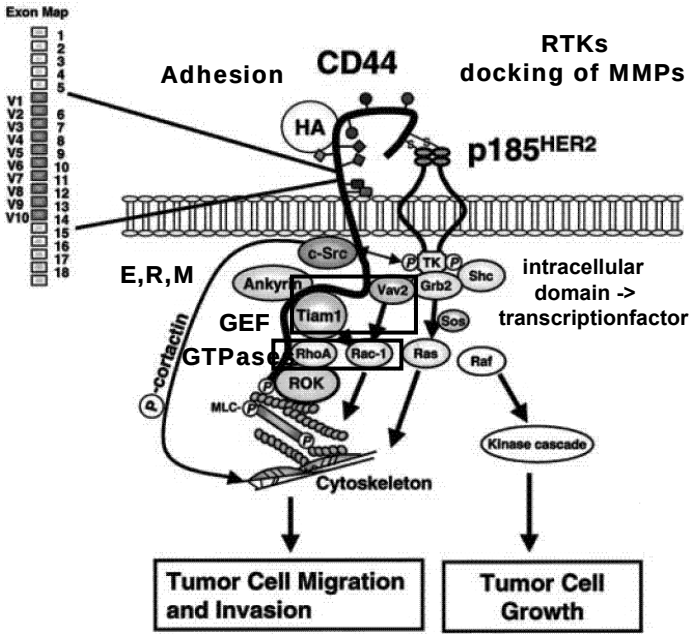
Remodeling of adhesion and actin cytoskelton in EGF induced HCC



ACTN1: actinin, alpha 1, **ARPC1B**: actin related protein 2/3 complex, subunit 1B, **CAPN3**: calpain 3, (p94), **CCND1**: cyclin D1, **CD14** : CD14 antigen, **CLDN**: claudin 1, **FGFR2**: fibroblast growth factor receptor 2, **IGF1**: insulin-like growth factor 1, **IQGAP1**: IQ motif containing GTPase activating protein 1, **JUN**: v-jun sarcoma virus 17 oncogene homolog, **LMO7**: LIM domain 7, **MYH9**: myosin, heavy polypeptide 9, **PDGFA**: platelet-derived growth factor alpha polypeptide, **ROCK2**: Rho-associated, coiled-coil containing protein kinase 2, **RRAS**: related RAS viral (r-ras) oncogene homolog, **TCF7**: transcription factor 7, **VASP**: vasodilator-stimulated phosphoprotein

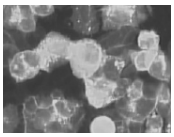
Altered focal adhesion in hepatocellular carcinomas





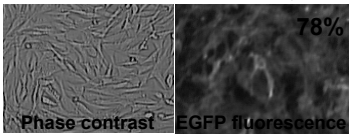
retroviral CD44-miRNA cassette delivery in murine tumor cells

packaging via HEK293T

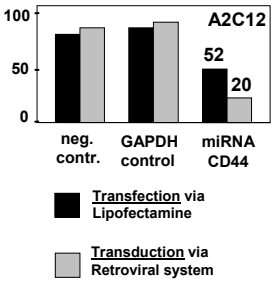


transduction of A2C12
MLV-derived retrovirus

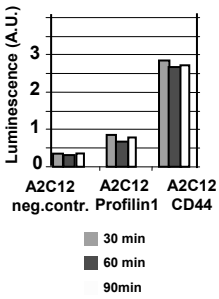
transduction via "three-vector system"



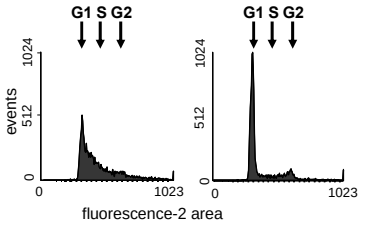
gene expression qRT-PCR CD44



caspase-3/7-assay

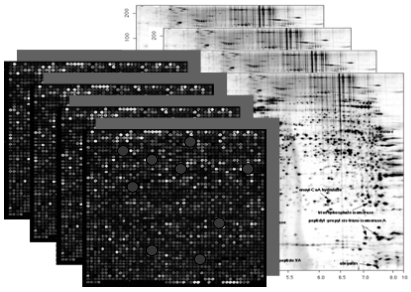


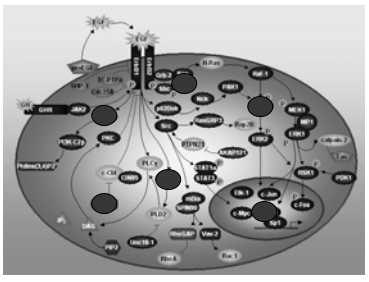
cell cycle analysis



	G1	S	G2
A2C12 miRNA neg. control	35.76	47.21	17.03
A2C12 CD44-miRNA transduced	91.84	6.20	1.96

The challenge: Pathway mapping



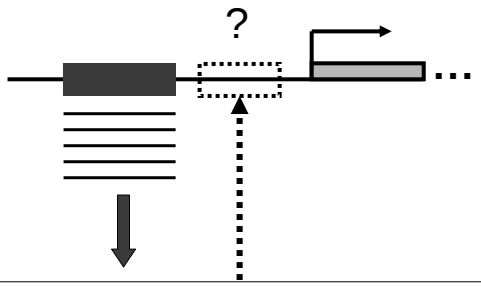


Differentially expressed genes/proteins

Mapping on pathways

Cause of disease ??

To reduce complexitiy of gene expression data common TF binding sites of regulated genes are studied with PWMs

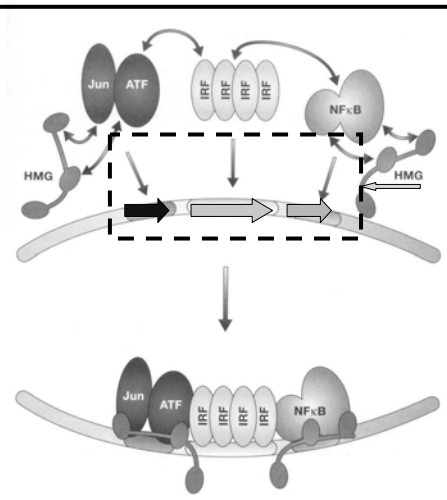


A	9	2	1	0	1	0	0	0	0	1	**15**	**13**	**13**	7
C	8	3	1	1	**13**	3	**29**	0	**22**	**8**	**9**	1	4	8
G	4	2	2	2	**15**	**26**	0	**29**	7	**17**	3	**7**	**9**	8
T	8	**22**	**25**	**26**	0	0	0	0	0	3	2	**8**	3	6
	N	**T**	**T**	**T**	**S**	**G**	**C**	**G**	**C**	**S**	**M**	**D**	**R**	**N**

$$q = \frac{\sum_{i=1}^{l} I(i) f(b_i, i) - \sum_{i=1}^{l} I(i) f^{\min}(i)}{\sum_{i=1}^{l} I(i) f^{\max}(i)} \qquad (1)$$

$$I(i) = \sum_{b \in \{A,T,G,C\}} f(b,i) \ln(4 f(b,i)) \qquad (2)$$

Enhanceosome

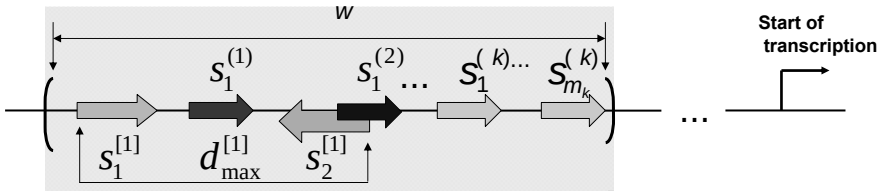


Composite Module (CM)

FIGURE 3.3. The human interferon-β enhanceosome. HMG represents HMGI/Y, a ubiquitous protein that binds cooperatively with the three activators. HMGI/Y both bends the DNA and contacts the activators. Each of the transcription factors shown is a member of a family of related activators.

(Mark Ptashne, Alexander Gann Genes and Signals, 2002)

Composite Modules (CM) are genetic algorithms to find site combinations of different TFs for the prediction of a putative enhancesosome



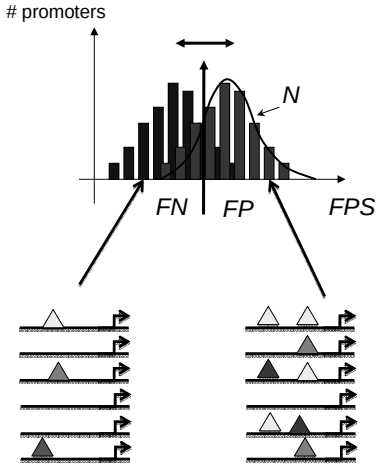
$$d_{\max}^{[1]} \quad d_{\max}^{[1]} \quad \dots \quad d_{\max}^{[R]}$$
$$q_{cut-off}^{(1)} \quad q_{cut-off}^{(2)} \quad \dots \quad q_{cut-off}^{(k)}$$
$$\phi^{(1)} \quad \phi^{(2)} \quad \dots \quad \phi^{(k)}$$

Parameters of the model to be estimated by GA

K= number of PWM; R = limit of distance between matches of matrix pairs, o = relative improtance; d = score of the context

Fitness function of the Genetic-Regression Algorithm (GRA)

$$F = \alpha \cdot R + \beta \cdot (1 - FN) + (1 - \beta) \cdot (1 - FP) + \gamma \cdot T + \delta \cdot N - \mu \cdot k$$



R – linear regression

FN – false negatives

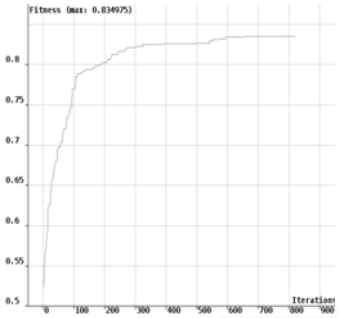
FP – false positives

T – T-test (difference between mean values)

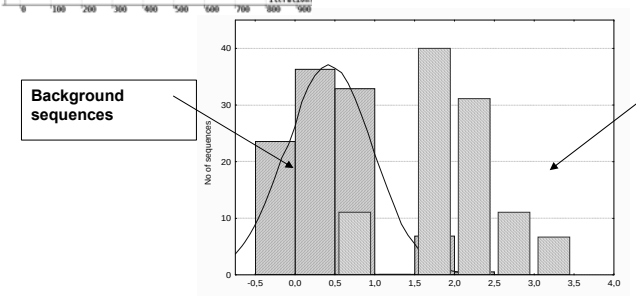
N – normal likeness

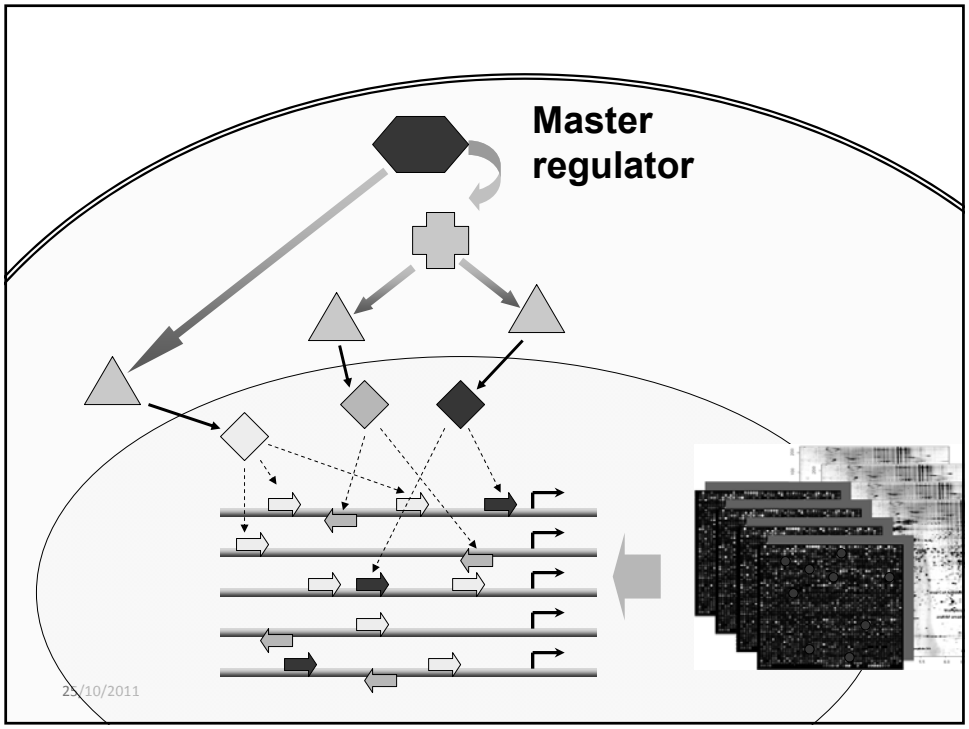
k – number of free parameters

Composite module in promoters of cell cycle-related genes



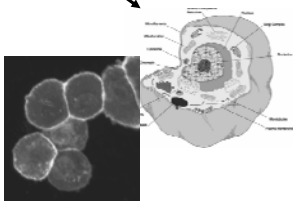
Weight: ϕ	$q_{cut-off}$	TF matrix
1.000000	0.840072	V$E2F_19
0.954483	0.737637	V$TATA_01
0.888064	0.939687	V$CREB_01
0.816179	0.941583	V$SP1_Q6
0.039746	0.839702	V$TAL1BETAE47_01



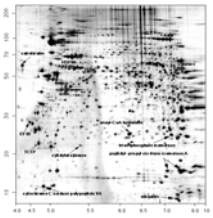


Human epidermoid carcinoma A431 cells treated with
epidermal growth factor (EGF)

EGF ?



320 differentially expressed proteins



Master regulator analysis

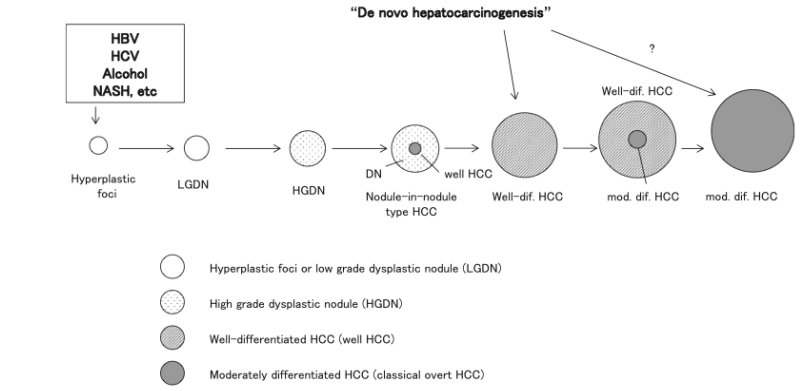
EGF was still not in the list !

Mapping differentially expressed
proteins to canonical signal transduction pathways

Pathway name	#Hits in group	Hit names	Group size	p-value
Caspase network	6	K18; E1; Cytochrome C; Hsp10; Ku70; Cdc42	104	0.00201348
CHIP ---/ Pael-R	2	E1; Hsc70	12	0.01177937
p53 pathway	4	E1; L23; Cytochrome C; Ku70	79	0.02072214
beta-catenin ---/ KAI1	1	Reptin52	5	0.06701759
Aurora-A cell cycle regulation	2	Ubc5B; E1	34	0.07924485
JNK pathway	3	E1; 14-3-3zeta; Trx1	75	0.0813304
parkin associated pathways	2	E1; Hsc70	40	0.10447487
beta-catenin:E-cadherin complex phosphorylation and dissociation	1	alpha-catenin	9	0.11739049
stress-associated pathways	3	E1; 14-3-3zeta; Trx1	100	0.15476
hypoxia pathways	1	Trx1	24	0.2849595
TNF-alpha pathway	1	Trx1	36	0.39594524
EGF pathway	1	E1	103	0.57615756

- Pathways are far..far...far from being fully understood!
- Also, there is network plasticity

We like to think in a linear fashion!



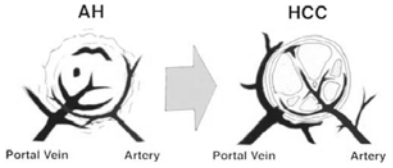
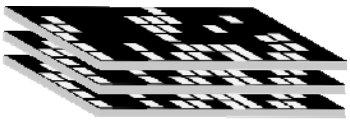


Fig. 1. Schematic representation of multistep progression of human hepatocarcinogenesis

Fig. 2. Schematic representation of blood supplies shifting from LGDN to hepatocellular carcinoma (*HCC*). *AH*, adenomatous hyperplasia

Epidermal Growth Factor induced Carcinogenicity

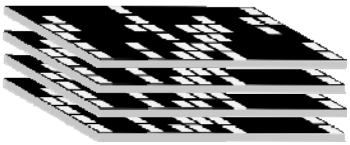
transgenic



transgenic/normal

Tumorigenic switch ?

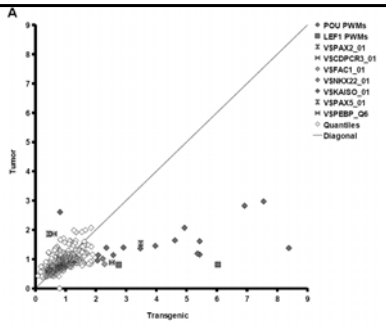
small tumor as opposed to large tumors



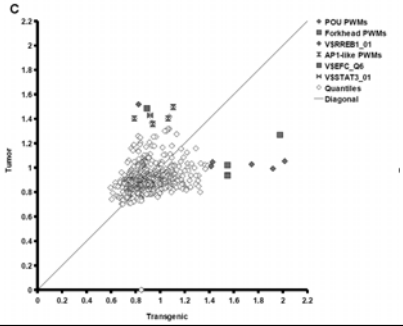
small tumors progress to large tumors

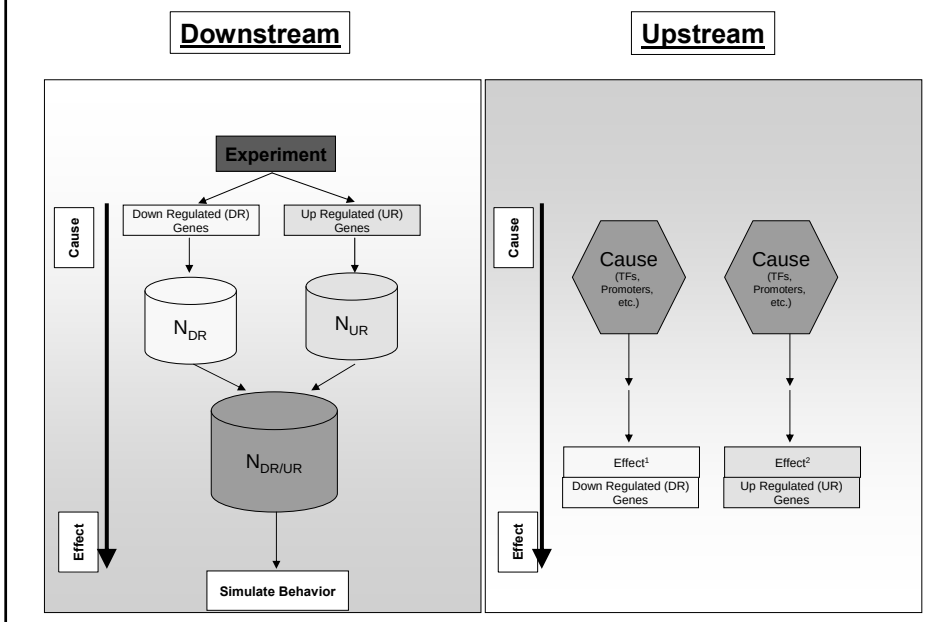
Differentially expressed genes

Down-regulated genes



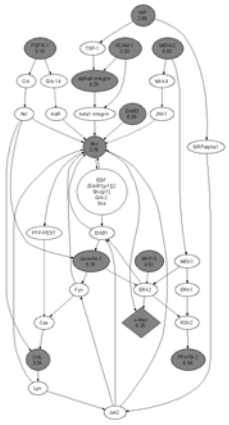
Up-regulated genes



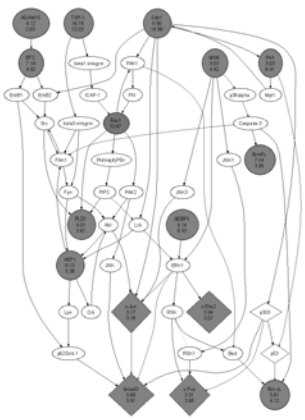


Network clusters in EGF induced hepatocellular carcinomas

Transgenic cells only

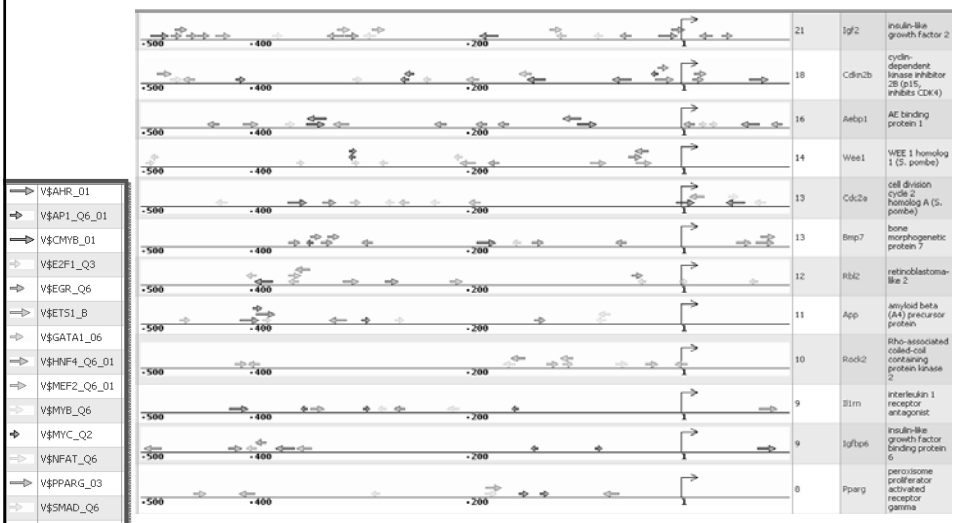


Transgenic & tumor cells

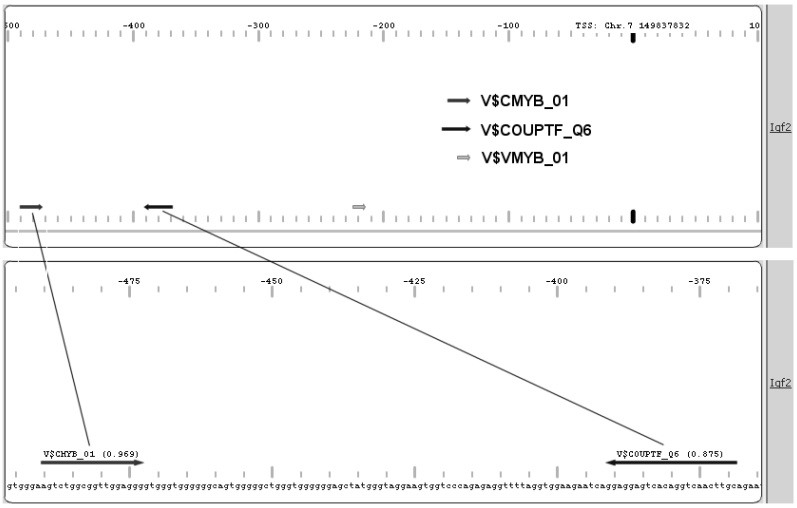


Tumor cells only

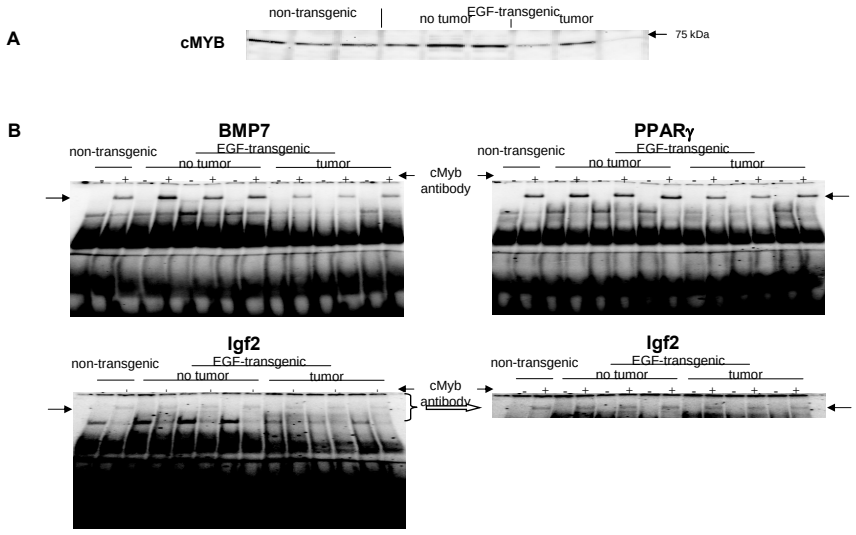
Predicted TF sites in promoters of up- and down-regulated genes



Promoter sequence of insulin growth factor gene



Western blotting and EMSA band shift assays of de novo identified TF and targeted promoters



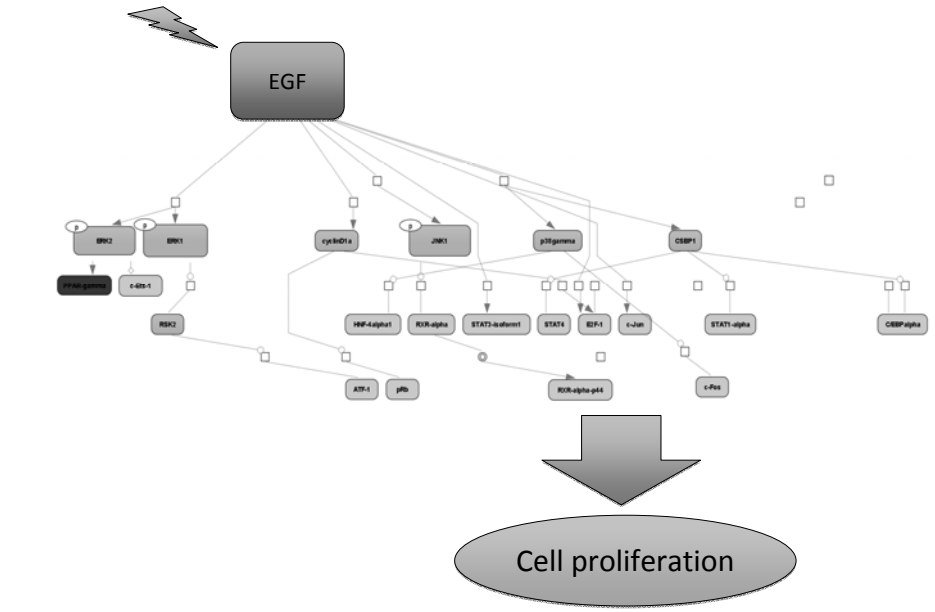
Experimental validation by EMSA

Gene	Fold change Transgenic	Fold change Tumor	PPAR-gamma	c-Myb	c-Ets-2	STAT5B	GATA-1	Mef2a	p53	HNF-4 gamma (alpha - antibody)	GR	C/EBP alpha
Igf2	2.90	25.98	-	+	-			-				
Il1rn	2.75	8.49		-	-	+						
Igfbp6	0.83	7.84	+	-	+	+						
Pparg	0.40	5.76		+	-	+	+	-				
Bmp7	0.92	4.64		+	-	-						
Zbtb7b	3.69	1.11							+			+
Foxc1	4.67	1.43								+	-	+
Xlr	2.90	0.90						-	+			
Erbb3	6.99	2.35	-						+			
Itga4	2.76	1.03							-			+
Th	4.08	1.59						+	-			+
Nr2f1	6.26	2.62							+	+		
Defcr6	13.86	6.05								+	-	+
Nr3c1	3.27	1.44							-		-	+
Cav1	5.35	2.49							-		+	+
Sprr2i	3.24	1.59							+			+
Mgp	6.67	3.31						-	-			
Fgf18	4.07	2.02	-						+		-	
Egf	0.63	0.31							-	+		+

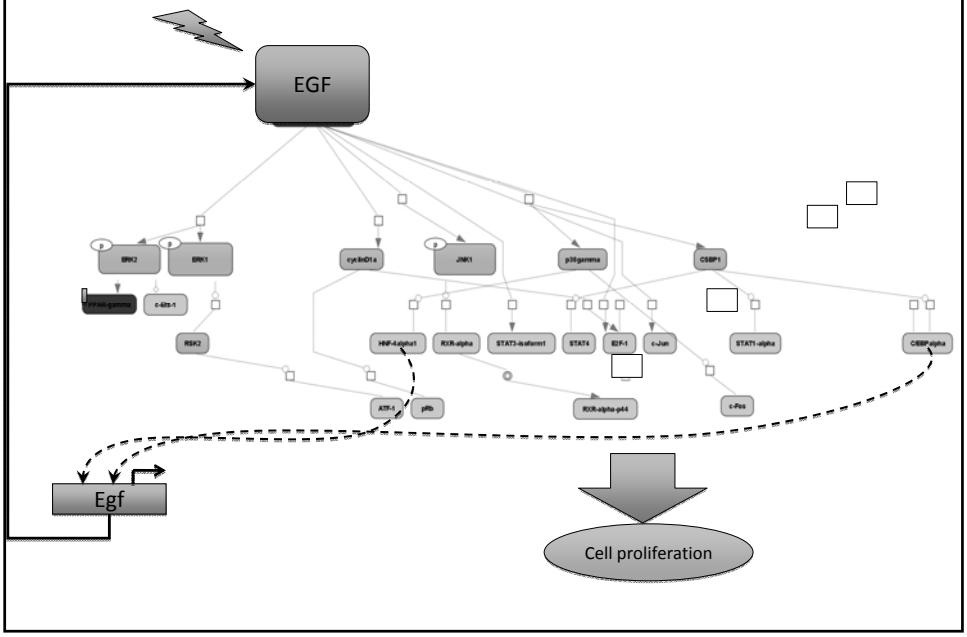
Description

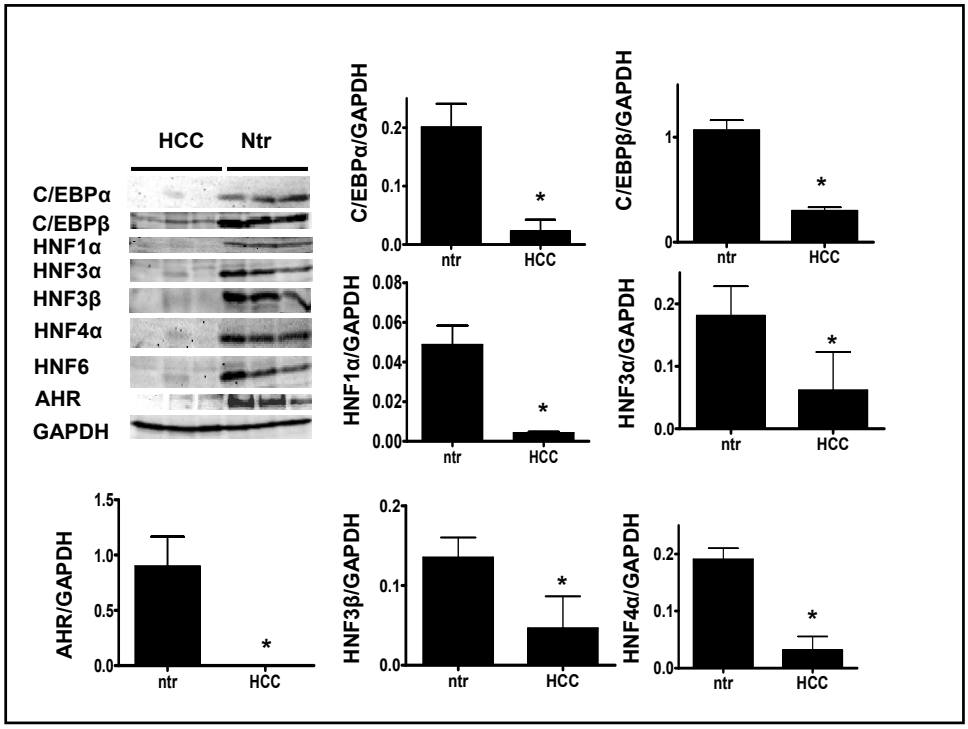
"+" EMSA confirmed predicted site, "-" predicted site was not confirmed

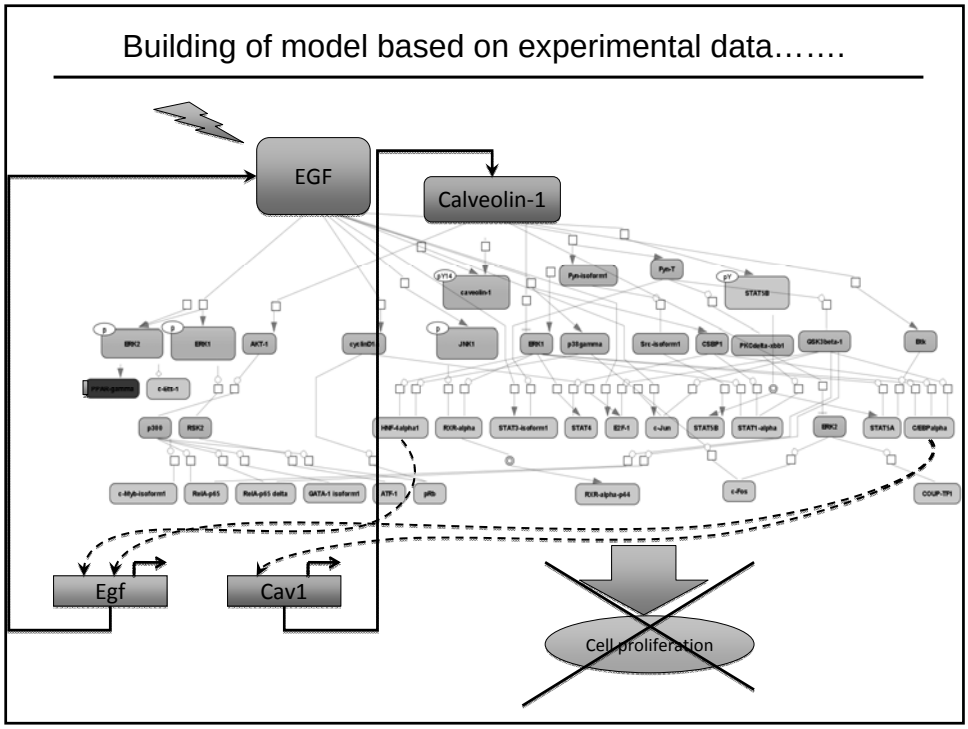
Building of model based on experimental data…….



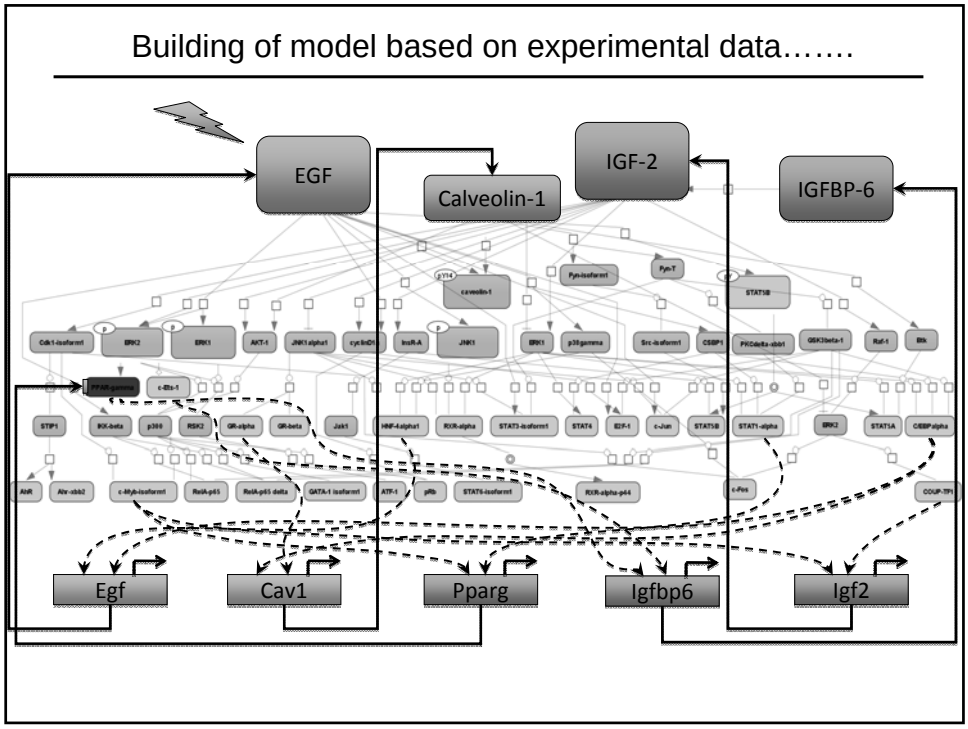
Building of model based on experimental data…….



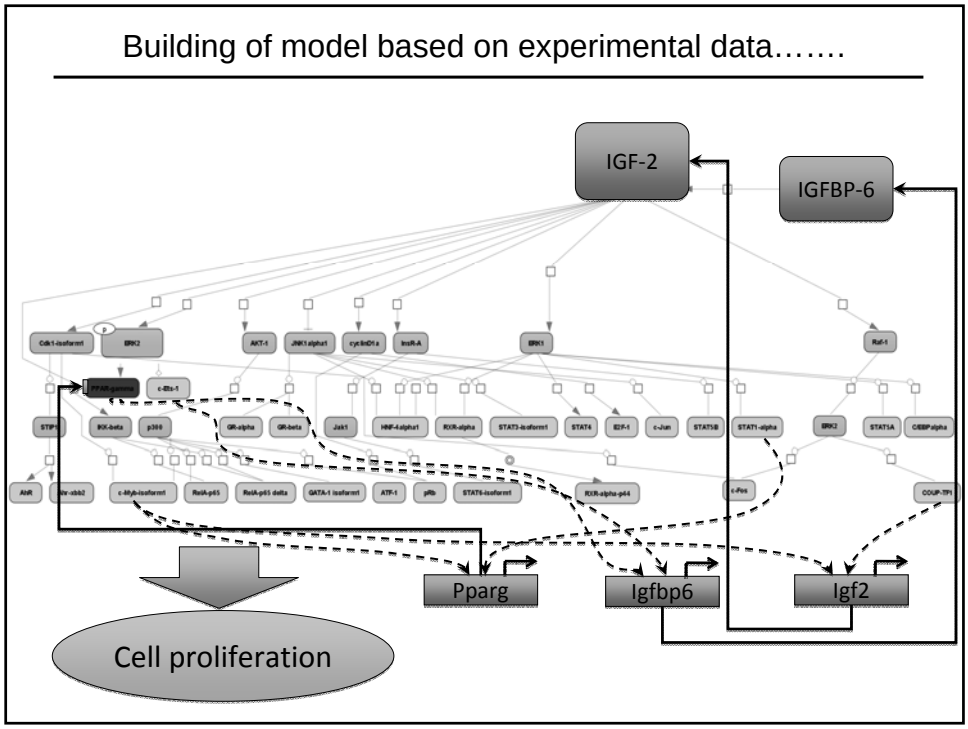

HCC
Ntr
C/EBPα
C/EBPβ
HNF1α
HNF3α
HNF3β
HNF4α
HNF6
AHR
GAPDH
C/EBPα/GAPDH
C/EBPβ/GAPDH
HNF1α/GAPDH
HNF3α/GAPDH
AHR/GAPDH
HNF3β/GAPDH
HNF4α/GAPDH
ntr
HCC


Building of model based on experimental data…….
EGF
Calveolin-1
Egf
Cav1
Cell proliferation

Building of model based on experimental data…….



Building of model based on experimental data…….



Overall summary and conclusions

- Based on a combination of sequence analysis and entrained genetic algorrithms gene regulatory networks were constructed to identify de novo TF, processes, key nodes and molecules to connect as yet unknown interacting partners at the level of protein-DNA interactions.
- A switch in autocrine signaling was identified to foster tumor growth that was triggered by EGF and sustained by IGF signaling.