# Fuzzy modeling reveals a dynamic self-sustaining network of the GLI transcription factors controlling important metabolic regulators in adult mouse hepatocytes

Wolfgang Schmidt-Heck*[a], Madlen Matz-Soja*[b] Susanne Aleithe*[b], Eugenia Marbach[b], Reinhard Guthke[a], Rolf Gebhardt[b]

Abstract

The GLI transcription factors, GLI1, GLI2, and GLI3, transduce Hedgehog and non-hedgehog signals and are involved in regulating development and tumorgenesis. Surprisingly, they were recently found to modulate important functions of mature liver. However, less is known about their mutual interactions and possible target genes in mature hepatocytes. To get a deeper insight into these interactions cultured mouse hepatocytes were transfected with siRNAs against each GLI factor. RNA was extracted at different times and the expression levels of the genes of interest were determined by quantitative real-time PCR. The time-dependent data were analysed by a Fuzzy Logic-based modelling approach. The results indicated that the GLI factors constitute an interconnected network. GLI2 inhibited GLI1 expression and was coupled with GLI3 by a positive feedback loop. The regulatory activity between GLI1 and GLI3 was more complex switching between a positive and a negative feedback loop depending on whether the level of GLI2 is low or high, respectively. Generally, this network structure enables a dynamic behaviour. When GLI2 is low, it may keep GLI1 and GLI3 activity balanced favouring the appropriate modulation of transcription factors like the Ppars and Srebp1. When GLI2 is high, it may prevent an uncontrolled amplification that may lead to cancer. In conclusion, the three GLI factors in mature hepatocytes form an interactive transcriptional network that is involved in the control of target genes associated with metabolic zonation as well as with lipid and drug metabolism. Its structure in mature cells seems different from embryonic cells.

Graphical abstract
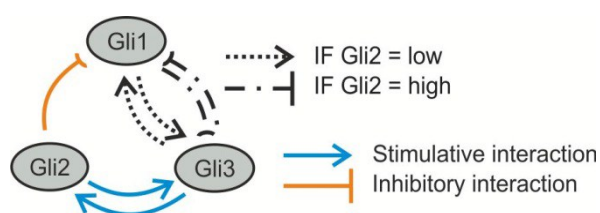
The Hedgehog signalling-driven Gli transcription factors in hepatocytes form a regulatory network identified by a fuzzy-logic modelling approach. The network explains dynamic features important for hepatocyte function and fate.

Introduction

The Hedgehog (Hh) signalling pathway is known to play an important role in embryonic development, organogenesis and tissue homeostasis.[1] In the adult stage, its activity is downregulated in most organs, but can be reactivated in physiological and pathological processes such as tissue regeneration and cancer.[2,3] Despite its low activity in mature organs and cells, current investigation has revealed important functions of this pathway in metabolic control.[4] Most likely, these activities are exerted via both, non-canonical and canonical Hh signalling. As the canonical cascade of events initiated by the Hh ligands has been extensively reviewed elsewhere [2,5,6], only a brief description of the key regulators and their functions is described herein. Secreted Hh ligands bind to Patched 1 (Ptch1) and Patched 2 (Ptch2) receptors at the plasma membrane of receiving cells and alleviate Patched-mediated suppression of Smoothened (Smo). Activated Smo induces a cascade of downstream reactions resulting in the final stabilization of the GLI (glioma associated oncogene) family of transcriptional factors, their translocation into the nucleus and transcription of Hh target genes. Each of the three vertebrate GLI proteins is endowed with a distinct biochemical activity.[3] GLI2 and GLI3 are bifunctional transcription factors that act as full-length activators

(GLI2A/GLI3A) only in the presence of Hh ligands, and otherwise undergo proteolytic cleavage to truncated repressors (GLI2R/GLI3R).[7] In contrast, GLI1 was identified as a direct transcriptional target of GLI2A/GLI3A that functions only as an activator.[8] The complexity of the GLI factor network based on these distinct features has led to the concept of the GLI code.[9] Recently, we discovered that the Hh signalling is active and of considerable functional importance in normal hepatocytes of mature liver.[10] Using transgenic mice with hepatocyte-specific conditional deletion of Smoothened (Smo) and RNAi experiments on cultured hepatocytes we could demonstrate that Hh signalling participates in the transcriptional regulation of various genes including IGF-I and IGFBP-1. Flanking in vitro studies revealed that most effects are transmitted by GLI3 followed by GLI1, while GLI2 plays a less decisive role.[10] In course of these studies we realized that the three GLI factors showed a dynamic expression pattern even in vitro in cultured hepatocytes, i.e. in the absence of external stimuli. In part, these changes seem to reflect a diurnal rhythm of the Hh signalling pathway (Marbach, Matz-Soja, Gebhardt, unpublished observation). The involvement of mutual interactions of the GLI factors, however, cannot be excluded. So far, the interactions of the three GLI factors have been investigated only in the developmental context. In Xenopus neural plate differentiation evidence was collected that GLI3 and Sonic Hedgehog (SHH) repress each other leading to opposite gradients.11 Since GLI2, like GLI1, were found to be targets of SHH signalling and both, GLI2 and GLI3, antagonize GLI1 function, a regulatory feedback loop triggered by SHH is established.[11] Likewise, a positive regulation of GLI1 by SHH and a negative interaction between SHH and GLI3 was found during limb development.[12] During somite myogenesis in the mouse embryo GLI2 and GLI3 were recognized as essential for Gli1 expression in somites.[13] These findings suggest complex direct and indirect interactions among the GLI transcription factors which, however, remain poorly understood. Whether they can be applied to the adult stage and, particularly, to the adult liver is doubtful for a number of reasons. In the present study we aimed at elucidating the GLI factor network in adult hepatocytes. Since robust readouts or antibodies that can discriminate between the activating versus repressing form of GLI proteins are currently unavailable, we focused exclusively on the transcriptional level and used an RNAi experimental approach to provoke network responses. In order to analyse the data in the most efficient way and to improve its understanding mathematical modeling was applied. Since the Hh signalling pathway has been modelled mainly in a completely different context (developmental aspects)[14] and frequently does not distinguish between the different GLI factors, we applied a fuzzy modelling approach able to cope with the expected complexity and uncertainties. The model obtained not only matched the measured expression levels of all three GLI factors, but also led to the surprising result that the model structure switches between two distinct modes of interactions between GLI1 and GLI3 depending upon the abundance of GLI2.

Results and discussion

Changes in GLI mRNA levels in cultured hepatocytes suggest an interactive GLI factor network

Studying the GLI transcription factors in healthy mature mouse hepatocytes is challenging as their mRNA levels are orders of magnitude lower than those observed in embryogenesis.[15] When primary mouse hepatocytes are cultured in Williams medium E under standard conditions, the mRNA levels of GLI1, GLI2, and GLI3 are not following a specific trend, but apparently show mild fluctuations within the recorded period of 72 h. Although it is rather challenging to precisely measure these levels with qRT-PCR, it is unlikely that these fluctuations result from imprecise measurements, but instead reflect some hidden dynamic behaviour. This possibility prompted us to approach the phenomenon by performing transfection experiments with siRNA against each GLI factor combined with extensive time-series measurements starting at 24 h or 48 h of cultivation and lasting till 72 h. In order to exclude the disturbance of these measurements by a possible circadian rhythm of the Hh signalling pathway, samples were taken at exactly the same time each day in the morning (ZT3 = 9 AM). The time courses of the mRNA levels of the three GLI factors under the influence of the three different siRNAs are depicted relative to the respective nonsense oligo RNA in Fig. S1, ESI† for up to 20 independent hepatocyte cultures. The efficiency of each siRNA is illustrated by the almost complete knockdown of Gli1 after 72 h, of Gli2 after 48 and 72 h, and of Gli3 after 72 h in Figs. S1A, S1E, and S1I, respectively, ESI†. As shown in Fig. 1A, the knockdown of Gli1 after 72 h resulted in the significant increase of the mRNA values of Gli2 and Gli3. The siRNA-mediated knockdown of Gli2 was accompanied by an up-regulation of Gli1 and Gli3 (Fig. 1B) and that of Gli3 by a significant down-regulation of Gli1 (Fig. 1C).

Thus, each GLI factor seemed to be regulated to some extend by the other two, but the precise way of how this happens cannot easily be predicted intuitively from such diagrams. Similar interactions among the GLI factors have not been described before in such detail for mature cells. Although they reflect some interactions described in embryonic systems[11-13], the GLI factor network characterized herein seems to be specific for hepatocytes.
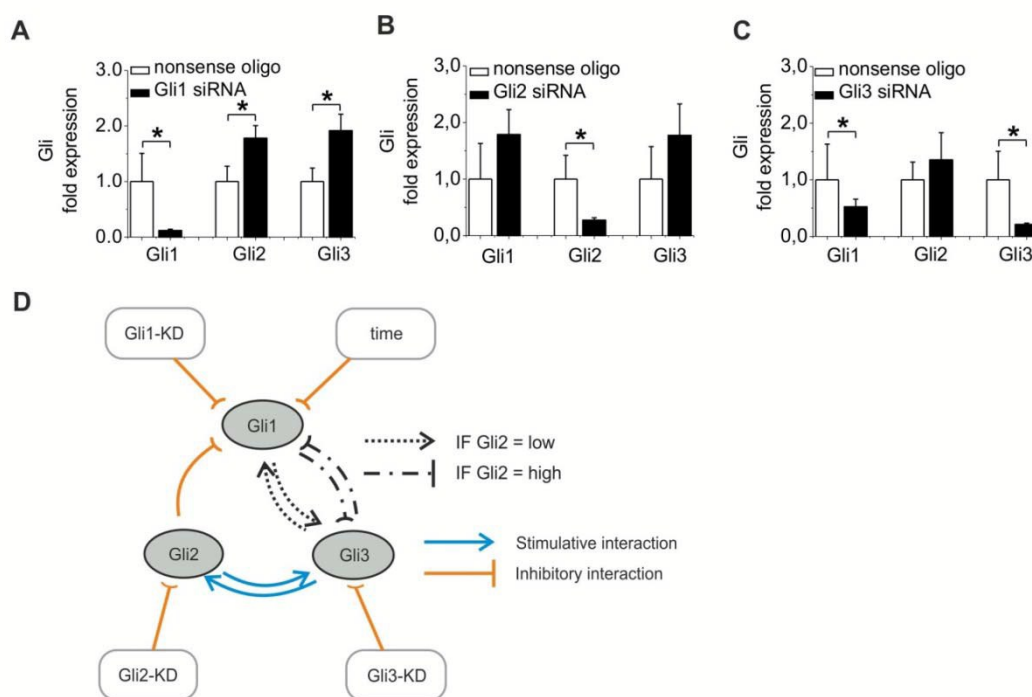
Fig. 1 Response of GLI factor expression to factor-specific siRNA-mediated knockdown. Relative expression of GLI transcription factor (Gli1, Gli2, Gli3) mRNA determined by qRT-PCR in hepatocytes 72 h post transfection with **A** Gli1 siRNA (black bars) (n=10-12), **B** Gli2 siRNA (black bars) (n=6-13) and **C** Gli3 siRNA (black bars) (n=11-20) compared to nonsense oligo transfection (white bars) (n=6-20). Values are presented as means + SEM; *, $p < 0.05$. **D** Synopsis of all state-dependent fuzzy gene regulatory networks generated as described in Methods. The network shows that the GLI-specific gene knockdown suppresses the expression of the respective Gli gene. GLI2 suppresses the expression of Gli1. Furthermore, Gli2 and Gli3 are coupled by a positive feedback loop. The regulatory activity between GLI1 and GLI3 can be described by either a positive or a negative feedback loop depending on the level of GLI2. (Further explanations see text)

Fuzzy modeling of the GLI factor network

In order to obtain a better insight into the transcriptional interactions of the three GLI factors, we applied a fuzzy logic-based modelling approach to this gene regulatory network (GRN) (see Methods). The nodes of this network model are represented in a fuzzy inference system (FIS). FIS provides a method to model complex, nonlinear input-output relationships. A second advantage is that this modeling approach can cope with uncertainties in experimental data sets and "missing values" (irrespective whether these were not measured or were below detection level). First, membership functions of fuzzy interference systems (FIS) variables were generated as described in Methods (Fig. S2, ESI†). Then, an ensemble of FISs with all possible combinations of input variables and fuzzy values was generated for each network node (Fig. 2) together with the mean squared error calculated for each possible FIS based on one hundred randomly selected learn and test data sets.

For each node, the FIS`s with the smallest test-error are highlighted (grey) (Fig. 2) and selected for the mining of the rule set of the Fuzzy-GRN (Fig. S2). It was found that the optimal number of input variables (N) for GLI1 is 4 (time, Gli1-KD, GLI2, GLI3) and for GLI2 or GLI3 equals 3 (Fig. 2). A maximum of 2N rules related to the highest values of the coefficient lift was selected for the respective rule sets provided that this value is greater than 1 (Fig. S2, ESI†). Further analysis for determining and visualizing the kind of regulatory activity between the components of the network was performed as described in Methods. Fig. 1D shows the resulting synopsis of all state-dependent Fuzzy-GRNs depicted in Fig. S3, ESI†. This network reflects all the regulatory results. Furthermore, a declining effect of the cultivation time on GLI1 was found. More interesting regulatory activities were detected between GLI2 and GLI3 on the one hand and between GLI1 and GLI3 on the other hand. The regulatory activity between GLI2 and GLI3 is described by a positive feedback loop. The regulatory activity between GLI1 and GLI3 is more complex and can be described by a positive or a negative feedback loop depending on the level of GLI2. Of note, apparently there is no unambiguous effect of GLI1 on the transcription of GLI2. The validity of the model is depicted in Fig. 3 representing a scatterplot of the measured and predicted values by the Fuzzy-GRN. The regression lines for Gli1 and Gli3 are nearly ideal, while that of Gli2 with a slope of 0.8531 is slightly less accurate.

| Antecedent				Consequent	Evaluation criterion			
#1	#2	#3	#4		mean of mse(train)	SD of mse(train)	mean of mse(test)	SD of mse(test)
time				Gli1	1.1036	0.0256	1.1526	0.4694
Gli1-KD					1.1229	0.0195	1.2337	0.3798
Gli2					1.2059	0.0437	1.3246	0.4565
Gli3					1.0757	0.0271	1.2524	0.5425
Gli2	Gli3				0.8721	0.0363	1.0872	0.4785
Gli1-KD	Gli3				0.8544	0.0303	1.2594	0.5962
Gli1-KD	Gli2				0.9916	0.0275	1.2241	0.4571
time	Gli3				0.9496	0.0234	1.1102	0.4243
time	Gli2				0.8925	0.0280	0.8829	0.3884
time	Gli1-KD				0.7967	0.0269	0.8971	0.4713
time	Gli1-KD	Gli2			0.7494	0.0253	1.1039	0.3973
time	Gli1-KD	Gli3			0.6046	0.0242	0.8397	0.4802
time	Gli2	Gli3			0.7902	0.0206	0.9927	0.3860
Gli1-KD	Gli2	Gli3			0.7458	0.0458	0.9579	0.4324
time	**Gli1-KD**	**Gli2**	**Gli3**		0.5868	0.0371	**0.8362**	0.4829
time				Gli2	0.9413	0.0180	0.9842	0.3216
Gli2-KD					0.8970	0.0187	0.8740	0.3560
Gli1					0.9380	0.0235	0.9762	0.3669
Gli3					0.7337	0.0192	0.7240	0.3588
Gli1	Gli3				0.6826	0.0230	0.9427	0.4309
Gli2-KD	Gli3				0.6431	0.0210	0.7457	0.3972
Gli2-KD	Gli1				0.8833	0.0286	1.1206	0.4227
time	Gli3				0.6795	0.0220	0.7218	0.4265
time	Gli1				0.8581	0.0336	1.0840	0.6021
time	Gli2-KD				0.8136	0.0227	0.9620	0.3681
time	Gli2-KD	Gli1			0.7867	0.0255	1.1718	0.4767
time	Gli2-KD	Gli3			0.5704	0.0341	0.8536	0.4716
time	Gli1	Gli3			0.6221	0.0254	0.8699	0.5132
Gli2-KD	**Gli1**	**Gli3**			0.5815	0.0207	**0.7189**	0.3990
time	Gli2-KD	Gli1	Gli3		0.5624	0.0289	0.9107	0.4292
time				Gli3	1.5300	0.0317	1.6713	0.5938
Gli3-KD					1.1299	0.0290	1.2446	0.5482
Gli1					1.5433	0.0277	1.6550	0.5402
Gli2					1.1936	0.0430	1.3002	0.8105
Gli1	Gli2				0.8848	0.0285	1.2298	0.5993
Gli3-KD	Gli2				0.9050	0.0259	1.2969	0.5533
Gli3-KD	Gli1				0.9562	0.0271	1.1398	0.4994
time	Gli2				1.1219	0.0455	1.5550	0.8842
time	Gli1				1.3614	0.0679	1.6235	0.5159
time	Gli3-KD				1.0568	0.0272	1.2230	0.5203
time	Gli3-KD	Gli1			0.8790	0.0294	1.2621	0.5624
time	Gli3-KD	Gli2			0.8164	0.0394	1.2133	0.5556
time	Gli1	Gli2			0.8189	0.0334	1.2024	0.6579
Gli3-KD	**Gli1**	**Gli2**			0.6355	0.0264	**0.9030**	0.5017
time	Gli3-KD	Gli1	Gli2		0.6185	0.0518	0.9267	0.4162

Fig. 2: Mean squared error (mse) of learn- and test data of fuzzy inference systems (FIS) generated with all possible combinations of input variables as described in Methods in detail. The FIS's with the smallest test-error are highlighted and selected for the fuzzy gene regulatory network.
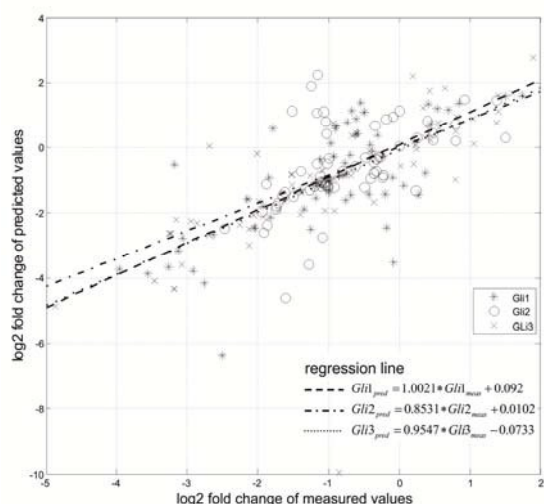
Fig. 3 Scatterplot of the measured and predicted values by the Fuzzy-GRN. The measured data are taken from Fig. S1, ESI†. The regression lines for GLI1, GLI2, and GLI3 are depicted. The insert shows the parameters of the respective regression lines

Applied to the measured data shown in Figs. 1A to 1C the Fuzzy-GRN correctly predicts the results of each knockdown of a GLI factor when GLI2 is considered "high" as in Figs. 1A and 1C or "low" as in Fig. 1B.

From this point of view, the Fuzzy logic-based modelling approach appears ideal in coping with the experimental situation due to the low mRNA levels of the GLI factors in mature hepatocytes. The Fuzzy-GRN presented herein, allows reasonable quantitative predictions about the mRNA levels of the different GLI factors in the absence of external stimuli. Relatively few known parameters are sufficient to perform rough calculations predicting the status of the whole network. Apparently, low numbers of GLI proteins within the cells and, particularly, within the nuclei guarantee the most specific interactions with respective high-affinity binding sites on the DNA. This seems to be a general principle effective in controlling hepatocyte functional zonation[16,17]. Furthermore, a network with so much feedback loops is prone to a dynamic behaviour. This feature may explain the observed fluctuations in non-stimulated hepatocytes in culture. A second surprising feature of the GRN is the dependence of the structure of the network on the level of GLI2. In normal mouse liver the level of GLI2 is usually very low, even lower than that of GLI1 and GLI3 (unpublished observation). Sometimes, it may even be undetectable with commercially available antibodies.[18] At such a low level of GLI2 (which may also be reached in the presence of Gli2 siRNA), GLI1 activates the expression of GLI3 in accordance with the activator functions of GLI1. The positive feedback loop between the two ensures a certain balance between GLI1 and GLI3 that may even hold when GLI3 levels are increased due to enhanced canonical Hh signalling or non-canonical stimulation (see below). In contrast, moderate to high levels of GLI2 in mature hepatocytes may occur under nutritional challenge and have been detected particularly under diseased conditions such as alcoholic steatosis[19]. Under such conditions, a switch from the positive feedback loop to a negative one seems reasonable, because otherwise the GLI factors could mutually amplify each other in an almost uncontrolled manner. The unidirectional inhibition of the GLI1 level by GLI2 may serve the same purpose. Altogether, these inhibitory functions may prevent strongly enhanced expression levels of all GLI factors, a condition that seems frequently associated with cancer.[5]

The negative feedback loop between GLI1 and GLI3 occurring in the latter case fits to the observation that GLI1 can antagonize GLI3 in mouse development.[20] However, it is problematic as the GLI1 protein lacks a repressor function. Therefore, the direct inhibition of Gli3 expression is impossible and has to be replaced by an indirect one via the activation of a distinct repressor (Fig. 4).
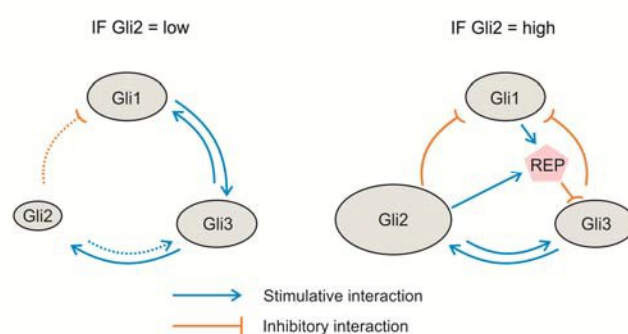

Fig. 4 Scheme of possible molecular interactions between the three GLI factors in the two states (GLI2 = low; dotted arrows) and (GLI2 = high; solid arrows) according to Fig. 1d. The scheme on the right illustrates that a hypothetical repressor (REP) which is stimulated by GLI1 and GLI2 has to be inferred in order to explain how the activator GLI can exert a GLI2-dependent repression of GLI3. (Further details see text)

So far, no factor is known that simultaneously inhibits GLI3 expression, is positively regulated by GLI1 and responds only, if GLI2 level is high. Recently, however, an interesting cooperative action of E2F2 and CREB1 was described in T-lymphocytes that results in the enhanced repression of a subset of E2F2 target genes.[21] Why could this mechanism be a promising candidate for GLI-mediated repression? First, according to the MotifMap database[22], GLI3 is the only GLI-factor with E2F2 binding sites in the promoter region. Thus, E2F2 might act as a repressor of GLI3. Second, CREB1 is a transcriptional target of GLI1 and GLI2 (see MotifMap,[22]). Thus, CREB1 could be induced to sufficiently high concentrations only when both transcription factors are active. Third, according to Laresgoiti and co-workers CREB1 reinforces the repressive function of E2F2.[21] This reinforcement could be the final trigger for switching the interaction mode between GLI1 and GLI3. Whether such a cooperative repressive mechanism is active in hepatocytes and could actually be both, target and regulator of GLI function remains to be investigated.

The GLI factors control the expression of different sets of target genes in hepatocytes

As we have reviewed before[16], several publications on GLI target genes in other mature tissues indicate the involvement of these transcription factors in the regulation of genes associated with lipid metabolism. For instance, Suh and coworkers reported the control of certain lipotrophic genes (e.g. Ppar-γ) by the hedgehog signalling pathway mediated by GATA

factors.[23] Moreover, Pospisilik et al. could show that at least 34 genes of adipogenesis and lipid metabolism have more than one predicted binding site for one or more the three GLI factors of the Hh signalling pathway within 2500 bp upstream of their transcription starting point.[24] Therefore, we performed further qRT-PCR measurements of the mRNA levels of some of these genes in the RNAi experiments (Fig. 5A), in order to find out whether the three GLI factors of the hepatocellular network act via separate sets or a common set of target genes. This question does not seem trivial as embryological studies have suggested that GLI1 can rescue all the SHH signaling defects of Gli2 mutant mice.[20] Our measurements revealed that each GLI factor regulates a (small) distinct subset of metabolic genes but also joins one or both of the other GLI factors in regulating different sets of genes (Fig. 5B).
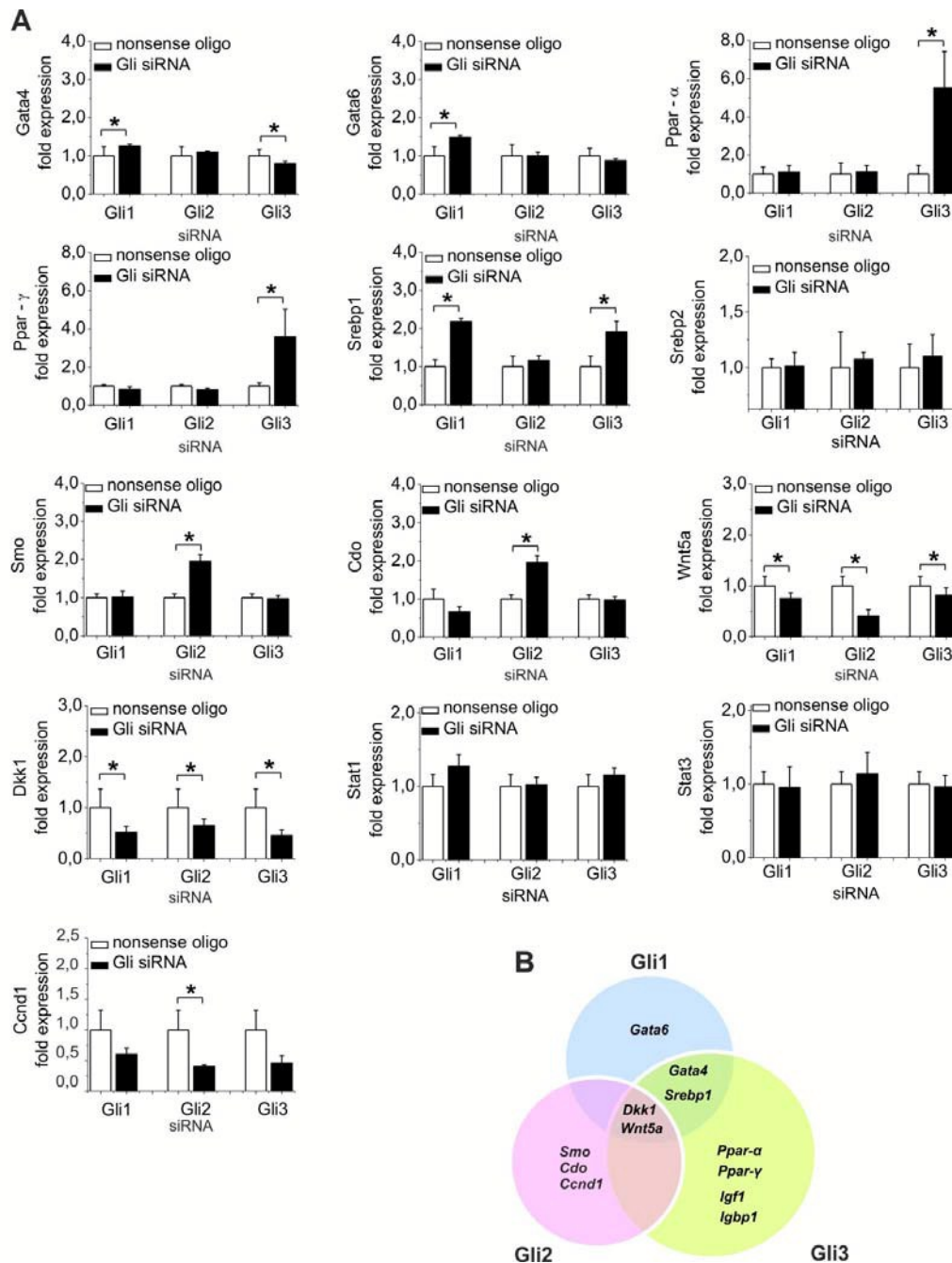
Fig. 5 Response of the expression of selected genes to GLI factor-specific knockdown. **A** Relative gene expression was determined by qRT-PCR in hepatocytes from male mice 72 h post transfection with siRNA (black bars; n=3-10) to Gli1, Gli2 and Gli3 compared to nonsense oligo transfection (white bars; n=3-10). Values are presented as means + SEM; *, p<0.05. **B** Venn diagram of GLI factor-specific regulated genes.

For instance, GLI1 controls Gata4 (GATA binding protein 4) and Gata6 (GATA binding protein 6), whereas GLI3 controls only Gata4. Interestingly, the effect of GLI3 knockdown on Gata4 (suppression) is different from that of GLI1 (enhancement). Likewise, the transcription factor Srebp1 (sterol regulatory element-binding protein 1) is controlled by both GLI1 and GLI3, while none of them seems to regulate Srebp2 (sterol regulatory element-binding protein 2). Ppar-α (peroxisome proliferator-activated receptor α) and Ppar-γ (peroxisome proliferator-activated receptor γ) are up-regulated approximately 5-fold and 3-fold, respectively, by Gli3 knock-down only (Fig. 5A). In contrast, GLI2 may participate in the regulation of proliferation indicated by its stimulating effect on Ccnd1 (Cyclin D1) expression (Fig. 5A). That these findings are not only restricted to the selective knock-down of the GLI factors but reflect the hepatocyte response to Hh signalling is indicated by the fact that these effects can be reversed by activation of the pathway, e.g. through siRNA-mediated knock-down of Ptch2 or Sufu (Fig. S5, ESI†). Even these few results support our hypothesis about the role of Hh signalling in hepatic lipid metabolism.[15] Of course, however, a detailed microarray study is needed in order to show how many genes can be attributed to the different subsets.

In this regard the functional consequences of the GLI factor network for the regulation of the IGF axis should be noted. In our previous publication we have reported that Igf-I is a transcriptional target of GLI3, while Igfbp-1 is repressed by this transcription factor.[10] In contrast, Gli2 knockdown in hepatocytes does affect none of the metabolic genes mentioned above, but selectively up-regulates the Hh signalling genes Smo and Cdo (Cell adhesion associated oncogene regulated) (Fig. 5A). These results are interesting because, on the one hand, they partially exclude GLI2 from the metabolic regulation in hepatocytes, a view compatible with the findings mentioned above. On the other hand, they support the idea that extremely low levels of GLI2 (as after siRNA-mediated knockdown) indirectly stabilize the function of hepatocellular Hh signalling by up-regulating the two co-receptors Smo and Cdo, compatible with reported functions of Cdo in other tissues.[25]

The expression of the transcription factors Stat1 (Signal transducer and activator of transcription 1) and Stat3 is not affected by the GLI factor network in mature hepatocytes (Fig. 5A). This finding is in line with the healthy phenotype of the hepatocytes. In contrast, all three GLI factors positively influence the transcription of Wnt5a (Wingless-type MMTV integration site family, member 5a) and Dkk1 (Dickkopf1) which are secreted regulators of Wnt signaling (Fig. 5A). These results show that the GLI network participates in the regulation of two signalling pathways, namely the Wnt/beta-catenin and the Hh pathway, and therefore seems to be strongly involved in the control of liver zonation.[16]

Finally, it should be noted that the self-sustaining GLI network described here must be considered as kind of a regulatory sub-module in the framework of regulatory signals that influence GLI factor activity. In particular, it can be viewed as part of the GLI code, the sum of all positive and negative functions of all GLI proteins.[25] Besides the canonical Hh signalling pathway that is controlled by the different Hh ligands, a bunch of non-Hh signals (as diverse as TGF-β, EGF, and oncogenic load) acting via non-canonical signalling are able to converge on GLI regulation.[26] The GLI factor network module integrates these signals and may combine them with additional context-dependent inputs.[27] To find out whether the hepatocellular GLI network maintains its structure under the influence of regulatory signals, whether it can be adapted by adding further conditional switches, or whether it completely loses its cooperative structure remains an interesting task for the future.

Conclusions

Collectively, our findings illustrate that the fuzzy logic-based modeling approach reveals novel features of the dynamic behavior of the GLI transcriptions factor network in mature hepatocytes that are distinct from the features observed in embryonic cells. In particular, the network may provide more robustness in case of external challenge. Apparently each GLI factor of the hepatocellular network regulates a different set of target genes with only few overlap. Since GLI1 and GLI3 dominate over GLI2 in mature hepatocytes, genes associated with metabolic zonation as well as with lipid and drug metabolism seem to be the primary targets.

Materials and methods

Maintenance of mice and feeding

C57BL/6-N mice or WT SAC mice[10] were maintained according to the German guidelines and local regulations for the care and safe use of experimental animals in a pathogen-free facility as described[10].

Isolation and cultivation of primary hepatocytes

Primary hepatocytes were isolated, plated onto 12-well plates pre-coated with collagen type 1, and cultured in serum-free medium as described.[10]

8

RNA preparation and quantitative real-time PCR (qRT-PCR)

Total RNA from hepatocytes was extracted using TRIzol (peqlab, Erlangen, Germany). Quantitative real-time PCR was performed as described in detail.[10] Primers are listed in Additional file 1: Table S1, ESI†. For normalization, β-actin was amplified from each sample.

RNA-interference

GliI1-, Gli2- and Gli3-, Sufu- and Ptch2-specific siRNAs and respective nonsense control siRNAs were purchased from Invitrogen, Darmstadt, Germany. Transfection of siRNAs (25 nmol for each GLI-factor, 10 nmol for Sufu- and Ptch2-siRNA) was performed as described.[10] Sequences for siRNA primers are listed in Additional file 1: Table S2, ESI†. Changes in gene expression were analysed by qRT-PCR at 24 h, 48 h and 72 h of cultivation.

Fuzzy Logic Gene Regulatory Network Model

For the description of interactions in a gene regulatory network (GRN), a fuzzy logic based network modeling approach is used. The nodes of this network model are represented in a fuzzy inference system (FIS). FIS provides a method to model complex, nonlinear input-output relationships. The structure of a FIS is composed of the input and output variables and the rule base, which represent the input-output relationship. The input variables can be divided into two groups, experimental conditions (cultivation and information about siRNA-knockdown) and current state of gene regulation, i.e. the current gene expression levels. The output is the new state of gene expression.

Membership function of fuzzy variables

Before learning the FIS, missing data were imputed using the nearest-neighbour method.[28] The learning of the membership function of the FIS is done in two steps. First, the measured values for the input / output fuzzy variables (representing the expression intensity of the genes) were clustered using the fuzzy c-means algorithm.[29] The number of clusters for each variable is pre-set to two, labelled by the fuzzy values 'low' and 'high'. Finally, the obtained membership values of each cluster were fitted by a sigmoidal function to get for each fuzzy variable the membership function of the FIS (Fig. S1, ESI†).

Fuzzy rule base

To describe the input-output behavior of the FIS, a fuzzy rule base (FRB) was created (Fig. S4, ESI†). The fuzzy rules of the FRB were learned by the fuzzy *a priori* algorithm[30] using the measured gene expression values. The received rules have the following form:

IF A1 is I1 THEN C1 is O1

where the antecedent A1 and consequent C1 are related to measured fuzzy variables and I1 and O2 are called fuzzy sets (of the above mentioned fuzzy values) that qualify A1 and C1, respectively, in a semi-quantitative manner. Three evaluation criteria of association rules, called 'support', 'confidence' and 'lift' are used to assess the performance of the rules (Fig. S4 ESI†).[31] For the knowledge base only rules with a 'lift' greater than 1 are taken into account. A 'lift' greater than 1 indicates that G1 and G2 appear more often together than randomly expected and that G1 is positively correlated with G2. The 'lift' value of a rule is computed by dividing the 'confidence' by the 'support' of a rule, where the 'support' of the rule is the quotient of the sum of votes satisfying the respective rule (G1→G2) and the total number of measured data. The 'confidence' is the ratio of the sum of votes satisfying the rule (G1→G2) and the 'support' of the antecedent G1.

Fuzzy interference process

To get a computationally efficient solution, the SUGENO [32] method for defuzzification was used. This method uses a linear model of the fuzzy variables of the antecedent to calculate a crisp result value for the consequent. The parameters of the linear model were fitted minimizing the mean square error of the model fit to the measured data of the consequent by using unconstrained nonlinear optimization.[33]

Structure of Gene Regulatory Network Model

In order to identify the most robust network structure different combinations of FIS were generated and analysed using the following procedure. For each network node an ensemble of FISs with all possible combination of input variables are generated. Each system was learned with randomly selected learn data 100 times. Thereby the whole data set was divided into 95% learn data and 5% test data. After each learn process the mean squared error (mse) of the learn and test data was calculated. The system with the smallest average test error for a given combination of input variables is used for the gene regulatory network model. Next, to determine and visualize the kind of regulatory activity between the components of the network, the whole network with high complexity was segregated in a set of state-dependent sub-networks with lower complexity, i.e. for each fuzzy variable and fuzzy value one specific sub-network was generated. Each sub-network comprises edges valid and relevant only for the respective

condition. The total number of state-dependent sub-networks is two times that of the number of genes, since it discriminates for each gene between the fuzzy values 'low' and 'high' for the expression level (Fig. S3, ESI†). The kind of the regulation was derived from the rule set of the state-dependent sub-network. The interactions without contradiction in the individual state-dependent networks were included in the resulting synopsis of the network.

Statistical Analysis

Values are expressed as means ± standard deviation of the mean (SEM) and sample size (n) is given in each figure or table. Statistical evaluation was performed using the Wilcoxon matched pairs test for the siRNA experiments in cultured hepatocytes. The null hypothesis was rejected at the $p<0.05$ (*) and $p<0.01$ (**) levels.

Acknowledgements

Notes

* contributed equally to this manuscript

[a] Leibniz Institute for Natural Product Research and Infection Biology – Hans Knoell Institute, Beutenbergstr. 11a, 07745 Jena, Germany

[b] Institute of Biochemistry, Faculty of Medicine, University of Leipzig, Johanissallee 30, 04103 Leipzig, Germany

† Electronic supplementary information (ESI) available. Fig S1, S2, S3, S4, S5 and Add file 1 (tables)

References

1   M. T. Barakat, E. W. Humke and M. P. Scott, *Trends Mol. Med.*, 2010, **16**, 337–348.

2   J. Jiang and C.-C. Hui, *Developmental Cell*, 2008, **15**, 801–812.

3   C.-C. Hui and S. Angers, *Annu. Rev. Cell. Dev. Biol.*, 2011, **27**, 513–537.

4   R. Teperino, F. Aberger, H. Esterbauer, N. Riobo and J. A. Pospisilik, *Seminars Cell Dev. Biol.*, 2014, **33**, 81–92.

5   J. Briscoe and P. P. Thérond, *Nature reviews. Mol. Cell Biol.*, 2013, **14**, 416–429.

6   D. J. Robbins, D. L. Fei and N. A. Riobo, *Sci. Signal.*, 2012, **5**, re6.

7   a) Y. Pan, C. B. Bai, A. L. Joyner and B. Wang, *Mol. Cell. Biol.*, 2006, **26**, 3365–3377; b) B. Wang, J. F. Fallon and P. A. Beachy, *Cell*, 2000, **100**, 423–434;

8   H. Sasaki, Y. Nishizaki, C. Hui, M. Nakafuku and H. Kondoh, *Development*, 1999, **126**, 3915–3924.

9   Ruiz i Altaba, A, *Development*, 1999, **126**, 3205–3216.

10 M. Matz-Soja, S. Aleithe, E. Marbach, J. Böttger, K. Arnold, W. Schmidt-Heck, J. Kratzsch and R. Gebhardt, *Cell Commun.. Sig.*, 2014, **12**, 11.

11 Ruiz i Altaba, A, *Development*, 1998, **125**, 2203–2212.

12 D. Buscher and U. Ruther, *Dev. Dyn.*, 1998, **211**, 88–96.

13 A. McDermott, M. Gustafsson, T. Elsam, C.-C. Hui, Emerson, Charles P Jr and A.-G. Borycki, *Development*, 2005, **132**, 345–357.

14 E. R. Boykin and W. O. Ogle, *Mol. BioSyst.*, 2010, **6**, 1993–2003.

15 M. Matz-Soja, A. Hovhannisyan and R. Gebhardt, *Med. Hyp.*, 2013, **80**, 589–594.

16 R. Gebhardt and M. Matz-Soja, *World J. Gastroenterol.*, 2014, **20**, 8491–8504.

17 R. Gebhardt and A. Hovhannisyan, *Dev. Dyn.*, 2010, **239**, 45–55.

18 J. K. Sicklick, Y.-X. Li, S. S. Choi, Y. Qi, W. Chen, M. Bustamante, J. Huang, M. Zdanowicz, T. Camp, M. S. Torbenson, M. Rojkind and A. M. Diehl, *Lab. Invest.*, 2005, **85**, 1368–1380.

19 Y. Jung, K. D. Brown, R. P. Witek, A. Omenetti, L. Yang, M. Vandongen, R. J. Milton, I. N. Hines, R. A. Rippe, L. Spahr, L. Rubbia-Brandt and A. M. Diehl, *Gastroenterology*, 2008, **134**, 1532–1543.

20 C.B. Bai, A.L. Joyner, Development, 2001, **128**, 5161-5172.

20 U. Laresgoiti, A. Apraiz, M. Olea, J. Mitxelena, N. Osinalde, J. A. Rodriguez, A. Fullaondo and A. M. Zubiaga, *Nucleic Acids Res.*, 2013, **41**, 10185–10198.

21 K. Daily, V. R. Patel, P. Rigor, X. Xie and P. Baldi, *BMC Bioinformatics*, 2011, **12**, 495.

22 J. M. Suh, X. Gao, J. McKay, R. McKay, Z. Salo and J. M. Graff, *Cell. Metab.*, 2006, **3**, 25–34.

23   J. A. Pospisilik, D. Schramek, H. Schnidar, Cronin, Shane J F, N. T. Nehme, X. Zhang, C. Knauf, P. D. Cani, K. Aumayr, J. Todoric, M. Bayer, A. Haschemi, V. Puviindran, K. Tar, M. Orthofer, G. G. Neely, G. Dietzl, A. Manoukian, M. Funovics, G. Prager, O. Wagner, D. Ferrandon, F. Aberger, C.-C. Hui, H. Esterbauer and J. M. Penninger, *Cell*, 2010, **140**, 148–160.

24   W. Zhang, J.-S. Kang, F. Cole, M.-J. Yi and R. S. Krauss, *Developmental Cell*, 2006, **10**, 657–665.

25   B. Stecca and Ruiz I Altaba, Ariel, *J Mol Cell Biol*, 2010, **2**, 84–95.

26   F. Aberger and Ruiz I Altaba, Ariel, *Semin. Cell Dev. Biol.*, 2014, **33**, 93–104.

27   O. Troyanskaya, M. Cantor, G. Sherlock, P. Brown, T. Hastie, R. Tibshirani, D. Botstein and R. B. Altman, *Bioinformatics*, 2001, **17**, 520–525.

28   J. C. Bezdek, *Pattern recognition with fuzzy objective function algorithms*, Plenum Press, New York, ©1981.

29   C. M. Kuok, A. Fu and M. H. Wong, *SIGMOD Rec.*, 1998, **27**, 41–46.

30   H. Chen and C. C. Yang, eds., *ACM SIGKDD Workshop*.

31   T. Takagi and M. Sugeno, *IEEE Trans. Syst., Man, Cybern.*, 1985, **SMC-15**, 116–132.

32   J. C. Lagarias, J. A. Reeds, M. H. Wright and P. E. Wright, *SIAM J. Optim.*, 1998, **9**, 112–147.