Tissue-specific Disallowance of Housekeeping Genes: the other Face of Cell Differentiation

running head: Tissue-specific Disallowance of Housekeeping Genes

Lieven Thorrez[1,2,4], Ilaria Laudadio[3], Katrijn Van Deun[4], Roel Quintens[1,4], Nico Hendrickx[1,4], Mikaela Granvik[1,4], Katleen Lemaire[1,4], Anica Schraenen[1,4], Leentje Van Lommel[1,4], Stefan Lehnert[1,4], Cristina Aguayo-Mazzucato[5], Rui Cheng-Xue[6], Patrick Gilon[6], Iven Van Mechelen[4], Susan Bonner-Weir[5], Frederic Lemaigre[3], and Frans Schuit[1,4,$]

[1] Gene Expression Unit, Dept. Molecular Cell Biology, Katholieke Universiteit Leuven, 3000 Leuven, Belgium
[2] ESAT-SCD, Department of Electrical Engineering, Katholieke Universiteit Leuven, 3000 Leuven, Belgium
[3] Université Catholique de Louvain, de Duve Institute, 1200 Brussels, Belgium
[4] Center for Computational Systems Biology, Katholieke Universiteit Leuven, 3000 Leuven, Belgium
[5] Section of Islet Transplantation and Cell Biology, Joslin Diabetes Center, Harvard University, Boston, MA 02215, US
[6] Unité d'Endocrinologie et Métabolisme, University of Louvain Faculty of Medicine, 1200 Brussels, Belgium

[$] To whom correspondence should be addressed. Email: frans.schuit@med.kuleuven.be
Phone : +32 16 347227 , Fax : +32 16 345995

Abstract

Specialized tissues are usually perceived as being the developmental endpoint of a process that hinges on expression of tissue-specific genes. Here we propose that specialization is also supported by tissue-specific repression of particular housekeeping genes whose product would interfere with or even destroy the specialized tissue function. This newly discovered phenomenon that we call "disallowance" has specific examples in most human and mouse tissues. We developed statistical tools to retrieve such genes from multi-tissue mRNA expression databases. Furthermore, we show that disallowance is established during fetal or postnatal maturation of differentiating tissues via mechanisms involving histone H3 lysine 27-trimethylation, loss of H3 lysine 9-acetylation and binding of tissue-specific microRNAs to the 3' untranslated region. With two detailed examples in metabolism, exercise-induced inappropriate insulin release and interference with ketogenesis in liver, we illustrate that disallowance of genes is important for tissue function.

Author Summary

Differentiation of specialized tissues has been explained by expression of tissue-specific genes. Here we explain how specialization is also achieved by tissue-specific repression of certain housekeeping genes, a new phenomenon we propose to name disallowance. We depart from a novel probabilistic approach to identify disallowed genes in adult mouse tissues, demonstrate that their repression is established during tissue maturation and show that disallowance is achieved via chromatin modification and microRNAs. With two metabolic examples we provide a rationale why specialized tissues avoid certain processes occurring in all other tissues, as allowance would compromise their function and can result in pathology.

Introduction

Expression of housekeeping genes is generally believed to be required for the maintenance of general tasks that are common to all cells. A textbook example is that all living cells continuously need to produce sufficient ATP to stay alive. Most ATP is produced via oxidative phosphorylation in the mitochondrial inner membrane, a process requiring oxygen. When oxygen supply is low, such as during ischemia, cells can survive through ATP production via anaerobic glycolysis, which depends on lactate dehydrogenase (LDHA) (reducing pyruvate to lactate) and members of the monocarboxylic acid transporter (MCT) protein family to transport lactate and pyruvate across the plasma membrane. A curious deviation from this principle is that glycolysis in rat pancreatic insulin-producing beta cells is exclusively aerobic [1]. This exceptional metabolic behavior may be explained by the fact that beta cells in the islets of Langerhans are intensively vascularized,[2,3] so that ischemia is very unlikely. Biochemically, beta cells have vanishingly low expression levels of both the lactate/pyruvate transporter MCT1 [4] and lactate dehydrogenases [1,5]. Better insight into the rationale of avoidance of anaerobic glycolysis was provided by the discovery of physical exercise-induced hyperinsulinism, a Mendelian dominant disease that is caused by gain-of-function mutations of the MCT1 promoter [6]. Patients with this disease secrete insulin when plasma pyruvate rises, i.e. after physical exercise, thereby reducing glucose availability exactly when glucose is needed, resulting in hypoglycemia [7]. Beta cells should secrete insulin only in response to rising glucose levels after a meal, and pyruvate is an intermediate product in the glucose sensing. When MCT1 is expressed, lactate and pyruvate generated by muscles during physical exercise can enter the beta cells and disrupt the correct glucose sensing. This is dangerous to the organism since insulin secretion at already low blood glucose levels would lower blood glucose to levels that disturb brain

functioning and can even cause irreversible brain damage. This example illustrates that better knowledge of expression of housekeeping genes is needed to understand tissue function. We specifically asked if such repression of the housekeeping genes *Mct1* (official gene name *Slc16a1)* and *Ldha* is a typical phenotypic trait of the differentiated beta cell, and how this repression is achieved during beta cell differentiation.

Differentiation of cells requires the successful unfolding of a developmental program that depends on multiple interactions between the products of gene regulatory networks. It is generally assumed that the most important result of this program is the restricted use of genes that mediate cell-specific functional tasks [8]. In this study, we show for the first time that a developmentally regulated process in pancreas and in liver involves the cell-selective inactivation of genes that are expressed everywhere else, a phenomenon we call "disallowance".

Results

Identification of repressed genes

Based on a multi-tissue mRNA expression profiling in mouse, we noted for certain genes a strongly decreased or even undetectable expression in one specific tissue, contrasting with expression in all other tissues. In a first analysis we scored mRNA transcript abundance in pancreatic islets as compared to the other tissues in the panel. In addition to the expected existence of islet-specific genes (significantly higher in islets than in each of the other tissues), we found evidence for genes of which expression was significantly lower in islets than in each of the other tissues, although they were less in number than the tissue-specific genes (Fig. 1A). This phenomenon was not restricted to islets but was also observed with other tissues, as illustrated by comparing liver (baseline) to the panel of other tissues (Fig. 1A). We next established a statistical framework to screen systematically for housekeeping genes that are specifically lower in one tissue as compared to each of the others (Fig. 1B). Both ANOVA and multiple comparison procedures (e.g. Dunnett or Bonferroni) are unsuitable because a significant result means that at least one pair of tissues differs significantly in expression, while we wanted evidence that expression in the baseline tissue is lower than in *all* other tissues. Therefore, we applied the intersection-union test [9,10](explained in greater detail in Materials and Methods). We did so for each of the 17,334 unique genes that are represented in the mRNA expression analysis and with each of the 21 tissues in turn acting as the baseline. This analysis largely confirmed the genes that were found repressed in pancreatic islets and liver in the sum score analysis (Fig. 1A), excluding the possibility that the outcome was an artifact of the filtering strategy. We used information from redundant probe sets (present for >7750 genes) to reduce genes found repressed due to a probeset downstream of an alternative polyadenylation site (example in Fig. S1). The distribution of the

total number of 1074 repressed genes over the 21 mouse tissues is shown in Fig. 1C and a full list is provided in Supporting Information (Table S1). Testis accounts for about half of the genes that were selected as being tissue-specifically repressed. We tested this set of 1074 repressed genes for associations with cellular and molecular functions (Fig. 1D) and found significant enrichment for genes involved in processes determining tissue development and function. Further analysis of the degree of suppression in the concerning tissue indicated that some genes are still expressed, though at a significantly lower level than in the other tissues, while expression of certain repressed genes was not or barely detectable. An example of this distinction is shown for genes found repressed in islets (Fig. S2). When adding a second criterion for strong repression in the target tissue (at least 30-fold lower compared to the mean of the tissue panel), we identified 35 strongly repressed genes, seven of which are repressed in islets (Table S2).

Disallowance in islets of Langerhans

One of these seven genes is *Mct1*, encoding the monocarboxylate transporter MCT1 which mediates the transport of lactate and pyruvate across cell membranes and is present in all tissues except adult beta cells [6,11]. RNA from pancreatic islets mainly (±80%) derives from beta cells. These are specialized cells as they are unique to produce, store and secrete insulin in a manner that is highly dependent on the feeding status and intracellular glucose metabolism [12]. Recently, a disease with autosomal dominant Mendelian inheritance was discovered, in which patients develop hypoglycemia after physical exercise [7]. This impaired metabolic control is caused by inappropriate insulin release as a result of expression of *Mct1* in beta cells. This example illustrates that loss of metabolic control can be caused by a failure of specific repression of a housekeeping gene in the beta cell. Therefore we propose to name these specifically repressed housekeeping genes "*disallowed*". Besides *Mct1*, *Ldha* was the most strongly repressed

gene in islets. Adult beta cells have a strictly aerobic mode of glycolysis with low rates of pyruvate reduction and high rates of anaplerosis and pyruvate oxidation [13]. Therefore, expression of pyruvate carboxylase (*Pc*), which catalyzes the anaplerotic conversion of pyruvate into oxaloacetate [1,14] is high in beta cells as compared to other non-lipogenic and non-neoglucogenic tissues (Fig. 2A,B). LDHA catalyzes the conversion of lactate to pyruvate (Fig. 2B). This reaction should not occur in beta cells since lactate should not enter metabolism even if it would enter the beta cell.

Cellular heterogeneity affects detection of disallowed genes

Because pancreatic islets are composed of several endocrine and non-endocrine cell types, the possibility exists that the data obtained from whole islets even underestimate the degree of repression in the beta cell. This was further investigated by comparing mRNA abundance of repressed genes in whole islets with acinar cells and FACS-purified beta cells (>99% pure β-cells)(Fig. S3A,B). Expression of *Mct1* and *Mgst1* (microsomal glutathione S-transferase 1) was determined by microarray (Fig. S3A) and Q-PCR (Fig. S3B). For both genes, an even lower expression can be observed in purified beta cells versus whole islets. This indicates that for a number of repressed genes, a low-level expression may be due to contaminating cell types in which the gene is not repressed. This contamination is in the order of 1-5%, consistent with our observation that the expression of the acinar transcription factor *Ptf1a* in isolated islets is at 2% of expression in isolated acini (Fig. S3A).

Disallowance of Oxct1 in liver

Because both *Mct1* and *Ldha* were found to be strongly co-repressed and as this repression is important for beta cell function [6,11], it is conceivable that the other 33 deeply

repressed genes we found in this study are disallowed for other reasons and have biological and/or medical interest as well. As an example, we further examined repression of 2-oxoacid CoA transferase mRNA (*Oxct1*) in mouse liver. The encoded protein is an enzyme that plays a key role in ketone body degradation by catalyzing the reversible transfer of coenzyme A from succinyl-CoA to acetoacetate. Ketone body degradation provides an alternative energy source to many tissues during fasting (Fig 2D). Disallowance in liver seems appropriate as this tissue is specialized in ketone production in the fasted state to support metabolism of other tissues, in particular the brain [15]. In fact, the specific repression of *Oxct1* in the mouse adult liver mirrors the expression of 3-hydroxy 3-methylglutaryl-CoA synthase 2 (*Hmgcs2*), which is required for ketogenesis and is highly expressed in adult liver, even in the fed state (Fig. 2C). Other enzymes necessary for ketogenesis (such as *Acat1-3* and *Hmgcl*) were also highly expressed in adult liver (data not shown).

Disallowance is established during tissue maturation

Next, we analyzed whether the tissue-specific disallowance of MCT1, LDHA and OXCT1 was linked to the development of islets and liver. In the pancreas, the beta cells expand rapidly in the first month of life [16] and only become truly functional in their response to glucose about weaning, when insulin release becomes controlled by oral food intake. As is shown in Fig. 3A neonatal rat islets progressively induce expression of typical beta cell genes during this period. For instance, the expression of insulin (*Ins2*) as well as pyruvate carboxylase rises five- to tenfold. In contrast, during neonatal beta cell maturation the mRNA expression of both *Mct1* and *Ldha* decreases progressively (Fig. 3B). In addition, *c-Maf* which is another deeply repressed gene in pancreatic islets, displays a similar neonatal time-dependent down regulation (Fig. S4A). This phenomenon is not restricted to maturating pancreatic islets, but also exists in liver. In

mouse embryonic and neonatal liver we observed a parallel pronounced upregulation of *Hmgcs2* and progressive loss of expression of *Oxct1* (Fig. 3C,D), again suggesting that during differentiation cell-specific genes are turned on while repressed housekeeping genes are turned off. The reduction of *Hmgcs2* mRNA in rat liver after weaning (Fig. 3C) has been reported before in rodents [17].

Mechanisms for disallowance

Several mechanisms may be responsible for the observed tissue-specific gene repression. First, we tested whether the differentiation-dependent repression of certain housekeeping genes is dependent on epigenetic changes in chromatin structure. Histone modifications are well known to be stably introduced during cellular differentiation [18] and profoundly alter the transcriptional activity of genes. Histone H3 trimethylation at lysine 27 (H3K27me3) is associated with dense packaging of the nucleosome and low transcriptional activity, while acetylation at lysine 9 (H3K9ac) is associated with loosely packaged DNA and enhanced transcriptional activity. To examine a relationship between nucleosome structure and tissue-specific repression, we tested H3K9ac and H3K27me3 of the *Mct1* and *Ldha* genes in chromatin from mouse islets and liver. As is shown in Figure 4A, H3K27me3 content of chromatin associated with the *Mct1* and *Ldha* genes was higher in islets than in liver. For the *Mct1* gene in islets we found a gradient of H3K27me3 along the length of the gene, with the highest level around the promoter (data not shown). In contrast, the signal for H3K9ac was significantly higher in liver than in islets for both genes (Fig. 4A). Therefore, the low *Mct1* and *Ldha* mRNA levels in mature islets correlate to epigenetic silencing that is expected to be established when beta cells mature postnatally. Also *c-Maf* and *Mgst1* were associated with dense packaging (H3K27me3) in pancreatic islets, in contrast to liver where an open nucleosome structure (H3K9ac) was predominant (Fig. S4B,C).

We next looked at the *Oxct1* gene, which is repressed in the liver and highly expressed in many other tissues such as heart, kidney, muscles and brain (Fig. 2C). In line with the results found in islets, nucleosomal chromatin around the promoter of *Oxct1* was more trimethylated and less acetylated in liver than in brain (Fig. 4A). Together, these data suggest that epigenetic changes in chromatin structure can explain in part the selective disallowance of housekeeping genes in specialized cell types.

Do other levels of control, besides epigenetic suppression, exist to induce disallowance? A possible mechanism could occur via microRNAs, short non-coding RNAs with widespread influence on mRNA stability and mRNA translation [19,20]. In line with this idea, the 3' untranslated region (UTR) of mouse *Oxct1* mRNA is a predicted target for miR-122 (Fig. S5A), which is specifically expressed in mouse liver and which is the most abundant liver microRNA species (Fig. S5B,C)[21,22]. During the course of mouse embryonic liver differentiation, miR-122 expression increased 12-fold between E15.5 and birth (Fig. 4B). In parallel, *Oxct1* expression gradually decreased from E15.5 until 2 weeks after birth (Fig. 3D). In order to assess whether the expression of miR-122 is responsible for repression of *Oxct1* mRNA, we performed experiments with a bipotential mouse embryonic liver (BMEL) cell line, which is derived from e14.5 hepatoblasts and can be differentiated to hepatocyte-like cells [23]. During differentiation of cultured BMEL cells, miR-122 expression increased 9-fold (data not shown). Expression of miR-122 could be inhibited by the addition of miR-122 antagomir (Fig. 4B), which was specific since a mutated antagomir abolished the inhibition (Fig. 4B) and a nontarget microRNA (miR-20) was not affected by either antagomirs (data not shown). In differentiated BMEL cells, *Oxct1* expression increased 1.6-fold by adding miR-122 antagomir, while the mutated control antagomir did not alter *Oxct1* expression (Fig. 4B). Next, we investigated microRNA expression in pancreatic islets, where miR-124 is most abundant (Fig. S6A). This microRNA has multiple

predicted target sites on the 3'UTR of *Mct1* (Fig. S6B) which are conserved among human, rat, mouse, dog and chicken. In addition to miR-124 binding sites, other microRNAs abundant in islets are also predicted to recognize the *Mct1* 3'UTR (Fig. S6B). This indicates that also for *Mct1* an additional layer of posttranscriptional control may be present to ensure an appropriate repression in islets. More generally, if repressed genes are downregulated by microRNAs, we expect to find binding sites for microRNAs in the 3'UTR of these genes. We would then expect these binding sites to correspond to microRNAs which are abundant in the specific tissue. Therefore, we checked for overrepresentation of tissue-specific microRNA binding sites in the 3' UTRs of the repressed genes for the particular tissue (where such info was available). For several tissues, we found such strong overrepresentation (Table S3), indicating that at least a subset of the repressed genes can be downregulated by tissue-specific and abundant microRNAs.

Evolutionary conservation

We verified the expression of the 35 most strongly repressed mouse genes in human tissue microarrays. A dataset was used encompassing 17 human tissues equivalent to the murine tissues. Since not all murine tissues were represented and not all murine genes have a human orthologue, only 27 deeply repressed genes could be checked. Of these, 15 had the lowest or background expression in the same human tissue as the murine tissue in which it was found to be repressed (Table S4). Taken together with a similar repression of *Mct1* and *Ldha* in rat islets (Fig. 3B) these results suggest that tissue-specific repression of most of the genes identified in this study is evolutionary conserved.

Discussion

Unlike what is sometimes claimed, expression of housekeeping genes can widely differ among different tissues [24]. In the present article, we propose that profound repression of certain housekeeping genes in specialized tissues is possible and necessary and that this repression is the opposite of expression of cell-specific genes. Moreover, we propose that this severe repression is, in fact, a newly-discovered facet of cellular differentiation. We support this novel concept with different arguments that will be discussed in detail below. First, some housekeeping genes can be turned off completely in a tissue and this type of repression is conserved in several mammalian species such as mice, rats and human. Second, at least two different gene regulatory mechanisms have developed to repress these genes in a single cell type. Finally, failure to achieve this repression during tissue formation may be a mechanism of disease.

Tissue-specific gene repression is known to occur during development and requires the spatial and temporal organization of an interacting network of transcriptional repressors and activators. Since specific absence of one gene in a particular tissue has not been systematically studied so far, few genes are known to be expressed as such. One well-studied example is the repressor element 1 silencing transcription factor (REST), which prevents the transcription of neuronal genes in non-neuronal tissues. REST is repressed in differentiating neurons but also in pancreatic islets [25] and was therefore not detected in our screen.

Immature cells that are the direct precursors of the studied differentiated cells may not have fully exercised the level of disallowance; therefore, specific elements that are induced during the developmental program are responsible for repression in the differentiated cell. Chromatin condensation and the resulting gene silencing plays a key role in differentiation. A hallmark for transcriptional silencing is trimethylation of lysine27 on histone H3 (H3K27me3), a

process that is catalyzed by the methyl transferase activity of the Polycomb complex PCR2 [26,27] and antagonized by the demethylases Utx and Jmjd3 [28–31]. Polycomb target genes are preferentially activated upon induction of differentiation [32]. The group of tissue-specifically repressed housekeeping genes we present in this paper is expressed in embryonic stem cells and is distinct from the genes associated with polycomb complexes in embryonic stem (ES) cells [32]. Polycomb target genes are preferentially upregulated during ES cell differentiation, whereas disallowed genes are downregulated during differentiation. Certain promoters in ES cells carry a bivalent mark, with large regions of H3K27me3 harboring smaller regions of H3 lysine 4 methylation [33]. This mark serves to poise key developmental genes for lineage-specific activation or repression [34]. Most promoters of the repressed genes in this paper were found to be negative for H3K27me3 in ES cells [34]. These data fit with our observations that the expression of the disallowed genes is turned off only during late fetal liver maturation and early postnatal islet maturation. Cell-specific combinations of transcription factors might activate the PCR2 complex specifically to repress expression of disallowed genes during tissue differentiation.

By applying a custom statistical intersection-union test [9,10] on a set of mouse tissue microarrays, we identified genes with significantly lower expression in one tissue versus all other (20) tissues. Many of these genes are associated with tissue development and function. For a few examples we provide a logical explanation for their repression, since presence of the gene product would interfere with correct tissue function. Repression is established during tissue maturation and is subsequently maintained through adulthood. This mechanism is maintained by association with silent chromatin and expression of microRNAs targeting the transcript 3'UTR. Our present identification approach has two technical limitations. First, the 3' mRNA expression microarrays we used to identify repressed genes have probesets pre-designed mostly towards the 3' end of the

transcript. Therefore, if the transcript has an alternative termination site upstream of the probeset-binding site that is used in only one tissue and no other probesets are available for this gene, a false-positive hit may occur (example Fig. S1). Second, our screen is based on mRNA extracts from whole tissues that are composed of a number of different cell types. This mixed population may result in the situation that genes that are specifically repressed in a certain cell type are not detected in the screen because other cell types in the same tissue do express the gene(s) and thus mask the repression. If this phenomenon were re-examined at the level of primary differentiated cell types, a more extensive degree of repression may be found. The fact that we have underestimated the repression of some of the islet genes is illustrated by the analysis of highly-purified islet beta cells (Figure S3).

This study is to our knowledge the first to propose that disallowance of housekeeping genes is required for tissue function. A known example of a human disease illustrates that ignoring this disallowance can destroy specific tissue function [6]. We focused mainly on two examples of energy metabolism to illustrate this novel concept, but other examples outside the realm of metabolism likely exist. For instance in spleen, a lymphoid organ, ring finger protein 128 (*Rnf128*), also known as "gene related to anergy in lymphocytes" (GRAIL), is deeply repressed (expression is 59-fold lower as compared to the average of other tissues, Table S2). *Rnf128* encodes a ubiquitin ligase and is associated with T cell suppression [35,36]. Disallowance of *Rnf128* in spleen and expression in other tissues may help to shape the balance in central and peripheral T-cell reactivity.

In summary, tissue-specific disallowance of housekeeping genes supports the specialized function of differentiated tissues. Profound repression is maintained through a multi-tier mechanism comprising chromatin remodeling and microRNAs. We propose that this novel phenomenon has widespread consequences in cell biology and can lead to disease when disrupted.

Materials and Methods

Preparation of tissues and purified cells

All experiments based upon laboratory animals were approved by committees for animal welfare at the Katholieke Universiteit Leuven and Harvard University. Following tissues were hand dissected from 10-12 week old C57Bl6 mice: liver, gastrocnemius muscle, brain, heart, adrenal gland, eye, small intestine, thymus, epididymal adipose tissue, pituitary gland, kidney, parotid gland, spleen, lung, bone marrow, testis, and seminal vesicles (males); ovary and placenta (females). Fetal tissue was isolated at day 16. Tissues were rinsed in phosphate-buffered saline, frozen in liquid nitrogen and stored at -80°C. Islets of Langerhans were isolated from the male and female adult (12 wk) mouse pancreata after injecting collagenase solution into the pancreatic duct. For adult rats, islets from one pancreas were used as a sample. For liver time series experiments, livers were isolated from CD1 mice, stored at 4°C overnight in RNAlater (Ambion) and then stored at -80°C. Mouse pure β-cells were purified from islets of RIPYY mice expressing EYFP under the rat insulin promoter, i.e. specifically in β-cells [37]. Additional details on tissue preparation and RNA extraction in Supporting Information.

mRNA expression analysis via microarray

Cellular mRNA was reverse transcribed into cDNA (SuperScript Choice System Invitrogen, Carlsbad, CA) using oligo-dT primers and a T7 RNA polymerase promoter site. Two μg of total RNA was used to prepare biotinylated cRNA with IVT labeling kit (Affymetrix, Santa Clara, CA) according to the Genechip expression analysis technical manual 701025 Rev.5, except for islets, adrenal gland and pituitary gland where 1 μg of total RNA was used. The concentration of labeled cRNA was measured using the NanoDrop ND-1000 spectrophotometer. Labeled cRNA

was fragmented in a fragmentation buffer during 35 min at 94°C. The quality of labeled and fragmented cRNA was analyzed using the Agilent bioanalyzer 2100. Fragmented cRNA was hybridized to mouse 430 2.0 arrays (Affymetrix) during 16h at 45°C. The arrays were washed and stained in a fluidics station (Affymetrix) and scanned using the Affymetrix 3000 GeneScanner.

We used a microarray dataset consisting of 21 different murine tissues, with 3-5 replicates for each tissue (in total 69 microarrays). Quality controls of the arrays were according to manufacturer's criteria. All CEL files were analyzed using GCOS (Affymetrix GeneChip Operating Software) and the affy library [38] of the BioConductor project [39]. Data were processed with the Robust Multichip Average (RMA) algorithm with default parameters (RMA background correction, probe-level quantile normalization and average difference summarization). We independently applied the MAS5 algorithm using global scaling to 150. All data were log2 transformed for normalization and all further data analysis was performed on log2 transformed data. The data files are accessible through the NCBI Gene Expression Omnibus (GEO, http://www.ncbi.nlm.nih.gov/geo/) with accession number GSE9954. For the human data, we selected 69 arrays from GSE7307, representing 17 human tissues (3-5 arrays/tissue) equivalent to the mouse tissues.

Statistical analysis

To reduce multiple testing, we defined a core of 17334 probesets, each corresponding to a unique gene symbol, where only the probeset with the highest average expression was selected when multiple probesets were targeting the gene. To find those genes that are significantly repressed in a particular baseline tissue compared to *each* of the other 20 tissues, we initially used a sum score, which is the summation of the number of times the gene is more (+1), less (-1) or equally

(0) expressed in the baseline tissue versus the 20 other members in the tissue panel, using unpaired Student's t-tests and P<0.001 as a threshold for significance. To control the type I error, we improved this test statistic by applying the testing procedure described by Berger [9], also known as the intersection-union test [10]. This tests the null hypothesis that at least one other tissue scores lower than or equal to the baseline against the alternative that expression in each of the other tissues is larger than in the baseline tissue (see Supporting Information). When available, we used redundant probesets as an additional control by also requiring the redundant probesets to have the lowest expression in baseline tissue when not in background range. For significance testing in time course experiments, we used one-way ANOVA with Dunnett post-hoc testing. *: $p<0.05$, **: $p<0.01$, ***: $p<0.001$. We also used contrasts to test for increasing or decreasing trends. The remaining comparisons were based on t-tests with the standard error from ANOVA. All statistical analysis was performed in SAS 9.2.

Bioinformatics

Benjamini Hochberg corrected p-values for associations with molecular and cellular functions were calculated with Ingenuity Pathway Analysis. Prediction of microRNA targets in 3'UTRs was based on context scoring [40] performed with TargetScanMouse 5.0. We identified the most abundant microRNAs in various mouse tissues via a public microRNA expression database [41] or literature [42–44]. Overrepresentation of repressed genes in the target list was assessed with the hypergeometric test.

Chromatin immunoprecipitation

ChIP was performed according to a modified protocol of Upstate (Millipore) EZ-ChIP. Detailed method provided in Supporting Information. Precleared chromatin was incubated overnight at 4°C using rabbit anti-acetyl-histone H3 (Lys9), rabbit anti-trimethyl-histone H3 (Lys27) or normal rabbit IgG (Upstate). Antibody-chromatin complexes were collected by adding blocked protein A-TSK beads for 1 h at 4°C and washed. Immune complexes were eluted in 1% SDS, 0.1M $NaHCO_3$ at room temperature. Samples were treated with RNase A and RNase T1 (Fermentas) for 30 min at 37°C and 0.6 U proteinase K (Fermentas) for 1 h 30 min at 45°C. DNA was purified using spin columns (Geneclean Turbo; Qbiogene) and eluted in 60 µl DNase-free water. ChIP samples were analyzed with probe-based quantitative PCR (Rotor-Gene 3000; Corbett). PCRs were carried out in a 25 µl reaction volume containing ABsolute QPCR mix (Abgene), 1.5 µl DNA, 300 nM or 900 nM forward and reverse primer and 50 nM dual-labeled fluorogenic probe. Primers and probes were obtained from Sigma-Aldrich. Amplification conditions: 15 min 95°C, 45 cycles of 20 s 95°C and 45 s 60°C. Enrichment of DNA was calculated using the equation: 2^(CtInput-CtAb), where Ct is mean threshold cycle of PCR done in triplicates. Input is 1% input sample and Ab is specific antibody or IgG sample.

Cell culture and antagomir treatment

BMEL cells were cultured as previously described *[45]* plus glutamine and hepatocyte-like differentiation was obtained by inducing the formation of floating aggregates of BMEL cells [23]. For miR-122 inhibition we used antagomirs (Eurogentec). Antagomir sequences are in Supporting Information. BMEL cells cultured in aggregates were exposed to 10 µg/ml antagomirs and RNA was extracted after 3 days.

Reverse transcription and Q-PCR

Rat islets

Reverse transcription was done in 25 µL containing 500 ng total RNA, 5 µL superscript buffer, 0.1 M dithiothreitol (DTT), 50 ng random hexanucleotide primers, 10 mM dNTP, 200 units of RNAsin, and 200 U Superscript II reverse transcriptase (Invitrogen), incubated for 10 minutes at 25°C, 50 minutes at 42°C and 15 minutes at 72°C. cDNA was diluted to 3 ng/µl in nuclease-free water and stored at -20°C. Q-RT-PCR was performed using the ABI7000 (Applied Biosystems, Foster City, CA). Primers were synthesized by MWG-Biotech (High Point, NC) and specificity was confirmed by dissociation curve analysis. Reaction mix contained 10 µL SYBR green master mix (Applied Biosystems), 1 µL of each 5′ and 3′ oligonucleotides (10 pmol/µL) and 1 µL cDNA. Reaction conditions for PCR were 15 s at 95°C, 30 s at 55-58 °C (depending on the primers) and 30 s at 72 °C for 40 cycles. After normalization of the gene of interest to a control gene (*S25*) the comparative CT (threshold cycle) method was used to calculate relative gene expression levels for the selected genes.

Mouse islets

Quantitative RT-PCR was performed as earlier described [6]. mRNA expression levels were normalized using polymerase II alpha (*RNA Pol II*). For *Mgst1* a TaqMan Gene Expression Assay (Applied Biosystems) was used.

Mouse liver

Reverse transcription on total RNA was performed in 10-25 µl containing 1X First Strand Buffer (Invitrogen), 10 mM DTT, 200 U M-MLV reverse transcriptase (Invitrogen), 20-40 U RnaseOUT

(Invitrogen), 4 mM dNTPs (Amersham Pharmacia Biotech) and 200-400 ng of random hexamers (Roche) for 10 min at room temperature and 2 h at 37 °C (mRNA) or 1 h at 37°C and 5min at 85°C (*18S* RNA). For miR-122, 50nM stem-loop RT primer (CTCAACTGGTGTCGTGGAGTCGGCAATTCAGTTGAGACAAACAC) was used instead of random hexamers and incubation was 30 min at 16°C, 30 min at 42°C and 5 min at 85°C. QPCR was performed with SYBR Green Master Mix Reagent (Invitrogen) on a IQ cycler (Bio-Rad). *Oxct1* and *Hmgcs2* were normalized to *β-actin* and miR-122 to *18S* RNA.

Reference List

1. Schuit F, De Vos A., Farfari S, Moens K, Pipeleers D, et al. (1997) Metabolic fate of glucose in purified islet cells. Glucose-regulated anaplerosis in beta cells. J Biol Chem 272: 18572-18579.

2. Bonner-Weir S, Orci L (1982) New perspectives on the microvasculature of the islets of Langerhans in the rat. Diabetes 31: 883-889.

3. Jansson L, Hellerstrom C (1981) A rapid method of visualizing the pancreatic islets for studies of islet capillary blood flow using non-radioactive microspheres. Acta Physiol Scand 113: 371-374.

4. Zhao C, Wilson MC, Schuit F, Halestrap AP, Rutter GA (2001) Expression and distribution of lactate/monocarboxylate transporter isoforms in pancreatic islets and the exocrine pancreas. Diabetes 50: 361-366.

5. Sekine N, Cirulli V, Regazzi R, Brown LJ, Gine E, et al. (1994) Low lactate dehydrogenase and high mitochondrial glycerol phosphate dehydrogenase in pancreatic beta-cells. Potential role in nutrient sensing. J Biol Chem 269: 4895-4902.

6. Otonkoski T, Jiao H, Kaminen-Ahola N, Tapia-Paez I, Ullah MS, et al. (2007) Physical exercise-induced hypoglycemia caused by failed silencing of monocarboxylate transporter 1 in pancreatic beta cells. Am J Hum Genet 81: 467-474.

7. Otonkoski T, Kaminen N, Ustinov J, Lapatto R, Meissner T, et al. (2003) Physical exercise-induced hyperinsulinemic hypoglycemia is an autosomal-dominant trait characterized by abnormal pyruvate-induced insulin release. Diabetes 52: 199-204.

8. Davidson EH, Rast JP, Oliveri P, Ransick A, Calestani C, et al. (2002) A genomic regulatory network for development. Science 295: 1669-1678.

9. Berger RL (1982) Multi-Parameter Hypothesis-Testing and Acceptance Sampling. Technometrics 24: 295-300.

10. Berger RL, Hsu JC (1996) Bioequivalence trials, intersection-union tests and equivalence confidence sets. Statistical Science 11: 283-302.

11. Quintens R, Hendrickx N, Lemaire K, Schuit F (2008) Why expression of some genes is disallowed in beta-cells. Biochem Soc Trans 36: 300-305.

12. Matschinsky FM, Collins HW (1997) Essential biochemical design features of the fuel-sensing system in pancreatic beta-cells. Chem Biol 4: 249-257.

13. Schuit FC (1996) Factors determining the glucose sensitivity and glucose responsiveness of pancreatic beta cells. Horm Res 46: 99-106.

14. Hasan NM, Longacre MJ, Stoker SW, Boonsaen T, Jitrapakdee S, et al. (2008) Impaired anaplerosis and insulin secretion in insulinoma cells caused by small interfering RNA-mediated suppression of pyruvate carboxylase. J Biol Chem 283: 28048-28059.

15. Cahill GF, Jr., Veech RL (2003) Ketoacids? Good medicine? Trans Am Clin Climatol Assoc 114: 149-161.

16. Bonner-Weir S (2000) Perspective: Postnatal pancreatic beta cell growth. Endocrinology 141: 1926-1929.

17. Serra D, Bellido D, Asins G, Arias G, Vilaro S, et al. (1996) The expression of mitochondrial 3-hydroxy-3-methylglutaryl-coenzyme-A synthase in neonatal rat intestine and liver is under transcriptional control. Eur J Biochem 237: 16-24.

18. Li E (2002) Chromatin modification and epigenetic reprogramming in mammalian development. Nat Rev Genet 3: 662-673.

19. Selbach M, Schwanhausser B, Thierfelder N, Fang Z, Khanin R, et al. (2008) Widespread changes in protein synthesis induced by microRNAs. Nature 455: 58-63.

20. Baek D, Villen J, Shin C, Camargo FD, Gygi SP, et al. (2008) The impact of microRNAs on protein output. Nature 455: 64-71.

21. Landgraf P, Rusu M, Sheridan R, Sewer A, Iovino N, et al. (2007) A mammalian microRNA expression atlas based on small RNA library sequencing. Cell 129: 1401-1414.

22. Lagos-Quintana M, Rauhut R, Yalcin A, Meyer J, Lendeckel W, et al. (2002) Identification of tissue-specific microRNAs from mouse. Curr Biol 12: 735-739.

23. Strick-Marchand H, Morosan S, Charneau P, Kremsdorf D, Weiss MC (2004) Bipotential mouse embryonic liver stem cell lines contribute to liver regeneration and differentiate as bile ducts and hepatocytes. Proc Natl Acad Sci U S A 101: 8360-8365.

24. Thorrez L, Van Deun K., Tranchevent LC, Van LL, Engelen K, et al. (2008) Using ribosomal protein genes as reference: a tale of caution. PLoS ONE 3: e1854.

25. Martin D, Allagnat F, Chaffard G, Caille D, Fukuda M, et al. (2008) Functional significance of repressor element 1 silencing transcription factor (REST) target genes in pancreatic beta cells. Diabetologia 51: 1429-1439.

26. Vire E, Brenner C, Deplus R, Blanchon L, Fraga M, et al. (2006) The Polycomb group protein EZH2 directly controls DNA methylation. Nature 439: 871-874.

27. Shen X, Liu Y, Hsu YJ, Fujiwara Y, Kim J, et al. (2008) EZH1 mediates methylation on histone H3 lysine 27 and complements EZH2 in maintaining stem cell identity and executing pluripotency. Mol Cell 32: 491-502.

28. Agger K, Cloos PA, Christensen J, Pasini D, Rose S, et al. (2007) UTX and JMJD3 are histone H3K27 demethylases involved in HOX gene regulation and development. Nature 449: 731-734.

29. Lee MG, Villa R, Trojer P, Norman J, Yan KP, et al. (2007) Demethylation of H3K27 regulates polycomb recruitment and H2A ubiquitination. Science 318: 447-450.

30. De Santa F., Totaro MG, Prosperini E, Notarbartolo S, Testa G, et al. (2007) The histone H3 lysine-27 demethylase Jmjd3 links inflammation to inhibition of polycomb-mediated gene silencing. Cell 130: 1083-1094.

31. Lan F, Bayliss PE, Rinn JL, Whetstine JR, Wang JK, et al. (2007) A histone H3 lysine 27 demethylase regulates animal posterior development. Nature 449: 689-694.

32. Boyer LA, Plath K, Zeitlinger J, Brambrink T, Medeiros LA, et al. (2006) Polycomb complexes repress developmental regulators in murine embryonic stem cells. Nature 441: 349-353.

33. Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, et al. (2006) A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell 125: 315-326.

34. Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, et al. (2007) Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature 448: 553-560.

35. Kostianovsky AM, Maier LM, Baecher-Allan C, Anderson AC, Anderson DE (2007) Up-regulation of gene related to anergy in lymphocytes is associated with Notch-mediated human T cell suppression. J Immunol 178: 6158-6163.

36. MacKenzie DA, Schartner J, Lin J, Timmel A, Jennens-Clough M, et al. (2007) GRAIL is up-regulated in CD4+ CD25+ T regulatory cells and is sufficient for conversion of T cells to a regulatory phenotype. J Biol Chem 282: 9696-9702.

37. Quoix N, Cheng-Xue R, Guiot Y, Herrera PL, Henquin JC, et al. (2007) The GluCre-ROSA26EYFP mouse: a new model for easy identification of living pancreatic alpha-cells. FEBS Lett 581: 4235-4240.

38. Gautier L, Cope L, Bolstad BM, Irizarry RA (2004) affy--analysis of Affymetrix GeneChip data at the probe level. Bioinformatics 20: 307-315.

39. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, et al. (2004) Bioconductor: open software development for computational biology and bioinformatics. Genome Biol 5: R80.

40. Grimson A, Farh KK, Johnston WK, Garrett-Engele P, Lim LP, et al. (2007) MicroRNA targeting specificity in mammals: determinants beyond seed pairing. Mol Cell 27: 91-105.

41. Betel D, Wilson M, Gabow A, Marks DS, Sander C (2008) The microRNA.org resource: targets and expression. Nucleic Acids Res 36: D149-D153.

42. Sood P, Krek A, Zavolan M, Macino G, Rajewsky N (2006) Cell-type-specific signatures of microRNAs on target mRNA expression. Proc Natl Acad Sci U S A 103: 2746-2751.

43. Babak T, Zhang W, Morris Q, Blencowe BJ, Hughes TR (2004) Probing microRNAs with microarrays: tissue specificity and functional inference. RNA 10: 1813-1819.

44. Ro S, Park C, Sanders KM, McCarrey JR, Yan W (2007) Cloning and expression profiling of testis-expressed microRNAs. Dev Biol 311: 592-602.

45. Plumb-Rudewiez N, Clotman F, Strick-Marchand H, Pierreux CE, Weiss MC, et al. (2004) Transcription factor HNF-6/OC-1 inhibits the stimulation of the HNF-3alpha/Foxa1 gene by TGF-beta in mouse liver. Hepatology 40: 1266-1274.

Financial Disclosure

This study was supported by research grants from the K.U.Leuven (GOA/2004/11, GOA/2005/04, GOA/2009/10, and CoE EF/05/007), from the Belgian Science Policy (Interuniversity Attraction Poles Program (PAI 6/20 and 6/40), from the D.G. Higher Education and Scientific Research of the French Community of Belgium, by the Juvenile Diabetes Research Foundation (JDRF grants 2006-182, 2007-685) and by the National Institutes of Health (NIH R01 DK 66056). PG is Research Director at the FNRS and IL holds a fellowship from Télévie (Belgium). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interest

None declared.

Abbreviations

H3K27me3	Histone H3 trimethylation at lysine 27
H3K9ac	Histone H3 acetylation at lysine 9
BMEL	Bipotential Mouse Embryonic Liver
UTR	UnTranslated Region

Figure legends

Figure 1. Tissue-specific repression of housekeeping genes. A: Frequency distribution of the sum score for islets (green) and liver (brown) as baseline. Sum score numbers are calculated genewise and are the addition of the number of pairwise t-tests where expression in baseline is significantly lower (-1) or higher (+1) as compared to the other tissues. Liver- and islet-specific genes reside in the category with the highest possible sum score (right), whereas genes of which expression is specifically repressed in islets and liver are found in the category with the lowest possible sum score (left). **B:** Strategy to identify genes that are selectively repressed in one tissue, based on mRNA profiling of 21 mouse tissues. Expression of a gene in a baseline tissue was compared to expression of that gene in all other tissues. When significantly lower compared to all other tissues, this gene was termed repressed in that tissue. This procedure was applied to all genes and repeated for all tissues in the panel. Statistical testing methods are described in materials and methods. **C:** Pie-chart distribution of the number of specifically repressed genes in tissues. **D**: Significant associations of specifically repressed genes (in all tissues) with molecular and cellular functions. The Benjamini-Hochberg corrected P-value threshold of 0.05 is indicated by a dashed line.

Figure 2. Allowed and disallowed genes. A: Microarray expression profiles of pyruvate carboxylase (*Pcx*) versus lactate dehydrogenase A (*Ldha*) which are, respectively highly expressed and deeply repressed in adult mouse pancreatic beta cells. Mean mRNA expression signals ± SEM of 3 to 5 biological replicate microarray experiments. **B:** Glucose levels in the blood are measured by the beta cell through conversion of glucose to pyruvate which is further metabolized in the mitochondria for ATP generation and anaplerosis. Downstream signals lead to

insulin secretion. In many other cells, pyruvate does not necessarily come from glucose uptake. Extracellular pyruvate can be directly taken up by monocarboxylate transporter 1 (*Mct1*) or lactate can be converted to pyruvate by *Ldha*. However, when this happens in beta cells, the glucose sensing mechanism is disrupted, leading to inappropriate insulin release [6]. **C:** Microarray expression profiles of 3-hydroxy 3-methylglutaryl-CoA synthase 2 (*Hmgcs2*) versus 2-oxoacid CoA transferase (*Oxct1*) which are, respectively highly expressed and deeply repressed in adult liver. Mean mRNA expression signals ± SEM of 3 to 5 biological replicate microarray experiments. **D:** In the fasted state, ketone bodies are generated in the liver. The first and rate-limiting enzyme for ketogenesis is *Hmgcs2*. Ketone bodies are transported in the blood and can be used by many organs as an alternative energy source. *Oxct1* plays a central role in extrahepatic ketone body catabolism by catalyzing the reversible transfer of coenzyme A from succinyl-CoA to acetoacetate.

Figure 3. Tissue-specific gene expression and repression of certain housekeeping genes during maturation of rat islets (A,B) and mouse liver (C,D). A: mRNA expression level in extracts from rat pancreatic islets from age 1 day (P1) to 4 weeks (P28) after birth measured by real-time Q-PCR; data were normalized for signal in islets of adult animals (10 weeks; dashed line). Maturation-dependent upregulation of genes that are important for beta cells, such as insulin 2 (*Ins2*) and pyruvate carboxylase (*Pc*) was observed during this period. **B:** In parallel we observed maturation-dependent repression of monocarboxylate transporter 1 (*Mct1*) and lactate dehydrogenase (*Ldha*), genes which are profoundly repressed in adult mouse islets (Table S2). **C,D:** mRNA expression level in extracts from mouse fetal and neonatal liver between embryonic day E15.5 and 6 weeks after birth. A progressive upregulation (linear increasing trend $P<0.001$) of *Hmgcs2* (C) mirrors a progressive loss (linear decreasing trend $P<0.001$) of expression of

Oxct1 (D). Data in panels A-D are means ± SEM of 3-4 individual experiments. Statistical significance of the difference between postnatal and adult expression signal was calculated by one-way ANOVA with Dunnett post-hoc test. *: P<0.05; **: P<0.01; ***: P<0.001.

Figure 4. Mechanisms for gene repression. A: Epigenetic mechanism for islet- and liver-specific silencing of repressed genes. Histone H3 lysine9-acetylation (green bars) and lysine27-trimethylation (red bars) of chromatin around promoter region of repressed genes were compared in adult mouse tissues with and without observed repression. For *Ldha* and *Mct1*, a high level of H3 lysine27-trimethylation is seen in islets but not in liver, whereas H3 lysine9-acetylation is more extensive in the liver. This difference parallels differences between liver and islets for mRNA expression (see also Fig. 2). Histone H3 trimethylation in the *Oxct1* promoter is significantly higher in liver than in brain, while acetylation at the lysine9 residue is lower. Nucleosomes of the beta actin (*Actb*) promoter which is active in all tissues, display a high level of histone H3 lysine9 acetylation and a low level of histone 3 lysine 27 trimethylation in islets, brain and liver. Black bars represent IgG immunoprecipitation control. Data are means ± SEM of 3-4 QPCRs.

B: MicroRNA 122 (miR-122) contributes to liver-specific silencing of Oxct1 expression. Left panel: Time course of miR-122 expression *in vivo* during liver maturation as measured by Q-PCR. MiR-122 increases with a linear trend, P=0.01. Statistical significance of the difference between P42 and other time points was calculated by one-way ANOVA with Dunnett post-hoc test. *: P<0.05; **: P<0.01; ***: P<0.001. **Middle panel.** Fold change in miR-122 expression *in vitro* in differentiated BMEL cells as measured by Q-PCR. MiR-122 expression decreases 6-fold by treatment with antagomiR-122 (α-miR-122). A mutated antagomiR (α-miR-122 mut) control did not significantly influence miR-122 expression. **Right panel.** Fold change in *Oxct1*

expression in differentiated BMEL cells as measured by Q-PCR. *Oxct1* expression is increased by miR-122 knockdown with antagomiR-122, indicating that miR-122 inhibits *Oxct1* expression. A mutated antagomiR did not significantly alter *Oxct1* expression. Data are means ± SEM of 3 individual experiments.

Supporting Information

Figure S1. Repressed genes versus alternative use of 3' UTR. By screening on the probeset level, data can indicate that a gene is repressed, whereas alternative termination may be the real cause. For example, we initially found *Fscn1* repressed in bone marrow based on the expression data of probeset 1416514_a_at. However, when taking other probesets into account, it is clear that *Fscn1* transcript is present in bone marrow (even slightly higher than the average in other tissues). This indicates that in bone marrow, transcription of *Fscn1* terminates before the binding region of probeset 1416514_a_at (alternative termination site indicated by green vertical bar). In subsequent analysis, we took the information of redundant probesets into account in order to filter out this alternative termination effect. Binding region of the probesets is indicated on the gene structure scheme in a corresponding color.

Figure S2. Expression of 33 repressed genes in pancreatic islets of Langerhans, compared to average expression of these genes in all other tissues. Although significantly lower in islets, expression of a number of these genes is not necessarily close to zero (e.g. *Uqcrb, Rpl24, Rpl36, Cox5a, C1qbp*).

Figure S3. Cell population purity impacts detection of repressed genes. Expression of two genes that are repressed in islets (*Mct1* and *Mgst1*) was measured in whole islets, FACS purified beta cells and acini by microarray (A) and QPCR (B). QPCR expression levels are shown relative to pancreatic acini. Each of the repressed genes displays its highest expression in acini. Expression in pure beta cells is even lower than in islets since inherent to the isolation procedure islets are slightly (1-5% of cells) contaminated with acini. Two tailed t-tests were used to

determine differences between expression in beta cells versus islets or acini. To demonstrate contamination with acini, expression of the exocrine marker *Ptf1a* is shown. While expression of this marker in purified beta cells is negligible, a signal at about 2% of expression in acini can be observed for the whole islet preparations.

Figure S4. A: Timecourse of *c-Maf* mRNA expression in pancreatic islets from neonatal rats (1-28 days old) measured by real-time Q-PCR; data were normalized for signal in islets of adult animals (10 weeks; dashed line). *C-Maf* expression displays a linear decreasing trend ($P<0.001$). Statistical significance of the difference between postnatal and adult expression signal was calculated by one-way ANOVA with Dunnett post-hoc test (***: $P<0.001$). **B,C**: ChIP data further Supplemental chromatin structure around the promotor regions of repressed genes. Histone H3 lysine9-acetylation (green bars) and lysine27-trimethylation (red bars) of *c-Maf* (B) and *Mgst1* (C) in mouse liver and pancreatic islets. For both genes, which are repressed in islets, a high level of H3 lysine27-trimethylation is seen in islets but not in liver, whereas H3 lysine9-acetylation is more extensive in the liver. Black bars represent IgG immunoprecipitation control. Data are means ± SEM of 4 QPCRs.

Figure S5. *Oxct1* mRNA is a predicted target for mouse miR-122. A: Complementarity of a 7-mer seed region of miR-122 to the 3'UTR of mouse *Oxct1* is shown (grey box). Context dependent parameters (described in detail in [40]) are shown as predicted by TargetScanMouse 5.0.

B: Expression profile of miR-122 in different tissues shows its liver-specificity. **C:** miR-122 is the most abundant liver microRNA. Expression values are relative to the most abundant microRNA.

Figure S6. MicroRNAs abundant in islets target islet-repressed genes. A: Expression of the most abundant microRNAs in islets, relative to miR-124. **B:** Predicted target sites of above microRNAs on the *Mct1* (gene symbol *Slc16a1*) 3'UTR. Five binding sites were predicted to be targeted by three different microRNAs that are abundant in islets. Both mir-29a and miR-124 binding sites are conserved in human, mouse, rat, dog and chicken and miR-124 has the highest predicted site efficacy (P_{CT}) of all predicted microRNA binding sites. Predicted binding sites for less abundant microRNAs in islets are not displayed. Grey box designates seed region.

Table S1. Full list of tissue-specifically repressed genes. This table contains all 1074 genes which were found to be repressed in one of the 21 tissues studied.

Table S2. Subset of 35 tissue-specifically repressed genes with fold change >30 versus tissue panel. Data show average mRNA expression signals of the 35 mouse genes that are selectively and most deeply repressed in one target tissue. The mean expression signal averaged over the other 20 tissues in the panel is at least 30 times higher than in the tissue in which the gene is repressed. Fold change was calculated based on linear scale data.

Table S3. Overrepresentation of repressed genes in predicted target gene sets of microRNAs most abundant in tissues. We used only those tissues for which we obtained information on the most abundant microRNAs and had >3 repressed genes (11/21 tissues). For 7 of these tissues, we found a significant overrepresentation of repressed genes within the set of predicted target genes for the most abundant microRNAs. p-value was calculated with the

hypergeometric test. MicroRNA specificity information was obtained via microrna.org or literature, as specified in the last column.

Table S4. Overlap of mouse deeply repressed genes with genes found repressed in human tissues. Genes found repressed in a murine tissue were checked whether they were also expressed the lowest in the same human tissue. Only 17 human tissues were used that are equivalent to the mouse tissue. Data from pure human islets were not available so genes from islets could not be verified. Not all genes have a known human orthologue, so only 27 of the 35 deeply repressed genes could be verified in the human dataset. Fifteen of those, displayed in yellow, had the lowest (or background) expression in the equivalent human tissue.

Supporting Information: Methods

Preparation of tissues and purified cells

Digestion of mouse pancreata was 3 min at 37°C after which islets and acini were separated at 0°C by three rounds of sedimentation (1 x g) and handpicking under a dissection microscope.
For rat islet isolation , P1, P7, P15, P21, P28 and adult 10 week old rats were anesthesized and their islets isolated by collagenase digestion with rodent Liberase RI (Roche, Indianapolis IN). Islets of 10 animals were pooled as one sample for both P1 and P7; those from 2-3 animals were pooled for one sample for P15, P21 and P28.
Rat islets from all ages were handpicked under a stereomicroscope to ensure high purity of islet preparations.
Single cells from islets of RIPYY mice were obtained by dispersing the islets with trypsin, and a population containing > 99% pure β-cells was obtained by sorting the cells with a Beckton Dickinson FACSVantage SE (USA) (excit: 488 nm; emission: 530; 50 µm nozzle; 45 psi, 81 KHz).

RNA extraction

Total RNA was extracted using TRIzol Reagent according to the manufacturer's protocol (Invitrogen, Carlsbad, CA), followed by a cleanup procedure with RNeasy columns (Qiagen, Cologne, Germany). Total RNA from pituitary gland, adrenal gland and islets was extracted using the Absolutely RNA microprep from Stratagene (CA). RNA quantity and quality was determined using the NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, DW) and the 2100 Bioanalyzer (Agilent, Waldbronn, Germany), respectively. Total RNA profiles of all tested samples were similar with sharp 18S and 28S rRNA peaks on a flat baseline. Total RNA of rat islets was isolated using the picoRNA extraction kit from Arcturus or QIAGEN RNeasy Plus Mini Kit depending on the quantity of islets.

Intersection-union test

According to this test, a gene is declared specifically repressed (with type I error at level α) in a baseline tissue if it holds that the null hypothesis of equal or higher expression in the baseline tissue can be rejected at level α for *each* of the 20 baseline tissue – other tissue pairs. The test statistic we used is the t-test as calculated within an ANOVA setting (which implies using the information of all 21 tissues for the calculation of the mean squared error) and we calculated this statistic on the log 2 transformed expression levels obtained after RMA pre-processing. To account for the fact that we perform Berger's test 17344 times (once for each gene), we used Sidak's multiple comparison procedure. For an overall error rate (α_{EW}) of .05 the procedure sets the significance level for each probeset at $\alpha_{CW}=1-(1-(\alpha_{EW})^{1/17344}$, meaning that Berger's procedure was performed for each gene at a significance level of 2.96e-06.

Antagomir treatment

The sequences of antagomir-122 (α-miR-122) and control antagomir (α-miR-122-mut) were: 5'-a*c*aaacaccauugucacacu*c*c*a*-Chol-3' and 5'-a*c*acacaacacugucacauu*c*c*a*-Chol-3', respectively. All the bases were 2' OMe modified; the star indicates phosphorothioate linkage; Chol represents 3'-TEG Cholesteryl.

Chromatin immunoprecipitation

ChIP was performed according to a modified protocol of Upstate (Millipore) EZ-ChIP. Tissue was cross-linked in 1% formaldehyde for 10 min at room temperature. The reaction was stopped by addition of glycine to a final concentration of 0.125 M and samples were washed once with cold PBS. Cell membranes were lysed by pestle homogenization and incubation for 15 min on ice in cell lysis buffer (5mM PIPES [pH 8.0], 85 mM KCl, 0.5% NP-40) supplemented with 1 mM

PMSF and protease inhibitors (Complete protease inhibitor cocktail; Roche). Nuclei were isolated by centrifugation at 3000 x g for 5 min at 4°C and incubated on ice for 15 min in nuclei lysis buffer (1% SDS, 10 mM EDTA, 50 mM Tris-HCl, [pH 8.1], 1mM PMSF, protease inhibitors). Isolated DNA was sheared by sonication for 20 cycles of 30 sec on/off at 200 W (Bioruptor; Diagenode) to an average length of 200-500 bp and centrifuged at 20 000 x g for 30 min at 4°C. DNA in supernatant was diluted 10-fold in dilution buffer. Chromatin was precleared for 1h at 4°C by addition of protein A-TSK (Affiland) beads blocked with salmon sperm DNA (1 mg/ml) and BSA (1 mg/ml). 1% of total input (10 µl of precleared chromatin) was set aside at 4°C until elution. Precleared chromatin (5 µg DNA per ml) was incubated overnight at 4°C using 2 µg of rabbit anti-acetyl-histone H3 (Lys9), rabbit anti-trimethyl-histone H3 (Lys27) or normal rabbit IgG (all antibodies purchased from Upstate). Antibody-chromatin complexes were collected by adding blocked protein A-TSK beads for 1 h at 4°C and washed sequentially in low-salt buffer, high-salt buffer, LiCl buffer and twice in TE buffer. Immune complexes were eluted 2 times for 15 min in 100 µl 1% SDS, 0.1M $NaHCO_3$ at room temperature and cross-links were reversed by overnight incubation at 65°C with 200 mM NaCl. Samples were treated with 10 U RNase A and 5 U RNase T1 (RNase A/T1 mix; Fermentas) for 30 min at 37°C and 0.6 U proteinase K (Fermentas) for 1 h 30 min at 45°C. DNA was purified using spin columns (Geneclean Turbo; Qbiogene) and eluted in 60 µl DNase-free water. For assays involving islets samples (*Mct1*, *Ldha*), input was 5 µl of precleared chromatin, 650 ng DNA was incubated in 500 µl and 1 µg antibody was used and elution was in 30 µl DNase-free water.

Q-PCR primers

Rat islets

	Forward primer	Reverse primer
S25	GTGGTCCACACTACTCTCTGAGTTTC	GACTTTCCGGCATCCTTCTTC
Ins2	TCTTCTACACACCCATGTCCC	GGTGCAGCACTGATCCAC
Pc	TTGAAGGATGTGAAGGGCC	ACCTTTCGGATAGTGCCCTC
Ldha	AGACTTGGCCGAGAGCATAATG	ATAGAGACCCTTAATCATGGTGGAA
Mct1	AGTGCAACGACCAGTGAAGTGT	AGGACCTCCGGCATACATGA

Mouse islets

RNA Pol II	FW	GCACCACGTCCAATGATATTGTG
	Rev	GGAGATGACATGGTACAGTTCTCG
	Probe	(6-Fam)CTTCCGCACAGCCTCAATGCCCAGT(Tamra)
Mct1	FW	GCTTGGTGACCATTGTGGAATG
	Rev	CCCAGTACGTGTATTTGTAGTCTCC
	Probe	(6-Fam)CCCTGTCCTCCTAGGGCCACCACT(Tamra)

Mouse liver

	Forward primer	Reverse primer
Oxct1	GAGCCATGCAGGTTTCTAAGTATG	GCTCCTCCCATTCCTTTCAC
Hmgcs2	TGGTGGATGGGAAGCTGTCTA	TTCTTGCGGTAGGCTGCATAG
Actb	TCCTGAGCGCAAGTACTCTGT	CTGATCCACATCTGCTGGAAG
miR-122	ACACTCCAGCTGGGTAACACTGTCT GGTAA	CTCAACTGGTGTCGTGGAGTCGGCA ATTCAGTTGAGCTACCTGC
18S	GGCGCCCCCTCGATGCTCTTAG	GCTCGGGCCTGCTTTGAACACTCT

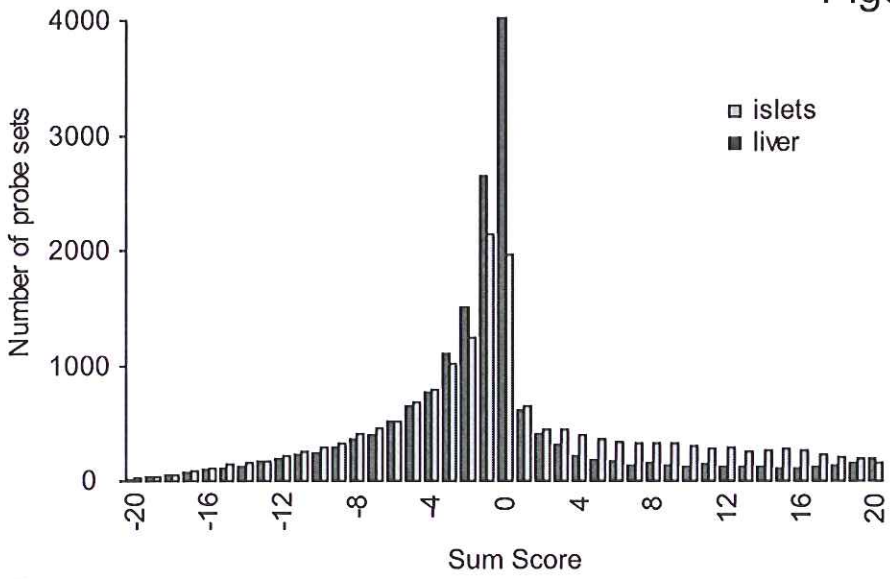
ChIP Primers/probes:

Gene	**Amplicon location**	**Oligonucleotide sequence (5' to 3')**
Mct1	-324 to -233	TTTCTCCTCCAGAGCCTGA
		ACATCCTTATCAGCGCTCTGA
		(6-FAM)GGACGCCATCGTGGGCCTCA(Tamra)
Ldha	-109 to -39	CCAGCCTACACGTGGGTT
		TTAAATGGAAGCTCCGCGCT
		(6-FAM)TCCGCTGGGCTCCCACTCTGA(Tamra)
Actb	+29 to +134	CTTCTTTGCAGCTCCTTCGTTG
		CCCTGCAGTGAGGTACTAGC

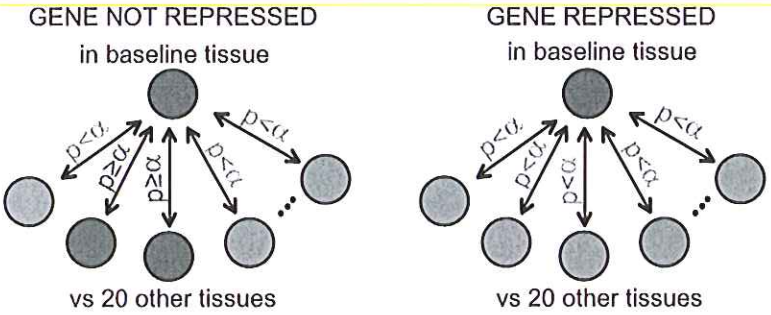
		(6-FAM)CCACACCCGCCACCAGGTAAGCAG(Tamra)
Oxct1	-121 to -2	TCTCGCTTCTTGACAAGCCA AAGAGAGAAGGCAAGGCGA (6-FAM)ACCGCTCTGAGGTCCCGAGGAC(Tamra)
c-Maf	-133 to -7	GTG TGC ACG TTC GAG CTT TC CAG ATG GGC TGC AGG AGA (6-FAM)CCG CTG GCC ACC CAG CAC AG(Tamra)
Mgst1	-161 to -76	TGC CAA GAT CCT TAG CAG AGA AA CCT TCG GAG CTT GCC AAA T (6-FAM) ACC CTC ACC AGC CAC AGC AGT TTC G (Tamra)

Figure 1

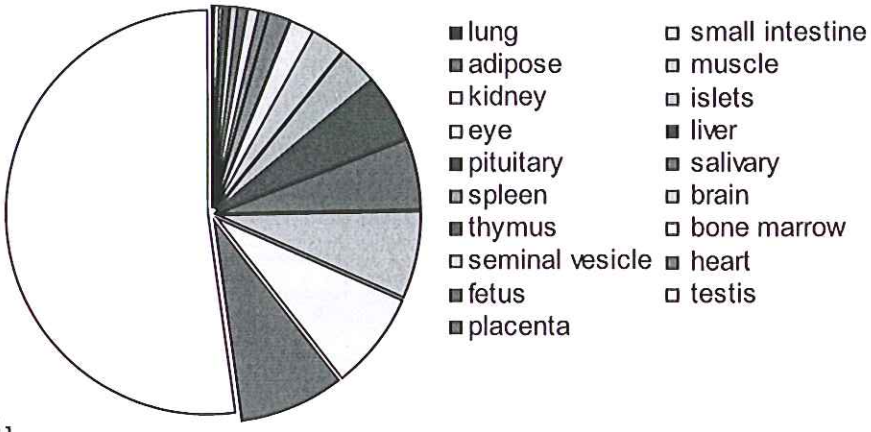
A



B



C



D

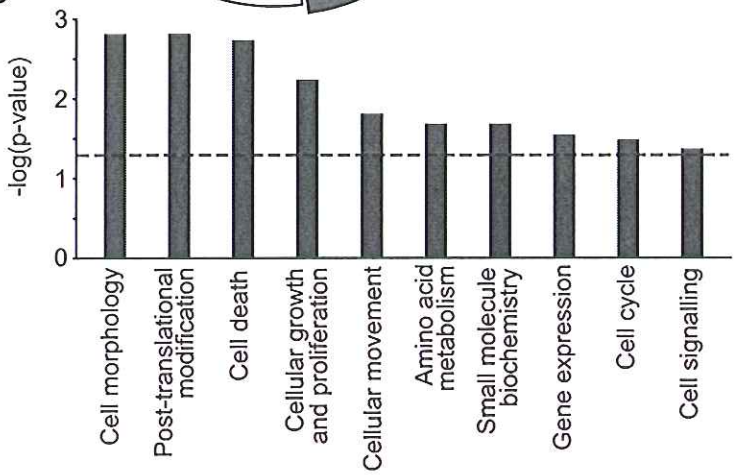


Figure 2

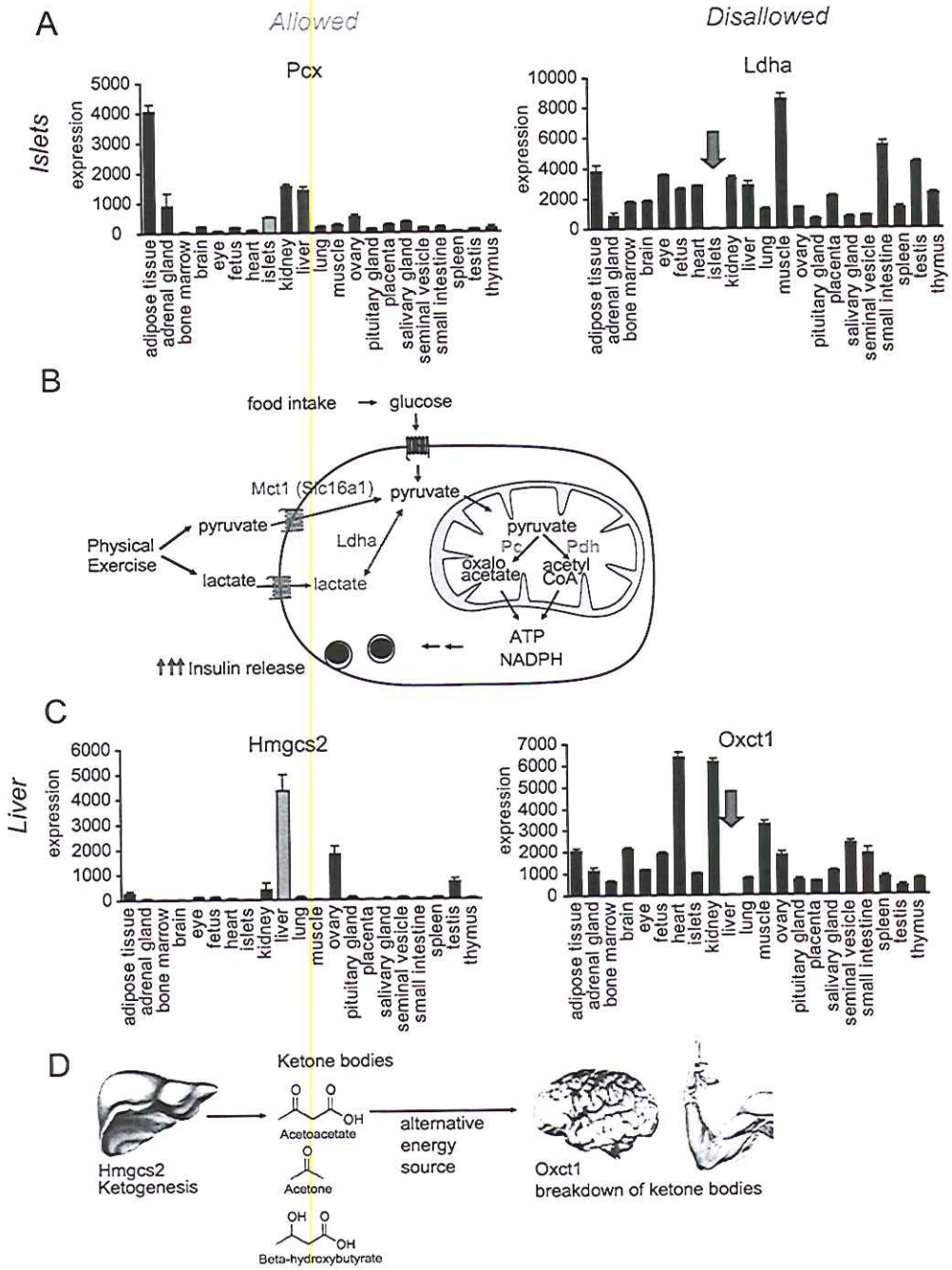

A
Allowed
Disallowed
Pcx
Ldha
Islets
expression
adipose tissue
adrenal gland
bone marrow
brain
eye
fetus
heart
islets
kidney
liver
lung
muscle
ovary
pituitary gland
placenta
salivary gland
seminal vesicle
small intestine
spleen
testis
thymus
B
food intake
glucose
Mct1 (Slc16a1)
pyruvate
Physical Exercise
lactate
Ldha
Pc
Pdh
oxalo acetate
acetyl CoA
ATP
NADPH
Insulin release
C
Hmgcs2
Oxct1
Liver
D
Ketone bodies
Acetoacetate
Acetone
Beta-hydroxybutyrate
alternative energy source
Hmgcs2 Ketogenesis
Oxct1 breakdown of ketone bodies

Figure 3

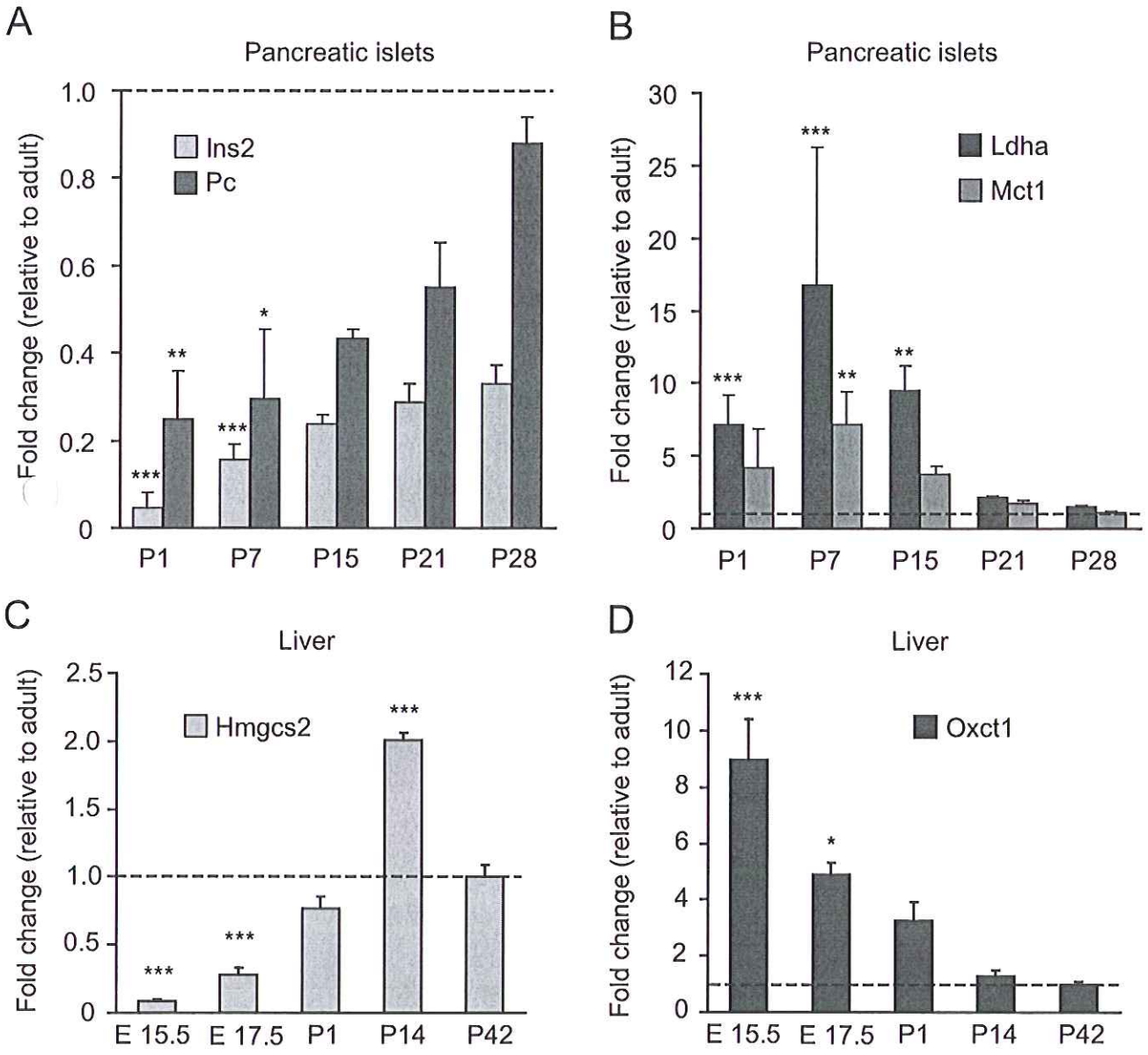
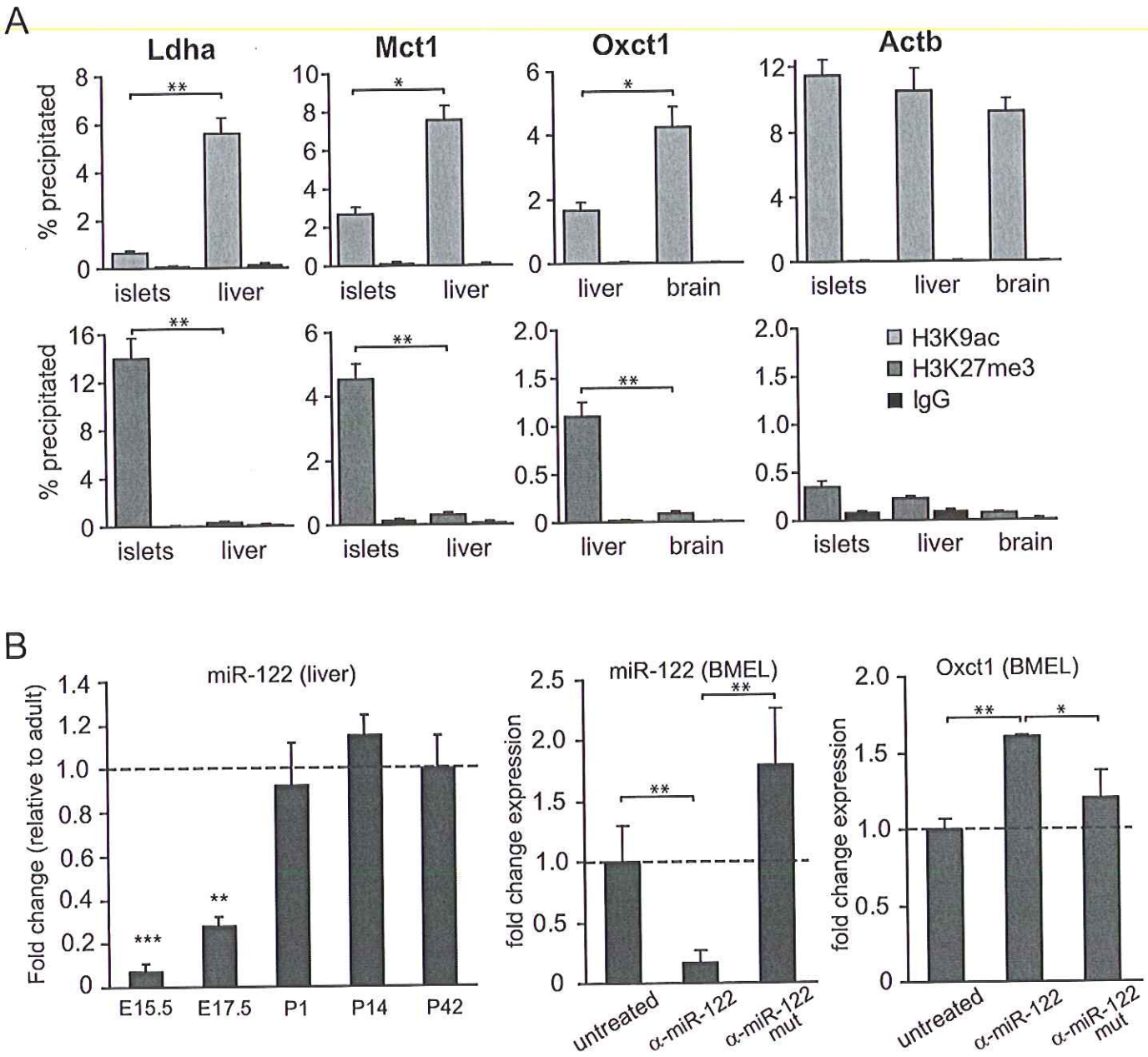

A
Pancreatic islets
Fold change (relative to adult)
Ins2
Pc
P1
P7
P15
P21
P28
B
Pancreatic islets
Ldha
Mct1
C
Liver
Hmgcs2
E 15.5
E 17.5
P1
P14
P42
D
Liver
Oxct1

Figure 4


A
Ldha
Mct1
Oxct1
Actb
% precipitated
islets
liver
brain
H3K9ac
H3K27me3
IgG
B
miR-122 (liver)
Fold change (relative to adult)
E15.5
E17.5
P1
P14
P42
miR-122 (BMEL)
fold change expression
untreated
α-miR-122
α-miR-122 mut
Oxct1 (BMEL)