SUPPLEMENTAL MATERIAL

"Autoregulation: Key to Stable Control of Mammalian Transcriptional Programs?"

Odom D.T. *et al. Mol. Systems Biol. (2005) submission*

List of Supplemental Figures

List of Supplemental Tables

List of Supplemental Sections

Website

Further experimental details are available on the authors' website at
http://web.wi.mit.edu/young/autoregulation/
username: autoregulation *password:* liver
(Available anonymously through http://www.anonymizer.com)

Data Accession Number

Complete, unprocessed data from these experiments have been deposited into the public database ArrayExpress (http://www.ebi.ac.uk/arrayexpress/)
with accession number: **E-WMIT-9**

Figure S1A. The HNF4α gene has an enhancer located 6 kb upstream from the P1 transcriptional start site (Bailly et al. 2001). HNF1α and HNF4α ChIP enrichment is shown as a trace relative to genomic position (graphic below). The known HNF binding sites in the HNF4A P1 promoter are located within 1 kb of the TSS and in the enhancer region 6 kb from the transcriptional start site, and are shown as color shaded ovals along the genome track.

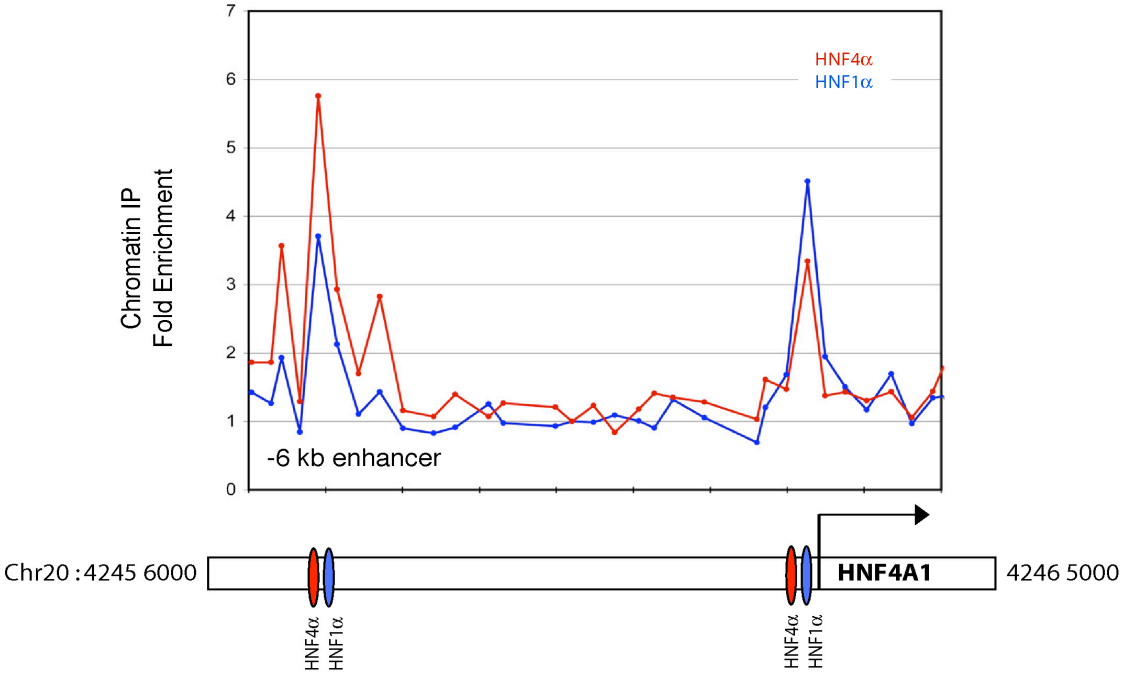


Figure S1B. The TTR gene is bound by a number of HNF factors in vivo at previously known sites within its immediate proximal promoter (reviewed in Costa et al. 2003). The known binding sites in the TTR promoter are located within 250 bp of the TSS, and are shown as color shaded ovals.

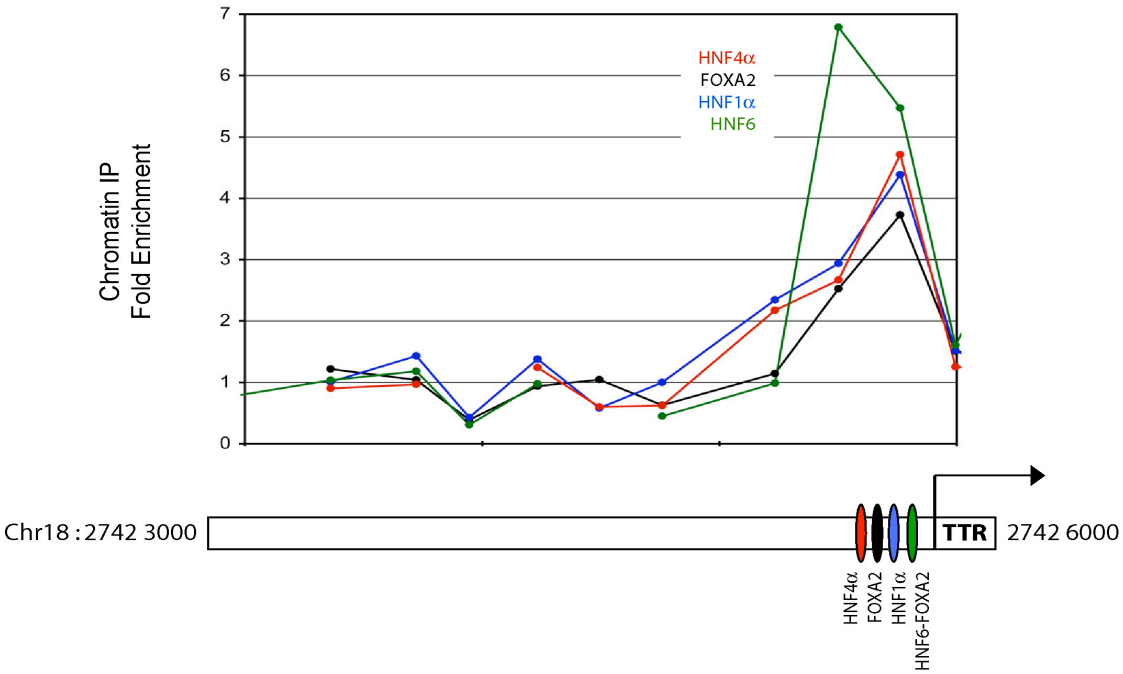


Table S1. Hepatocyte master regulators profiled in this study.*

Regulator	Function	Antisera	Gene Symbol	Regulator Class	Accession
HNF1α	Metabolic control	sc-6547 (Santa Cruz)	TCF1	Homeodomain	NM_000545
HNF4α	Development, metabolism	sc-8987 (Santa Cruz)	NR2A1	Nuclear Receptor	NM_008261
HNF6	Development	sc-13050 (Santa Cruz)	ONECUT1	Homeodomain	NM_004498
HNF3β	Development	R. Costa	FOXA2	Winged Helix	NM_021784
CREB1	Nutrient response	pCREB-5322 (Montminy)	CREB1	bHLH	NM_004379
USF1	Glucose, Lipid metabolism	USF1 (Upstate)	USF1	bHLH	NM_007122

*** References**

Cereghini, S. (1996). Liver-enriched transcription factors and hepatocyte differentiation. Faseb J 10, 267-282.

Hayashi, Y., Wang, W., Ninomiya, T., Nagano, H., Ohta, K., and Itoh, H. (1999). Liver enriched transcription factors and differentiation of hepatocellular carcinoma. Mol Pathol 52, 19-24.

Zaret, K. S. (2002). Regulatory phases of early liver development: paradigms of organogenesis. Nat Rev Genet 3, 499-512.

Costa, R. H., Kalinichenko, V. V., Holterman, A. X., and Wang, X. (2003). Transcription factors in liver development, differentiation, and regeneration. Hepatology *38*, 1331-1347.

Duncan, S. A., Navas, M. A., Dufort, D., Rossant, J., and Stoffel, M. (1998). Regulation of a transcription factor network required for differentiation and metabolism. Science 281, 692-695.

Krivan, W., and Wasserman, W. W. (2001). A predictive model for regulatory sequences directing liver-specific transcription. Genome Res 11, 1559-1566.

HNF1α

Boj, S. F., Parrizas, M., Maestro, M. A., and Ferrer, J. (2001). A transcription factor regulatory circuit in differentiated pancreatic cells. Proc Natl Acad Sci U S A 98, 14481-14486.

Hansen, S. K., Parrizas, M., Jensen, M. L., Pruhova, S., Ek, J., Boj, S. F., Johansen, A., Maestro, M. A., Rivera, F., Eiberg, H., et al. (2002). Genetic evidence that HNF-1alpha-dependent transcriptional control of HNF-4alpha is essential for human pancreatic beta cell function. J Clin Invest 110, 827-833.

Hiraiwa, H., Pan, C. J., Lin, B., Akiyama, T. E., Gonzalez, F. J., and Chou, J. Y. (2001). A molecular link between the common phenotypes of type 1 glycogen storage disease and HNF1alpha-null mice. J Biol Chem 276, 7963-7967.

Pontoglio, M., Sreenan, S., Roe, M., Pugh, W., Ostrega, D., Doyen, A., Pick, A. J., Baldwin, A., Velho, G., Froguel, P., et al. (1998). Defective insulin secretion in hepatocyte nuclear factor 1alpha-deficient mice. J Clin Invest 101, 2215-2222.

Shih, D. Q., Bussen, M., Sehayek, E., Ananthanarayanan, M., Shneider, B. L., Suchy, F. J., Shefer, S., Bollileni, J. S., Gonzalez, F. J., Breslow, J. L., and Stoffel, M. (2001). Hepatocyte nuclear factor-1alpha is an essential regulator of bile acid and plasma cholesterol metabolism. Nat Genet 27, 375-382.

HNF4α

Bailly, A., Torres-Padilla, M. E., Tinel, A. P., and Weiss, M. C. (2001). An enhancer element 6 kb upstream of the mouse HNF4alpha1 promoter is activated by glucocorticoids and liver-enriched transcription factors. Nucleic Acids Res *29*, 3495-3505.

Hayhurst, G. P., Lee, Y. H., Lambert, G., Ward, J. M., and Gonzalez, F. J. (2001). Hepatocyte nuclear factor 4alpha (nuclear receptor 2A1) is essential for maintenance of hepatic gene expression and lipid homeostasis. Mol Cell Biol 21, 1393-1403.

Gupta, R. K., and Kaestner, K. H. (2004). HNF-4alpha: from MODY to late-onset type 2 diabetes. Trends Mol Med 10, 521-524.

Gupta, R. K., Vatamaniuk, M. Z., Lee, C. S., Flaschen, R. C., Fulmer, J. T., Matschinsky, F. M., Duncan, S. A., and Kaestner, K. H. (2005). The MODY1 gene HNF-4alpha regulates selected genes involved in insulin secretion. J Clin Invest 115, 1006-1015.

Lee, Y. H., Sauer, B., and Gonzalez, F. J. (1998). Laron dwarfism and non-insulin-dependent diabetes mellitus in the Hnf-1alpha knockout mouse. Mol Cell Biol 18, 3059-3068.

Parviz, F., Matullo, C., Garrison, W. D., Savatski, L., Adamson, J. W., Ning, G., Kaestner, K. H., Rossi, J. M., Zaret, K. S., and Duncan, S. A. (2003). Hepatocyte nuclear factor 4alpha controls the development of a hepatic epithelium and liver morphogenesis. Nat Genet 34, 292-296.

HNF6

Jacquemin, P., Lemaigre, F. P., and Rousseau, G. G. (2003). The Onecut transcription factor HNF-6 (OC-1) is required for timely specification of the pancreas and acts upstream of Pdx-1 in the specification cascade. Dev Biol 258, 105-116.

Lemaigre, F. P., Durviaux, S. M., Truong, O., Lannoy, V. J., Hsuan, J. J., and Rousseau, G. G. (1996). Hepatocyte nuclear factor 6, a transcription factor that contains a novel type of homeodomain and a single cut domain. Proc Natl Acad Sci U S A 93, 9460-9464.

Rausa, F., Samadani, U., Ye, H., Lim, L., Fletcher, C. F., Jenkins, N. A., Copeland, N. G., and Costa, R. H. (1997). The cut-homeodomain transcriptional activator HNF-6 is coexpressed with its target gene HNF-3 beta in the developing murine liver and pancreas. Dev Biol 192, 228-246.

Samadani, U., and Costa, R. H. (1996). The transcriptional activator hepatocyte nuclear factor 6 regulates liver gene expression. Mol Cell Biol 16, 6273-6284.

FOXA2/HNF3β

Rausa, F. M., Tan, Y., Zhou, H., Yoo, K. W., Stolz, D. B., Watkins, S. C., Franks, R. R., Unterman, T. G., and Costa, R. H. (2000). Elevated levels of hepatocyte nuclear factor 3beta in mouse hepatocytes influence expression of genes involved in bile acid and glucose homeostasis. Mol Cell Biol 20, 8264-8282.

Wolfrum, C., Asilmaz, E., Luca, E., Friedman, J. M., and Stoffel, M. (2004). Foxa2 regulates lipid metabolism and ketogenesis in the liver during fasting and in diabetes. Nature 432, 1027-1032.

Zhang, L., Rubins, N. E., Ahima, R. S., Greenbaum, L. E., and Kaestner, K. H. (2005). Foxa2 integrates the transcriptional response of the hepatocyte to fasting. Cell Metab 2, 141-148.

CREB1

Montminy, M., Koo, S. H., and Zhang, X. (2004). The CREB family: key regulators of hepatic metabolism. Ann Endocrinol (Paris) 65, 73-75.

Sharma, C., Pradeep, A., Pestell, R. G., and Rana, B. (2004). Peroxisome proliferator-activated receptor gamma activation modulates cyclin D1 transcription via beta-catenin-independent and cAMP-response element-binding protein-dependent pathways in mouse hepatocytes. J Biol Chem 279, 16927-16938.

Koo, S. H., Flechner, L., Qi, L., Zhang, X., Screaton, R. A., Jeffries, S., Hedrick, S., Xu, W., Boussouar, F., Brindle, P., et al. (2005). The CREB coactivator TORC2 is a key regulator of fasting glucose metabolism. Nature in press.

USF1

Casado, M., Vallet, V. S., Kahn, A., and Vaulont, S. (1999). Essential role in vivo of upstream stimulatory factors for a normal dietary response of the fatty acid synthase gene in the liver. J Biol Chem 274, 2009-2013.

Iynedjian, P. B. (1998). Identification of upstream stimulatory factor as transcriptional activator of the liver promoter of the glucokinase gene. Biochem J 333 (Pt 3), 705-712.

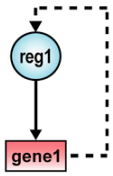
Pajukanta, P., Lilja, H. E., Sinsheimer, J. S., Cantor, R. M., Lusis, A. J., Gentile, M., Duan, X. J., Soro-Paavonen, A., Naukkarinen, J., Saarela, J., et al. (2004). Familial combined hyperlipidemia is associated with upstream transcription factor 1 (USF1). Nat Genet 36, 371-376.

Ribeiro, A., Pastier, D., Kardassis, D., Chambaz, J., and Cardot, P. (1999). Cooperative binding of upstream stimulatory factor and hepatic nuclear factor 4 drives the transcription of the human apolipoprotein A-II gene. J Biol Chem 274, 1216-1225.

Vallet, V. S., Casado, M., Henrion, A. A., Bucchini, D., Raymondjean, M., Kahn, A., and Vaulont, S. (1998). Differential roles of upstream stimulatory factors 1 and 2 in the transcriptional response of liver genes to glucose. J Biol Chem 273, 20175-20179.

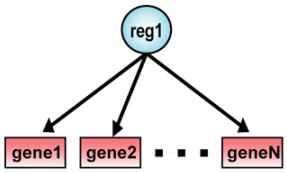
Table S2. Regulatory network motifs in human liver.

Autoregulation



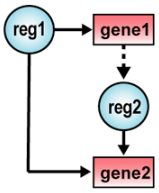
HNF1A
HNF3B
HNF4A
HNF6
CREB1

Single Input



Regulators	# genes
HNF1A	102
HNF3B	177
HNF4A	1829
HNF6	299
CREB1	672
USF1	671

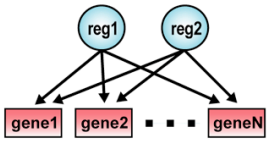
Feedforward Loop



Regulator 1	Regulator 2	# genes
HNF1A	HNF4A	232
HNF3B	HNF4A	130
HNF3B	HNF6	35
HNF3B	CREB1	22
HNF4A	HNF1A	232
HNF6	USF1	41
USF1	HNF4A	239

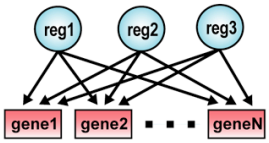
Table S2 (cont'd). Regulatory network motifs in human liver.

2-Factor Multi-input



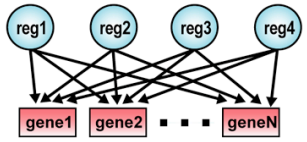
Regulators		# genes
HNF1A	HNF3B	3
HNF1A	HNF4A	232
HNF1A	HNF6	15
HNF1A	CREB1	8
HNF1A	USF1	19
HNF3B	HNF4A	130
HNF3B	HNF6	35
HNF3B	CREB1	22
HNF3B	USF1	33
HNF4A	HNF6	267
HNF4A	CREB1	680
HNF4A	USF1	239
HNF6	CREB1	52
HNF6	USF1	41
CREB1	USF1	82

3-Factor Multi-input



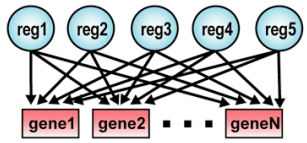
Regulators			# genes
HNF1A	HNF3B	HNF4A	62
HNF1A	HNF3B	HNF6	1
HNF1A	HNF3B	CREB1	1
HNF1A	HNF3B	USF1	4
HNF1A	HNF4A	HNF6	55
HNF1A	HNF4A	CREB1	138
HNF1A	HNF4A	USF1	50
HNF1A	HNF6	USF1	3
HNF1A	CREB1	USF1	2
HNF3B	HNF4A	HNF6	92
HNF3B	HNF4A	CREB1	28
HNF3B	HNF4A	USF1	34
HNF3B	HNF6	CREB1	8
HNF3B	HNF6	USF1	11
HNF4A	HNF6	CREB1	99
HNF4A	HNF6	USF1	47
HNF4A	CREB1	USF1	134
HNF6	CREB1	USF1	13

4-Factor Multi-input



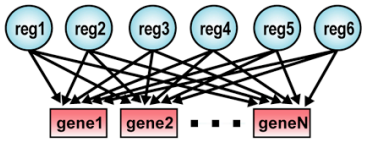
Regulators				# genes
HNF1A	HNF3B	HNF4A	HNF6	59
HNF1A	HNF3B	HNF4A	CREB1	27
HNF1A	HNF3B	HNF4A	USF1	24
HNF1A	HNF3B	CREB1	USF1	1
HNF1A	HNF4A	HNF6	CREB1	20
HNF1A	HNF4A	HNF6	USF1	25
HNF1A	HNF4A	CREB1	USF1	65
HNF1A	HNF6	CREB1	USF1	1
HNF3B	HNF4A	HNF6	CREB1	27
HNF3B	HNF4A	HNF6	USF1	19
HNF3B	HNF4A	CREB1	USF1	10
HNF4A	HNF6	CREB1	USF1	20

5-Factor Multi-input



Regulators					# genes
HNF1A	HNF3B	HNF4A	HNF6	CREB1	24
HNF1A	HNF3B	HNF4A	HNF6	USF1	21
HNF1A	HNF3B	HNF4A	CREB1	USF1	12
HNF1A	HNF3B	HNF6	CREB1	USF1	1
HNF1A	HNF4A	HNF6	CREB1	USF1	26
HNF3B	HNF4A	HNF6	CREB1	USF1	9

6-Factor Multi-input



Regulators						# genes
HNF1A	HNF3B	HNF4A	HNF6	CREB1	USF1	15

Table S3. *S. cerevisiae* transcription factors with autoregulatory loops*

Yeast regulator	Process
STB5	Multidrug resistance
SWI4	Cell cycle
STE12	Mating
SUM1	Sporulation
TEC1	Pseudohyphal growth
PDR3	Multidrug resistance
RCS1	Iron homeostasis
RDS1	Multidrug resistance
ROX1	An/aerobic growth
SMP1	Osmotic stress
SUT1	Hypoxia response
YAP6	Salt tolerance
YAP7	Stress response
ZAP1	Heavy metal exposure
ARG81	Arginine metablism
ARO80	Aromatic amino acid regulation
CBF1	Centromere stability
HAP1	Aerobic respiration
NRG1	Glucose regulation
RAP1	Protein biosynthesis
INO2	Phospholipid biosynthesis
IME1	Meiosis

Autoregulation among S. cerevisiae master regulators

A master regulator is a transcription factor that controls cellular processes using a combination of direct and indirect means. We used the yeast S. cerevisiae as a model to explore the correlation between the presence of autoregulation and identification of a transcription factor as a master regulator. This comparison was made possible because a complete set of binding interactions is available for all known and suspected transcriptional regulators. We analyzed the two cases of direct and indirect regulation separately, though both mechanisms can be present in varying degrees. We reasoned that in order to directly control a cellular process, which can include scores of genes, master regulators as a class would be expected to bind a disproportionately large fraction of yeast promoter regions. In contrast, if a master regulator were controlling a cellular process through intermediary transcription factors, and was present at the top of a hierarchy of factors, we would expect to observe the master regulator binding to a disproportionate number of promoter regions of other transcription factors, which would then act downstream. We used these definitions to test whether autoregulatory loops were correlated with master regulators of yeast cellular functions.

First, we selected transcriptional regulators that bind large numbers of promoters by ordering the 204 transcription factors profiled in rich media by number of bound genes, and taking the top fifth. This cutoff reflects a minimum of 50 bound genes. These 39 regulators bind a combined total of 3955 genes, out of the 6253 genes bound in the complete data set. Comparison to list of regulators that contain autoregulatory loops revealed that 10 of these 39 factors bound their own promoters (hypergeometric p-value of $<1.8 \times 10^{-7}$), and we found that

varying the cutoff impacted the p-value only slightly. Thus, autoregulatory loops are highly associated with transcriptional regulators that bind relatively large numbers of protomers in the yeast genome.

Second, we ranked the 204 transcriptional factors by the number of other transcription factors whose promoters they bind in vivo, and again taking the top fifth. This cutoff affords a separate collection of 39 transcription factors which bind to the promoters of at least 3 and at most 13 transcriptional regulators. Hence each regulator in this class binds on average the promoters for 5.0 transcriptional regulators. This set of regulators are partially overlapping with the prior set of regulators determined to have large numbers of targets. The set of 39 captured 11 autoregulatory events, for a hypergeometric p-value of 4.9×10^{-4}. Again, varying the number of promoter regions used as a minimum cutoff around this range only slightly altered the probability scores. Thus, autoregulatory loops are highly associated with transcriptional regulators that bind many other transcriptional regulators in the yeast genome.

*** References**

Ashburner, B. P., and Lopes, J. M. (1995). Autoregulated expression of the yeast INO2 and INO4 helix-loop-helix activator genes effects cooperative regulation on their target genes. Mol Cell Biol 15, 1709-1715.

Delahodde, A., Delaveau, T., and Jacq, C. (1995). Positive autoregulation of the yeast transcription factor Pdr3p, which is involved in control of drug resistance. Mol Cell Biol 15, 4043-4051.

Shefer-Vaida, M., Sherman, A., Ashkenazi, T., Robzyk, K., and Kassir, Y. (1995). Positive and negative feedback loops affect the transcription of IME1, a positive regulator of meiosis in Saccharomyces cerevisiae. Dev Genet 16, 219-228.

Ren, B., Robert, F., Wyrick, J. J., Aparicio, O., Jennings, E. G., Simon, I., Zeitlinger, J., Schreiber, J., Hannett, N., Kanin, E., et al. (2000). Genome-wide location and function of DNA binding proteins. Science 290, 2306-2309.

Eiznhamer, D. A., Ashburner, B. P., Jackson, J. C., Gardenour, K. R., and Lopes, J. M. (2001). Expression of the INO2 regulatory gene of Saccharomyces cerevisiae is controlled by positive and negative promoter elements and an upstream open reading frame. Mol Microbiol 39, 1395-1405.

Simon, I., Barnett, J., Hannett, N., Harbison, C. T., Rinaldi, N. J., Volkert, T. L., Wyrick, J. J., Zeitlinger, J., Gifford, D. K., Jaakkola, T. S., and Young, R. A. (2001). Serial regulation of transcriptional regulators in the yeast cell cycle. Cell 106, 697-708.

Lee, T. I., Rinaldi, N. J., Robert, F., Odom, D. T., Bar-Joseph, Z., Gerber, G. K., Hannett, N. M., Harbison, C. T., Thompson, C. M., Simon, I., et al. (2002). Transcriptional regulatory networks in Saccharomyces cerevisiae. Science 298, 799-804.

Zeitlinger, J., Simon, I., Harbison, C. T., Hannett, N. M., Volkert, T. L., Fink, G. R., and Young, R. A. (2003). Program-specific distribution of a transcription factor dependent on partner transcription factor and MAPK signaling. Cell 113, 395-404.

Harbison, C. T., Gordon, D. B., Lee, T. I., Rinaldi, N. J., Macisaac, K. D., Danford, T. W., Hannett, N. M., Tagne, J. B., Reynolds, D. B., Yoo, J., et al. (2004). Transcriptional regulatory code of a eukaryotic genome. Nature 431, 99-104.

Table S4. Selected master regulators known to directly autoregulate*§

Regulator	Formal name	Reference *
OCT4	POU5F1	Boyer et al. 2005 and references therein
NANOG	NANOG	Boyer et al. 2005 and references therein
SOX2	NR5A1	Boyer et al. 2005 and references therein
MYOD	MYOD1	Blais et al. 2005 and references therein
MYOG	MYOG1	Blais et al. 2005 and references therein
PU.1	SPI1	Chen et al. 1995
HNF3β	FOXA2	Pani et al. 1992
HNF4α	NR2A1	Briancon et al. 2004

*** References**

Pani, L., Quian, X. B., Clevidence, D., and Costa, R. H. (1992). The restricted promoter activity of the liver transcription factor hepatocyte nuclear factor 3 beta involves a cell-specific factor and positive autoactivation. Mol Cell Biol 12, 552-562.

Chen, H., Ray-Gallet, D., Zhang, P., Hetherington, C. J., Gonzalez, D. A., Zhang, D. E., Moreau-Gachelin, F., and Tenen, D. G. (1995). PU.1 (Spi-1) autoregulates its expression in myeloid cells. Oncogene 11, 1549-1560.

Briancon, N., Bailly, A., Clotman, F., Jacquemin, P., Lemaigre, F. P., and Weiss, M. C. (2004). Expression of the alpha7 isoform of hepatocyte nuclear factor (HNF) 4 is activated by HNF6/OC-2 and HNF1 and repressed by HNF4alpha1 in the liver. J Biol Chem 279, 33398-33408.

Blais, A., Tsikitis, M., Acosta-Alvear, D., Sharan, R., Kluger, Y., and Dynlacht, B. D. (2005). An initial blueprint for myogenic differentiation. Genes Dev 19, 553-569.

Boyer, L. A., Lee, T. I., Cole, M. F., Johnstone, S. E., Levine, S. S., Zucker, J. P., Guenther, M. G., Kumar, R. M., Murray, H. L., Jenner, R. G., et al. (2005). Core transcriptional regulatory circuitry in human embryonic stem cells. Cell 122, 947-956.

§ A more comprehensive list is available on the supporting website.

Table S5. Mammalian transcription factors that do not directly autoregulate*

Regulator	Formal name	Reference
USF1	USF1	This work
YY1	YY1	Yao et al. 1998
SF1	NR5A1	Woodson et al. 1997
MR	NR3C2	Meyer et al. 1994
LEF1	LEF1	Vadlamudi et al. 2005
PAX3	PAX3	Borycki et al. 1999
TBX1	TBX1	Ataliotis et al. 2005

*** References**

Borycki, A. G., Li, J., Jin, F., Emerson, C. P., and Epstein, J. A. (1999). Pax3 functions in cell survival and in pax7 regulation. Development 126, 1665-1674.

Meyer, A. S., and Schmidt, T. J. (1994). In contrast to glucocorticoid receptors, mineralocorticoid receptors are not autoregulated in rat distal colon epithelia. Endocrinology 134, 1163-1172.
Vadlamudi, U., Espinoza, H. M., Ganga, M., Martin, D. M., Liu, X., Engelhardt, J. F., and Amendt, B. A. (2005). PITX2, beta-catenin and LEF-1 interact to synergistically regulate the LEF-1 promoter. J Cell Sci 118, 1129-1137.
Woodson, K. G., Crawford, P. A., Sadovsky, Y., and Milbrandt, J. (1997). Characterization of the promoter of SF-1, an orphan nuclear receptor required for adrenal and gonadal development. Mol Endocrinol 11, 117-126.
Yao, Y. L., Dupont, B. R., Ghosh, S., Fang, Y., Leach, R. J., and Seto, E. (1998). Cloning, chromosomal localization and promoter analysis of the human transcription factor YY1. Nucleic Acids Res 26, 3776-3783.

MATERIALS AND METHODS

Complete protocols describing all materials and methods can be downloaded from http://web.wi.mit.edu/young/autoregulation/

Human tissues. Normal human hepatocytes were obtained through the Liver Tissue Procurement and Distribution System, Pittsburgh, Pennsylvania, which was funded by NIH Contract DK-92310. Cells were chemically crosslinked by the addition of a final concentration of 1% formaldehyde solution for 15 minutes at room temperature. Cells were rinsed twice with PBS or HBSS, frozen in liquid nitrogen, and stored at –80°C prior to use. Most experiments reported here used hepatocyte mixtures created using healthy donor material obtained from multiple individuals, where the donor combinations were selected to maximize diversity of age and gender (see supporting website for more information).

Regulator	Liver Used	Slide identifier format	Liver Used	Slide identifier format
HNF1α	Liver Mix 1	050828_HNF1A_hLiver_1_HP01a_slidecode	Liver Mix 2	050905_HNF1A_hLiver_2_HP01a_slidecode
HNF4α	Liver Mix 1	050826_HNF4A_hLiver_1_HP01a_slidecode	Liver Mix 2	050830_HNF4A_hLiver_2_HP01a_slidecode
FOXA2	Liver Mix 5	050920_FOXA2_hLiver_1_HP01a_slidecode	Liver Mix 2	050924_FOXA2_hLiver_2_HP01a_slidecode
HNF6	Liver Mix 1	050607_HNF6_hLiver_1_HP01a_slidecode	Liver Mix 2	050710_HNF6_hLiver_3_HP01a_slidecode
USF1	Liver Mix 3	050922_USF1_hLiver_1_HP01a_slidecode		
CREB1	Liver Mix 3	050926_pCREB_hLiver_1_HP01a_slidecode		

Chromatin immunoprecipitation protocol. Protocols describing materials and methods are available at http://web.wi.mit.edu/young/autoregulation/ and have been previously reported (Odom et al. 2004, Boyer et al. 2005). In short, frozen cells were resuspended, and lysed as previously described (Odom et al. 2004). Sonication conditions were determined for each cell batch to maximize chromatin sizes between 300-700 bp, and sonications were performed using a Misonix Sonicator 3000, typically at power 7 for ten x 30 second pulses (90 second pause between pulses) at 4°C in an ice bath. The lysate was clarified, and incubated 8 hours at 4°C with 100 μl of Dynal Protein G or Protein A magnetic beads and 10 μg of the appropriate antibody (Table S1). The purified DNA was blunted and ligated to unidirectional linker and amplified using a two-stage PCR protocol (Odom et al. 2004, Boyer et al. 2005). Amplified DNA was labeled and purified using Invitrogen Bioprime random primer labeling kits. Labeled DNA was combined (5 – 6 μg each of immunoenriched Cy5 labelled and input Cy3 labelled DNA) and hybridized with tumbling to arrays in Agilent hybridization chambers for 40 hours at 40°C. Arrays were then washed and scanned.

Microarray design. To more accurately identify the transcriptional regulatory circuitry in human hepatocytes, we have investigated genome-wide occupancy of transcriptional regulators at considerably higher resolution than prior experimental designs that have been used to explore liver-specific gene regulation. Previous designs have used self-printed DNA microarrays with one or two DNA elements to capture signals from each promoter region, covering up to 3 kb at approximately 3,000 transcription start sites, or 1 kb at approximately 13,000 transcription start sites (Friedman et al. 2004, Phuc Le et al. 2005, Rubins et al. 2005, Odom et al. 2004, Zhang et al. 2005). These arrays represented substantially less than whole genome coverage, and were either biased towards promoter regions identified as being present in cDNA libraries isolated from specific tissues (Friedman et al. 2004, Phuc Le et al. 2005, Rubins et al. 2005) or they represented promoters of earlier and less complete genomic annotations derived from NCBI databases (Odom et al. 2004, Zhang et al. 2005).

To overcome these limitations, we used commercial DNA microarrays with custom designs that represent a comprehensive collection of regulatory regions that control most predicted and essentially all known genes (Boyer et al. 2005). The results we obtain with this approach provide a more complete picture of transcription factor occupancy in human hepatocytes by key master regulators, and produce the first high-resolution maps of transcriptional regulation by a core set of master regulators in primary human hepatocytes.

A comprehensive collection of well-characterized transcription start sites was determined by identifying all transcription start sites described in five different databases: RefSeq, Ensembl, MGC, VEGA (www.vega.sanger.ac.uk) and Broad (www.broad.mit.edu). The first three are publicly available databases for gene annotation, the VEGA and Broad data sets are manually annotated databases. We identified all the transcription start sites present in any two of these five databases from the May 2004 build of the human genome.

We then designed DNA microarrays that contain 60-mer oligonucleotide probes covering the region from –8 kb to +2 kb relative to the transcript start sites for these 17,917 annotated human genes (Boyer et al. 2005). Although transcription factors are known to regulate genes from distances greater than 8 kb, most known binding sites for human transcription factors occur within 8 kb of the transcription start site. The sites occupied by transcription factors afforded peaks of ChIP-enriched DNA that span neighboring probes (examples Figure S1). These promoter regions averaged one 60-mer for each 250 bases of sequence, and allowed identification of in vivo binding sites to within tens of bases of position in the genome.

Identification of bound regions. To determine bound regions on the arrays, we developed an algorithm to incorporate information from neighboring probes (described in detail in Boyer et al. 2005). For each 60-mer, we calculated the average probability score of the 60-mer and its two immediate neighbors. If a feature was flagged as abnormal during scanning, we assigned it a neutral contribution to the average probability score. Similarly, if the nearest adjacent feature was beyond 1000 bp, we assigned a neutral contribution. The distance threshold of 1000 bp was determined based on the maximum size of labeled DNA fragments hybridized to the array using the shear distribution of the DNA amplified from the chromatin immunoprecipitation. The maximum fragment size ranged from 500 to 700 bp, depending on the particular experiment.

This set of averaged values gave us a distribution that was subsequently used to calculate p-values of average probe sets. The probe set p-value was determined for each factor based on a number of criteria, including capture of known positives and statistical measurements of the distribution of noise among control probes in the array data (see supporting website).

Data availability. Complete raw data from these experiments have been deposited into the public database ArrayExpress with accession number: **E-WMIT-9**

Materials and Methods References

Friedman, J. R., Larris, B., Le, P. P., Peiris, T. H., Arsenlis, A., Schug, J., Tobias, J. W., Kaestner, K. H., and Greenbaum, L. E. (2004). Orthogonal analysis of C/EBPbeta targets in vivo during liver proliferation. Proc Natl Acad Sci U S A 101, 12986-12991.

Odom, D. T., Zizlsperger, N., Gordon, D. B., Bell, G. W., Rinaldi, N. J., Murray, H. L., Volkert, T. L., Schreiber, J., Rolfe, P. A., Gifford, D. K., et al. (2004). Control of pancreas and liver gene expression by HNF transcription factors. Science 303, 1378-1381.

Boyer, L. A., Lee, T. I., Cole, M. F., Johnstone, S. E., Levine, S. S., Zucker, J. P., Guenther, M. G., Kumar, R. M., Murray, H. L., Jenner, R. G., et al. (2005). Core transcriptional regulatory circuitry in human embryonic stem cells. Cell 122, 947-956.

Phuc Le, P., Friedman, J. R., Schug, J., Brestelli, J. E., Parker, J. B., Bochkis, I. M., and Kaestner, K. H. (2005). Glucocorticoid receptor-dependent gene regulatory networks. PLoS Genet 1, e16.

Rubins, N. E., Friedman, J. R., Le, P. P., Zhang, L., Brestelli, J., and Kaestner, K. H. (2005). Transcriptional networks in the liver: hepatocyte nuclear factor 6 function is largely independent of Foxa2. Mol Cell Biol 25, 7069-7077.

Zhang, X., Odom, D. T., Koo, S. H., Conkright, M. D., Canettieri, G., Best, J., Chen, H., Jenner, R., Herbolsheimer, E., Jacobsen, E., et al. (2005). Genome-wide analysis of cAMP-response element binding protein occupancy, phosphorylation, and target gene activation in human tissues. Proc Natl Acad Sci U S A 102, 4459-4464.