HHS Public Access

Author manuscript

Cell. Author manuscript; available in PMC 2020 April 18.

Published in final edited form as:

Cell. 2019 April 18; 177(3): 722–736.e22. doi:10.1016/j.cell.2019.02.030.

Insulin receptor associates with promoters genome-wide and regulates gene expression

Melissa L. Hancock[1,5], **Rebecca C. Meyer**[1,8], **Meeta Mistry**[2,8], **Radhika S. Khetani**[2,8], **Alexandre Wagschal**[1,3,6], **Taehwan Shin**[1], **Shannan J. Ho Sui**[2], **Anders M. Näär**[1,3,7], and **John G. Flanagan**[1,4,9,*]

[1]Department of Cell Biology, Harvard Medical School, Boston, MA 02115, USA

[2]Bioinformatics Core, Harvard T.H. Chan School of Public Health, Boston, MA 02115, USA

[3]Massachusetts General Hospital Cancer Center, Charlestown, MA 02129, USA

[4]Program in Neuroscience, Harvard Medical School, Boston, MA 02115, USA

[5]Present address: John A. Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, USA

[6]Present address: Exonics Therapeutics, Cambridge, MA, USA

[7]Present address: Department of Nutritional Sciences & Toxicology, University of California, Berkeley, CA 94720, USA

[8]These authors contributed equally

[9]Lead Contact

Summary

Insulin receptor (IR) signaling is central to normal metabolic control and dysregulated in prevalent chronic diseases. IR binds insulin at the cell surface, and transduces rapid signaling via cytoplasmic kinases. However, mechanisms mediating long-term effects of insulin remain unclear. Here, we show that IR associates with RNA polymerase II in the nucleus, with striking enrichment at promoters genome-wide. The target genes were highly enriched for insulin-related functions including lipid metabolism and protein synthesis, and diseases including diabetes, neurodegeneration and cancer. IR chromatin binding was increased by insulin, and impaired in an insulin-resistant disease model. Promoter binding by IR was mediated by coregulator host cell factor-1 (HCF-1) and transcription factors, revealing an HCF-1-dependent pathway for gene

*Correspondence: flanagan@hms.harvard.edu.

Author Contributions

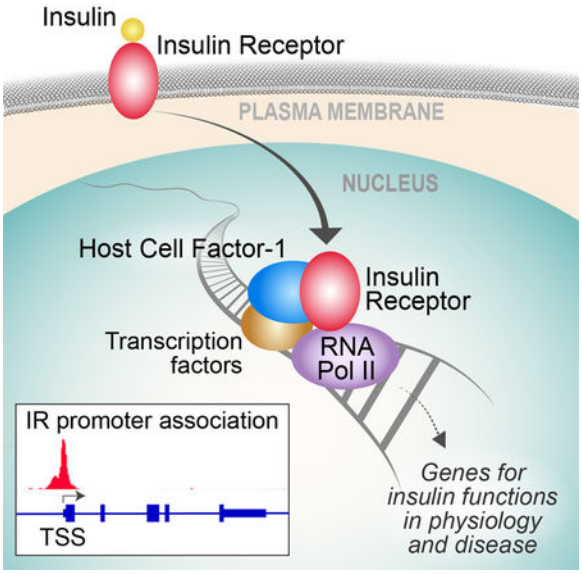
M.L.H. and J.G.F. conceived the project; M.L.H., A.M.N. and J.G.F. designed and interpreted the study; M.L.H., R.C.M., A.W. and T.S. performed experiments; M.M., R.S.K., M.L.H. and S.H.S. performed computational analyses; M.L.H. and J.G.F. wrote the paper, with comments from A.M.N. and the other authors.

Declaration of Interests

The authors declare no competing interests.

regulation by insulin. These results show that IR interacts with transcriptional machinery at promoters, and identify a pathway regulating genes linked to insulin's effects in physiology and disease.

Graphical Abstract



In brief:

Insulin receptor translocates from the cell surface to the nucleus, where it associates with transcriptional machinery at promoters, and regulates genes linked to insulin functions.

Introduction

Insulin is the key hormone that promotes anabolic conversion of precursors into macromolecules, including carbohydrates, lipids, and proteins, leading to nutrient storage and cell growth (Boucher et al., 2014). Insulin binds at the cell surface to the insulin receptor (IR), a receptor tyrosine kinase (RTK). In addition to its critical role in normal physiology, IR plays a central part in diseases that are among the most prevalent and growing health problems worldwide. Resistance to insulin signaling is central to disorders such as type 2 diabetes (T2D), obesity and metabolic syndrome (Boucher et al., 2014; Czech, 2017; Titchenell et al., 2017). While the classic insulin-responsive targets are viewed as liver, muscle and adipose tissue, since they mediate whole-body nutrient homeostasis, IR is expressed ubiquitously. Notably, insulin has powerful effects on the nervous system, including cognition as well as neuron growth and survival, and insulin resistance correlates strongly with neurodegenerative diseases (Heni et al., 2015; de la Monte, 2017). Additionally, excessive insulin signaling contributes to a range of cancers (Vigneri et al., 2016).

The canonical IR signaling pathways involve cascades of kinase enzymes. The major pathway is the phosphoinositide-3-kinase (PI3K)-AKT pathway, which regulates multiple

target proteins to acutely regulate processes in the cytoplasm, including glucose transporter activity, glycogen synthesis, and protein translation (Boucher et al., 2014). AKT also exerts longer-term control of gene expression by phosphorylating transcription factors, including FOXO1, which regulates genes involved in carbohydrate metabolism (Kang et al., 2016). However, key long-term aspects of insulin action and insulin resistance remain poorly understood (Kang et al., 2016; Czech, 2017), suggesting still-unknown downstream mechanisms.

While IR is a cell surface receptor, evidence of its presence in the nucleus has been reported for more than four decades, leading to hypotheses that this might account for long-term effects of insulin including transcriptional regulation (Goldfine and Smith, 1976; Podlecki et al., 1987). More generally, there has been accumulating evidence that several other RTKs that bind ligands at the cell surface can also be found in the nucleus (Carpenter and Liao, 2013; Song et al., 2013). However, the mechanisms, biological functions, and potential genome-wide targets of RTKs in the nucleus generally remain little understood. In the case of IR, it remains unknown whether nuclear IR participates in regulating gene expression, and if so, what molecular partners and mechanisms are involved.

Host Cell Factor-1 (HCF-1) is a ubiquitously-expressed nuclear protein discovered as a host cell factor required for vIRα1 gene transcription (Wilson et al., 1993). HCF-1 functions as a transcriptional coregulator that is directed to gene promoters by binding to a set of DNA sequence-specific transcription factors. Genome-wide analysis shows that HCF-1 targets a large set of actively-transcribed genes, and it has been studied primarily as a factor that promotes cell proliferation by regulating transcription of genes for the cell cycle and cell growth (Wysocka and Herr, 2003; Dejosez et al., 2010; Michaud et al., 2013; Parker et al., 2014).

The present study was initiated by a screen of IR-associated proteins, which unexpectedly revealed that among the most prominent proteins associated with IR was RNA Polymerase II (Pol II). Genome-wide analysis by chromatin immunoprecipitation with high-throughput sequencing (ChIP-seq) identified an IR distribution with a high degree of localization to gene promoters, and the target genes were strongly enriched for insulin-related functions. The ChIP-seq analysis also identified a highly enriched DNA motif, which led us to identify HCF-1 as a protein mediating binding of IR to promoters and insulin regulation of target genes. These results provide a new level of insight into RTK functions in the nucleus, reveal a pathway mediating physiological functions of insulin, and outline a set of potential therapeutic targets for insulin-related diseases.

Results

IR associates with RNA Polymerase II on chromatin

Our previous studies, showing cell surface receptor association with translation machinery (Tcherkezian et al., 2010), led us to further investigate the interactions of cell surface receptors with intracellular partners, using a broad and unbiased approach designed to capture potential associations in all cellular compartments. Testing for receptor-associated proteins by immunoprecipitation and mass spectrometry, we were surprised to find an

association of IR with Pol II (Figures 1A and S1A–C). This interaction appeared prominent and robust: Pol II was at or near the top of the list from whole cells or isolated nuclei, several Pol II subunits were detected, and the association was seen in both liver and brain (Figures 1A and S1A–B). Moreover, it was specific for IR, since no association was detected for two other cell surface receptors tested in parallel (Figures S1B–C). The protocol involved sonication and nuclease treatment, which may explain why association of IR with Pol II was detected prominently here, yet was not noted in previous studies. Indeed when these steps were omitted, no prominent Pol II interaction was seen, consistent with a chromatin-bound complex (Figures 1B and S1C). The interaction was confirmed by co-immunoprecipitation using different antibodies against endogenous and epitope-tagged proteins (Figures S2A–E). Interestingly, IR associated preferentially with Pol II hyperphosphorylated on its C-terminal domain (CTD)(Figures 1C and S2F). This further supports specificity, and since CTD phosphorylation is characteristic of Pol II actively progressing through the transcription cycle (Buratowski, 2009), it also supports a potential role in transcriptional regulation.

For some other RTKs that translocate to the nucleus this involves proteolytic liberation of the cytosolic region, while in other cases it involves the entire receptor holoprotein (Carpenter and Liao, 2013; Song et al., 2013). While evidence for nuclear IR has been described since the 1970s (Goldfine and Smith, 1976; Purrello et al., 1983; Podlecki et al., 1987; Wong et al., 1988; Nelson et al., 2011; Carpenter and Liao, 2013), past studies have not given consistent conclusions about the form of IR in the nucleus. IR is produced as a single polypeptide proform, which is cleaved to α and β chains that remain in a complex. By cellular fractionation, we found both the α and β chains but not the pro-form in the nucleus, supporting translocation of the mature holoprotein (Figure 1D). Other approaches gave results in agreement. Mass spectrometry of nuclear IR included peptides from both α and β chains (Figure S2G). Immunogold electron microscopy detected both α and β chains, with at least a subset within a distance consistent with direct molecular interaction, in the nucleoplasm with no evident enrichment at the nuclear envelope (Figure 1E). Co-precipitation showed association of both α and β chains with Pol II (Figures S2C and S2D). In addition to mouse liver, IRβ was confirmed to be present in the nuclear fraction of human liver (Figure 1F).

Consistent with its nuclear import, IR was found to interact with importin α (KPNA2; Figures S1C and S2H). Translocated proteins may bind importin α directly or indirectly (Miyamoto et al., 2016), and although IR contains potential NLS motifs, we were not able to localize a single sequence required for nuclear import by site-directed mutagenesis (data not shown). To identify additional proteins that might participate in nuclear import, we performed sequential affinity purification with IR and KPNA2, and identified 26 associated proteins including several heat shock proteins (Figure S2I). Interaction of IR and Hsp70 was confirmed by co-immunoprecipitation and electron microscopy (Figures S2J and S2K), which revealed a close proximity consistent with a molecular complex, in both cytoplasm and nucleus. Heat shock proteins mask protein hydrophobic regions, and are proposed to sequester the transmembrane region of the EGFR holoprotein during nuclear import (Carpenter and Liao, 2013). While future studies would be needed to fully characterize the nuclear translocation mechanisms for RTKs in general, these results appear consistent with studies on other receptors (Carpenter and Liao, 2013).

Genome-wide analysis reveals high enrichment of IR on gene promoters

We next performed ChIP-seq to study the distribution of IR genome-wide. ChIP-seq was done in parallel for Pol II phosphorylated at the CTD Ser-5 position (Pol II S5P), a form preferentially associated with the transcription start site (TSS) and first few hundred nucleotides of the gene body (Buratowski, 2009). ChIP-seq for IR identified 3,976 high-confidence peaks ($P \leq 0.001$) in HepG2 human hepatocellular carcinoma cells cultured with insulin. Strikingly, IR peaks were highly enriched near TSSs (Figure 2A), with 82% of IR peaks in gene promoter regions (demarcated here as ±500 bp from the TSS; Figure 2B), an extremely high enrichment as promoter annotation is ~1% of the genome. 73% of IR peaks overlapped Pol II S5P peaks (Figure 2C), with a distribution centered on a similar location near the TSS (Figures 2A and 2D). Validation for individual IR targets by ChIP-qPCR confirmed strong IR enrichment at promoters, in HepG2 cells and primary mouse hepatocytes (Figures 2F, 2G, and S3A–E). *De novo* consensus motif discovery identified a highly enriched DNA consensus motif in IR peaks (E-value 1.4e–442; Figure 2E and Table S1), further showing the extremely high specificity of the IR chromatin association.

Other aspects of the ChIP-seq data further supported a potential role of nuclear IR in transcriptional regulation. Insulin serves physiologically as a homeostatic factor that modulates gene expression in response to nutritional status (Kang et al., 2016), so IR target genes may be predicted to be in chromatin that has previously been opened up by developmental factors. Consistent with this, IR peaks overlapped extensively with active transcription marks such as H3K4me3 (94%) and H3K27ac (78%), and much less with repressive marks such as H3K9me3 (0.1%) and H3K27me3 (21%)(Figures 2F, 2H, and S3C). Additionally, the average expression of IR-bound genes was substantially higher than total genes ($P < 2.2e–16$; Figure 2I), further emphasizing that IR preferentially binds genes that are being actively transcribed.

IR target genes are highly enriched for insulin-related functions

We next explored whether genes with IR-bound promoters have insulin-related functions. Functional pathway analysis revealed extremely strong enrichment for insulin-related functions, with the top categories in Figures 3A and 3B having *P*-values in the range of 1e-10 to 1e-32. The top Reactome canonical pathways (Figure 3A) were characteristic of well-known regulatory roles of insulin in physiology and insulin-resistant disease states, including lipid synthesis and storage (Meikle and Summers, 2017; Titchenell et al., 2017), protein synthesis (Laplante and Sabatini, 2012), immunity (Winer and Winer, 2012), and transcription (Cai et al., 2017). The top disease-related categories showed a striking relationship to disorders associated with insulin signaling, including diabetes (Boucher et al., 2014), cancer (Vigneri et al., 2016), and neurodegeneration including Huntington's, Parkinson's and Alzheimer's diseases (Heni et al., 2015; de la Monte, 2017)(Figure 3B). (Interestingly, the categories also included viral infection – see Discussion.) These results show that IR-bound genes are strongly enriched for insulin-related functions, arguing with high probability that binding of IR to promoters is functionally relevant to the actions of insulin in biology and major diseases.

Insulin regulation of IR chromatin binding, and dysregulation in insulin resistance

Our results suggested a signaling model where insulin would promote IR translocation to chromatin. As predicted, when mice were administered glucose to induce a physiologically relevant rise in blood insulin, this increased liver chromatin-bound IR ($P<0.01$; Figures 4A, 4B, S4A and S4B). Insulin injection had a comparable effect, causing an ~2-fold increase ($P<0.01$; Figures 4C, 4D and S4C). Obese (*ob/ob*) mice were also tested, as a model of insulin resistance (Kennedy et al., 2010), and compared with wild-type littermates, liver chromatin-bound IR was profoundly reduced, ~30-fold, both in the fasting state and after glucose administration ($P<0.01$; Figures 4A and 4B). Membrane-associated IR levels were also somewhat reduced in *ob/ob* animals, in line with previous observations (Ludwig et al., 1988), but the reduction of chromatin-bound IR was far stronger (Figure S4D). These results support a signaling model where insulin increases IR levels on chromatin, and also show that the nuclear IR pathway is highly sensitive to dysregulation in insulin resistance.

To further investigate mechanisms downstream of insulin addition, we used HepG2 cells. Insulin increased nuclear IR (Figures 4E and S4E), consistent with the results in mice, and also increased the association between IR and Pol II S5P (Figures 4F and 4G). When intact HepG2 cells were treated with a membrane-impermeable biotinylation reagent, the results showed that at least some, and perhaps all, of the nuclear IR originated from the cell surface (Figure 4H). Together, these results support a model where insulin binds IR at the cell surface, and induces translocation to the nucleus where it associates with Pol II on chromatin at gene promoters.

Insulin regulates expression of IR-bound genes

We also assessed the effect of insulin on gene expression. In response to 4h insulin treatment of HepG2 cells, RNA-seq identified 3,367 differentially expressed genes (DEGs)(Figures 4I, S4F, S4G and S4H). Combining the RNA-seq and ChIP-seq data, 27% of the genes with IR-bound promoters met the criteria for regulation by insulin (Figure S4G). Genes that were insulin-regulated but lacked detectable IR peaks may be regulated by other mechanisms, or by IR binding sites that did not meet the identification threshold. IR-bound genes that were not detectably insulin regulated might represent non-functional binding sites, or the genes may be regulated by insulin under different conditions, or more subtly than detectable here. Consistent with this last explanation, insulin had moderate effects on a large number of genes (Figure 4I), consistent with previous reports (Cai et al., 2017), and with insulin's role as a homeostatic factor broadly modulating metabolism and cell growth.

We next used Binding and Expression Target Analysis (BETA), a computational program that integrates ChIP-seq and RNA-seq data, to identify whether binding sites are directly involved in transcriptional activation or repression (Wang et al., 2013). The program takes into account quantitative factors including the extent of up- or down-regulation, comparing this for each gene with the number of binding peaks and their distance from the TSS. The results provided strong support for a direct role of IR binding sites in insulin-promoted gene regulation (P=2.4e–13). In contrast, there was no significant relationship of IR binding sites to insulin-repressed genes (P=0.185)(Figure 4J). Consistent with the BETA analysis, the average IR distribution on promoters was somewhat higher and sharper on insulin-

upregulated than-downregulated genes (Figure S4I). While these results do not rule out a direct role of IR binding in repression, they instead suggest a model where either the repressive effects are caused by some other insulin-triggered pathway, or where some of the upregulated genes encode repressive factors that indirectly down-regulate other genes. Our subsequent analyses focused primarily on insulin-upregulated genes. Functional pathway analysis of these insulin-upregulated IR-bound genes confirmed their enrichment for insulin-related functions, notably lipid metabolism and particularly lipid anabolic pathways (Figures S4J and S4K), matching up well with insulin's role in promoting lipid synthesis and storage.

To further assess mechanisms for regulation by insulin, we compared ChIP-seq data from cells in the presence or absence of insulin. Insulin increased IR binding at the TSS of insulin-upregulated genes (area under the curve increased 1.5-fold; Figure 4K), and increased the number of detected IR binding peaks (from 2,988 to 3,976). An increase was validated for individual genes by ChIP-qPCR (Figure S4L). Insulin treatment also caused a 1.5-fold increase in the Pol II S5P peak at the TSS of insulin-upregulated genes, without an obvious shift in distribution toward the gene body (Figure 4K). This indicates increased recruitment of Pol II to the TSS, rather than a potential alternative of increased release and transcription elongation by pre-bound paused Pol II (Adelman and Lis, 2012). Together, these results suggest a mechanistic model where insulin treatment increases IR binding at the TSS, which increases Pol II recruitment and transcription initiation.

IR binding to promoters shows cell-type specificity

To test whether IR associates with different genes in different cell types, we performed ChIP-seq on a human neural cell line, neuroblastoma SH-SY5Y. While fewer IR peaks were detected in SH-SY5Y compared to HepG2 cells (572 versus 3,976), peaks were similarly enriched near TSSs (Figure 5A), with 78% of peaks located within promoters (Figure 5B). The IR peaks in SHSY-5Y cells also contained a highly enriched top consensus motif (E-value 3.4e-718; Figure 5C) similar to that in HepG2 cells (Figure 2E), with additional motifs unique to each cell type (Table S2). Interestingly, in addition to genes in common, a subset of IR-bound genes was unique to each cell type (Figures 5D and 5E). Moreover, the genes unique to the hepatocyte-derived cells were enriched for functions such as lipid metabolism that are characteristic of liver, whereas those unique to the neural cells were enriched for vesicular transport and RNA splicing (Figure 5F) – functions with special importance in neurons (Licatalosi and Darnell, 2006; Sudhof, 2012). These results provide further confirmation that the IR-bound genes are related to functions of insulin, and show that IR binds to specific gene sets with tissue specificity.

IR interacts physically and functionally with transcriptional coregulator HCF-1

To identify proteins that might link IR to promoters, we carried out DNA affinity purification of nuclear proteins, using a biotinylated double-stranded oligonucleotide containing the highly enriched IR site consensus motif (Figures 2E and 5C), followed by mass spectrometry. The resulting list of proteins (Table S3) was combined with our earlier IR co-precipitation (Figure S1C), narrowing down a short list of ten candidates (Figure S5A). Of these, HCF-1 stood out as especially interesting due to its known role as a transcriptional coregulator.

The IR-HCF-1 interaction was confirmed by co-immunoprecipitation (Figures 6A and S5B–E). Moreover, HCF-1-bound peaks in HepG2 cells (ENCODE data set ENCSR529JYA) contained a *de novo* motif similar to that reported in HeLa cells (Michaud et al., 2013), and strikingly similar to the top IR consensus motif identified here (Figure 6B). The distribution of HCF-1 and IR at promoters was nearly identical (Figure 6C), and 64% of promoters with IR ChIP-seq peaks also had HCF-1 peaks (P=8.9e-495; Figure 6D). Overlap of IR and HCF-1 was confirmed by ChIP-qPCR for individual promoters (Figures 6E, S5F and S5G). Furthermore, re-ChIP analysis, by sequential immunoprecipitation for IR and HCF-1, showed their co-association at target promoters (Figures 6F and S5H).

Our working model was that HCF-1 mediates chromatin binding of IR. Supporting this, HCF-1 knockdown abrogated IR binding to promoters (Figures 6G and S5I). In contrast, HCF-1 binding to target promoters was not affected by knockdown of IR (Figure S6A). However, supporting our overall model for insulin-induced gene expression, insulin treatment strongly increased association between HCF-1 and Pol II S5P (Figure S6B).

To address the role of HCF-1 in insulin-regulated gene expression, we used a reporter assay based on the human *LARS* promoter, a target co-bound by IR and HCF-1. As predicted, insulin increased transcription driven by the wild-type *LARS* promoter, and site-directed mutations within the IR/HCF-1 consensus motif reduced this insulin response (P<0.001; Figure 6H). HCF-1 knockdown decreased wild-type *LARS* promoter activity, and prevented the insulin response (P<0.001; Figure 6H). Expression of a kinase-dead IR mutant also abrogated the insulin response (Figure S6C), consistent with our models, although we cannot conclude whether this involves enzyme activity in the nucleus since kinase activity is required for internalization from the plasma membrane (McClain et al., 1987). Further confirming the reporter results, HCF-1 was required for insulin-induced expression of the endogenous *Lars* gene in primary mouse hepatocytes (P<0.001; Figures 6I and S6D).

We also assessed other proteins that might associate with IR in the nucleus. IR coimmunoprecipitated with O-GlcNAc transferase (OGT)(Figure S5D), a protein that binds and regulates HCF-1, and is involved in glucose sensing (Yang et al., 2008; Ruan et al., 2012). We were particularly interested in transcription factors, since HCF-1 does not bind DNA directly. Accordingly, IR-bound promoters showed highly enriched peak overlap with THAP11, YY1, GABPA, and ZNF143 (Figure S6E), all transcription factors known to bind HCF-1 (Michaud et al., 2013). In contrast, IR-bound genes showed minimal overlap with a dataset liver genes bound by Foxo1 (Figure S6F)(Shin et al., 2012), a transcription factor downstream of the canonical IR signaling pathway (Boucher et al., 2014).

We studied THAP11 further, since it recognizes a DNA consensus (Dejosez et al., 2010) nearly identical to the top motifs for IR and HCF-1 (Figure 6B), and since 80% of IR peaks overlapped THAP11 ChIP-seq peaks (Figure S6E), while 18% contained a *de novo* identified THAP11 motif (Figure S6G). Co-immunoprecipitation showed that THAP11 interacts with IR (Figure S6H). Moreover, HCF-1 and THAP11 showed a coincident peak at the TSS (Figure S6I), comparable to IR and Pol II S5P (Figure S4I), consistent with formation of a common complex. When THAP11 was functionally tested by knockdown, this abrogated IR binding to target promoters (Figure S6J). Thus, the association of IR with

HCF-1 and THAP11 provides a mechanistic model accounting, at least in part, for its sequence-selective binding at promoters.

Since 64% of IR-bound promoters contain HCF-1 binding peaks (Figures 6D and S6E), HCF-1 could account for a high proportion, at least, of IR binding sites. Sites where overlap was not seen could either reflect statistical detection limits, or potentially HCF-1-independent IR binding at some sites. We tested for DNA sequences that might account for potentially differential binding of IR and HCF-1 at some sites, and although there were some differences in motif enrichment, we did not see decisive differences between sites in HepG2 cells (Table S4). Intriguingly, though, when we compared our HepG2 and SH-SY5Y ChIP datasets, 63% of IR peaks in both cell types contained a THAP11 motif, versus 9% and 21% of peaks unique to HepG2 or SH-SY5Y cells, respectively (Figure S6K). In addition to the THAP11 consensus, IR binding sites in HepG2 and SH-SY5Y cells contained other enriched motifs, some of which are consensus sequences for transcription factors that bind HCF-1 (Tables S1, S2, Figures S6E and S6G). Overall, these results suggest a model where, in addition to THAP11, other transcription factors also likely contribute to binding, potentially with cell-type specific contributions.

HCF-1-dependent signaling pathway mediates downstream effects of insulin

We next wanted to find a way to block IR binding to promoters, to assess the downstream functional effects – but without blocking the canonical PI3K-AKT pathway (see Figure 7A). We had already found that HCF-1 knockdown blocks IR binding to promoters (Figure 6G), so we tested its affects on the PI3K-AKT and MAP kinase pathways, and found it gave no evident impairment of these cytoplasmic kinase pathways downstream of IR (Figures 7B and S7A).

Our regulatory model (Figure 7A) predicts that insulin-regulated expression of genes with promoters co-bound by IR and HCF-1 should be impaired by depletion of HCF-1. As a negative control, we tested the *IGFBP1* gene, a known transcriptional readout of PI3K-AKT signaling via FOXO1 (Boucher et al., 2014) that is not bound by HCF-1 or IR. No effect of HCF-1 knockdown was seen on insulin regulation of *IGFBP1* (Figure S7A), consistent with our models, and further confirming the lack of effect of HCF-1 knockdown on PI3K-AKT signaling. When the effect of HCF-1 knockdown was tested by RNA-seq, a substantial proportion of insulin-upregulated genes were found to show HCF-1-dependent gene expression (Figures S7B and S7C; insulin-upregulated genes enriched with P=8.4e–102). We also tested individual insulin-regulated genes with promoters co-bound by IR and HCF-1, and found that HCF-1 knockdown blocked insulin-upregulated gene expression (Figures 7C and S7D). Moreover, the effects of either HCF-1 or IR knockdown were comparable, indicating that the HCF-1-dependent IR signaling pathway has a major role in insulin regulation of these genes (Figure 7C).

Having found effects on expression of IR target genes, we were also interested in biological functions downstream of these genes. The functional categories identified for insulin-regulated genes co-bound by IR and HCF-1 (Figure 7D) led to a prediction that the IR/HCF-1 pathway should mediate functions including lipid metabolism, protein translation, and cell cycle regulation. HCF-1 is already known to be involved in cell cycle regulation

(Dejosez et al., 2010; Parker et al., 2014), though a relationship to insulin has not been investigated. Consistent with our predictions, HCF-1 depletion strongly impaired insulin-induced cell proliferation (Figures 7E and S7E), insulin-induced protein translation (Figure 7F), and insulin-induced triglyceride accumulation (Figure 7G). Moreover, importantly, in each case the effect of depleting HCF-1 was similar to the effect of depleting IR itself, indicating that HCF-1-dependent signaling has a major role in regulating cellular functions downstream of insulin.

We also tested effects on lipid metabolism in animals, since lipid metabolism was especially prominent in our pathway analyses (Figures 3A, 7D, and S4J), and since insulin plays a key role in lipid regulation in physiology and disease, and liver-specific depletion of IR is known to reduce liver triglycerides (Michael et al., 2000). IR or HCF-1 were knocked down in mouse liver with shRNA-expressing adeno-associated virus vectors (Figures S7F), and two weeks later liver lipids were assessed. As predicted, HCF-1 knockdown reduced liver triglycerides and free fatty acids, and the effect of depleting either HCF-1 or IR was comparable (Figure 7H). Assessing liver lipid composition by mass spectrometry, 123 lipid species were significantly affected by knockdown of HCF-1 and IR, and as predicted by our model, in the great majority of these cases (113/123) the changes were in the same direction after knockdown of IR or HCF-1 (Figure S7G and Table S5). While other mechanisms undoubtedly contribute, the results in animals and cultured cells show a role for HCF-1-dependent signaling in insulin regulation of lipid metabolism.

Discussion

IR signaling plays a key role in physiology, as well as the most prevalent chronic diseases, including classic nutrition-related disorders such as T2D, obesity, and cardiovascular disease, as well as neurodegenerative disorders and cancer. IR signaling has therefore been studied intensively, yet key aspects of the mechanisms for long-term effects such as gene expression, cell growth, and dyslipidemia remain poorly understood (Kang et al., 2016; Czech, 2017). Here, we identify a pathway where IR associates with promoters and regulates gene expression. Since the target genes were highly enriched for functions important for long-term effects of insulin, and our pathway had a substantial impact on downstream biological outputs, it is likely that these mechanisms have significant roles in insulin-related physiology and disease.

Our results lead to an overall model for a direct effector pathway of transcriptional regulation, initiated by insulin binding at the cell surface. The IR nuclear translocation step appears analogous to ErbB1/EGFR, as both receptors are imported as a holoprotein, in association with heat shock proteins that have been proposed to sequester the hydrophobic transmembrane sequence (Carpenter and Liao, 2013). Within the nucleus, IR was seen by EM in the nucleoplasm, with no evident enrichment at the nuclear envelope, matching at least some previous studies (Podlecki et al., 1987), and suggesting that IR associates with chromatin without simultaneously being membrane associated.

Studies of signal transduction by RTKs at the cell surface lead to a general model where ligand binding triggers receptor autophosphorylation, and this leads to docking of

downstream proteins, and regulation of biological outputs by mechanisms involving induced proximity. By analogy, nuclear functions of IR may involve enhancing physical associations among a complex of nuclear components. Very consistent with this model, we found that treating cells with insulin increased complex formation among a set of components at promoters, including IR, HCF-1 and Pol II.

Other mechanisms could potentially contribute to regulation at IR-bound promoters. In principle, associated nuclear proteins could be phosphorylated directly by the IR tyrosine kinase or indirectly via the canonical kinase cascades. Also, HCF-1 recruits chromatin modifying enzymes to promoters (Dejosez et al., 2010; Michaud et al., 2013; Parker et al., 2014), and these may act upstream of IR by opening genes up for insulin regulation, or might act downstream of IR. Worth noting in this context, while our results clearly showed effects in insulin-promoted gene expression, it also remains possible that IR binding to promoters might repress expression of some genes, and if so, recruitment of repressive chromatin modifying enzymes could provide a potential mechanism. Another protein we found in association with IR was OGT. OGT binds and regulates HCF-1, and glycosylates numerous target proteins, acting as a sensor for glucose levels (Yang et al., 2008; Ruan et al., 2012), so these results suggest intriguing possibilities to regulate gene expression by integrating signals from glucose and insulin.

IR binding to chromatin was highly specific, with a striking localization to gene promoters, and very high enrichment of specific nucleotide sequences. While it is conceivable that IR could interact with DNA directly, the receptor contains no obvious nucleic acid binding domains. Instead, our results indicate that association of IR with a high proportion of its target promoters was mediated by transcriptional coregulator HCF-1. Although IR and HCF-1 co-precipitated in nuclear extracts, we did not detect direct binding of recombinant proteins *in vitro* (IRβ intracellular domain and HCF-1; data not shown), suggesting either that they associate through intermediary proteins, or that binding requires full-length IR or other conditions in the native cell context. HCF-1 itself binds DNA indirectly via DNA sequence-specific transcription factors, and accordingly we found that IR and HCF-1 form a complex on chromatin with THAP11, a transcription factor that binds a DNA consensus almost identical to the top consensus for IR. IR binding sites were also enriched in consensus motifs for other transcription factors, some of which bind HCF-1, and it is likely that multiple transcription factors contribute to binding, potentially conferring cell type specificity and acting cooperatively or combinatorially.

Another key feature of the data is that IR target genes were highly enriched for functional categories with a strong correspondence to distinctive known roles of IR in physiology and disease. Intriguingly, the enriched categories also included vIRαl infections. This seems notable as HCF-1 is a host cell factor required for the herpes simplex virus lytic mode of infection (Wysocka and Herr, 2003), suggesting an evolutionary model where the virus hijacked an insulin-regulated anabolic cell program that serves the needs of virus production. Recent work has found that some viruses encode insulin-like peptides (Altindis et al., 2018), and that interference with IR signaling can block vIRαl infections (Haqshenas et al., 2019), aligning with our work in identifying links between IR signaling and vIRαl infections.

Since insulin resistance is a core feature of T2D and other diseases, we tested an animal model of insulin resistance, the *ob/ob* mouse, revealing a dramatic decrease in the levels of chromatin-bound IR. These results show that the nuclear IR pathway is especially sensitive to the insulin resistant state, although it is important to note that the specific nature and direction of effects on this pathway may depend on the time course and causes of insulin resistance. A pathological feature of insulin resistant states is dyslipidemia, and in our functional classifications of IR-bound genes in liver-derived cells, lipid metabolism was consistently at the top of the lists. In contrast, carbohydrate metabolism did not appear among the top categories, and a lack of detectable IR binding was confirmed for individual genes such as *PCK1* with key roles in insulin regulation of carbohydrate metabolism. Also consistent with a selective role for nuclear IR in lipid metabolism, our IR-bound gene set showed minimal overlap with a set of liver genes bound by Foxo1, which primarily regulates carbohydrate metabolism. A key unresolved question about T2D is why it involves selective effects on insulin regulation of carbohydrate versus lipid metabolism (Czech, 2017; Titchenell et al., 2017), and nuclear IR provides a promising pathway for future investigations of lipid metabolism and insulin resistance. IR binds genes for other cellular processes such as RNA and protein synthesis, as well as genes such as *SOCS1* and *TSC2* with potential negative feedback roles for insulin signaling. Our results open up new avenues of investigation into mechanisms for the long-term effects of insulin in physiology and disease.

More generally, the regulation of gene expression by ligand-regulated IR nuclear translocation is reminiscent of nuclear hormone receptors, which are ligand-regulated transcription factors (Cheung and Kraus, 2010), and signaling molecules such as Notch and β-catenin, which respond to extracellular signals by translocating from the plasma membrane to the nucleus to regulate transcription (Willert and Jones, 2006). Over the last decades, a number of RTKs that bind ligands at the cell surface have been found in the nucleus (Carpenter and Liao, 2013; Song et al., 2013; Mikula et al., 2016; Papadopoulos et al., 2018). By identifying a mechanism for insulin-regulated IR interaction with transcriptional machinery at key regulatory elements genome-wide, targeting a set of genes with a clear relationship to distinctive functions of the receptor, our results provide clear validation for a general principle where RTKs translocate to the nucleus and widely regulate gene expression.

STAR METHODS

Contact for Reagent and Resource Sharing

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, John G. Flanagan (flanagan@hms.harvard.edu).

Experimental Model and Subject Details

In vivo animal studies—All animal procedures were approved by the Harvard Medical School or Massachusetts General Hospital Institutional Animal Care and Use Committee (IACUC). Male C57BL/6J and B6.Cg-*Lep*^{ob}/J mice were purchased from Jackson Laboratory, Bar Harbor, ME. All animals were kept under a 12h light/dark cycle with *ad*

libitum access to pelleted chow and tap water. Mice were allowed to acclimate for at least two weeks in our animal facility before handling. For insulin stimulation, 14 week old male C57BL/6J mice were fasted for 4h prior to intraperitoneal administration of insulin (Humulin R; Eli Lilly) at 0.00075 U/g weight or saline solution. One hour after insulin injection, livers were collected and used for subcellular fractionation. For glucose stimulation, male B6.Cg-*Lep*^{ob}/J mice, either heterozygous or homozygous for Lep^{ob}, were used at 11 weeks old. Mice were fasted overnight for 15h prior to intraperitoneal administration of D-(+)-glucose (Sigma-Aldrich) at 2.5 mg/g weight or saline solution. Fifteen minutes after glucose injection, livers were collected and used for subcellular fractionation studies. Blood glucose concentrations were measured with a regular clinical-grade glucometer at baseline, before insulin injection and upon euthanasia. For shRNA knockdown, Adeno-Associated Virus serotype 8 (AAV8) encoding shRNA against IR (targeting sequence: 5'-CCCTGAAGGATGGAGTCTTTA −3'), HCF-1 (targeting sequence: 5'-TGGCTATCAAGGAGCTTATAG-3') or control shRNA (5'-TGGCTATCAAGGAGCTTATAG −3') were obtained from System Biosciences Inc. (SBI). Sixteen week old male C57BL/6J mice were tail vein injected with 7×10^{10} GC/g of mouse diluted in 100 μl PBS. Two weeks post-infection, livers were collected and flash frozen for further analyses. For immunoprecipitation followed by mass spectrometry, liver from 12 week old male C57BL/6J mice or brain from C57BL/6J mouse embryos (E14.5), both male and female, were used. For the number of animals used per condition in each experiment, please see the legends for Figures 4, 6, 7, S3, S4, S6, and S7. For all experiments, age-matched animals were used.

For cytoplasmic, membrane, soluble nuclear, and chromatin-bound fractionation, mouse tissue was fractionated using Protein Subcellular Fractionation Kit for Tissues (Thermo Fisher Scientific) according to manufacturer's instruction. Tissue was homogenized in Cytoplasmic Extraction Buffer and filtered through a tissue strainer to remove debris. Resulting supernatant was centrifuged at 500 x g at 4°C for 5 min. Cytosolic supernatant was collected and pellet was resuspended in Membrane Extraction Buffer and vortexed briefly to resuspend and then incubated at 4°C for 10 min while rotating. Samples were centrifuged at 3000 x g for 5 min and membrane supernatant was collected. Pellet was resuspended in Nuclear Extraction Buffer and vortexed before incubation at 4°C for 60 min. Samples were centrifuged at 5000 x g for 5 min and soluble nuclear supernatant was collected. The chromatin pellet was resuspended in Chromatin Extraction Buffer including Micrococcal Nuclease and incubated at 37°C for 15 min. Samples were centrifuged at 17,900 x g at 4°C for 5 min and chromatin supernatant was collected. Measurements of triglyceride or free fatty acid from liver were obtained using Triglyceride Colorimetric Assay Kit (Cayman Chemical) or Free Fatty Acid Quantification Colorimetric/Fluorometric Kit (BioVision), respectively, according to manufacturer's instruction.

Human liver—Post-mortem human liver tissues were obtained flash frozen from Sekisui XenoTech LLC, Kansas City, KS. Tissue was obtained by Sekisui XenoTech LLC in accordance with the policies and regulations of the United Network for Organ Sharing (UNOS), and informed consent was obtained from all subjects prior to organ acquisition. Donors were a 36 year old female (donor H0923) and a 46 year old male (donor H1296),

with BMIs under 22, and diagnosed as Normal. We did not perform any analysis of the influence of sex, gender identity, or both, on the results of the study due to the limited sample size. For isolation of non-nuclear and nuclear fractions, human liver tissues were homogenized on ice, strained through 250 mm filters, and resuspended in lysis buffer (50 mM NaCl, 10 mM HEPES, pH 8.0, 500 mM sucrose, 1 mM EDTA, 0.2% Triton X-100, 0.5 mM 2-mercaptoethanol, protease inhibitor (cOmplete; Roche), and phosphatase inhibitor (PhosSTOP; Sigma-Aldrich). After centrifugation at 1,500 g at 4°C for 10 min, the supernatant was collected as the non-nuclear fraction. Pelleted nuclei were washed with sucrose buffer (250 mM sucrose, 6 mM MgCl2, 0.5% Triton-X, and 10 mMTris-HCl, pH 7.4) and resuspended in nuclear lysis buffer (350 mM NaCl, 10 mM HEPES, pH 8.0, 25% glycerol, 0.1 mM EDTA, 0.5 mM 2-mercaptoethanol, protease inhibitor (cOmplete; Roche), and phosphatase inhibitor (PhosSTOP; Sigma-Aldrich)). To release chromatin-bound proteins, nuclear lysis buffer was supplemented with micrococcal nuclease (3000 units/ml; Thermo Fisher Scientific). After incubation on ice and centrifugation at 17,900 g at 4°C for 10 min, the supernatant was collected as the nuclear fraction.

Cultured cells—HepG2 cells (ATCC), originally derived from human liver tissue of a 15 year old male, were cultured in EMEM medium and used before passage number 12. SH-SY5Y cells (ATCC), originally derived from human bone marrow of a four year old female, were cultured in a 1:1 mixture of EMEM and F-12K medium and used before passage number 15. HEK293 cells (ATCC), originally derived from human embryonic kidney, were cultured in DMEM (Thermo Fisher Scientific) and used before passage 25. Cell lines were not authenticated by us, but they were obtained directly from American Type Culture Collection Cell, a trusted source for authentic cell lines. Cells were maintained in respective medium, supplemented with 10% final concentration of fetal bovine serum (FBS; ATCC) and 1% penicillin-streptomycin in a humidified incubator with 5% CO_2 at 37°C. Cells were tested for mycoplasma contamination using MycoFluor Mycoplasma Detection Kit (Thermo Fisher Scientific) and showed no signs of infection. Mouse primary hepatocytes were cultured by perfusing C57BL/6J male mice at 12 weeks old with liver digest medium (Thermo Fisher Scientific). Primary hepatocytes were isolated by filtration with a 70 μm mesh filter, followed by centrifugation in a Percoll (Sigma-Aldrich) gradient. Primary hepatocytes were seeded at 4.5×10^5 cells per well in six-well plates in DMEM medium supplemented with 10% FBS, 2 mM sodium pyruvate, 1% penicillin-streptomycin, 1 μM dexamethasone, and 100 nM insulin. After four hours incubation, medium was changed to maintenance medium (DMEM, 0.2% BSA, 2 mM sodium pyruvate, 1% penicillin-streptomycin, 0.1 μM dexamethasone, and 1 nM insulin).

For serum starvation of cells, medium was removed, cells were washed three times with PBS, and medium supplemented with 0.1% BSA was added to cells for 16h. Cells were stimulated with 1, 10, or 100 nM insulin or IGF1 (IGF1 was used for Figure S7E only) for indicated times. After stimulation, cells were washed in ice-cold PBS, and cell pellets were either scraped in lysis buffer or collected for further analysis. For cytoplasmic, membrane, soluble nuclear, and chromatin-bound fractionation, cells were fractionated using Subcellular Protein Fractionation Kit for Cultured Cells (Thermo Fisher Scientific) according to manufacturer's instruction. Cell viability (proliferation) assays were performed

using Cell Counting Kit-8 (CCK8) colorimetric assay (Dojindo Molecular Technologies), according to manufacturer's instruction. Measurement of cellular triglyceride level of cells plated in 6-well plates was performed using Triglyceride Colorimetric Assay Kit (Cayman Chemical) following the manufacturer's instruction. Triglyceride levels were normalized to cellular protein concentration, which was obtained using Ionic Detergent Compatibility Reagent and Pierce 660 nm Protein Assay Reagent (Thermo Fisher Scientific).

Method Details

Molecular cloning—pShuttle-CMV-hIR-B was a gift from Dr. Ronald Kahn, and full-length human IR-B cDNA was subcloned into the vector pcDNA3.1/myc-His, which has a C-terminal c-Myc tag, using *Hind*III and *Xba*I restriction enzyme sites. For the insertion of an N-terminal HA tag, PCR was used to amplify an HA tag, which was inserted at the N-terminal of IR to generate pcDNA3.1-HA-IR-Myc/His using TagMaster Site-directed Mutagenesis Kit (GM Bioscience). Primer pairs for HA tag amplification are listed in the Key Resource Table. To generate the *LARS* promoter reporter plasmid, human *LARS* promoter was subcloned into pGL4.19[*luc2CP*/Neo] vector (Promega). To generate the mutant *LARS* promoter, site-directed mutagenesis was performed using Q5 Site-Directed Mutagenesis Kit (New England BioLabs). Primer pairs for *LARS* promoter amplification and mutagenesis are listed in Table S6. Plasmid DNA was transfected with Lipofectamine 2000 (Thermo Fisher Scientific) according to manufacturer's instruction.

siRNA transfection—For siRNA transfection, HepG2 cells were reverse transfected using siPORT NeoFX (Thermo Fisher Scientific) according to the manufacturer's instruction. The transfection procedure was repeated two days later, and cell assays were performed two days after the second transfection. HEK293 cells were transfected as described above with Lipofectamine RNAiMAX (Thermo Fisher Scientific) according to manufacturer's instruction. Cell health was assessed by Trypan Blue labeling, and only cells showing minimal labeling were used for cellular assays. Cells appeared healthy within four days of siRNA expression. Primary hepatocytes were transfected with siRNA using Lipofectamine RNAiMax (Thermo Fisher Scientific) four hours after seeding, when medium was changed to maintenance medium. The morning after transfection, medium was changed to fresh maintenance medium, and the transfection procedure was repeated in the evening. Primary hepatocytes were harvested the morning after the second transfection. Oligonucleotide sequences are listed in Key Resource Table and Table S6.

Luciferase reporter assays—For the HCF-1 knockdown experiment, two days after control or HCF-1 siRNA transfection, HEK293 cells were co-transfected with WT or mutant *LARS* promoter-pGL4.10[*luc2*] and pGL4.73[*hRluc*/SV40] (for normalization). One day later, cells were serum starved and incubated in serum-free media alone or supplemented with 10 nM insulin for 24h. In a separate experiment to assess the kinase activity of IR, cells were co-transfected with wild-type or kinase dead IR-FLAG, *LARS* promoter-pGL4.10[*luc2*] and pGL4.73[*hRluc*/SV40]. One day after transfection, cells were serum starved and incubated in serum-free media alone or supplemented with 10 nM insulin for 24h. Cells were processed using Dual-Glo Luciferase Assay System (Promega), and luminescence was measured using an EnVision luminometer (Perkin Elmer). Firefly

luminescence was normalized to *Renilla* luminescence. Luciferase reporter assay results presented in Figures 6H and S6C were confirmed in two or more independent experiments. Oligonucleotide sequences are listed in Table S6.

Cellular cap-dependent translation assays were performed using a bicistronic reporter plasmid pRL-5'-HCV-IRES-FL (Krüger et al., 2001), which directs cap-dependent translation of Renilla luciferase (RL) and also cap-independent hepatitis C virus (HCV) IRES-mediated translation of firefly luciferase (FL) gene. Firefly luciferase was used as an internal control for normalization. HepG2 cells were first transfected with siRNAs using siPORT NeoFX (Thermo Fisher Scientific), then two days later transfected with pRL-5'-HCV-IRES-FL using Lipofectamine 2000 (Thermo Fisher Scientific). The next morning cells were washed with PBS, and cells were incubated in medium supplemented with 0.1% BSA (control) or 0.1% BSA and 10 nM insulin for 24h. Cells were processed using Dual-Glo Luciferase Assay System (Promega), and luminescence was measured using an EnVision luminometer (Perkin Elmer).

Liquid chromatography-tandem mass spectrometry (LC-MS/MS)—Proteins were separated by 10% SDS-PAGE gel, and excised bands were used for mass spectrometry. Excised gel bands were cut and subjected to an in-gel trypsin digestion procedure, and peptides were analyzed by liquid chromatography-tandem mass spectrometry (LC-MS/MS) by the Taplin Mass Spectrometry Facility at Harvard Medical School (Boston, MA) using an LTQ Orbitrap Velos Pro ion-trap mass spectrometer (Thermo Fisher Scientific). Proteins immunoprecipitated with IgG were considered background and not included, and proteins with fewer than two peptide hits were not included. DAVID Bioinformatics Database (v6.7) was used to identify proteins within the Nuclear Part category of GO Cellular Compartment shown in Figure S1C. Protein network analysis in Figure S2I was visualized using STRING (v10)(Szklarczyk et al., 2015).

For extraction of lipid metabolites, 600 μl LC/MS grade methanol was added to 30 mg of liver tissue on ice. Tissue was homogenized using a Pellet Pestle (Fisher), and 300 μl LC/MS grade water followed by ice-cold 400 μl HPLC grade chloroform were added. The top layer containing polar metabolites was separated from the bottom containing non-polar lipid metabolites by centrifugation. The bottom layer was extracted and dried using a speedvac, and stored at −80°C. LCMS/MS and lipid analysis were performed by the Metabolite Profiling Core Facility at Whitehead Institute for Biomedical Research (Cambridge, MA). Dried lipid fractions were resuspended in 50 μl 65/30/5 acetonitrile/isopropanol/water immediately prior to analysis. MS instrument parameters and lipid separation procedures were previously described. Conditions were held constant for both positive and negative ionization mode acquisitions, and MS spectra of lipids were acquired in full-scan/data-dependent MS^2 mode as previously described. High-throughput identification and relative quantification of lipids was performed separately for positive and negative ionization mode data using LipidSearch software (Thermo Fisher Scientific/Mitsui Knowledge Industries) with default parameters for QExactive Product Search and Alignment. After alignment, raw peak areas for all identified lipids were filtered according to the following pre-determined quality control criteria: Rej ("Reject" parameter calculated by LipidSearch software) equal to 0; PQ ("Peak Quality" parameter calculated by LipidSearch software) greater than 0.85;

CV (standard deviation/average peak area across triplicate injections of a representative [pooled] biological sample) below 0.4; R (linear correlation across a three-point dilution series of the representative [pooled] biological sample) greater than 0.9. Typically ~70% of identified lipids passed all four quality control criteria. Raw peak areas of the filtered lipids were added to generate a “total lipid signal” for each sample, and individual lipid peak areas were normalized to this total signal as a control for extraction efficiency and sample loading. Dendrogram of hierarchical clustering of lipids based on log2 transformed fold-change (vs control shRNA) was generated using the HierarchicalClustering module in GenePattern (v3.9.10)(Afgan et al., 2016).

Immunogold Electron Microscopy—HepG2 cells grown in complete medium were collected and fixed in 4% paraformaldehyde and 0.1% glutaraldehyde and incubated in 2.3 M sucrose in PBS containing 0.2 M glycine. Frozen samples were sectioned at −120°C and transferred to formvar-carbon-coated copper grids. Double immunogold labeling was carried out using rabbit anti-IRβ (sc-711), mouse anti-IRα (Thermo Fisher Scientific MA513759), mouse anti-HSP70 (ab5439), and rabbit anti-mouse bridging antibody and 10 nm or 15 nm protein-A gold. Labeled grids were examined in a JEOL 1200EX electron microscope and images were recorded with an AMT 2k CCD camera. Labeled proteins within a distance of 40 to 60 nm is consistent with being associated in a molecular complex (Hermann et al., 1996).

Cell Surface Biotin Labeling—Cell surface proteins were labeled using EZ-Link Sulfo-NHS-SS-Biotin (Thermo Fisher Scientific) according to manufacturer’s instruction. Cells were washed three times with PBS, and incubated in the biotin/PBS solution at RT for 10 min. After labeling, biotin solution was removed and cells were incubated with serum-free EMEM medium alone or supplemented with 10 nM insulin at 37°C for 10 min. Cells were washed with PBS and fractionated using Subcellular Protein Fractionation Kit for Cultured Cells (Thermo Fisher Scientific). Biotin-labeled proteins in either membrane or soluble nuclear fractions were captured by incubation with streptavidin-coupled dynabeads (Thermo Fisher Scientific) at 4°C overnight. Beads were washed in lysis buffer (described below) and proteins were eluted off beads in sample buffer (containing DTT) at 70°C for 10 min.

Immunoprecipitation and Immunoblotting—Tissue or cells were resuspended in lysis buffer (50 mM Tris pH 7.5, 150 mM NaCl, 0.5 mM EDTA, and 1% Triton X-100 supplemented with protease inhibitor (cOmplete; Roche) and phosphatase inhibitor (PhosSTOP; Sigma-Aldrich) and homogenized on ice with a Dounce glass pestle homogenizer. Lysates were sonicated with a Bioruptor UCD-200 (Diagenode) at maximum power for two rounds of 10 min with 30s on and 30s off. Lysates were incubated with RQ1 Dnase I (Promega) and Benzonase (Sigma-Aldrich) at 37°C for 10 min, followed by centrifugation at 17,900 g at 4°C for 15 min to collect the supernatant. Protein concentrations were determined using Pierce 660 nm Protein Assay Reagent. For immunoprecipitation of nuclear extract, proteins were first concentrated using Amicon Ultra 0.5 ml centrifugal filters (Millipore Sigma) at 4°C, then diluted 1:20 in lysis buffer. Lysates were precleared with Protein A or G dynabeads (Thermo Fisher Scientific) coated with 0.1% BSA for 1h at 4°C. Primary antibody and cleared lysate were incubated overnight at 4°C,

and immunoprecipitated complexes were captured by incubation with Protein A or G dynabeads for 2h at 4°C. Beads were washed with lysis buffer two times, followed by two washes with Buffer A (50 mM Tris pH 8.0, 250 mM NaCl, 0.1 mM EDTA, 0.1% NP-40, 10% Glycerol, and 1 mM DTT), and one wash with Buffer B (50 mM Tris pH 8.0, 150 mM NaCl, 0.1 mM EDTA, 0.1% NP-40, 10% Glycerol, and 1 mM DTT). Proteins were eluted in sample buffer at 70°C for 10 min and separated by SDS-PAGE and transferred to nitrocellulose membrane for immunoblotting. Quantitation of immunoblotting band intensity on developed film was performed using ImageJ (version 1.43r), and results shown in Figures 4D, 4G, and S2H were confirmed in two or more independent experiments. The following primary antibodies were used for immunoprecipitation: goat anti-DCC (Santa Cruz sc-6535), mouse anti-FLAG M2 magnetic beads (Sigma-Aldrich M8823), goat anti-GABBR1 (Santa Cruz sc-7338), mouse anti-HA magnetic dynabeads (Thermo Fisher Scientific 88836), rabbit anti-HCF-1 (Bethyl A301–399A), rabbit anti-IGF1Rβ (CST 3027), rabbit IgG (Santa Cruz sc-2027), mouse IgG (Santa Cruz sc-3877), rabbit anti-IRβ (Bethyl A303–712A, for Figure S3A), rabbit anti-IRβ (CST 3025, for Figure S2B); rabbit anti-IRβ (Santa Cruz sc-711, for Figures 1A, 1C, 4F, S1A–C, S2A, S2F, S2G, S3A, and S5B), mouse anti-IRβ Ab-5 (Thermo Fisher Scientific MS-635-P0, for Figures S2J, S5D, and S7F), mouse anti-c-Myc magnetic dynabeads (Thermo Fisher Scientific 88842), and mouse anti-RPB1 CTD (Ser5P, 4H8)(CST 2629, for Figure S2D). The following primary antibodies were used for immunoblotting: rabbit anti-phospho-AKT Ser473 (CST 9271), rabbit anti-AKT (CST 9272), rabbit anti-calnexin (CST 2433), rabbit anti-CD71 (Transferrin receptor; CST 13208), rabbit anti-CREB (CST 9197), rabbit anti-FLAG (CST 2368), rabbit anti-GAPDH (CST 2118), mouse anti-HA.11 (Covance MMS-101P, for Figure S5C), rabbit anti-HA-Tag (CST 3724, for Figures S2C, S5E, and S6B), rabbit anti-Histone H3 (CST 9717), rabbit anti-HCF-1 (CST 69690, for Figures 6A, 7B, S5D, S5I, and S7F), mouse anti-HCF-1 (Santa Cruz sc-390950, for Figures S5B, S5C, and S6J), rabbit anti-HSP70 (Abcam ab5439), mouse anti-IGF1Rβ (Millipore 05–656), rabbit anti-IRα (Santa Cruz sc-710), rabbit anti-IRβ (Santa Cruz sc-711, for Figures 1B, 1D, 1F, 4A, 4C, 4H, S2D, S2J, S3B, S4E, and S5D), mouse anti-IRβ (Ab-6)(Thermo Fisher Scientific MS-636-P0, for Figures 1C, 4F, 7B, S2A, S2B, S2F, and S5B), rabbit anti-KPNA2 (CST 14372), rabbit anti-phospho-p44/42 MAPK Thr202/Tyr204 (CST 4377), rabbit anti-MAPK (CST 9102), rabbit anti-mCherry (Abcam ab167453), mouse anti-c-Myc tag (Thermo Fisher Scientific MA1–980, for Figure S5C), rabbit anti-Myc-tag (CST 2278, for Figure S6H), rabbit anti-OGT (CST 24083), rat anti-Pol II unphosphorylated (ChromoTek 1C7), rat anti-Pol II Ser2P (ChromoTek 3E10), rat anti-Pol II Ser5P (ChromoTek 3E8, for Figures 6A, S2C, S2D, and S6B), rat anti-Pol II Ser7P (ChromoTek 4E12), rat anti-Pol II Tyr1P (Millipore 3D12), mouse anti-RPB1 CTD (Ser5P, 4H8)(CST 2629, for Figures 1C, 4C, 4F, S2A, and S2B), mouse anti-Pol II (Santa Cruz sc-56767, for Figures 1B and 1D), rabbit anti-phospho-p70 S6 kinase Thr389 (CST 9205), rabbit anti-p70 S6 kinase (CST 9202), rabbit anti-T7 tag (CST 13246), and rabbit anti-THAP11 (CST 12305). Primary antibodies are also listed in Key Resource Table.

For immunoprecipitation followed by mass spectrometry, liver from 12 week old male C57BL/6J mice or brain from C57BL/6J mouse embryos (E14.5) were lysed, and proteins were immunoprecipitated as described above. For liver nuclear extracts, liver tissue was

fractionated using Protein Subcellular Fractionation Kit for Tissues (Thermo Fisher Scientific) according to manufacturer's instructions. Soluble nuclear and chromatin-bound fractions were combined and used for immunoprecipitation with rabbit IgG or rabbit anti-IRβ (Santa Cruz sc-711) antibody as described above. HEK293 cells co-expressing IR-FLAG and KPNA2-T7 were processed for sequential immunoprecipitation followed by mass spectrometry. Briefly, IR-FLAG was immunoprecipitated from cell lysate using anti-FLAG M2 beads, and bound proteins were eluted off washed beads with the addition of 100 μg/ml 3X FLAG peptide (Sigma-Aldrich). The eluate was used for a second round of immunoprecipitation with rabbit anti-T7 antibody (13246; CST) and Protein A beads, and proteins were eluted off washed beads in sample buffer at 70°C for 10 min. Protein network analysis in Figure S2I was visualized using STRING (v10)(Szklarczyk et al., 2015).

DNA affinity purification—Nuclear extract from HEK293 cells was incubated with 5 μg WT or scrambled control biotinylated double-stranded DNA oligonucleotides (1:1 mixture of FW 5' biotin/RV unlabeled and FW 3'biotin/RV unlabeled; see Table S6 for oligonucleotide sequences), 30 ng/μl poly(dI-dC)(Sigma-Aldrich), 40 μl of streptavidin-coupled dynabeads (Thermo Fisher Scientific) that were preblocked in 5% milk, 200 μl of 5X binding buffer (50 mM Tris pH 7.5, 250 mM KCl, 10 mM $MgCl_2$, 0.25% NP-40, 0.5 mM EDTA, 5 mM DTT, and 25% glycerol), and water to bring the final volume to 1 ml. Samples were incubated overnight at 4°C. Beads were washed four times with 1X binding buffer, resuspended in sample buffer, and heated at 95°C for 5 min. Eluted proteins were separated on 10% SDS-PAGE gels and gel bands were excised and processed for mass spectrometry, as described above. Proteins with more than two peptide hits, and ≥2-fold over scrambled control were included. Two independent experiments were performed, and identified proteins are listed in Table S3.

RT-qPCR—Total RNA from cells or tissue was extracted using the RNeasy Plus Mini kit (Qiagen). RNA amount and quality were measured using a Nanodrop instrument (Thermo Fisher Scientific). RNA was reverse transcribed using Superscript III Reverse Transcriptase and random hexamer primers (Thermo Fisher Scientific), and cDNA was measured using a QuantStudio 7 Flex real-time PCR machine (Applied Biosystems) with Power SYBR green PCR master mix (Thermo Fisher Scientific). Fold changes in mRNA levels were determined using the using 2-delta cycle threshold method normalized to *HPRT* or *TBP* mRNA. RT-qPCR results presented in Figures 6I, S4H, and S6D were confirmed in at least two independent experiments. Primer pair sequences were obtained from PrimerBank (Wang et al., 2012) and listed in Table S6.

ChIP and reChIP Analyses—Cells cultured in 15-cm dishes were washed with PBS and treated with formaldehyde (1% final concentration) for 10 min to cross-link proteins to DNA before harvesting. Crosslinking was quenched with glycine, and cells were scraped off the plate and washed with ice-cold PBS three times. Cells were resuspended in hypotonic buffer (25 mM HEPES pH 7.9, 1.5 mM $MgCl_2$, 10 mM KCl, 0.1% NP-40, 1 mM DTT, and protease inhibitors) and homogenized on ice with 20 strokes of a Dounce glass pestle homogenizor. Nuclei were spun down, resuspended in SDS-sonication buffer (0.3% SDS, 50 mM HEPES pH 7.9, 140 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% Nadeoxycholate,

and protease inhibitors), and chromatin extract was sonicated using a Bioruptor UCD-200 instrument (Diagenode) at 4°C to obtain a DNA smear with an average fragment size of 200–500 bp. After centrifugation, supernatant was diluted 1:3 with immunoprecipitation buffer (50 mM HEPES pH 7.9, 140 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% Na-deoxycholate, and protease inhibitors). Lysate was precleared with Protein A dynabeads preblocked in 0.5% BSA at 4°C for 1h, and at this point input samples were taken for input controls. Precleared chromatin extract was incubated at 4°C overnight with primary antibody (1 μg per 10 μg chromatin), and Protein A dynabeads were added to capture immunoprecipitated complexes for 2h at 4°C. Beads were washed twice with immunoprecipitation buffer, followed by two washes in Buffer A (50 mM HEPES pH 7.9, 500 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% SDS, 0.1% Na-deoxycholate, and protease inhibitors), two washes in Buffer B (20 mM HEPES pH 7.9, 250 mM LiCl, 1 mM EDTA, 0.5% NP-40, 0.5% Na-deoxycholate, and protease inhibitors), and one wash in TE buffer (10 mM Tris pH 8.0 and 1 mM EDTA). Complexes were eluted off the beads in Elution Buffer (50 mM Tris pH 8.0, 1 mM EDTA, 1% SDS, and 50 mM $NaHCO_3$) at 65°C for 15 min. Elution buffer was added to input control samples, and they were processed in parallel. Samples were adjusted to 200 mM NaCl and incubated at 65°C overnight to reverse crosslinks. Samples were sequentially treated with Rnase A at 37°C for 1h, followed by EDTA and Proteinase K at 42°C for 2h, and DNA was purified using MinElute PCR purification kit (Qiagen) according to manufacturer's instruction. ChIP and input DNA were analyzed by RT-qPCR or used for constructing sequencing libraries. For ChIP-qPCR, the value of enrichment was calculated relative to input and (where indicated) the ratio to enrichment at distal control regions. ChIP-qPCR results presented in Figures 2G, 6G, S3D, S4L, and S5G were confirmed in two or more independent experiments. The following primary antibodies were used for ChIP-qPCR: rabbit anti-IRβ (Santa Cruz sc-711, for Figures 2G, 6F, 6G, S3B, S3D, S3E, S4L, S5G–I, and S6J), rabbit anti-IRβ (Bethyl A303–712A, for Figures 2G and S3D), rabbit anti-IGFRβ (CST 3027), anti-rabbit HCF-1 (Bethyl A301–399A), anti-rabbit Pol II S5P (CST 2629), or anti-rabbit IgG (sc-2027). Rabbit anti-IRβ (Santa Cruz sc-711) or mouse anti-RPB1 CTD (Ser5P, 4H8)(CST 2629) primary antibodies were used for ChIP-seq. For ChIP-qPCR, primer pair sequences and positions of amplicons relative to TSS are provided in Table S6.

ReChIP was conducted as previously described (Parker et al., 2012), with modifications. ChIP using rabbit IgG or rabbit anti-IRβ (Santa Cruz sc-711) antibody was performed as described above, but after the last TE wash, beads were resuspended in reChIP elution buffer (10 mM Tris pH 7.6, 1 mM EDTA, 2% SDS, 20 mM DTT) and heated at 37°C for 30 min. Samples were diluted 20-fold in immunoprecipitation buffer (10 mM Tris pH 7.6, 100 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% BSA), and samples were split and incubated in rabbit IgG or rabbit anti-HCF-1 (Bethyl A301–399A) at 4°C overnight. The reChIP samples were collected, washed, and eluted as described above for ChIP. The value of enrichment was measured by RT-qPCR and shown as relative to input of the first round of ChIP.

ChIP Sequencing Analysis—For control conditions without insulin, medium was removed and cells were cultured in serum-free media for 16h. For insulin treatment, cells were stimulated 100 nM insulin for 20 min. Two biological replicates were used for each

condition, and rabbit anti-IRβ (Santa Cruz sc-711) or mouse anti-RPB1 CTD (Ser5P, 4H8) (CST 2629) primary antibodies were used for ChIP (described above). Figures 2 and 5 show ChIP-seq data from insulin-treated cells. Libraries were constructed using Thruplex DNAseq kit (Rubicon) according to manufacturer's instruction and single-end sequenced on an Illumina NextSeq 500 instrument at the Molecular Biology Core Facilities (MBCF) at Dana-Farber Cancer Institute (DFCI). Data quality was evaluated with FASTQC. High quality reads were mapped to the human genome UCSC build hg19 using Bowtie2 2.2.8 (Langmead and Salzberg, 2012). Alignments were filtered to retain only uniquely mapping reads. Peaks were called with MACS2 2.1.1 (Zhang et al., 2008) using default parameters, and a narrow peak cutoff of $P<0.001$. To assess the reproducibility of peaks across replicates, samples were subjected to IDR analysis (Babu et al., 2011), in which peak consistency was validated using pooled pseudo-replicates. The pooled pseudo-replicates resulted peak numbers within a factor of two of the original results suggesting highly reproducible replicates. As such, the consensus peakset (minimum 1 bp overlap) was retained as confident peaks. Coordinates for the first replicate were used for downstream analysis.

For publicly available HepG2 ChIP-seq datasets, peak coordinates were obtained from the ENCODE Project Consortium (https://www.encodeproject.org/): H3K4me3 (ENCSR575RRX), H3K9me3 (ENCSR000ATD), H3K27ac (ENCSR000AMO), H3K27me3 (ENCSR000AOL), THAP11-FLAG (ENCSR562POI), GABPA (ENCSR000BJK), and YY1 (ENCSR000BNT). For HCF-1 (ENCSR529JYA) and ZNF143 (ENCSR101FJT), FASTQ sequence files obtained from the ENCODE Consortium were processed using our quality control and IDR analysis pipelines, described above, to obtain peak coordinates.

For genomic annotation of ChIP-seq peaks in Figure 2B, HOMER v4.6 (Heinz et al., 2010) was used with default parameters using RefSeq annotations for the human genome hg19. Promoter regions were defined as ±500 bp from the TSS; transcription termination sites by default are defined from −100 bp to +1 kb from the TTS. Peak overlaps between various groups were determined with BedTools 2.24.0. Heatmaps and average profile plots were generated with deepTools2.0 (Ramírez et al., 2016), using computeMatrix and plotHeatmap or plotProfile tools. Bigwigs were normalized to input using the bamCompare tool from deepTools 2.0, setting normalization to bins per million (BPM) and the ratio parameter to subtract. Bigwigs were visualized using the integrative genomics viewer (IGV)(v2.3.75). Mouse Foxo1 ChIP-seq (Shin et al., 2012) region coordinates were converted to human (hg19) coordinates using LiftOver. Genomic regions enrichment of annotations tool (GREAT)(McLean et al., 2010) was used to assign genomic regions to genes. Genes were functionally annotated using the Molecular Signature Database (MsigDB v6.2)(Subramanian et al., 2005) and the Reactome database (v57)(Fabregat et al., 2016). To simplify the resulting output, we plotted only the top Reactome categories of the second and third hierarchical levels for Figures 3A, 7E, and S4J.

De novo motif analysis was performed using the MEME suite of tools (Bailey et al., 2009) with default parameters. Tomtom was used with default settings to match discovered motifs to the human Hocomoco v11 database. For user specified motifs, sequences were scanned using PWMScan (Ambrosini et al., 2018) with default parameters. Position weight matrices

(PWMs) for core promoter elements were from the Promoter Motifs library; THAP11 was from the human Hocomoco v11 library.

RNA Sequencing Analysis—HepG2 cells were incubated in medium without serum overnight, then treated with insulin (10 nM) or PBS (control condition) for 4 hours (nine biological replicates per condition). For HCF-1 knockdown experiments, HepG2 cells were transfected with HCF-1 siRNA or control siRNA, as described above (eight biological replicates per condition). Total RNA was extracted using the Rneasy Plus Mini kit (Qiagen). RNA-seq libraries were constructed using TruSeq Stranded mRNA (Illumina) according to manufacturer's instruction and single-end sequenced on an Illumina NextSeq 500 instrument at the DFCI MBCF. RNA samples were processed using the bcbio RNA-seq pipeline (v0.9.1a; https://bcbio-nextgen.readthedocs.org). Reads were aligned to the hg19 genome using STAR (v2.4.1d)(Dobin et al., 2013). Reads were quantified with featureCounts (v1.4.4)(Liao et al., 2014). Differential gene expression was analyzed using DESeq2 (v1.16.1)(Love et al., 2014).

To correlate IRβ promoter targets with expression levels (Figure 2I), ChIP-seq targets were intersected with the RNA-seq results from HepG2 cells cultured in media supplemented with insulin, ranked by expression value. Log2 "baseMean" values from DESeq2 (v1.16.1) of IRβ targets (2,528 genes) and the full list of genes (total) expressed in HepG2 cells were plotted in the boxplots. Statistical significance was determined using the Mann Whitney Wilcoxon test.

Significant associations between IRβ ChIP-seq sites and DEGs from RNA-seq of cells ±insulin were identified using Binding and Expression Target Analysis (BETA) v1.0.7 (Wang et al., 2013) with the BETA-plus sub-protocol. The final target gene lists include genes differentially expressed at FDR < 0.001 that contain IRβ binding sites within 100 kb of the TSS.

Quantification and Statistical Analysis

One-way analysis of variance (ANOVA) with Dunnett's post-hoc analysis, two-way ANOVA with Tukey's post-hoc analysis, and two-tailed t-tests were performed using Prism (v7). For the liver lipidomics analysis, two-tailed t-test followed by FDR method of Benjamini and Hochberg was performed to determine statistical significance of lipids relative to control liver using Prism (v7). The threshold for the heatmap generated in Figure S7G was determined by p-values (P<0.05) to show the broad pattern of increased and decreased lipids, and FDR for each lipid is shown in Table S5. Significance of gene overlap between datasets was calculated using the cumulative distribution function of the hypergeometric test (http://systems.crump.ucla.edu/hypergeometric/). Human protein-coding genes [20,576 genes; RefSeq annotation release 107] were used for background in the hypergeometric tests. There was no explicit randomization or blinding procedure for animals. Sample sizes were based on comparable prior studies showing samples sizes necessary to achieve significance. All n-values represent biological replicates and are stated in the figure legends. In the legends for Figures 4B, 4D, 7H, S4A, S4C, S4D, S7F, and S7G, n-values represent the number of animals used per condition in each experiment.

Experimental significance was set to $P<0.05$, and n.s. = no significance. Variance was similar between compared groups, and all graphs with error bars show standard error relative to mean.

Data and Software Availability

Sequencing data files have been deposited in NCBI with GEO accession number GEO: GSE107336. Raw data files for images are available at Mendeley (http://dx.doi.org/10.17632/6tsjbsyyt7.1).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank Steve Buratowski, Karen Adelman, Malcolm Whitman, Pere Puigserver, Alex Banks, Steve Gygi, Clary Clish, Shirley Liu, James Kadonaga, Sudha Biddinger, David Van Vactor, Stephen Liberles and Jeffrey Flier for helpful discussions; and Maria Ericsson, Seung-Joo Lee, Hyunsuk Suh, Kfir Sharabi, Jianhua Shi, Bowen Tan, Ross Tomaino and Caroline Lewis for experimental help and advice. This work was supported by NIH grants R01EY011559 and R01NS069913 to JGF; R01DK084459, R01HL116391 and U54HL119145 to AMN, also an MGH Research Scholar; T32CA009361 to RCM; F32NS073307 to MLH; and U54HL119145 to AW; and support from the Harvard NeuroDiscovery Center and the HMS Tools and Technology Committee to SHS.

References

Adelman K and Lis JT (2012). Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. Nat Rev Genet 13, 720–731. [PubMed: 22986266]

Afgan E, Baker D, van den Beek M, Blankenberg D, Bouvier D, Cech M, Chilton J, Clements D, Coraor N, Eberhard C, et al. (2016). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. Nucleic Acids Res 44, W3–W10. [PubMed: 27137889]

Altindis E, Cai W, Sakaguchi M, Zhang F, GuoXiao W, Liu F, De Meyts P, Gelfanov V, Pan H, DiMarchi R, Kahn CR (2018) VIRαl insulin-like peptides activate human insulin and IGF-1 receptor signaling: A paradigm shift for host-microbe interactions. Proc Natl Acad Sci USA 115, 2461–2466. [PubMed: 29467286]

Ambrosini G, Groux R, and Bucher P (2018). PWMScan: a fast tool for scanning entire genomes with a position-specific weight matrix. Bioinformatics 34, 2483–2484. [PubMed: 29514181]

Babu M, Gagarinova A, and Emili A (2011). Array-based synthetic genetic screens to map bacterial pathways and functional networks in Escherichia coli. Methods Mol Biol 781, 99–126. [PubMed: 21877280]

Bailey TL, Bodén M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, and Noble WS (2009). MEME SUITE: tools for motif discovery and searching. Nucleic Acids Research, 37, W202–W208. [PubMed: 19458158]

Boucher J, Kleinridders A, and Kahn CR (2014). Insulin receptor signaling in normal and insulin-resistant states. Cold Spring Harb Perspect Biol 6.

Buratowski S (2009). Progression through the RNA polymerase II CTD cycle. Mol Cell 36, 541–546. [PubMed: 19941815]

Cai W, Sakaguchi M, Kleinridders A, Gonzalez-Del Pino G, Dreyfuss JM, O'Neill BT, Ramirez AK, Pan H, Winnay JN, Boucher J, et al. (2017). Domain-dependent effects of insulin and IGF-1 receptors on signalling and gene expression. Nat Commun 8, 14892. [PubMed: 28345670]

Carpenter G, and Liao HJ (2013). Receptor tyrosine kinases in the nucleus. Cold Spring Harb Perspect Biol 5, a008979. [PubMed: 24086039]

Cheung E, and Kraus WL (2010). Genomic analyses of hormone signaling and gene regulation. Annu Rev Physiol 72, 191–218. [PubMed: 20148673]

Czech MP (2017). Insulin action and resistance in obesity and type 2 diabetes. Nat Med 23, 804–814. [PubMed: 28697184]

de la Monte SM (2017). Insulin Resistance and Neurodegeneration: Progress Towards the Development of New Therapeutics for Alzheimer's Disease. Drugs 77, 47–65. [PubMed: 27988872]

Dejosez M, Levine SS, Frampton GM, Whyte WA, Stratton SA, Barton MC, Gunaratne PH, Young RA, and Zwaka TP (2010). Ronin/Hcf-1 binds to a hyperconserved enhancer element and regulates genes involved in the growth of embryonic stem cells. Genes Dev 24, 1479–1484. [PubMed: 20581084]

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, and Gingeras TR (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21. [PubMed: 23104886]

Fabregat A, Sidiropoulos K, Garapati P, Gillespie M, Hausmann K, Haw R, Jassal B, Jupe S, Korninger F, McKay S, et al. (2016). The Reactome pathway Knowledgebase. Nucleic Acids Res 44, D481–487. [PubMed: 26656494]

Goldfine ID, and Smith GJ (1976). Binding of insulin to isolated nuclei. Proc Natl Acad Sci U S A 73, 1427–1431. [PubMed: 179086]

Haqshenas G, Terradas G, Paradkar PN, Duchemin J, McGraw EA, and Doerig C (2019). A role for the insulin receptor in the suppression of dengue virus and Zika virus in Wolbachia-infected mosquito cells. Cell Reports 26, 529–535. [PubMed: 30650347]

Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, and Glass CK (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell 38, 576–589. [PubMed: 20513432]

Heni M, Kullmann S, Preissl H, Fritsche A, and Haring HU (2015). Impaired insulin action in the human brain: causes and metabolic consequences. Nat Rev Endocrinol 11, 701–711. [PubMed: 26460339]

Hermann R, Walther P, and Muller M (1996). Immunogold labeling in scanning electron microscopy. Histochem Cell Biol 106, 31–39. [PubMed: 8858365]

Huang DW, Sherman BT, and Lempicki RA (2009). Systematic and integrative analysis of large gene lists using DAVID Bioinformatics Resources. Nat Protoc 4, 44–57. [PubMed: 19131956]

Kang S, Tsai LT, and Rosen ED (2016). Nuclear Mechanisms of Insulin Resistance. Trends Cell Biol 26, 341–351. [PubMed: 26822036]

Kelley JB, Talley AM, Spencer A, Gioeli D, and Paschal BM (2010). Karyopherin alpha7 (KPNA7), a divergent member of the importin alpha family of nuclear import receptors. BMC Cell Biol 11, 63. [PubMed: 20701745]

Kennedy AJ, Ellacott KL, King VL, and Hasty AH (2010). Mouse models of the metabolic syndrome. Dis Model Mech 3, 156–166. [PubMed: 20212084]

Kim W, Lao Q, Shin YK, Carlson OD, Lee EK, Gorospe M, Kulkarni RN, and Egan JM (2012). Cannabinoids induce pancreatic beta-cell death by directly inhibiting insulin receptor activation. Sci Signal 5, ra23. [PubMed: 22434934]

Krüger M, Beger C, Welch PJ, Barber JR, Manns MP, and Wong-Staal F (2001). Involvement of proteasome alpha-subunit PSMA7 in hepatitis C virus internal ribosome entry site-mediated translation. Mol Cell Biol 21, 8357–8364. [PubMed: 11713272]

Langmead B, and Salzberg SL (2012). Fast gapped-read alignment with Bowtie 2. Nat Methods 9, 357–359. [PubMed: 22388286]

Laplante M, and Sabatini DM (2012). mTOR signaling in growth control and disease. Cell 149, 274–293. [PubMed: 22500797]

Liao Y, Smyth GK, and Shi W (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 30, 923–930. [PubMed: 24227677]

Licatalosi DD, and Darnell RB (2006). Splicing regulation in neurologic disease. Neuron 52, 93–101. [PubMed: 17015229]

Love MI, Huber W, and Anders S (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15, 550. [PubMed: 25516281]

Ludwig S, Muller-Wieland D, Goldstein BJ, and Kahn CR (1988). The insulin receptor gene and its expression in insulin-resistant mice. Endocrinology 123, 594–600. [PubMed: 2838265]

Mao YS, Sunwoo H, Zhang B, and Spector DL (2011). Direct visualization of the cotranscriptional assembly of a nuclear body by noncoding RNAs. Nat Cell Biol 13, 95–101. [PubMed: 21170033]

McClain DA, Maegawa H, Lee J, Dull TJ, Ulrich A, and Olefsky JM (1987). A mutant insulin receptor with defective tyrosine kinase displays no biologic activity and does not undergo endocytosis. J Biol Chem 262, 14663–14671. [PubMed: 3312193]

McLean CY, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB, Wenger AM, and Bejerano G (2010). GREAT improves functional interpretation of cis-regulatory regions. Nat Biotechnol 28, 495–501. [PubMed: 20436461]

Meikle PJ, and Summers SA (2017). Sphingolipids and phospholipids in insulin resistance and related metabolic disorders. Nat Rev Endocrinol 13, 79–91. [PubMed: 27767036]

Michael MD, Kulkarni RN, Postic C, Previs SF, Shulman GI, Magnuson MA, and Kahn CR (2000). Loss of insulin signaling in hepatocytes leads to severe insulin resistance and progressive hepatic dysfunction. Mol Cell 6, 87–97. [PubMed: 10949030]

Michaud J, Praz V, James Faresse N, Jnbaptiste CK, Tyagi S, Schutz F, and Herr W (2013). HCFC1 is a common component of active human CpG-island promoters and coincides with ZNF143, THAP11, YY1, and GABP transcription factor occupancy. Genome Res 23, 907–916. [PubMed: 23539139]

Mikula M, Skrzypczak M, Goryca K, Paczkowska K, Ledwon JK, Statkiewicz M, Kulecka M, Grzelak M, Dabrowska M, Kuklinska U, et al. (2016). Genome-wide co-localization of active EGFR and downstream ERK pathway kinases mirrors mitogen-inducible RNA polymerase 2 genomic occupancy. Nucleic Acids Res 44, 10150–10164. [PubMed: 27587583]

Miyamoto Y, Yamada K, and Yoneda Y (2016). Importin alpha: a key molecule in nuclear transport and non-transport functions. J Biochem 160, 69–75. [PubMed: 27289017]

Nelson JD, LeBoeuf RC, and Bomsztyk K (2011). Direct recruitment of insulin receptor and ERK signaling cascade to insulin-inducible gene loci. Diabetes 60, 127–137. [PubMed: 20929976]

Papadopoulos N, Lennartsson J, and Heldin CH (2018). PDGFRb translocates to the nucleus and regulates chromatin remodeling via TATA element-modifying factor 1. J Cell Biol 217, 1701–1717. [PubMed: 29545370]

Parker JB, Palchaudhuri S, Yin H, Wei J, and Chakravarti D (2012). A transcriptional regulatory role of the THAP11-HCF-1 complex in colon cancer cell function. Mol Cell Biol 32, 1654–1670. [PubMed: 22371484]

Parker JB, Yin H, Vinckevicius A, and Chakravarti D (2014). Host cell factor-1 recruitment to E2F-bound and cell-cycle-control genes is mediated by THAP11 and ZNF143. Cell Rep 9, 967–982. [PubMed: 25437553]

Podlecki DA, Smith RM, Kao M, Tsai P, Huecksteadt T, Brandenburg D, Lasher RS, Jarett L, and Olefsky JM (1987). Nuclear translocation of the insulin receptor. A possible mediator of insulin's long term effects. J Biol Chem 262, 3362–3368. [PubMed: 3546306]

Purrello F, Burnham DB, and Goldfine ID (1983). Insulin receptor antiserum and plant lectins mimic the direct effects of insulin on nuclear envelope phosphorylation. Science 221, 462–464. [PubMed: 6346487]

Ramírez F, Ryan DP, Grüning B, Bhardwaj V, Kilpert F, Richter AS, Heyne S, Dündar F, Manke T (2016). deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res 44, W160–W165. [PubMed: 27079975]

Ruan HB, Han X, Li MD, Singh JP, Qian K, Azarhoush S, Zhao L, Bennett AM, Samuel VT, Wu J, et al. (2012). O-GlcNAc transferase/host cell factor C1 complex regulates gluconeogenesis by modulating PGC-1alpha stability. Cell Metab 16, 226–237. [PubMed: 22883232]

Shin DJ, Joshi P, Hong SH, Mosure K, Shin DG, and Osborne TF (2012). Genome-wide analysis of FoxO1 binding in hepatic chromatin: potential involvement of FoxO1 in linking retinoid signaling to hepatic gluconeogenesis. Nucleic Acids Res 40, 11499–11509. [PubMed: 23066095]

Song S, Rosen KM, and Corfas G (2013). Biological function of nuclear receptor tyrosine kinase action. Cold Spring Harb Perspect Biol 5.

Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, et al. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 102, 15545–15550. [PubMed: 16199517]

Sudhof TC (2012). The presynaptic active zone. Neuron 75, 11–25. [PubMed: 22794257]

Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, Simonovic M, Roth A, Santos A, Tsafou KP, et al. (2015). STRING v10: protein-protein interaction networks, integrated over the tree of life. Nucleic Acids Res 43, D447–452. [PubMed: 25352553]

Tcherkezian J, Brittis PA, Thomas F, Roux PP, and Flanagan JG (2010). Transmembrane receptor DCC associates with protein synthesis machinery and regulates translation. Cell 141, 632–644. [PubMed: 20434207]

Titchenell PM, Lazar MA, and Birnbaum MJ (2017). Unraveling the Regulation of Hepatic Metabolism by Insulin. Trends Endocrinol Metab.

Vigneri R, Goldfine ID, and Frittitta L (2016). Insulin, insulin receptors, and cancer. J Endocrinol Invest 39, 1365–1376. [PubMed: 27368923]

Wang S, Sun H, Ma J, Zang C, Wang C, Wang J, Tang Q, Meyer CA, Zhang Y and Liu XS (2013). Target analysis by integration of transcriptome and ChIP-seq data with BETA. Nat Protoc 8, 2502–2515. [PubMed: 24263090]

Wang X, Spandidos A, Wang H, and Seed B (2012). PrimerBank: a PCR primer database for quantitative gene expression analysis, 2012 update. Nucleic Acids Res 40, D1144–1149. [PubMed: 22086960]

Willert K, and Jones KA (2006). Wnt signaling: is the party in the nucleus? Genes Dev 20, 1394–1404. [PubMed: 16751178]

Wilson AC, LaMarco K, Peterson MG, and Herr W (1993). The VP16 accessory protein HCF is a family of polypeptides processed from a large precursor protein. Cell 74, 115–125. [PubMed: 8392914]

Winer S, and Winer DA (2012). The adaptive immune system as a fundamental regulator of adipose tissue inflammation and insulin resistance. Immunol Cell Biol 90, 755–762. [PubMed: 22231651]

Wong KY, Hawley D, Vigneri R, and Goldfine ID (1988). Comparison of solubilized and purified plasma membrane and nuclear insulin receptors. Biochemistry 27, 375–379. [PubMed: 2831959]

Wysocka J, and Herr W (2003). The herpes simplex virus VP16-induced complex: the makings of a regulatory switch. Trends Biochem Sci 28, 294–304. [PubMed: 12826401]

Yang X, Ongusaha PP, Miles PD, Havstad JC, Zhang F, So WV, Kudlow JE, Michell RH, Olefsky JM, Field SJ, et al. (2008). Phosphoinositide signalling links O-GlcNAc transferase to insulin resistance. Nature 451, 964–969. [PubMed: 18288188]

Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, et al. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol 9, R137. [PubMed: 18798982]

Highlights:

- Cell-surface IR translocates to the nucleus and associates with promoters genome-wide
- IR interaction with DNA is mediated by coregulator HCF-1 and transcription factors
- IR associates with Pol II and regulates gene expression
- Target genes are characteristic of insulin functions in physiology and disease

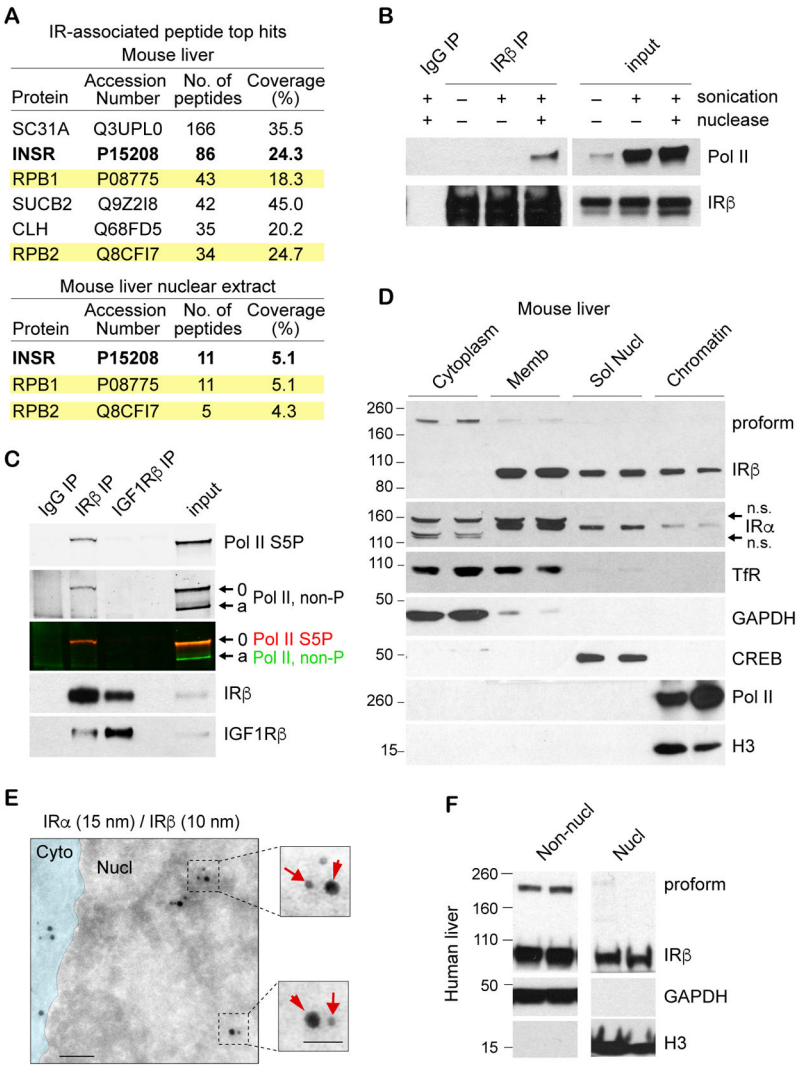


Figure 1. IR associates with RNA Polymerase II on chromatin.

(A) Proteins that co-IP with IRβ in mouse liver, or liver nuclear extract. The list shows the top hits identified by mass spectrometry, highlighting Pol II subunits RPB1 and RPB2 in yellow.

(B) Co-immunoprecipitation of IRβ and Pol II in HepG2 cells with or without sonication and nuclease treatment.

(C) Co-immunoprecipitation showing preferential association of IRβ with hyperphosphorylated Pol II. The CTD has 52 heptad repeats: the blot is probed with antibodies recognizing Ser5P (red) or non-phosphorylated (green) repeats, and shows hyperphophorylated (0) and hypophosphorylated (a) forms. The related receptor IGF1Rβ serves as a negative control that showed no evident association with Pol II.

(D) Mouse liver fractionation shows IRα and β in cytoplasmic, membrane, soluble nuclear, and chromatin-bound fractions. Other proteins are markers confirming effective fractionation. While Pol II is highly enriched in chromatin, a longer exposure shows its presence also in the soluble nuclear fraction. IRα non-specific bands are marked n.s.

(E) Immunogold EM labeling of nuclear IRβ (arrows) and IRα (arrowheads) proximity in HepG2 cells. Scale bar: 100 nm; 50 nm for insets.
(F) Human liver fractionation shows IRβ in non-nuclear and nuclear fractions. See also Figures S1 and S2.

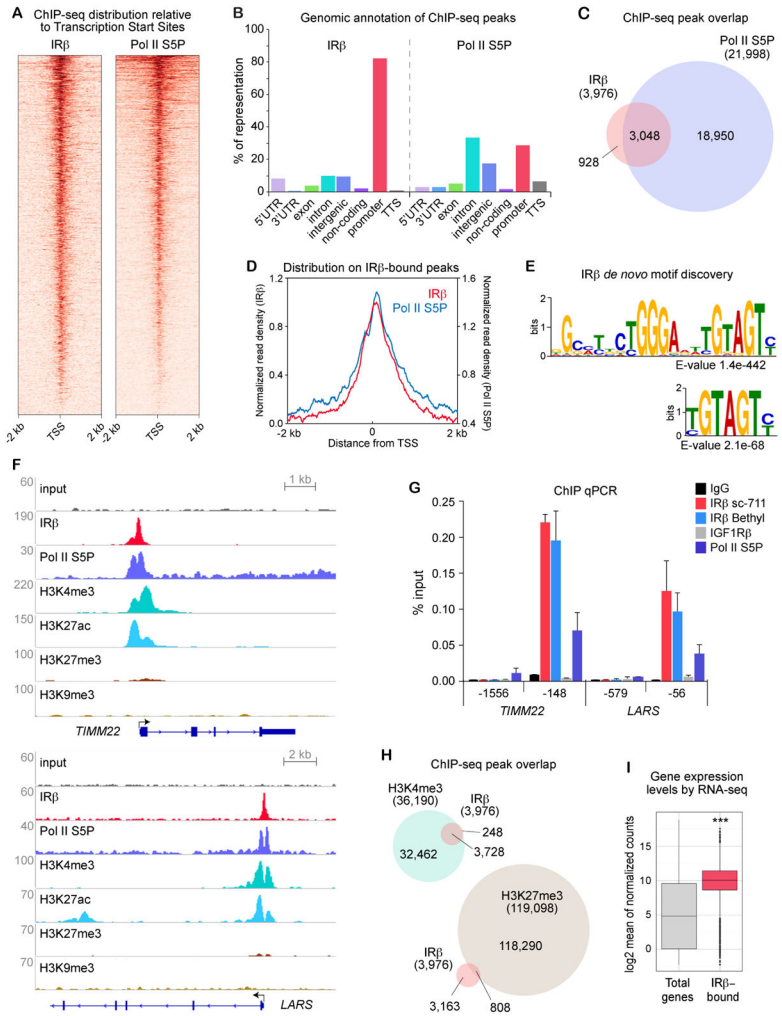


Figure 2. Genome-wide analysis reveals high enrichment of IR on gene promoters.

(A) Heatmaps of IRβ and Pol II S5P ChIP-seq peaks near the TSS in HepG2 cells. Raw read densities were used, and each horizontal line shows a separate IRβ-bound gene locus.

(B) IRβ ChIP-seq peaks classified by human genomic annotations (hg19).

(C) Overlap of IRβ and Pol II S5P ChIP-seq peaks.

(D) ChIP-seq density plot for IRβ and Pol II S5P at IRβ-bound loci.

(E) Top consensus sequences identified by *de novo* motif discovery at IRβ sites within promoters.

(F) ChIP-seq distribution for IRβ, Pol II S5P, and chromatin modifications, at representative gene loci, *TIMM22* and *LARS*. Histone modification data are from ENCODE Consortium.

(G) ChIP-qPCR confirmation of IRβ promoter binding for representative genes. Amplified DNA fragment positions (5' ends) are shown relative to the TSS. n=4.

(H) Overlap of IRβ with H3K4me3 or H3K27me3 ChIP-seq peaks genome-wide.

(I) Expression level of IRβ-bound targets compared to average total gene expression based on our RNA-seq data in HepG2 cells. ***P<0.001 (Mann Whitney Wilcoxon test).

See also Figure S3 and Table S1.

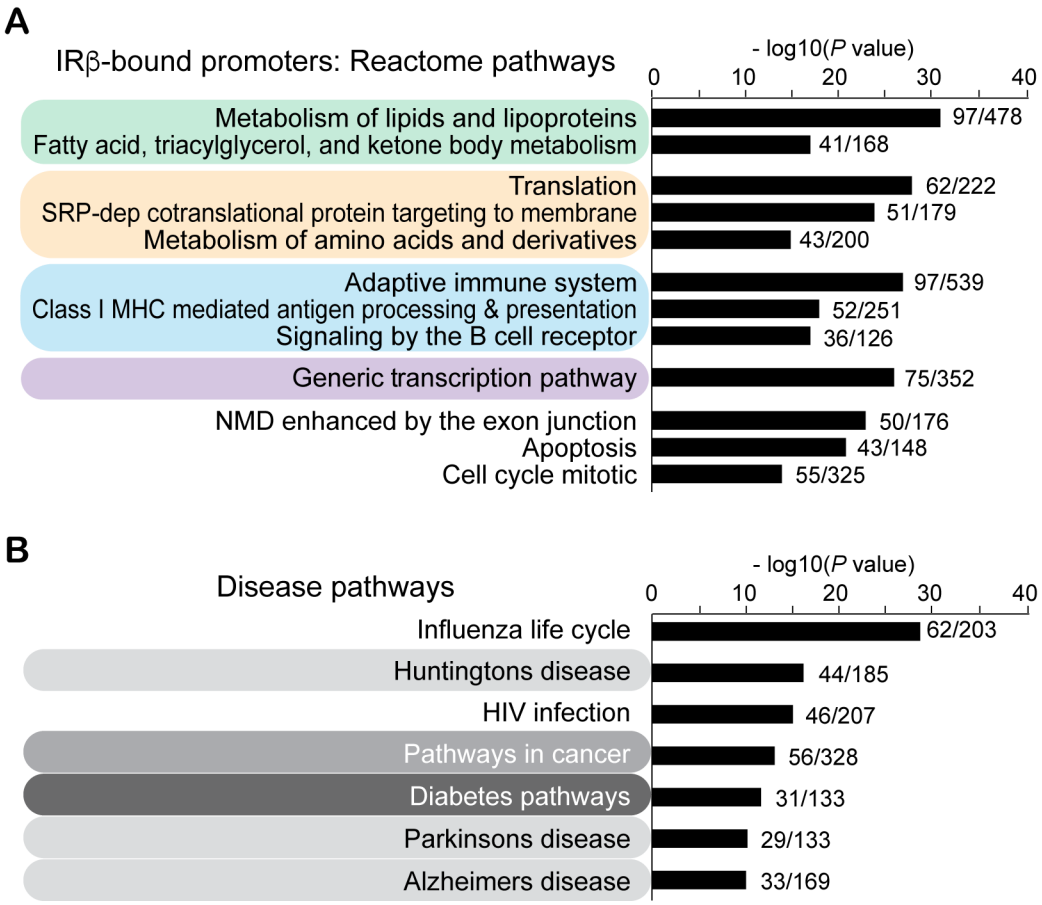


Figure 3. IR target genes are highly enriched for insulin-related functions.

(A) Top functional pathways of genes with IRβ-bound promoters, in Reactome database hierarchical levels 2 and 3, grouped according to related functional categories.

(B) Top disease pathways of genes with IRβ-bound promoters in MSigDB database. Numbers of IRβ-bound and total genes within each pathway are shown.

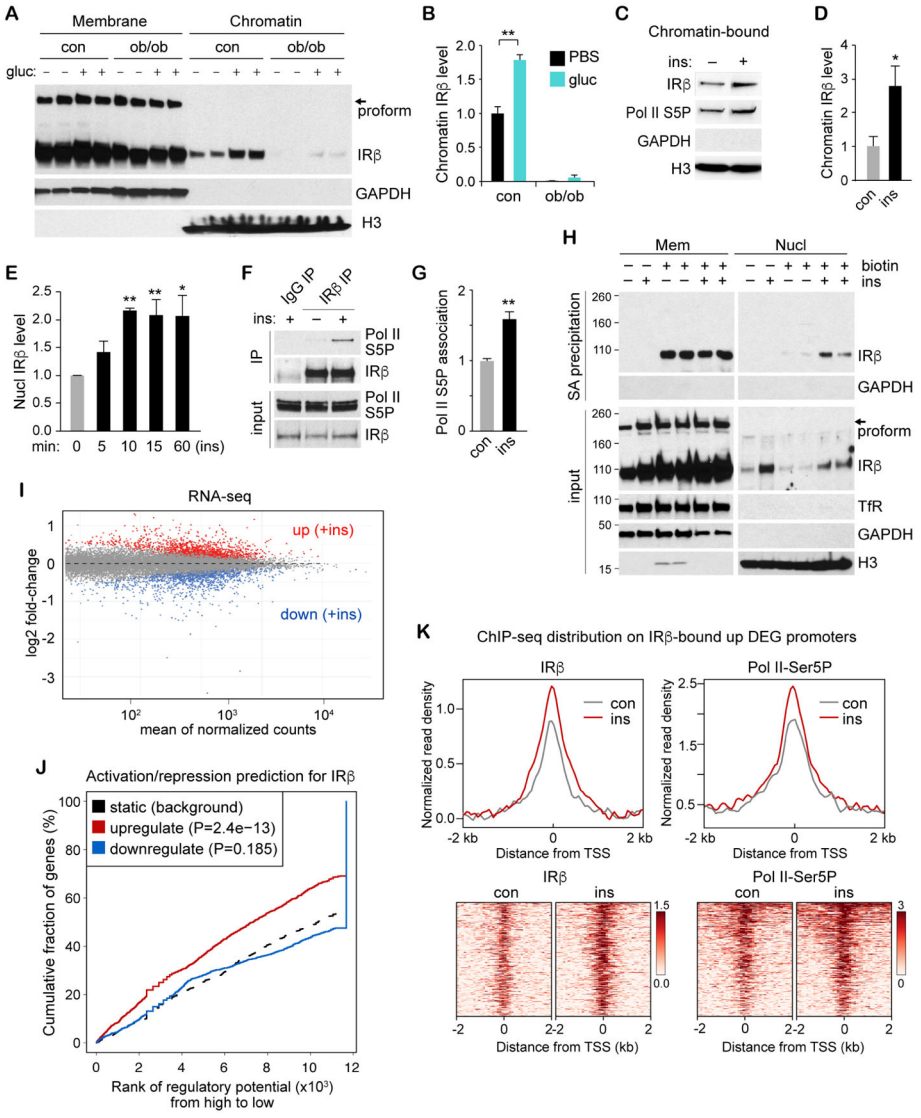


Figure 4. Insulin regulation of IR chromatin binding, dysregulation in insulin resistance, and insulin-regulated expression of IR-bound genes.

(A) Western analysis of IRβ in liver membrane or chromatin-bound fraction from wild-type or *ob/ob* mice injected with glucose or saline control solution.

(B) Quantitation of IRβ in chromatin-bound fraction in response to glucose. n=3, ** $P<0.01$ (two-tailed t-test).

(C) Western analysis of IRβ in liver chromatin-bound fraction from mice injected with insulin or saline control solution.

(D) Quantitation of chromatin-bound IRβ in response to insulin. n=5, * $P<0.05$ (two-tailed t-test).

(E) Quantitation of nuclear IRβ in response to insulin treatment over time in HepG2 cells. (See Figure S4E for immunoblot image.) n=3, * $P<0.05$, ** $P<0.01$ (vs. 0 min control, one-way ANOVA with Dunnett's post hoc analysis).

(F) Co-immunoprecipitation of IRβ and Pol II S5P in HepG2 cells with 10 min insulin or control treatment.

(G) Quantitation of Pol II S5P associated with IRβ. n=4, $^{**}P<0.01$ (two-tailed t-test).
(H) Western analysis of cell surface-biotinylated IRβ in membrane and nuclear fractions of HepG2 cells with 10 min insulin or control treatment.
(I) MA plot of normalized counts showing differentially expressed genes (DEGs) with FDR<0.001 in response to 4h insulin treatment.
(J) Activating/repressive function prediction for IRβ by Binding and Expression Target Analysis (BETA). IRβ ChIP-seq sites are integrated with RNA-seq gene expression data from HepG2 cells with control or insulin treatment. Red and blue lines represent up-regulated and down-regulated genes; dashed line represents non-DEGs as background. Genes are ranked based on the regulatory potential scores of their IR binding sites, and significance of up- or down-regulated gene distributions compared to non-DEGs is determined by the Kolmogorov-Smirnov test.
(K) ChIP-seq peak distribution for IRβ and Pol II S5P on IRβ-bound genes that are up-regulated by insulin. Graphs shows average distribution, while horizontal lines on the heatmaps show individual promoters.
See also Figure S4.

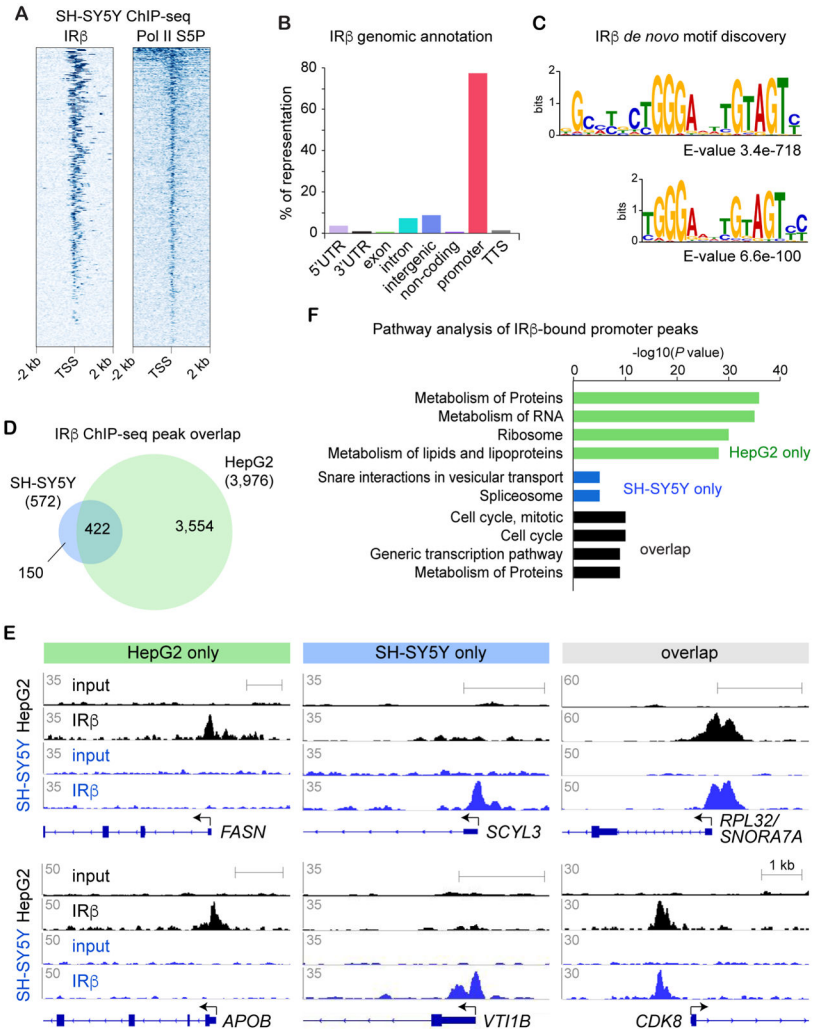


Figure 5. IR binding to promoters in SH-SY5Y neuroblastoma cells.

(A) Heatmaps of IRβ and Pol II S5P ChIP-seq peaks near the TSS in SH-SY5Y cells. Each horizontal line shows a separate IRβ-bound gene locus.

(B) IRβ ChIP-seq peaks classified based on human genomic annotations (hg19).

(C) Top consensus sequences identified by *de novo* motif discovery at IRβ sites within promoters.

(D) Overlap of IRβ ChIP-seq peaks between SH-SY5Y and HepG2 cells.

(E) ChIP-seq distribution for IRβ in HepG2 (black) and SH-SY5Y (blue) cells, showing examples of cell-type specific peaks, as well as those present in both cell types.

(F) Pathway analysis for genes with IRβ ChIP-seq promoter peaks (± 500 bp from TSS) that are unique to HepG2 or SH-SY5Y cells, or in both cell lines (overlap).

See also Table S2.

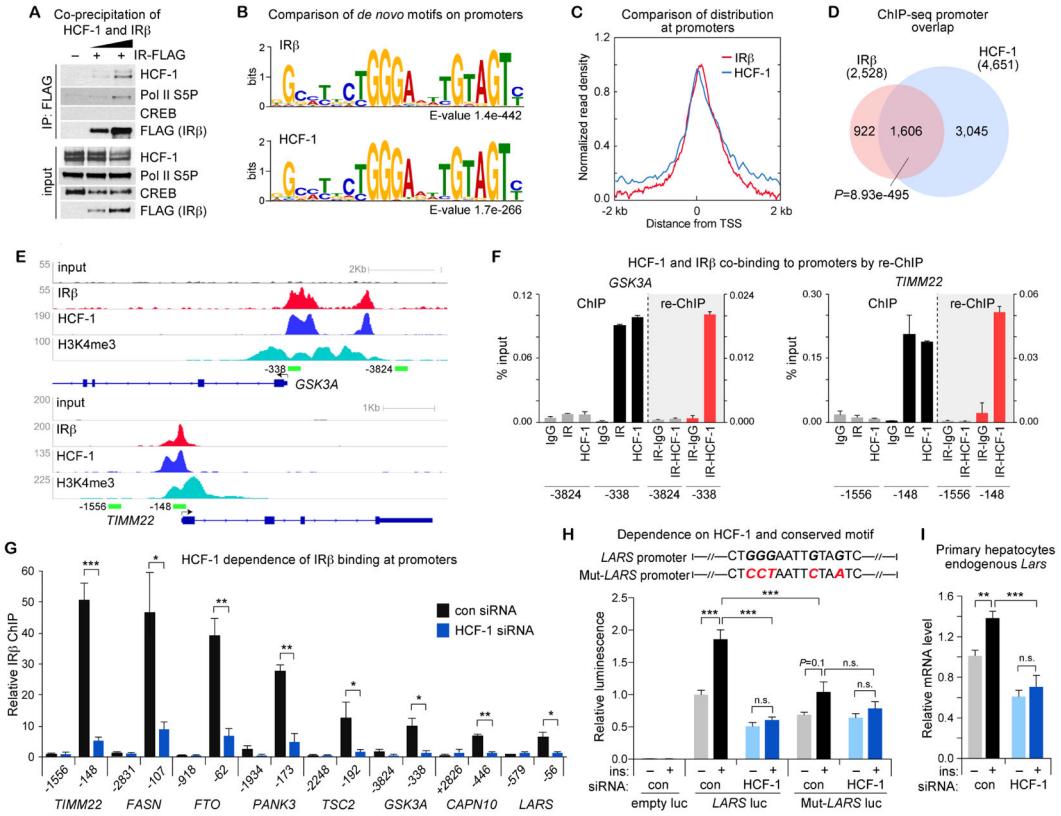


Figure 6. IR interacts physically and functionally with transcriptional coregulator HCF-1.

(A) Co-immunoprecipitation of IR (C-terminal FLAG tag), endogenous HCF-1 (antibody against the N-terminus, CST 69690), and endogenous Pol II S5P in cells expressing increasing concentrations of IR-FLAG. No association was seen with CREB negative control.

(B) Comparison of top consensus sequences identified by *de novo* motif discovery at HCF-1 or IRβ sites within promoters in HepG2 cells.

(C) ChIP density plots for IRβ and HCF-1 at IRβ-bound loci in HepG2 cells.

(D) Overlap of genes bound to IRβ and HCF-1 within promoter regions (±500 bp from TSS) in HepG2 cells. Significance calculated by hypergeometric test.

(E) ChIP-seq peak distribution for IRβ, HCF-1, and H3K4me3 at representative gene loci *GSK3A* and *TIMM22*.

(F) Sequential ChIP-qPCR using antibody against IRβ, followed by antibody against IgG or HCF-1 at *GSK3A* and *TIMM22* promoters and negative control distal regions (regions underlined in green in panel E). A control locus that binds HCF-1 but not IRβ is in Figure S5H. n=3.

(G) IRβ ChIP-qPCR in HepG2 cells transfected with HCF-1 or control siRNA. For each position, ChIP binding normalized to input is shown as fold-change to *TIMM22* negative control. n=3, **P*<0.05, ***P*<0.01, ****P*<0.001 (two-tailed t-test).

(H) *LARS* promoter-driven luciferase reporter in cells expressing HCF-1 or control siRNA, in response to 24h insulin or control treatment. *LARS* promoter mutations altered the IRβ and HCF-1 consensus motif as shown. n=5, **P*<0.05, ****P*<0.001 (Two-way ANOVA with Tukey's post hoc analysis).

(I) RT-qPCR of *Lars* expression in primary mouse hepatocytes expressing HCF-1 or control siRNA, and with 3h insulin or control treatment. n≥4, $^{**}P<0.01$, $^{***}P<0.001$ (Two-way ANOVA with Tukey's post hoc analysis).

See also Figures S5 and S6 and Tables S3 and S4.

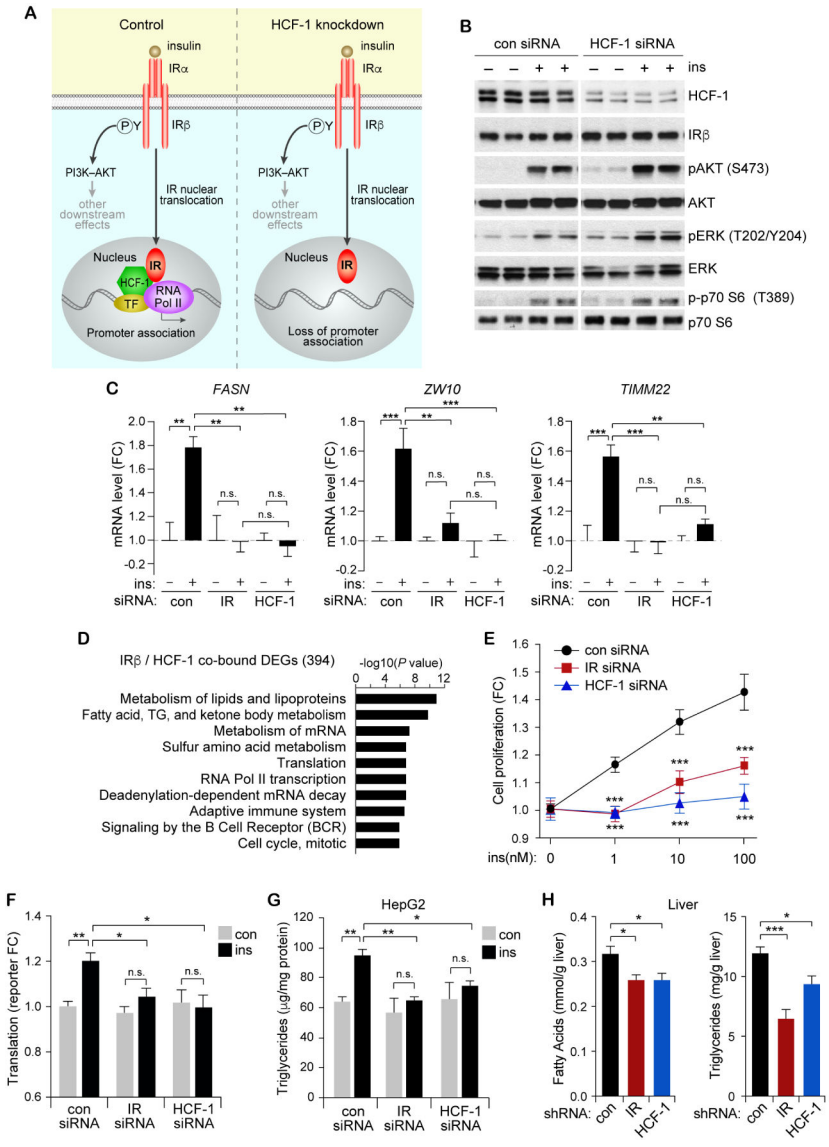


Figure 7. HCF-1-dependent signaling pathway mediates downstream effects of insulin.

(A) Left: In the model illustrated, insulin binding to IR activates canonical kinase signaling, as well as IR nuclear translocation and association with transcription machinery at gene promoters. Interactions of IR within the complex may potentially be direct or indirect. Right: Loss of HCF-1 prevents IR association with promoters, without causing obvious impairment of canonical PI3K-AKT kinase signaling.

(B) Western analysis of protein phosphorylation in HepG2 cells expressing HCF-1 or control siRNA, with 10 min insulin treatment. (See Figure S7A for quantitation.)

(C) Gene expression measured by RT-qPCR in response to 24h insulin treatment in HepG2 cells expressing IR, HCF-1, or control siRNA. Results are graphed for the top 3 insulin-responsive genes from Figure S7D. n=3, $^{**}P<0.01$, $^{***}P<0.001$ (Two-way ANOVA with Tukey's post hoc analysis).

(D) Top functional pathways of insulin-induced genes with promoters co-bound by IRβ and HCF-1 in HepG2 cells. Shown are categories in Reactome database hierarchical levels 2 and 3.

(E) Cell proliferation rates in HepG2 cells transfected with IR, HCF-1, or control siRNAs, and treated with increasing concentrations of insulin for 24h. n=10, ***P<0.001 (vs. control siRNA, Two-way ANOVA with Dunnett's post hoc analysis).

(F) Cap-dependent translation measured with a bicistronic luciferase reporter in cells transfected with IR, HCF-1, or control siRNA, in response to 24h insulin treatment. n=5, *P<0.05, **P<0.01 (two-tailed t-test).

(G) Triglyceride level in HepG2 cells expressing IR, HCF-1, or control siRNA, with 24h insulin treatment. *P<0.05, **P<0.01 (two-tailed t-test).

(H) Free fatty acid or triglyceride level in livers from mice injected with adeno-associated virus expressing IR, HCF-1, or control shRNA. n=5, *P<0.05, ***P<0.001 (vs control shRNA; One-way ANOVA with Dunnett's post hoc analysis).

See also Figure S7 and Table S5.

Key Resource Table

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
rabbit anti-phospho-AKT Ser473	Cell Signaling Technology	Cat# 9271; RRID:AB_329825
rabbit anti-AKT	Cell Signaling Technology	Cat# 9272; RRID:AB_329827
rabbit anti-calnexin	Cell Signaling Technology	Cat# 2433; RRID:AB_2243887
rabbit anti-CD71	Cell Signaling Technology	Cat# 13208
rabbit anti-CREB	Cell Signaling Technology	Cat#9197; RRID:AB_331277
goat anti-DCC	Santa Cruz Biotechnology	Cat# sc-6535; RRID:AB_2245770
rabbit anti-dykddddk (FLAG)	Cell Signaling Technology	Cat# 2368; RRID:AB_2217020
mouse anti-FLAG M2 magnetic beads	Sigma-Aldrich	Cat# M8823; RRID:AB_2637089
goat anti-GABBR1	Santa Cruz Biotechnology	Cat# sc-7338; RRID:AB_640741
rabbit anti-GAPDH	Cell Signaling Technology	Cat#2118; RRID:AB_561053
mouse anti-HA.11	Covance	Cat#MMS-101P; RRID:AB 2314672
rabbit anti-HA-Tag (71D10)	Cell Signaling Technology	Cat# 3724; RRID:AB_1549585
mouse anti-HA magnetic dynabeads	Thermo Fisher Scientific	Cat# 88836
rabbit anti-Histone H3	Cell Signaling Technology	Cat# 9717; RRID:AB_331222
rabbit anti-HCF-1	Bethyl Laboratories	Cat# A301–399A; RRID:AB_961012
rabbit anti-HCF-1	Cell Signaling Technology	Cat# 69690
mouse anti-HCF-1	Santa Cruz Biotechnology	Cat# sc-390950
rabbit anti-HSP70	Abcam	Cat# ab5439; RRID:AB_304888
rabbit anti-IGF1Rβ	Cell Signaling Technology	Cat# 3027; RRID:AB_2122378
mouse anti-IGF1Rβ	Millipore	Cat# 05–656; RRID:AB_309881
normal rabbit IgG	Santa Cruz Biotechnology	Cat# sc-2027
normal mouse IgG	Santa Cruz Biotechnology	Cat# sc-3877
rabbit anti-IRα	Santa Cruz Biotechnology	Cat# sc-710; RRID:AB_631106
mouse anti-IRα	Thermo Fisher Scientific	83–14; Cat#MA5–13759; RRID:AB_10985953
rabbit anti-IRβ	Bethyl Laboratories	Cat# A303–712A; RRID:AB_11205634
rabbit anti-IRβ	Santa Cruz Biotechnology	Cat# sc-711; RRID:AB_631835

REAGENT or RESOURCE	SOURCE	IDENTIFIER
rabbit anti-IRβ	Cell Signaling Technology	Cat# 3025; RRID:AB_2280448
mouse anti-IRβ (Ab-5)	Thermo Fisher Scientific	Cat# MS-635-P0; RRID:AB_142338
mouse anti-IRβ (Ab-6)	Thermo Fisher Scientific	Cat# MS-636-P0; RRID:AB_142351
rabbit anti-KPNA2	Cell Signaling Technology	Cat# 14372
rabbit anti-phospho-p44/42 MAPK Thr202/Tyr204	Cell Signaling Technology	Cat# 4377; RRID:AB_331775
rabbit anti-MAPK	Cell Signaling Technology	Cat# 9102; RRID:AB_330744
rabbit anti-mCherry	Abcam	Cat# ab167453; RRID:AB_2571870
mouse anti-c-Myc Tag	Thermo Fisher Scientific	9E10; Cat#MA1–980; RRID:AB_558470
rabbit anti-Myc-Tag	Cell Signaling Technology	Cat# 2278 RRID: AB_490778
mouse anti-c-Myc magnetic dynabeads	Thermo Fisher Scientific	Cat# 88842
Rabbit anti-OGT	Cell Signaling Technology	Cat# 24083 RRID: AB_2716710
rat anti-Pol II unphosphorylated	ChromoTek	Cat# 1C7; RRID:AB_2631402
rat anti-Pol II Ser2P	ChromoTek	Cat# 3E10; RRID:AB_2631403
rat anti-Pol II Ser5P	ChromoTek	Cat# 3E8; RRID:AB_2631404
rat anti-Pol II Ser7P	ChromoTek	Cat# 4E12; RRID:AB_2631443
rat anti-Pol II Tyr1P	Millipore	3D12;Cat#MABE350
mouse anti-RPB1 CTD (Ser5P, 4H8)	Cell Signaling Technology	4H8; Cat# 2629; RRID:AB_2167468
mouse anti-Pol II (8WG16)	Santa Cruz Biotechnology	Cat# sc-56767 RRID:AB_785522
rabbit anti-phospho-p70 S6 kinase Thr389	Cell Signaling Technology	Cat# 9205; RRID:AB_330944
rabbit anti-p70 S6 kinase	Cell Signaling Technology	Cat# 9202; RRID:AB_331676
rabbit anti-T7-Tag	Cell Signaling Technology	Cat# 13246
rabbit anti-THAP11	Cell Signaling Technology	Cat# 12305
mouse anti-THAP11	R&D Systems	Cat# MAB5727; RRID:_0717826
Bacterial and Virus Strains		
Adeno-Associated Virus serotype 8 (AAV8) encoding shRNA against IR: CCCTGAAGGATGGAGTCTTTA	This paper	System Biosciences Inc.
Adeno-Associated Virus serotype 8 (AAV8) encoding shRNA against HCF-1: TGGCTATCAAGGAGCTTATAG	This paper	System Biosciences Inc.
Adeno-Associated Virus serotype 8 (AAV8) encoding scrambled control shRNA TGGCTATCAAGGAGCTTATAG	This paper	System Biosciences Inc.
Biological Samples		

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Human liver, flash frozen biopsy	Sekisui XenoTech LLC	Sample ID H0923
Human liver, flash frozen biopsy	Sekisui XenoTech LLC	Sample ID H1296
Chemicals, Peptides, and Recombinant Proteins		
D-(+)-Glucose	Sigma-Aldrich	Cat# G7528
human insulin	Sigma-Aldrich	Cat# I9278
insulin (Humulin R)	Eli Lilly	N/A
human IGF-1	Sigma-Aldrich	Cat# I3769
Dexamethasone	Sigma-Aldrich	Cat# D4902
Liver digest medium	Thermo Fisher Scientific	Cat# 17703034
Percoll	Sigma-Aldrich	Cat# P1644
EMEM medium	ATCC	Cat# 30–2003
F-12K medium	ATCC	Cat# 30–2004
DMEM medium	Thermo Fisher Scientific	Cat# MT10013CV
Fetal bovine serum	ATCC	Cat# 30–2020
Opti-MEM	Thermo Fisher Scientific	Cat# 31985070
EZ-Link Sulfo-NHS-SS-Biotin	Thermo Fisher Scientific	Cat# 21331
Dynabeads Protein A	Thermo Fisher Scientific	Cat# 10001D
Dynabeads Protein G	Thermo Fisher Scientific	Cat# 10003D
Streptavidin-coupled dynabeads	Thermo Fisher Scientific	Cat# 11205D
3X FLAG peptide	Sigma-Aldrich	Cat# F4799
complete protease inhibitor	Roche	Cat# 11697498001
PhosSTOP phosphatase inhibitor	Sigma-Aldrich	Cat# 4906845001
RQ1 Dnase I	Promega	Cat# M6101
Benzonase	Sigma-Aldrich	Cat# E8263
Micrococcal nuclease	Thermo Fisher Scientific	Cat# EN0181
poly(dl-dC)	Sigma-Aldrich	Cat# 10108812001
Ionic Detergent Compatibility Reagent	Thermo Fisher Scientific	Cat# 22663
Pierce 660 nm Protein Assay Reagent	Thermo Fisher Scientific	Cat# 22660
siPORT NeoFX	Thermo Fisher Scientific	Cat# AM4510
Lipofectamine RNAiMax	Thermo Fisher Scientific	Cat# 13778150
Lipofectamine 2000	Thermo Fisher Scientific	Cat# 11668019
Critical Commercial Assays		
Subcellular Protein Fractionation Kit for Cultured Cells	Thermo Fisher Scientific	Cat# 78840
Protein Subcellular Fractionation Kit for Tissues	Thermo Fisher Scientific	Cat# 87790
MinElute PCR purification kit	Qiagen	Cat# 28004
Thruplex DNAseq kit	Rubicon	Cat# R400427
Truseq stranded mRNA kit	Illumina	Cat# 20020594
RNeasy Plus Mini kit	Qiagen	Cat# 74134
Superscript III First Strand Synthesis System	Thermo Fisher Scientific	Cat# 18080051

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Power SYBR green PCR master mix	Thermo Fisher Scientific	Cat# 4367659
Q5 Site-Directed Mutagenesis Kit	New England BioLabs	Cat# E0554S
TagMaster Site-directed Mutagenesis Kit	GM Biosciences	Cat# GM7002
Dual-Glo Luciferase Assay System	Promega	Cat# E2920
Cell Counting Kit-8 (CCK8) colorimetric assay	Dojindo Molecular Technologies	Cat# CK04–01
Triglyceride Colorimetric Assay Kit	Cayman Chemical	Cat# 10010303
Total Cholesterol/Cholesteryl Ester Quantitation Kit II	BioVision	Cat# K623
Free Fatty Acid Quantification Colorimetric/Fluorometric Kit	BioVision	Cat# K612
Deposited Data		
ChIP-seq and RNA-seq sequencing data	This paper	GEO:GSE107336
Raw data files for images	This paper	http://dx.doi.org/10.17632/6tsjbsyyt7.1
HepG2 ChIP-seq dataset: H3K4me3	ENCODE Project Consortium	ENCSR575RRX
HepG2 ChIP-seq dataset: H3K9me3	ENCODE Project Consortium	ENCSR000ATD
HepG2 ChIP-seq dataset: H3K27ac	ENCODE Project Consortium	ENCSR000AMO
HepG2 ChIP-seq dataset: H3K27me3	ENCODE Project Consortium	ENCSR000AOL
HepG2 ChIP-seq dataset: THAP11-FLAG	ENCODE Project Consortium	ENCSR562POI
HepG2 ChIP-seq dataset: GABPA	ENCODE Project Consortium	ENCSR000BJK
HepG2 ChIP-seq dataset: YY1	ENCODE Project Consortium	ENCSR000BNT
HepG2 ChIP-seq dataset: HCF-1	ENCODE Project Consortium	ENCSR529JYA
HepG2 ChIP-seq dataset: ZNF143	ENCODE Project Consortium	ENCSR101FJT
Mouse ChIP-seq dataset: Foxo1	Shin et al., 2012	N/A
Experimental Models: Cell Lines		
Human HepG2 cells	ATCC	Cat# HB-8065
Human SH-SY5Y cells	ATCC	Cat# CRL-2266
Human HEK293 cells	ATCC	Cat# CRL-1573
Experimental Models: Organisms/Strains		
Mouse: C57BL/6J	The Jackson Laboratory	Stock# 000664; RRID:IMSR_JAX:000664
Mouse: B6.Cg-*Lep*^{ob}/J	The Jackson Laboratory	Stock # 000632; RRID:IMSR_JAX:000632
Oligonucleotides		
Primer sequences for ChIP-qPCR	Table S6	N/A
Primer sequences for RT-qPCR	Table S6	N/A
Primer sequences for molecular cloning, DNA affinity pulldown, and siRNA	Table S6	N/A
Human HCF-1 27mersiRNA duplexes	OriGene Technologies	Cat# SR302076
Human THAP11 27mersiRNA duplexes	OriGene Technologies	Cat# SR311427
Mouse HCF-1 27mer siRNA duplexes	OriGene Technologies	Cat# SR423371
Recombinant DNA		
pShuttle-CMV-hlR-B	Dr. Ronald Kahn laboratory	N/A

REAGENT or RESOURCE	SOURCE	IDENTIFIER
vector pcDNA3.1/myc-His	Thermo Fisher Scientific	
pcDNA3.1-IR-Myc/His, with C-terminal myc tag	This paper	N/A
pcDNA3.1-HA-IR-Myc/His, with N-terminal HA tag and C-terminal myc tag	This paper	N/A
pSV2-mCherry-Pol2 plasmid	Dr. David Spector (Mao et al., 2011)	N/A
C-terminal FLAG-tagged human IR	Dr. Josephine Egan (Kim et al., 2012)	N/A
C-terminal FLAG-tagged human kinase dead IR	Dr. Josephine Egan (Kim et al., 2012)	N/A
pEF1 a-Ronin-Flag-lres-Neo	Dejosez et al.,2010	Addgene #28020
pCGN-HCF-1 fl, with HA and c-myc tags	Wilson et al., 1993	Addgene #53309
pCMVTNT-T7-KPNA2	Kelley et al.,2010	Addgene #26678
pGL4.19[*luc2CP*/Neo] vector	Promega	Cat# E6741
pGL4.73[*hRluc*/SV40]	Promega	Cat# E6911
pRL-5'-HCV-IRES-FL	Krüger et al., 2001	N/A
Software and Algorithms		
ImageJ version 1.43r	U. S. National Institutes of Health	https://imagej.nih.gov/ij/
STRING v10	Szklarczyk et al., 2015	https://string-db.org/
GenePattem v3.9.10	Afgan et al., 2016	http://genepattern.org/
Bowtie2 2.2.8	Langmead and Salzberg, 2012	https://sourceforge.net/projects/bowtie-bio/files/bowtie2/2
MACS2 2.1.1	Zhang et al., 2008	https://pypi.python.org/pypi/MACS2
HOMER v4.6	Heinz et al., 2010	http://homer.ucsd.edu/homer/
MEME SUITE	Bailey et al., 2009	http://meme-suite.org/doc/meme-chip.html
PWMScan	Ambrosini et al., 2018	https://ccg.vital-it.ch/pwmtools/pwmscan.php
deepTools2.0	Ramirez et al., 2016	https://deeptools.readthedocs.io/en/latest/content/installat
Genomic regions enrichment of annotations tool (GREAT) v3.0.0	McLean et al., 2010	http://great.stanford.edu/public/html/
Molecular Signature Database (MsigDB) v6.2	Subramanian et al., 2005	http://software.broadinstitute.org/gsea/msigdb/annotate.js
Reactome database v57	Fabregat et al., 2016	http://www.reactome.org/
DAVID Bioinformatics Resources v6.7	Huang et al., 2009	https://david.ncifcrf.gov/home.jsp
bcbio-nextgen project v0.9.1a	bcbio	https://bcbio-nextgen.readthedocs.org/en/latest/
STARv2.4.1d	Dobin et al., 2013	https://github.com/alexdobin/STAR
featureCounts v1.4.4	Liao et al.,2014	http://subread.sourceforge.net/
DESeq2	Love et al., 2014	http://bioconductor.org/packages/release/bioc/html/DESe
BETA v1.0.7	Wang et al., 2013	http://cistrome.org/BETA/
Prism v7	Graph Pad Software	https://www.graphpad.com/scientific-software/prism/