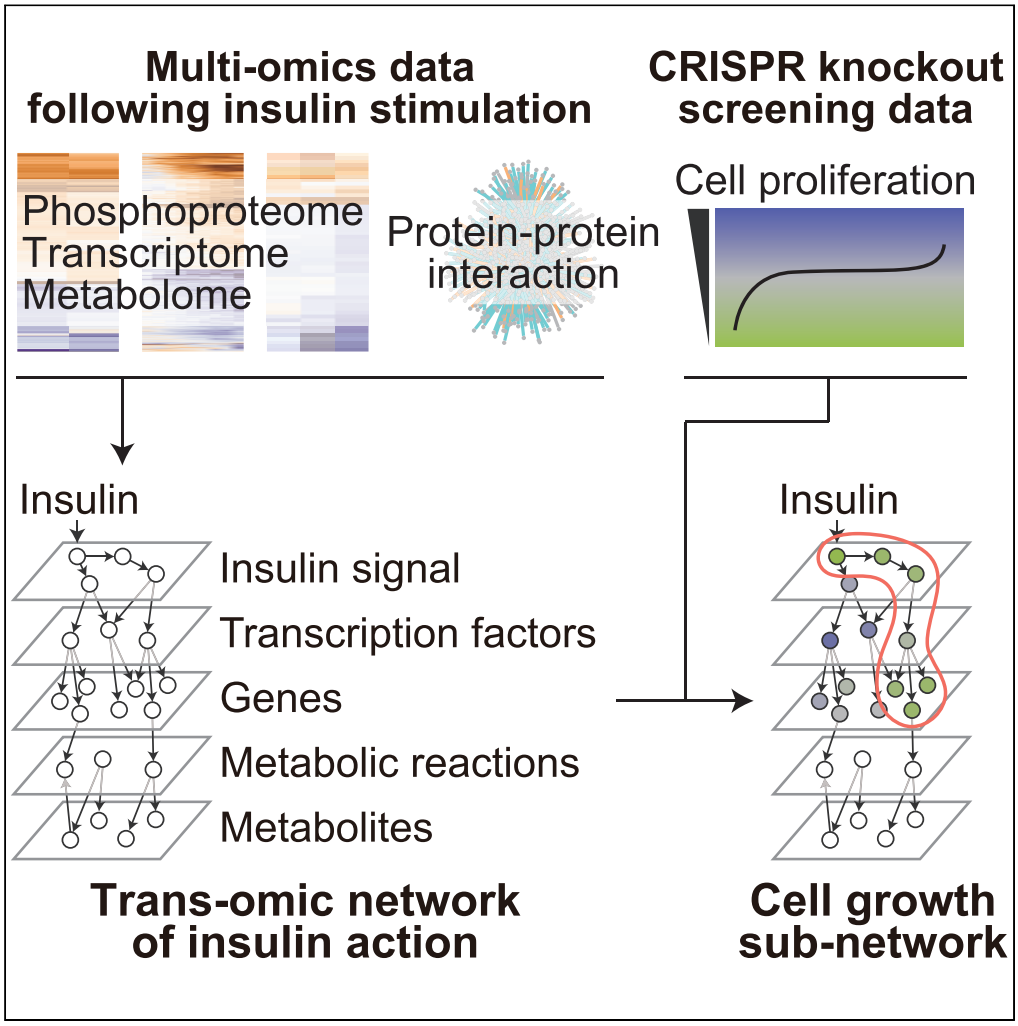
iScience

CellPress
OPEN ACCESS

Article

Trans-omics analysis of insulin action reveals a cell growth subnetwork which co-regulates anabolic processes



Akira Terakawa, Yanhui Hu, Toshiya Kokaji, ..., Martha L. Bulyk, Norbert Perrimon, Shinya Kuroda

perrimon@genetics.med.harvard.edu (N.P.)
skuroda@bs.s.u-tokyo.ac.jp (S.K.)

Highlights

A *trans*-omic network of insulin action in *Drosophila* cells was constructed

Insulin co-regulates various anabolic processes in a time-dependent manner

The *trans*-omic network and a CRISPR screen for cell proliferation were integrated

A Myc-mediated subnetwork promoting anabolic processes is required for cell growth

Terakawa et al., iScience *25*, 104231
May 20, 2022
https://doi.org/10.1016/j.isci.2022.104231

iScience

CellPress
OPEN ACCESS

Article

Trans-omics analysis of insulin action reveals a cell growth subnetwork which co-regulates anabolic processes

Akira Terakawa,[1,15] Yanhui Hu,[2,3,15] Toshiya Kokaji,[1,4] Katsuyuki Yugi,[1,5,6] Keigo Morita,[1] Satoshi Ohno,[1,7] Yifei Pan,[8] Yunfan Bai,[8] Andrey A. Parkhitko,[2,9] Xiaochun Ni,[2,10] John M. Asara,[11,12] Martha L. Bulyk,[10,13] Norbert Perrimon,[2,14,*] and Shinya Kuroda[1,7,8,16,*]

SUMMARY

Insulin signaling promotes anabolic metabolism to regulate cell growth through multi-omic interactions. To obtain a comprehensive view of the cellular responses to insulin, we constructed a trans-omic network of insulin action in *Drosophila* cells that involves the integration of multi-omic data sets. In this network, 14 transcription factors, including Myc, coordinately upregulate the gene expression of anabolic processes such as nucleotide synthesis, transcription, and translation, consistent with decreases in metabolites such as nucleotide triphosphates and proteinogenic amino acids required for transcription and translation. Next, as cell growth is required for cell proliferation and insulin can stimulate proliferation in a context-dependent manner, we integrated the trans-omic network with results from a CRISPR functional screen for cell proliferation. This analysis validates the role of a Myc-mediated subnetwork that coordinates the activation of genes involved in anabolic processes required for cell growth.

INTRODUCTION

The insulin signaling pathway regulates growth and metabolism and is highly evolutionarily conserved from *Drosophila* to mammals (Oldham and Hafen, 2003; Teleman, 2009). In both *Drosophila* and mammals, mutations in components of the insulin pathway are associated with diabetes and growth defects (Baker and Thummel, 2007; Oldham and Hafen, 2003; Teleman, 2009). Furthermore, dysregulation of insulin signaling causes systemic disorders such as dyslipidemia, hypertension, cardiovascular disease, stroke, blindness, kidney disease, female infertility, and neurodegeneration (White, 2003).

To regulate growth, insulin signaling coordinately promotes various anabolic metabolic processes such as glycogenesis, lipid synthesis, nucleic acid synthesis, and protein synthesis (Saltiel and Kahn, 2001; Valvezan and Manning, 2019; Zhu and Thompson, 2019). Insulin activates signaling molecules including the Akt kinase, extracellular-signal-regulated kinase (Erk), and mechanistic target of rapamycin (mTOR) by phosphorylation and protein–protein interactions (PPIs), and regulates downstream transcriptional and translational events required for anabolic metabolism. This growth program requires precise co-regulation of the number of metabolic enzymes through insulin signaling-dependent transcriptional and translational mechanisms.

There have been many attempts to characterize the insulin-regulated network of signaling molecules, transcription factors (TFs), metabolic enzymes, and other proteins regulating cellular functions such as proteins synthesis, and metabolites. Various "omic" studies of insulin action have been reported focusing on the phosphoproteome (Humphrey et al., 2013, 2015; Kawata et al., 2018, 2019; Krüger et al., 2008; Krycer et al., 2017; Matsuzaki et al., 2021; Monetti et al., 2011; Ohno et al., 2020; Vinayagam et al., 2016; Zhang et al., 2017), PPIs (Friedman et al., 2011; Glatter et al., 2011; Vinayagam et al., 2016), the transcriptome (Dupont et al., 2001; Hectors et al., 2012; Kawata et al., 2018; Kim and Lee, 2014; Matsuzaki et al., 2021; Rome et al., 2003; Sano et al., 2016; Versteyhe et al., 2013), and the metabolome (Everman et al., 2016; Kawata et al., 2018; Krycer et al., 2017; Matsuzaki et al., 2021; Noguchi et al., 2013; Ohno et al., 2020; Yugi et al., 2014). To provide a more comprehensive view than what can be gained from a single type of omic data

[1]Department of Biological Sciences, Graduate School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

[2]Department of Genetics, Blavatnik Institute, Harvard Medical School, 77 Avenue Louis Pasteur, Boston, MA 02115, USA

[3]Drosophila RNAi Screening Center, Department of Genetics, Harvard Medical School, 77 Avenue Louis Pasteur, Boston, MA 02115, USA

[4]Data Science Center, Nara Institute of Science and Technology, 8916-5 Takayama, Ikoma, Nara, Japan

[5]Laboratory for Integrated Cellular Systems, RIKEN Center for Integrative Medical Sciences, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

[6]Institute for Advanced Biosciences, Keio University, Fujisawa, 252-8520, Japan

[7]Molecular Genetics Research Laboratory, Graduate School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, Japan

[8]Department of Computational Biology and Medical Sciences, Graduate School of Frontier Sciences, The University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8562, Japan

[9]Aging Institute of UPMC and the University of Pittsburgh, Pittsburgh, PA, USA

[10]Division of Genetics, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA 02115, USA

Continued

alone, we have previously proposed "trans-omics" as a discipline for constructing molecular interaction networks from multi-omic data sets using direct molecular interactions rather than indirect statistical relationships (Yugi and Kuroda, 2017; Yugi et al., 2014, 2016). For example, in our previous studies, we constructed trans-omic networks responding to insulin stimulation in mammalian cells (Kawata et al., 2018; Kokaji et al., 2020; Matsuzaki et al., 2021; Ohno et al., 2020; Yugi et al., 2014), and responding to glucose administration in the liver of healthy and obese mice (Kokaji et al., 2020). Although these studies have expanded our view of insulin signaling, functional aspects (i.e., effects on phenotypes) remain to be explored for a large part of the network components (i.e., molecules and the regulatory relationships between them) and their relationships with phenotypes.

Insulin signaling is required for cell growth and proliferation in both *Drosophila* and mammals in a context-dependent manner (Straus, 1981; Wu et al., 2007). For example, some of the effects of insulin on cell proliferation in culture require non-physiological concentrations of insulin (Straus, 1981), and hormones and nutrients in the culture medium can affect the growth stimulatory effects of insulin (Straus, 1981; Wu et al., 2007). In *Drosophila* cells, insulin signaling can promote cell proliferation by inhibiting Foxo, which is a TF that causes cell-cycle arrest (Puig et al., 2003). However, insulin stimulation significantly promotes cell proliferation only when Pvr, a receptor tyrosine kinase regulating cell proliferation, is deficient in *Drosophila* Kc cells (Sopko et al., 2015). Furthermore, although insulin promotes cell growth (March and Bentley, 2006; Wu et al., 2007), insulin inhibits cell cycle progression through G2/M (Wu et al., 2007), indicating that insulin regulates cell proliferation in a context-dependent manner.

Here, we constructed a trans-omics network of insulin action that regulates gene expression and metabolism in *Drosophila* S2R + cells by integrating PPI, phosphoproteomic, transcriptomic, and metabolomic data following insulin stimulation. Next, as cell growth is required for cell proliferation and insulin can stimulate proliferation in a context-dependent manner, we developed a framework for integrating the trans-omic network with data from a genome-scale cell proliferation screen using CRISPR (Viswanatha et al., 2018). This analysis highlights, in particular, the role of the Myc TF in cell growth through coordinated activation of genes involved in anabolic processes.

RESULTS

Construction of the trans-omic network regulated by insulin and the integration with CRISPR knockout screening data

Using time-series data of PPIs, phosphoproteome, transcriptome, and metabolome (Figure 1A), together with various bioinformatic resources (Figure 1B), we constructed a regulatory trans-omic network for insulin-responsive gene expression and metabolic reactions in *Drosophila* S2R + cells ("trans-omic network of insulin action"; Figure 1C). The trans-omic network is composed of five layers: insulin signaling molecules (insulin signal), TFs, insulin-responsive genes (IRGs), metabolic reactions, and insulin-responsive metabolites (IRMs), with intra/inter-layer regulatory connections between molecules in the same and different layers (Figure 1C). We analyzed time-series multi-omic data (PPI, phosphoproteomic, transcriptomic, and metabolomic) measured in insulin-stimulated *Drosophila* S2R + cells (Figure 1D). Whereas PPI, phosphoproteomic, and transcriptomic data sets were obtained from previous studies [PPIs of 20 insulin signaling molecules by affinity purification-MS (AP-MS) (Vinayagam et al., 2016); phosphoproteome LC-MS data (Vinayagam et al., 2016), and transcriptome (Zirin et al., 2019)], we also generated a new dataset of metabolome data by liquid chromatography-mass spectrometry (LC-MS) following insulin stimulation (STAR Methods). The metabolome and transcriptome were measured up to 180 min after insulin stimulation and the phosphoproteome and the PPIs were measured up to 30 min after insulin stimulation (Figure 1D). We defined as "insulin-responsive" molecules that were quantitatively changed by insulin stimulation in the phosphoproteome, transcriptome, or metabolome data sets. We also determined the direction of insulin-responsive molecules as either increased or decreased in the amount based on the maximum and minimum absolute $\log_2$ fold change values across time-series compared to time 0 (Figure 1E).

Construction of the trans-omic network involved five steps (Figure 1C). In Step I, we identified IRGs from RNA-seq data sets. In Step II, we predicted TFs regulating clusters of IRGs using information from a TF binding motif database and a motif scanning software including Cis-BP and FIMO (Grant et al., 2011; Weirauch et al., 2014). In Step III, we predicted upstream signaling pathways regulating the TFs using phosphoproteomic data (Vinayagam et al., 2016), PPI data of 20 insulin signaling molecules (Vinayagam et al., 2016), the MIST PPI database (Hu et al., 2018), and the NetPhorest software for kinase prediction (Horn et al., 2014;

[11]Division of Signal Transduction, Beth Israel Deaconess Medical Center, Boston, MA 02115, USA

[12]Department of Medicine, Harvard Medical School, Boston, MA 02175, USA

[13]Department of Pathology, Brigham & Women's Hospital and Harvard Medical School, 77 Avenue Louis Pasteur, Boston, MA 02115, USA

[14]Howard Hughes Medical Institute, 77 Avenue Louis Pasteur, Boston, MA 02115, USA

[15]These authors contributed equally

[16]Lead contact

*Correspondence: perrimon@genetics.med.harvard.edu (N.P.), skuroda@bs.s.u-tokyo.ac.jp (S.K.)

https://doi.org/10.1016/j.isci.2022.104231

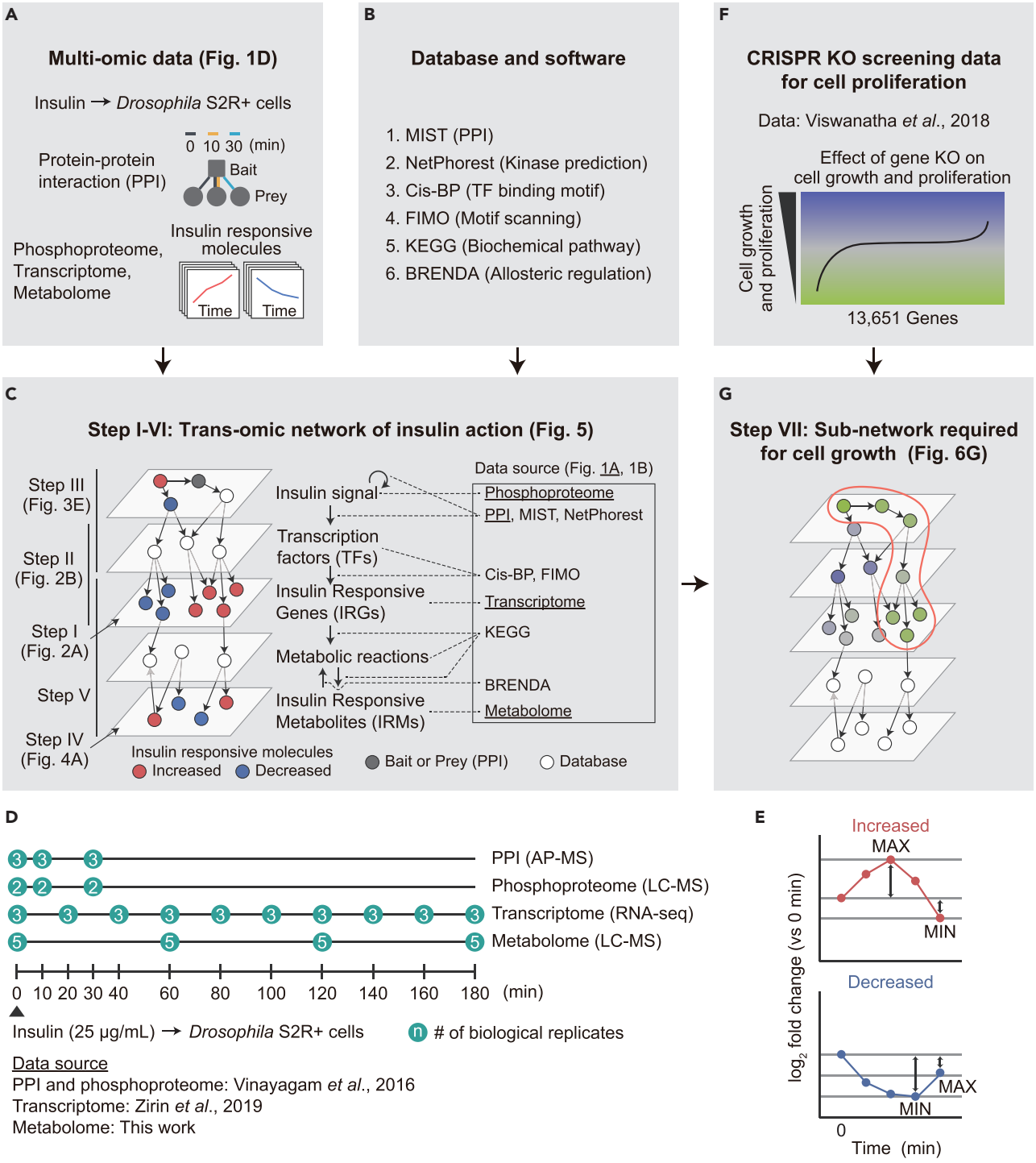


Figure 1. Construction of the trans-omic network regulated by insulin and the integration with CRISPR knockout screen data

(A–C) We used the time-series multi-omic data measured in insulin-stimulated *Drosophila* S2R + cells (A) and database and software (B) for constructing the trans-omic network of insulin action that involved Steps I–VI (C).

(D) Time-series multi-omic data measured in insulin-stimulated *Drosophila* S2R + cells. PPI data were measured for 20 insulin signaling molecules by affinity purification-mass spectrometry (Table S2) (Vinayagam et al., 2016).

Figure 1. *Continued*

(E) Definition of changes in the insulin-responsive molecules. We defined an insulin-responsive molecule as "increased" (red) when the absolute value of maximum (MAX) $\log_2$ fold change compared to time 0 (without insulin stimulation) across the time-series is larger than the absolute value of minimum (MIN) $\log_2$ fold change, and as "decreased" (blue) otherwise.

(F and G) We integrated the trans-omic network of insulin action and the CRISPR screen data for cell proliferation in *Drosophila* S2R + cells (Viswanatha et al., 2018) (F) to identify subnetworks required for cell growth (G).

Miller et al., 2008). In Step IV, we identified IRMs by metabolomic analysis. Finally, in Step V, we connected the IRMs/IRGs and the metabolic reactions using the KEGG database of metabolic pathways (Kanehisa et al., 2017) and connected the IRMs and metabolic reactions using the BRENDA database for allosteric regulations (Chang et al., 2021).

Next, we integrated the trans-omic network with results from a genome-wide pooled CRISPR knockout screen in *Drosophila* S2R + cells that identified 1,235 genes essential for cell proliferation (Figure 1F) (Viswanatha et al., 2018) to identify potential subnetworks required for cell growth (Figure 1G).

Step I: identification of IRGs

A previous RNA-seq study of insulin-stimulated *Drosophila* S2R + cells identified IRGs that were either increased or decreased following insulin stimulation (Table S1) (Zirin et al., 2019). This study focused on 163 genes, which is a subset of the IRGs that overlapped with 750 genes affecting nucleolar size from a previous RNAi screen (Neumüller et al., 2013). To further extend this study, we analyzed all the 1,212 IRGs (Figure 2A; Table S1). Based on our analysis (Figure 1E), 58.8% of the IRGs (713 genes) are increased, whereas 41.2% of the IRGs (499 genes) are decreased (Figure 2A).

Next, we clustered the IRGs by hierarchical clustering with Euclidean distance and Ward's method (Figure 2B; Table S1). We analyzed all clusters containing 100 genes or more in hierarchical clustering. Clusters 2 and 3 were obtained by dividing the dendrogram at the top level (Figure 2B). The majority of the genes in cluster 2 (96%) are increased IRGs. By contrast, the majority of the genes in cluster 3 (94%) are decreased IRGs. We categorized the IRG clusters into (i) cluster 2 and its sub-clusters as the increased clusters, and (ii) cluster 3 and its sub-clusters as the decreased clusters.

To identify the functional characteristics of each cluster, we performed a KEGG pathway enrichment analysis (q value <0.05, Fisher's exact test; Table S1). Because the clusters generated in this study were not mutually exclusive, a KEGG pathway can be enriched in overlapping clusters. Therefore, we selected non-overlapping and significant clusters for each KEGG pathway based on the significance (q value) of Fisher's exact test and the hierarchical structure of the dendrogram (STAR Methods) (Buehler et al., 2004). Increased (clusters 7, 9, 15, and 16) and decreased (clusters 3, 8, and 13) clusters were enriched for distinct pathways (upper right in Figure 2B). Increased clusters were enriched for multiple anabolic pathways such as Ribosome biogenesis in eukaryotes (cluster 7), Protein processing in endoplasmic reticulum (cluster 9), Spliceosome (cluster 9), Purine metabolism (cluster 15), Pyrimidine metabolism (cluster 15), RNA polymerase (cluster 15), and D-Glutamine and D-glutamate metabolism (cluster 16). Consistent with these observations, the insulin/mTOR pathway has been reported to promote the expression of ribosome biogenesis genes in both mammals and *Drosophila* using transcriptomic analyses of mutant or overexpressed cells or animals of components of the insulin/mTOR signaling pathway (PI3K, Foxo, Myc, S6K, or Tor), starved and fed conditions, or rapamycin-treated cells (Chauvin et al., 2014; Guertin et al., 2006; Li et al., 2010; Teleman et al., 2008). Decreased clusters were enriched for DNA replication (cluster 3), Tryptophan metabolism (cluster 8), and FoxO signaling pathway (cluster 13). This result is consistent with a previous report that Foxo is an evolutionarily conserved TF that is negatively regulated by insulin signaling (Puig et al., 2003).

Interestingly, increased and decreased clusters are enriched in distinct exclusive pathways, suggesting that IRGs belonging to the same pathway show a similar time course and are co-regulated by insulin signaling. To quantitatively evaluate whether IRGs in the same KEGG pathway are co-regulated in a time-dependent manner, we examined whether average Spearman correlation coefficients between the IRGs in the same KEGG pathway are higher than those of randomly sampled pairs from all the IRGs (Table S1; STAR Methods). Strikingly, the average Spearman correlation coefficients of the IRG pairs in the same KEGG pathways were significantly higher than randomly sampled pairs from the IRGs (Figure 2C). We also

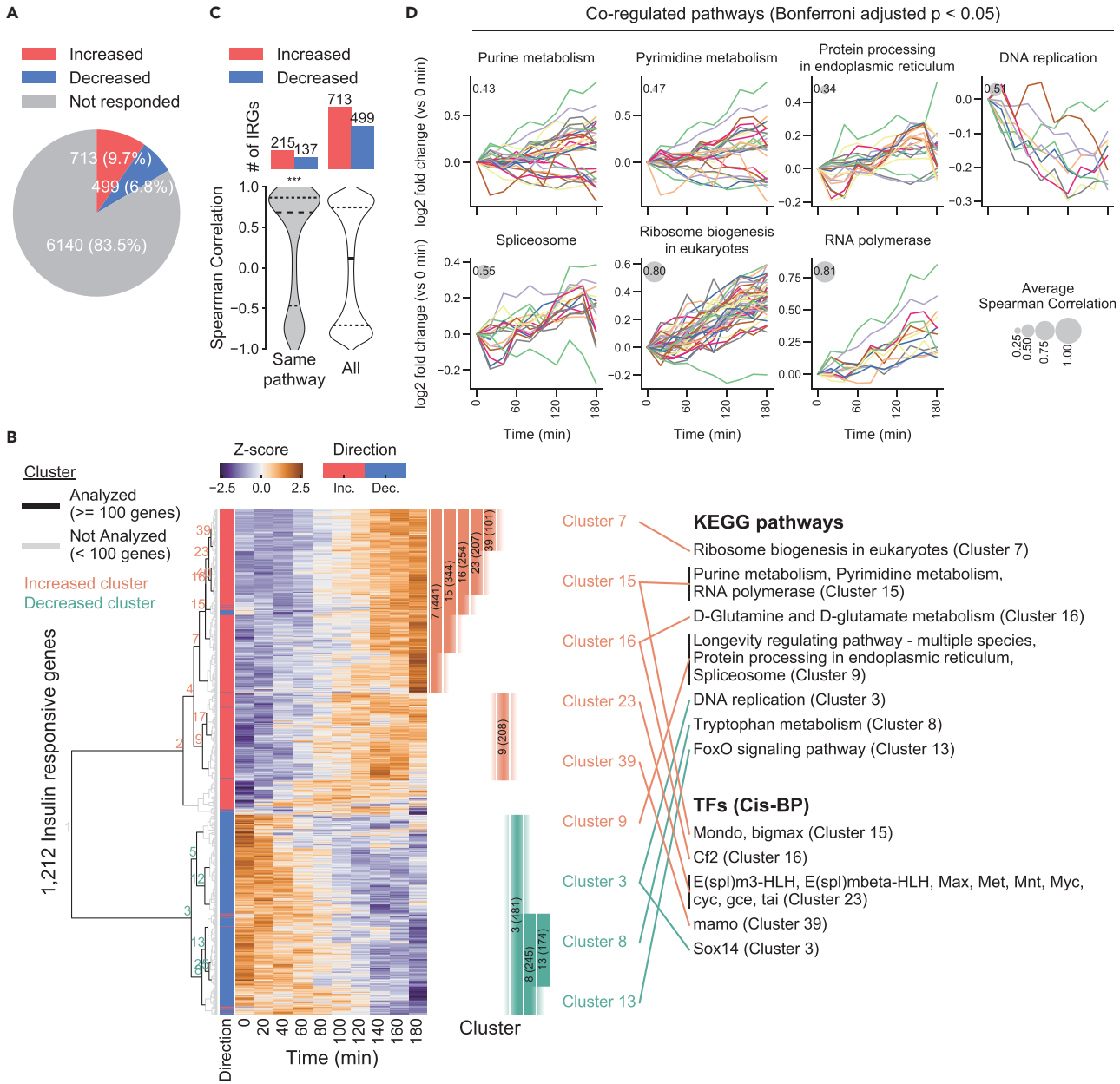


Figure 2. Identification of IRGs and prediction of TFs that regulate the IRGs

(A) Numbers of increased and decreased IRGs and not responsive genes. We divided the IRGs into increased and decreased IRGs as shown in Figure 1E. See also Table S1.

(B) Heatmap and hierarchical clustering of the *Z*-score normalized time course of gene expression of the IRGs. We performed hierarchical clustering using Euclidean distance and Ward's method. Numbers on the tree diagram indicate the cluster identity. The significantly enriched KEGG pathways and TF binding motifs are shown on the right. See also Table S1 and STAR Methods.

(C) Distribution of the Spearman correlation coefficients of gene pairs among the IRGs in the same KEGG pathway. We tested whether the average Spearman correlation coefficients of IRG pairs in the same KEGG pathway were significantly higher than randomly sampled IRG pairs. $*p < 0.05$, $**p < 0.01$, $***p < 0.001$. See also Table S1 and STAR Methods.

(D) Time course of the IRGs in significantly co-regulated KEGG pathways. Only significantly enriched pathways for any of the IRG clusters are shown. Data are shown as the $\log_2$ fold change values compared to time 0 that are calculated from the mean of biological replicates. Different line colors indicate different IRGs. We tested whether the average Spearman correlation coefficients of the IRG pairs in each of the KEGG pathways were significantly higher than randomly sampled IRG pairs. We analyzed KEGG pathways containing 10 or more IRGs. See also Figure S1A, Table S1, and STAR Methods.

examined the distributions of Spearman correlation coefficients between IRGs within individual KEGG pathways (Figures 2D and S1A; Table S1). For most of the anabolic pathways enriched for increased clusters such as Purine metabolism, Pyrimidine metabolism, Protein processing in the endoplasmic reticulum, Spliceosome, Ribosome biogenesis in eukaryotes, and RNA polymerase, the average Spearman correlations of the IRG pairs within the pathways were significantly higher than those of randomly sampled IRG pairs (Bonferroni-adjusted p value <0.05). Although we identified co-regulated pathways following insulin stimulation using the correlation analysis of the IRG pairs within the same KEGG pathway, it is possible that we underestimated the average Spearman correlation coefficients of some pathways that contain multiple functional modules (*e.g.* modules related to synthesis and degradation of a product within a metabolic pathway). Thus, we analyzed the Spearman correlation coefficients of IRG pairs within KEGG modules. Interestingly, both the average and median Spearman correlation coefficients of the IRG pairs within some KEGG modules were significantly higher than those within the KEGG pathways (Figure S1B; Table S1), suggesting the existence of co-regulated modules showing distinct time-series expression patterns from other modules within the same KEGG pathway. For instance, whereas the average Spearman correlation within the Purine and Pyrimidine metabolism pathways was relatively low (0.13 and 0.17, respectively; Figure 2D), we observed high Spearman correlation coefficients of IRG pairs within the Inosine monophosphate biosynthesis module in Purine metabolism and the Uridine monophosphate biosynthesis module in Pyrimidine metabolism, which are related to *de novo* purine synthesis and *de novo* pyrimidine synthesis, respectively (Figures S1C and S1D; Table S1). These results suggest that IRGs belonging to the same pathway are temporally co-regulated by insulin. Altogether, insulin signaling coordinately increased IRGs involved in anabolic pathways and coordinately decreased IRGs involved in DNA replication in a time-dependent manner.

Step II: prediction of TFs that regulate IRGs

Next, we predicted TFs regulating the IRG clusters using a motif enrichment analysis (Figure 2B; Table S1; STAR Methods). If a TF binding motif was enriched in the regions from −1,000 bp to +100 bp of the transcription start site of the genes in a cluster, we predicted the regulatory relationships between the TF and the genes in that cluster. We selected non-overlapping and significant clusters for each TF binding motif based on the significance (q value) of Fisher's exact test and the hierarchical structure of the dendrogram (STAR Methods) (Buehler et al., 2004).

We predicted TFs for the IRG clusters (lower right in Figure 2B). Interestingly, Myc was identified as a putative regulator of increased cluster 23 that is a sub-cluster of cluster 7 where ribosome biogenesis genes are enriched. This result is consistent with previous reports that Myc is regulated by insulin signaling in flies and ribosome biogenesis is one of the main cellular functions of Myc target genes (Grewal et al., 2005; Li et al., 2010; Orian et al., 2003; Teleman et al., 2008). Among the predicted TFs for any of the IRG clusters, gce, Myc, and tai were contained in the IRGs. To validate whether the predicted regulatory relationships are detected in ChIP-seq data, we compared our findings with the TF target gene regulatory relationships available in the ChIP-Atlas database (Table S1) (Oki et al., 2018). ChIP-Atlas contains ChIP-seq information for Myc and Max in *Drosophila* cell lines among the predicted TFs. The regulatory relationships registered in ChIP-Atlas for both factors significantly overlapped with our predictions (Bonferroni-adjusted p value < 0.05, Fisher's exact test; Figure S2). We also confirmed that KEGG pathways and TF motifs enriched in any of the analyzed clusters remained similar regardless of a subtle change of the threshold of cluster size around 100 genes (Figure S3). These results indicate that the TF prediction in our analysis is valid and robust.

Step III: prediction of upstream signaling pathways regulating the predicted TFs

We connected the TFs to insulin signaling using phosphoproteomic and PPI data (Vinayagam et al., 2016), NetPhorest software (Horn et al., 2014; Miller et al., 2008), and MIST PPIs (Hu et al., 2018) (Figure 3). We used phosphoproteomic data to infer possible kinases responsible for phosphorylated proteins with NetPhorest and Fisher's exact test. We used PPIs, detected by AP-MS measurement, to predict PPIs of TFs and insulin signaling molecules. Finally, we integrated the predicted kinase-substrate interactions, PPIs, and MIST PPIs, and constructed a merged network that allowed us to extract paths from the insulin receptor to the TFs.

We predicted the kinases responsible for the phosphorylated peptides. Specifically, we identified 268 phosphopeptides (193 proteins), whose phosphorylation levels were increased or decreased following

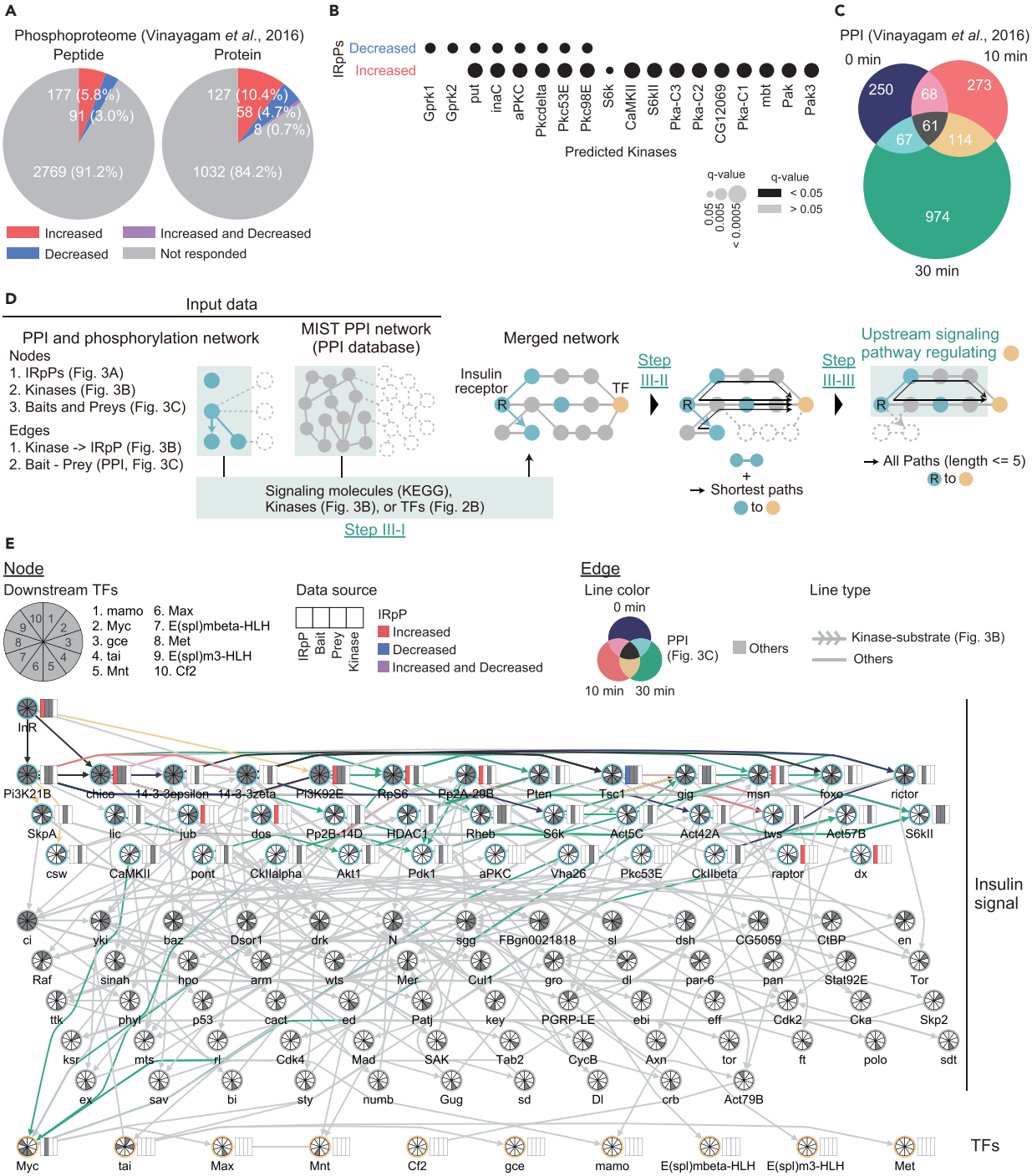

A
Phosphoproteome (Vinayagam et al., 2016)
Peptide
Protein
177 (5.8%)
91 (3.0%)
2769 (91.2%)
127 (10.4%)
58 (4.7%)
8 (0.7%)
1032 (84.2%)
Increased
Decreased
Increased and Decreased
Not responded
B
IRpPs
Decreased
Increased
Gprk1
Gprk2
put
inaC
aPKC
Pkcdelta
Pkc53E
Pkc98E
S6k
CaMKII
S6kII
Pka-C3
Pka-C2
CG12069
Pka-C1
mbt
Pak
Pak3
Predicted Kinases
q-value
0.05
0.005
< 0.0005
q-value
< 0.05
> 0.05
C
PPI (Vinayagam et al., 2016)
0 min
10 min
30 min
250
68
273
61
67
114
974
D
Input data
PPI and phosphorylation network
Nodes
1. IRpPs (Fig. 3A)
2. Kinases (Fig. 3B)
3. Baits and Preys (Fig. 3C)
Edges
1. Kinase -> IRpP (Fig. 3B)
2. Bait - Prey (PPI, Fig. 3C)
MIST PPI network
(PPI database)
Signaling molecules (KEGG),
Kinases (Fig. 3B), or TFs (Fig. 2B)
Step III-I
Merged network
Insulin
receptor
R
TF
Step
III-II
Shortest paths
to
Step
III-III
Upstream signaling
pathway regulating
All Paths (length <= 5)
to
E
Node
Downstream TFs
1. mamo
2. Myc
3. gce
4. tai
5. Mnt
6. Max
7. E(spl)mbeta-HLH
8. Met
9. E(spl)m3-HLH
10. Cf2
Data source
IRpP
Bait
Prey
Kinase
IRpP
Increased
Decreased
Increased and Decreased
Edge
Line color
0 min
10 min
30 min
PPI
(Fig. 3C)
Others
Line type
Kinase-substrate (Fig. 3B)
Others
InR
Pi3K21B
chico
14-3-3epsilon
14-3-3zeta
Pi3K92E
RpS6
Pp2A-29B
Pten
Tsc1
gig
msn
foxo
rictor
SkpA
lic
jub
dos
Pp2B-14D
HDAC1
Rheb
S6k
Act5C
Act42A
tws
Act57B
S6kII
csw
CaMKII
pont
CkIIalpha
Akt1
Pdk1
aPKC
Vha26
Pkc53E
CkIIbeta
raptor
dx
ci
yki
baz
Dsor1
drk
N
sgg
FBgn0021818
sl
dsh
CG5059
CtBP
en
Raf
sinah
hpo
arm
wts
Mer
Cul1
gro
dl
par-6
pan
Stat92E
Tor
ttk
phyl
p53
cact
ed
Patj
key
PGRP-LE
ebi
eff
Cdk2
Cka
Skp2
ksr
mts
rl
Cdk4
Mad
SAK
Tab2
CycB
Axn
tor
ft
polo
sdt
ex
sav
bi
sty
numb
Gug
sd
Dl
crb
Act79B
Insulin
signal
Myc
tai
Max
Mnt
Cf2
gce
mamo
E(spl)mbeta-HLH
E(spl)m3-HLH
Met
TFs
F
Hippo signaling pathway - multiple species
Hippo signaling pathway -fly
Mnt
q-value
0.05
0.005
< 0.0005
q-value
< 0.05
> 0.05

Figure 3. Prediction of upstream signaling pathways regulating the predicted TFs

(A) Number of phosphorylated peptides (left) and phosphorylated proteins (right) measured in the phosphoproteome obtained from Vinayagam et al. (2016). We divided the IRpPs into increased (red) and decreased (blue) IRpPs as shown in Figure 1E. See also Table S2.

(B) Enrichment analysis of possible substrates of each kinase for the increased and decreased IRpPs (q value <0.05, Fisher's exact test). See also Table S2 and STAR Methods.

(C) Number of PPIs measured at three different time points after insulin stimulation (0, 10, and 30 min after insulin stimulation) (Vinayagam et al., 2016). AP-MS data were measured for the 20 canonical insulin signaling proteins (Table S2). Note that the number of interactions, not the number of the preys, are shown. See also Table S2.

(D) Method used for predicting upstream signaling pathways regulating the predicted TFs. See also STAR Methods.

(E) Upstream signaling pathways regulating the predicted TFs. We predicted the signaling pathways from the insulin receptor to each of the TFs following the procedure shown in Figure 3D and obtained the insulin signal layer by merging the pathways. The border colors of the nodes in the layer correspond to the colors shown in Figure 3D. The pie charts on the nodes indicate the downstream TFs of the nodes. Data sources are shown on the right side of the nodes contained in the PPI and phosphorylation network with the color-coded for increased or decreased IRpPs (Figure 3A). The edges derived from the PPIs of Vinayagam et al. (2016) are colored by the detected time point(s) (Figure 3C), and the others are shown in grey. The network diagram was created using Cytoscape (Shannon et al., 2003). See also Table S2.

(F) KEGG signaling pathway enrichment analysis of the upstream signaling pathways of each predicted TF. See also Table S2 and STAR Methods.

insulin stimulation among the 3,037 phosphorylated peptides (1,225 proteins) measured in the phosphoproteomic data (Figure 3A; Table S2) (Vinayagam et al., 2016). According to our previous study, phosphopeptides that showed an absolute $\log_2$ fold change larger than 0.5 at any time point compared to time 0 were defined as the insulin-responsive phosphopeptides (IRpPs). We divided the IRpPs selected by this criterion into increased and decreased IRpPs using the method shown in Figure 1E, of which 177 are increased IRpPs (135 proteins), whereas 91 are decreased IRpPs (66 proteins). Eight proteins contained both increased and decreased IRpPs. Based on the amino acids sequences of the phosphopeptides measured in the phosphoproteomic data set, we predicted possible kinases that regulate the phosphopeptides using NetPhorest, a software for predicting kinase classifiers for input phosphopeptide sequences (Table S2; STAR Methods) (Horn et al., 2014; Miller et al., 2008). Using these predicted kinase–substrate relationships (KSRs), we tested whether the predicted substrates of each kinase were enriched in the increased and decreased IRpPs to identify significantly activated or inhibited kinases, respectively (q value <0.05, Fisher's exact test). In sum, 16 and 8 kinases were enriched in the increased and decreased IRpPs, respectively, including S6k, which is among the baits of the AP-MS measurement, and CamKII, which is a prey identified by the AP-MS measurement (Figure 3B). If the predicted target phosphopeptides of a kinase were enriched in the increased or decreased IRpPs, then we selected the KSRs between the kinase and its predicted substrates contained in the increased or decreased IRpPs, respectively, and used them for the following analysis (Table S2; STAR Methods).

Next, to infer PPIs of TFs and insulin signaling molecules, we identified PPIs of insulin-stimulated *Drosophila* S2R + cells from our previous study (Figure 3C) (Vinayagam et al., 2016). PPIs were measured by AP-MS for the 20 insulin signaling proteins at baseline (without insulin treatment) and after insulin stimulation for 10 or 30 min. The PPIs consist of 555 proteins and 1,807 interactions. We classified the PPIs into seven types based on the detected time point(s) according to our previous study (Figure 3C) (Vinayagam et al., 2016).

To construct the merged network, we integrated the IRpPs (Figure 3A), the predicted regulatory relationships from the kinases to the IRpPs (Figure 3B), PPIs (Figure 3C), and the MIST PPI network into a new network of insulin signaling from receptor to downstream TFs (Figure 3D). We first merged all the data from these data sources ("Input data" in Figure 3D) and extracted the subnetworks containing the signaling molecules extracted from KEGG, the kinases predicted by NetPhorest, or the predicted TFs to capture the comprehensive regulatory relationships from the insulin receptor to the TFs through the phosphorylation and the PPIs ("Step III-I" in Figure 3D).

We extracted paths from the insulin receptor to the TFs from the merged network ("Step III-II" and "Step III-III" in Figure 3D; Table S2; STAR Methods). We considered the extracted paths from the insulin receptor to a TF as the upstream signaling pathways of the TF because the intracellular insulin signaling originates from the activation of the insulin receptor and is transmitted through its downstream signaling cascade to the TF. We denoted nodes and edges in the union of the upstream signaling pathways for the TFs as the insulin signal. Among the 14 predicted TFs, 10 TFs were connected to the upstream signaling pathways ("TFs" in Figure 3E). The resulting insulin signal layer consists of 101 nodes and 357 edges (Table S2). In the insulin signal layer, we can distinguish the common upstream signaling molecules for all 10 TFs and the specific

upstream signaling molecules for a subset of the TFs by counting the number of downstream TFs for each of the nodes in the layer. In sum, 8 molecules (6.9%) were found in the upstream paths of all 10 TFs, whereas 65 molecules (64.4%) were upstream of two or more TFs, and 36 molecules (35.6%) were upstream of only one TF. The molecules that are commonly found in the upstream paths of all 10 TFs include InR, chico (*Drosophila* ortholog of insulin receptor substrates), Pi3K21B, and Pi3K92E. Importantly, our predictions captured the regulatory relationships that have been reported in previous studies. In particular, in *Drosophila* cells, Myc is stabilized by insulin/mTOR signaling followed by increased phosphorylation of Sgg/Gsk3-beta (Parisi et al., 2011). Consistent with this study, Sgg is present in the pathway upstream of Myc (Figures 3E, S4A, and S4B). We further analyzed which upstream signaling molecules directly regulate the TFs, as our method involves direct and indirect associations (interactions) between upstream signaling molecules and the TFs (Figure S4A). We found that RpS6, Msn, Act57B, CaMKII, CkIIalpha, Pont, and Vha26 directly interact with Myc only at 30 min following insulin stimulation (Figures S4A and S4B), suggesting that Myc is regulated by insulin signaling through the changes in the PPIs with these proteins. To identify the pathways enriched in the upstream signaling pathways of each TF, we performed pathway enrichment analysis using KEGG signaling pathway annotation (Fisher's exact test, q value <0.05; Table S2; STAR Methods). Interestingly, the Hippo pathway is enriched in the upstream signaling pathway of Mnt (Figure 3F), suggesting that insulin possibly regulates TF activities through non-canonical insulin signaling pathways.

Step IV: identification of IRMs

We measured the metabolome in *Drosophila* S2R + cells following insulin stimulation using LC-MS/MS (Table S3; STAR Methods). In sum, 236 metabolites were measured in three replicates or more at all time points (0, 60, 120, and 180 min). We confirmed the high reproducibility of our metabolomic measurements (>0.98 Pearson correlation coefficient between any two biological replicates; Figure S5), and identified 53 IRMs with 26.4% (14) and 73.6% (39) metabolites that increased and decreased, respectively (Figure 4A; Table S4). Many IRMs showed a monotonic increase or decrease (Figure 4B).

Next, we mapped the metabolites on the KEGG metabolic pathways (Figure 4C). Out of 236 metabolites (45 out of 53 IRMs), 200 were mapped on KEGG, and pathway enrichment analysis identified the metabolic pathways containing a significantly large number of the IRMs (Table S4). Although no pathways were significantly enriched in the increased or decreased IRMs (Fisher's exact test, q value <0.05), metabolites in purine metabolism, a sub-metabolic pathway of nucleotide metabolism, showed a moderate enrichment in the decreased IRMs (Fisher's exact test, uncorrected p value <0.05), and the 12 IRMs contained in this pathway were the highest of all the KEGG pathways (Figure S6). Also, eight IRMs mapped to the pyrimidine metabolism pathway that is another sub-pathway of nucleotide metabolism. In addition, IRMs contained TCA cycle intermediates (succinyl-CoA and 2-oxoglutarate decreased, and oxaloacetate increased), urea cycle intermediates (L-arginine, L-argininosuccinate, and L-ornithine decreased), proteinogenic amino acids (L-lysine, L-phenylalanine, and L-arginine decreased), and lactate (increased) (Figure 4C).

IRMs in purine and pyrimidine metabolism (nucleotide metabolism) pathways accounted for 12/32 (the IRMs in the pathway/the measured metabolites in the pathway) and 8/28 of the metabolites detected in each pathway (Figure S6). In addition, enzymes of purine and pyrimidine metabolism were enriched in the increased IRG clusters (Figure 2B). The majority of the IRMs in the nucleotide metabolism pathway decreased (15/20 metabolites), which contain nucleosides (Guanine, cytosine), NDP (GDP), NTPs (ATP, CTP, GTP, TTP, and UTP), and dNTPs (dATP, dCTP, dGTP, and dTTP). Increased IRMs in the nucleotide metabolism pathway contained dNMP (dTMP), dNDP (CDP), and N-carbamoyl-L-aspartate, which is a metabolite involved in *de novo* pyrimidine synthesis catalyzed by carbamoyl-phosphate synthetase2, aspartate transcarbamylase, dihydrooratase (CAD). In mammalian cells, it has been reported that insulin increases the metabolic flux of the reaction of CAD and the abundance of N-carbamoyl-L-aspartate (Ben-Sahra et al., 2013; Robitaille et al., 2013). These results indicate that NTPs and the proteinogenic amino acids, required for transcription and protein synthesis, decreased, which is consistent with the increase in gene expression of the enzymes involved in these processes.

Step V: connection of the IRMs/IRGs and metabolic reactions

We constructed the trans-omic network by connecting IRGs, metabolic reactions, and IRMs based on the annotation of enzymatic reactions in the KEGG database, and allosteric regulations from the BRENDA database (STAR Methods) (Chang et al., 2021).

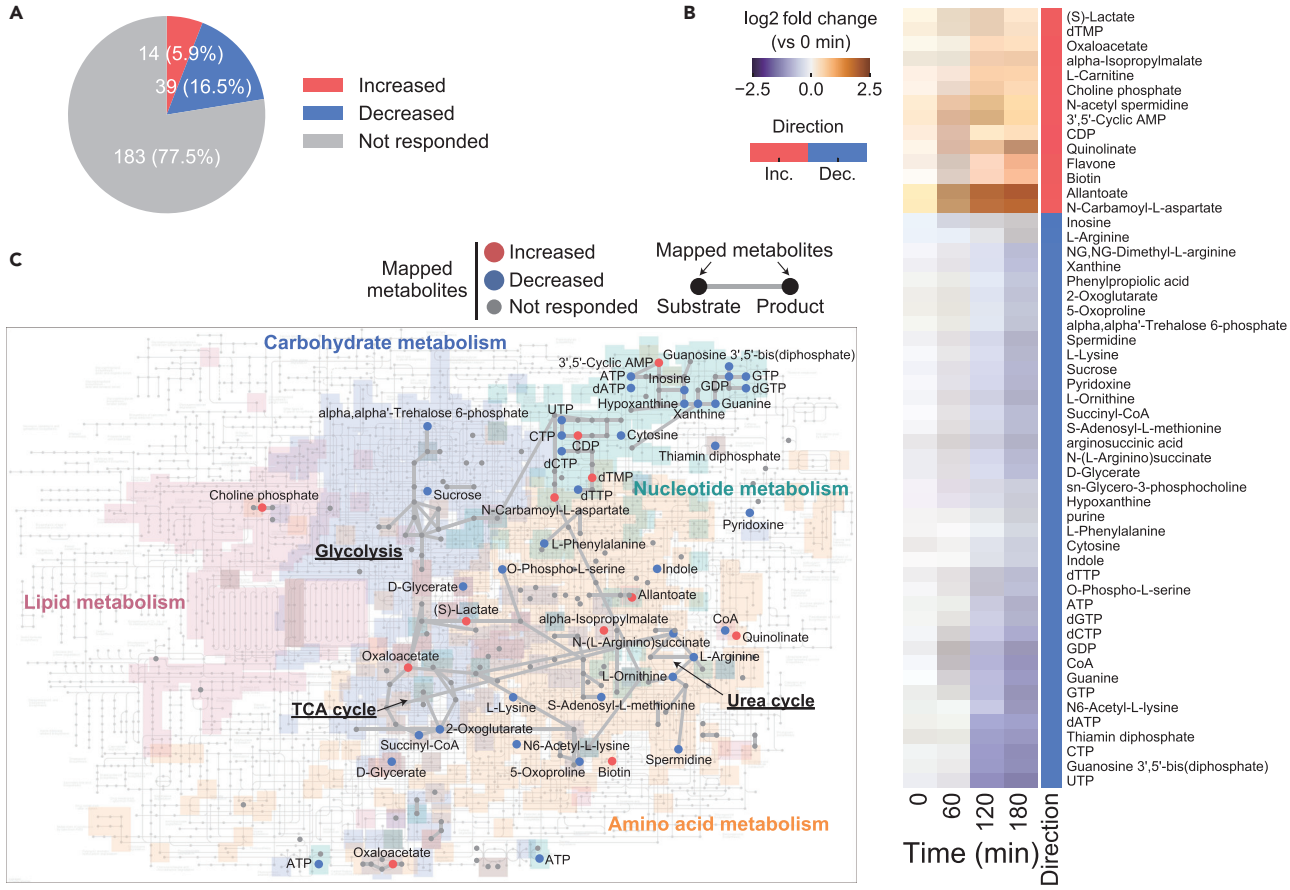


Figure 4. Identification of IRMs and connection of the IRMs/IRGs and metabolic reactions

(A) Number of metabolites measured in the metabolomic data. Metabolites that showed an absolute $\log_2$ fold change larger than 0.5 and an FDR-adjusted *p* value (*q* value) of less than 0.1 at any time point were defined as IRMs. The *q* values were calculated by Storey's procedure (Storey and Tibshirani, 2003). We divided the IRMs into increased and decreased IRMs as shown in Figure 1E. See also Table S4.

(B) Heatmap of the $\log_2$ scaled fold change (relative to time 0) normalized time courses of the IRMs.

(C) Increasing, decreasing, and non-responsive metabolites projected onto the KEGG metabolic pathways. Substrate-product relationships between the mapped metabolites (increased, decreased, and not responsive metabolites) were extracted from a KEGG Markup Language file (dme01100) downloaded from KEGG.

Using enzymatic reactions in the KEGG database, we connected IRGs, metabolic reactions, and IRMs to examine the effect of transcriptional regulation on metabolism (Table S4). We identified 120 IRGs encoding metabolic enzymes. These contain the increased IRG *rudimentary* (*r*, CAD ortholog of *Drosophila*), which is a rate-limiting metabolic enzyme of *de novo* pyrimidine synthesis, consistent with the increase of N-carbamoyl-L-aspartate, which is the product of the reaction catalyzed by *rudimentary*. The phosphorylation level of the rudimentary metabolic enzyme was also increased by insulin stimulation (Tables S2 and S4), consistent with previous studies on mammals (Ben-Sahra et al., 2013; Robitaille et al., 2013).

Using allosteric regulations extracted from the BRENDA database, we connected IRMs to metabolic reactions (Table S4). In total, we identified 1,543 allosteric regulations from IRMs to metabolic reactions. Altogether, our analysis identified inter-omic transcriptional and allosteric regulation for metabolic reactions that can affect the responses of the IRMs.

Step VI: construction of the trans-omic network of insulin action

We integrated the networks of Steps I-V and constructed a regulatory trans-omic network for insulin-responsive gene expression and metabolic reactions consisting of five layers, which are the insulin signal, TFs, IRGs, metabolic reactions, and IRMs (Figure 5; Table S5). The connections between the layers are

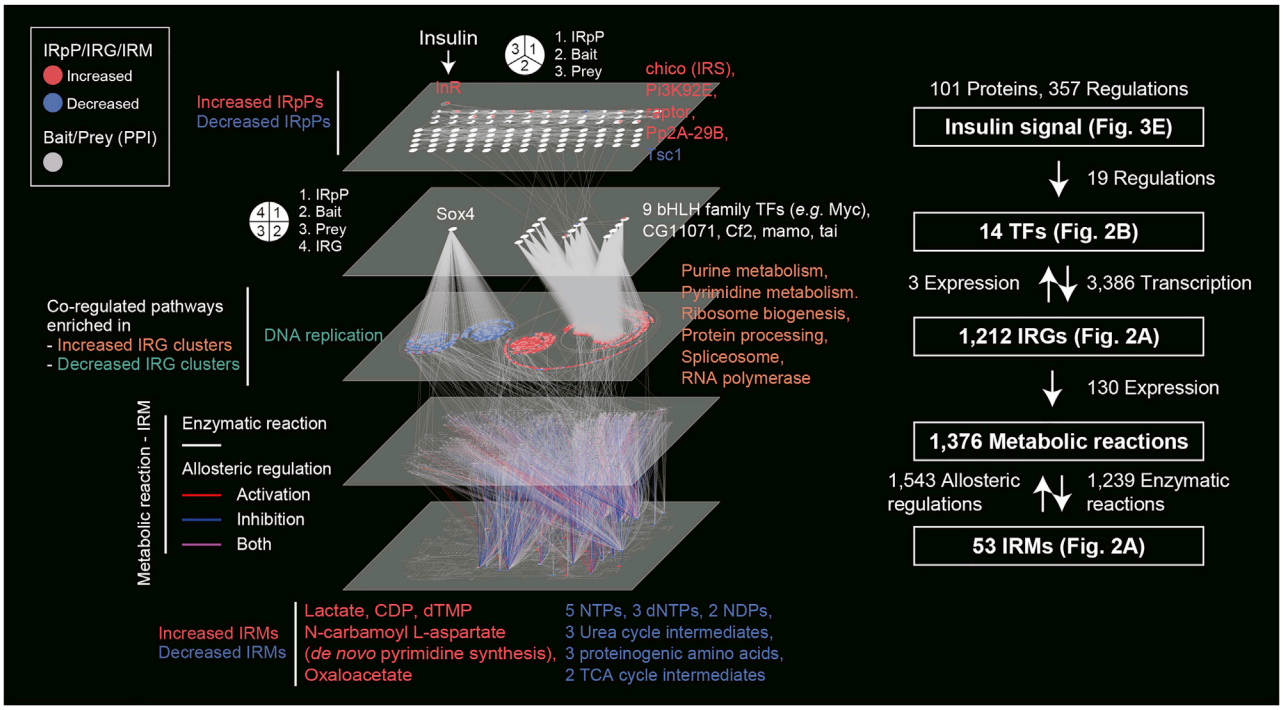


Figure 5. Construction of the trans-omic network of insulin action

The trans-omic network contains five layers and the regulatory relationships among them. The representative insulin-responsive molecules or significantly co-regulated pathways enriched in the IRG clusters in the trans-omic network are shown. The numbers of each type of insulin-responsive node and edge are shown on the right of the network diagram. The network diagram was created using Cytoscape (Shannon et al., 2003). See also Table S5.

representing regulatory events. The insulin signal layer contains the signaling molecules on the predicted signaling pathways from the insulin receptor to the TFs (Figure 3E). The TF layer contains the 14 predicted TFs regulating the IRGs (Figure 2B). The IRG layer contains the 1,212 IRGs (Figure 2A). The metabolic reaction layer contains 1,376 metabolic reactions regulated by the IRMs and/or the IRGs. The IRM layer contains 53 IRMs (Figures 4A and 4B). We then determined intra- and inter-layer regulatory connections between insulin-responsive molecules. The intra-layer regulatory connections of the insulin signal layer and the inter-layer regulatory connections from the insulin signal layer to the TF layer were the predicted signaling pathways from the insulin receptor to the candidate TFs (Figure 3E). The inter-layer regulatory connections from the TF layer to the IRG layer were determined by the predicted regulatory connections between TFs and IRGs (Figure 2B). The inter-layer regulatory connections from the IRG to the metabolic reaction layer were determined by matching metabolic reactions to the corresponding metabolic enzymes encoded by IRGs according to KEGG annotation. The inter-layer regulatory connections between the metabolic reaction layer and the IRM layer consisted of two types: (1) regulatory connections mediated by allosteric regulators assigned according to BRENDA and (2) regulatory connections mediated by the substrate or product of the reaction according to KEGG.

In our network, insulin activates the insulin receptor in the insulin signal layer and regulates 10 TFs in the TF layer, which contains Myc that is connected through the signaling pathway components layer that includes InR, chico, Pi3K, S6K, Akt, and Erk. In the IRG layer, the increased and decreased clusters of the IRGs are enriched for the distinct KEGG pathways and the functionally similar IRGs are significantly co-regulated (Figures 2B–2D). The multiple anabolic pathways related to nucleotide synthesis, transcription, and translation are enriched for the increased clusters of the IRGs and the genes in each of those pathways are co-regulated. On the other hand, genes related to DNA replication are enriched for the decreased clusters and co-regulated in a time-dependent manner. Insulin regulates the metabolic reactions in the metabolic reaction layer through the regulation of the expression levels of the metabolic enzyme genes in the IRG layer and changes the concentration of the metabolites in the IRM layer. For example, insulin increases the expression level of the *rudimentary* metabolic enzyme gene, which is a rate-limiting metabolic enzyme

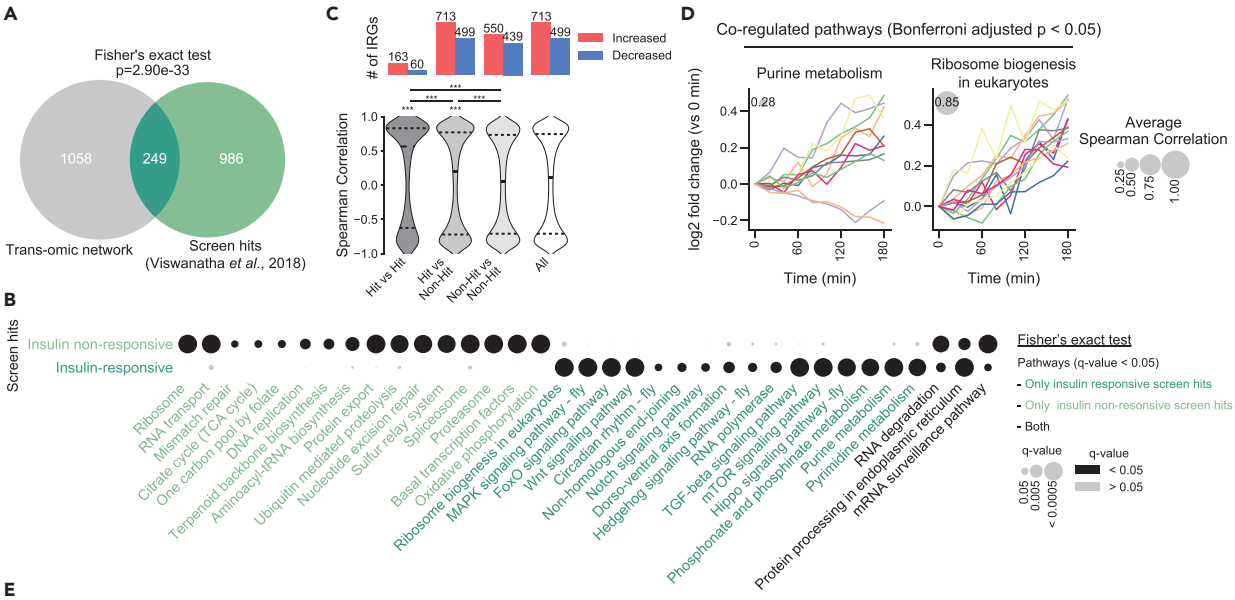
of *de novo* pyrimidine synthesis, and the increase of *rudimentary* is consistent with the increase in the concentration of N-carbamoyl-aspartate, which is catalyzed by the rudimentary metabolic enzyme (Figure 4C). In the IRM layer, the five NTPs and the three proteinogenic amino acids (L-Arginine, L-Lysine, and L-Phenylalanine) are decreased, which is consistent with increases in the expression of genes related to transcription (e.g., RNA polymerase and spliceosome) and translation (e.g., ribosome biogenesis). It is also possible that the IRMs regulate various metabolic processes through allosteric regulations. Altogether, the construction of the trans-omic network of insulin action revealed the functional and temporal characteristics of insulin-responsive molecules and the regulatory relationships among them.

Step VII: integration of the trans-omic network of insulin action with a CRISPR screen for cell proliferation

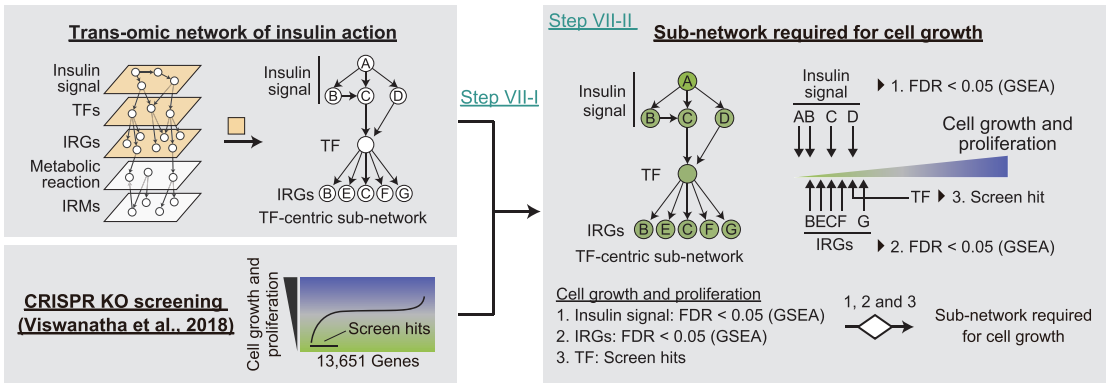
We previously performed a genome-wide pooled CRISPR knockout screening for cell proliferation in *Drosophila* S2R + cells and identified 1,235 genes essential for cell proliferation (hereafter denoted as screen hits) (Table S6) (Viswanatha et al., 2018). Thus, as cell growth is required for cell proliferation and insulin can stimulate proliferation in a context-dependent manner, we hypothesized that part of the trans-omic network of insulin action includes the essential genes identified in the CRISPR screen for cell proliferation. Next, we compared the screen hits and the molecules in the trans-omics network, specifically the molecules in the insulin signal, TFs, and IRGs layers. As expected, the screen hits and the molecules in the trans-omic network were overlapping significantly (249 genes; p value = 2.90e-33, Fisher's exact test; Figure 6A).

To identify cellular functions involved in cell growth and proliferation that are regulated in an insulin-dependent manner in S2R + cells, we analyzed the KEGG pathways enriched for the screen hits overlapping with the trans-omic network (hereafter denoted as insulin-responsive screen hits) as well as the remaining screen hits that are not overlapping with the trans-omic network (hereafter denoted as insulin non-responsive screen hits) (Table S6). In sum, 35 pathways were enriched in either the insulin-responsive screen hits or the insulin non-responsive screen hits (q value <0.05, Fisher exact test; Figure 6B). Only three pathways (mRNA surveillance pathway, Protein processing in endoplasmic reticulum, and RNA degradation) were enriched in both insulin-responsive and non-responsive screen hits (shown in black in Figure 6B). In sum, 16 pathways were specifically enriched for the insulin-responsive screen hits (shown in dark green in Figure 6B), among which are the signaling pathways and multiple anabolic pathways such as Purine metabolism, Pyrimidine metabolism, RNA polymerase, and Ribosome biogenesis in eukaryotes. All of those anabolic pathways were enriched in the increased IRG clusters (Figure 2B). This is consistent with previous studies that those anabolic pathways are activated by insulin/IGF signals to support cell growth and proliferation (Valvezan and Manning, 2019; Wullschleger et al., 2006; Zhu and Thompson, 2019). On the other hand, 16 pathways were specifically enriched for the insulin non-responsive screen hits (shown in light green in Figure 6B). Some of these pathways are involved in glucose and energy metabolism (TCA cycle, and Oxidative phosphorylation), protein catabolism (Proteasome and Ubiquitin mediated proteolysis), DNA metabolism, RNA metabolism, and protein synthesis. In contrast to the pathways specifically enriched in the insulin-responsive screen hits, the pathways specifically enriched in insulin non-responsive screen hits are relevant to catabolic processes and DNA replication, none of which are signaling pathways. Some of the anabolic pathways that were over-represented only for insulin non-responsive screen hits were also enriched for the preys of the PPI data that were not used for the network construction (i.e., the preys that are not contained in signaling pathways of the KEGG database) (Figure S7; Table S6). Those pathways contain Ribosome and Spliceosome, suggesting that these pathways are regulated at either post-transcriptional or post-translational levels. The association of ribosomal proteins and mTORC2, which is a component of the signaling pathways regulated by growth factors such as insulin, has previously been reported (Zinzalla et al., 2011). The result of pathway enrichment analysis indicates that the insulin signal regulates a large part of the genes involved in signaling pathways and anabolic pathways that are needed for cell growth and proliferation in S2R + cells. In addition, the regulations by insulin signal happen at both the transcriptional level and the post-transcriptional, or post-translational level in S2R + cells.

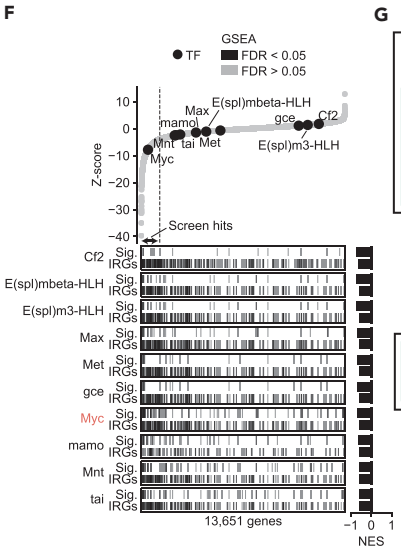
In this transcriptomic analysis, we found that the expression levels for the gene pairs of the IRGs in the same KEGG pathways show significantly higher Spearman correlation coefficients, i.e., the functionally similar IRGs tend to be co-regulated by insulin signaling (Figure 2D). We further investigated whether this trend was also found in the gene pairs from the screen hits. We hypothesize that the pairs between the screen hits should show significantly higher Spearman correlation coefficients if co-regulation of the functionally


A
Fisher's exact test
p=2.90e-33
1058
249
986
Trans-omic network
Screen hits
(Viswanatha et al., 2018)
C
# of IRGs
163
60
713
499
550
439
713
499
Increased
Decreased
Spearman Correlation
Hit vs Hit
Hit vs Non-Hit
Non-Hit vs Non-Hit
All
D
Co-regulated pathways (Bonferroni adjusted p < 0.05)
Purine metabolism
0.28
Ribosome biogenesis in eukaryotes
0.85
log2 fold change (vs 0 min)
Time (min)
Average Spearman Correlation
B
Screen hits
Insulin non-responsive
Insulin-responsive
Ribosome
RNA transport
Mismatch repair
Citrate cycle (TCA cycle)
One carbon pool by folate
DNA replication
Terpenoid backbone biosynthesis
Aminoacyl-tRNA biosynthesis
Protein export
Ubiquitin mediated proteolysis
Nucleotide excision repair
Sulfur relay system
Spliceosome
Proteasome
Basal transcription factors
Oxidative phosphorylation
Ribosome biogenesis in eukaryotes
MAPK signaling pathway - fly
FoxO signaling pathway
Wnt signaling pathway
Circadian rhythm - fly
Non-homologous end-joining
Notch signaling pathway
Dorso-ventral axis formation
Hedgehog signaling pathway - fly
RNA polymerase
TGF-beta signaling pathway
mTOR signaling pathway
Hippo signaling pathway - fly
Phosphonate and phosphinate metabolism
Purine metabolism
Pyrimidine metabolism
RNA degradation
Protein processing in endoplasmic reticulum
mRNA surveillance pathway
Fisher's exact test
Pathways (q-value < 0.05)
- Only insulin responsive screen hits
- Only insulin non-resonsive screen hits
- Both
q-value
q-value
< 0.05
> 0.05

E


Trans-omic network of insulin action
Insulin signal
TFs
IRGs
Metabolic reaction
IRMs
TF
TF-centric sub-network
CRISPR KO screening (Viswanatha et al., 2018)
Cell growth and proliferation
Screen hits
13,651 Genes
Step VII-I
Step VII-II
Sub-network required for cell growth
1. FDR < 0.05 (GSEA)
Cell growth and proliferation
TF
3. Screen hit
2. FDR < 0.05 (GSEA)
Cell growth and proliferation
1. Insulin signal: FDR < 0.05 (GSEA)
2. IRGs: FDR < 0.05 (GSEA)
3. TF: Screen hits
1, 2 and 3
Sub-network required for cell growth

F


GSEA
TF
FDR < 0.05
FDR > 0.05
Z-score
E(spl)mbeta-HLH
Max
mamo
gce
Cf2
Mnt
tai
Met
E(spl)m3-HLH
Myc
Screen hits
Sig. IRGs
13,651 genes
NES

G

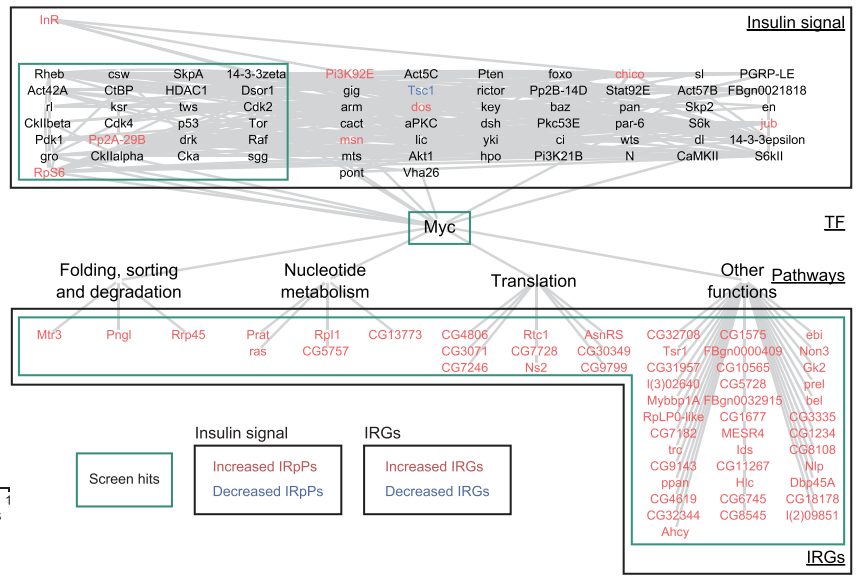

InR
Insulin signal
Rheb Act42A rl CkIIbeta Pdk1 gro RpS6
csw CtBP ksr Cdk4 Pp2A-29B CkIIalpha
SkpA HDAC1 tws p53 drk Cka
14-3-3zeta Dsor1 Cdk2 Tor Raf sgg
Pi3K92E gig arm cact aPKC msn lic Akt1 mts pont Vha26
Act5C Tsc1 dos
Pten rictor key dsh yki hpo
foxo Pp2B-14D baz Pkc53E ci Pi3K21B
chico Stat92E pan par-6 wts N
sl Act57B Skp2 S6k dl CaMKII
PGRP-LE FBgn0021818 en jub 14-3-3epsilon S6kII
Myc
TF
Folding, sorting and degradation
Nucleotide metabolism
Translation
Other functions
Pathways
Mtr3 Pngl Rrp45 Prat ras Rpl1 CG5757 CG13773 CG4806 CG3071 CG7246 Rtc1 CG7728 Ns2 AsnRS CG30349 CG9799 CG32708 Tsr1 CG31957 l(3)02640 Mybbp1A RpLP0-like CG7182 trc CG9143 ppan CG4619 CG32344 Ahcy CG1575 FBgn0000409 CG10565 CG5728 FBgn0032915 CG1677 MESR4 Ids CG11267 Hlc CG6745 CG8545 ebi Non3 Gk2 prel bel CG3335 CG1234 CG8108 Nlp Dbp45A CG18178 I(2)09851
IRGs
Screen hits
Insulin signal
Increased IRpPs
Decreased IRpPs
IRGs
Increased IRGs
Decreased IRGs

Figure 6. Integration of the trans-omic network of insulin action with a CRISPR screen for cell proliferation

(A) Numbers of molecules in the trans-omic network and screen hits. The *p* value of Fisher's exact test is shown above the Venn diagram. See also Table S6.

(B) KEGG pathway enrichment analysis for the insulin-responsive and insulin non-responsive screen hits. See also Table S6.

(C) Distribution of Spearman correlation coefficients of the pairs within hit-IRGs, between hit-IRGs and non-hit IRGs, and within non-hit IRGs. To test co-regulation of IRGs within the above categories, we examined whether the average Spearman correlation coefficients of the IRG pairs in each of the categories were significantly higher than randomly sampled IRG pairs. We also compared the distribution of Spearman correlation between the categories by the Wilcoxon rank-sum test. $*p < 0.05$, $**p < 0.01$, $***p < 0.001$. See also STAR Methods and Table S6.

(D) Time course of the hit-IRGs in significantly co-regulated KEGG pathways. Data are shown as the $\log_2$ fold change values compared to time 0 that are calculated from the mean of biological replicates. Different line colors indicate different hit-IRGs. We tested the significance of the average Spearman correlation in each group as described in the STAR Methods. We analyzed KEGG pathways containing 10 or more hit-IRGs. See also Table S6.

(E) We predicted subnetworks required for cell growth in an insulin-dependent manner in the following two steps (Step VII-I and VII-II). In Step VII-I, we extracted TF-centric subnetworks each containing a TF, its upstream signaling pathways, and its target IRGs from the trans-omic network of insulin action. We mapped the *Z*-score obtained from the CRISPR KO screen to the extracted subnetworks. In Step VII-II, we selected subnetworks required for cell growth by the following three criteria: (i) the results of Gene Set Enrichment Analysis (GSEA) (Subramanian et al., 2005) is significant for upstream signaling molecules of a TF (FDR <0.05); (ii) the result of GSEA is significant for the target IRG for the TF (FDR <0.05); and (iii) the TF itself is a screen hit.

(F) The scatterplot shows the *Z*-score of the CRISPR screen for each of the screened genes. A smaller *Z*-score indicates a lower cell proliferation activity. Genes are sorted by the *Z*-score. The barcode plots below the scatterplot show the distribution of the upstream signaling molecules ("Sig.") of the TFs (Figure 3E), and the target IRGs ("IRGs") regulated by the TFs (Figure 2B). Bar graphs indicate the normalized enrichment score (NES) of GSEA and the color of a bar indicates the significance of the NES (FDR< 0.05). Labels of TFs included in subnetworks involved in cell growth defined in Figure 6E are shown in red.

(G) The predicted subnetwork required for cell growth consists of a TF Myc, its upstream signaling molecules (insulin signal; Figure 3E), and its target hit-IRGs (IRGs; Figures 2B and 6A). The TF layer contains the Myc TF that is required for cell growth as shown in red in Figure 6F. The pathway layer contains KEGG pathways that contain three or more members that overlap with the hit-IRGs regulated by Myc. Pathways containing two or fewer members that overlap with the hit-IRGs regulated by Myc were categorized as "Other functions." See also Table S6.

similar IRGs is important in the context of cell growth and proliferation. We first divided the IRGs into two groups as (1) "hit-IRGs," which contain IRGs overlapping with the screen hits; (2) "non-hit IRGs," which are the IRGs not overlapping with the screen hits. Then we calculated Spearman correlation coefficients for the gene pairs within each group, and the gene pairs between hit IRGs and non-hit IRGs groups as well as randomly sampled pairs of all IRGs. The average and median Spearman correlation coefficients of the pairs within the hit-IRGs (0.19 and 0.56, respectively) were the highest among all the three groupings shown in Figure 6C and the average Spearman correlation coefficients of pairs within hit-IRGs were significantly higher than randomly sampled IRG pairs (Bonferroni-adjusted *p* value <0.05; Table S6; STAR Methods). In addition, the distributions of the correlation coefficients were significantly different between the groups (Bonferroni-adjusted *p* value <0.05, Wilcoxon rank-sum test; Figure 6C; Table S6). These results indicate that the gene pairs within hit-IRGs are more likely to be co-regulated than other pairs, suggesting that the co-regulation of IRGs is important in the context of cell growth and proliferation. We examined which genes of the KEGG pathways were co-regulated among hit-IRGs. Only Purine metabolism and Ribosome biogenesis in eukaryotes contained 10 or more hit-IRGs. Both of these pathways significantly show high average Spearman correlation coefficients calculated within the hit-IRGs in each of the pathways (Bonferroni-adjusted *p* value <0.05; Figure 6D; Table S6; STAR Methods).

Molecules in the trans-omics network and the screen hits significantly overlap (Figure 6A), suggesting that at least a part of the trans-omics network is required for cell growth in the context of cell proliferation. Therefore, we aimed to identify subnetworks that are required for cell growth in an insulin-dependent manner. We hypothesized that TFs regulating cell growth in an insulin-dependent manner are likely to be regulated by genes that negatively/positively affect cell growth and likely regulate them (Figure 6E). We predicted subnetworks that are potentially required for cell growth as shown in Figure 6E (STAR Methods). We identified the subnetwork containing Myc, the TF, and its upstream signaling pathways, as well as its target IRGs as a potential subnetwork required for cell growth in an insulin-dependent manner (Figures 6F and 6G; Table S6). Myc has been reported to regulate cell growth in both *Drosophila* and mammals (Dang, 2013; Johnston et al., 1999), and in addition, Myc has also been reported regulated by insulin/IGF signaling in *Drosophila* (Demontis and Perrimon, 2009; Parisi et al., 2011; Teleman et al., 2008). In the predicted subnetwork surrounding Myc, the upstream signaling molecules contain Rheb and Tor, which are components of mTOR signaling regulating cell growth (Oldham and Hafen, 2003; Wullschleger et al., 2006), rl (*Drosophila* ortholog of Erk) and pdk1, which are well-known components of the insulin signaling pathway (Teleman, 2009), and Sgg/Gsk3-beta, which has been reported as a potential Myc regulator (Parisi et al., 2011). Pathways regulated by Myc contained anabolic pathways such as protein folding, nucleotide metabolism, and translation. These pathways also contain Prat, a rate-limiting enzyme of the purine metabolism pathway, and genes involved in ribosome biogenesis that were increased and highly co-regulated by

insulin at the gene expression levels (Figures 2B, 2D, and 6D). Thus, integration of the trans-omic network and the CRISPR screening data suggests that insulin is connected to cell growth through a subnetwork containing Myc, its upstream signaling pathway, and its target IRGs involved in anabolic processes.

DISCUSSION

In this study, we constructed a trans-omic network of insulin action in *Drosophila* S2R + cells by integrating time course PPI, phosphoproteomic, transcriptomic, and metabolomic data following insulin stimulation with bioinformatics resources. In the trans-omic network, 14 TFs, including Myc, and their upstream signaling pathways coordinately upregulated multiple anabolic processes such as nucleotide synthesis, transcription, and translation in a time-dependent manner. These responses may contribute to the decrease in metabolites such as NTPs and proteinogenic amino acids that are required for transcription and translation. We further analyzed how insulin signaling regulates cell growth in the context of cell proliferation by integrating the trans-omic network and the results from a previous CRISPR knockout screen for cell proliferation. Among the subnetworks surrounding various TFs in the trans-omic network, we identified a subnetwork including Myc, its upstream signaling pathways, and its downstream target genes as being involved in anabolic processes regulating cell growth in an insulin-dependent manner.

We identified a coordinated upregulation of IRGs in multiple anabolic processes such as nucleotide synthesis, transcription, and translation in the time series. The IRGs related to these processes were also enriched in the screen hits of the CRISPR screen for cell proliferation. To precisely regulate cell growth, cells need to coordinately promote macromolecular syntheses of the substrates (*e.g.* nucleotides and amino acids), as well as the components of the transcriptional and translational machinery. Imbalances in these processes can have detrimental effects. In our previous study (Zirin et al., 2019), we showed that inhibition of aminoacyl-tRNA synthases selectively kills Myc-overexpressed HMEC cells, which has been known to increase the activity of ribosome biogenesis. In addition, it has been reported that PI3K activation, which is known to activate protein synthesis, together with knockdown of either protein synthesis- or protein catabolism-related genes, causes synthetic lethality. However, simultaneous inhibition of both pathways causes no lethality (Davoli et al., 2016). In another study, MYC-overexpressing human cells have been reported to show higher sensitivity to splicing inhibition than normal cells (Hsu et al., 2015; Kessler et al., 2012). The increase in the IRGs in the multiple anabolic pathways that are essential for cell growth and proliferation and the co-regulation of the IRGs within each of the anabolic pathways are reasonable regulatory strategies to balance the activities between anabolic pathways and maintain stoichiometry of the genes within each anabolic pathway. The significant co-regulation of the hit-IRGs further supports this interpretation. Although this study alone does not provide a causal relationship between the co-regulation of the IRGs related to macromolecule synthesis and the capacity of cell growth and proliferation, our results may reflect a design principle of the regulation of genes related to macromolecule synthesis under the constraints described above.

The trans-omic network and the CRISPR knockout screening data are complementary. The trans-omic network provides the mechanistic regulatory relationships between insulin-responsive molecules, however, it does not contain the functional information on how each node in the network affects phenotypes, such as cell growth and proliferation. On the other hand, the CRISPR screen data provide quantitative information about the phenotypic effects of gene knockout on cell proliferation, but it does not directly provide the underlying molecular mechanisms. Therefore, integrative analysis of the trans-omic network and the CRISPR screen data can provide insights into molecular mechanisms. This approach is useful for narrowing down for example important subnetworks for the regulation of cell growth. Importantly, loss-of-function phenotypes are less likely to be buffered by functionally redundant paralogs in *Drosophila* than in mammals (Ewen-Campen et al., 2017; Viswanatha et al., 2018). In this study, the subnetwork surrounding Myc, which contains potential downstream effectors involved in anabolic pathways such as protein processing/synthesis and nucleotide metabolism, was predicted as key for the regulation of cell growth in the context of cell proliferation. Myc is an evolutionarily conserved TF that regulates cell growth and proliferation from *Drosophila* to mammals and is regulated by insulin/mTOR signaling (Bellosta and Gallant, 2010). Our data-driven approach recapitulated the role of Myc as an important regulator of cell growth in an insulin-dependent manner, supporting the versatility of our approach for identifying evolutionarily conserved networks regulating phenotypes. In addition, our network provides a resource for the identification of novel regulatory relationships. Finally, the pipeline constructed in this study can be applied to other stimulatory signals and phenotypes.

In our previous studies, we have analyzed the insulin action on the metabolism in mammalian cells (Kawata et al., 2018; Krycer et al., 2017; Noguchi et al., 2013; Ohno et al., 2020; Yugi et al., 2014). In these studies, glycolytic intermediates such as glucose-1-phosphate, glucose-6-phosphate, fructose-6-phosphate (F6P), fructose-1,6-bisphosphate (F1,6BP), dihydroxyacetone phosphate (DHAP), 3-phosphoglycerate (3PG), and phosphoenolpyruvate (PEP), were reported to either increase or decrease following insulin stimulation. On the other hand, although all the above glycolytic intermediates were also measured in our metabolomic data, none of them were found to respond to insulin stimulation. In *Drosophila*, insulin, as is the case in mammals (Saltiel and Kahn, 2001), has been shown to regulate glucose metabolism (Teleman, 2009). Furthermore, we found that in our metabolomic analysis, that lactate, a product of glycolysis, increased after insulin stimulation, suggesting that the glycolytic flux may be increased by insulin stimulation in S2R + cells. Thus, one possible reason for the small number of IRMs belonging to the glycolytic pathway in our study may be owing to the time point measurements of our metabolome data. We have previously shown that, in rat FAO cells, some glycolytic intermediates (*e.g.* F1,6BP, DHAP, 2PG, 3PG, and PEP) transiently increased following insulin stimulation and returned to basal levels 60 min afterward (Yugi et al., 2014). Here, we measured the metabolome at 60, 120, and 180 min after insulin stimulation, which may not have captured the transient response occurring within 60 min. Further, the glycolytic flux can increase without changes in metabolite concentration. Metabolic flux analysis using ^{13}C-labeled glucose would address this issue in the future. To elucidate the rapid responses of metabolites and regulatory mechanisms of glucose metabolism in *Drosophila*, additional data in a shorter timescale than 60 min will be required.

In our transcriptomic analysis, the IRGs in the pathways related to nucleotide synthesis, transcription, and translation were upregulated by insulin stimulation. In addition, the gene expression and phosphorylation level of *rudimentary*, a rate-limiting metabolic enzyme of *de novo* pyrimidine metabolism, and the abundance of N-carbamoyl-L-aspartate, which is the product of the reaction catalyzed by *rudimentary*, increased after insulin stimulation. Consistent with our findings, in mammals, syntheses of macromolecules such as nucleic acids, lipids, and proteins are promoted by mTORC1 depending on growth factors (*e.g.* insulin and IGF) and nutrient and energy status (*e.g.* amino acids, glucose, and ATP) (Valvezan and Manning, 2019; Wullschleger et al., 2006; Zhu and Thompson, 2019). These results suggest that macromolecule syntheses increased after insulin stimulation in S2R + cells. NTPs and proteinogenic amino acids are consumed by transcription and protein synthesis, and the intermediate metabolites of the TCA cycle are used as substrates for nucleotide and amino acid synthesis. The decrease in the NTPs, the amino acids, and the TCA cycle intermediates observed in our metabolomic analysis may reflect the increased activity of transcription and protein synthesis through the increased gene expression of these pathways. Furthermore, in *Drosophila* KC cells, insulin stimulation has been reported to activate the pentose phosphate pathway that is an important source of nucleotide synthesis (Ceddia et al., 2003). Altogether, these data suggest that the trans-omic network constructed in this study mainly captures the anabolic effects of insulin signaling that are important for cell and organismal growth from *Drosophila* to mammals (Oldham and Hafen, 2003).

In this study, we extended the framework of the trans-omic analyses of insulin action developed in our previous studies mainly on two points (Kawata et al., 2018; Yugi et al., 2014). First, we previously constructed trans-omic networks involving kinase-substrate interactions and networks in the KEGG database without PPIs (Kawata et al., 2018; Yugi et al., 2014). Here, we newly integrated into the trans-omic network, the PPI network built using AP-MS following insulin stimulation. Therefore, we expect that our *Drosophila* trans-omic study captures context-specific interactions that we were not able to address previously. Second, we explored the mechanisms regulating cell growth in an insulin-dependent manner by integrating the trans-omic network and the CRISPR cell proliferation screen, providing new insights into the functional aspects of the trans-omic network. Finally, as insulin signaling is evolutionarily conserved, we anticipate that many aspects of the *Drosophila* insulin network presented here will be relevant to the corresponding mammalian network.

Limitations of the study

In this study, our integrative analysis of the trans-omic network of insulin action and the CRISPR screening data for cell proliferation successfully identified Myc as a key regulator of cell growth in an insulin-dependent manner, and identified the potential novel upstream regulators and target genes of Myc. However, further experiments are needed to validate the causal relationships of the changes in the activity of Myc

and these molecules following insulin stimulation. Furthermore, our study does not directly capture some important aspects of insulin action such as changes in metabolic flux, post-transcriptional regulations, and epigenetic regulations. Future studies will be required to expand the trans-omic network to include other data types. Regardless of these limitations, we provide a framework for integrating a trans-omic network and CRISPR screen data.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - Targeted mass spectrometry and data analyses
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Step I: identification of IRGs
  - Step II: prediction of TFs that regulate IRGs
  - Step III: prediction of upstream signaling pathways regulating the predicted TFs
  - Step IV: identification of IRMs
  - Step V: connection of the IRMs/IRGs and metabolic reactions
  - Step VII: integration of the trans-omic network of insulin action with a CRISPR screen for cell proliferation

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2022.104231.

ACKNOWLEDGMENTS

We thank our laboratory members for critically reading this manuscript and for their technical assistance with the experiments. This work was supported by the Japan Society for the Promotion of Science (JSPS) KAKENHI (JP17H06300, JP17H06299, JP18H03979, JP21H04759), the Japan Science and Technology Agency (JST) (JPMJCR2123), and The Uehara Memorial Foundation. T.K. receives funding from a Grant-in-Aid for Early-Career Scientists (JP21K16349). K.Y. receives funding from JSPS KAKENHI Grant Number JP18H05431. K.M. receives funding from a Grant-in-Aid for Early-Career Scientists (JP21K15342). S.O. receives funding from a Grant-in-Aid for Early-Career Scientists (JP17K14864 and JP21K14467). A.P. receives funding from NIA R00 AG057792. This work was funded by grants from the National Institutes of Health (R01 AR057352 and P01 CA120964, to N.P., and R01 DK088718 to M.L.B. and N.P.). N.P. is an investigator with the Howard Hughes Medical Institute.

AUTHOR CONTRIBUTIONS

A.T., Y.H., N.P., and S.K. conceived the project. A.T., Y.H., T.K., K.Y., K.M., S.O., Y.P., Y.B., A.A.P., and X.N. analyzed the data. A.A.P and N.P. designed the experiments. A.A.P. and X.N. performed the experiments. A.A.P. and J.M.A. performed the metabolome measurements. Y.H., X.N., and M.L.B. provided technical assistance on statistics for data analyses. M.L.B., N.P., and S.K. supervised the study. A.T., Y.H., T.K., K.Y., A.A.P., J.M.A., N.P., and S.K. wrote the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: January 11, 2022
Revised: March 9, 2022
Accepted: April 6, 2022
Published: May 20, 2022

REFERENCES

Alvarez, M.J., Shen, Y., Giorgi, F.M., Lachmann, A., Ding, B.B., Ye, B.H., and Califano, A. (2016). Functional characterization of somatic mutations in cancer using network-based inference of protein activity. Nat. Genet. *48*, 838–847.

Bailey, T.L., Johnson, J., Grant, C.E., and Noble, W.S. (2015). The MEME suite. Nucleic Acids Res. *43*, W39–W49.

Baker, K.D., and Thummel, C.S. (2007). Diabetic larvae and obese flies-emerging studies of metabolism in Drosophila. Cell Metab. *6*, 257–266.

Bellosta, P., and Gallant, P. (2010). Myc function in Drosophila. Genes Cancer *1*, 542–546.

Ben-Sahra, I., Howell, J.J., Asara, J.M., and Manning, B.D. (2013). Stimulation of de novo pyrimidine synthesis by growth signaling through mTOR and S6K1. Science *339*, 1323–1328.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. *57*, 289–300.

Buehler, E.C., Sachs, J.R., Shao, K., Bagchi, A., and Ungar, L.H. (2004). The CRASSS plug-in for integrating annotation data with hierarchical clustering results. Bioinformatics *20*, 3266–3269.

Buljan, M., Ciuffa, R., van Drogen, A., Vichalkovski, A., Mehnert, M., Rosenberger, G., Lee, S., Varjosalo, M., Pernas, L.E., Spegg, V., et al. (2020). Kinase interaction network expands functional and disease roles of human kinases. Mol. Cell *79*, 504–520.e9.

Ceddia, R.B., Bikopoulos, G.J., Hilliker, A.J., and Sweeney, G. (2003). Insulin stimulates glucose metabolism via the pentose phosphate pathway in Drosophila Kc cells. FEBS Lett. *555*, 307–310.

Chang, A., Jeske, L., Ulbrich, S., Hofmann, J., Koblitz, J., Schomburg, I., Neumann-schaal, M., Jahn, D., and Schomburg, D. (2021). BRENDA , the ELIXIR core data resource in 2021 : new developments and updates. Nucleic Acids Res. *49*, D498–D508.

Chauvin, C., Koka, V., Nouschi, A., Mieulet, V., Hoareau-Aveilla, C., Dreazen, A., Cagnard, N., Carpentier, W., Kiss, T., Meyuhas, O., et al. (2014). Ribosomal protein S6 kinase activity controls the ribosome biogenesis transcriptional program. Oncogene *33*, 474–483.

Dale, R.K., Pedersen, B.S., and Quinlan, A.R. (2011). Pybedtools: a flexible python library for manipulating genomic datasets and annotations. Bioinformatics *27*, 3423–3424.

Dang, C.V. (2013). MYC, metabolism, cell growth, and tumorigenesis. Cold Spring Harb. Perspect. Med. *3*, a014217.

Davoli, T., Mengwasser, K.E., Duan, J., Chen, T., Christensen, C., Wooten, E.C., Anselmo, A.N., Li, M.Z., Wong, K.-K., Kahle, K.T., et al. (2016). Functional genomics reveals that tumors with activating phosphoinositide 3-kinase mutations are dependent on accelerated protein turnover. Genes Dev. *30*, 2684–2695.

Demontis, F., and Perrimon, N. (2009). Integration of Insulin receptor/Foxo signaling and dMyc activity during muscle growth regulates body size in Drosophila. Development *136*, 983–993.

Dugourd, A., and Saez-Rodriguez, J. (2019). Footprint-based functional analysis of multiomic data. Curr. Opin. Syst. Biol. *15*, 82–90.

Dupont, J., Khan, J., Qu, B.-H., Metzler, P., Helman, L., and LeRoith, D. (2001). Insulin and IGF-1 induce different patterns of gene expression in mouse fibroblast NIH-3T3 Cells: identification by cDNA microarray Analysis. Endocrinology *142*, 4969–4975.

Everman, S., Meyer, C., Tran, L., Hoffman, N., Carroll, C.C., Dedmon, W.L., and Katsanos, C.S. (2016). Insulin does not stimulate muscle protein synthesis during increased plasma branched-chain amino acids alone but still decreases whole body proteolysis in humans. Am. J. Physiol. Metab. *311*, E671–E677.

Ewen-Campen, B., Mohr, S.E., Hu, Y., and Perrimon, N. (2017). Accessing the phenotype gap: enabling systematic investigation of paralog functional complexity with CRISPR. Dev. Cell *43*, 6–9.

Freschi, L., Osseni, M., and Landry, C.R. (2014). Functional divergence and evolutionary turnover in mammalian phosphoproteomes. PLoS Genet. *10*, e1004062.

Friedman, A.A., Tucker, G., Singh, R., Yan, D., Vinayagam, A., Hu, Y., Binari, R., Hong, P., Sun, X., Porto, M., et al. (2011). Proteomic and functional genomic landscape of receptor tyrosine kinase and ras to extracellular signal–regulated kinase signaling. Sci. Signal. *4*, rs10.

Glatter, T., Schittenhelm, R.B., Rinner, O., Roguska, K., Wepf, A., Jünger, M.A., Köhler, K., Jevtov, I., Choi, H., Schmidt, A., et al. (2011). Modularity and hormone sensitivity of the Drosophila melanogaster insulin receptor/target of rapamycin interaction proteome. Mol. Syst. Biol. *7*, 547.

Grant, C.E., Bailey, T.L., and Noble, W.S. (2011). FIMO: scanning for occurrences of a given motif. Bioinformatics *27*, 1017–1018.

Grewal, S.S., Li, L., Orian, A., Eisenman, R.N., and Edgar, B.A. (2005). Myc-dependent regulation of ribosomal RNA synthesis during Drosophila development. Nat. Cell Biol. *7*, 295–302.

Guertin, D.A., Guntur, K.V.P., Bell, G.W., Thoreen, C.C., and Sabatini, D.M. (2006). Functional genomics identifies tor-regulated genes that control growth and division. Curr. Biol. *16*, 958–970.

Hansson, J., Rafiee, M.R., Reiland, S., Polo, J.M., Gehring, J., Okawa, S., Huber, W., Hochedlinger, K., and Krijgsveld, J. (2012). Highly coordinated proteome dynamics during reprogramming of somatic cells to pluripotency. Cell Rep. *2*, 1579–1592.

Hectors, T.L.M., Vanparys, C., Pereira-Fernandes, A., Knapen, D., and Blust, R. (2012). Mechanistic evaluation of the insulin response in H4IIE hepatoma cells: new endpoints for toxicity testing? Toxicol. Lett. *212*, 180–189.

Horn, H., Schoof, E.M., Kim, J., Robin, X., Miller, M.L., Diella, F., Palma, A., Cesareni, G., Jensen, L.J., and Linding, R. (2014). KinomeXplorer: an integrated platform for kinome biology studies. Nat. Methods. *11*, 603–604.

Hsu, T.Y.-T., Simon, L.M., Neill, N.J., Marcotte, R., Sayad, A., Bland, C.S., Echeverria, G.V., Sun, T., Kurley, S.J., Tyagi, S., et al. (2015). The spliceosome is a therapeutic vulnerability in MYC-driven cancer. Nature *525*, 384–388.

Hu, Y., Flockhart, I., Vinayagam, A., Bergwitz, C., Berger, B., Perrimon, N., and Mohr, S.E. (2011). An integrative approach to ortholog prediction for disease-focused and other functional studies. BMC Bioinformatics *12*, 357.

Hu, Y., Vinayagam, A., Nand, A., Comjean, A., Chung, V., Hao, T., Mohr, S.E., and Perrimon, N. (2018). Molecular interaction search tool (MIST): an integrated resource for mining gene and protein interaction data. Nucleic Acids Res. *46*, D567–D574.

Humphrey, S.J., Yang, G., Yang, P., Fazakerley, D.J., Stöckli, J., Yang, J.Y., and James, D.E. (2013). Dynamic adipocyte phosphoproteome reveals that akt directly regulates mTORC2. Cell Metab. *17*, 1009–1020.

Humphrey, S.J., Azimifar, S.B., and Mann, M. (2015). High-throughput phosphoproteomics reveals in vivo insulin signaling dynamics. Nat. Biotechnol. *33*, 990–995.

Johnston, L.A., Prober, D.A., Edgar, B.A., Eisenman, R.N., and Gallant, P. (1999). Drosophila myc regulates cellular growth during development. Cell *98*, 779–790.

Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y., and Morishima, K. (2017). KEGG: new perspectives on genomes, pathways, diseases and drugs. Nucleic Acids Res. *45*, D353–D361.

Kawata, K., Hatano, A., Yugi, K., Kubota, H., Sano, T., Fujii, M., Tomizawa, Y., Kokaji, T., Tanaka, K.Y., Uda, S., et al. (2018). Trans-omic analysis reveals selective responses to induced and basal insulin across signaling, transcriptional, and metabolic networks. iScience *7*, 212–229.

Kawata, K., Yugi, K., Hatano, A., Kokaji, T., Tomizawa, Y., Fujii, M., Uda, S., Kubota, H., Matsumoto, M., Nakayama, K.I., et al. (2019). Reconstruction of global regulatory network from signaling to cellular functions using phosphoproteomic data. Genes Cells *24*, 82–93.

Kessler, J.D., Kahle, K.T., Sun, T., Meerbrey, K.L., Schlabach, M.R., Schmitt, E.M., Skinner, S.O., Xu, Q., Li, M.Z., Hartman, Z.C., et al. (2012). A SUMOylation-dependent transcriptional subprogram is required for myc-driven tumorigenesis. Science *335*, 348–353.

Kim, H.S., and Lee, N.K. (2014). Gene expression profiling in osteoclast precursors by insulin using microarray analysis. Mol. Cells *37*, 827–832.

Kokaji, T., Hatano, A., Ito, Y., Yugi, K., Eto, M., Morita, K., Ohno, S., Fujii, M., Hironaka, K., Egami, R., et al. (2020). Transomics analysis reveals

allosteric and gene regulation axes for altered hepatic glucose-responsive metabolism in obesity. Sci. Signal. *13*, eaaz1236.

Krüger, M., Kratchmarova, I., Blagoev, B., Tseng, Y.H., Kahn, C.R., and Mann, M. (2008). Dissection of the insulin signaling pathway via quantitative phosphoproteomics. Proc. Natl. Acad. Sci. U S A *105*, 2451–2456.

Krycer, J.R., Yugi, K., Hirayama, A., Fazakerley, D.J., Quek, L.-E., Scalzo, R., Ohno, S., Hodson, M.P., Ikeda, S., Shoji, F., et al. (2017). Dynamic metabolomics reveals that insulin primes the adipocyte for glucose metabolism. Cell Rep. *21*, 3536–3547.

Larkin, A., Marygold, S.J., Antonazzo, G., Attrill, H., dos Santos, G., Garapati, P.V., Goodman, J.L., Sian Gramates, L., Millburn, G., Strelets, V.B., et al. (2021). FlyBase: updates to the Drosophila melanogaster knowledge base. Nucleic Acids Res. *49*, D899–D907.

Li, L., Edgar, B.A., and Grewal, S.S. (2010). Nutritional control of gene expression in Drosophila larvae via TOR, Myc and a novel cis-regulatory element. BMC Cell Biol. *11*, 7.

March, J.C., and Bentley, W.E. (2006). Engineering eukaryotic signal transduction with RNAi: enhancing Drosophila S2 cell growth and recombinant protein synthesis via silencing of TSC1. Biotechnol. Bioeng. *95*, 645–652.

Matsuzaki, F., Uda, S., Yamauchi, Y., Matsumoto, M., Soga, T., Maehara, K., Ohkawa, Y., Nakayama, K.I., Kuroda, S., and Kubota, H. (2021). An extensive and dynamic trans-omic network illustrating prominent regulatory mechanisms in response to insulin in the liver. Cell Rep. *36*, 109569.

Miller, M.L., Jensen, L.J., Diella, F., Jørgensen, C., Tinti, M., Li, L., Hsiung, M., Parker, S.A., Bordeaux, J., Sicheritz-Ponten, T., et al. (2008). Linear motif atlas for phosphorylation-dependent signaling. Sci. Signal. *1*, ra2.

Mina, M., Jurman, G., and Furlanello, C. (2015a). CIDER: a pipeline for detecting waves of coordinated transcriptional regulation in gene expression time-course data. Preprint at BioRxiv. https://doi.org/10.1101/012518.

Mina, M., Magi, S., Jurman, G., Itoh, M., Kawaji, H., Lassmann, T., Arner, E., Forrest, A.R.R., Carninci, P., Hayashizaki, Y., et al. (2015b). Promoter-level expression clustering identifies time development of transcriptional regulatory cascades initiated by ERBB receptors in breast cancer cells. Sci. Rep. *5*, 11999.

Monetti, M., Nagaraj, N., Sharma, K., and Mann, M. (2011). Large-scale phosphosite quantification in tissues by a spike-in SILAC method. Nat. Methods *8*, 655–658.

Neumüller, R.A., Gross, T., Samsonova, A.A., Vinayagam, A., Buckner, M., Founk, K., Hu, Y., Sharifpoor, S., Rosebrock, A.P., Andrews, B., et al. (2013). Conserved regulators of nucleolar size revealed by global phenotypic analyses. Sci. Signal. *6*, ra70.

Noguchi, R., Kubota, H., Yugi, K., Toyoshima, Y., Komori, Y., Soga, T., and Kuroda, S. (2013). The selective control of glycolysis, gluconeogenesis and glycogenesis by temporal insulin patterns. Mol. Syst. Biol. *9*, 664.

Ohno, S., Quek, L.-E., Krycer, J.R., Yugi, K., Hirayama, A., Ikeda, S., Shoji, F., Suzuki, K., Soga, T., James, D.E., et al. (2020). Kinetic trans-omic analysis reveals key regulatory mechanisms for insulin-regulated glucose metabolism in adipocytes. iScience *23*, 101479.

Oki, S., Ohta, T., Shioi, G., Hatanaka, H., Ogasawara, O., Okuda, Y., Kawaji, H., Nakaki, R., Sese, J., and Meno, C. (2018). ChIP-Atlas: a data-mining suite powered by full integration of public ChIP-seq data. EMBO Rep. *19*, e46255.

Oldham, S., and Hafen, E. (2003). Insulin/IGF and target of rapamycin signaling: a TOR de force in growth control. Trends Cell Biol. *13*, 79–85.

Orian, A., van Steensel, B., Delrow, J., Bussemaker, H.J., Li, L., Sawado, T., Williams, E., Loo, L.W.M., Cowley, S.M., Yost, C., et al. (2003). Genomic binding by the Drosophila Myc, Max, Mad/Mnt transcription factor network. Genes Dev. *17*, 1101–1114.

Parisi, F., Riccardo, S., Daniel, M., Saqcena, M., Kundu, N., Pession, A., Grifoni, D., Stocker, H., Tabak, E., and Bellosta, P. (2011). Drosophila insulin and target of rapamycin (TOR) pathways regulate GSK3 beta activity to control Myc stability and determine Myc expression in vivo. BMC Biol. *9*, 65.

Puig, O., Marr, M.T., Ruhf, M.L., and Tjian, R. (2003). Control of cell number by Drosophila FOXO: downstream and feedback regulation of the insulin receptor pathway. Genes Dev. *17*, 2006–2020.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841–842.

R Core Team (2021). R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing).

Robitaille, A.M., Christen, S., Shimobayashi, M., Cornu, M., Fava, L.L., Moes, S., Prescianotto-Baschong, C., Sauer, U., Jenoe, P., and Hall, M.N. (2013). Quantitative phosphoproteomics reveal mTORC1 activates de novo pyrimidine synthesis. Science *339*, 1320–1323.

Rome, S., Clément, K., Rabasa-Lhoret, R., Loizon, E., Poitou, C., Barsh, G.S., Riou, J.-P., Laville, M., and Vidal, H. (2003). Microarray profiling of human skeletal muscle reveals that insulin regulates approximately 800 genes during a hyperinsulinemic clamp. J. Biol. Chem. *278*, 18063–18068.

Saltiel, A.R., and Kahn, C.R. (2001). Insulin signalling and the regulation of glucose and lipid metabolism. Nature *414*, 799–806.

Sano, T., Kawata, K., Ohno, S., Yugi, K., Kakuda, H., Kubota, H., Uda, S., Fujii, M., Kunida, K., Hoshino, D., et al. (2016). Selective control of up-regulated and down-regulated genes by temporal patterns and doses of insulin. Sci. Signal. *9*, ra112.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. *13*, 2498–2504.

So, J., Pasculescu, A., Dai, A.Y., Williton, K., James, A., Nguyen, V., Creixell, P., Schoof, E.M., Sinclair, J., Barrios-Rodiles, M., et al. (2015). Integrative analysis of kinase networks in TRAIL-induced apoptosis provides a source of potential targets for combination therapy. Sci. Signal. *8*, rs3.

Sopko, R., Lin, Y.B., Makhijani, K., Alexander, B., Perrimon, N., and Brückner, K. (2015). A systems-level interrogation identifies regulators of Drosophila blood cell number and survival. PLoS Genet. *11*, e1005056.

Storey, J.D., and Tibshirani, R. (2003). Statistical significance for genomewide studies. Proc. Natl. Acad. Sci. U S A *100*, 9440–9445.

Straus, D.S. (1981). Effects of insulin on cellular growth and proliferation. Life Sci. *29*, 2131–2139.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., 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.

Tan, C.S.H., Bodenmiller, B., Pasculescu, A., Jovanovic, M., Hengartner, M.O., Claus, J., Bader, G.D., Aebersold, R., Pawson, T., and Linding, R. (2009). Comparative analysis reveals conserved protein phosphorylation networks implicated in multiple diseases. Sci. Signal. *2*, ra39.

Teleman, A.A. (2009). Molecular mechanisms of metabolic regulation by insulin in Drosophila. Biochem. J. *425*, 13–26.

Teleman, A.A., Hietakangas, V., Sayadian, A.C., and Cohen, S.M. (2008). Nutritional control of protein biosynthetic capacity by insulin via Myc in Drosophila. Cell Metab. *7*, 21–32.

Valvezan, A.J., and Manning, B.D. (2019). Molecular logic of mTORC1 signalling as a metabolic rheostat. Nat. Metab. *1*, 321–333.

Versteyhe, S., Klaproth, B., Borup, R., Palsgaard, J., Jensen, M., Gray, S.G., and De Meyts, P. (2013). IGF-I, IGF-II, and insulin stimulate different gene expression responses through binding to the IGF-I receptor. Front. Endocrinol. (Lausanne). *4*, 98.

Vinayagam, A., Kulkarni, M.M., Sopko, R., Sun, X., Hu, Y., Nand, A., Villalta, C., Moghimi, A., Yang, X., Mohr, S.E., et al. (2016). An integrative analysis of the InR/PI3K/Akt network identifies the dynamic response to insulin signaling. Cell Rep. *16*, 3062–3074.

Viswanatha, R., Li, Z., Hu, Y., and Perrimon, N. (2018). Pooled genome-wide CRISPR screening for basal and context-specific fitness gene essentiality in Drosophila cells. Elife *7*, e36333.

Weirauch, M.T., Yang, A., Albu, M., Cote, A.G., Montenegro-Montero, A., Drewe, P., Najafabadi, H.S., Lambert, S.A., Mann, I., Cook, K., et al. (2014). Determination and inference of eukaryotic transcription factor sequence specificity. Cell *158*, 1431–1443.

White, M.F. (2003). Insulin signaling in Health and disease. Science *302*, 1710–1711.

Wu, C., MacLeod, I., and Su, A.I. (2013). BioGPS and MyGene.info: organizing online, gene-centric information. Nucleic Acids Res. *41*, D561–D565.

Wu, M.Y., Cully, M., Andersen, D., and Leevers, S.J. (2007). Insulin delays the progression of Drosophila cells through G2/M by activating the dTOR/dRaptor complex. EMBO J. *26*, 371–379.

Wullschleger, S., Loewith, R., and Hall, M.N. (2006). TOR signaling in growth and metabolism. Cell *124*, 471–484.

Xin, J., Mark, A., Afrasiabi, C., Tsueng, G., Juchler, M., Gopal, N., Stupp, G.S., Putman, T.E., Ainscough, B.J., Griffith, O.L., et al. (2016). High-performance web services for querying gene and variant annotation. Genome Biol. *17*, 91.

Yugi, K., and Kuroda, S. (2017). Metabolism-centric trans-omics. Cell Syst. *4*, 19–20.

Yugi, K., Kubota, H., Toyoshima, Y., Noguchi, R., Kawata, K., Komori, Y., Uda, S., Kunida, K., Tomizawa, Y., Funato, Y., et al. (2014). Reconstruction of insulin signal flow from phosphoproteome and metabolome data. Cell Rep. *8*, 1171–1183.

Yugi, K., Kubota, H., Hatano, A., and Kuroda, S. (2016). Trans-omics: how to reconstruct biochemical networks across multiple "Omic" layers. Trends Biotechnol. *34*, 276–290.

Zhang, Y., Zhang, Y., and Yu, Y. (2017). Global phosphoproteomic analysis of insulin/Akt/mTORC1/S6K signaling in rat hepatocytes. J. Proteome Res. *16*, 2825–2835.

Zhu, J., and Thompson, C.B. (2019). Metabolic regulation of cell growth and proliferation. Nat. Rev. Mol. Cell Biol. *20*, 436–450.

Zinzalla, V., Stracka, D., Oppliger, W., and Hall, M.N. (2011). Activation of mTORC2 by association with the ribosome. Cell *144*, 757–768.

Zirin, J., Ni, X., Sack, L.M., Yang-Zhou, D., Hu, Y., Brathwaite, R., Bulyk, M.L., Elledge, S.J., and Perrimon, N. (2019). Interspecies analysis of MYC targets identifies tRNA synthetases as mediators of growth and survival in MYC-overexpressing cells. Proc. Natl. Acad. Sci. U S A *116*, 14614–14619.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, peptides, and recombinant proteins		
Insulin from bovine pancreas	Sigma-Aldrich	Cat#I6634
Deposited data		
Metabolomic data	This paper	Table S3
Transcriptomic data	Zirin et al. (2019)	GEO: GSE129292
Phosphoproteomic data	Vinayagam et al. (2016)	Table S4
Affinity purification MS data	Vinayagam et al. (2016)	Table S1
CRISPR knockout screening data	Viswanatha et al. (2018)	https://doi.org/10.7554/eLife.36333.012
Experimental models: Cell lines		
D. melanogaster: Cell line S2R+	Laboratory of Norbert Perrimon	N/A
Software and algorithms		
Source code	This paper	https://doi.org/10.5281/zenodo.6414309
Python version 3.8.3	Python Software Foundation	https://www.python.org; RRID:SCR_008394
R version 3.6.3	R Core Team (2021)	https://www.r-project.org/
FlyBase version FB2018_04	Larkin et al. (2021)	https://flybase.org/; RRID:SCR_006549
KEGG	Kanehisa et al. (2017)	http://www.kegg.jp/; RRID:SCR_012773
Cis-BP version 1.02	Weirauch et al. (2014)	http://cisbp.ccbr.utoronto.ca; RRID:SCR_017236
MEME suite version 4.12.0	Bailey et al. (2015)	http://meme-suite.org/; RRID:SCR_001783
FIMO version 4.12.0	Grant et al. (2011)	https://meme-suite.org/meme/meme_5.3.2/doc/fimo.html
ChIP-atlas	Oki et al. (2018)	https://chip-atlas.org/; RRID:SCR_015511
BedTools version 2.21.0	Quinlan and Hall (2010)	https://github.com/arq5x/bedtools2; RRID:SCR_006646
PyBedTools version 0.8.1	Dale et al. (2011)	https://daler.github.io/pybedtools/#; RRID:SCR_021018
NetPhorest human version 2.1	Horn et al. (2014); Miller et al. (2008)	https://netphorest.info/
NetWorKIN version 3.0	Linding et al. (2008)	http://networkin.info/; RRID:SCR_007818
MyGene.info version 3	Xin et al. (2016); Wu et al. (2013)	https://mygene.info/; RRID:SCR_018660
DIOPT Ortholog Finder version 8	Hu et al. (2011)	https://www.flyrnai.org/cgi-bin/DRSC_orthologs.pl
MIST version 4.0	Hu et al. (2018)	https://fgrtools.hms.harvard.edu/MIST/
Cytoscape version 3.8.2	Shannon et al. (2003)	https://cytoscape.org/; RRID:SCR_003032
BRENDA version 2021.2	Chang et al. (2021)	http://www.brenda-enzymes.org; RRID:SCR_002997
GSEA version 4.0.3	Subramanian et al. (2005)	http://www.broadinstitute.org/gsea/; RRID:SCR_003199

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Shinya Kuroda (skuroda@bs.s.u-tokyo.ac.jp).

Materials availability

This study did not generate new unique reagents.

Data and code availability

The metabolomic data generated in this study is available in Tables S3. Publicly available data used in the study are listed in the key resources table.

The code used for the analysis in this paper is available online at https://doi.org/10.5281/zenodo.6414309.

Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Drosophila S2R + cells (sex: male) were cultured in Schneider's *Drosophila* medium (21,720–024; Thermo Fisher Scientific) supplemented with 10% fetal bovine serum at 25 °C. For insulin treatment, cells were incubated overnight in serum-free Schneider's *Drosophila* medium.

METHOD DETAILS

Targeted mass spectrometry and data analyses

Drosophila S2R + cells were incubated overnight in serum-free Schneider's *Drosophila* medium (21,720–024; Thermo Fisher Scientific). Cells were then treated with 25 μg/mL insulin from bovine pancreas (I6634; Sigma-Aldrich) for 0, 60, 120, or 180 min 1 × 10^6 cells per sample (5 biological replicates) were collected on ice and rapidly snap frozen in liquid nitrogen. The intracellular metabolites were extracted using 1 mL of cold (−80 °C) 80% (v/v) aqueous methanol. The insoluble material in the lysates was centrifuged at 5,000 *g* for 5 min. The resulting supernatant was evaporated using a speed vac. Samples were re-suspended using 20 μL HPLC grade water for mass spectrometry. 10 μL were injected and analyzed using a 5500 QTRAP triple quadrupole mass spectrometer (AB/SCIEX) coupled to a Prominence UFLC HPLC system (Shimadzu) via selected reaction monitoring (SRM) of a total of 287 endogenous water-soluble metabolites for steady-state analyses of the samples. Some metabolites were targeted in both the positive and negative ion mode for a total of 287 SRM transitions using positive/negative polarity switching. ESI voltage was +4900 V in the positive ion mode and −4500 V in the negative ion mode. The dwell time was 3 ms per SRM transition, and the total cycle time was 1.55 s. Approximately 10-14 data points were acquired per detected metabolite. Samples were delivered to the mass spectrometer via normal phase chromatography using a 4.6 mm i.d x 10 cm Amide Xbridge HILIC column (Waters Corp.) at 350 μL/min. Gradients were run starting from 85% buffer B (HPLC grade acetonitrile) to 42% B from 0-5 min; 42% B to 0% B from 5-16 min; 0% B was held from 16-24 min; 0% B to 85% B from 24-25 min; 85% B was held for 7 min to re-equilibrate the column. Buffer A was comprised of 20 mM ammonium hydroxide/20 mM ammonium acetate (pH = 9.0) in 95:5 water:acetonitrile. Peak areas from the total ion current for each metabolite SRM transition were integrated using Multi-Quant v2.0 software (AB/SCIEX).

QUANTIFICATION AND STATISTICAL ANALYSIS

Step I: identification of IRGs

Identification of IRGs

We identified IRGs from time-series RNA-seq data obtained in insulin-stimulated *Drosophila* S2R + cells as described in our previous study (Zirin et al., 2019). Briefly, we modeled the expression levels of each gene as the added sum of 3 parts: (i) the time effect after insulin stimulation, modeled as a cubic polynomial of time t; (ii) potential confounding components, such as batch effects, extracted using surrogate variables; and (iii) random white noise, modeled as a normal distribution with a mean of 0. We applied an F-test statistical framework to test each gene for the null hypothesis that the gene is not differentially expressed over time vs. the alternative hypothesis that the gene is temporally differentially expressed.

Hierarchical clustering of the IRGs

We subtracted the effects of the surrogate variables estimated by the statistical model used in the identification of the IRGs from the $\log_2$ scaled normalized read count for each IRG. We calculated the Z-score from the obtained values and performed hierarchical clustering using Euclidean distance and Ward's method. We analyzed all clusters containing 100 genes or more in hierarchical clustering.

KEGG pathway enrichment analysis

The enrichment of the genes in each pathway was determined using one-tailed Fisher exact test, and KEGG pathways with q values of less than 0.05 were defined as significantly enriched. The q values were calculated by the Benjamini-Hochberg procedure (Benjamini and Hochberg, 1995). We used the genes measured in the RNA-seq data as a background. Because the clusters analyzed in this study were not mutually exclusive, a pathway can be enriched in overlapping clusters. We selected non-overlapping and significant clusters based on the significance (q value) of Fisher exact test and the hierarchical structure of the dendrogram (see "Finding non-overlapping and statistically significant clusters of a hierarchical clustering").

Finding non-overlapping and statistically significant clusters of a hierarchical clustering

We selected non-overlapping and significant clusters for each of the pathways or TF motifs based on the significance (q value) of Fisher exact test and the hierarchical structure of the dendrogram. If the genes included in each of the two clusters overlapped, we defined the pair of clusters as "overlapping clusters" (*e.g.* clusters 1 and 2 in Figure 2B). If not, we defined a pair of clusters as "non-overlapping clusters" (*e.g.* clusters 2 and 3 in Figure 2B). We used "computational recognition and analysis of statistically significant subtrees (CRASSS)", which is an algorithm for finding the most statistically significant and non-overlapping clusters of hierarchical clustering (Buehler et al., 2004).

Identification of significantly co-regulated groups of IRG pairs

To identify significantly co-regulated groups of IRG pairs (*e.g.* pairs among the IRGs within a KEGG pathway), we tested whether the average Spearman correlation coefficient of IRG pairs was significantly higher or not. The same number of IRG pairs as the one that is used for hypothesis testing were randomly sampled 1000 times from all the pairs among the IRGs. For each sampling, the average Spearman correlation of the sampled IRG pairs was calculated. We obtained a p value by calculating the percentage of random sampling exhibiting larger average Spearman correlations than that calculated from the tested pairs. We defined a group of IRG pairs with Bonferroni-corrected p value less than 0.05 as significantly co-regulated. A similar method has been used in a previous study (Hansson et al., 2012). *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

Step II: prediction of TFs that regulate IRGs

Prediction of TF binding motif and inference of regulatory connections between TFs and genes

We predicted TFs regulating the IRG clusters based on the method developed in previous studies (Mina et al., 2015a; 2015b). The sequences of the flanking regions of genes were downloaded from FlyBase (version FB2018_04) (Larkin et al., 2021). The region from −1000 bp to +100 bp of the transcription start site was defined as the flanking region. Although distal regulatory DNA elements can affect gene expression levels, information on interaction with such regions is available only for a subset of the genes in the *Drosophila* genome. Furthermore, a prediction of TF binding motifs for all potential distal regulatory regions can increase false positives. For these reasons, we used the regions from −1000 bp to +100 bp of the transcription start site for our analysis. The TF binding motifs in each flanking region were predicted using Cis-BP (version 1.02), a TF binding motif database, and FIMO (version 4.12.0), a TF binding motif prediction tool (Grant et al., 2011; Weirauch et al., 2014). The information on TF binding motifs in Cis-BP was downloaded from the MEME suite (version 4.12.0) (Bailey et al., 2015). We used FIMO with the default parameters. We used only the binding motifs of TFs detected in any of the omic data analyzed in this study (PPI, phosphoproteome, and transcriptome). For the prediction of regulatory connections between TFs and IRGs, we performed TF motif enrichment analysis of the genes in each cluster. The enrichment of TF binding motif in the flanking regions of genes in each cluster was determined by one-tailed Fisher exact test, and TF binding motifs with q values less than 0.05 were defined as significantly enriched. The q values were calculated by the Benjamini-Hochberg procedure (Benjamini and Hochberg, 1995). We used the genes measured in the RNA-seq data as a background. If a TF binding motif was enriched in the promoter regions of the genes in a cluster, then we inferred the regulatory connections between the corresponding TF and the genes in the cluster. Because the clusters analyzed in this study were not mutually exclusive, a TF binding motif can be enriched for overlapping clusters. We selected the non-overlapping and significant clusters based on the significance (q value) of Fisher exact test and the hierarchical structure of the dendrogram (see "Finding non-overlapping and statistically significant clusters of a hierarchical clustering").

Confirmation of the TF predictions using data from ChIP-Atlas

For the validation of the predicted regulatory connections, we examined the overlap between the predicted target genes of each TF and those predicted from experimental ChIP data from the ChIP-Atlas database (Oki et al., 2018). Genes whose flanking region around the transcription start sites were detected in ChIP-sequencing peaks were obtained using BedTools (v2.21.0) with PyBedTools (Dale et al., 2011; Quinlan and Hall, 2010). We used the flanking regions from −1000 bp to +100 bp of the transcription start sites. The overlap between the predicted genes and genes from ChIP data was determined by one-tailed Fisher exact test, and those with Bonferroni adjusted p value less than 0.05 were defined as significant.

Step III: prediction of upstream signaling pathways regulating the predicted TFs

Identification of IRpPs

We extracted phosphopeptides whose phosphorylation levels were increased or decreased by insulin stimulation from phosphoproteome data measured in our previous study (Vinayagam et al., 2016). Accordingly, phosphopeptides that showed an absolute $\log_2$ fold change larger than 0.5 at any time point compared to time 0 were defined as IRpPs.

Identification of PPIs from the AP-MS data

PPIs and their detected time points were extracted from the AP-MS data of insulin-stimulated *Drosophila* S2R + cells measured in our previous study (Vinayagam et al., 2016).

Prediction of protein kinases for protein phosphorylation

We predicted kinases regulating the IRpPs using the NetPhorest software and Fisher exact test. First, we predicted possible kinases for the phosphopeptides measured in the phosphoproteomic dataset based on the amino acid sequences of the proteins corresponding to the phosphopeptides using a standalone version of NetPhorest for humans with the default parameters (Horn et al., 2014; Miller et al., 2008). The input data for NetPhorest are fly protein sequences in FASTA format. The outputs for NetPhorest are posterior probabilities of an amino acid residue being recognized by a protein kinase classifier (kinases with similar substrate recognition motifs). Among the candidate classifiers, we selected the classifier with the posterior probability value larger than 0.035 as well as the posterior to be higher than the prior as the kinase classifier related to the amino acid sequence. The threshold of the posterior probability value was applied from previous studies (Buljan et al., 2020; Freschi et al., 2014; So et al., 2015; Tan et al., 2009). A predicted KSR is represented as an edge between a kinase classifier as one node and a phosphorylation site as the other node. We also extracted individual kinases within each classifier from a table provided by NetworKIN software and defined these kinases as possible protein kinases regulating the predicted phosphosites for the kinase classifier. We converted the protein IDs (Ensembl) of the responsible protein kinases to the gene IDs (HGNC) using MyGene.info (version 3) (Wu et al., 2013; Xin et al., 2016) (http://mygene.info/). We converted the obtained human gene IDs of the responsible protein kinases to the fly gene IDs by DIOPT (DIOPT score >2) (Hu et al., 2011). Next, we performed Fisher exact test of the predicted target phosphopeptides of each kinase for the increased and decreased IRpPS. If the predicted target phosphopeptides of a kinase were enriched in the increased or decreased IRpPs (q value <0.05), then we selected the KSRs between the kinase and its predicted substrates contained in the increased or decreased IRpPs, respectively. The q values were calculated by the Benjamini-Hochberg procedure (Benjamini and Hochberg, 1995).

Prediction of upstream signaling pathways regulating the predicted TFs

We predicted upstream insulin signaling pathways regulating the predicted TFs using the IRpPs, the PPI network following insulin stimulation extracted from our previous study (Vinayagam et al., 2016), the kinases-IRpPs relationships predicted by NetPhorest (Horn et al., 2014; Miller et al., 2008) and Fisher's exact test, and the MIST PPI network (Hu et al., 2018). We performed the prediction in the following three steps as shown in Figure 3D. In Step III-I, we constructed a network by merging the IRpPs (Figure 3A), the PPI network following insulin stimulation (Figure 3C), and the predicted KSRs (Figure 3B), and the predicted TFs (Figure 2B). Hereafter, we denoted the obtained network as the "PPI and phosphorylation network". We also merged the MIST PPI network and the PPI and phosphorylation network. To obtain the subnetwork involved in signal transduction, we extracted the subnetworks that consisted only of the proteins in signaling pathways in the KEGG database, the predicted kinases, or the predicted TFs from the obtained network. The signaling pathways in the KEGG database were defined as pathways including the character string of "signaling pathway" in their names. We denoted this subnetwork as the "merged network". In

Step III-II, we extracted the shortest paths from the nodes in the PPI and phosphorylation network to each of the TFs from the merged network. Then we merged the PPI and phosphorylation network and the network consists of the extracted shortest paths. In Step III-III, to predict the signaling pathways activated/inhibited in an insulin-dependent manner, we extracted the paths from the insulin receptor to each of the TFs in the network obtained in Step III-II. We defined the paths from the insulin receptor to a TF as the insulin signaling pathway regulating the TF. In this study, the maximum shortest path length from the insulin receptor to each TF is 5. We extracted the paths whose lengths were 5 or less to reduce the false positives while maximizing the number of the TFs with predicted upstream signaling pathways.

KEGG signaling pathway enrichment analysis for the upstream signaling pathways of each TF

The significance of the overlap between proteins in the signaling pathways of the KEGG database and the upstream signaling pathways of each TF was determined by one-tailed Fisher exact test, and signaling pathways with q values less than 0.05 were defined as significantly enriched. The q values were calculated by the Benjamini-Hochberg procedure (Benjamini and Hochberg, 1995). We used the genes in any of the signaling pathways in the KEGG database as a background.

Step IV: identification of IRMs

Metabolites that were detected in less than 60% of replicates (3 replicates) at any time point following insulin stimulation (0, 60, 120, and 180 min) were removed from the analysis. To correct the bias among samples, the area under the peak obtained by the LC-MS measurement of each metabolite was normalized by the median of the area under the peak of all metabolites for each sample. The significance of the change at each time point compared to time 0 was tested by two-tailed Welch's t-test for each metabolite. Metabolites that showed an absolute $\log_2$ fold change larger than 0.5 and an FDR-adjusted p value (q value) less than 0.1 at any time point were defined as IRMs. The q values were calculated by Storey's procedure (Storey and Tibshirani, 2003).

Step V: connection of the IRMs/IRGs and metabolic reactions

Identification of allosteric regulation

We identified IRMs that function as allosteric regulators for metabolic enzymes using the BRENDA database, which is a database with information regarding allosteric effectors and their target enzymes (Chang et al., 2021). A metabolite can operate as an activator for some enzymes and as an inhibitor for others. We obtained the entries for metabolic enzymes from the KEGG database and extracted their allosteric effector (activator and inhibitor) information, as reported for at least one organism in BRENDA. Then, we associated the standard compound names of allosteric effectors used in BRENDA with metabolite names that were used in KEGG to obtain the KEGG compound ID related to each allosteric effector.

Identification of substrates, products, and metabolic enzyme genes involved in metabolic reactions

We extracted the substrates, products, and metabolic enzymes related to each metabolic reaction from the KEGG database. Because the reversibility of metabolic reactions was not available in a comprehensive manner, metabolic reactions were presumed to be regulated by both the substrate and product.

Step VII: integration of the trans-omic network of insulin action with a CRISPR screen for cell proliferation

Identification of screen hits of the CRISPR knockout screen for cell proliferation

We identified screen hits that significantly affect cell proliferation from the CRISPR screen for cell proliferation measured in *Drosophila* S2R + cells as described in our previous study (Viswanatha et al., 2018).

Identification of differences in the distributions of Spearman correlation coefficients between groups of IRG pairs

Differences in the distribution of Spearman correlation coefficients between groups of IRG pairs were tested using the Wilcoxon rank-sum test. We defined distributions between groups with Bonferroni-corrected p values less than 0.05 as significantly different. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

Prediction of subnetworks required for cell growth in an insulin-dependent manner

We aimed to identify subnetworks that are required for cell growth in an insulin-dependent manner by using the trans-omic network of insulin action and the CRISPR screen data for cell proliferation (Viswanatha et al., 2018). We hypothesized that both the upstream signaling molecules and the target IRGs of a TF which is required for cell growth in an insulin-dependent manner are likely to negatively/positively affect cell growth. We predicted such TFs by GSEA (Subramanian et al., 2005) using the Z-scores of the CRISPR screening and the information of the upstream signaling molecules and the target IRGs of each TF in the trans-omic network. We first extracted TF-centric subnetworks each containing a TF, its upstream signaling molecules (Figure 3E), and its target IRGs (Figure 2B). Next, we performed GSEA using the Z-score of the CRISPR screening for both the upstream signaling molecules and the target IRGs of the TF in each TF-centric subnetwork. When the results of GSEA were significant for both the upstream signaling molecules and the target IRGs for a TF (FDR <0.05), and the TF itself is a screen hit, we considered the subnetwork potentially involved in cell growth. It has been proposed that the activity of a regulatory protein (*e.g.* a TF and a kinase) can be estimated based on the statistics (*e.g.* abundance and fold change) of their target molecules (*e.g.* target genes regulated by a TF and phosphopeptides regulated by a kinase) (Alvarez et al., 2016; Dugourd and Saez-Rodriguez, 2019), which is the foundation for the method used in this study to estimate the importance of each TF to cell growth by using the Z-scores of the CRISPR screening of the upstream molecules and downstream target genes of the TFs.

Implementation

Statistical analyses and trans-omic network analysis were performed using Python 3.8.3 (https://www.python.org) or R 3.6.3 (R Core Team, 2021). Network diagrams were visualized using Cytoscape 3.8.2 (Shannon et al., 2003).