Supporting Information

Figure Legends

Supplementary Figure 1. Multiple amino acid alignment of activation loop sequences of ERK2 homologs from several eukaryotes. In addition to the canonical threonine and tyrosine activation sites (blue boxes), the threonine at position 188 (by rat ERK2 numbering) is also highly conserved (red box). In addition to the canonical threonine and tyrosine activation sites (blue boxes), the threonine at position 188 (by rat ERK2 numbering) is also highly conserved (red box). Serine sometimes occurs in the comparable position[1,2].

Supplementary Figure 2. MD simulation of changes of T188-OH distance to ATP over a 25ns window.

Supplementary Figure 3. ERK2 differentially binds a range of oligonucleotides. (A) Gel shift assays with activated ERK2 and canonical phosphorylation site-mutants with radiolabeled oligonucleotides. (B) Gel shift assay with IRDye 700 5'-tagged oligonucleotide (mIns(19)) that maps to the insulin promoter and contains the motif GAAAC.

Supplementary Figure 4. Molecular model of potential conformational changes in T188-OH distance to ATP in ERK2.

Supplementary Figure 5. Chromatin immunoprecipitation (ChIP) results from H358 cells indicating ERK1/2 associates with target genes in cells stimulated with epidermal growth factor (EGF) for 15 min. Recruitment is severely reduced when cells were treated with the MEK1/2 inhibitor PD0325901 3 hours prior to EGF stimulation.

Supplemental Methods

Expression and purification of ERK2. BL21 DE 3 competent cells were transformed with NpT7 plasmids containing N-terminal 6xHistidine-tagged rat ERK2 (His_6-ERK2) cDNA and grown in Lennox Broth at 37°C until cultures reached an OD_{600} of 0.8 – 0.9. Cultures were induced with 0.5 mM isopropyl thiogalactopyranoside (IPTG) and grown for 14-16 h at 30°C. Cells were harvested in a sonication

buffer of 50 mM $NaHPO_4$ pH 8.0, 300 mM NaCl, with protease inhibitors and pellets were flash frozen in liquid nitrogen and stored at -80°C until purification. Pellets were thawed and lysozyme was added on ice. Following sonication, the lysate was centrifuged in a Ti-45 rotor at 105,000xg, for 1 h at 4°C. The clarified lysate was mixed with nickel resin (Clontech) overnight at 4°C with stirring. The resin was washed 3X with ice-cold sonication buffer containing protease inhibitors and 3X with 50 mM Tris, pH 7.4, 150 mM NaCl, 20 mM imidazole, pH 7.5. Proteins were eluted from the resin with a 20-200 mM imidazole gradient; fractions were pooled based on ERK2 protein profile on gels and diluted with 50 mM Tris, pH 8.0, 20 mM NaCl, 1 mM dithiothreitol (DTT), 0.5 mM EDTA. Proteins were used in partially purified form or further purified on MonoQ (see below).

Preparation from Origami bacterial cells. Origami competent cells (Millipore) were transformed with NpT7 plasmids containing the same ERK2 construct and grown in Terrific Broth at 30°C until cultures reached an OD_{600} of 0.55 – 0.65. Cultures were induced as above and pellets were harvested. Following sonication and clarification as above, the lysate was mixed with nickel resin for 2 h at 4°C. An additional wash with 5 ml 50 mM HEPES pH 8.0, 300 mM NaCl, 50 mM imidazole was included before elution. Proteins were eluted from beads with 5 ml 50 mM HEPES pH 8.0, 300 mM NaCl, 250 mM imidazole. Eluates were collected from sedimented beads and dialyzed into either 20 mM NaCl for subsequent ion exchange chromatography or into 20 mM HEPES pH 7.4, 1 mM EGTA, 1 mM DTT, 10% glycerol, flash frozen, and stored at -80°C.

MonoQ purification of 6XHis tagged ERK2. Proteins exchanged into 20 mM NaCl were applied to MonoQ and eluted with a linear gradient of 20-1000 mM NaCl, 7.5-column volumes, at a flow rate of 0.4 ml/min. Pooled and individual fractions were subjected to mass spectroscopy or dialyzed into the buffer above, flash frozen and stored at -80°C until further use.

Computational methods

System setup and equilibration. Protein coordinates were retrieved from the Protein Data Bank (PDB), PDB codes 1ERK for inactive ERK2, 2ERK for active ERK2. 2ZOQ was used as a template to model a monophosphorylated ERK2 at Y185; this was generated using MODELLER. Once the three initial structures were obtained, the following thermalization and equilibration protocol was performed in explicit solvent for all of the systems: hydrogens were added with the t-leap module of the AMBER 12 program, and the amber99-SB force field was used for the simulations. Standard protonation states were assigned to tritable residues (D and Q were assigned negative charge, K and R positive charge). Histidine protonation was assigned favoring formation of hydrogen bonds in the crystal structure. The parameters for phosphotyrosine and phosphoserine were taken from Craft et al. [3] and the ATP parameters from Meagher et al [4]. Each protein was immersed in a truncated octahedral box of TIP3P water [5]. After the solvation the systems consisted of 361 amino acids, over 10,000 water molecules, and two Mg^{2+} ions and ATP for ERK2-ATP. Each system was optimized using a conjugate gradient algorithm for 1.000,000 steps. This optimization was followed by 100 ps long constant volume molecular dynamics (MD) thermalization where the temperature of the system was slowly raised to 300K. The heating was followed by 400 ps of constant temperature and constant pressure. MD simulation was used to equilibrate the system's density. During these processes the protein Cα atoms were restrained by a 10 kcal/mol harmonic potential for the thermalization and a 1 kcal/mol harmonic potential for the density equilibration.

MD Simulations. For each system 25 ns long constant temperature and pressure MD simulations were performed with no restraint for the Cα atoms. Pressure and temperature were kept constant with the langevin thermostat. All simulations were performed with periodic boundary conditions. SHAKE was used to constrain bonds involving hydrogen atoms, and the time step was 2.0 fs. The particle mesh Ewald (PME) algorithm was employed to calculate the long-range electrostatics. The non-bonded cutoff was set to Å. During the sampling process, the coordinates were saved every 0.02 ps.

Umbrella sampling. The initial structure was relaxed by 25 ns MD simulations from the canonical ensemble. The umbrella sampling MD simulations were carried out with the pmemd program in AMBER12 to simulate the conformational transition path between inactive and reactive state. During the simulations, a conformational parameter, the distance ATP γ phosphate- T188 OG1, was selected as the reaction coordinate of the reaction path for the conformational change. Here, the whole simulation consists of 25 simulation windows by steadily decreasing the distance of T188-ATP from 9.50 Å to 3.50 Å with a step of 0.25 Å. For each sampling window, the distance was restrained to the new value by an artificial harmonic biasing potential with a force constraint of 40 kcal/mol/Å2. The MD simulations of 1500 ps were performed for each sampling window, and a constant temperature of 300K was controlled by the weak coupling algorithm. SHAKE was applied to fix all bonds involving hydrogen atoms, and the time step was 2.0 fs. The last snapshot of each sampling window was used as the starting coordinates for the next sampling window. The construction of the PMF was performed with WHAM.

References

(1) Manning, G., Whyte, D. B., Martinez, R., Hunter, T., and Sudarsanam, S. (2002) The protein kinase complement of the human genome, *Science 298*, 1912-1934.

(2) Goldsmith, E. J., Akella, R., Min, X., Zhou, T., and Humphreys, J. M. (2007) Substrate and docking interactions in serine/threonine protein kinases, *Chem Rev. 107*, 5065-5081.

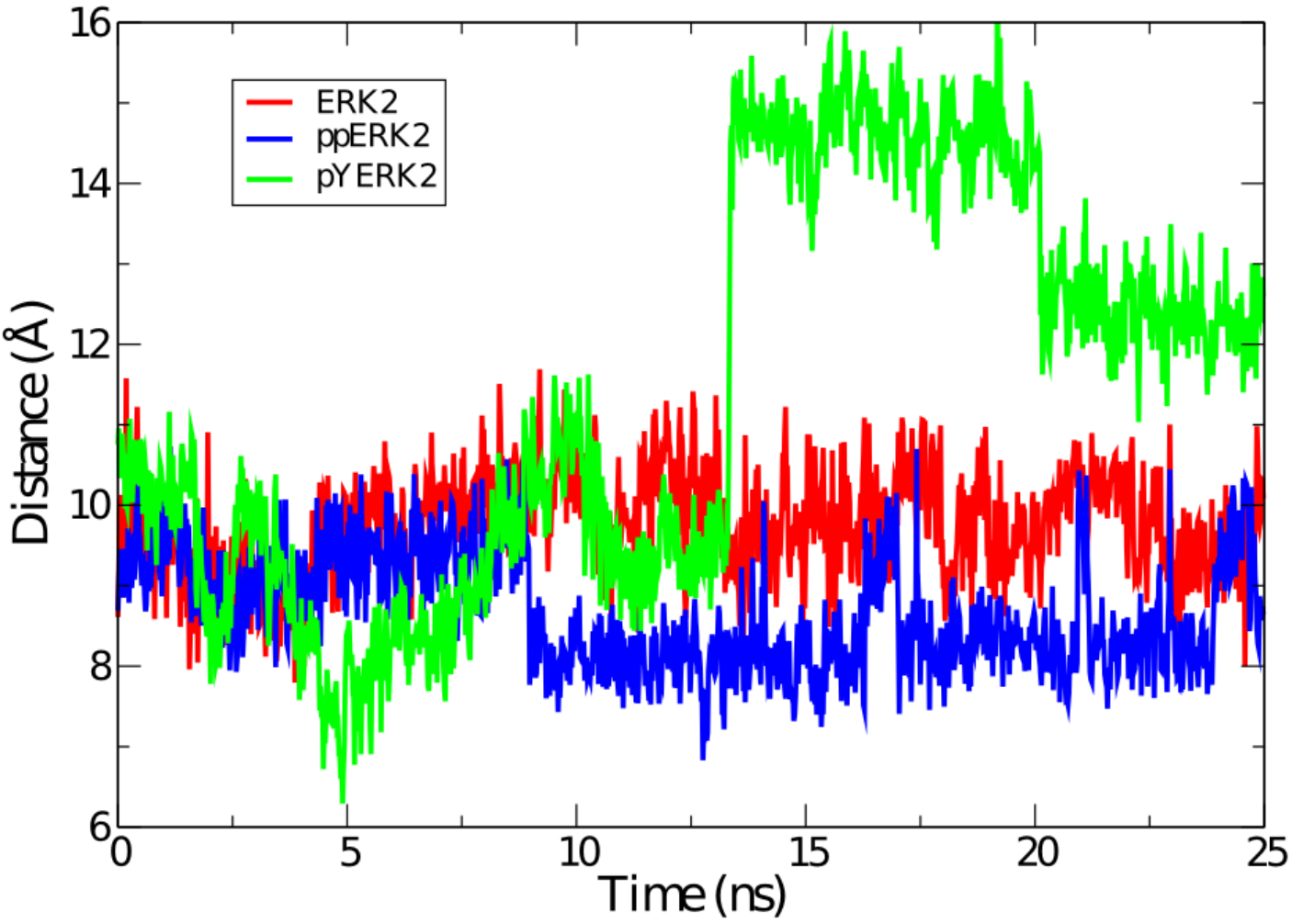
(3) Perez, M. M. and Turjanski, A. G. (2015) MAPK's networks and their capacity for multistationarity due to toric steady states, *Math. Biosci. 262*, 125-137.

(4) Meagher, K. L., Redman, L. T., and Carlson, H. A. (2003) Development of polyphosphate parameters for use with the AMBER force field, *J. Comput. Chem. 24*, 1016-1025.

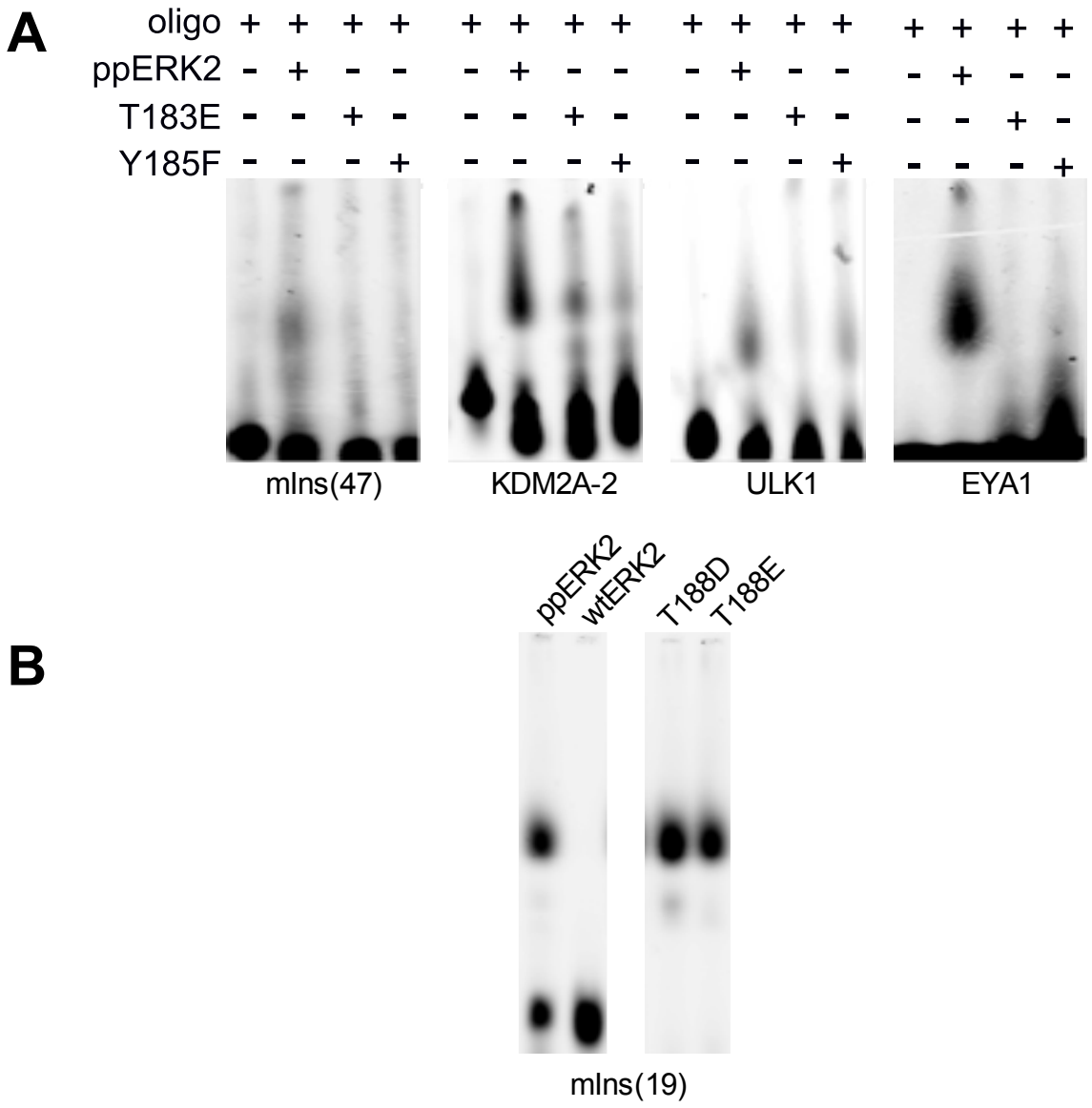
(5) Craft, J. W., Jr. and Legge, G. B. (2005) An AMBER/DYANA/MOLMOL phosphorylated amino acid library set and incorporation into NMR structure calculations, *J. Biomol. NMR 33*, 15-24.

R. norvegicus	181-	F	L	T	E	Y	V	A	T	R	W	-190
M. musculus	181-	F	L	T	E	Y	V	A	T	R	W	-190
C. griseus	190-	F	L	T	E	Y	V	A	T	R	W	-199
H. sapiens	183-	F	L	T	E	Y	V	A	T	R	W	-192
D. melanogaster	196-	F	L	T	E	Y	V	A	T	R	W	-205
D. rerio	192-	F	L	T	E	Y	V	A	T	R	W	-201
C. elegans	254-	F	L	T	E	Y	V	A	T	R	W	-263
Human GSK3β	213-	P	N	V	S	Y	I	C	S	R	Y	-222

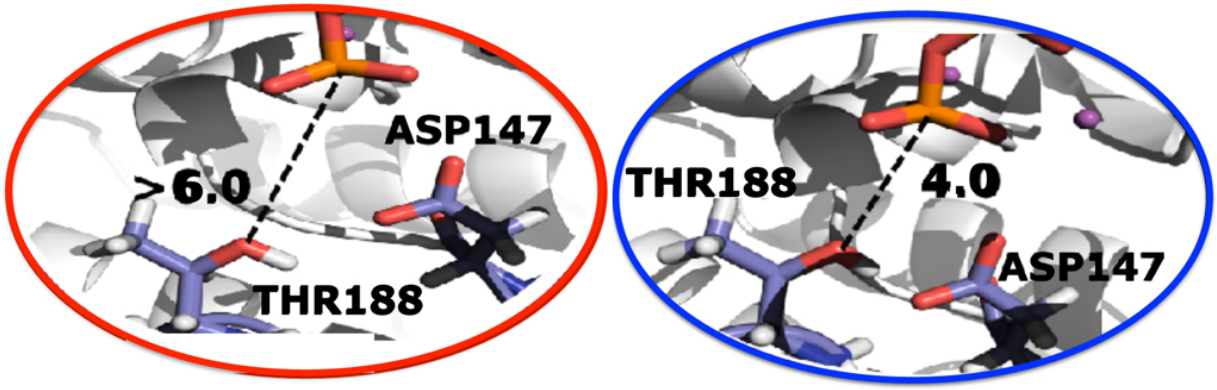
Supplementary Figure 1.



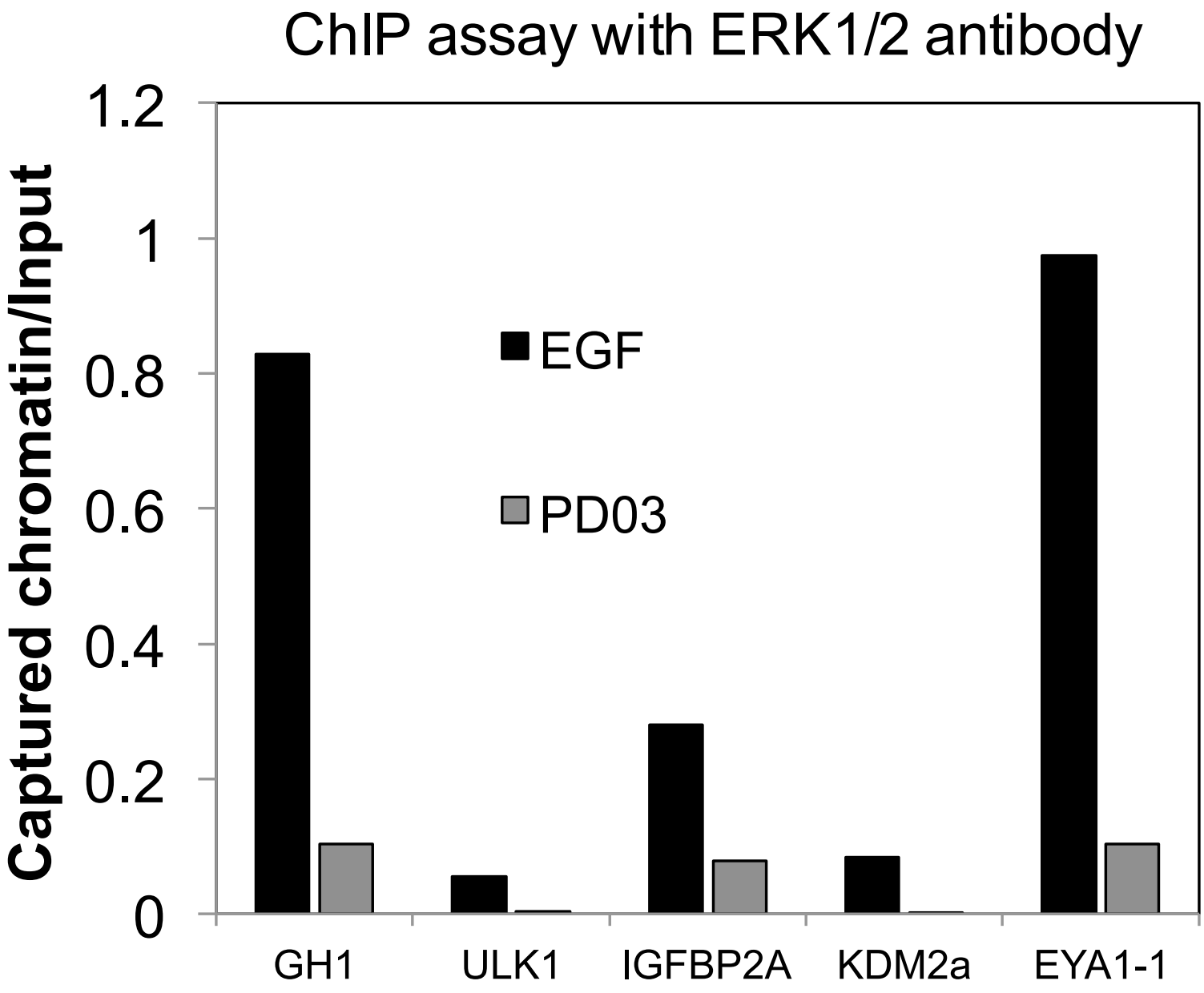
Supplementary Figure 2.



Supplementary Figure 3.



Supplementary Figure 4.



Supplementary Figure 5.

Table 1. Sequences of oligonucleotides used in this study

Oligonucleotide	Sequence (duplexes)
mIns(47)	GGAAGTGTTTGGAAACTGCAGCTTCAGCCCCTCTGGCCATCTGCTGAC
mIns(19)	GGAAACTGCAGCTTCAGCC
KDM2a-2	GAGGCGGGGCCAGGCCGCTGCACCGAGGAGCCCCCGGGCCCAGGGCAGACGGGCCCCGTG
EYA1	GGTTGAATGAAAATGCACATCAAAGAT
GH1	AGGCTCAGCTAATTTTTGTTTTTTTGGTAGAGACGGGGTTTCAC
GH1	GGATTACAGGCGTGAACCACTGCTCCCTTCCCTGTCCTTCTGATTTTAAAATAACTATAC
ULK1	GGAAAAGGAGGGAGGAGGACGGAGAGAGGGGAAAGG
IGFBP2	AGAAATCTAAAGTGAAGAATGCTCCAGGAACAATGGGATGGGCAAAGC

Table 2. Associated gene regions of oligonucleotides used in this study

Oligonucleotide	Associated gene	Region
mIns(47)	mouse insulin 1 and 2, rat insulin 2	promoter
mIns(19)	mouse insulin 1 and 2, rat insulin 2	promoter
KDM2a-2	human lysine demethylase 2a-2	3' UTR
EYA1	human eyes absent homolog 1	intronic
GH1	human growth hormone (#1)	promoter
GH1	human growth hormone (#2)	promoter
ULK1	Unc51-like autophagy activating kinase	promoter
IGFBP2	human insulin-like growth factor binding protein 2	promoter