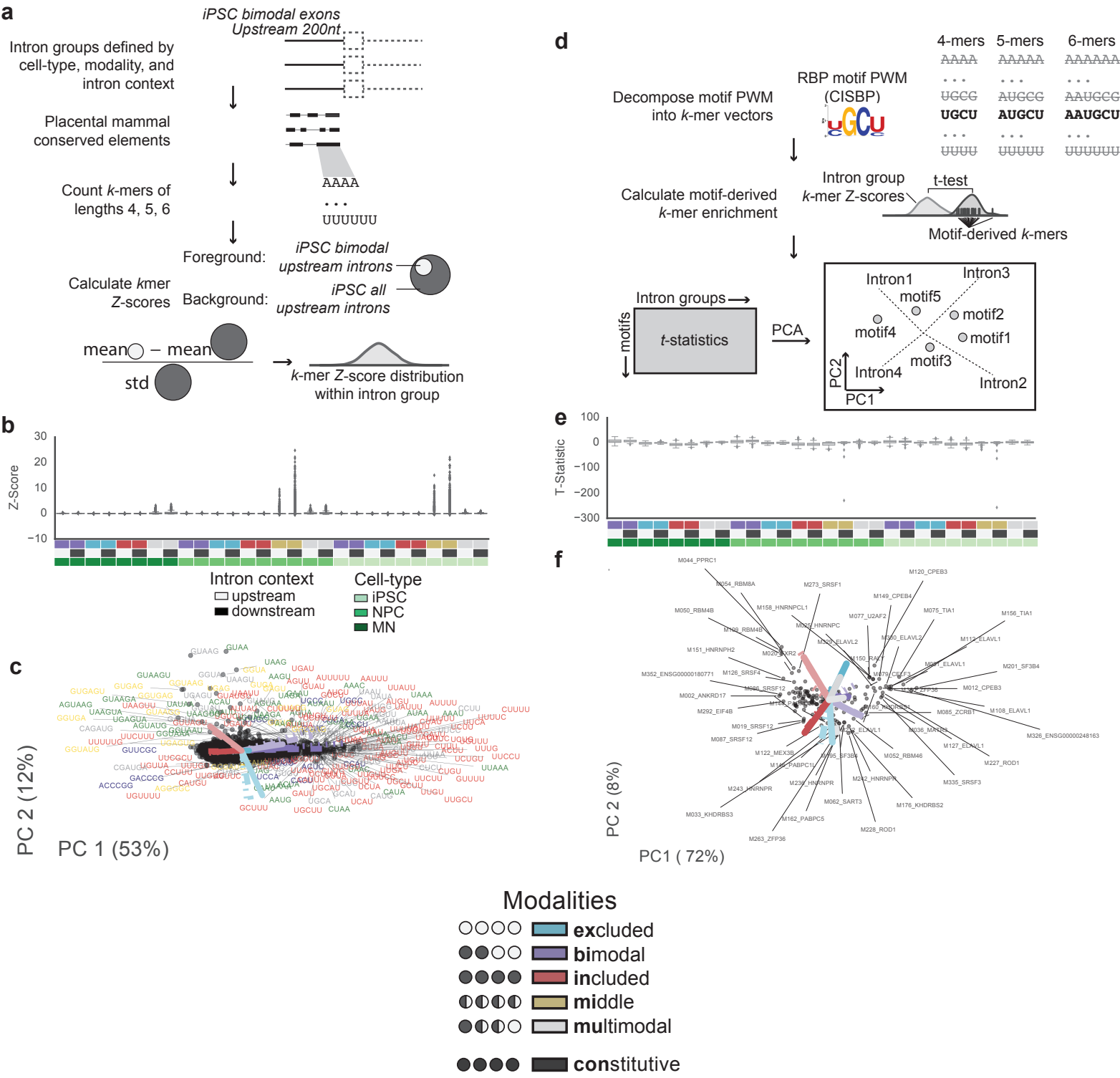
a
iPSC bimodal exons
Upstream 200nt
Intron groups defined by cell-type, modality, and intron context
Placental mammal conserved elements
Count k-mers of lengths 4, 5, 6
AAAA
. . .
UUUUUU
Calculate kmer Z-scores
Foreground:
iPSC bimodal upstream introns
Background:
iPSC all upstream introns
mean – mean
std
k-mer Z-score distribution within intron group
b
Z-Score
Intron context
upstream
downstream
Cell-type
iPSC
NPC
MN
c
PC 2 (12%)
PC 1 (53%)
d
Decompose motif PWM into k-mer vectors
RBP motif PWM (CISBP)
4-mers
5-mers
6-mers
AAAA
UGCG
UGCU
UUUU
AAAAA
AUGCG
AUGCU
UUUUU
AAAAAA
AAUGCG
AAUGCU
UUUUUU
Calculate motif-derived k-mer enrichment
Intron group k-mer Z-scores
t-test
Motif-derived k-mers
Intron groups
motifs
t-statistics
PCA
Intron1
Intron2
Intron3
Intron4
motif1
motif2
motif3
motif4
motif5
PC1
PC2
e
T-Statistic
f
PC 2 (8%)
PC1 (72%)
Modalities
excluded
bimodal
included
middle
multimodal
constitutive