A versatile mouse model of epitope-tagged histone H3.3 to study epigenome dynamics

Mahesh Bachu[1], Tomohiko Tamura[1,2], Chao Chen[1], Ankur Narain[1,3], Vishal Nehru[1], Naoyuki Sarai[1,4], Sukhendu B Ghosh[1,5], Anu Ghosh[1,5], Raghuveer Kavarthapu[6], Maria L. Dufau[6], Keiko Ozato[1*]

From the [1]Division of Developmental Biology, [6]Section on Molecular Endocrinology, Division of Developmental Biology, National Institute of Child Health and Human Development, National Institutes of Health, Bethesda MD, 20892, [2]Department of Immunology, Yokohama City University Graduate School of Medicine, Yokohama, Japan.

Running title: *Epitope tagged H3.3 mouse model*

Present Address: [3]Centre for Chromosome Biology Biochemistry, School of Natural Sciences National University of Ireland, Galway, Ireland, [5]Bhabha Atomic Research Center, Trombay, Mumbai, India

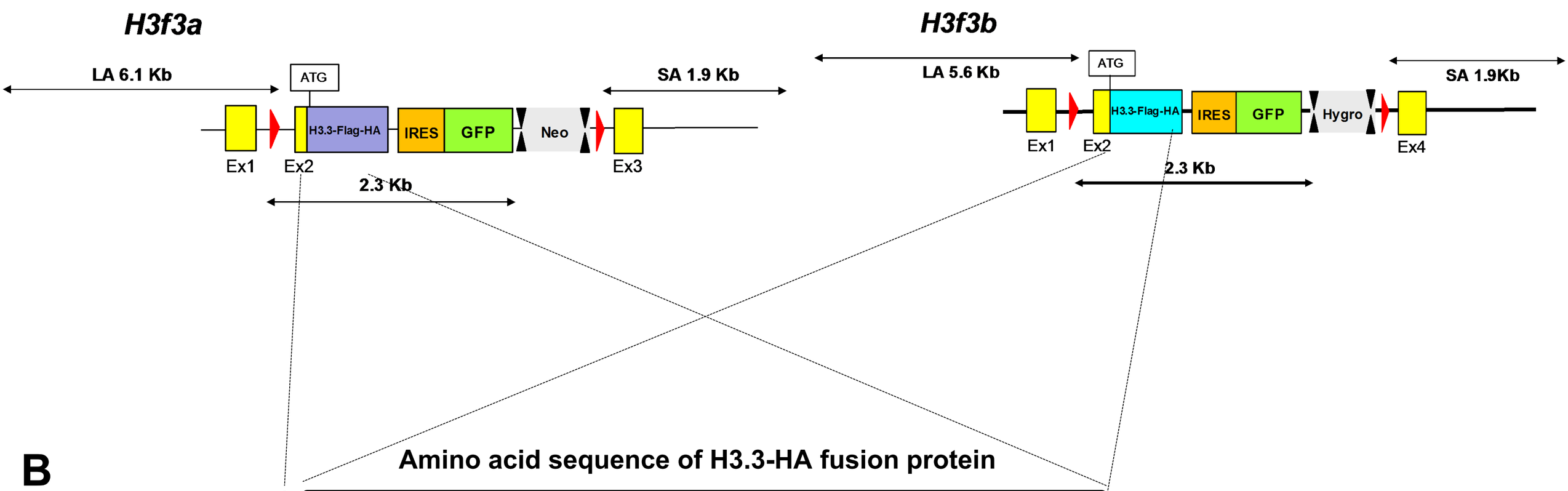
*To whom correspondence should be addressed: Keiko Ozato: SMGI, NICHD, NIH, ozatok@nih.gov

List:

1. Supporting Information 1
2. Supporting Information 2
3. Supporting Information 3
4. Supporting Information 4
5. Supporting Information 5
6. Supporting Information 6
7. Table S1
8. Table S2
9. Table S3

Supporting Information 1

A



B

Amino acid sequence of H3.3-HA fusion protein

H3.3

MARTKQTARK STGGKAPRKQ LATKAARKSA P**S**TGGVKKPH RYRPG — H3.3-Tail region

TVALREIRRY QKSTELLIRK LPFQRLVREI AQDFKTDLRF QS**A**A**I**

GALQEA**S**EAY LVGLFEDTNL CAIHAKRVTI MPKDIQLARR IRGERA — H3.3-globular region

AAAGG**DYKDD** **DDK**SAAGG**YP** **YDVPDYA*** 2.77 kD

FLAG-tag HA-tag

C

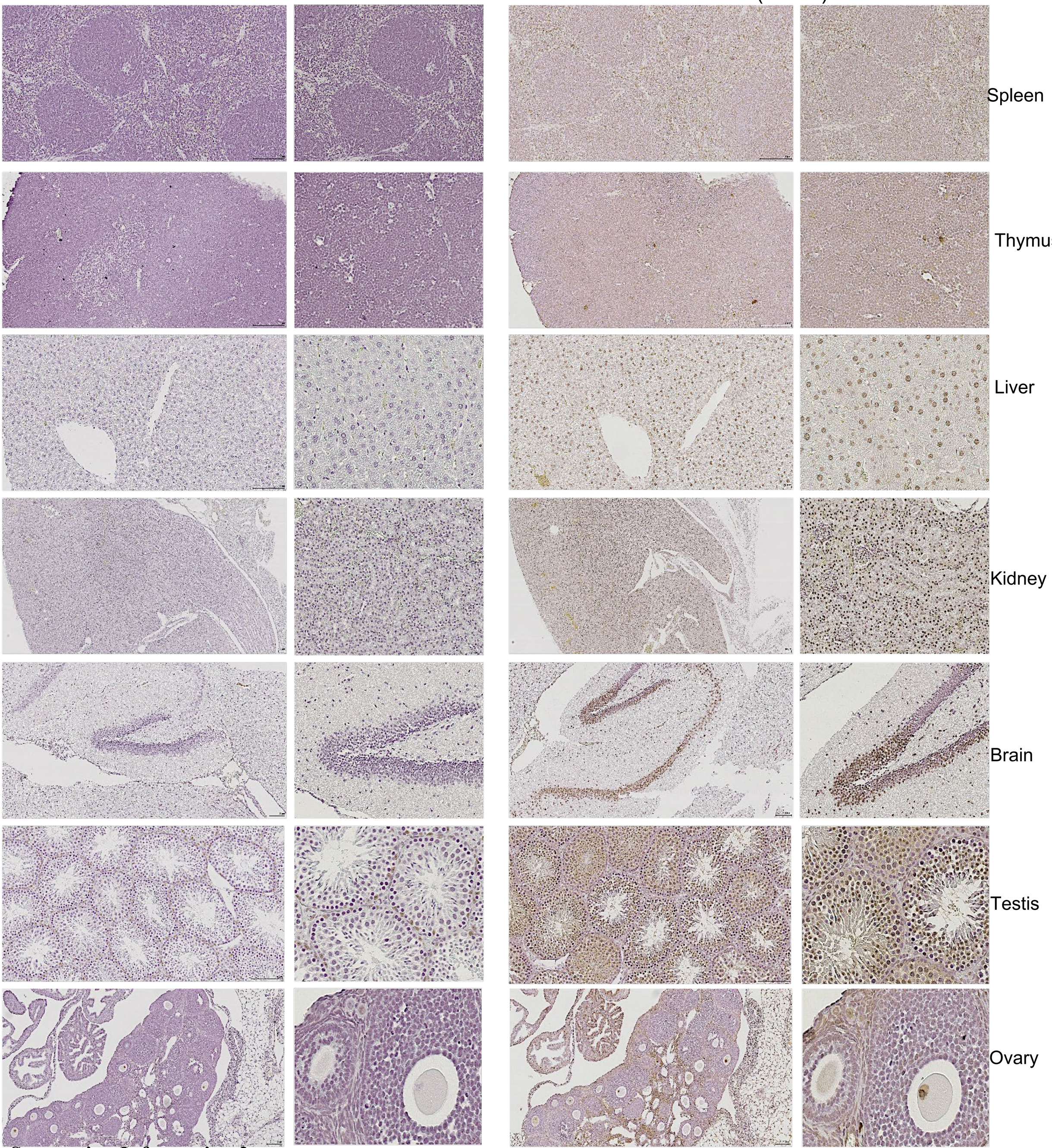
Primer sequence for genotyping H3f3a and H3f3b knock-in mice

Primer	5' ------> 3'
A_Common (RP)	GACATCAGCCATTAGAGGGCACAAGTTGCC
A_Common (FP)	GCACACACAGCCCATCAATCTGTGCTAAGACC
A_GFP (FP)	GGCGACGGCCCCGTGCTGCTGCCCGACAAC
B_Unique (RP)	GCCCTTCCTTCTTCTGGTGACTGCAGGCCAGG
B_Common (FP)	GCAACTTGTCACTCCTGAGCCACAGTGCTC
B_GFP (FP)	GGAGTTCGTGACCGCCGCCGGGATCAC

Supporting Information 1

(A) Schematic representation of targeting vectors. The Yellow rectangles represent *H3f3a* and *H3f3b* coding regions. The H3.3-Flag-HA-IRES-GFP cassettes are shown by rectangles of different color. Solid line represents genomic sequence. ATG denotes the initiation codon, FRTs are shown by double black rectangles Neo and Hygro indicate Neomycin and Hygromycin resistance genes. LoxP sites are shown in red triangles. The length of replaced sequences are shown at the bottom. Sizes in the map are not in scale. (B) Amino acid sequence of the inserted H3.3-HA fusion protein. Highlighted at the N-terminal end is the Flag-HA sequence adding an additional 2.7 kDa mass to the H3.3 protein. C) Primer sequences used for PCR based genotyping for *H3f3a* and *H3f3b*.

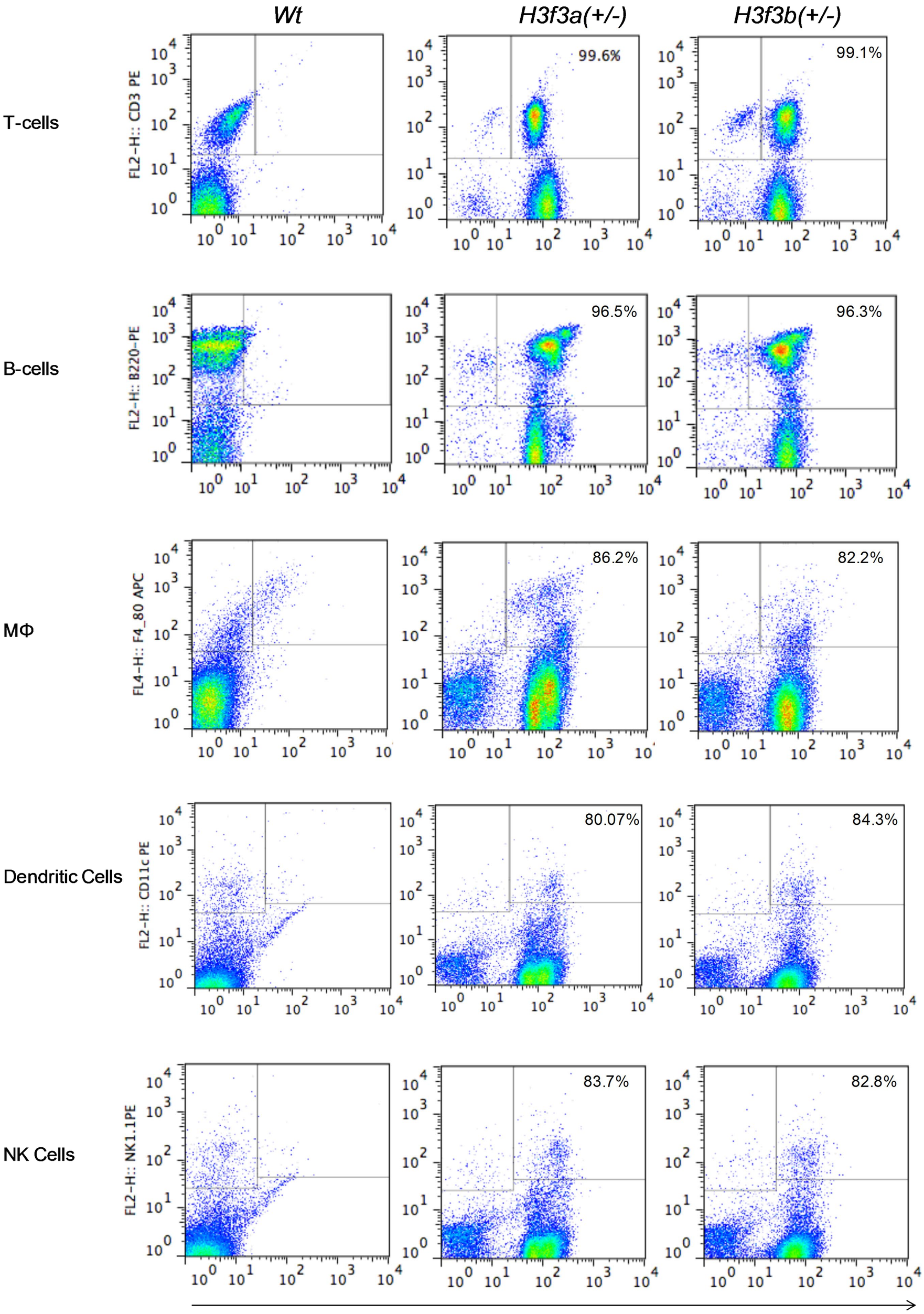
Supporting Information 2



Supporting Information 2

IHC analysis of H3f3b-HA expression in adult tissues. Tissue sections from 9 weeks old mice were examined in the same manner as Figure 2. Images corroborate IHC data of H3f3a-HA and show that H3f3b is expressed in most of indicated adult tissues. See Legends for Figure 2 for details

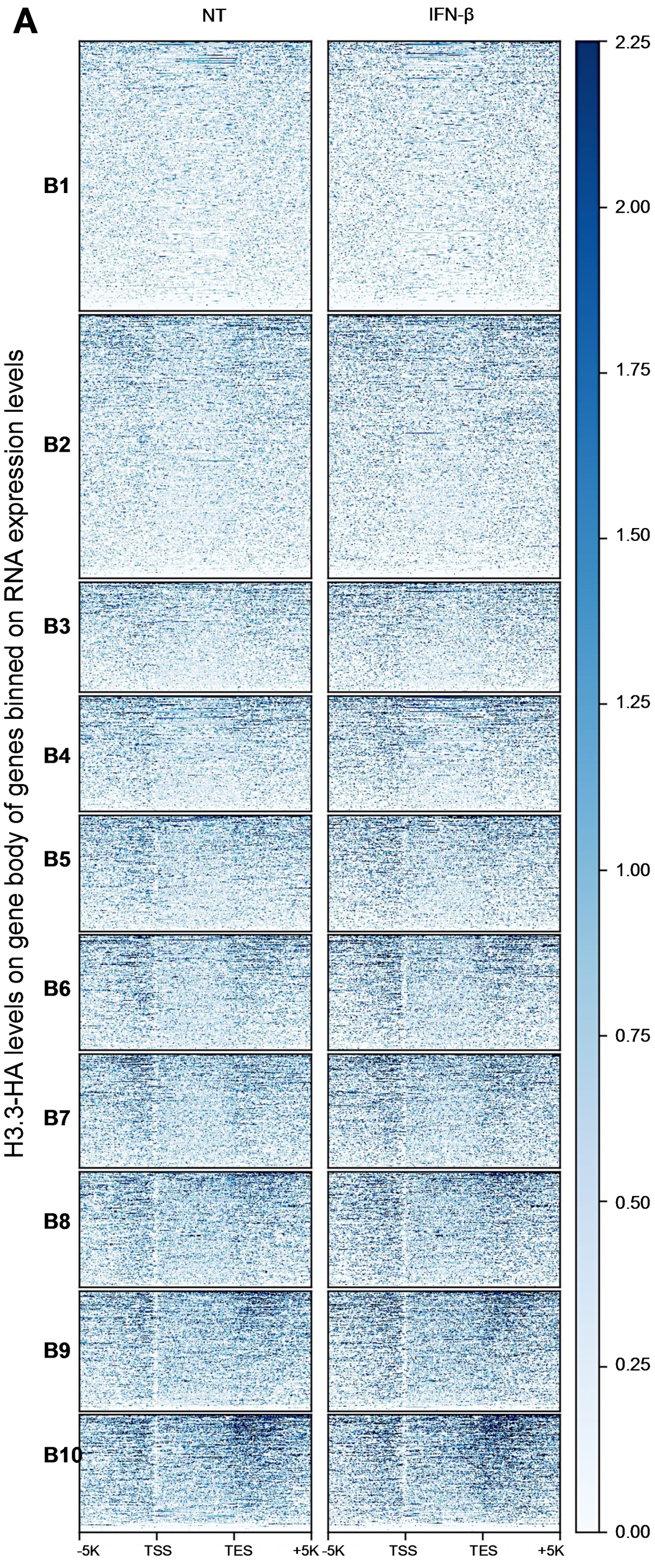
Supporting Information 3

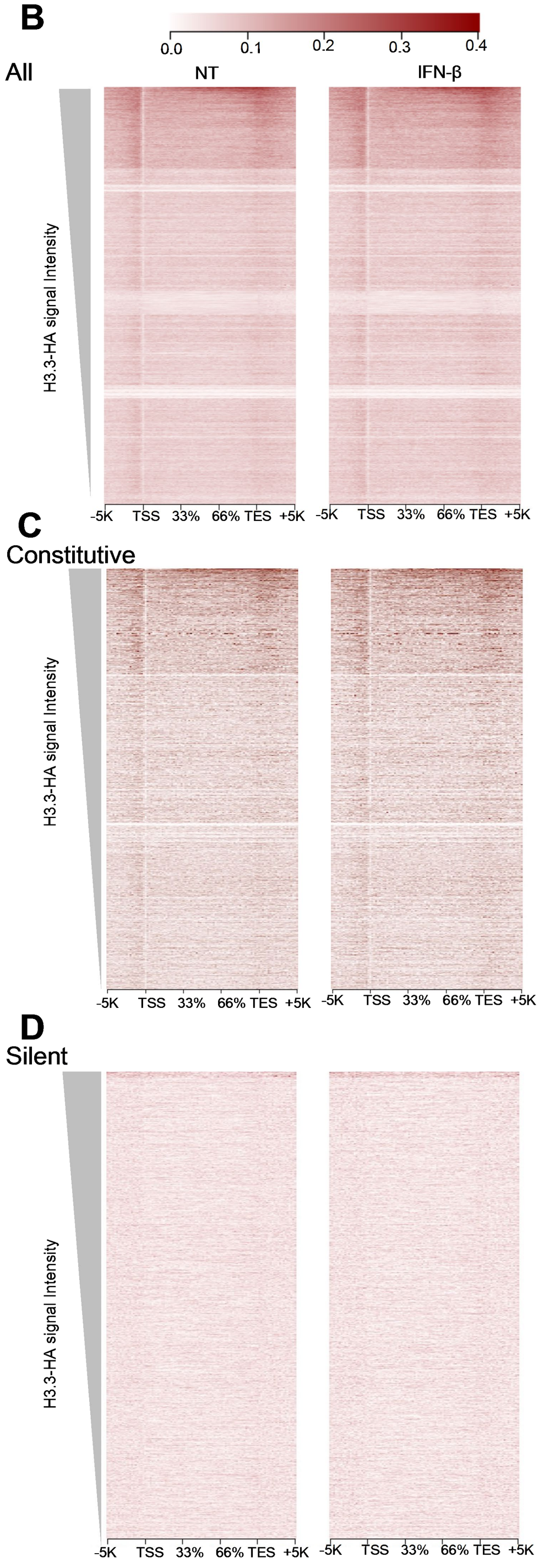


Supporting Information 3

Flow cytometry analysis of EGFP signals in immune cells. Spleen cells from H3f3a-HA and H3f3b-HA knock-in mice (8 or 9 weeks old) were analyzed for expression of surface markers (Y axis) and eGFP signals (X axis). T and B cells were distinguished by CD3 and B220, respectively. Dendritic cells, macrophages and NK cells were identified by Cd11c, F4/80 and NK1.1, respectively. Note that more than 80 % of cells were EGFP positive in all cell types.

Supporting Information 4



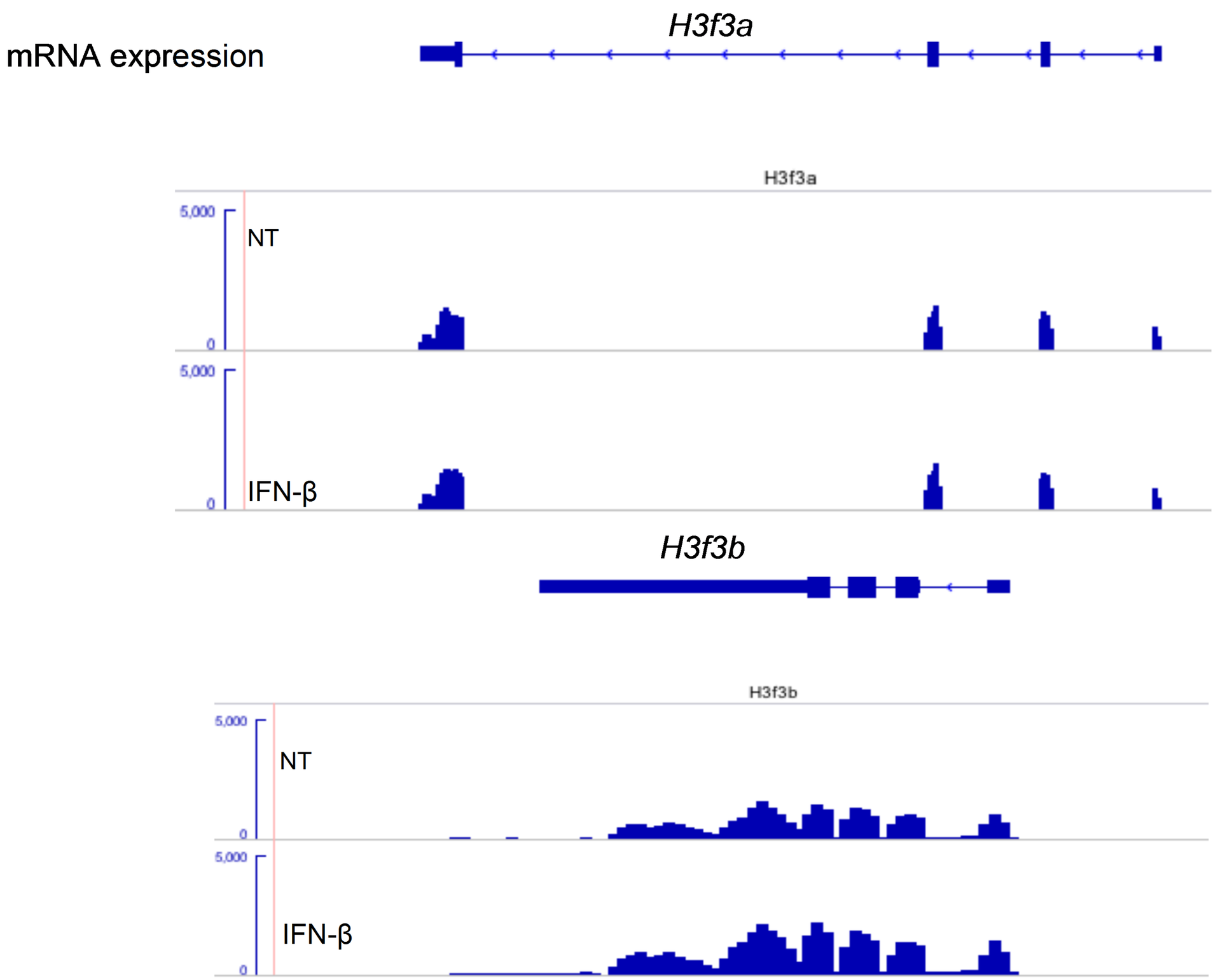


Supporting Information 4

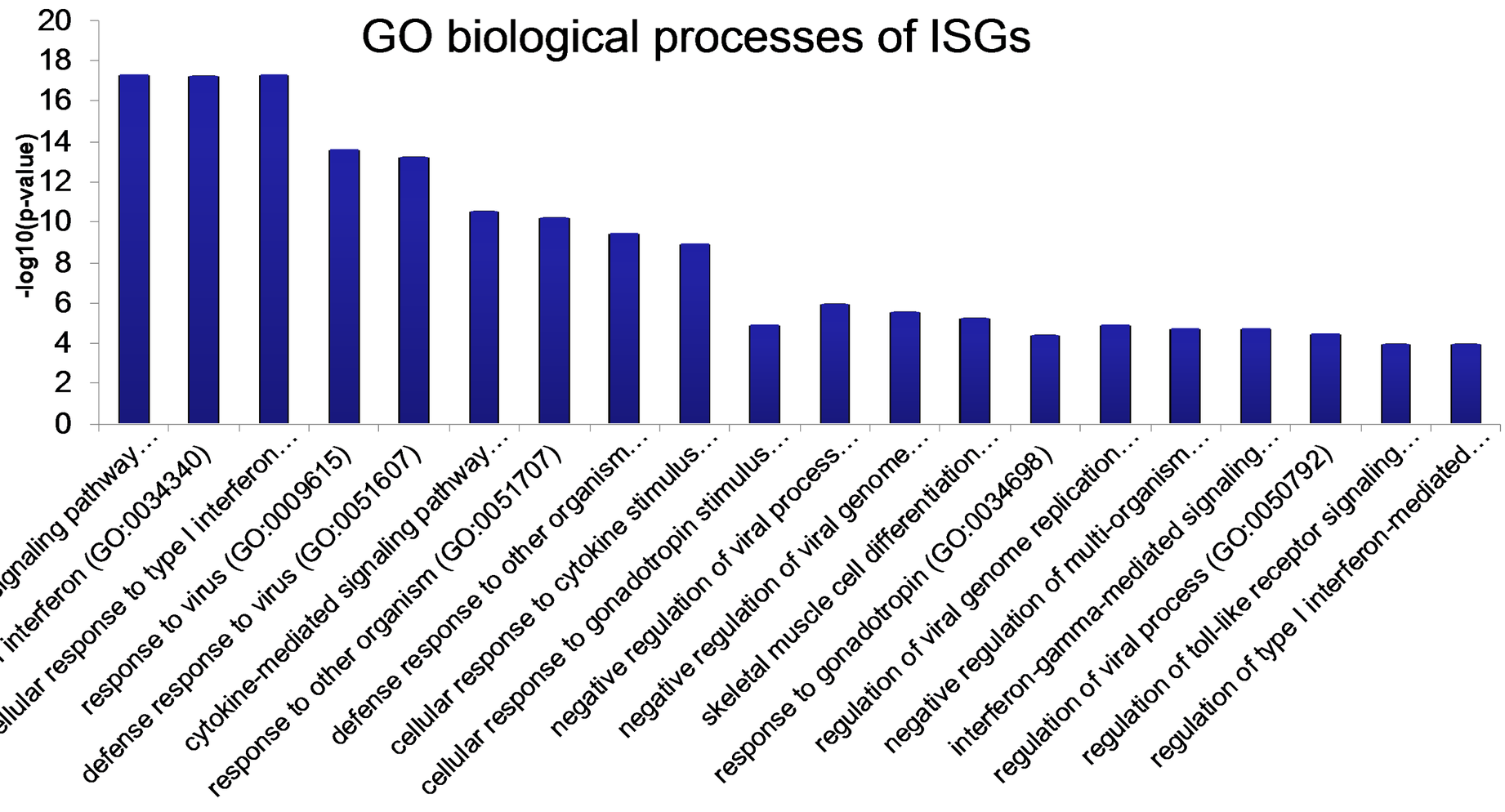
(A) Binned H3.3 ChIP-seq peaks are depicted in the heat map format. Genes (n= number of genes in each bin) in each bin were ranked according to H3.3 signal intensity. (B-D) Heat map of H3.3 peaks shown for all genes, constitutively expressed genes and silent genes in untreated (NT) and IFN treated cells.

Supporting Information 5

A



B

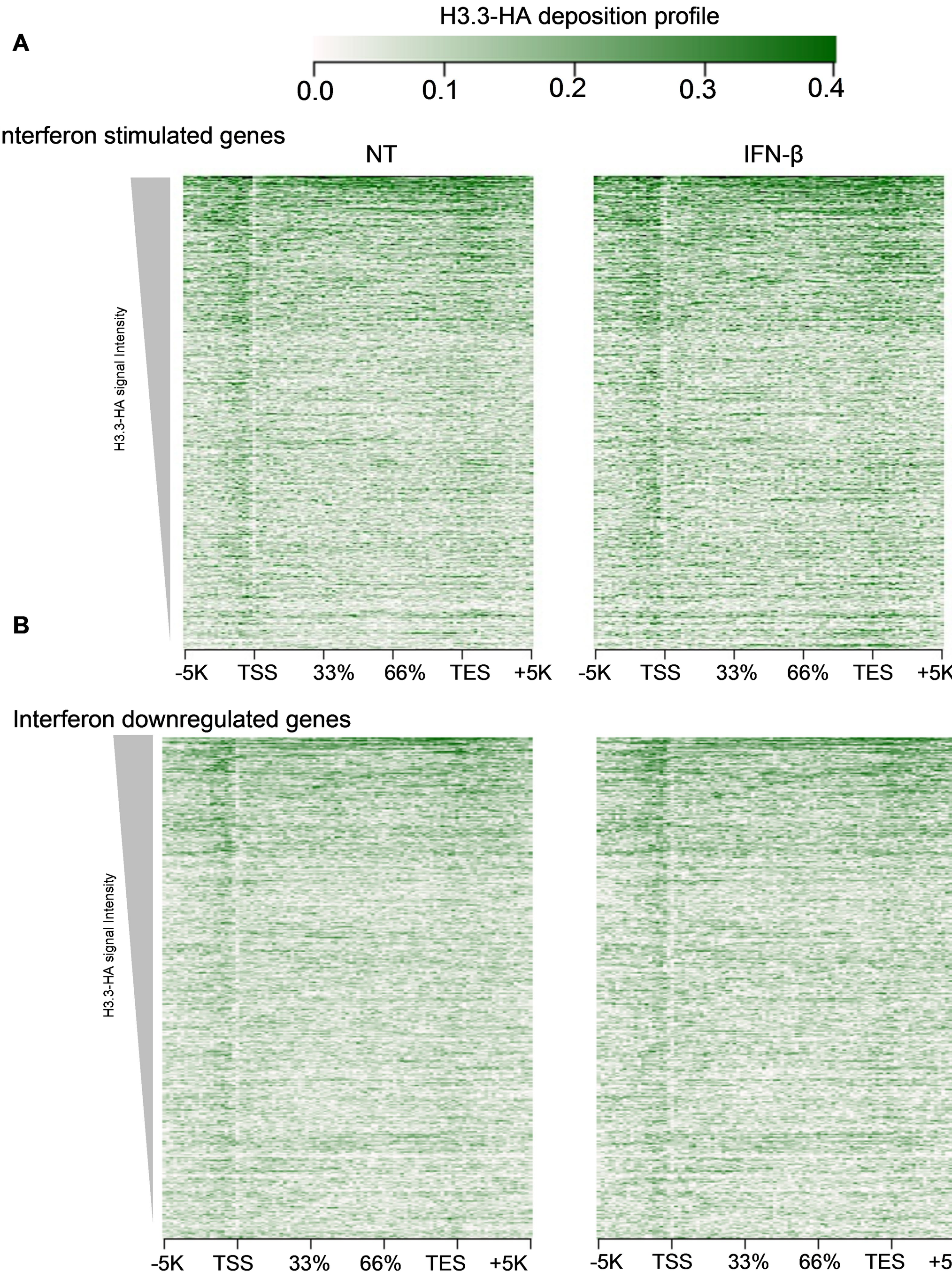


Supporting Information 5

IGV screenshots of *H3f3a* and *H3f3b* mRNA expression

(A) Normalized read coverage profiles of H3f3a and H3f3b transcripts in untreated (NT) and IFN treated MEFs. (B) GO analysis was performed for ISGs identified in *H3f3b* MEFs (Enrichr program, (http://amp.pharm.mssm.edu/Enrichr/). Y axis represent negative log of p-values. Enrichment in GO-terms related to type 1 IFN signaling is evident.

Supporting Information 6



Supporting Information 6

H3.3-HA peaks in all ISGs and genes downregulated by IFN treatment are shown in the heat map formant. Genes belonging to either group were aligned according to the H3.3 signal intensity. See the legend for Supporting information 4.