# Supporting Information

## Jenner et al. 10.1073/pnas.0909357106

### SI Methods

**Purification of Naïve Human CD4$^+$ T Cells and Differentiation into Th1 and Th2 Cells.** Human CD4$^+$ T cells were isolated from healthy donors as previously described (1) and were subjected to further rounds of purification by depletion of CD25+, CD45RO+, and HLA Class II (DR)+ cells by negative immunomagnetic selection (2). Naïve CD4$^+$ purified cells were routinely >95% CD4+CD45RA+CD45RO-DR-CD62L+.

Naïve T cells were then activated for 48–72 h by plate bound anti-CD3 and anti-CD28 monoclonal antibodies (both at 2 $\mu$g/mL, BD PharMingen) in the presence of rhIL-12 (10 ng/mL, R&D Systems) and anti-IL-4 (10 $\mu$g/mL, BD PharMingen) for Th1 polarization or rhIL-4 (10 ng/mL, R&D Systems) and anti-IFN-$\gamma$ (10 $\mu$g/mL, BD PharMingen) for Th2 polarization. rhIL-2 (National Cancer Institute, 100 U/mL) was added at day 3 and the cells were expanded for 10–12 days in the presence of the same polarizing cytokines.

Cells were harvested during primary stimulation for ChIP (see below), RNA analysis (TRIzol, Invitrogen) or Western blotting (RIPA buffer).

Secondary and tertiary stimulated Th1 and Th2 cell cultures were obtained by washing the primary cultures at day 7 and reculturing on plates coated with anti-CD3 and anti-CD28 for 24 h in the presence of the same skewing cytokines and antibodies as above. These cells were harvested in the same way as the primary cultures.

**Purification, Differentiation, and Retroviral Transduction of ïNaïve Murine CD4$^+$ T Cells.** BALB/c T-bet$^{-/-}$, C57BL/6 T-bet$^{-/-}$, DO11.10, DO11.10 × T-bet$^{-/-}$, and T-bet$^{-/-}$× IFN-$\gamma^{-/-}$ mice, were generated and used as indicated in the text and as previously described (2). All animals were maintained in SPF conditions in our animal facility in accordance with our institutional policies.

Naïve murine CD4$^+$ T cells were purified as previous described (2). Briefly lymph node cells from 4- to 6-week-old Balb/C mice were purified by fluorescence activated cell sorting into CD4+CD62L+CD25- T cells and activated with plate bound anti-CD3 and anti-CD28 (both 2 $\mu$g/mL, BD PharMingen) for 72 h in the presence of rmIL-12 (10 ng/mL, R&D Systems) and anti-IL-4 antibody (10 $\mu$g/mL, BD PharMingen) for Th1 skewing conditions. Further expansion was performed with the addition of fresh medium and rhIL2 (National Cancer Institute).

We used an ecotropic retrovirus to express GFP or T-bet (3). Retroviral transduction was performed at 36–48 h after activation in the presence of 4 $\mu$g/mL polybrene, as previously described (3, 4).

**Flow Cytometry.** Flow cytometry was performed as described (2). Cells were analyzed on a FACsCalibur (Becton Dickinson) using Cellquest software (Becton Dickinson) and further analyzed using FlowJo software (TreeStar, Inc). For the retroviral transduction experiments, live GFP+CD4$^+$ cells were gated for subsequent cytokine analysis.

**Antibodies for ChIP.** T-bet-bound genomic DNA was isolated from whole cell lysate using a rabbit polyclonal antibody (3). In separate experiments, we also used a rabbit IgG antibody (Upstate Biotechnology) as a control. GATA-3-bound DNA was isolated using a goat polyclonal IgG (D-16) antibody from Santa Cruz. Replicate GATA-3 ChIPs were performed with HG3–31 (Santa Cruz), a mouse monoclonal.

**Chromatin Immunoprecipitation.** We followed previously published protocols for ChIP-Chip (5–7).

We performed independent immunoprecipitations for each array analysis. Th1 and Th2 cells, pooled from multiple donors, were chemically crosslinked by the addition of one-tenth volume of fresh 11% formaldehyde solution for 20 min at room temperature. Cells were rinsed twice with 1× PBS and harvested using a silicon scraper and flash frozen in liquid nitrogen. Cells were stored at −80 °C before use.

For our replicate experiments, T-bet was immunoprecipitated from independent pools of donors with the antibody 9856 (biological replicates), as was GATA-3 using the D-16 and HG3–31 antibodies.

Cells were lysed in lysis buffers and sonicated to solubilize and shear crosslinked DNA. Sonication conditions vary depending on cells, culture conditions, crosslinking and equipment. We used a Misonix Sonicator 3000 and sonicated at 24 W for 10 × 30-s pulses (1 min pause between pulses). Samples were kept on ice at all times.

The resulting whole cell extract was incubated overnight at 4 °C with 100 $\mu$L Dynal Protein G magnetic beads that had been preincubated with approximately 10 $\mu$g appropriate antibody. For the T-bet antibody (9856), 10 $\mu$L supplied solution was used per reaction. The immunoprecipitation was allowed to proceed overnight.

Beads were washed five times with RIPA buffer and once with TE containing 50 mM NaCl. Bound complexes were eluted from the beads by heating at 65 °C with occasional vortexing and crosslinking was reversed by overnight incubation at 65 °C. Whole cell extract DNA (reserved from the sonication step) was also treated for crosslink reversal.

Immunoprecipitated DNA and whole cell extract DNA were then purified by treatment with RNase A, proteinase K, and multiple phenol:chloroform:isoamyl alcohol extractions. Purified DNA was blunted and ligated to linker and amplified using a two-stage PCR protocol. Amplified DNA was labeled and purified using Bioprime random primer labeling kits (Invitrogen, immunoenriched DNA was labeled with Cy5 fluorophore and whole cell extract DNA was labeled with Cy3 fluorophore).

Labeled DNA was mixed with Herring sperm DNA, yeast tRNA and Cot-1 human DNA and hybridized to Agilent oligonucleotide microarrays in Agilent hybridization chambers for 40 h at 40 °C [see (7) for details]. Arrays were then washed and scanned using an Agilent DNA microarray scanner BA and the data extracted using Agilent Feature Extractor.

**Array Design: Agilent Human 2 × 244 K Set.** We designed a two-slide set to cover 8 kb (approximately 4 kb upstream and 4 kb downstream) around the transcription start site of 18,450 RefSeq-annotated human genes (8). Probes were designed against build 35 of the human genome sequence according to previously published criteria (6). Each array contains approximately 244,000 60-mer oligonucleotide probes. Oligonucleotide probes were spaced approximately 250 bp along the genome, on average.

**Data Normalization and Analysis.** After scanning, data were extracted from the image file using Feature Extractor. We used the default CGH protocol with the following modifications: the

background was set to the average intensity of all negative control spots, rank consistent probes was used only to calculate the normalization factor (linear normalization) and spacial detrending was turned off. We then calculated the log of the ratio of intensity in the IP-enriched channel to intensity in the genomic DNA channel for each probe and used a whole chip error model (5, 6) to calculate confidence values for each spot on each array [single probe P value; see (6) for details].

***Identification of bound regions.*** To automatically determine bound regions in the datasets, we used an algorithm that incorporates information from neighboring probes (6). Briefly, for each 60-mer oligonucleotide, we calculated the average X score of the 60-mer and its two immediate neighbors. If a feature was flagged as abnormal during scanning, we assumed it gave a neutral contribution to the average X score. Similarly, if an adjacent feature was beyond a reasonable distance from the probe (1,000 bp), we assumed it gave a neutral contribution to the average X score. This set of averaged values gave us a distribution that was subsequently used to calculate *P* values of average X (probe set *P* values). If the probe set *P* value was less than 0.001, the three probes were marked as potentially bound. As most probes were spaced within the resolution limit of chromatin immunoprecipitation, we next required that multiple probes in the probe set provide evidence of a binding event. Candidate bound probe sets were required to pass one of two additional filters: two of the three probes in a probe set must each have single probe *P* values $<0.005$ or the center probe in the probe set has a single probe *P* value $<0.001$ and one of the flanking probes has a single point *P* value $<0.1$. Individual probe sets that passed these criteria were collapsed into bound regions if the center probes of the probe sets were within 1,000 bp of each other.

***Comparing bound regions to known genes.*** The locations of bound regions were compared relative to transcript start coordinates of known genes compiled from RefSeq, Mammalian Gene Collection (MGC) and Ensembl. All coordinate information was downloaded in January 2005 from the UCSC Genome Browser (NCBI build 35). Bound regions were assigned to a particular gene if it lay within 1 kb of the annotated transcription start site. We identified 1,009 genes bound by T-bet in Th1 cells in replicate experiments, of which we removed 177 because these were also identified in our IgG control experiment, leaving 832 specific targets. Of these 832 genes bound by T-bet in Th1 cells, 52 were also called in Th2 cells. For GATA-3, 411 genes were identified as bound in Th2 cells using the D-16 antibody, of which 67 were also called by our IgG experiment, leaving 344 specific targets. In Th1 cells, 1,416 genes were identified as bound using the D-16 antibody with a probe set *P* value below $10^{-3,}$ of which 199 were also called by our IgG experiment, leaving 1,217 specific targets.

***Verification of GATA-3 target genes using HG3–31 antibody.*** We performed replicate GATA-3 ChIP with another antibody, HG3–31. Examination of the enrichment levels achieved with this antibody shows that is not as efficient as D-16 in isolating GATA-3 bound DNA (Fig. S4). Of the 344 genes which we identified as GATA-3 targets in Th2 cells with the D-16 antibody, 122 were also identified using HG3–31. These therefore represent very high confidence targets. We believe that our failure to detect GATA-3 at the remaining genes is due to the poor performance of this antibody compared with D-16. Of the 1,217 genes that we identified as GATA-3 targets in Th1 cells, 383 were also identified with HG3–31.

**Overlap of T-bet and GATA-3 target genes.** Using the GATA-3 antibody D-16, we identified GATA-3 association with 39% of T-bet target genes in Th1 cells ($P < 10^{-16}$, hypergeometric; Fig. 4*A* and Table S1). Using the HG3–31 antibody, we identified GATA-3 association with 18% of T-bet target genes ($P < 10^{-16}$, hypergeometric). In Th2 cells, GATA-3 associated with 9% of T-bet target genes (T-bet associated with 22% of GATA-3 target genes, $P < 10^{-16}$, hypergeometric).

***Quantitative PCR of ChIP-Enriched DNA.*** We sought to verify enrichment of T-bet and GATA-3 target gene DNA in the ChIP fractions compared to whole cell extract DNA using quantitative real-time PCR (Applied Biosystems Prism 7000 Sequence Detection System). Using Primer Express (Applied Biosystems), we designed primers to amplify approximately 55–70 bp within the 500 bp upstream of the transcription start site for selected target genes; primer sequences available on request. Gene-specific PCR was performed upon LM-PCR amplified samples (6) and the amount of each sequence in 2 ng ChIP sample quantified relative to a dilution series of WCE sample using Sybrgreen (Applied Biosystems).

**Gene Expression Analysis.** ***RNA purification and labeling.*** RNA was harvested from primary and secondary stimulated Th1 and Th2 cells by precipitation after TRIzol extraction (Invitrogen) and further purification with RNeasy columns (Qiagen). For each experimental sample, RNA quality was assessed by RNA Nano LabChip analysis on an Agilent Bioanalyzer 2100. Under standard conditions, processing of RNAs for GeneChip Analysis was in accordance with methods described in the Affymetrix GeneChip Expression Analysis Technical Manual (rev. 1.3), as subsequently detailed. Synthesis of cDNA first and second strand was performed using the GeneChip Expression 3′-Amplification Reagents One-Cycle cDNA Synthesis Kit (P/N 900431). Cleanup of the double stranded product was carried out according to standard Affymetrix protocols using the Affymetrix GeneChip Cleanup Module (Affymetrix Catalog # 900371).

In vitro transcription (IVT) was performed using the GeneChip Expression Amplification Reagents kit-30 reactions (P/N 900449) and was carried out according to the standard Affymetrix protocols. Quantification of the IVT samples was carried out on a Bio-Tek UV Plate Reader.

***Hybridization to affymetrix arrays.*** Hybridization to Affymetrix human genome U133 plus 2.0 arrays or mouse genome 430 2.0 arrays was carried out according to the Affymetrix GeneChip Manual (rev. 1.3). Twenty micrograms of IVT material was the nominal amount used on the GeneChip arrays. Affymetrix hybridization ovens were used to incubate the arrays at a constant temperature of 45 °C overnight.

Preparation of microarrays for scanning was carried out with appropriate Affymetrix wash protocols matched to the specific chip type on a Model 450 Fluidics station. Affymetrix GeneChip Operating Software (GCOS) operating system controls the Fluidics station process. Scanning was carried out on an Affymetrix Model 3000 scanner with autoloader and the data handled in GCOS.

***Analysis for relative expression levels.*** We first generated median centered log2 expression ratios (Th1 primary stimulated/Th2 primary stimulated and Th1 secondary stimulated/Th2 secondary stimulated). We then filtered for probe sets with transcripts called "Present" in the cell type in which expression was calculated to increase. Data from multiple probe sets representing a single gene were then averaged to give one value per gene.

***Analysis for absent/present calls.*** Genes for which any probe set was called "Present" by the Affymetrix GCOS software was classified as expressed in that cell type.

***Comparing binding and relative expression data.*** Gene expression was measured in sorted $CD4^+$ cells from T-$bet^{-/-}$/Ifn$\gamma^{-/-}$ mice after infection with retroviral vectors expressing T-bet or GFP and data from triplicate arrays averaged. Mouse gene expression data were mapped onto human expression and ChIP-Chip data using Homologene. The fraction of genes overexpressed in human Th1 cells vs. Th2 cells (1- to 5-fold) and bound by T-bet that were upregulated by 1.5-fold or more in mouse $CD4^+$ T-cells expressing T-bet versus GFP were compared against the number

expected by chance using the Binomial distribution. The number expected by chance was derived from all T-bet target genes over expressed in Th1 cells for which we had mouse expression data.

**Gene Ontology Analysis.** We used GO (9), as implemented in DAVID (10), to find biological processes enriched in the sets of genes bound by T-bet and GATA-3. We used the set of all genes on the arrays as the background list to calculate the significance of enrichment. There were no significant differences in the enriched GO categories returned for the set of GATA-3 targets identified using the antibody D-16 alone compared to the set of targets identified with both D-16 and HG3–31.


1. Lord GM, et al. (1998) Leptin modulates the T-cell immune response and reverses starvation-induced immunosuppression. *Nature* 394:897–901.
2. Lord GM, et al. (2005) T-bet is required for optimal proinflammatory CD4+ T-cell trafficking. *Blood* 106:3432–3439.
3. Szabo SJ, et al. (2000) A novel transcription factor, T-bet, directs Th1 lineage commitment. *Cell* 100:655–669.
4. Hwang ES, Szabo SJ, Schwartzberg PL, Glimcher LH (2005) T helper cell fate specified by kinase-mediated interaction of T-bet with GATA-3. *Science* 307:430–433.
5. Odom DT, et al. (2004) Control of pancreas and liver gene expression by HNF transcription factors. *Science* 303:1378–1381.
6. Lee TI, et al. (2006) Control of developmental regulators by Polycomb in human embryonic stem cells. *Cell* 125:301–313.
7. Lee TI, Johnstone SE, Young RA (2006) Chromatin immunoprecipitation and microarray-based analysis of protein location. *Nat Protoc* 1:729–748.
8. Pruitt KD, Tatusova T, Maglott DR (2005) NCBI Reference Sequence (RefSeq): A curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res* 33:D501–504.
9. Ashburner M, et al. (2000) Gene ontology: Tool for the unification of biology. The Gene Ontology Consortium. *Nat Gen* 25:25–29.
10. Dennis G, Jr, et al. (2003) DAVID: Database for annotation, visualization, and integrated discovery. *Genome Biol* 4:P3.

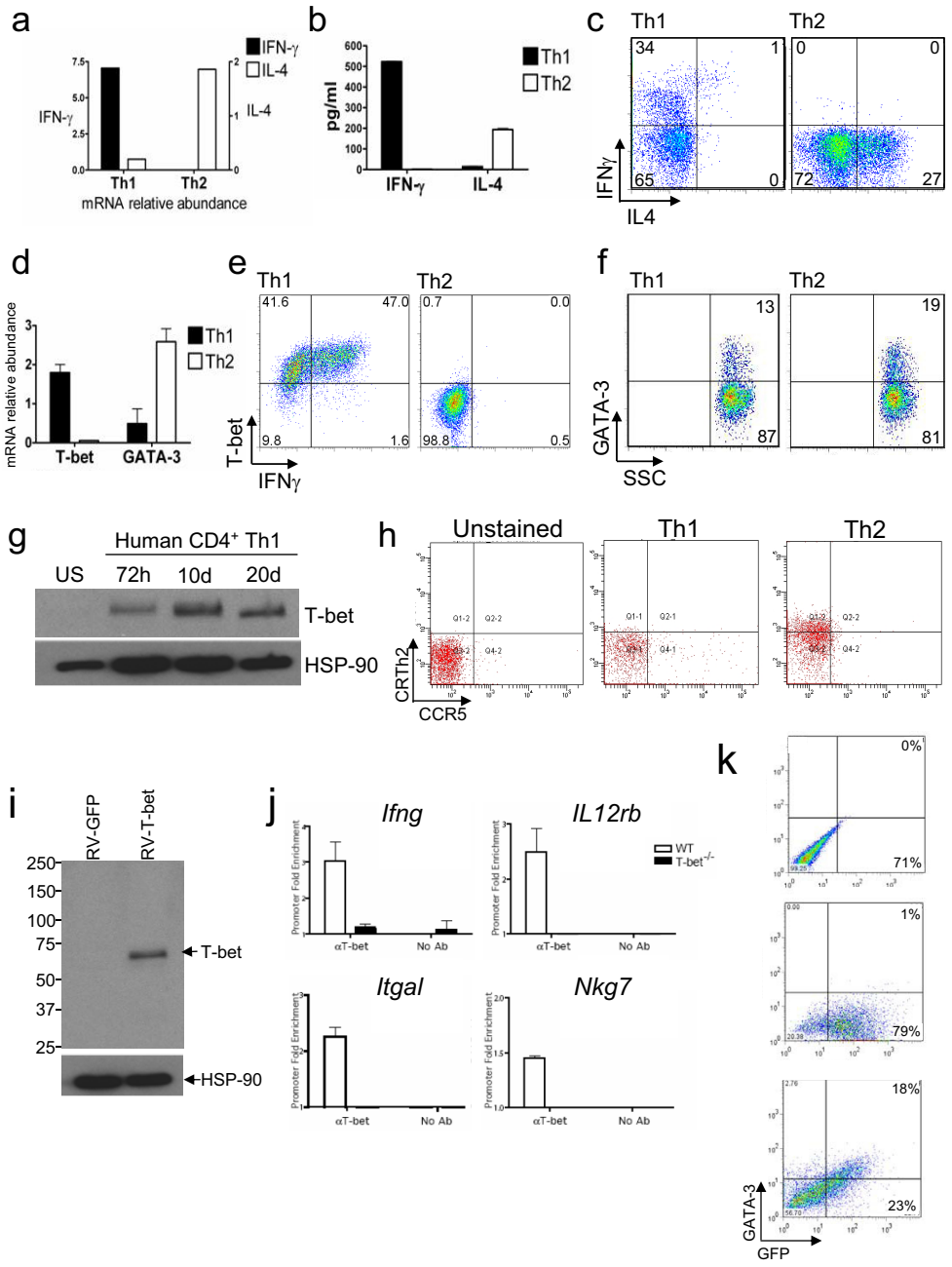


**Fig. S1.** T-cell polarization and antibody specificity. (*a* and *b*) IL-4 and IFN-γ expression measured by quantitative PCR (relative to beta-actin mRNA) (*a*) and ELISA (b). (*c*) IL-4 (*x*-axis) and IFN-γ (*y*-axis) expression in primary human CD4$^+$ Th1 cells (*Left*) and Th2 cells (*Right*) measured by flow cytometry. The percentage of cells in each quadrant is indicated. Lines dividing quadrants were positioned according to the experiments with isotype controls. (*d*) T-bet and GATA-3 expression measured by quantiative PCR (relative to beta-actin mRNA). (*e*) IFN-γ (*x*-axis) and T-bet (*y*-axis) expression in primary human CD4$^+$ Th1 cells (*Left*) and Th2 cells (*Right*) measured by flow cytometry. (*f*) GATA-3 expression (*y*-axis) in primary human CD4$^+$ Th1 cells (*Left*) and Th2 cells (*Right*). (*g*) Expression of T-bet protein during human Th1 differentiation. Western blot of T-bet protein expression during differentiation of naïve CD4 T-cells. A blot for HSP-90 is shown below as a control. (*h*) Surface staining for CCR5 (*x* axis, labeled with PE) and CRTh2 (*y* axis, labeled with APC) in Th1 and Th2 cells, measured by flow cytometry. (*i*) Western blot of T-bet protein expression in 293 cells infected with retroviral vectors encoding GFP or T-bet. A blot for HSP-90 is shown below. (*j*) Fold enrichment of DNA promoter sequences in T-bet ChIP material from Th1 cells derived from wild-type mice and from T-bet-deficient mice relative to whole genomic DNA measured by quantitative PCR. Error bars show standard deviations ($n = 3$). (*k*) GFP expression (*x* axis) and immunocytochemical staining for GATA-3 (PE, *y* axis) using the antibody HG3–31 in 293 cells, 293 cells infected with a lentiviral vector encoding GFP, and 293 cells infected with a vector encoding GATA-3 and GFP.

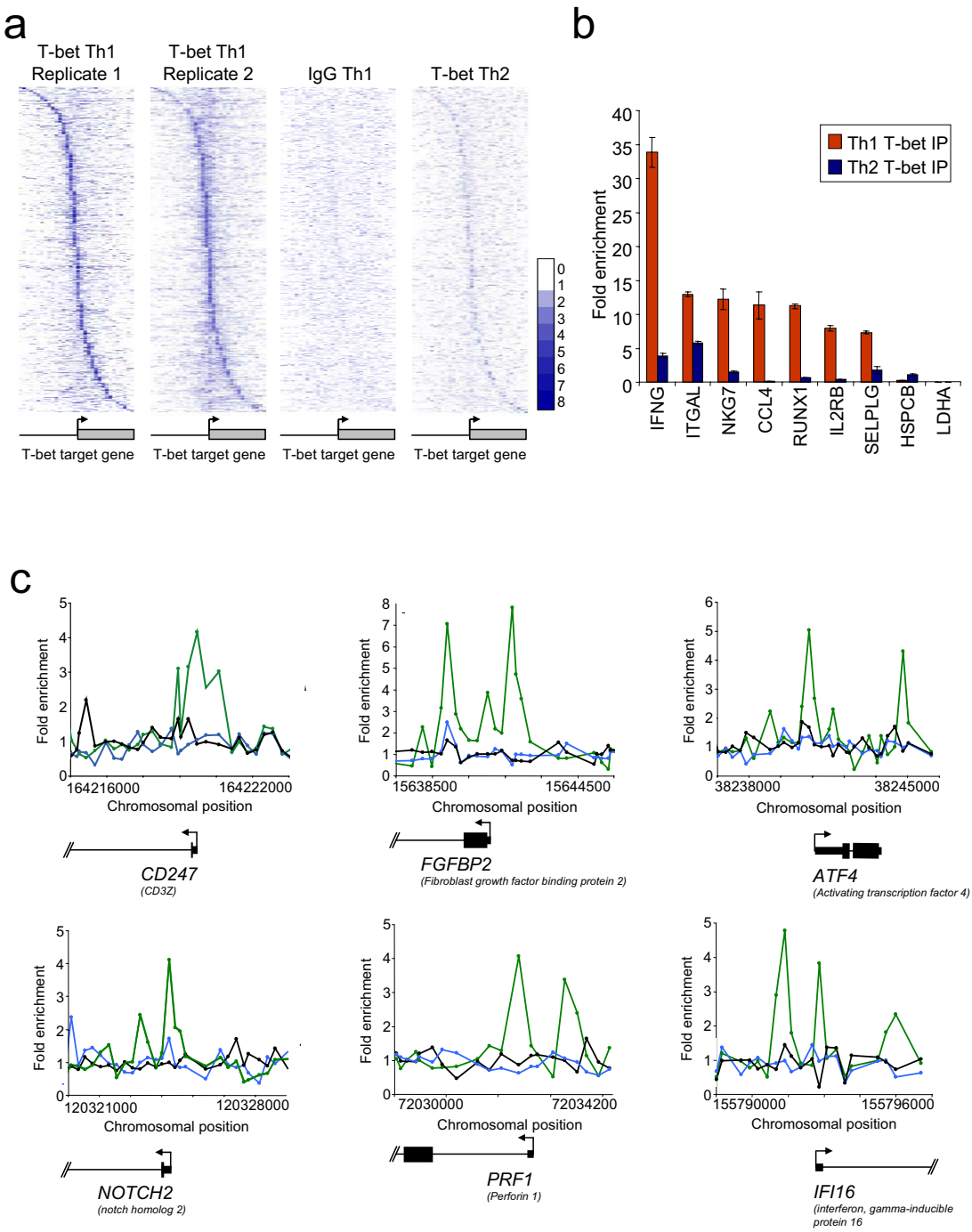


**Fig. S2.** T-bet binding in Th1 cells. (*a*) Heat maps showing enrichment of T-bet at its target genes in Th1 cells. The figure shows the results from two replicate experiments from independent batches of Th1 cells (derived from different donors) and an IgG control. Each row represents one gene considered to be occupied by T-bet according to our error model (the same genes are shown in each of the three panels) and each column represents the data from one oligonucleotide probe. Oligos are ordered by their position relative to the transcription start site, as shown by the diagram below. A scale for the enrichment ratios is shown on the right. (*b*) Verification of T-bet ChIP-Chip results using quantitative PCR. Fold enrichment of DNA promoter sequences in T-bet ChIP material relative to whole genomic DNA measured by quantitative PCR. Error bars, standard deviations ($n = 3$). (*c*) Examples of T-bet ChIP signals at protein-coding genes. The plots show unprocessed enrichment ratios for all probes within a genomic region (ChIP vs. whole genomic DNA) for T-bet (green) in Th1 cells, T-bet in Th2 cells and an IgG control (black) in Th1 cells. Chromosomal positions are from NCBI build 35 of the human genome. Genes are shown to scale below plots (exons are represented by vertical bars).

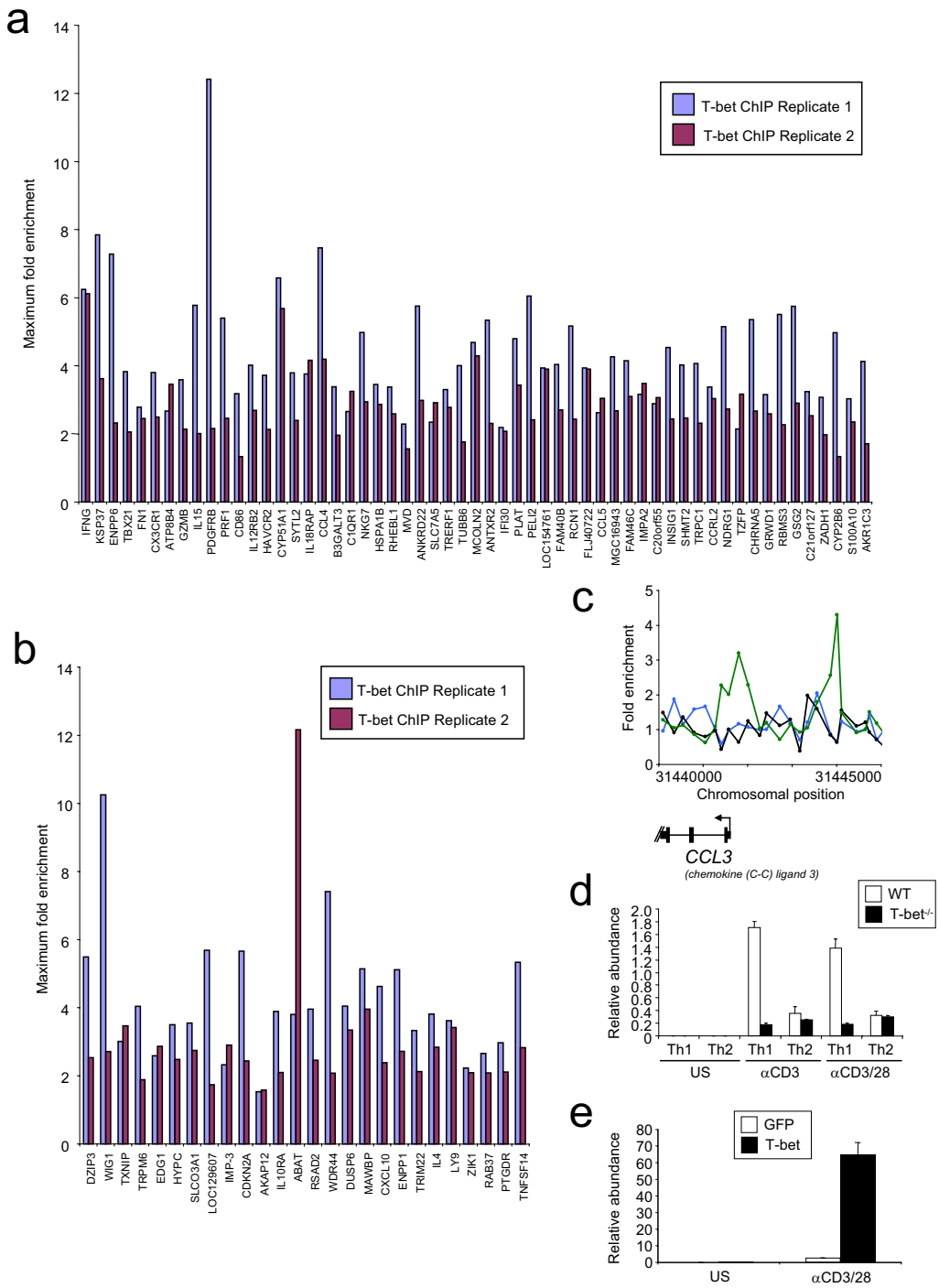


**Fig. S3.** Genes differentially expressed between Th1 and Th2 cells that are targeted by T-bet. Fold enrichment of T-bet ChIP DNA versus input DNA at genes that are overexpressed (*a*) and underexpressed (*b*) in Th1 cells versus Th2 cells by an average of at least 2-fold in replicate experiments. (*C*) T-bet ChIP signals at *CCL3* in human Th1 cells (green) and Th2 cells (blue) and signal from an IgG control IP in Th1 cells (black). Details as for Fig. 1*B*. (*D*) Real time PCR for *CCL3* in WT (open bars) and T-bet$^{-/-}$ (filled bars) murine T cells. Cells were unstimulated (u/s), stimulated with anti-CD3 antibodies (αCD3), or stimulated with anti-CD3 and anti-CD28 antibodies (αCD3/28). RNA abundance is relative β-actin mRNA. (*E*) Real time PCR for *CCL3* in CD4$^{+}$ T cells from T-bet$^{-/-}$ × IFN-γ$^{-/-}$ murine CD4$^{+}$ T cells transduced with empty vector (open bars) or T-bet (filled bars) expressing retrovirus.

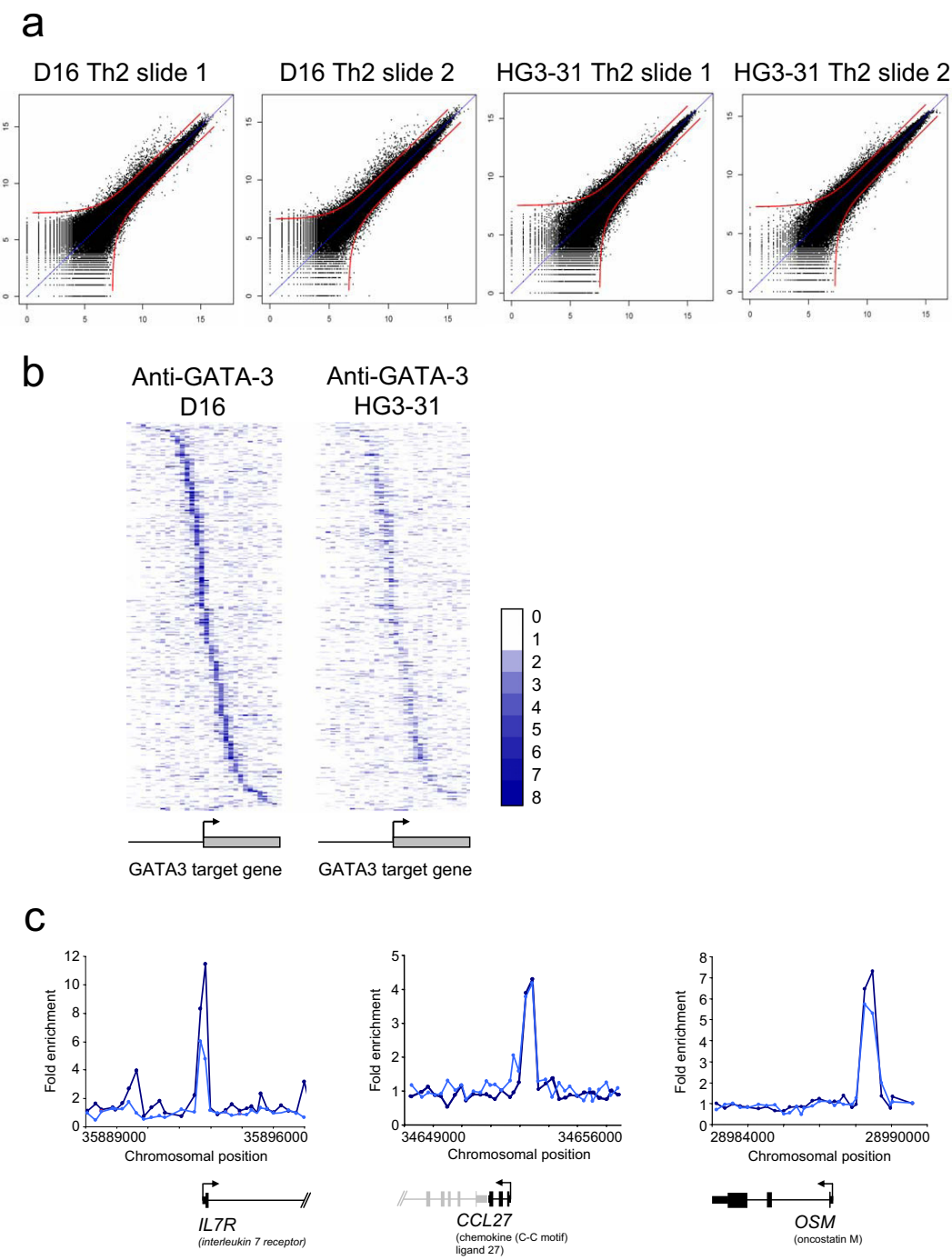


**Fig. S4.** GATA-3 binding in Th2 cells. (*a*) Scatter plots showing enrichment of genomic DNA sequences by ChIP with GATA-3 antibodies D16 and HG3–31. The *x* axis shows the hybridization signal intensity for input whole cell extract (WCE) DNA (log10 Cy3 background-subtracted signal). The *y* axis shows the hybridization signal intensity for ChIP-enriched DNA, using the antibodies named above each panel. Oligonucleotide probes tiling approximately −4 kb to + 4 kb around 18,450 human genes are split over two 244,000 element arrays (slide 1 and slide 2). (*b*) Heat maps showing the results from two replicate GATA-3 ChIP experiments from independent batches of Th2 cells with two different antibodies, D-16 and HG3–31. Each row represents one gene considered to be occupied by GATA-3 from our D-16 ChIP data according to our error model and each column represents the data from one oligonucleotide probe. Oligos are ordered by their position relative to the transcription start site, as shown by the diagram below. A scale for the enrichment ratios is shown on the right. (*c*) Examples of GATA-3 ChIP signals in Th2 cells. The plots show unprocessed enrichment ratios for all probes within a genomic region for three GATA-3 targets (ChIP vs. whole genomic DNA). D-16 ChIP data in dark blue, HG3–31 ChIP data in light blue. Chromosomal positions are from NCBI build 35 of the human genome. Genes are shown to scale below plots (exons are represented by vertical bars).

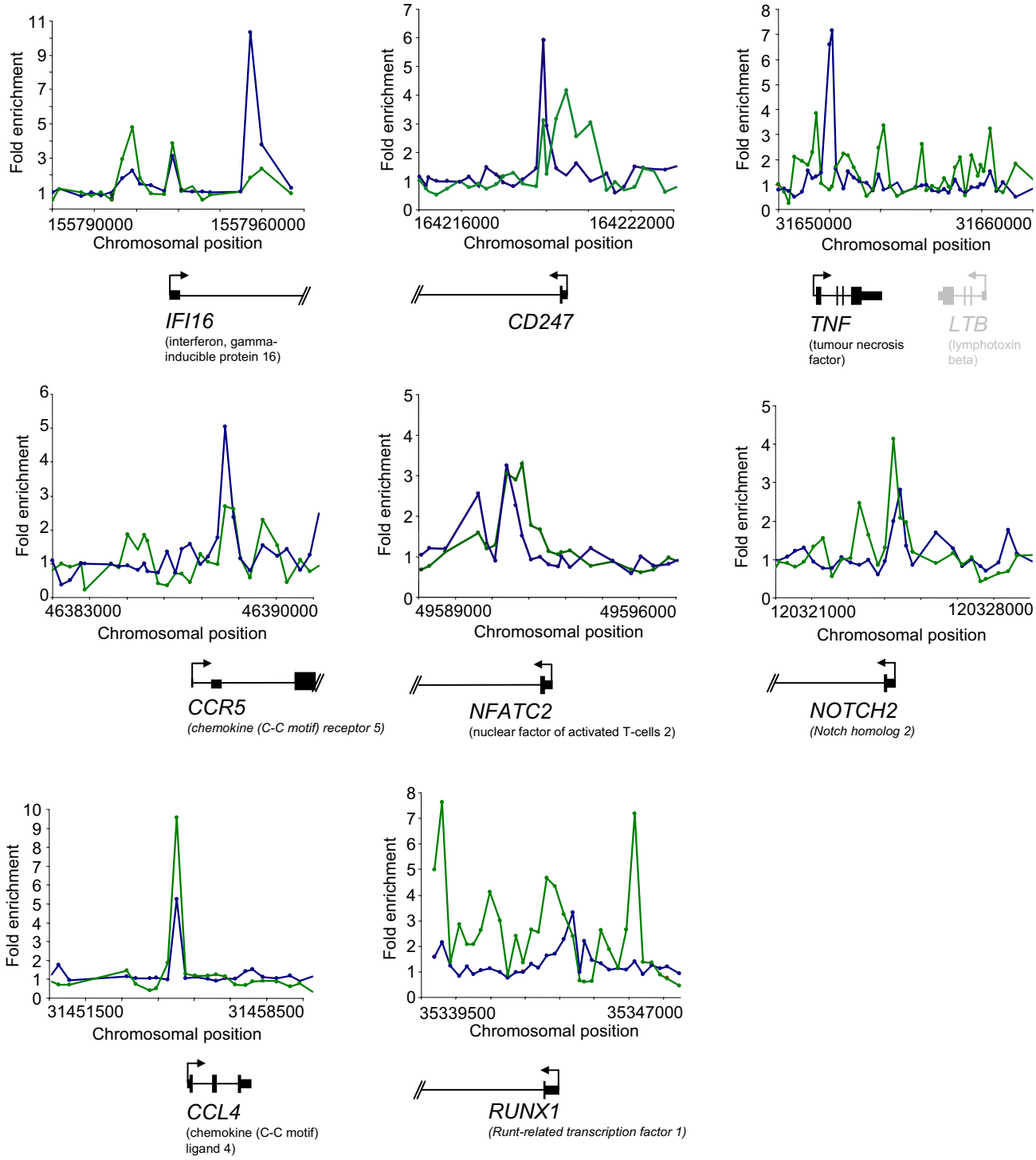


**Fig. S5.** Examples of genes targeted by T-bet in Th1 cells and GATA-3 in Th2 cells. The plots show unprocessed enrichment ratios for all probes within a genomic region (ChIP vs. whole genomic DNA) for T-bet in Th1 cells (green) and GATA-3 in Th2 cells (D16 anibody, blue). Chromosomal positions are from NCBI build 35 of the human genome. Genes are shown to scale below and aligned with the plots by chromosomal position (exons are represented by vertical bars, the start and direction of transcription by an arrow).

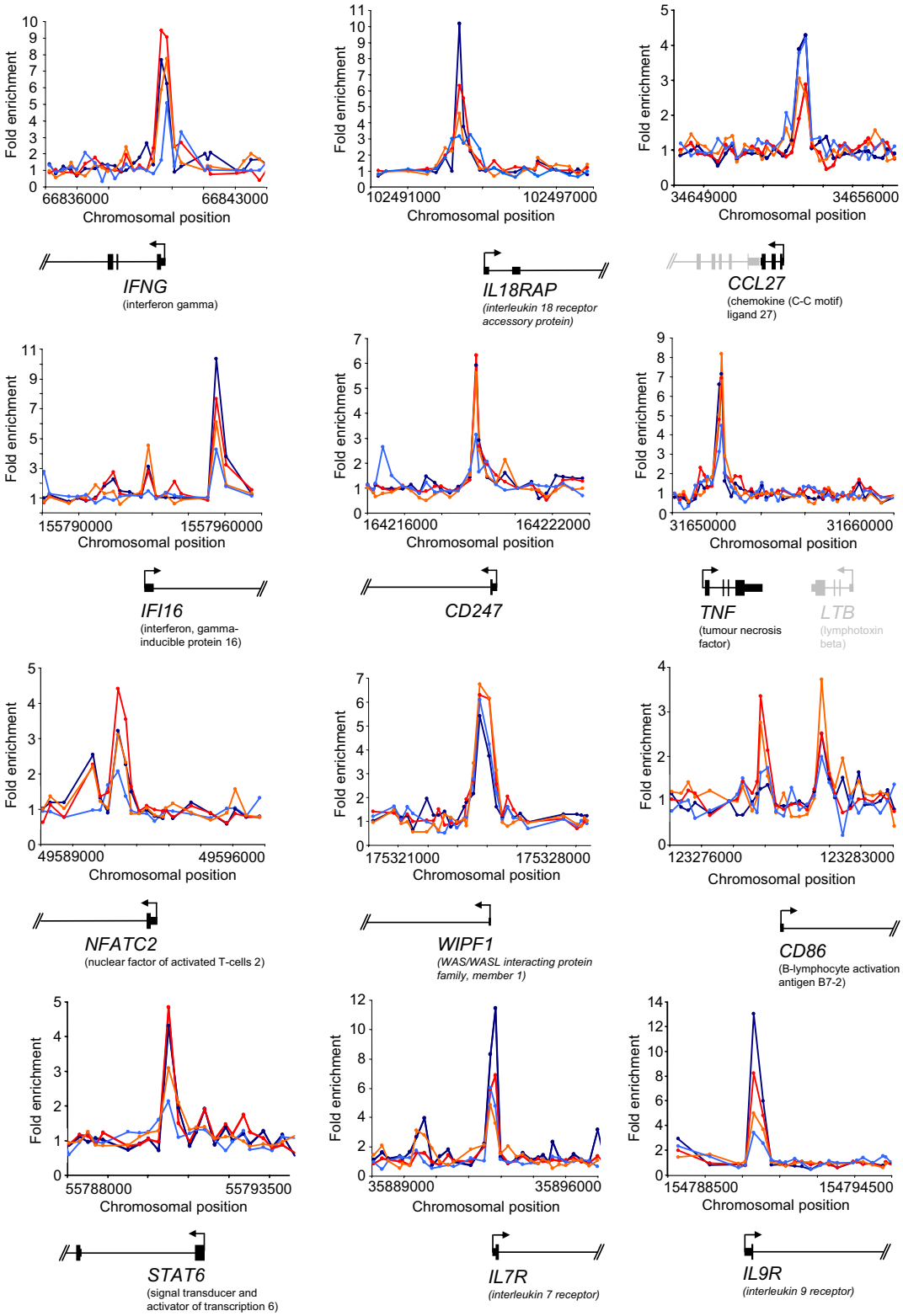


**Fig. S6.** Examples of GATA-3 ChIP signals in both Th2 and Th1 cells. The plots show unprocessed enrichment ratios for all probes within a genomic region (ChIP vs. whole genomic DNA) in Th2 cells (D16 ChIP data in dark blue, HG3–31 in light blue) and Th1 cells (D16 ChIP data in red, HG3–31 in orange). Chromosomal positions are from NCBI build 35 of the human genome. Genes are shown to scale below and aligned with the plots by chromosomal position (exons are represented by vertical bars, the start and direction of transcription by an arrow).

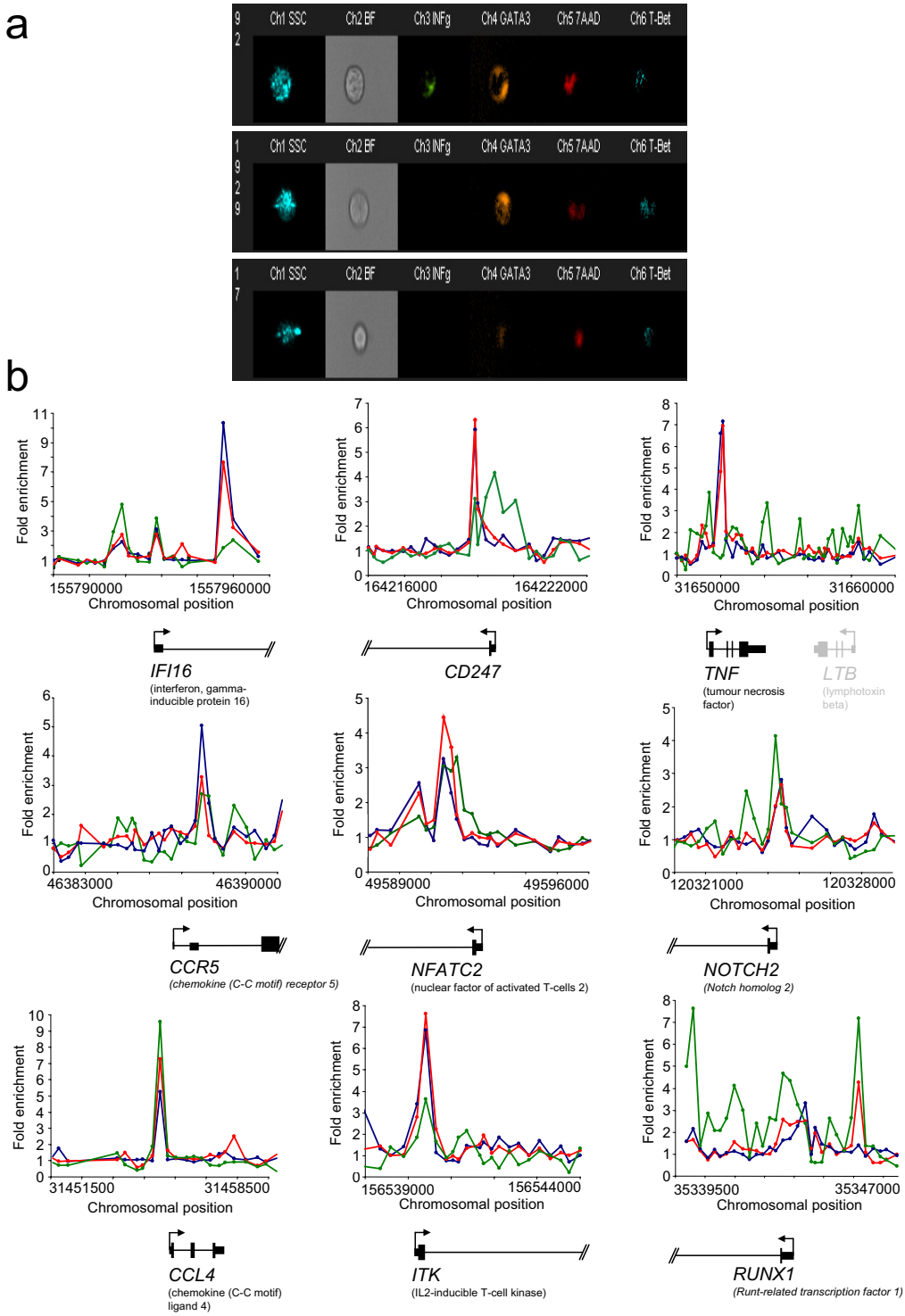


**Fig. S7.** Co-occupancy of genes by T-bet and GATA-3 in Th1 cells. (*a*) T-bet, GATA-3 and IFN-γ in human CD4$^+$ T cells. Image Stream analysis of human CD4$^+$ T cells at day 5 of polarisation. Channel (Ch) 1 – Side scatter, Ch2 – bright field, Ch3 – IFN-γ, Ch4 GATA-3, Ch5 - 7-Aminoactinomycin D, Ch6 - T-bet. (*b*) Examples of genes occupied by T-bet and GATA-3 in Th1 and Th2 cells. The plots show unprocessed enrichment ratios for all probes within a genomic region (ChIP vs. whole genomic DNA) for T-bet in Th1 cells (green), GATA-3 in Th1 cells (D16 ChIP, red) and GATA-3 in Th2 cells (D16 ChIP, blue). Chromosomal positions are from NCBI build 35 of the human genome. Genes are shown to scale below plots (exons are represented by vertical bars, the start and direction of transcription by an arrow).

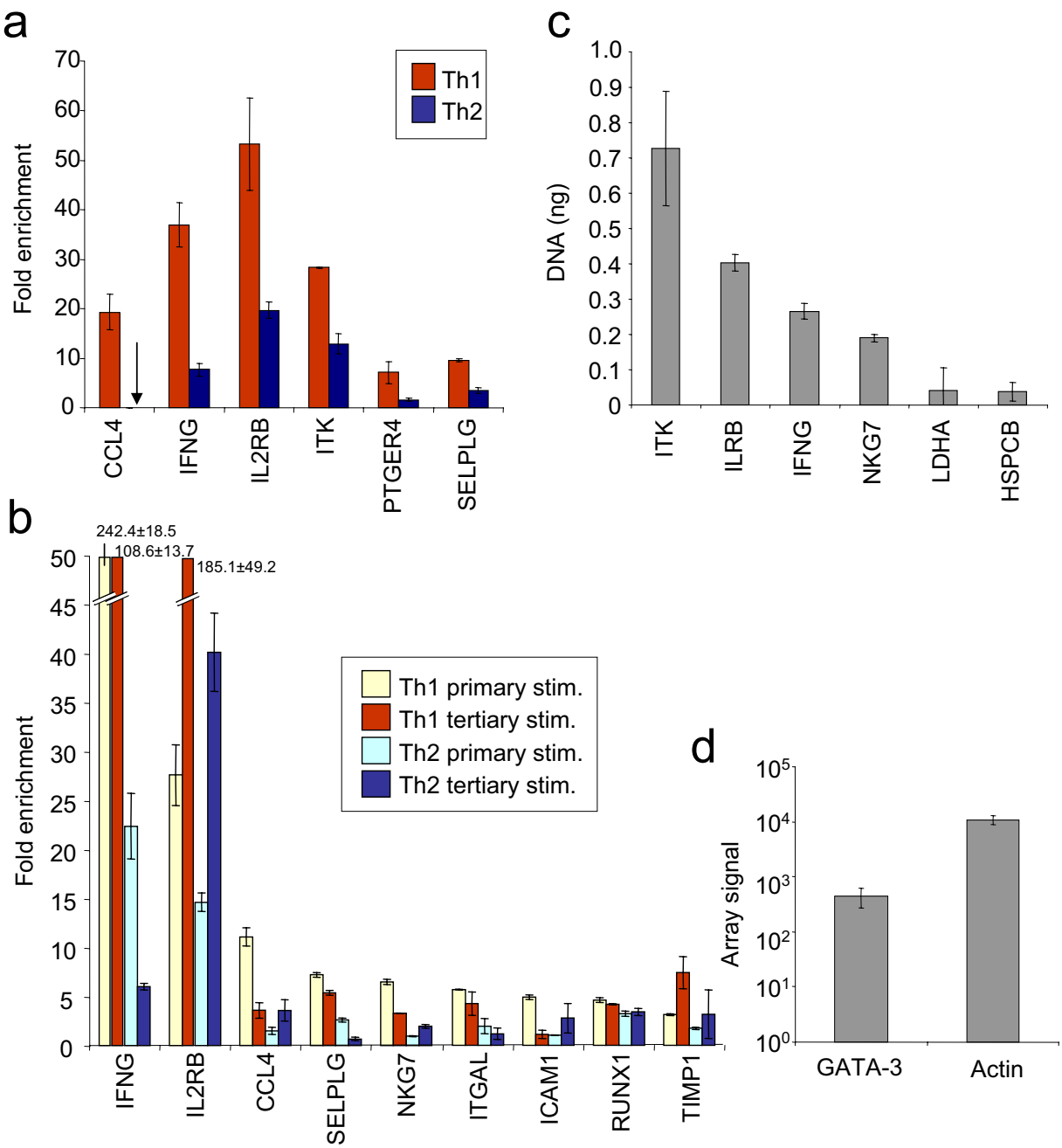


**Fig. S8.** Quantitative PCR of GATA-3 ChIP samples. (*a*) ChIP from primary stimulated Th1 and Th2 cells with HG3–31 antibody. Fold enrichment of DNA promoter sequences in amplified GATA-3 ChIP material relative to whole genomic DNA. Error bars show standard deviations ($n = 3$). (*b*) ChIP from primary and tertiary stimulated Th1 and Th2 cells with D-16 antibody. Details as for (*a*). (*c*) ChIP from a Th1 cell clone with the GATA-3 D-16 antibody. The *y* axis shows the amount of DNA reported to be present in the unamplified ChIP sample compared to a standard curve of whole genomic DNA. GATA-3 target genes ITK, IL12RB, IFNG, and NKG7 show an enrichment relative to control genes LDHA and HSPCB. (*d*) Detection of GATA-3 in murine $CD4^+$ T-cells before infection with retroviral vectors encoding T-bet or GFP.

## Other Supporting Information Files

Table S1