Genome-Wide Analysis Reveals Unique Regulation of Transcription of Th2-Specific Genes by GATA3

Shu Horiuchi, Atsushi Onodera, Hiroyuki Hosokawa, Yukiko Watanabe, Tomoaki Tanaka, Sumio Sugano, Yutaka Suzuki and Toshinori Nakayama

*J Immunol* published online 2 May 2011
http://www.jimmunol.org/content/early/2011/05/02/jimmunol.1100179

Supplementary Material
http://www.jimmunol.org/content/suppl/2011/05/02/jimmunol.1100179.DC1

Subscription
Information about subscribing to *The Journal of Immunology* is online at:
http://jimmunol.org/subscription

Permissions
Submit copyright permission requests at:
http://www.aai.org/About/Publications/JI/copyright.html

Email Alerts
Receive free email-alerts when new articles cite this article. Sign up at:
http://jimmunol.org/alerts
Genome-Wide Analysis Reveals Unique Regulation of Transcription of Th2-Specific Genes by GATA3

Shu Horiuchi,*,1 Atsushi Onodera,*,1 Hiroyuki Hosokawa,* Yukiko Watanabe,* Tomoaki Tanaka,† Sumio Sugano,‡ Yutaka Suzuki,‡ and Toshinori Nakayama*

Differentiation of naive CD4 T cells into Th2 cells is accompanied by chromatin remodeling and increased expression of a set of Th2-specific genes, including those encoding Th2 cytokines. IL-4-mediated STAT6 activation induces high levels of transcription of GATA3, a master regulator of Th2 cell differentiation, and enforced expression of GATA3 induces Th2 cytokine expression. However, it remains unclear whether the expression of other Th2-specific genes is induced directly by GATA3. A genome-wide unbiased chromatin immunoprecipitation assay coupled with massive parallel sequencing analysis revealed that GATA3 bound to 1279 genes selectivity in Th2 cells, and 101 genes in both Th1 and Th2 cells. Simultaneously, we identified 26 highly Th2-specific STAT6-dependent inducible genes by DNA microarray analysis-based three-step selection processes, and among them 17 genes showed GATA3 binding. We assessed dependency on GATA3 for the transcription of these 26 Th2-specific genes, and 10 genes showed increased transcription in a GATA3-dependent manner, whereas 16 genes showed no significant responses. The transcription of the 16 GATA3-nonresponding genes was clearly increased by the introduction of an active form of STAT6, STAT6VT. Therefore, although GATA3 has been recognized as a master regulator of Th2 cell differentiation, many Th2-specific genes are not regulated by GATA3 itself, but in collaboration with STAT6.

The Journal of Immunology, 2011, 186: 000–000.

*Department of Immunology, Graduate School of Medicine, Chiba University, Chiba 260-8670, Japan; †Department of Clinical Cell Biology, Graduate School of Medicine, Chiba University, Chiba 260-8670, Japan; and ‡Laboratory of Functional Genomics, Department of Medical Genome Sciences, Graduate School of Frontier Sciences, University of Tokyo, Chiba 277-8562, Japan

1S.H. and A.O. contributed equally to this study.

Received for publication January 24, 2011. Accepted for publication March 24, 2011.

This work was supported by the Global COE Program (Global Center for Education and Research in Immune System Regulation and Treatment); the Ministry of Health, Labor, and Welfare (Japan); the Uehara Memorial Foundation; and the following grants from the Ministry of Education, Culture, Sports, Science, and Technology, Japan: Grants-in-Aid for Scientific Research on Priority Areas (17016010), Scientific Research (B) Grant 21390147, Start-up Grant 19890041, and Start-up Grant 2056131.

The sequences presented in this article have been submitted to the Gene Expression Omnibus database (http://www.ncbi.nlm.nih.gov/geo) under accession number GSE28292.

Address correspondence and reprint requests to Dr. Toshinori Nakayama, Department of Immunology, Graduate School of Medicine, Chiba University, 1-8-1 Inohana, Chuo-ku, Chiba 260-8670, Japan. E-mail address: tnkayama@faculty.chiba-u.jp

The online version of this article contains supplemental material.

Abbreviations used in this article: ChIP, chromatin immunoprecipitation; ChIP-chip, ChIP with microarray technology; ChIP-Seq, ChIP assay coupled with massive parallel sequencing; HPRT, hypoxanthine phosphoribosyltransferase; IP, immunoprecipitation; NGFR, nerve growth factor receptor; qPCR, quantitative PCR; qRT-PCR, quantitative RT-PCR; RS, recovery score; TES, transcriptional end site; TSS, transcriptional start site; WT, wild-type.

Copyright © 2011 by The American Association of Immunologists, Inc. 0022-1767/11/$16.00
and histone modifications by ChIP-Seq analysis in Th1 and Th2 cells. Simultaneously, 26 highly Th2-specific inducible genes were identified by a DNA microarray-based three-step selection process. In addition, we examined expression of GATA3 and constitutively active STAT6 (STAT6VT in STAT6-deficient Th2 cells to determine direct effects of GATA3 and STAT6 on Th2-specific gene expression. Among the Th2-specific genes, transcription of nine genes was directly regulated by GATA3 in the absence of STAT6. However, other Th2-specific genes were regulated not by GATA3 itself, but in collaboration with STAT6.

**Materials and Methods**

**Mice**

C57BL/6 mice were purchased from CLEA (Tokyo, Japan). STAT6-deficient mice (27) were provided by S. Akira (Osaka University, Osaka, Japan). All mice used in this study were maintained under specific pathogen-free conditions and ranged from 6 to 8 wk of age. All animal care was performed in accordance with the guidelines of Chiba University.

**Reagents**

Recombinant mouse IL-12 was purchased from BD Pharmingen, and recombinant mouse IL-4 was from TOYOBO (Osaka, Japan).

**Abs**

The Abs used for the ChIP assay were anti-GATA3 mAb (Santa Cruz Biotechnology), anti-acetylated histone H3-K9 (Upstate; 06-599), anti-trimethylhistone H3-K4 (LP Bio; AR-0169), and anti-trimethylhistone H3-K27 (Upstate; 07-449). The Abs used for cytoplasmic staining were anti–IFN-γ FITC (XMG1.2) and anti–IL-4 PE (11B11).

**The generation of effector Th1 and Th2 cells**

Effector Th1/Th2 cells were generated, as previously described (32). Splenic CD4 T cells were prepared using a magnetic cell sorter (AutoMACS; Milenyi Biotec) yielding a purity of >98%. Where indicated, cells from C57BL/6 mice were stimulated with immobilized anti-TCR mAb (H57-597; 3 μg/ml) and anti-CD28 mAb under Th1 or Th2 culture conditions for 5 d in vitro. Th1 conditions were as follows: 25 U/ml IL-2, 10 μM IL-12, and anti-IL-4 mAb. Th2 conditions were as follows: 25 U/ml IL-2 and 100 μM IL-4. These cells were used as either effector Th1 or Th2 cells, respectively.

**Microarray data collection and analysis**

Total cellular RNA from cells cultured under Th1 and Th2 conditions was extracted with TRIzol reagent (Invitrogen), according to the manufacturer’s instructions. Approximately 5 μg RNA was labeled and hybridized to GeneChip Mouse Genome 430 2.0 arrays (Affymetrix), according to the manufacturer’s protocol. Expression values were determined with GeneChip Operating Software v1.2 software (Affymetrix).

**Quantitative RT-PCR**

Total RNA was isolated using the TRIzol reagent (Invitrogen). cDNA was synthesized using oligo(dT) primer and Superscript II RT (Invitrogen). Quantitative RT-PCR was performed, as described previously, using an ABI PRISM 7500 Sequence Detection System (32). The primers and TaqMan probes for the detection of the indicated genes and hypoxanthine phosphoribosyltransferase (HPRT) were purchased from Applied Biosystems and Roche, respectively. The expression was normalized by the HPRT signal. For quantitative RT-PCR (qRT-PCR) of cytokine genes, mRNA from Th1 or Th2 cells that were restimulated by anti-TCR mAb (H57-597; 3 μg/ml) for 4 h was used.

**ChIP assay**

ChIP was performed using ChIP assay kits (Upstate Biotechnology), as previously described (32). Real-time quantitative PCR (qPCR) analysis was performed on an ABI Prism 7500 real-time PCR machine with TaqMan probes and primers. To calculate the enrichment of each protein to a particular target DNA, values obtained (via the standard curve method) for each target were divided by the amount of the corresponding target in the input fraction. Enrichments obtained from mock immunoprecipitations (IPs) performed in parallel with normal IgG were then subtracted from the enrichment values obtained with specific Abs (specific Ab ChIP − control Ig ChIP/input DNA). All the enrichments are expressed as a function of the highest enrichment obtained on the locus (set to 10) (33).

**ChIP-Seq and Illumina sequencing**

For ChIP-Seq analysis, IP and input samples were prepared using ChIP-Seq Sample Prep kit (Illumina). Adaptor-ligated DNA fragments were size fractionated by 12% acrylamide gel, and the 170- to 250-bp fraction was recovered. DNA thus obtained was amplified by 18 cycles of PCR. One nanogram of DNA was used for the sequencing reaction of the Illumina GAfix, according to the manufacturer’s instructions. A total of 170,000–250,000 clusters was generated per tile, and 36 cycles of the sequencing reactions were performed. Short-read sequences were aligned to the mouse genome sequences (mm9 from University of California, Santa Cruz Genome Browser; http://genome.ucsc.edu/) using the Eland program. Sequences allowing no more than two mismatches per sequence were used for the analysis. To enumerate GATA3-bound genes, at least one peak (with a 5-fold increase in signal intensity compared with input DNA) detected on the gene locus was selected. MEME (34) was used for performing a de novo search of consensus-binding motifs for GATA3.

**Retroviral vectors and infection**

The pMX-IRES-human nerve growth factor receptor (NGFR) plasmid was generated, as previously described (16). The infected cells were enriched

**FIGURE 1.** Schematic view of ChIP-Seq experiments and screening procedure of Th2-specific STAT6-dependent inducible genes. Freshly isolated CD4 T cells from C57BL/6 WT mice were stimulated with immobilized anti-TCR mAb and anti-CD28 mAb under Th1 or Th2 culture conditions for 5 d in vitro. These cells were used for genome-wide GATA3 binding and histone modification (H3-K4Me3, H3-K9Ac, and H3-K27Me3) analysis by ChIP-Seq and Illumina sequencing.
by MACS with anti-human NGFR (clone C40-1457; BD Pharmingen), and were subjected to a qRT-PCR assay. GATA3 or STAT6 dependency was evaluated by the recovery score (RS), which is defined as the linear equation below:

\[ RS(\%) = 100 \times \frac{(C - B)}{(A - B)} \]

where A indicates the signal intensity of Th2, B indicates the signal intensity of STAT6KO Th2, and C indicates the signal intensity of STAT6KO Th2 with overexpression of GATA3 or STAT6VT.

**Accession number**

Normalized microarray raw data are available in the Gene Expression Omnibus database (http://www.ncbi.nlm.nih.gov/geo) under accession number GSE28292.

**Results**

**Genome-wide analysis of GATA3-binding genes and histone modifications in Th1 and Th2 cells**

To globally identify the GATA3-binding genes and their histone modification status in Th1 and Th2 cells, we performed ChIP-Seq analysis using Abs specific for GATA3, histone H3K9 acetylation (H3-K9Ac), histone H3K4 trimethylation (H3-K4Me3), and histone H3K27 trimethylation (H3-K27Me3). The ChIP DNA fragments were sequenced with an Illumina-Solexa 1G Genome Analyzer, as described (30, 35). A total of 5,400,000–9,000,000 sequence tags was generated for Th1 and Th2 cells and mapped to the mouse genome (Fig. 1). We identified 510 GATA3-bound peaks in Th1 and 2,258 GATA3-bound peaks in Th2 cells. To enumerate GATA3-bound genes, the genes having at least one significant peak (>5-fold increase in signal intensity compared with input DNA) detected in the gene locus were selected (hereafter defined as GATA3-bound genes). Each gene locus was defined as a region from 5 kb upstream of the transcriptional start site (TSS) to 5 kb downstream of transcriptional end site (TES). Using these criteria, we identified 250 genes selectively bound by GATA3 in Th1 cells and 1,279 genes in Th2 cells. GATA3 binding was detected in both Th1 and Th2 cells at 101 gene loci (Fig. 2A).

We next compiled the localization of H3-K9Ac, H3-K4Me3, and H3-K27Me3 marks for the GATA3-bound genes in relation to each gene TSS ± 10 kb to visualize the tag density profile for each mark (Fig. 2B). Histone H3-K9Ac and H3-K4Me3 are permissive marks for chromatin, and H3-K27me3 is a mark for the closed configuration of chromatin (30). In both Th1 and Th2 cells, H3-K9Ac (in blue) and H3-K4Me3 (in red) were highly enriched around the TSS, as previously reported (30, 36). We also found that H3-K27Me3 marks (in green) were relatively depleted around the TSS of the GATA3-bound genes. Of note, the levels of H3-K9Ac and H3-K4Me3 marks were lower at GATA3-bound genes in Th1 cells, compared with those in Th2 cells. We also compiled the localization of GATA3 peaks for all target genes in the context of the TSS ± 10 kb to again visualize the tag density profile, and a marked enrichment for binding was observed around the TSS (Fig. 2B, lower panels). To determine whether epigenetic marks are globally aligned with GATA3 occupancy, we plotted cumulative sequence frequencies associated with GATA3-bound sites for each of the histone marks (Fig. 2C). The analysis revealed the enrichment of H3-K9Ac (in blue) and H3-K4Me3 (in red) around the GATA3 peak, whereas H3-K27Me3 marks (in green) were not enriched.

Previous in vitro studies have indicated that GATA family proteins bind to a WGATAR motif (37). To determine whether such an element was identifiable in our genome-wide analysis of GATA3 binding, we carried out de novo searches of consensus motifs of GATA3-bound regions in Th1 and Th2 cells using MEME (34). In the GATA3 peaks detected in Th2 cells, we found 2257 motifs that include a GATA motif (Fig. 2D). For the GATA3

![Image](http://www.jimmunol.org/)

**FIGURE 2.** Genome-wide distribution of GATA3 binding sites and identification of consensus-binding motifs. A, Venn diagram showing the number of genes bound by GATA3 uniquely in Th1, Th2, or genes bound in both Th1 and Th2 cells. GATA3-bound genes were identified if at least one peak (>5-fold increase in signal intensity compared with input DNA) of binding was present within the gene locus. B, Compiled tag density profiles across the TSS ± 10-kb flanking regions with 100-bp resolution for H3-K9Ac, H3-K4Me3, and H3-K27Me3 and GATA3 binding at GATA3-bound genes in Th1 and Th2 cells are shown. C, Compiled tag density profiles across the GATA3 peak ± 10-kb flanking regions at GATA3-bound genes with 100-bp resolution for H3-K9Ac, H3-K4Me3, and H3-K27Me3 in Th1 and Th2 cells are shown. D, MEME (34) was used to perform a de novo search of consensus-binding motifs for GATA3 in Th2 cells.
peaks in Th1 cells and those overlapped in Th1 and Th2 cells, we did not identify a consensus motif for GATA3 as a WGATAR (Supplemental Fig. 1A). We also identified the ETS-binding motif (GGAA) on GATA3 peaks in Th2 cells in the MEME analysis (Supplemental Fig. 1B).

**Histone modification patterns in the GATA3-bound genes**

To assess the correlation between direct binding of GATA3 and histone modifications, we evaluated global histone modification patterns at the 1380 GATA3-binding gene loci using the results of ChIP-Seq analyses. We quantified the total tag counts of H3-K9Ac, H3-K4Me3, and H3-K27Me3 at the promoter region, which is defined as a region from 5 kb upstream to 3 kb downstream relative to the TSS in this analysis. The tag counts were normalized by the total number of input DNA tags at the corresponding region. We clustered GATA3-bound genes on the basis of the ratio of tag counts in Th1 and Th2 cells (Th2/Th1 ratio) for each of the epigenetic marks (Fig. 3A). From this analysis, four epigenetic patterns for GATA3-bound genes emerged. We identified GATA3-bound gene clusters with a Th2/Th1 ratio >2 (in red) for H3K9Ac (denoted H3K9Ac high, 7.6%), and H3K4Me3 (H3K4Me3 high, 4.2%), and <0.5-fold Th2/Th1 ratio (in green) for H3K27Me3 (H3K27Me3 low, 11.6%) (Fig. 3B). For the largest cluster (63% of the total GATA3-bound genes), GATA3 binding had minimal effects (between 1- and 2-fold Th2/Th1 ratio, in pink; between 0.5- and 1-fold Th2/Th1 ratio, in white) on any epigenetic modifications we investigated (denoted as indeterminate). We also obtained similar results from the GATA3-bound genes in Th1 cells and those detected in both Th1 and Th2 cells (Supplemental Fig. 1C, 1D). These results indicate that GATA3 binding correlates very well with changes in histone modifications between Th1 and Th2 cells in a set of genes, but also shows that in the majority of cases, GATA3 binding and changes in histone modifications do not coincide.

**Selection of Th2-specific genes**

The genome-wide surveys on GATA3 binding and histone modifications indicate a possibility that GATA3 binding is not always associated with changes in histone modifications, which may correlate with the status of transcription. An emerging concept coming from genome-scale studies is that only a fraction of factor-binding genes showed clear functional dependence, as evaluated by gene expression changes (13, 38). Therefore, we next sought to determine the extent to which the status of transcription correlates with both the binding of GATA3 and histone modifications at specific gene loci. To investigate the role of GATA3 during Th2 cell differentiation, we focused on Th2-specific genes in which the level of transcription was upregulated during Th2 cell differentiation and was preferentially higher in Th2 cells as compared with Th1 cells. To identify Th2-specific genes, we adopted three-step selection processes (Fig. 1). On the first step, we assessed gene expression profiles in in vitro generated Th1 and Th2 cells by a DNA microarray analysis, and compared the results with those reported previously (28). A list of the potential Th2-specific genes (185 genes) was obtained. On the second step, the 185 selected genes were validated using ΔΔ cycle threshold qRT-PCR with mRNA from wild-type (WT) and STAT6-deficient Th2 cells (Supplemental Fig. 2). mRNA levels of 92 of the 185 candidate genes were increased by at least 2-fold in WT Th2 cells compared with Th2 cells lacking STAT6. Ninety-two STAT6-dependent

**FIGURE 3.** GATA3-bound genes in Th2 cells form clusters that share common epigenetic signatures. A, The total tag count of each epigenetic modification (H3K9Ac, H3K4Me3, and H3K27Me3) was computed across the promoter of each GATA3-bound gene in Th2 versus Th1 cells. Red indicates >2-fold Th2/Th1 ratio; pink indicates 1- to 2-fold Th2/Th1 ratio; white indicates 0.5- to 1-fold Th2/Th1 ratio; and green indicates <0.5-fold Th2/Th1 ratio. B, The ratio of tag counts for three epigenetic modifications was used to cluster GATA3-bound genes in four patterns, as follows: H3K9Ac-high, H3K4Me3-high, H3K27Me3-low, and an indeterminate pattern.
genes were identified in Th2 cells. Then qRT-PCR analysis was performed using Th1, Th2, and STAT6-deficient Th2 cells (the third step). mRNA levels of 38 genes of the 92 genes were increased by at least 4-fold in WT Th2 cells compared with Th1 or STAT6-deficient Th2 cells, whereas mRNA levels for 10 of these 38 genes were not increased by >2-fold in Th2 cells compared with fresh CD4 T cells, and so were also excluded (data not shown). We also excluded two genes that were not described as the RefSeq genes (mm9 from University of California, Santa Cruz Genome Browser; http://genome.ucsc.edu/). As a result, 26 genes were selected as highly Th2-specific STAT6-dependent inducible genes. The preferential expression of these genes in Th2 cells is shown in comparison with Th1 cells and STAT6-deficient Th2 cells (Fig. 4A) or freshly prepared CD4 T cells (Fig. 4B). We focused on these 26 genes for further study.

H3-K9Ac, H3-K4Me3, and H3-K27Me3 modifications at the 26 Th2-specific STAT6-dependent inducible gene loci

The chromatin status at the 26 previously selected Th2-specific gene loci was determined using the results of the ChIP-Seq analysis on histone modifications in Th1 and Th2 cells. We quantified the total tag counts of H3-K9Ac, H3-K4Me3, and H3-K27Me3 at each gene locus (5 kb upstream of the TSS to 5 kb downstream of TES) (Fig. 5A). The raw tag count patterns are also shown (Supplemental Fig. 3). The histone modification tag counts were also quantified at each gene promoter, which is defined as a region from 5 kb upstream to 3 kb downstream relative to the TSS (Fig. 5B). The tag counts were normalized by the total number of input tags at the corresponding region. Epas1, GATA3, NFIL3, GZMA, IL-4, IL-5, IL-13, IL-24, CCR8, Ecm1, IL1r2, Itgb3, RNF128, TNFRSF8, TMTC2, Cyp11a1, S100a1, and Tube1 gene loci had higher levels of H3-K9Ac by at least 1.5-fold in Th2 cells (Fig. 5A, top). Cell surface molecules tended to have the lowest levels of H3-K9Ac and H3-K4Me3 when the whole gene locus was analyzed (Fig. 5A). Overall, the results of H3-K9Ac and H3-K4Me3 levels at the whole gene locus were similar to those of the promoter region, although the levels were slightly higher in the promoter region (Fig. 5). The levels of H3-K27Me3 at the Epas1, GATA3, IL-4, IL-5, IL-13, IL-24, CCR8, Ecm1, TNFRSF8, Cyp11a1, DUSP4, and TANC2 gene loci were higher by at least 1.5-fold in Th1 cells as compared with Th2 cells. These results indicate that increased permissive histone modifications tend to associate with Th2-specific gene expression. However, changes in gene expression are not always accompanied by epigenetic changes in histone modifications, and this is particularly evident in the cell surface molecule group.

Analysis of GATA3 binding at the 26 Th2-specific STAT6-dependent inducible gene loci

To assess the possible correlation between the direct binding of GATA3 and transcription of the Th2-specific genes, we next searched GATA3 binding at the 26 Th2-specific STAT6-dependent inducible genes using the results of GATA3 ChIP-Seq analysis. Of the 26, GATA3 binding was detected at 17 gene loci (Fig. 6A), leaving 9 gene loci with no discernible GATA3 binding (data not shown). Thirteen genes showed a single peak, and 4 genes showed more than one peak at their gene loci. GATA3 binding was strongly detected at the region named conserved GATA3 response element, which is located 1.6 kb upstream of the IL-13 gene TSS (39). The GATA3 binding site in the GATA3 gene locus was very close to the previously identified STAT6 binding site (13, 14). These GATA3 binding sites detected by ChIP-Seq analysis were

**FIGURE 4.** Th2-specific genes evaluated by qRT-PCR. A and B. The gene expression levels in Th1, Th2, and STAT6-deficient Th2 cells (A), or in freshly prepared CD4 T cells and Th2 cells (B) were determined by qRT-PCR. The relative intensity (relative expression/HPRT; highest signal intensity = 10) (mean of two samples with SD) is shown.
then confirmed by conventional ChIP assay with qPCR (ChIP-qPCR) in Th1, Th2, and STAT6-deficient Th2 cells, and Th2-specific increased binding of GATA3 was detected at all of the testable regions (Fig. 6B).

**Preferential GATA3 binding and H3K9Ac modification in the 26 Th2-specific STAT6-dependent inducible genes**

Next, we studied GATA3 binding and histone H3K9Ac in the 92 Th2-specific and STAT6-dependent genes shown in Supplemental Fig. 2A. We divided the 92 genes into two groups, as follows: the 26 inducible genes and the 66 other genes, and compared the frequency of the genes with GATA3 binding (Supplemental Fig. 4A) and increased levels of H3K9Ac (Th2/Th1 ratio; Supplemental Fig. 4B). The percentages of GATA3-binding target genes and also genes that displayed high H3K9Ac (Th2/Th1 ratio >1.5) were higher in the 26 STAT6-dependent inducible genes as compared with the other 66 genes (GATA3 target genes, 65.4 versus 31.8%, and high H3K9Ac genes, 73.1 versus 30.0%).

**Enforced expression of GATA3 induces transcription of a subset of Th2-specific genes in STAT6-deficient Th2 cells**

Next, we assessed whether the enforced expression of GATA3 or STAT6 induces the transcription of the 26 Th2-specific genes using a retrovirus gene introduction system. To examine GATA3 dependency semiquantitatively, we sorted WT and STAT6-deficient Th2 cells overexpressing GATA3 into three populations by a cell sorter, as follows: NGFR\(^{2-}\) (expressing no retrovirus-induced GATA3, G1), NGFR\(^{\text{low}}\) (expressing low levels of retrovirus-induced GATA3, G2), and NGFR\(^{\text{high}}\) (expressing high levels of retrovirus-induced GATA3, G3) (Fig. 7A). The mRNA levels of the 26 Th2-specific genes including 17 GATA3-bound genes (Fig. 7B) and 9 GATA3-nonbound genes (Fig. 7C) were then assessed in these sorted populations. The mRNA expression level of GATA3 in the NGFR\(^{\text{high}}\) population of the STAT6-deficient GATA3 retrovirus-infected group was similar to that of WT Th2 cells (data not shown) (23). In GATA3 high-expressing cells (G3), the mRNA levels of IL-13, Cyp11a1, Asb2, JDP2, IL-5, TMTC2, IL-4, Ccnj1, IL-24, and PTGIR were greatly increased (>40% recovery based on the RS, see Materials and Methods) (Fig. 7D), and this increase occurred in a gene dosage-dependent manner. In contrast, the other 16 Th2-specific genes showed little or no recovery even in the GATA3 high-expressing cells (G3) (see GATA3-low genes in Fig. 7B–D). Next, a constitutively active form of STAT6, STAT6VT, was introduced into STAT6-deficient CD4 T cells cultured under Th2 conditions (Fig. 8). Among the 16 Th2-specific genes that showed little or no recovery even in the GATA3 high-expressing cells (G3), STAT6VT introduction resulted in >40% recovery of mRNA expression in all (GATA3, Itg\(\beta3\), GZMA, Eps1, Crem, NFIL3, CCR8, TNFRSF8, S100a1, Ecml, Dusp4, Tube1, IL1r2, RNF128, F2R, and TANC2) (see Fig. 8B, GATA3-low genes). Furthermore, among these 16 genes, we detected GATA3 binding on 8 genes, as follows: GATA3, Itg\(\beta3\), GZMA, Eps1, Crem, NFIL3, CCR8, and TNFRSF8 (Fig. 6A). Taken together, these results indicate that among Th2-specific...
STAT6-dependent genes, some genes are bound by GATA3, and their transcription is directly regulated by GATA3, whereas others are either not directly bound by GATA3 or not regulated by GATA3 itself, but in collaboration with other molecules, including STAT6 (Table I).

**Discussion**

We performed ChIP-Seq analyses to assess genome-wide GATA3 binding and histone modifications in Th1 and Th2 cells. Simultaneously, 26 genes that are selectively induced in developing Th2 cells were identified. The recruitment of GATA3 and histone modifications at these 26 gene loci were analyzed in detail using the results of ChIP-Seq analysis, quantitative ChIP, and qRT-PCR assays. Although GATA3 has been recognized as a master regulator of Th2 cell differentiation, we found that the transcription of many Th2-specific genes was not regulated by GATA3 itself, and therefore GATA3 is not the only factor responsible for the induction of the transcription of a set of Th2-specific genes.

In this study, we investigated in detail the dependency of 26 Th2-specific STAT6-dependent inducible genes on GATA3. To select the genes that are preferentially induced in developing Th2 cells, comparisons between WT Th2 cells and Th1 cells and between WT Th2 cells and STAT6-deficient Th2 cells were adopted (Fig. 1). It is known that the IL-4/IL-4R/STAT6 pathway is not the only downstream
pathway of IL-4R signaling (40, 41). Therefore, the current study may exclude IL-4/IL-4R–dependent and STAT6-independent genes. Although it has not been well studied whether these genes are important for Th2 cell function, the use of IL-4–deficient and/or IL-4R–deficient T cells may allow us to address this issue.

We examined GATA3 binding and H3K9Ac levels in the 92 Th2-specific STAT6-dependent genes, and compared the frequency between the 26 Th2-specific inducible genes and the remaining 66 genes (Supplemental Fig. 4). Preferential GATA3 binding and increased H3K9Ac levels were detected in the 26 Th2-specific inducible genes. These results may indicate a critical role for GATA3 in the increased expression of Th2-specific genes during Th2 cell differentiation.

Several interesting features in the regulation of transcription of Th2-specific genes have been revealed. The summary is depicted in Table I. We identified nine GATA3-bound target genes (IL-13, Cyp11a1, Asb2, JDP2, IL-5, TMTC2, IL-4, Ccnj1, IL-24) whose mRNA expression was highly dependent on GATA3 in the absence of STAT6 (GB-1 group in Table I). For these genes, the induction of GATA3 appears to be necessary and sufficient for the upregulation of expression during Th2 cell differentiation. We also identified eight GATA3-bound target genes (GATA3, Itgβ3, GZMA, Epas1, Crem, NFIL3, CCR8, and TNFRSF8) whose mRNA expressions were strongly induced by STAT6VT introduction, but not by GATA3 (GB-2 group in Table I). For these eight genes, GATA3 may not be the only factor that regulates transcription, but may cooperate with STAT6 or STAT6-dependent genes. Furthermore, we also identified GATA3-nonbinding Th2-specific genes whose expression was not recovered very well by GATA3 introduction, but strongly induced by STAT6VT introduction (S100a1, Ecm1, Dusp4, Tube1, IL1r2, RNF128, F2r, and TANC2) (GN-2 group in Table I). For these eight genes, STAT6 could play a critical role in transcription regardless of the expression levels of GATA3. Finally, an interestingly behaving gene PTGIR was identified (GN-1 group in Table I). Although PTGIR was categorized as a GATA3 nontarget gene, the transcription of PTGIR was highly dependent on GATA3. Alternatively, GATA3 may bind to a distant gene enhancer region that is beyond the gene boundaries used in our global analysis (from 5 kb upstream of TSS to 5 kb downstream of TES). As for the PTGIR gene, the nearest GATA3 peak was located 872 kb upstream of the PTGIR gene TSS (S. Horiuchi, A. Onodera, Y. Suzuki, and T. Nakayama, unpublished observations).
GATA3, Itgb3, GZMA, Eps1, Crem, NFIL3, CCR8, TNFRSF8, Ecm1, Dusp4, Tube1, IL1r2, and TANC2 were classified as GATA3-low/STAT6-high (GB-2 and GN-2 in Table I), and each of these have been reported previously as STAT6-bound genes (second column, Table I) (13). We further confirmed a significant STAT6-dependent recovery of these genes (Table I, right column). Thus, STAT6 binding is most likely important for the transcription for these 13 genes. Among these 13 genes, GATA3, Itgb3, GZMA, Eps1, Crem, NFIL3, CCR8, and TNFRSF8 were GATA3-bound genes (GB-2 in Table I), and therefore, GATA3 may collaborate with STAT6 to regulate transcription. Indeed, each GATA3 peak was located within 500 bp of the STAT6 peak in these GB-2 gene loci (13). In contrast, S100a1, RNF128, and F2R were neither GATA3-bound genes nor STAT6 bound. The Th2-specific increase in transcription of these three genes could be regulated by other genes that were directly regulated by STAT6 in a similar fashion to that proposed for GATA3 and PTGIR.

We selected 26 Th2-specific genes that include previously reported Th2-specific genes (IL-4, IL-13, IL-5, IL-24, GATA3, CCR8, and TNFRSF8) (13, 42–44), and also new genes that were not reported previously (Fig. 4). The newly identified genes were phenotype-defining transcription factors (Crem, Eps1, JDP2, and NFIL3), a cytokine (GZMA), cell surface molecules (Ecm1, F2r, IL1r2, Itgb3, PTGIR, RNF128, and TMTC2), signaling molecules, and others with unknown functions (Asb2, Ccnj1, Cyp11a1, Dusp4, S100a1, TANC2, and Tube1). The IL-4-inducible gene NFIL3 is a transcription factor that is required for NK cell development and maintenance (45, 46). The c-Jun dimerization protein 2, JDP2, is a DNA-binding protein that is known to be involved in cell differentiation process in skeletal muscle cells and osteoclasts (47, 48).

Of the GATA3 peaks detected in Th2 cells, we found WGATAR motifs, which include GATA (Fig. 2D). Furthermore, we identified the ETS-binding motif (GGAA) among GATA3 peaks in Th2 cells.
Role of GATA3 in Th2-specific gene induction

In this study, integration of the three following biological readouts was used for the identification of gene clusters: 1) GATA3 binding; 2) changes in gene expression by enforced expression of GATA3; and 3) changes in gene expression by enforced expression of STAT6.VT. These represent various actions of GATA3 and STAT6 involved in the transcription of Th2-specific genes. GATA3 or STAT6 dependency of the 26 genes was determined by the RS, as described in Materials and Methods.

In the MEME analysis (Supplemental Fig. 1B), indicating a possibility that ETS family members bound to the sequence close to the GATA3-binding motif. Indeed, ETS-1 is reported to play an important role in regulating the expression of Th2 cytokines in Th2 cells (49). A consensus motif for GATA3 was not identified in the GATA3 peaks detected in Th1 cells or those that overlapped in Th1 and Th2 cells (Supplemental Fig. 1A). However, because the expression of GATA3 is lower in Th1 cells in comparison with Th2 cells, the binding peak analysis may incorrectly identify nonspecific binding more frequently in Th1 cells.

Our present genome-wide surveys on GATA3 binding and histone modifications indicate that GATA3 binding does not always correlate with changes in the active form of histone modifications in Th2 cells. From this result, there is a possibility that GATA3 acts not only as an activator, but also as a repressor in both Th1 and Th2 cells. Thus, the set of GATA3-binding genes that has no active form not only as an activator, but also as a repressor in both Th1 and Th2 cells. From this result, there is a possibility that GATA3 acts

Th2 cell differentiation is accompanied by histone modifications, such as H3-K9/14 acetylation and H3-K4 methylation at the IL-4, IL-13, and IL-5 loci, which are dependent on GATA3 expression (19, 20, 39). There is a previous report on genome-wide GATA3 target analysis in primary human T cells using a ChIP with microarray technology (ChIP-chip) analysis (50). We compared the results of this study with our own, and an interesting difference was noted. IL-4 was identified as a GATA3 target in both analyses. However, many GATA3 target genes identified in the current study, including IL-5 and IL-13, were not listed in the previous genome-wide ChIP-chip analysis in human T cells (50). This discrepancy may be explained by the difference in experimental systems, as follows: 1) the difference between human and mouse; 2) the different analysis of ChIP-chip and ChIP-Seq; 3) the difference in the time course of gene expression in vitro differentiated human and mouse T cells; and 4) the use of a different anti-GATA3 Ab. In fact, in the genome-wide analyses of STAT6 in human and mice, STAT6 binding to certain genes, including GATA3, was different between the two species (13, 31). The difference in the results between ChIP-chip and ChIP-Seq analyses was also reported (51, 52). Human Th2 cells were prepared 7 d after TCR stimulation (50). This procedure may not induce high-level expression of many Th2-specific genes because our previous studies pointed out that two cycles of TCR stimulation

Table I. Overview of data presented in this study, highlighting representative genes that belong to each cluster

<table>
<thead>
<tr>
<th>RefSeq Accession</th>
<th>Gene Name</th>
<th>Gene Product</th>
<th>GATA3</th>
<th>STAT6 (13)</th>
<th>GATA3-dep. Recovery</th>
<th>STAT6-dep. Recovery</th>
</tr>
</thead>
<tbody>
<tr>
<td>GATA3 target genes (GB)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NM_008355</td>
<td>IL-13</td>
<td>IL-13</td>
<td>o</td>
<td>x</td>
<td>201.35</td>
<td>218.76</td>
</tr>
<tr>
<td>NM_019779</td>
<td>Cyp11a1</td>
<td>Cytochrome P450, family 11, subfamily a, polypeptide 1</td>
<td>o</td>
<td>o</td>
<td>122.42</td>
<td>130.58</td>
</tr>
<tr>
<td>NM_023049</td>
<td>Asb2</td>
<td>Ankyrin repeat and SOCS box-containing 2</td>
<td>o</td>
<td>o</td>
<td>92.92</td>
<td>48.75</td>
</tr>
<tr>
<td>NM_030893</td>
<td>Jdp2</td>
<td>Jun dimerization protein 2</td>
<td>o</td>
<td>o</td>
<td>78.41</td>
<td>68.47</td>
</tr>
<tr>
<td>NM_010558</td>
<td>IL-5</td>
<td>IL-5</td>
<td>o</td>
<td>x</td>
<td>75.90</td>
<td>11.89</td>
</tr>
<tr>
<td>NM_177368</td>
<td>Tmtc2</td>
<td>Transmembrane and tetratricopeptide repeat containing 2</td>
<td>o</td>
<td>o</td>
<td>55.56</td>
<td>42.60</td>
</tr>
<tr>
<td>NM_021283</td>
<td>IL-4</td>
<td>IL-4</td>
<td>o</td>
<td>o</td>
<td>50.77</td>
<td>33.53</td>
</tr>
<tr>
<td>NM_001045530</td>
<td>Ccnd1</td>
<td>Cyclin D1-like</td>
<td>o</td>
<td>o</td>
<td>42.66</td>
<td>36.09</td>
</tr>
<tr>
<td>NM_053095</td>
<td>IL-24</td>
<td>IL-24</td>
<td>o</td>
<td>o</td>
<td>41.26</td>
<td>111.00</td>
</tr>
<tr>
<td>GATA3 low genes (GB-2)</td>
<td>STAT6 high genes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NM_008091</td>
<td>Gata3</td>
<td>GATA-binding protein 3</td>
<td>o</td>
<td>o</td>
<td>12.35</td>
<td>131.76</td>
</tr>
<tr>
<td>NM_016780</td>
<td>lgf3</td>
<td>Integrin β3</td>
<td>o</td>
<td>o</td>
<td>26.34</td>
<td>142.46</td>
</tr>
<tr>
<td>NM_010360</td>
<td>Gzma</td>
<td>Granzyme A</td>
<td>o</td>
<td>o</td>
<td>5.74</td>
<td>113.82</td>
</tr>
<tr>
<td>NM_010117</td>
<td>Eps1</td>
<td>Endothelial PAS domain protein 1</td>
<td>o</td>
<td>o</td>
<td>0.06</td>
<td>88.21</td>
</tr>
<tr>
<td>NM_013498</td>
<td>Crem</td>
<td>cAMP-responsive element modulator</td>
<td>o</td>
<td>o</td>
<td>27.90</td>
<td>85.55</td>
</tr>
<tr>
<td>NM_017373</td>
<td>Nfil3</td>
<td>NF, IL-3, regulated</td>
<td>o</td>
<td>o</td>
<td>2.32</td>
<td>71.37</td>
</tr>
<tr>
<td>NM_007720</td>
<td>Ccr8</td>
<td>Chemokine (C-C motif) receptor 8</td>
<td>o</td>
<td>o</td>
<td>19.44</td>
<td>68.65</td>
</tr>
<tr>
<td>NM_009401</td>
<td>Tnfrsf8</td>
<td>TNFR superfamily, member 8</td>
<td>o</td>
<td>o</td>
<td>16.18</td>
<td>66.62</td>
</tr>
<tr>
<td>GATA3 nontarget genes (GN)</td>
<td>GATA3 high genes (GN-1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NM_008967</td>
<td>Ptgir</td>
<td>PG I receptor (IP)</td>
<td>x</td>
<td>x</td>
<td>110.02</td>
<td>126.88</td>
</tr>
<tr>
<td>GATA3 high genes (GN-1)</td>
<td>STAT6 low genes (GN-2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NM_011309</td>
<td>S100a1</td>
<td>S100 calcium-binding protein A1</td>
<td>x</td>
<td>x</td>
<td>18.57</td>
<td>659.10</td>
</tr>
<tr>
<td>NM_007899</td>
<td>Ecm1</td>
<td>Extracellular matrix protein 1</td>
<td>x</td>
<td>o</td>
<td>2.16</td>
<td>289.46</td>
</tr>
<tr>
<td>NM_176933</td>
<td>Dusp4</td>
<td>Dual specificity phosphatase 4</td>
<td>x</td>
<td>o</td>
<td>8.20</td>
<td>209.64</td>
</tr>
<tr>
<td>NM_028006</td>
<td>Tub1</td>
<td>α-tubulin 1</td>
<td>x</td>
<td>o</td>
<td>21.52</td>
<td>175.88</td>
</tr>
<tr>
<td>NM_010555</td>
<td>Il1r2</td>
<td>IL-1R, type II</td>
<td>x</td>
<td>o</td>
<td>21.22</td>
<td>146.32</td>
</tr>
<tr>
<td>NM_023270</td>
<td>Rnf128</td>
<td>Ring finger protein 128</td>
<td>x</td>
<td>x</td>
<td>20.85</td>
<td>118.21</td>
</tr>
<tr>
<td>NM_010169</td>
<td>F2r</td>
<td>Coagulation factor II (thrombin) receptor</td>
<td>x</td>
<td>x</td>
<td>29.09</td>
<td>109.11</td>
</tr>
<tr>
<td>NM_181071</td>
<td>Tanc2</td>
<td>Tetratricopeptide repeat, ankyrin repeat, and coiled-coil containing 2</td>
<td>x</td>
<td>x</td>
<td>26.89</td>
<td>56.80</td>
</tr>
</tbody>
</table>

In this study, integration of the three following biological readouts was used for the identification of gene clusters: 1) GATA3 binding; 2) changes in gene expression by enforced expression of GATA3; and 3) changes in gene expression by enforced expression of STAT6.VT. These represent various actions of GATA3 and STAT6 involved in the transcription of Th2-specific genes. GATA3 or STAT6 dependency of the 26 genes was determined by the RS, as described in Materials and Methods. ○, binding is detected; ×, binding is not detected; dep., dependent
were required for the high-level expression of IL-4 and IL-5 (53, 54). We are not able to determine the reason for the discrepancy at present, and further studies are required.

In summary, our study confirms the notion that ChIP-Seq is an excellent technology to facilitate the analysis of genome-wide binding of transcription factors to target genes in a nonbiased fashion, and also indicates that additional accurate gene expression analyses are required to address the true biological function of the transcription factor at the target genes. The present study provides a platform for understanding how GATA3 regulates the transcription of Th2-specific genes during Th2 cell differentiation.

Acknowledgments

We are grateful to Drs. Damon Tumes and Ellen Rothenberg for helpful comments and constructive criticisms in the preparation of the manuscript. We thank Kaoru Sugaya, Hikari K. Asou, Miki Katou, and Toshihiro Ito for excellent technical assistance.

Disclosures

The authors have no financial conflicts of interest.

References


