Single-cell RNA-Seq analysis identifies a noncoding interleukin 4 (IL-4) RNA that post-transcriptionally up-regulates IL-4 production in T helper cells

Weijie Yin, Yan Song, and Xing Chang

From the Chinese Academy of Sciences Key Laboratory of Tissue Microenvironment and Tumor, Shanghai Institute of Nutrition and Health, Shanghai Institutes for Biological Sciences, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Shanghai 200031, China and the Department of Cellular and Molecular Medicine, University of California at San Diego, La Jolla, California 92037

Edited by Ronald C. Wek

High-throughput sequencing has revealed a tremendous complexity of cellular transcriptomes, which is partly due to the generation of multiple alternative transcripts from a single gene locus. Because alternative transcripts often have low abundance in bulk cells, the functions of most of these transcripts and their relationship with their canonical counterparts remain unclear. Here we applied single-cell RNA-Seq to analyze the transcriptome complexity of in vitro-differentiated, murine type 2 T helper (Th2) cells. We found that cytokine gene transcripts contribute most of the intercellular heterogeneity, with a group of universal cytokines, including interleukins 1α, 2, 3, and 16, being bimodally expressed. At the single-cell level, use of alternative promoters prevalently generated alternative transcripts. For instance, although undetectable in bulk cells, a noncoding RNA isoform of IL-4 (IL4nc), which was driven by an intronic promoter in the IL-4 locus, was predominantly expressed in a subset of Th2 cells. IL4nc displayed distinct temporal expression patterns compared with the canonical IL-4 mRNA and post-transcriptionally promoted the production of IL-4 protein in Th2 cells. In conclusion, our findings reveal a mechanism whereby minor noncanonical transcripts post-transcriptionally regulate expression of their cognate canonical genes.

Upon antigen stimulation, naïve helper T cells (Th)2 differentiate into various effector subsets (i.e. Th1, Th2, Th17, and regulatory T cells (Treg)) to orchestrate a proper immune response against foreign pathogens (1). Each subset of T helper cells expresses their characteristic cytokines (i.e. Th1 cells express IFN-γ but not IL-4, whereas Th2 cells express IL-4 but not IFN-γ) and master transcription factors (i.e. Tbet for Th1 cells and GATA-3 for Th2 cells) (2–4). In addition to these lineage-specific factors, effector Th cells also express common cytokines shared among virtually all effector T cells (i.e. IL-2 and tumor necrosis factor α) to carry out their physiological functions. The variable combinations of the unique and common factors further increase the diversity of effector Th cells. Although the lineage decision and functions of Th subsets have been intensively studied (2, 5, 6), the intra-subset diversity and its origin remain to be fully elucidated.

As a prototype of the Th subset, Th2 cells are induced by extracellular pathogens to initiate the humoral immune response. Th2 cells can be in vitro differentiated from naïve CD4+ T cells by antigen and IL-4 stimulation (7–10). Effector activities of Th2 cells depend on an array of secreted proteins, including IL-4, IL-5, IL-9, and IL-13 (collectively called type 2 cytokines) (11, 12) that have shared and unique functions. IL-4 plays a key role in both humoral and cellular immunity, including promoting proliferation and survival of activated B cells, and inducing B cell class switching to IgG1 and IgE (13–15). IL-13, on the other hand, promotes mucus production by goblet cells and airway hyperresponsiveness, a hallmark of allergic asthma (2, 11–13). In addition to type 2 cytokines, Th2 cells also secrete universal cytokines shared with other Th subsets, such as IL-2, IL-1α, and IL-3, although their regulation and function in Th2 cells are largely undetermined.

Recent progress in high-throughput sequencing has revealed the great complexity of cellular transcriptomes, partly contributed by previously underestimated noncoding transcripts. For instance, long noncoding RNAs can be generated from either noncoding DNA or from coding genes through alternative splicing or alternative promoter usage (16, 17). Furthermore, many enhancers are actively transcribed into enhancer RNAs, which are often unspliced, nonpolyadenylated, and associated with active gene expression (18). However, the functions of most noncoding transcripts and their relation to the coding counterparts remain poorly understood, partially because of their low abundance at the bulk cell level and the difficulty to separate them from transcriptional noise.

In this study, we analyzed the transcriptome complexity of in vitro differentiated Th2 cells using single-cell RNA-Seq.
Our analysis revealed that cytokines were the most variably expressed genes among Th2 cells, with universal cytokines, rather than Th2 specific cytokines, bimodally expressed among single cells. Furthermore, we identified a noncoding RNA abundantly transcribed from the IL4 locus in a subset of Th2 cells, which is driven by an enhancer within the second intron of the mouse IL4 gene. We demonstrated that induction of this noncoding isoform preceded the canonical IL-4 mRNA and promoted IL-4 translation following T cell antigen receptor (TCR) restimulation. Thus, our study charted the intra-subset diversity of Th2 cells and revealed a novel mechanism whereby translation of an immune effector cytokine is regulated by an alternative noncoding transcript from the same locus.

Results

Single-cell RNA-Seq reveals transcriptome heterogeneity of Th2 cells

To explore transcriptome diversification during T cell differentiation, we analyzed in vitro differentiated Th2 cells using single-cell RNA-Seq (scRNA-Seq). Purified naïve CD4 T cells (CD4⁺CD62LhiCD44loCD25⁻) were differentiated into the Th2 lineage in vitro by anti-CD3, anti-CD28, IL-4, and anti-IFN-γ stimulation (7, 8). Five days after differentiation, flow cytometric analysis identified that ~25% cells expressed IL-4 protein following TCR restimulation (Fig. S1A), confirming successful differentiation of Th2 cells (7, 8). 56 single cells were then captured from the restimulated Th2 cells for scRNA-Seq analysis using a Fluidigm C1 platform (19, 20). As a control, pooled 10,000 Th2 cells were subjected to a standard RNA-Seq library preparation protocol. Each library was subsequently sequenced to an average depth of 24 million reads with a single-end 100-bp read length (Fig. S1B), which is well above the sequence depth required to cover the transcriptomes of single Th2 cells (Fig. S2). To validate that our scRNA-Seq assay truly reflected cellular heterogeneity, we compared global gene expression between single cells and pooled cells. The average global gene expression of single cells was very similar to the pooled sample (Pearson correlation, r = 0.85) (Fig. 1A, top panel). The transcriptomes of any two single cells, however, differed greatly (average Pearson correlation, r < 0.6), suggesting a large degree of cellular heterogeneity of in vitro differentiated Th2 cells (Fig. 1, A center panel, and B). The pooled sample also correlated poorly with almost all single cells, indicating that traditional RNA-Seq analysis was not able to truly portray transcriptomes of single cells (Fig. 1A, bottom panel, and B). Further hierarchical cluster analysis revealed two distinct populations within the Th2 cells and one outlier, which was excluded from the subsequent analysis (Fig. 1B and Fig. S1C). The scRNA-Seq analysis revealed cellular heterogeneity of in vitro differentiated Th2 cells, which was masked in bulk cell RNA-Seq analysis.

To quantitate variation of gene expression among single cells, we calculated coefficient variation (CV) of 10,565 genes detected in Th2 cells (TPM ≥ 10 in at least one single cell or pooled sample). Consistent with a previous scRNA-Seq analysis in immune cells (21), many housekeeping genes were expressed abundantly (average TPM (μ) > 8) and invariably among single Th2 cells (CV < 0.2). In contrast, although a number of cytokines were also abundantly expressed (μ > 8), their expression varied greatly among individual cells (CV > 0.28) (Figs. 1C and 2A). Indeed, gene ontology analysis revealed that cytokine genes were enriched in a class of the most variably expressed genes (μ > 8, CV > 0.28) in Th2 cells (Fig. 1D). As expected, cellular housekeeping functions, including RNA binding and ribosome components, were enriched in the most consistently expressed genes. Together, these data indicate that cytokine genes are major contributors of cellular heterogeneity of differentiated Th2 cells.

Universal cytokines are bimodally expressed in Th2 cells

Given the highly variable expression and the importance of cytokines in physiological functions of Th2 cells, we further analyzed the expression and distribution of each cytokine gene among single cells. To our surprise, some cytokines exhibited a bimodal expression pattern in which the gene was highly expressed in a subpopulation of cells but minimally expressed
in the remaining cells (Fig. 2A). Most of these bimodally expressed genes were able to be produced by virtually all Th subsets (e.g. IL2, IL1α, and IL3, called universal cytokines hereafter). For instance, IL1α was highly expressed in 25 cells (μ = 10.2) but minimally expressed in the remaining 30 cells (μ = 2.7). In contrast, most of the Th2-specific cytokines (e.g. IL-4), transcription factors (e.g. GATA-3), and housekeeping genes were expressed uniformly (Fig. 2A). For example, IL-4 RNA was abundantly expressed in most cells (μ = 10.9). Similarly, GATA-3, a lineage-determining transcription factor of Th2 cells, was uniformly and abundantly expressed among single cells (22), which also confirmed the successful Th2 differentiation in our experiments.

To quantitate the bimodal distribution of gene expression among single cells, we calculated a bimodal index (BI) for every cytokine gene (23), which uses a two-component mixture model to measure the separation of two normal distributions. Setting BI > 1.8 as a threshold, over 50% of universal cytokines were bimodally expressed, in contrast to 25% of the Th2-specific cytokines (Fig. 2B). The average BI of universal cytokines (BI = 1.46) was also significantly higher than that of the Th2-specific cytokines (BI = 0.92) (Fig. 2C). Together, these data demonstrate that the universal cytokines more likely exhibit a unique bimodal expression pattern among single cells and contribute most to the cell-to-cell variation in Th2 cells.

**Figure 2. Universal cytokines are bimodally expressed among Th2 cells.** A, violin plots depicting expression distribution of housekeeping genes (first row), universal cytokines (second row), Th2-specific cytokines (third row), and Th2-specific transcription factors (bottom row) among 55 single cells. The BI is indicated above the violin plots for each gene. Th2-specific genes were defined as those with microarray expression value in Th2 cells being at least 4-fold higher than in either Th1 or Th17 cells. Universal cytokines were defined as being abundantly expressed in all Th subsets (microarray expression value ≥ 16 in Th1, Th2, and Th17 cells). The microarray data of gene expression of Th1, Th2, and Th17 cells were derived from GSE11924 (40). B and C, universal cytokines preferentially exhibit a bimodal expression pattern compared with Th2-specific cytokines. B, percentages of cytokine genes with BI > 1.8 were calculated in universal cytokines or Th2-specific cytokines. C, BI of the Th2-specific or universal cytokines were plotted. The p value was calculated using a Mann-Whitney test. *p < 0.05. Error bars stand for the standard deviation of the mean.

**Figure 3. AFE promotes cellular heterogeneity.** A and B, genome browser tracks of scRNA-Seq results of IL4 (A) or CAPG (B) in 55 Th2 single cells and the 10,000-cell population (pool). Canonical and alternative first exons are annotated with arrows. C, global AFE was determined using MISO (Mixture of Isoforms) (41), and distribution of Ψ (percent spliced in) was plotted.

**Alternative promoter use contributes to cellular transcriptome heterogeneity**

Although our scRNA-Seq analysis revealed that IL-4 RNA was uniformly expressed among Th2 cells (Fig. 2A), IL-4 protein was only detectable in ~25% cells, as determined by flow cytometric analysis (Fig. S1A). To reconcile the discrepancy between IL-4 RNA and protein expression among single cells, we examined transcripts that originated from the Il4 locus in depth. Although over 80% of cells (45 of 55) had robust transcription from the Il4 locus (TPM > 100), only ~60% (34 of 55) abundantly expressed the full-length IL-4 mRNA (Il4fl) (Fig. 3A). The remaining cells predominantly expressed an IL-4 non-coding isoform (Il4nc) that utilized an alternative transcription start site between canonical exon 2 and exon 3 in the Il4 gene (Fig. 3A). The anti-IL-4 antibody was only able to detect IL4fl-expressing K562 cells but not IL4nc-expressing cells, confirming that IL4nc RNA is noncoding and undetectable by the anti-IL-4 antibody (Fig. S3, A and B). Sequence analysis further confirmed that, except for the first exon, the remaining sequence of IL4nc was identical to IL4fl. Indeed, six cells only expressed IL4nc but not IL4fl (IL4nc/IL4fl > 7). Thus, transcription of a noncoding RNA from the Il4 locus at least partially contributes to the fewer IL-4 protein–expressing cells compared with the RNA-expressing cells during *in vitro* Th2 cell differentiation.

Similar to *Il4*, the CAPG gene, which is specifically up-regulated during Th2 response to parasite infection (24), also utilized an alternative transcription start site in most (90%) cells (Fig. 3B). Analysis of alternative first exon (AFE) usage globally identified 1620 AFE events (percent spliced in, Ψ > 0.1 in at least 20% of single cells). Many of these alternative first exons were considerably utilized together with their canonical coun-
terparts (Ψ around 0.5) (Fig. 3C), indicating that utilization of alternative promoters is a general phenomenon for transcriptome diversification at the single-cell level.

**IL-4**<sub>nc</sub> **RNA promotes IL-4 production in Th2 cells**

The first exon of the **IL-4**<sub>nc</sub> RNA falls into an intragenic enhancer of the **Il4** locus (HSII) that regulates IL-4 but not IL-5 or IL-13 expression in T cells (25). Enhancer-derived RNAs have been widely reported and are usually associated with target gene activation (18, 26–28), although most of their functions remain unclear. The unique expression pattern of the **IL-4**<sub>nc</sub> RNA among single cells prompted us to explore its function in depth.

We first developed a TaqMan real-time PCR assay that was able to distinguish the **IL-4**<sub>nc</sub> RNA from the pre-mRNA of the IL-4 coding isoform (**IL-4**<sub>fl</sub>) by designing probes covering unique junctions between their first and the second exons, respectively (Fig. S3C). Reverse transcription using oligo(dT) primers indicated that the **IL-4**<sub>nc</sub> RNA was polyadenylated (Fig. S3D). Among all Th cell subsets, **IL-4**<sub>nc</sub> RNA was detectable in Th0 cells and Th2 cells but not in Th1, Th17, or Treg cells (Fig. 4A), consistent with reduced **Il4** locus accessibility in the latter lineages (29). Fractionation further showed that **IL-4**<sub>nc</sub> RNA predominantly localized in the cytosol rather than in the nucleus in either resting or restimulated Th2 cells (Fig. 4B). Thus, unlike most enhancer RNAs, **IL-4**<sub>nc</sub> RNA is a spliced and polyadenylated RNA mainly located in the cytosol.

One of the key characteristics of effector Th cells is their capacity to quickly produce a large amount of effector cytokines following antigen receptor stimulation. Prior to TCR restimulation, we found that the **IL-4**<sub>nc</sub> RNA was the predominant transcript from the **Il4** locus, which was expressed, on average, 4-fold higher than the **IL-4**<sub>fl</sub> RNA (Fig. 4C). Following TCR stimulation, induction of the **IL-4**<sub>nc</sub> RNA (peaked 1 h post-stimulation) also preceded the **IL-4**<sub>fl</sub> RNA (peaked 3 h post-stimulation). Despite quicker induction, **IL-4**<sub>nc</sub> RNA was only induced around 7-fold after TCR restimulation, much less than the magnitude of **IL-4**<sub>fl</sub> RNA (>200-fold induction) (Fig. 4D and Fig. S3E), rendering the **IL-4**<sub>nc</sub> isoform constituting 2% of total IL-4 transcripts 3 h after the TCR restimulation. These data together show that **IL-4**<sub>nc</sub> RNA is a primary transcript from the **Il4** locus in resting Th2 cells, whereas transcription shifts to **IL-4**<sub>fl</sub> RNA after TCR restimulation.

To examine connections between **IL-4**<sub>nc</sub> RNA and IL-4 production in Th2 cells, we transduced **IL-4**<sub>nc</sub> RNA in primary Th2 cells (Fig. 5A). **IL-4**<sub>nc</sub> promoted the production of IL-4 protein in Th2 cells over 40% compared with the control group, as determined by either flow cytometry (Fig. 5, B and C) or ELISA (Fig. 5D). Expression of other cytokines, including IL-13, IL-2, IL-10, and tumor necrosis factor α, was not significantly changed (Fig. 5, B, and C, and Fig. S3, F–H), indicating that **IL-4**<sub>nc</sub> RNA specifically regulates IL-4 production in Th2 cells. We further mapped sequence elements required for the **IL-4**<sub>nc</sub> RNA to promote IL-4 production. Only a mutant **IL-4**<sub>nc</sub> RNA missing the last exon, but not mutants missing either the first or the second exon, completely lost its capacity to facilitate IL-4 production (Fig. 5, E–G). Further truncation of the last exon showed that the sequence overlapping the 3' UTR of the coding isoform was necessary and sufficient for enhancing IL-4 production (Fig. S4). Interestingly, despite promoting IL-4 protein expression, the noncoding isoform did not affect the IL-4 mRNA level or its expression kinetics following TCR stimulation (Fig. 5, H and I), indicating that **IL-4**<sub>nc</sub> directly enhances translation of **IL-4**<sub>fl</sub> RNA.

Figure 4. **IL-4**<sub>nc</sub> is constitutively expressed in Th2 cells. A, **IL-4**<sub>nc</sub> was specifically expressed in Th2 and Th0 cells. Naïve CD4<sup>+</sup> T cells were activated with anti-CD3 and anti-CD28 and differentiated into various lineages under the respective conditions. Five days after differentiation, expression of **IL-4**<sub>nc</sub> was analyzed by real-time PCR either under steady state or after being activated by anti-CD3 and anti-CD28 or PAM and ionomycin for 5 h. B, **IL-4**<sub>nc</sub> is primarily located in the cytoplasm. D10.G4.1 cells were fractionated into the nuclear (N) fraction and cytoplasmic (C) fraction, and the presence of the **IL-4**<sub>nc</sub> isoform was determined by real-time PCR (left panel). The protein lysates from both fractions were further analyzed by immunoblotting for the cytoplasmic marker (B-actin) and nucleus marker (histone H3) (right panel). C, expression of either **IL-4**<sub>nc</sub> or **IL-4**<sub>fl</sub> in resting in vitro differentiated Th2 cells was determined by real-time PCR. D, induction of **IL-4**<sub>nc</sub> RNA precedes **IL-4**<sub>fl</sub> following TCR stimulation. Th2 cells were differentiated as described and restimulated with anti-CD3/anti-CD28 for the indicated time. Expression of either **IL-4**<sub>nc</sub> or **IL-4**<sub>fl</sub> was determined by real-time PCR. Data are representative (B) or the sum (A–D) of at least three independent experiments. Error bars stand for the standard deviation of the mean. **, p < 0.01 in paired Student's t test. ns, not significantly changed.
Conversely, abrogating IL4 nc RNA inhibited IL-4 protein production in Th2 cells. We designed two independent shRNAs targeting the unique region of IL4 nc RNA (its first exon) and transduced the two shRNAs into differentiating Th2 cells. Either shRNA decreased ~50% of IL-4 noncoding RNA without affecting the expression of IL-4–coding mRNA (Fig. 6D), confirming the specificity of these shRNAs. Despite an unaltered IL4 fl mRNA level, production of the IL-4 protein was reduced over 40% in shRNA-transduced cells compared with control cells, as determined by either flow cytometry (Fig. 6, A and B) or ELISA (Fig. 6C). Similarly, in murine Th2 clone D10 cells (30, 31), knockdown of IL4 nc suppressed IL-4 protein production without changing IL4 fl mRNA expression (Fig. S5, A–D).

Finally, we examined the translation efficiencies of IL4 fl mRNA in IL4 nc–deficient Th2 cells using polysomal profiling assays. Compared with control cells, IL-4 mRNA in polysomal fractions was decreased in the IL4 nc–deficient cells (Fig. 6E). Conversely, overexpression of the IL4 nc transcript in Th2 cells increased the ratios of the IL-4 mRNA in polysomal fractions (Fig. 6F). Distribution of ribosome RNAs was not affected by either IL4 nc knockdown or overexpression (Fig. S6). These data together indicate that IL4 nc RNA promotes production of IL-4 by enhancing its mRNA translation, consistent with its predominant cytoplasmic localization.

Discussion

In this study, we applied scRNA-seq to characterize the transcriptome complexity of Th2 cells. Our results revealed that variable cytokine expression contributed most to the cell-to-cell heterogeneity of in vitro differentiated Th2 cells. We also found that alternative promoter use was an underlying mechanism of cellular heterogeneity. Moreover, we identified that an intragenic enhancer drives expression of an IL-4 noncoding RNA.
IL-4 mRNA translation. Antigen stimulation shifts transcription of the noncoding isoform to the canonical coding IL-4 mRNA, and the noncoding isoform promotes translation of the coding IL-4 mRNA. Our study thus exemplifies that an enhancer-derived noncoding RNA promotes the translation of coding RNA from the same locus.

We chose to analyze the in vitro differentiated Th2 cells using the Fluidigm platform. Compared with recently published scRNA-seq results of in vivo differentiated Th2 cells (32, 33), our analysis likely reflects intrinsic cellular heterogeneity. One of the unexpected findings is that universal cytokines, but not Th2-specific cytokines, were bimodally expressed among single Th2 cells. Bimodal distribution among single cells often suggests distinct functions or response to external stimulation, suggesting that these cytokines respond digitally to the TCR stimulation. A similar digital response has been observed in signaling pathways downstream of TCR stimulation (i.e., phosphorylation of extracellular signal-regulated kinase) (34), suggesting that these universal cytokines are directly downstream of TCR signaling events. Other processes, including stochastic effects, the cell cycle, or the frequencies of cell division from naïve CD4+ T cells, may also influence the expression pattern of these universal cytokines.

Although the Fluidigm platform was only able to capture a small number of cells, it enabled analysis of full-length RNA at the single-cell level, which is unachievable with droplet-based platforms for single-cell RNA-Seq (35–37). For instance, IL4nc would be undistinguishable from IL-4 full-length mRNA if only the 3’ end of RNA had been sequenced. Even with a relatively small number of cells, our data revealed heterogeneity of transcripts from the IL-4 locus. Given that over 30% of Th2 cells expressed the IL4nc transcript, our current sequence depth and cell number sufficiently reflected the cellular heterogeneity of in vitro differentiated Th2 cells. The relatively abundant IL4nc transcript was also confirmed by our TaqMan real-time PCR assays.

Previous studies have shown that locus accessibility contributes greatly to the variation of Th2 gene expression (38, 39). Our study provides another layer of transcriptional complexity. An alternative promoter-derived noncoding RNA is constitutively expressed and able to sequester negative regulatory elements of RNA translation, which may enable rapid production of IL-4 protein right after antigen stimulation.

Our data also suggest that the long-observed lower efficiency of Th2 cell differentiation compared with Th1 is probably caused by the transcription of IL4nc in a subset of cells. Flow cytometry and intracellular staining of IL-4 protein have been the most widely used methods to detect IL-4-expressing cells, which, however miss IL4nc-expressing cells. Moreover, it has long been observed that naïve CD4+ T cells express IL-4 mRNA but not IL-4 protein after TCR stimulation, and our analysis suggests that one reason for this is that naïve CD4+ T cells only express IL4nc. It will be interesting to explore whether a similar mechanism exists for other cellular differentiation processes.

**Experimental procedures**

*Mice*

C57BL/6 mice were obtained from the Shanghai Laboratory Animal Center, Chinese Academy of Sciences. All mice were maintained under specific pathogen-free conditions and used at 8–10 weeks of age. All animal studies were performed in compliance with the Guide for the Care and Use of Laboratory Animals and approved by the Institutional Biomedical Research Ethics Committee of the Shanghai Institutes for Biological Sciences.

*Cell lines*

Plat-E cells were obtained from the Cell Bank of the Type Culture Collection of the Chinese Academy of Sciences (Shanghai, China) and cultured in Dulbecco’s modified Eagle medium (HyClone) with 10% FBS (HyClone). D10.G4.1 cells were obtained from the ATCC and cultured in RPMI 1640 medium (HyClone) with 10% FBS, 1× antibiotics, 1× nonessential amino acids, 50 μM β-mercaptoethanol, 10 ng/ml IL-1β, and antibiotics, 1× nonessential amino acids, 50 μM β-mercaptoethanol, 10 ng/ml IL-1β,
IL4 non-coding RNA promotes IL-4 mRNA translation

and 100 units/ml IL-2. K562 cells were obtained from the Cell Bank of the Type Culture Collection of the Chinese Academy of Sciences and cultured in Iscove’s modified Dulbecco’s medium (Gibco) with 10% FBS (Hyclone), 1× glutamine, 1× penicillin, 1× streptomycin, and 50 µM β-mercaptoethanol.

Single-cell and bulk RNA-Seq

Single cells were isolated with a Fluidigm C1 system. A total of 56 cells were collected and sequenced with an Illumina HiSeq system. For bulk RNA-Seq, 10,000 pooled cells were prepared using a standard protocol for the RNA-Seq library. All fastq files were deposited in the GEO database under accession number GSE103647. The microarray data of gene expression of Th1, Th2, and Th17 cells were deposited in the GEO database under accession number GSE11924.

Bimodal index analysis

The BI of gene expression was calculated as described previously (23). Briefly, if a gene is expected to have two populations in the RNA expression level by the model, then the mean expression values of the two populations (µ1 and µ2), the common standard deviation (σ), and the percentage of one subset (π) are predicted, and then the standardized distance between two subsets (δ) and BI are defined as

\[ \delta = \left( \frac{\mu_1 - \mu_2}{\sigma} \right) \] (Eq. 1)

\[ BI = \left[ \pi(1 - \pi) \right]^{1/2} \delta \] (Eq. 2)

RNA extraction, cDNA synthesis, and quantitative real-time PCR

Total RNA was isolated with the E.N.Z.A® HP Total RNA Kit (Omega), and cDNA was synthesized using the PrimeScript™ RT Reagent Kit with gDNA Eraser (Takara). Quantitative PCR was performed using SYBR Premix Ex Taq (Takara) on a StepOnePlus real-time PCR system (Applied Biosystems). Gene expression was normalized to glyceraldehyde-3-phosphate dehydrogenase.

Intracellular cytokine staining and flow cytometry

Cells were stained with phorbol 12-myristate 13-acetate (PMA, Sigma, 50 ng/ml) and ionomycin (Sigma, 500 ng/ml) in the presence of Golgi-plug (BD Biosciences) for 5 h at 37 °C. Cells were stained with the LIVE/DEAD® Fixable Violet Dead Cell Stain Kit (Thermo Fisher), fixed/permeabilized with the Cytofix/Cytoperm Kit (BD Biosciences), and then stained with antibodies against the indicated cytokines. Fluorescence-labeled anti-IL-4 and anti-IL-13 antibodies were from eBioscience.

CD4+ T cell isolation and differentiation

CD4+ T cells were purified from the spleens of C57BL/6 mice using a magnetic biotin selection kit (Stem Cell). Purified CD4+ T cells were activated with plate-bound anti-hamster IgG (MP Biomedicals) and soluble anti-CD3 (5 µg/ml, 145-2C11, Bio X Cell) and anti-CD28 (1 µg/ml, 37N, Bio X Cell) in T cell medium (Iscove’s modified Dulbecco’s medium, 10% FBS, 1× glutamine, 1× penicillin, 1× streptomycin, and 50 µM β-mercaptoethanol), and defined as the T4,0 condition. For other sub-sets, various combinations of antibodies and cytokines were added as follows: T4,1 conditions, anti-IL-4 (5 µg/ml, 11B11, Bio X Cell) and IL-12 (2 ng/ml, PeproTech); T4,2 conditions, anti-IFN-γ (5 µg/ml, Bio X Cell) and IL-4 (10 ng/ml, PeproTech); T4,17 conditions, transforming growth factor β1 (1 ng/ml, PeproTech), IL-6 (20 ng/ml, PeproTech), anti-IL-4 (5 µg/ml), and anti-IFN-γ (5 µg/ml); Treg conditions, IL-2 (100 units/ml, PeproTech), TGF-β1 (5 ng/ml), anti-IL-4 (5 µg/ml), and anti-IFN-γ (5 µg/ml). 2 days later, cells were suspended from plate and IL-2 (100 units/ml) were added for T4,0, T4,1 and T4,2; IL-1β (10 ng/ml, PeproTech) and IL-23 (10 ng/ml, PeproTech) were added for T4,17. Cells were cultured for 5 days.

Retroviral infection

Retrovirus was produced by transfecting Plat-E cells with the indicated plasmids, and supernatant containing fresh virus was collected 48 h and 72 h after transfection. To infect K562 cells and D10.G4.1 cells, cells were suspended with fresh retrovirus, spin-infected for 1.5 h at 900 × g at 37 °C in the presence of Polybrene (8 µg/ml), and cultured at 37 °C for another 6 h before recovering to culture medium. Then puromycin (3 µg/ml) was added for at least 3 days before analysis. To infect T4,2 cells, naïve CD4+ T cells were first stimulated with anti-CD3, anti-CD28, anti-IFN-γ, and IL-4 as mentioned above. For the first infection, 80% of the stimulation medium was replaced with medium containing fresh retrovirus 18 h after stimulation. Then the CD4+ T cells were spin-infected for 1.5 h at 900 × g at 37 °C in the presence of Polybrene (8 µg/ml) and cultured at 37 °C for another 2 h before recovering to the stimulation medium. The second infection was carried out 40 h after stimulation the same as the first infection.

ELISA assay

Cells were stimulated with PMA (Sigma, 50 ng/ml) and ionomycin (Sigma, 500 ng/ml) for 5 h at 37 °C. The IL-4 levels in the collected supernatants were analyzed using a mouse IL-4 ELISA kit (R&D Systems, DY404-05).

Plasmids

For knockdown of IL-4 noncoding RNA, shRNA sequences targeting mouse IL-4 noncoding RNA or a scramble sequence of the same length was inserted into the MSCV-LTRmiR30-PIG (LMP) vector. The LMP and pcl-eco plasmids were cotransfected into Plat-E to produce retrovirus. For the purpose of overexpression, the IL-4 coding sequence and IL-4 noncoding RNA and its truncated isoforms were inserted into the MSCV-PIG vector.

Polysomal profiling

After cycloheximide treatment (100 µg/ml), 2.5 × 107 Th2 cells were collected and lysed in 1.3 ml of lysis buffer (100 mM KCl, 5 mM MgCl2, 20 mM Tris-HCl (pH 7.5), 0.03% TritonX-100, 1 mM dithiothreitol, 40 units/ml RNase inhibitor (Takara), and protease inhibitor (Roche)) on ice for 30 min. Then cells were spun for 10 min at 12,000 rpm (13,800 × g) at 4 °C, and supernatant was collected. 10%, 20%, 30%, 40%, and 50% sucrose buffers were prepared in lysis buffer. To obtain a linear

296 J. Biol. Chem. (2019) 294(1) 290–298
gradient, 1.5 ml of each sucrose buffer was added into layers with the heavy sucrose solution at the bottom, and the gradient was preserved at 4 °C overnight. The cell lysate was added to the top of the gradient. After ultracentrifugation (35,000 rpm for 2 h in a SW40Ti rotor at 4 °C), 19 fractions from the sucrose gradient were collected with the same volume. The RNA extraction, cDNA synthesis, and quantitative real-time PCR were conducted as mentioned above.

Author contributions—W. Y. formal analysis; W. Y. and X. C. investigation; W. Y. and X. C. writing-original draft; W. Y. and X. C. project administration; Y. S. data curation; X. C. conceptualization; X. C. supervision; X. C. funding acquisition; X. C. writing-review and editing.

Acknowledgments—We thank Shiqi Tu and Ping Lin for help with analysis of RNA-Seq data and Yu Zhang for assisting with ultracentrifugation.

References

1. Zhu, J., Yamane, H., and Paul, W. E. (2010) Differentiation of effector CD4 T cell populations (*). Annu. Rev. Immunol. 28, 445–489 CrossRef Medline
2. Murphy, K. M., and Reiner, S. L. (2002) The lineage decisions of helper T cells. Nat. Rev. Immunol. 2, 933–944 CrossRef Medline
3. Zhu, J., Yamane, H., Cote-Sierra, J., Guo, L., and Paul, W. E. (2006) GATA-3 promotes Th2 responses through three different mechanisms: induction of Th2 cytokine production, selective growth of Th2 cells and inhibition of Th1 cell-specific factors. Cell Res. 16, 3–10 CrossRef Medline
4. Szabo, S. J., Kim, S. T., Costa, G. L., Zhang, X., Fathman, C. G., and Glimcher, L. H. (2000) A novel transcription factor, T-bet, directs Th1 lineage commitment. Cell 100, 655–669 CrossRef Medline
5. Abbas, A. K., Murphy, K. M., and Sher, A. (1996) Functional diversity of helper T lymphocytes. Nature 383, 787–793 CrossRef Medline
6. Ho, I. C., and Glimcher, L. H. (2002) Transcription: tantalizing times for T cells. Cell 109, S109–S120 CrossRef Medline
7. Le Gros, G., Ben-Sasson, S. Z., Seder, R., Finkelman, F. D., and Paul, W. E. (1990) Generation of interleukin 4 (IL-4)-producing cells in vivo and in vitro: IL-2 and IL-4 are required for in vitro generation of IL-4-producing cells. J. Exp. Med. 172, 921–929 CrossRef Medline
8. Swain, S. L., Weinberg, A. D., English, M., and Huston, G. (1990) IL-4 directs the development of Th2-like helper effectors. J. Immunol. 145, 3796–3806 Medline
9. Yamane, H., Zhu, J., and Paul, W. E. (2005) Independent roles for IL-2 and GATA-3 in stimulating naive CD4+ T cells to generate a Th2-inducing cytokine environment. J. Exp. Med. 202, 793–804 CrossRef Medline
10. Bancroft, A. I., Artis, D., Donaldson, D. D., Sypek, J. P., and Grecis, R. K. (2000) Gastrointestinal nematode expulsion in IL-4 knockout mice is IL-13-dependent. Eur. J. Immunol. 30, 2083–2091 CrossRef Medline
11. Licosa-Limón, P., Kim, L. K., Palm, N. W., and Flavell, R. A. (2013) TH2, allergy and group 2 innate lymphoid cells. Nature Immunol. 14, 536–542 CrossRef Medline
12. Kubo, M. (2017) T follicular helper and TH2 cells in allergic responses. Allergol Int. 66, 377–381 CrossRef Medline
13. Paul, W. E. (2015) History of interleukin-4. Cytokine 75, 3–7 CrossRef Medline
14. Kühn, R., Rajewsky, K., and Müller, W. (1991) Generation and analysis of interleukin-4-deficient mice. Science 254, 707–710 CrossRef Medline
15. Noben-Trauth, N., Shultz, L. D., Brombacher, F., Urban, J. F., Jr., Gu, H., and Paul, W. E. (1997) An interleukin 4 (IL-4)-independent pathway for CD4+ T cell IL-4 production is revealed in IL-4 receptor-deficient mice. Proc. Natl. Acad. Sci. U.S.A. 94, 10838–10843 CrossRef Medline
16. de Klerk, E., and ‘t Hoen, P. A. (2015) Alternative mRNA transcription, processing, and translation: insights from RNA sequencing. Trends Genet. 31, 128–139 CrossRef Medline
17. Quinn, J. I., and Chang, H. Y. (2016) Unique features of long non-coding RNA biogenesis and function. Nat. Rev. Genetics 17, 47–62 CrossRef Medline
18. Grubau, J., Diderich, K., Pruunia, S., Calzolari, O., and Fraser, P. (2000) Intragenic transcription and developmental remodeling of chromatin subdomains in the human β-globin locus. Mol. Cell 5, 377–386 CrossRef Medline
19. Fu, J. D., Stone, N. R., Liu, L., Spencer, C. I., Qian, L., Hayashi, Y., Delgado-Olguin, P., Ding, S., Bruneau, B. G., and Srivastava, D. (2013) Direct reprogramming of human fibroblasts toward a cardiomyocyte-like state. Stem Cell Rep. 1, 235–247 CrossRef Medline
20. Brenecke, P., Anders, S., Kim, J. K., Kolodziejczyk, A. A., Zhang, X., Proserpio, V., Baying, B., Benes, V., Teichmann, S. A., Marioni, J. C., and Heisler, M. G. (2013) Accounting for technical noise in single-cell RNA-seq experiments. Nat. Methods 10, 1093–1095 CrossRef Medline
21. Shalek, A. K., Satija, R., Adiconis, X., Gertner, R. S., Gaublomme, J. T., Raychowdhury, R., Schwartz, S., Yosef, N., Malboeuf, C., Lu, D., Trombetta, J. J., Gennert, D., G nikre, A., Goren, A., Hacohen, N., et al. (2013) Single-cell transcriptomics reveals bimodality in expression and splicing in immune cells. Nature 498, 236–240 CrossRef Medline
22. Zheng, W., and Flavell, R. A. (1997) The transcription factor GATA-3 is necessary and sufficient for Th2 cytokine gene expression in CD4 T cells. Cell 89, 587–596 CrossRef Medline
23. Wang, J., Wen, S., Symmans, W. F., Pusztai, L., and Coombes, K. R. (2009) The bimodality index: a criterion for discovering and ranking bimodal signatures from cancer gene expression profiling data. Cancer Inform. 7, 199–216 Medline
24. Kang, S. A., Park, M. K., Cho, M. K., Park, S. K., Jang, M. S., Yang, B. G., Kang, M. H., Kim, D. H., and Yu, H. S. (2014) Parasitic nematode-induced CD4+Foxp3+ T cells can ameliorate allergic airway inflammation. PLoS Negl. Trop. Dis. 8, e3410 CrossRef Medline
25. Tanaka, S., Motomura, Y., Suzuki, Y., Yagi, R., Inoue, H., Miyatake, S., and Kubo, M. (2011) The enhancer HS2 critically regulates GATA-3-mediated Il4 expression in Th2 cells. Nat. Immunol. 12, 77–85 CrossRef Medline
26. Li, W., Notani, D., and Rosenfeld, M. G. (2016) Enhancers as non-coding RNA transcription units: recent insights and future perspectives. Nat. Rev. Genet. 17, 207–223 CrossRef Medline
27. Onodera, C. S., Underwood, J. G., Katzman, S., Jacobs, F., Greenberg, D., Salama, S. R., and Haussler, D. (2012) Gene isoform specificity through enhancer-associated antisense transcription. PloS ONE 7, e35111 CrossRef Medline
28. Cech, T. R., and Steitz, J. A. (2014) The noncoding RNA revolution—trash or treasure? Cell 157, 77–94 CrossRef Medline
29. Ansell, K. M., Djuretic, I., Tanasa, B., and Rao, A. (2006) Regulation of Th2 differentiation and IL4 locus accessibility. Annu. Rev. Immunol. 24, 607–656 CrossRef Medline
30. Ojeda, G., Ronda, M., Ballester, S., Diez-Orejas, R., Feito, M., García-Albert, L., Rojo, J. M., and Montandon, R., et al. (2019) 294(1) 290–298 CrossRef Medline
31. Agarwal, S., and Rao, A. (1998) Modulation of chromatin structure regulates cytokine gene expression during T cell differentiation. Immunity 9, 765–775 CrossRef Medline
32. Proserpio, V., Piccolo, A., Haim-Vilmovsky, L., Kar, G., Lönning, T., Svensson, V., Pramanik, J., Natarajan, K. N., Zhai, W., Zhang, X., Donati, G., Kayikci, M., Kotar, J., McKenzie, A. N., Montandon, R., et al. (2016) Single-cell analysis of CD4+ T-cell differentiation reveals three major cell states and progressive acceleration of proliferation. Genome Biol. 17, 103 CrossRef Medline
33. Mahata, B., Zhang, X., Kolodziejczyk, A. A., Proserpio, V., Haim-Vilmovsky, L., Taylor, A. E., Hebenstreit, D., Dingler, F. A., Moingard, V., Göttgens, B., Arlt, W., McKenzie, A. N., and Teichmann, S. A. (2014) Single-cell RNA sequencing reveals T helper cells synthesizing steroids de novo to contribute to immune homeostasis. Cell Rep. 7, 1130–1142 CrossRef Medline
34. Altan-Bonnet, G., and Germain, R. N. (2005) Modeling T cell antigen discrimination based on feedback control of digital ERK responses. PLoS Biol. 3, e356 CrossRef Medline

IL4 non-coding RNA promotes IL-4 mRNA translation

J. Biol. Chem. (2019) 294(1) 290–298 297
35. Zheng, G. X., Terry, J. M., Belgrader, P., Ryykin, P., Bent, Z. W., Wilson, R., Ziraldo, S. B., Wheeler, T. D., McDermott, G. P., Zhu, J., Gregory, M. T., Shu, J., Montesclaros, L., Underwood, J. G., Masquelier, D. A., et al. (2017) Massively parallel digital transcriptional profiling of single cells. Nat. Commun. 8, 14049 CrossRef Medline

36. Klein, A. M., Mazutis, L., Akartuna, I., Tallapragada, N., Veres, A., Li, V., Peshkin, L., Weitz, D. A., and Kirschner, M. W. (2015) Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. Cell 161, 1187–1201 CrossRef Medline

37. Macosko, E. Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tiros, L., Bialas, A. R., Kamitaki, N., Martersteck, E. M., Trombetta, J. J., Weitz, D. A., Sanes, J. R., Shalek, A. K., Regev, A., and McCarroll, S. A. (2015) Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. Cell 161, 1202–1214 CrossRef Medline

38. Guo, L., Hu-Li, J., and Paul, W. E. (2004) Probabilistic regulation of IL-4 production in Th2 cells: accessibility at the Il4 locus. Immunity 20, 193–203 CrossRef Medline

39. Guo, L., Hu-Li, J., Zhu, J., Watson, C. J., Difilippantonio, M. J., Pannetier, C., and Paul, W. E. (2002) In TH2 cells the Il4 gene has a series of accessibility states associated with distinctive probabilities of IL-4 production. Proc. Natl. Acad. Sci. U.S.A. 99, 10623–10628 CrossRef Medline

40. Nurieva, R. I., Chung, Y., Hwang, D., Yang, X. O., Kang, H. S., Ma, L., Wang, Y. H., Watowich, S. S., Jetten, A. M., Tian, Q., and Dong, C. (2008) Generation of T follicular helper cells is mediated by interleukin-21 but independent of T helper 1, 2, or 17 cell lineages. Immunity 29, 138–149 CrossRef Medline

41. Katz, Y., Wang, E. T., Airoldi, E. M., and Burge, C. B. (2010) Analysis and design of RNA sequencing experiments for identifying isoform regulation. Nat. Methods 7, 1009–1015 CrossRef Medline