

[DOI: 10.1002/eji.201545646]

## Epigenetic analysis of regulatory T cells using multiplex bisulfite sequencing

A subset of regulatory CD4<sup>+</sup> T cells (Tregs) is characterized by the stable constitutive expression of the transcription factor FOXP3. Demethylation at a conserved region within intron 1 of *FOXP3*, the Treg-specific demethylated region (TSDR), is exclusive to this subset of Tregs: other immune cells that do not express FOXP3 or express the transcription factor transiently, such as activated effector T cells or TGF- $\beta$ -treated CD4<sup>+</sup> T cells, are methylated at the TSDR [1–4]. Hence, the TSDR provides a specific target to enumerate Tregs within biological samples; developing methods that enable robust Treg-cell quantitation with small numbers of cells from clinical samples remains an important goal. Current methods to measure *FOXP3* demethylation include qPCR [5, 6] and epigenetic sequencing methylation analysis [7] of Sanger sequencing traces [8], neither of which capture the methylation information at each CpG site on a single piece of DNA, but rather produce an average methylation over the whole region or an average at each site. Although several studies have used a PCR cloning and sequencing method to assess the methylation of each CpG site within the TSDR, this is a laborious method and normally less than 50 clones per sample are analyzed [3, 9, 10].

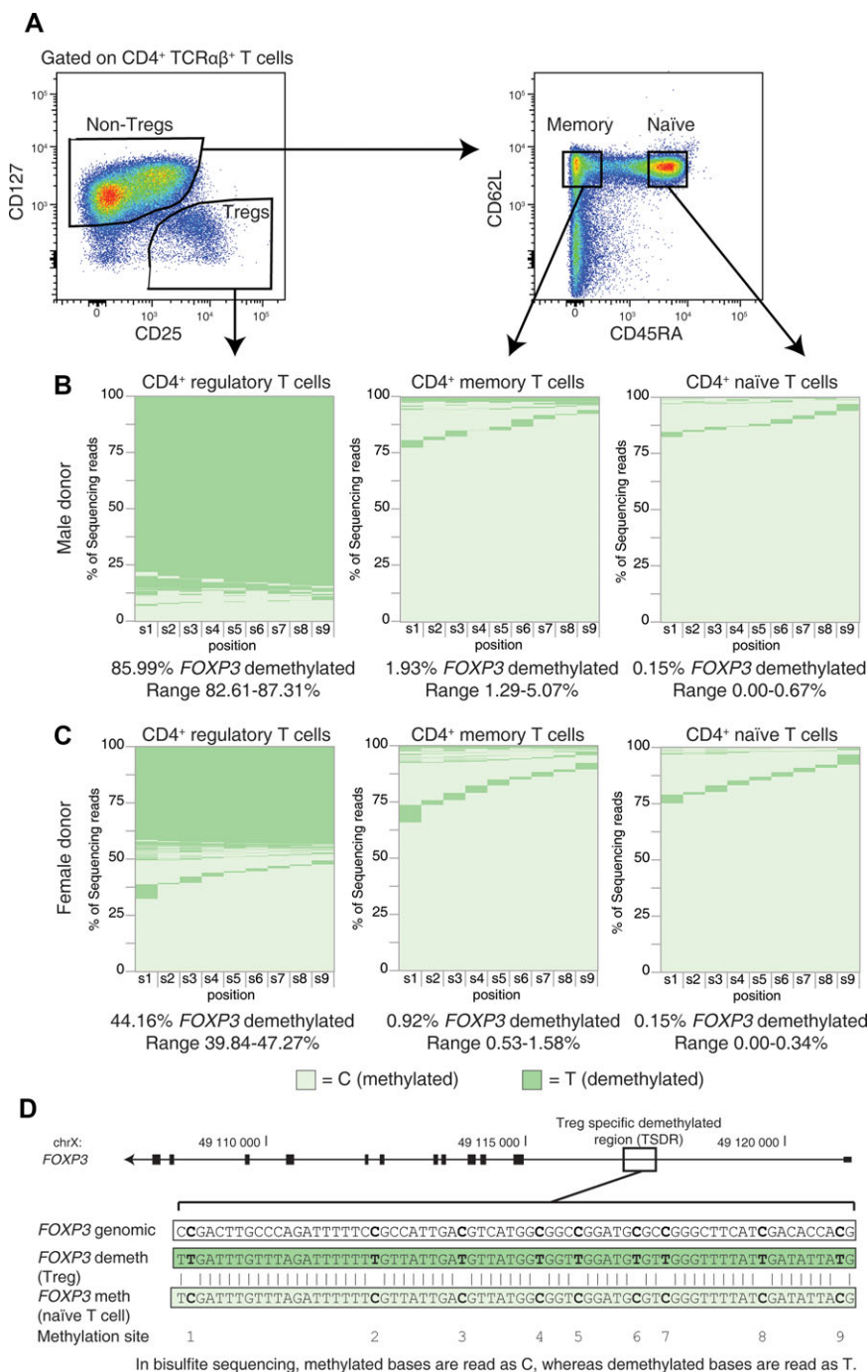
We therefore developed a next-generation sequencing (NGS) method to assess at single-base resolution the methylation status of the *FOXP3* TSDR. The method, which we originally developed to genotype rs1800521 single nucleotide polymorphism in the gene *AIRE* (Methods detailed as Supporting Information), reports at single base resolution the methylation of each DNA amplicon, provides hundreds to thousands of reads per replicate and can be multiplexed to analyze several DNA regions simultaneously.

Our method uses two-stage PCR: the first is a gene-specific PCR with primers having an adaptor sequence that is targeted by a second round PCR, which adds a unique index sequence enabling up to 960 different samples to be processed per sequencing run. We targeted nine CpG sites within the TSDR (Supporting Information Fig. 1), which have also been used in qPCR assays [6]. Our primers flank the TSDR alleviating the need for methylation-specific primers as well as the demethylated and methylated standards required in qPCR assays. We also developed a bioinformatics processing pipeline that reads the raw output from sequencing and reports the methylation status at each CpG (<https://github.com/XinYang6699/Methpup>). To validate our method, we sequenced the TSDR in three CD4<sup>+</sup> T-cell subsets, Tregs and naïve and memory effector cells, which were defined by surface molecules CD4, CD25, CD127 [11] and purified by flow sorting from two male and two female

donors (Fig. 1, Supporting Information Fig. 2, Supporting Information Table 1). We observed a linear relationship between the proportion of Tregs and the proportion of *FOXP3* demethylation in defined mixtures of Tregs and naïve CD4<sup>+</sup> T cells ( $r^2 = 0.99$ , Supporting Information Fig. 3). Cell inputs below 2000 cells/PCR replicate showed increasing assay variation so unless cells are extremely limiting, we typically utilize 2000–4000 cells per replicate and six replicates per cell sample (Supporting Information Fig. 4).

As expected based on previous publications where the TSDR was shown using multiple technologies to be demethylated in a majority of Tregs [1, 3, 8], 86% and 62% of the TSDR sequences were demethylated in Treg samples from male donors and 44% and 40% from female donors (Fig. 1, Supporting Information Fig. 2, Supporting Information Table 1). Owing to X inactivation, demethylation of the *FOXP3* TSDR occurs on only one copy of the X chromosome in females thereby explaining the ~50% lower demethylation level as compared to males. Although in Tregs the majority of highly demethylated sequencing reads are demethylated at all nine CpG sites considered, a proportion (<10%) of reads have only eight CpG sites demethylated. This incomplete demethylation could be due to biological effects or to incomplete bisulfite conversion. We favor the former possibility since bisulfite conversion efficiency as determined by cytosines that cannot be methylated was greater than 99% (Supporting Information Fig. 5A). Underscoring this biological heterogeneity, in naïve and memory effector cells demethylation of a single CpG among the nine sites was observed at each

\*These authors contributed equally to this work.

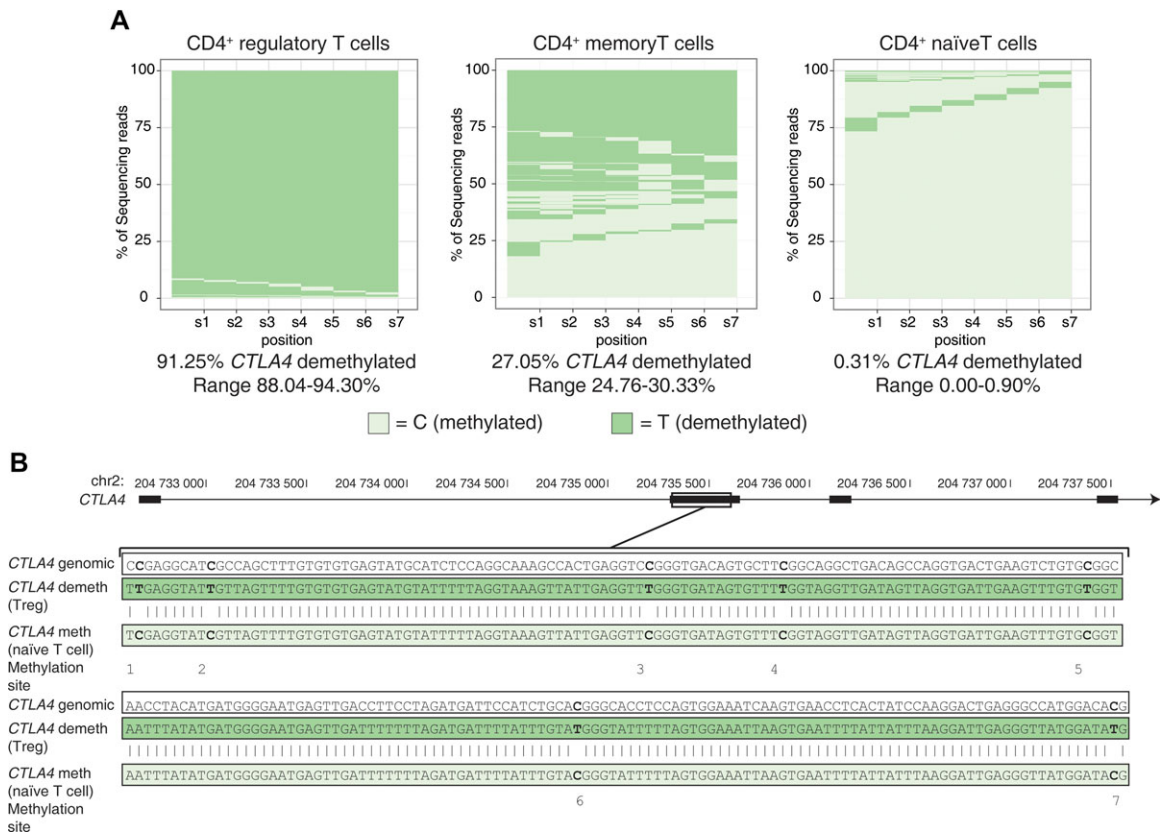


**Figure 1.** FOXP3 TSDR sequencing of purified human CD4<sup>+</sup> T-cell subsets. (A) Naïve (CD4<sup>+</sup>, CD127<sup>+</sup>, CD45RA<sup>-</sup>, CD62L<sup>+</sup>) and memory (CD4<sup>+</sup>, CD127<sup>+</sup>, CD45RA<sup>-</sup>, CD62L<sup>+</sup>) effector cells and Tregs (CD4<sup>+</sup>, CD25<sup>+</sup>, CD127<sup>-</sup>) were sorted from a male (B) and a female (C) donor by flow cytometry and analyzed for FOXP3 TSDR methylation (two representative donors out of a total of four, for more information see Supporting Information Fig. 2 and Supporting Information Table 1). See Supporting Information Fig. 7 for gating strategy defining CD4<sup>+</sup> TCRαβ<sup>+</sup> T cells. The x axis shows the nine methylation sites analyzed and each row indicates the methylation status of one copy of the sequenced TSDR. Light green represents a C (methylated) and dark green represents a T (demethylated). The y axis is the percentage of sequencing reads for each methylation/demethylation pattern. Six replicates of 3000 cells per replicate (5 ng DNA) were analyzed from a single donor and the median (with range) reads with eight or nine sites demethylated at FOXP3 for the replicates is shown below each graph. (D) A schematic of the FOXP3 gene and the TSDR sequence are shown.

position interrogated and, when combined, comprised ~20% of the reads. Thus the use of NGS to quantify demethylation has the advantage of revealing heterogeneity inherent in the biological process of gene regulation. As expected, and in contrast to Tregs, less than 2% of the TSDR reads from naïve and memory effector cells had a Treg-like TSDR methylation pattern. Although this low level of

Treg-specific sequences could be caused by suboptimal cell purification, Tregs with low levels of surface CD25 and a demethylated TSDR have been reported [12] and could account for the low percentage of Treg-specific reads observed in effector T-cell subsets sorted by cell surface markers. The NGS-based demethylation assay offers a great advantage in detecting small subpopulations of cells demethylated at

the FOXP3 TSDR as compared to methods that generate an average demethylation status at each residue since each NGS read simultaneously assesses the methylation status of all CpG sites from a single copy of the gene giving confidence in the interpretation of a demethylated read. Further characterization of rare subsets requires more detailed sorting in future studies including intracellular staining of



**Figure 2.** CTLA4 demethylation in human FACS-purified cells. Cells were purified and analyzed as detailed in Fig. 1 for CTLA4 methylation status. The median (with range) reads with all seven sites demethylated at CTLA4 for the replicates is shown below each graph. Data are representative of four donors (for more information see, Supporting Information Fig. 2 and Supporting Information Table 2).

FOXP3, a procedure that is compatible with NGS-based analysis of bisulfite converted DNA [13].

Because the current version of our NGS sequencing platform utilizes up to 960 unique index sequences per sequencing run of approximately 20 million reads, each index is associated with approximately 20 000 reads. In most cases, 20 000 reads per replicate is in excess of what is required for robust measurements of percent demethylation of any particular region thereby allowing for the multiplexing of gene regions within the same sample. We used multiplexing of up to six gene regions to search for an autosomal TSDR since female samples have 50% less demethylated FOXP3 TSDR reads than male samples. In mouse studies, specific regions in multiple genes in addition to *Foxp3* were identified that showed Treg-specific demethylation when compared to conventional CD4<sup>+</sup> T cells analyzed ex vivo: *Ctla4*, *Il2ra*, *Gitr*, *Eos*, and *Helios*

[9]. In comparing the human and mouse sequences in the regions reported, only one strong homology was observed, that in exon 2 of mouse *Ctla4* and human *CTLA4* (Supporting Information Fig. 6). We developed a sequencing assay for this region in the human gene, multiplexed with the FOXP3 TSDR sequencing assay, and demonstrated that demethylation of exon 2 of human *CTLA4* is not Treg-specific. Although >90% of Tregs analyzed ex vivo were demethylated at all seven CpG sites examined, 21–38% of memory CD4<sup>+</sup> T cells were demethylated at these same sites (Fig. 2, Supporting Information Fig. 2, Supporting Information Table 2). In addition, memory effector cells displayed a heterogeneous demethylation pattern with 37–48% of reads having an intermediate methylation pattern (3T, 4T, 5T, or 6T). Thus demethylation of the human *CTLA4* exon 2 region reflects more than Treg lineage commitment.

To conclude, we developed a high-throughput method for analyzing methylation at single CpG resolution that uses low DNA input and can be multiplexed. Although we have used the method to measure Treg-specific methylation, it can be easily adapted to verify genome-wide methylation results. Finally, given recent data showing that many disease-associated SNPs are located in enhancer regions [14], the presence of allele-specific methylation detected using our high-throughput method could provide a mechanism for disease-associated gene regulation.

Daniel B. Rainbow<sup>\*1</sup>, Xin Yang<sup>\*1</sup>,  
 Oliver Burren<sup>1</sup>, Marcin L. Pekalski<sup>1</sup>,  
 Deborah J. Smyth<sup>1</sup>,  
 Marcus D. R. Klarqvist<sup>1</sup>,  
 Christopher J. Penkett<sup>2</sup>, Kim Brugger<sup>3</sup>,  
 Howard Martin<sup>3</sup>, John A. Todd<sup>1</sup>,  
 Chris Wallace<sup>1,4</sup> and Linda S. Wicker<sup>1</sup>

<sup>1</sup> JDRF/Wellcome Trust Diabetes and Inflammation Laboratory, Department of Medical Genetics, National Institute for Health Research (NIHR) Cambridge Biomedical Research Centre,

Cambridge Institute for Medical Research,  
University of Cambridge, Cambridge Biomedical  
Research Campus, Cambridge, UK

<sup>2</sup> Department of Haematology, University of  
Cambridge, Long Road Cambridge, UK

<sup>3</sup> East Anglian Medical Genetics Service, Molecular  
Genetics Laboratories, Level 6 Addenbrooke's  
Treatment Centre, Addenbrooke's Hospital,  
Cambridge, UK

<sup>4</sup> MRC Biostatistics Unit, Cambridge Institute of  
Public Health, Cambridge, UK

**Acknowledgements:** This work was supported by Wellcome Trust Grant 091157, JDRF Grant 9-2011-253, the National Institute for Health Research Cambridge Biomedical Research Centre (BRC), and Award P01AI039671 (to L.S.W. and J.A.T.) from the National Institute of Allergy and Infectious Diseases (NIAID). C.W. is supported by the Wellcome Trust (089989). The content of this article is solely the responsibility of the authors and does not necessarily represent the official views of NIAID or the National Institutes of Health. The Cambridge Institute for Medical Research is in receipt of Wellcome Trust Strategic Award 100140. We gratefully acknowledge the participation of all NIHR Cambridge BioResource volunteers. We thank the Cambridge BioResource staff for their help with volunteer recruitment. We thank members of the Cambridge BioResource SAB and Management Committee for their support of our study and the National Institute for Health Research Cambridge Biomedical Research Centre for funding. We thank Fay Rodger and Ruth Littleboy for running the Illumina MiSeq in the Molecular Genetics Laboratories, Addenbrooke's Hospital, Cambridge. This research was supported by the Cambridge NIHR BRC Cell Phenotyping Hub. In particular, we wish to thank Anna Petrunina Harrison, Simon McCallum, Christopher Bowman, Natalia Savinykh, Esther Perez and Jelena Markovic Djuric for their advice and support in cell sorting. We also thank Helen Stevens, Pamela Clarke,

Gillian Coleman, Sarah Dawson, Jennifer Denesha, Simon Duley, Meeta Maisuria-Armer, and Trupti Mistry for acquisition and preparation of samples.

DBR performed all the practical work, D.B.R. and X.Y. performed the data processing and analysis, M.P. provided the sorted cells, X.Y., O.B., and C.P. developed the Bioinformatics pipeline, D.J.S. and M.D.R.K. performed the AIRE analysis, K.B. and H.M. designed the PCR index system, D.B.R., X.Y., and L.S.W. drafted the manuscript, all authors contributed to the editing of the manuscript and L.S.W., J.A.T., and C.W. supervised the work.

**Conflict of interest:** The authors declare no financial or commercial conflict of interest.

## References

- 1 Baron, U. DNA demethylation in the human FOXP3 locus discriminates regulatory T cells from activated FOXP3(+) conventional T cells. *Eur. J. Immunol.* 2007. **37**: 2378–2389.
- 2 Floess, S. Epigenetic control of the foxp3 locus in regulatory T cells. *PLoS Biol.* 2007. **5**: e38.
- 3 Miyara, M. Functional delineation and differentiation dynamics of human CD4+ T cells expressing the FoxP3 transcription factor. *Immunity* 2009. **30**: 899–911.
- 4 Toker, A. Active demethylation of the Foxp3 locus leads to the generation of stable regulatory T cells within the thymus. *J. Immunol.* 2013. **190**: 3180–3188.
- 5 Stockis, J. Comparison of stable human Treg and Th clones by transcriptional profiling. *Eur. J. Immunol.* 2009. **39**: 869–882.
- 6 Wiczorek, G. Quantitative DNA methylation analysis of FOXP3 as a new method for counting regulatory T cells in peripheral blood and solid tissue. *Cancer Res.* 2009. **69**: 599–608.
- 7 Lewin, J. Quantitative DNA methylation analysis based on four-dye trace data from direct sequencing of PCR amplicates. *Bioinformatics* 2004. **20**: 3005–3012.
- 8 Spreafico, R. A sensitive protocol for FOXP3 epigenetic analysis in scarce human samples. *Eur. J. Immunol.* 2014. **44**: 3141–3143.
- 9 Ohkura, N. T cell receptor stimulation-induced epigenetic changes and Foxp3 expression are independent and complementary events required for Treg cell development. *Immunity* 2012. **37**: 785–799.
- 10 Kim, H. P. CREB/ATF-dependent T cell receptor-induced FoxP3 gene expression: a role for DNA methylation. *J. Exp. Med.* 2007. **204**: 1543–1551.
- 11 Liu, W. CD127 expression inversely correlates with FoxP3 and suppressive function of human CD4+ T reg cells. *J. Exp. Med.* 2006. **203**: 1701–1711.
- 12 Bending, D. Hypomethylation at the regulatory T cell-specific demethylated region in CD25<sup>hi</sup> T cells is decoupled from FOXP3 expression at the inflamed site in childhood arthritis. *J. Immunol.* 2014. **193**: 2699–2708.
- 13 McClymont, S. A. Plasticity of human regulatory T cells in healthy subjects and patients with type 1 diabetes. *J. Immunol.* 2011. **186**: 3918–3926.
- 14 Zilbauer, M. Genome-wide methylation analyses of primary human leukocyte subsets identifies functionally important cell-type specific hypomethylated regions. *Blood* 2013. **122**: e52–e60.

**Abbreviations:** NGS: Next-generation sequencing · TSDR: Treg-specific demethylated region · Treg: Regulatory T cell · qPCR: Quantitative PCR

**Keywords:** Bisulfite sequencing · FOXP3 · Methylation · Next-generation sequencing · Regulatory T cell

**Full correspondence:** Prof. Linda S. Wicker  
**e-mail:** linda.wicker@cimr.cam.ac.uk

Received: 11/3/2015

Revised: 14/8/2015

Accepted: 1/9/2015

Accepted article online: 7/9/2015



The detailed *Materials and methods* for Technical comments are available online in the Supporting information