Expanded low gene body methylation defines cell-specific epigenome signatures in the adaptive immune system

Marcela K. Sjöberg[1*], Nicolas Walker[2*], Catalina Vallejos[3], Hui Shi[2], Martin Del Castillo-Velasco Herrera[1], Michael J. Booth[4], Martin Bachman[7], Kim Wong[1], Anastasiya Kazachenka[2], Richard Gunning[1], Mamunur Rashid[1], Lira Mamanova[6], Alexandra Ashcroft[1], Laurent Vanhille[5,] Salvatore Spicuglia[5], Shankar Balasubramanian[7], Sarah Teichmann[8], John Marioni[8], David J. Adams[1,+], Anne C. Ferguson-Smith[2,+]

[1]*Experimental Cancer Genetics, Wellcome Trust Sanger Institute, Hinxton, Cambridge, UK.*
[2]*Department of Genetics, University of Cambridge, Cambridge CB2 3EH, UK*
[3]*European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SD, UK.*
[4]*Department of Chemistry, University of Cambridge, Cambridge, UK.*
[5]*Inserm U1090, Technological Advances for Genomics and Clinics (TAGC), F-13009 Marseille, France; Aix-Marseille University UMR-S 1090, TAGC, F-13009 Marseille, France*
[6]*The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK.*
[7]*Department of Chemistry, University of Cambridge, Cambridge, UK; Cancer Research UK, Cambridge Research Institute, Li Ka Shing Centre, Cambridge, UK.*
[8]*European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SD, UK; Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK.*
**These authors contributed equally to this work.*
+co-corresponding authors

Abstract
Dynamic changes in methylation can be identified in gene bodies however their extent and impact on gene expression is not understood. As part of the BLUEPRINT epigenome project we performed genome-wide comparative epigenomics and transcriptomics on purified quiescent mouse naïve CD4+ T and B lymphocytes. Integrative analysis of single-nucleotide resolution 5-methylcytosine (5mC) and 5-hydroxymethylcytosine (5hmC), histone marks and RNA-seq profiles reveal contrasting epigenetic states between cell types. We identified T-cell-specific 5hmC at overall levels comparable to ES cells and observed different patterns of 5hmC distribution over gene bodies with some spanning broad domains not constrained by gene boundaries. Comparative analysis between the two cell types identified remarkable differences characterised by expanded regions of T-cell low gene body methylation (T-LMG) proximal to 5hmC domains. Expanded domains are conserved between human and mouse and are not a prerequisite for gene expression. Repressed and active low methylated genes contain expanded domains of H3K27me3 and H3K27Ac/H3K4me3 respectively. Single cell transcriptome analysis has been conducted and provides new insights into the relationships between epigenetic states and gene expression characteristics. These findings will be presented.