

RESEARCH

Open Access

Differential CpG DNA methylation in peripheral naïve CD4⁺ T-cells in early rheumatoid arthritis patients



R. Pitaksalee¹, A. N. Burska¹, S. Ajaib¹, J. Rogers¹, R. Parmar¹, K. Mydlova¹, X. Xie¹, A. Droop², J. S. Nijjar³, P. Chambers¹, P. Emery¹, R. Hodgett⁴, I. B. McInnes³ and F. Ponchel^{1*}

Abstract

Background: The genetic risk associated with rheumatoid arthritis (RA) includes genes regulating DNA methylation, one of the hallmarks of epigenetic re-programing, as well as many T-cell genes, with a strong MHC association, pointing to immunogenetic mechanisms as disease triggers leading to chronicity. The aim of our study was to explore DNA methylation in early, drug-naïve RA patients, towards a better understanding of early events in pathogenesis.

Result: Monocytes, naïve and memory CD4⁺ T-cells were sorted from 6 healthy controls and 10 RA patients. DNA methylation was assessed using a genome-wide Illumina 450K CpG promoter array. Differential methylation was confirmed using bisulfite sequencing for a specific gene promoter, ELISA for several cytokines and flow cytometry for cell surface markers. Differentially methylated (DM) CpGs were observed in 1047 genes in naïve CD4⁺ T-cells, 913 in memory cells and was minimal in monocytes with only 177 genes. Naïve CD4⁺ T-cells were further investigated as presenting differential methylation in the promoter of > 500 genes associated with several disease-relevant pathways, including many cytokines and their receptors. We confirmed hypomethylation of a region of the TNF-alpha gene in early RA and differential expression of 3 cytokines (IL21, IL34 and RANKL). Using a bioinformatics package (DMRcate) and an in-house analysis based on differences in β values, we established lists of DM genes between health and RA. Publicly available gene expression data were interrogated to confirm differential expression of over 70 DM genes. The lists of DM genes were further investigated based on a functional relationship database analysis, which pointed to an IL6/JAK1/STAT3 node, related to TNF-signalling and engagement in Th17 cell differentiation amongst many pathways. Five DM genes for cell surface markers (CD4, IL6R, IL2RA/CD25, CD62L, CXCR4) were investigated towards identifying subpopulations of CD4⁺ T-cells undergoing these modifications and pointed to a subset of naïve T-cells, with high levels of CD4, IL2R, and CXCR4, but reduction and loss of IL6R and CD62L, respectively.

Conclusion: Our data provided novel conceptual advances in the understanding of early RA pathogenesis, with implications for early treatment and prevention.

Keywords: Naïve CD4⁺ T-cells, DNA methylation, Rheumatoid arthritis, Interleukin-6, Illumina 450K array

* Correspondence: f.ponchel@leeds.ac.uk

¹Leeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds, Leeds, UK

Full list of author information is available at the end of the article



© The Author(s). 2020 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Background

Rheumatoid arthritis (RA) is a chronic condition with substantial impact on the lives of millions of people worldwide. RA patients present with symmetrical joint inflammation, which, if inadequately treated, results in severe disability and substantial pain, fatigue, depression, and adverse social consequences with a significantly increased risk of mortality [1, 2].

Epigenetic modifications are characterising diverse pathologies, including cancer and autoimmune diseases (AIDs). In cancer, CpG methylation patterns often affect large regions of DNA being activated or silenced. Such epimutations are considered as critical as genetic mutations but importantly, were shown to occur before the onset of the latter [3]. In AIDs, similar changes were identified, although on a smaller scale [4]. In RA, synovial fibroblasts and T-cells harbour alterations in DNA methylation profiles, while whole blood analysis associated changes in methylation (between baseline and 4 weeks) in 2 CpGs, with response to methotrexate in early RA [5–8]. We hypothesised that such changes may indeed occur at the earliest stages of RA progression, under the influence of environmental triggers notably inflammation [9], representing an opportunity to explore molecular mechanisms mediating transition to chronicity.

Thus, we investigated whether differential methylation (DM) could be detected in early RA patients with < 6 months symptom duration (drug naïve). We selected 2 circulating cell types implicated in RA pathogenesis [4], namely T-cells (segregating naïve and memory subsets) and monocytes, and compared DNA methylation patterns between RA and healthy controls (HC) using a 450K-CpG genome-wide array. Herein, we describe our dataset and then the application of strategies to prioritise DM-CpG sites, extrapolating these to genes and pathways modulated in early RA.

Results

Preliminary exploration of DNA methylation data

DNA methylation profile of 6 HC and 10 RA patients from naïve and memory CD4⁺ T-cells and monocytes was analysed using a workflow shown in Figure S1. Multidimensional scaling (MDS) established the main discriminants between samples, and clustered data tightly by cell types (Fig. 1a) but not by gender or between RA and HC (supplementary Figure S2). We explored levels of methylation for each CpG by t-tests (false discovery rate correction was not applied). Manhattan plots displaying *p* values in an ordered manner along chromosomes, identified thresholds of significance for *p* values, separating DM-CpGs from the background: high ($p \leq 0.0001$), medium ($0.0001 < p \leq 0.001$) and low ($0.001 < p \leq 0.01$; Fig. 1b, naïve T-cells, supplementary Figure S2, memory T-cell and monocyte). This preliminary

data mining strategy identified distinct numbers of DM-CpGs for the 3 cell types (Table 1).

DM CpG distributions in the 3 cell subsets

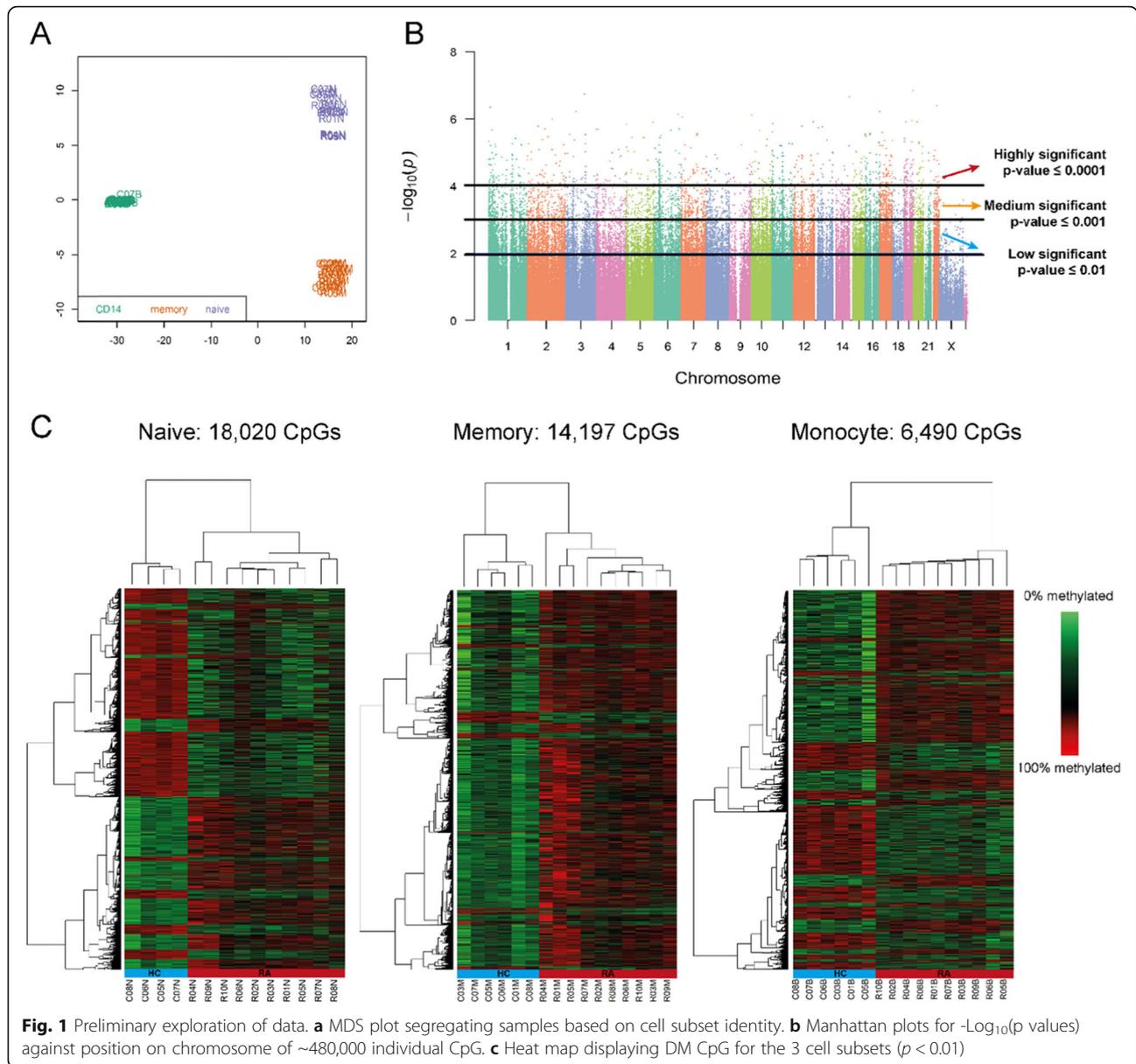
In naïve T-cells, the number of DM CpGs (18,020, $p \leq 0.01$) was the highest representing 4.09% of all tested CpGs. Of this number, 3.11%, 16.04% and 80.84% were categorised as of high, medium, and low significance, respectively. Using annotation, we analysed the distribution of DM CpGs with respect to their location in gene structures for (i) core of CpG islands (ii) shelves/shore of an island or (iii) in open sea (i.e. outside of a defined island). We observed approximately 1/3 of all DM CpG in each category. We also observed similar proportion of hypo and hyper-methylation, with a slight bias toward hypomethylation. In memory T-cells, the main difference observed was that DM preferably occurred in the core of islands (50%) and were mostly hyper-methylation (93% of DM CpGs) which may suggest more generalised gene silencing in RA memory T-cells. In monocytes, DM was of lower significance altogether (Table 1). Overall, these data demonstrate distinct methylation changes between cell types. This was confirmed using unsupervised hierarchical clustering displayed as heatmap (Fig. 1c), showing clear segregation of patients and HC, as well as major hyper-methylation in memory T-cells.

Differential methylation patterns

To evaluate the effect of DM on gene expression, we manually inspected selected CpGs, chosen from the top of the *p*-value lists. Three typical patterns were identified (supplementary Figure S3). First, DM at a single CpG usually located in the core of an island, possibly modulating transcription factor binding sites (island a). Second, patterns showing clusters of DM CpGs potentially mediating a cumulative local effect (island b). Third, an isolated CpG in open sea (not always associated with a specific gene) possibly part of an enhancer region (example c), although these were not meant to be targeted by the 450K bead array.

Exploration of clusters of DM CpG in the 3 cell subsets

We ranked clusters automatically using an in-house R-code, selecting only the highly significant DM CpGs and taking into account significantly DM CpGs in their proximity in a region of ± 1500 bp (supplementary Figure S3; R-code available on request). In monocytes, most clusters only showed the initial selecting CpG associated with only 1 other DM CpG (final score = 2). In contrast, in naïve T-cells, some clusters showed up to 7 selecting CpGs and scores up to 16, suggesting much larger effects on a wider region of the DNA. Intermediate results were observed in memory T-cells (score up to 9). The number of DM CpG clusters with a score of > 2 are listed in Table 1. DM CpG clusters and isolated DM CpG were associated with genes



via annotation and lists were drawn. Six hundred forty eight DM genes were observed in naïve T-cells (Table 1: 354 hypomethylation and 294 hypermethylation), 605 for memory T-cells and only 58 for monocytes. Of note, the second highest score for hypomethylated DM CpG cluster in naïve T-cells was for the tumour necrosis factor gene (TNF-alpha, Fig. 2, which presents no DM in memory cells or monocytes). For isolated DM CpGs (only $p < 0.001$, Table 1), using ranking based on p value, 266 hypomethylated genes and 133 hypermethylated for naïve T-cells. Full list of genes are available in [supplementary files](#) (Data S1-3).

Of note, we only observed 1 DM CpG cluster common to 3 cell types ([supplementary Figure S3](#)), associated with the 4-aminobutyrate aminotransferase (ABAT)

gene and none for isolated DM CpGs. We considered whether this finding may be an artefact but note that this gene is associated with clusters of 9 DM CpGs in naïve cells, 6 in memory cells and 4 in monocytes.

DM of cytokines/chemokines/receptors in the 3 cell subsets

Many of the DM genes described above were cytokines/chemokines (IL1beta, IL6, IL12, IL13, IL15, IL17, IL21 and more as well as TNF, TGF-beta1 and many more members of their families; several CCLs and CXCLs) as well as their receptors (IL2RA/RB, IL6R, IL10RA, IL15R, IL17Rs, TNFRs and more, chemokine receptors) all listed in Table 2. Many of these factors were showed to be upregulated at

Table 1 Summary of differential methylation at individual CpG level and summary of the prioritisation of clusters of DM CpG and associated genes

	Naive CD4 T-cell	Memory CD4 T-cell	Monocytes
Number (%) of differentially methylated / total probes tested (440 490), $p \leq 0.01$	18020 (4.09%)	14197 (3.22%)	6490 (1.47%)
Number (%)* of probes with			
High significance ($p \leq 0.0001$)	561 (3.11%)	440 (3.10%)	130 (2.00%)
Medium significance ($0.0001 < p \leq 0.001$)	2891 (16.04%)	1885 (13.28%)	672 (10.35%)
Significance ($0.001 < p \leq 0.01$)	14568 (80.84%)	11872 (83.62%)	5688 (87.64%)
Number and (%)* of probe associated with			
Core island	6141 (34.08%)	7120 (50.15%)	1985 (30.59%)
Shelves/shore island	5873 (32.59%)	3948 (27.81%)	2060 (31.74%)
Outside of CpG island	6006 (33.33%)	3129 (22.04%)	2445 (37.67%)
Hypermethylation in RA (%)*	8425 (46.75%)	13218 (93.10%)	3525 (54.31%)
Hypomethylation in RA (%)*	9595 (53.25%)	979 (6.9%)	2965 (45.69%)
Number of probe (%)** associated with an			
Island/shelve/shore			
Hypermethylated	1201 (34.79%)	1842 (79.23%)	209 (26.06%)
Hypomethylated	1111 (32.18%)	114 (4.9%)	310 (38.65%)
Outside of an island (Open sea)			
Hypermethylated	176 (5.10%)	299 (12.86%)	190 (23.69%)
Hypomethylated	964 (27.93%)	70 (3.01%)	93 (11.60%)
DM CpG clusters			
Number of CpG with a score ≥ 3			
Hypermethylated	143	305	6
Hypomethylated	197	0	15
Number of CpG with a score = 2			
Hypermethylated	277	414	15
Hypomethylated	223	7	33
Score \geq median (range)			
Hypermethylated	3.65 (3–9)	3.73 (3–9)	3 (3–4)
Hypomethylated	4.15 (3–16)	Na	3 (3–4)
Corresponding number of genes (all clusters)			
Hypermethylated	354	600	19
Hypomethylated	294	5	39
Isolated DM CpG			
Number of CpG/gene associated with			
Island/shelve/shore			
Hypermethylated	121	249	34
Hypomethylated	139	24	58
In Open sea			
Hypermethylated	12	18	14
Hypomethylated	127	17	13

* (%) of all probes with $p \leq 0.01$, ** (%) of all probes with $p \leq 0.001$

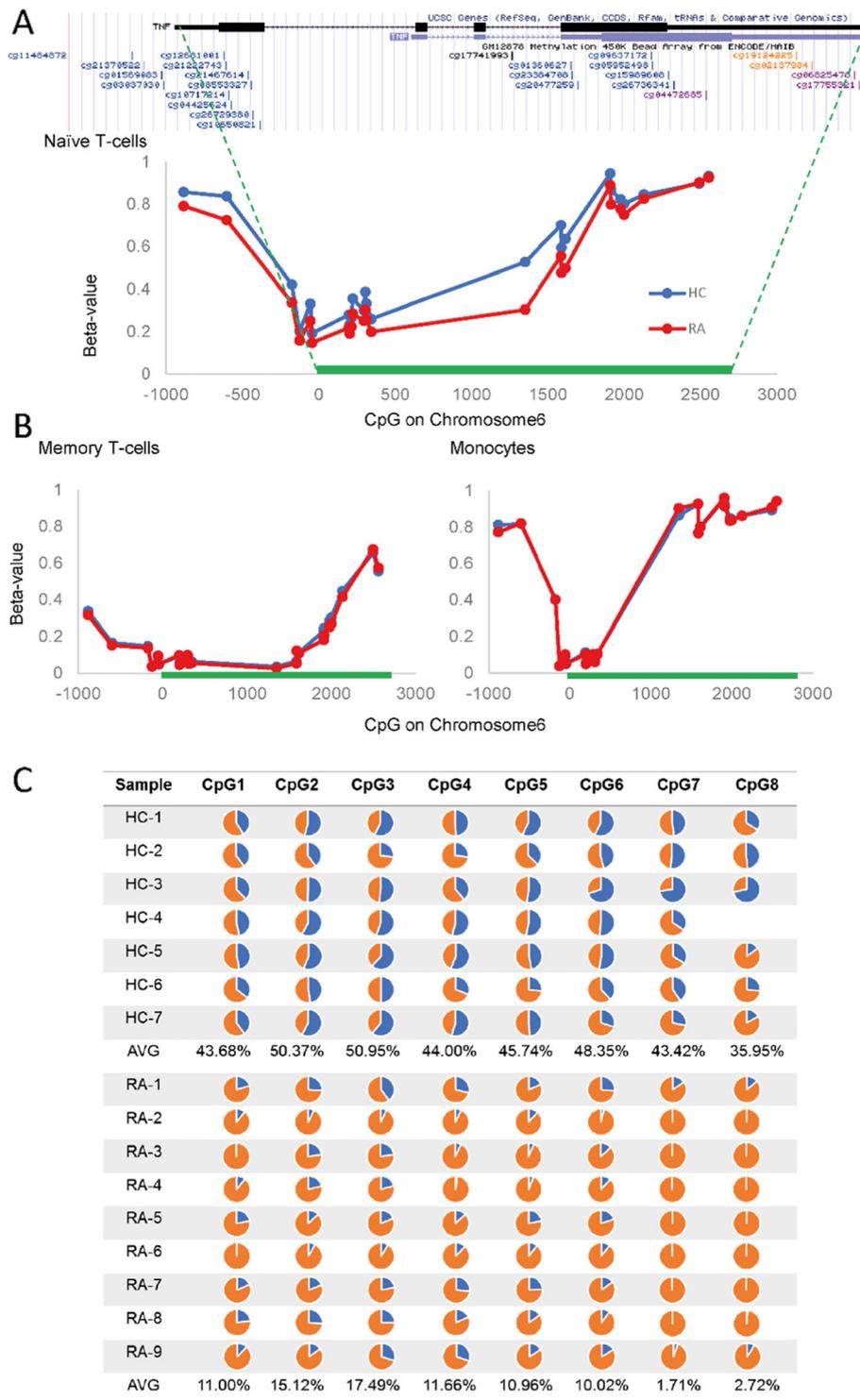


Fig. 2 DNA bisulfite sequencing of the TNF-alpha promoter region. **a** CpGs present in the TNF-alpha gene were ordered on Chromosome 6. At most CpG positions, the median β values in naive CD4⁺ T-cells show significant hypomethylation in RA (red) compared to HC (blue). **b** Median β values in the similar region of chromosome 6 in memory cells and monocytes. There was no DM between HC and RA in both cell types. **c** A region of 273 bp was amplified for direct bisulfite sequencing, containing 3 of the array CpGs. This region is highly demethylated in memory cells but highly methylated in monocytes. Results of the sequencing covering 8 CpG displayed as pie chart for the percentage of methylated (blue)/ demethylated (orange) DNA, showing on average ~45% demethylation in HC ($n = 7$) and ~90% in RA ($n = 9$)

the protein levels in early RA patients [10–12], notably including IL6 for which this was also associated with DM of the gene and mRNA levels [13]. We chose 3 cytokines for which no data were available in early RA (IL21, IL34 and RANKL) and measured protein levels in serum samples using ELISA. All 3 cytokines showed higher levels in early RA (all $p < 0.001$, data shown in supplementary Figure S4-A). We also have data similar to those widely reported, showing higher levels of IL1-beta, IL6, IL10, IL12, IL17A, IFN-gamma and TNF-alpha (all $p < 0.001$), recapitulated in supplementary Figure S4 B.

Validation of DM regions using publicly available R packages for the naïve T-cell dataset

There are R packages publicly available to analyse DNA methylation gene arrays. We ran our dataset using DMRcate, an R-package aiming to find DM region (DMR) [14], which uses moderated t-tests and a function to agglomerate neighbouring CpGs. It relies on 3 main parameters. The first one allows to modulate the acceptable FDR.

The second modulates the length of the region to consider. The third modulates a threshold in difference in β values ($\Delta\beta$ values), with the options to use the average (standard/default) or the maximum difference (notably in small data set). DMRs can then be annotated to the associated gene. Running this analysis using the recommended/default settings (FDR ≤ 0.05 , 1000 bp, average $\Delta\beta$ value ≥ 0.05), only 2 genes were DM in RA, CUTA (CutA divalent cation tolerance homolog), and B3GALT4 (Beta-1,3-Galactosyltransferase 4).

In assessing whether DM are relevant, $\Delta\beta$ values appear quite important. We therefore analysed our dataset considering the distribution of β value observed. We filtered out 3778 CpGs with poor reproducibility (i.e. $> 10\%$ difference between Min and Max β values). We then analysed the distribution of $\Delta\beta$ values for each GpG. Allowing for positive/negative differences, 2.5% on each side of the distribution suggested a cut-off at -0.04612 and $+0.04432$ $\Delta\beta$ values (supplement Figure S5A). We therefore used a ± 0.05 cut-off to customise

Table 2 DM genes for cytokine/chemokines and their receptor in early RA

Gene Symbol	Naive cells	Memory cells	Monocytes
Interleukin family			
Cluster	IL6, IL12A, IL13, IL21, IL25, IL31, IL34, IL36G, IL1RAPL2, IL5RA, IL6R, IL10RA, IL12RB1, IL15RA, IL17RC, IL17RE, IL17REL, IL27RA	IL1B, IL6, IL12A, IL15, IL17F, IL17D, IL24, IL37, IL1R2, IL4R, IL6R, IL12RB1, IL15RA, IL17RB, IL17RC, IL17RD, IL17REL, IL20RB, IL21RAS1	IL16, IL37, IL1RN, IL17RC
Isolated CpG	IL1B, IL5, IL10, IL15, IL16, IL17C, IL17D, IL19, IL25, IL36B, IL17RA, IL18RAP, IL18BP, IL1R1, IL1R2, IL1RN, IL2RA, IL2RB	IL17RA	IL12
Tumour necrosis growth factor family			
Cluster	TNF, TNFAIP8, TNFSF10, TNFSF12, TNFSF14, TNFRSF1A, TNFRSF1B, TNFRSF6, TNFRSF8, TNFRSF10B, TNFRSF13B,	TNFRSF10C, TNFAIP8, TNFAIP8L3, TNFAIP8L1, TNFRSF9, TNFRSF10B, TNFRSF18, TNFRSF19, TNFRSF25	TNFAIP2, TNFAIP3, TNFRSF18, TNFRSF19, TNFRSF25
Isolated CpG	TNFSF4, TNFSF10, TNFSF12, TNFRSF10A	TNFRSF13B, TNFRSF13C	TNFRSF1B
Interferon family			
Cluster	IFNA2, IFNGR2	IFNGR2, IFNGR1	
Isolated CpG	IFNL4		
Transforming growth factor family			
Cluster	TGFB1, TGFB2, TGFB3	TGFB3, TGFB1, TGFB2, TGFB3,	TGFB1, TGFB2AS1
Isolated CpG	TGFA, TGFB2-AS1	TGFB3L	TGFA
Chemokine families			
Cluster	CCL5, CCL16, CCL25, CCL27, CXCL12, CCR3, CCR5, CCR7, CCR9, CCR2, CXCR4	CXCL2, CXCL5	CCL24, CXCL3, CXCL12, CCR3, CCR9
Isolated CpG	CCL2, CCL8, CCL20, CCL24, CXCL11, CXCR1, CXCR5	CCL23, CCL25, CCL28, CXCL1, CXCL3, CXCL11, CXCL14, CXCL12, CXCL16, CCR3, CCR6, CCR9, CCR10, CCR2, CXCR1, CXCR2, CXCR4	CCL3, CXCL1, CXCL16, CCR7, CXCR1, CXCR4

DMRcate settings, allowing for small size sample by using maximum $\Delta\beta$ values. The package allowed for a list of genes to be drawn (FDR at ≤ 0.10), including 105 and 98 genes, hypo and hypermethylated respectively (LIST-1, [supplementary material](#)). The most immediately recognisable genes were TNF/TNFRs, some IFN signalling-related genes, HLA-related genes, STAT5, some integrin.

Alternatively, the top $\pm 2.5\%$ of the $\Delta\beta$ value distribution represents $\sim 20,000$ CpG (Supplementary Figure S5A, on each side of the distribution) of which $\sim 10\%$ were significantly DM ($p \leq 0.001$, 1297 hypomethylated and 336 hypermethylated). We annotated these for gene symbols (some had more than 1 DM CpG/gene), which resulted in LIST-2 comprising 1217 genes (952 hypo and 265 hypermethylated, including 26 microRNA). This LIST-2 showed an overlap of 122 genes with LIST-1 (supplementary Figure S6). A more stringent list for the most highly DM genes was further drawn, by using a cut-off for maximum $\Delta\beta$ values set at 0.10 (LIST-3), including 355 genes (262 hypo and 93 hypermethylated). This comprised again TNF/TNFR, IFN signalling-related genes, HLA-related genes, STAT5, integrin and, additionally, some cytokines/chemokines/receptors. LIST-1 and LIST-3 had 78 genes in common (supplementary Figure S6). For illustration, the volcano plot ($-\log p$ value vs $\Delta\beta$ values) for all CpG and the DM-CpG using a $\Delta\beta$ values cut-off set at 0.05 (LIST-2) and at 0.10 (LIST-3) can be found in supplementary Figure S5C.

Bisulfite sequencing of the TNF gene promoter in CD4⁺ T-cells

Some of the genes highlighted above (notably TNF [15]) were shown to be DM in RA (early or established) using bisulfite sequencing [13] and pyrosequencing [16], although not in early, drug-naïve disease. These included notably IL6, IL6R [13, 17, 18]. We selected the TNF gene for further validation based on its known relationship with RA pathogenesis. In naïve cells, the region between $-850/+2000$ bp from the transcriptional start of the TNF gene showed partial demethylation (Fig. 2a) with an average β value of 50% methylation). In RA naïve CD4⁺ T-cells, DM was observed with consistent hypomethylation of the whole region with an average of -7.1% difference in β values (range -2.3% to -20.8%). In contrast, this region is almost fully demethylated (average 22%) in memory cells (Fig. 2b), while in the monocyte, the region was also fully demethylated (average 8%) but on a much shortened scale between $-175/+343$ bp while the rest of the region was highly methylated.

A 273 bp regions encompassing 8 CpGs (including 3 CpGs from the array) was sequenced from total CD4⁺ T-cell DNA following cell sorting using magnetic beads ($n = 16$, average 97.5% purity). In HC 50%/50% methylated/demethylated DNA was observed (Fig. 2c, $n = 7$), suggesting 2 subpopulations of CD4⁺ T-cells, one with

methylated and one with unmethylated DNA. In RA ($n = 9$), 90% of unmethylated DNA was observed at all CpGs, showing that most CD4⁺ T-cells have altered their TNF gene, early in the RA disease process. In this region of the promoter, memory T-cells show highly demethylated CpGs with no significant difference between HC and RA but contributing to the intermediate levels of total demethylation observed (50%) when combined with naïve cells. In RA, our data therefore confirm that a large proportion of naïve T-cells have hypomethylated the TNF gene-promoter compared to HC, totalling 90% of demethylated DNA in that region.

Differential gene expression compared to differential gene methylation in CD4⁺ T-cells

We selected two gene expression datasets on CD4⁺ T-cell from early drug-naïve RA patient and HC [19, 20] (no dataset available for naïve CD4⁺ T-cells). After normalisation and aggregation of the 2 datasets (supplement Figure S7A), we obtained a list of differentially expressed genes (DEG; with adjusted p value ≤ 0.05 , fold change ≥ 1.5 , FDR ≤ 0.05) between HC and RA. These genes included JAK1, TNF-family, ICOS, CD69, several MAP-kinases and their regulators, TGF-beta1, c-FOS and JUN, HLA-related molecules, several IFN signalling genes (IRFs, IFITMs), some TLRs, cytokines/chemokines, their receptors and PADI4.

From the lists of DM genes (LIST-3), 70 gene symbols could be matched with DEGs (after removing microRNA and ambiguous symbols, supplement Figure S7B). Taking the top genes based on fold differences in gene expression between RA and HC, the DM/DEG genes associated with known RA pathological pathways pointed again to JAK1, STATs, TNF-family, IFN signalling genes.

In silico functional interactions between products of DM genes in naïve CD4⁺ T-cells

We next explored whether DM genes would point to specific pathways and/or functions in naïve CD4⁺ T-cells that could be further associated with pathogenesis. We selected the STRING database [21] for known and predicted physical interactions and/or functional associations between gene products (i.e. proteins) from knowledge databases (including experimental data, computational prediction methods and public text collections using a number of functional classification systems such as GO, Pfam and KEGG). We used the 70% confidence in interactions setting but rejected co-expression therefore allowing mainly for functional interactions.

We generated a network based on LIST-1. This pointed to an initial system of nodes based on TNF/TNF-R1, IFN-signalling (IFITM1, IRF4), STAT5, integrins (ITGB2, ITGAL), HLAs and HDAC; however, 123 of the 203 symbols submitted remained outside of the network.

LIST-2 was too large to fit into the STRING model. Using LIST-3 (excluding microRNA), the same system of nodes was generated however, further enforced by several novel association/interaction becoming visible: the TNF/TNF-R1 node included 3 more TNF-super family genes, creating further associations; the IFN signalling node was enriched with 4 more genes linked to several others; the STAT5 node was linked to IL2/IL2R and IL10R. Additional small networks were created related to ubiquitination and DNA methylation and a SMAD/TGF association appeared.

The software suggested manual addition of several genes to strengthen certain nodes: some of these genes were present on LIST-1/2, some were reported on the cytokine/receptor (Table 2). We therefore manually interrogated methylation data for all suggested genes. LIST-2 genes were accepted (for $\Delta\beta$ values ≥ 0.05 , FDR ≤ 0.10 , $p \leq 0.001$). We included or rejected other genes based on whether they had (i) a $\Delta\beta$ value of ≥ 0.05 but a $p \leq 0.005$ (mainly from Table 2) or (ii) whether they were central protein in a node bringing together more interaction with DM genes present on LIST-1 or 3 (3 genes, displayed as grey bubbles on the model). Overall, the added genes included cytokines/chemokines and their receptors, STATs molecules and a few others (final gene list in [supplementary files](#): STRING LIST-4).

Our final potential interaction model (Fig. 3) displays 3 main JAK1/STATs nodes, an inflammatory node associating IL1R/IL6-IL6R/JAK1/STAT3 signalling (green), a second node for JAK1/STAT2 linked to interferon signalling-related genes (orange), and a third centred on IL2R/IL15R/JAK1/STAT5 (dark blue). Many other associations were suggested, in relation with IL6, including TNF family (purple), the IL10R and DNA methylation-related genes and targets. An IL17/IL25 loop (duck green) was more directly related to JAK1 signalling, and last DNA modification enzymes themselves (HDAC and other in yellow). Other additions (not displayed on Fig. 3 to simplify the overall network) implicated a further node for IL18/JAK1/STAT4, added more members of the TNF superfamily, linked IL4R and IL13 to STAT5 and created a link between IL1 and IL36.

A similar analysis was performed on memory and monocyte. IL6 also appears central to the memory T-cell networks via a molecule called EP300 (E1A-associated protein p300, a histone acetyltransferase) associated with IL6-signalling [22] (Supplementary Figure S8). Although the IL6R/IL6 genes were not directly DM in monocytes (neither was TNF), again EP300 appears central. We subsequently analysed the IL6 and TNF signalling cascades in more detail (Supplementary Figure S9 A, B). Many components directly involved in the TNF signalling cascade showed DM CpGs, however not prioritized in the STRING network figure (Fig. 3), while similar and/or additional genes showed the same for the IL6

pathway. We also aligned DEG onto these two signalling cascades, which highlighted JAK1 and STAT3/4 and many more genes, with concomitant DM at DNA level ($p \leq 0.01$, Figure S9 C, D).

Characterisation of a subpopulation of cells based on DM of cell surface marker

Considering that methylation is a binary event at each CpG position, the differences in β values observed on the array suggests the emergence of a subpopulation of cells that have altered their methylation status at such positions. The amplitude of these differences (i.e. $\Delta\beta$ values) reflects the proportion of cells that have achieved that change in HC or RA. In order to identify such subpopulation, we hypothesised that these changes could affect cell surface molecules allowing for a subset of naïve CD4⁺ T-cells to be defined.

We analysed the total CD4⁺ T-cell gene expression dataset for HC ($n = 16$) for expressed gene with an associated protein localisation in the cell membrane. Using the top 25% quartile of the mRNA levels distribution, > 1000 genes were selected as highly expressed. Using the GO-term database, 302 of these were annotated as cell surface proteins. We then cross this list with LIST-2 and Table 2 and obtained 32 potential DM cell surface markers. These included CD4 itself, markers associated with the naïve/memory phenotype, cytokine receptors, chemokine receptor, TLRs and others, that could help identify subpopulation(s) of naïve cells.

Blood samples from 10 HC and 35 RA patients were collected. We first used CD45RA⁺ and CD45RO⁻ to identify naïve (Fig. 4a, red square) and memory (green circles) cells. Both populations were homogenous in HC and in RA (representative individuals). The expression of 5 surface markers was then analysed (Fig. 4b). CD4, IL6R, IL2R, CD62L and CXCR4 were tested (IL7R was used as no-DM control, data not shown) initially on CD45RA⁺ naïve cells, measuring mean fluorescence intensity (MFI) of expression. CD4, CXCR4 and IL2R expression were significantly higher in RA ($p < 0.0001$) but not IL6R, which expression was very variable compare to HC. The expression of CD62L was either positive (Fig. 4c red circle, best example RA patient displayed on figure) or negative (blue square). The percentage of CD62L⁻ naïve CD4⁺ T-cells was significantly higher in RA (median 1.3%, $p < 0.0001$) compared to HC (median 0.15%). This was particularly clear in 3 patients with high CRP (55, 75 and 178 mg/L) where the CD62L⁻ cells represented 11, 13 and 18% of the total CD4⁺ T-cells, respectively. This subpopulation of CD62L⁻ naïve cells (Fig. 4d, blue line) was then analysed in these 3 RA patients, compared to CD62L⁺ naïve cells (red line) and memory cells (green line) for 3 markers. The expression of CD4 showed no significant difference ($n = 3$) between

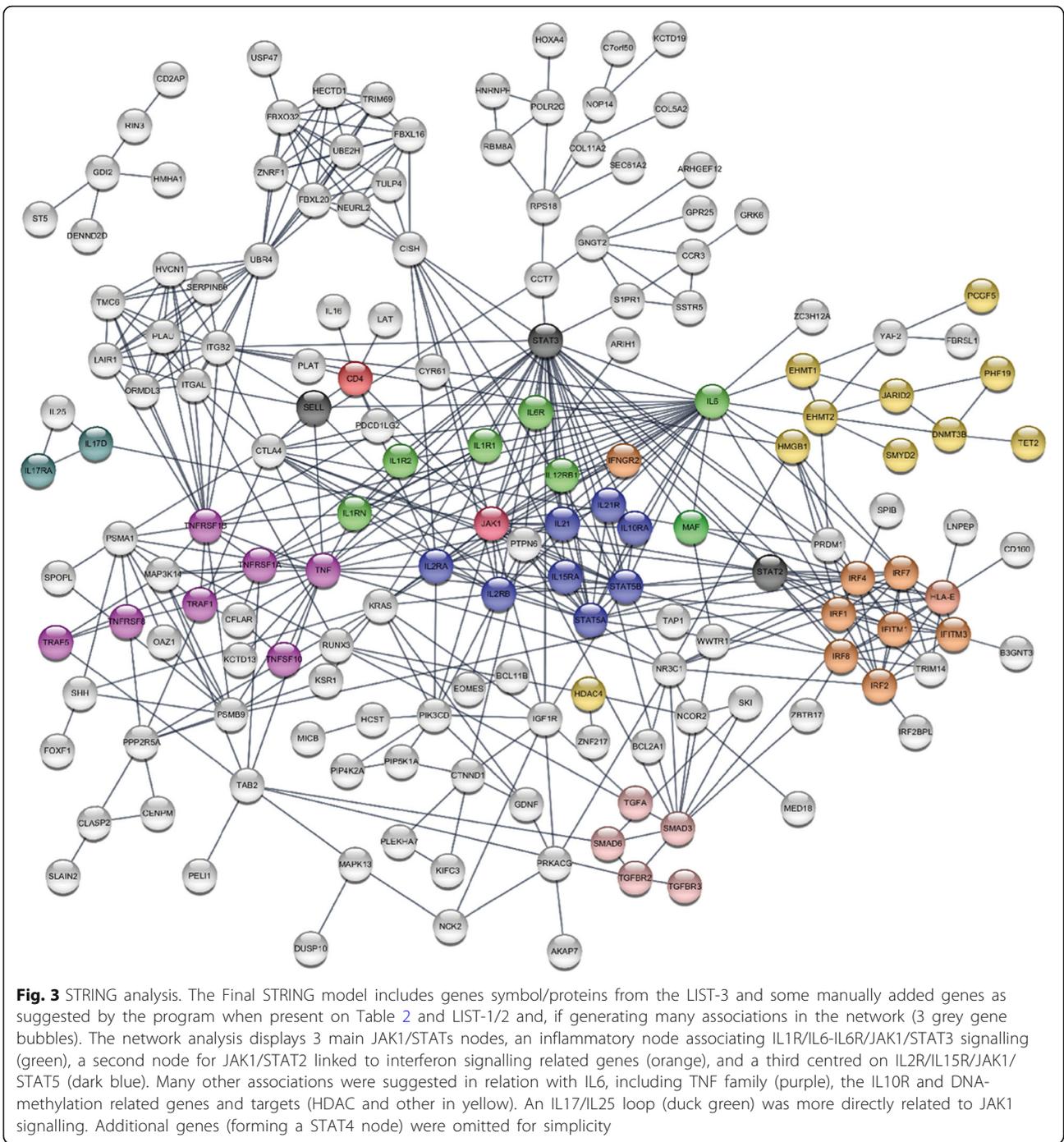


Fig. 3 STRING analysis. The Final STRING model includes genes symbol/proteins from the LIST-3 and some manually added genes as suggested by the program when present on Table 2 and LIST-1/2 and, if generating many associations in the network (3 grey gene bubbles). The network analysis displays 3 main JAK1/STATs nodes, an inflammatory node associating IL1R/IL6-IL6R/JAK1/STAT3 signalling (green), a second node for JAK1/STAT2 linked to interferon signalling related genes (orange), and a third centred on IL2R/IL15R/JAK1/STAT5 (dark blue). Many other associations were suggested in relation with IL6, including TNF family (purple), the IL10R and DNA-methylation related genes and targets (HDAC and other in yellow). An IL17/IL25 loop (duck green) was more directly related to JAK1 signalling. Additional genes (forming a STAT4 node) were omitted for simplicity

the 2 subpopulations of naïve cells or in memory cells, although the MFI was lower for CD62L⁻ naïve cells compared to naïve cells and memory cells (see details on Fig. 4d legend). The expression of the IL-6R was clearly reduced on CD62L⁻ compared to CD62L⁺ naïve cells as well as on memory cells. The IL2R expression was negative on CD62L⁺ naïve cells but presented 2 populations (- and +) for CD62L⁻ naïve cells. Memory cells were mainly positive.

Discussion

Our results demonstrate highly cell-specific changes in DNA methylation in early RA in circulating leukocytes. These differences are both qualitative (for example the excess of hypermethylation in memory T-cells) and quantitative (more marked in naïve, then memory then monocytes). The lack of gene commonly affected in the 3 subsets confirm that changes are cell specific [6]. These data were validated *in silico*, using publicly available data matching DM-

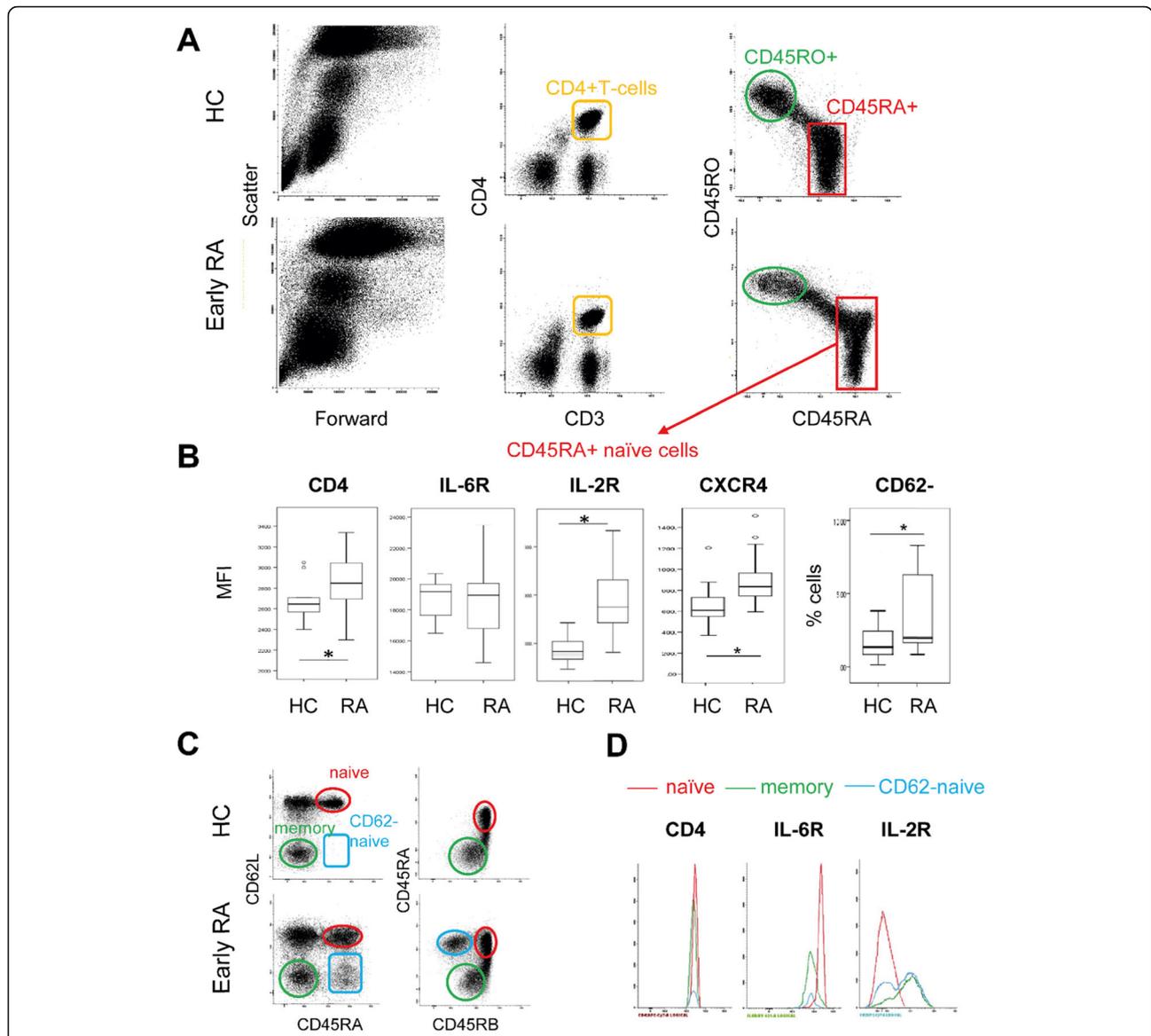


Fig. 4 Flow cytometry validation of differential expression. Flow cytometry was performed using standard cell surface staining protocol using fresh EDTA blood, following red cell lysis. **a** CD3⁺CD4⁺ T-cells (top left panels orange gate) were first gated. Naïve cells were then gated as CD45RA⁺/CD45RO⁻ (red square) and memory cells as CD45RA⁻/CD45RO⁺ (green circle) in a representative HC and RA patient. **b** The expression of CD4, IL6R, IL2R and CXCR4, were measured using Mean fluorescence intensity (MFI). Results are shown as box plot for 11 HC and 35 RA patients. CD62L was either positive or negative and percentage of naïve CD62L⁻ cells was recorded and displayed. Significant differences (Mann–Whitney *U*-test, *p* < 0.05) are highlighted by stars. **c** CD45RB and CD62L were further used to refine the phenotype of naïve CD4⁺ T-cells. CD45RB expression was consistently high in naïve cells but declined in experienced cells and was low in memory cells (green circles), with no major difference between HC and RA for this subset. CD62L expression is positive on naïve cells (red circle, consistently in HC) but was either positive (red circle) or negative (blue square) in RA defining an subpopulation of naïve CD62L⁻ cells also expressing reduced levels of CD45RB (blue circle). **d** Differential levels of expression for CD4, IL6R and IL2R are shown in a RA patients with a raised CD62L⁻ naïve cells subpopulation (best representative patient displayed) for naïve (red) memory (green) and IRC (blue) cells. Levels of CD4 were not significantly different (*n* = 3) but a lower MFI was observed for CD4 expression on CD62L⁻ naïve cells (2250) compared to CD62L⁺ naïve cells (2850) and memory cells (2700). The expression of the IL-6R was lower on CD62L⁻ (MFI 7300) compared to CD62L⁺ naïve cells (17,600) as well as on memory cells (11,400). The IL2R expression was negative on CD62L⁺ naïve cells but presented 2 populations (negative < 1000 fluorescence units and + fractions > 1000) for CD62L⁻ naïve cells. Memory cells were mainly positive (72% of cells)

gene and DEG and we also confirmed directly the DM of the TNF gene promoter and the higher expression of 3 DM-cytokine levels. Focussing on naïve CD4⁺ T-cells, a functional network analysis suggested a central role of IL1-

IL6/JAK1/STAT3 with links to TNF and IL17. A second set of genes suggests a role for IL2-IL15/JAK1/STAT5 and finally a JAK1/STAT2-interferon-signalling gene set node. Finally, we confirmed differential expression of CD4,

CXCR4 and CD25 on naïve CD4⁺ T-cells and reduced CD4, IL6R and increased CD25 expression on a subset of CD62⁻ naïve CD4⁺ T-cells.

Our DNA methylation workflow used standard procedures and an in-house analysis prioritising DM higher $\Delta\beta$ values. High-throughput genomic data usually corrects for FDR controlling the family-wise error. Correcting with a $FDR \leq 0.05$ in our dataset left very few CpGs, so we accepted a $FDR \leq 0.010$, considering our study exploratory with some false-positive results being the risk (LIST-1). The $\Delta\beta$ values relied on a t-test ($p \leq 0.001$) selecting the highest differences (LIST-3). Within the limitation of each analysis, the overlap between lists and Table 2 generated association of genes in several pathways which relevance to RA is high.

The only commonly DM gene to all 3 subsets was ABAT, 4-aminobutyrate aminotransferase enzyme, responsible for the catabolism of gamma-aminobutyric acid (GABA), an essential inhibitory neurotransmitter, reducing neuronal excitability and directly responsible for the regulation of muscle tone. T-cells express the GABA receptor and exposure to GABA is involved in regulatory loop reducing inflammation and promoting 'regulatory' responses [23, 24]. As such, GABA has been implicated in AIDs animal models including arthritis [25]. DM in ABAT may therefore contribute to alteration of GABA's regulatory effect in all blood cell types, promoting inflammation and autoimmunity.

Our data suggest that naïve T-cell in early RA are prompted towards differentially methylating the IL17/IL17R genes potentially towards Th17 development (more clearly observed in memory cells). The RORC gene itself was also MD in naïve cells ($> 5\%$ $\Delta\beta$ values at 2 CpG sites). It remained unclear how de novo differentiation of Th17-cells is stimulated in vivo, but in vitro, it can be induced by combinations of cytokines (IL1 β , IL21, IL23, IL6 with/without TGF- β) [26–30]. Importantly, IL6/IL21/IL21R and TGF-beta-1 were DM in RA naïve cells. Th2 differentiation appears intact in RA [31]; however, Th1 polarisation is compromised by a deficit in Tbet engagement [32, 33] (also confirmed in our data with hypermethylation 6.85% $\Delta\beta$ values). Deficient Th1 polarisation in early RA could be a mechanism resulting in Th17-cells developing preferentially.

An IFN signalling node was also highlighted. Dysregulation of type-I INFs are often observed in autoimmunity and chronic inflammatory diseases [34–37]. In a study of at-risk individual for RA, IFN signalling genes were indicative of progression to the inflammatory stage [38–40]; however, IFN signatures were no longer reported predictive later in the disease course [35, 41, 42]. As supported by our data, this suggests that IFN signalling is associated with early pathogenesis, independently of whether this can be exploited clinically later in the disease. Furthermore,

links between IL6 signalling/production and type-I IFN-gene signatures (and vice versa) were also observed in other inflammatory diseases [43, 44], supporting a possible link in early RA development.

IL6's importance is well recognised in RA [18, 45, 46]. Our data suggest a potentially central role for this pathway. Several studies have also reported specific DM in the IL6 and IL6R genes in established RA patients' PBMC, T-cells or synovial fibroblast [13, 17, 18]. IL6 has been shown to induce changes in DNA methylation in cancer and SLE [47–49]. Recent gene expression analysis in CD4⁺ T-cells also suggested a role for IL6 signalling in early RA [19]. IL6 levels in serum are also well known for being increased in early RA (shown here and by others [10–12]). The effects of IL6 on CD4⁺ T-cells have been explored extensively (reviewed in [50]). Specifically in naïve CD4⁺ T-cells, IL6 induces survival [51], proliferation [52] while memory CD4⁺ T-cells respond by expanding [53]. IL-6 also has a role in the balancing CD4⁺ T-cell differentiation between Treg/Th17 cells [54–56]. Importantly, T-cell migration into IL6-producing tissues is prevented by the expression of selectin-L (SELL/CD62L) [57], which is directly downregulated by IL6 [58], while favoured by upregulation of CXCR4/CXCR5/CCR3, which we previously showed to be increased on CD62-naïve CD4⁺ T-cells in RA [59]. The IL6R flow cytometry strategy that we used may have been impacted by the biology of the IL6R itself. Binding of IL6 to IL6R triggers a complex formation with gp130, leading to activation JAK/STAT signalling. However, it is also accompanied by the internalization of the IL-6/IL-6R/gp130-complex [60, 61]. Therefore, levels of the IL6R at the cell surface may reflect a balanced between recently and past IL6-activated cells, as well as resting cells. As such, it may not be surprising that we observed such a large distribution of results on total naïve cells. On the other hand, we showed clear differences in levels of expression on the subsets of CD62L⁻ naïve cells. From our analysis of DM surface molecules aiming at identifying the subpopulation of cells that underwent DM, we confirmed that the CD62L⁻ subpopulation of naïve cells also expressed different cell surface levels of DM genes (CD4, CD25). The SELL gene itself was too modestly DM to be considered on LIST-2, but it was directly functionally linked to the IL6/STAT3 node [62]. Furthermore, we previously reported on a similar CD62L⁻ subset differentiated from naïve CD4⁺ T-cell [63] which were hypothesised to results from exposure to IL6 (amplified by IL2/TNF) [58], with clinical significance in relation to the progression of RA from preclinical and early inflammatory stages [59, 64, 65].

Follicular T-cell (TFh) were also described recently [66]. These are also indirectly induced by IL6, via IL21. These TFh cells display CXCR5⁺/PD1^{high} phenotype [67, 68]. They have recently been observed in RA [69]. Both these markers were also modestly DM, and we previously

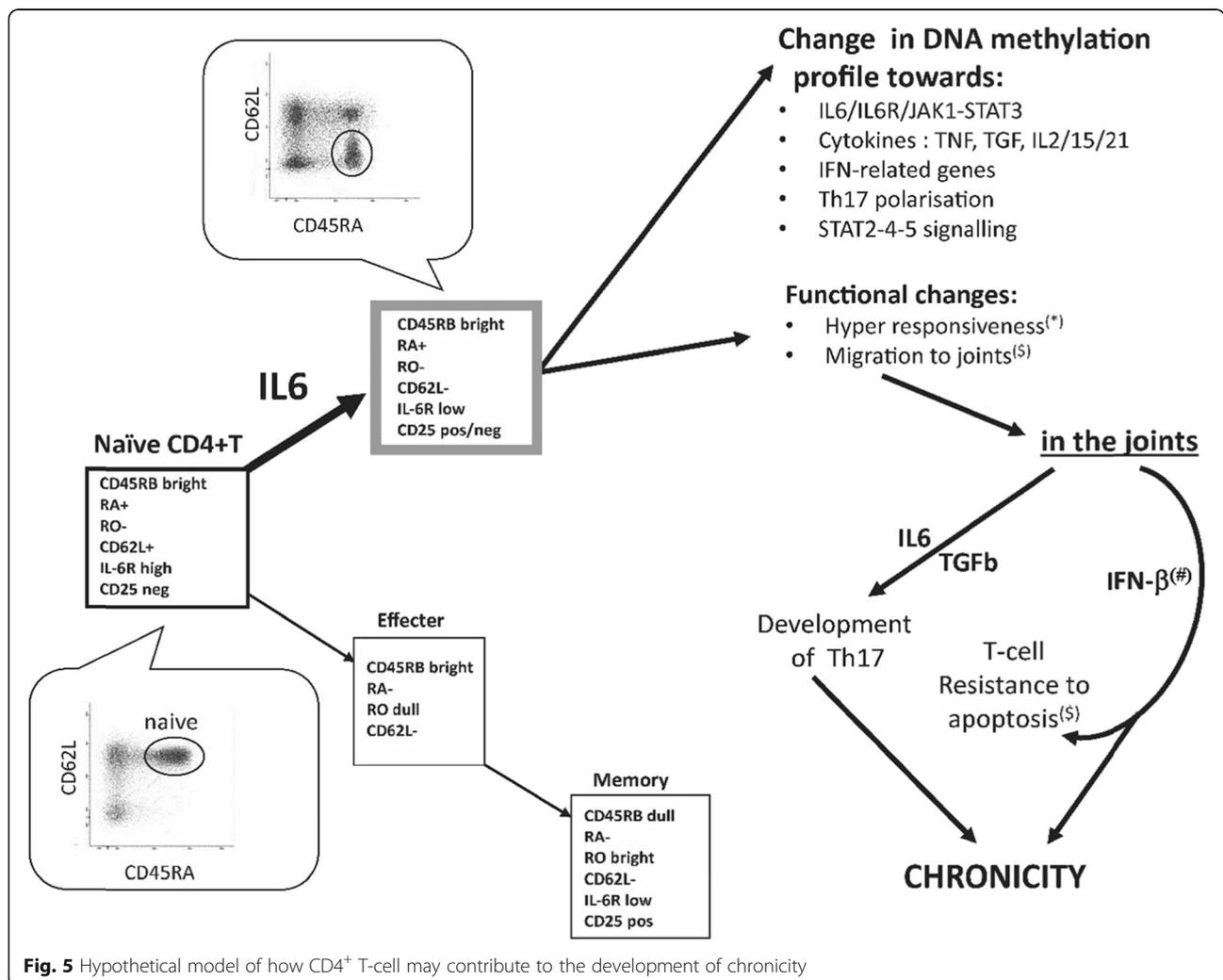
reported high CXCR5 expression specifically on CD62L⁻ naïve cells in RA [59]. These data support the potential of IL6/IL21 in generating CXCR5⁺ cells that may correspond to (i) CD62L⁻ naïve cells, and/or (ii) TFh also considering the IL21 ELISA results.

We therefore built an updated version of the model of T-cell differentiation proposed in 2002 (Fig. 5) in which we showed that differentiation was perturbed in early RA, possibly as a result of IL-6 activation of naïve T-cells [63], also observed by others [58, 70] and supported further [71]. We identified a subset of CD45RA⁺ cells which had lost CD62L expression (Fig. 5, grey box), in direct relationship with levels of inflammation [63]. This subset remains naïve with respect to antigen stimulation but showed hyper-responsiveness to stimulation (*) [63] and expression of chemokine receptors for trafficking to disease sites (\$), notably CXCR4/CXCR5/CCR3 [59]. The data presented here allows us to refine our original model, adding a possible mechanism for the IL-6 mediated effects via alterations in DNA methylation (as

reported in other conditions) [47, 58, 59, 63]. DM in genes may allow several pathways to become more accessible/primed (polarisation of Th17 cells), and modulate other signalling pathways (TNF, TGF, IL2/15/21). The effects of these changes combined with functional alteration in naïve CD62L⁻ T-cells may then allow these to migrate to the joints and contribute to the development of chronicity via the acquisition of resistance to apoptosis as previously suggested [72, 73] and the local maturation of Th17 cells.

Conclusion

Altogether, our data point to a role for IL6 signalling in early RA pathogenesis, with a central role in diversification towards other pathways (notably TNF, IFN-signalling, Th17 differentiation) very early in the disease course, which may already contribute to patient heterogeneity at this stage. Biological therapies approved for RA notably target TNF- α , IL6 and more recently JAKs. Our data showing a central role for JAK1/STAT3/IL6



also suggest that it may be relevant to use anti-IL6 and JAK-inhibitors as early as possible in the course of RA, maybe at a stage when it may still be possible to prevent the acquisition of further epigenetic changes.

Material and methods

Patients and ethics

Blood samples were obtained from patients recruited from an Early Inflammatory Arthritis clinic (Ethically approved REC: 09/H1307/98) including HCs. For the DNA methylation array, Patients were selected on the basis of fulfilling EULAR-2010 classification criteria, with < 6-month symptoms, naïve for anti-rheumatic drug, with active disease and at least 2 swollen joints and a CRP > 10 mg/L. For bisulfite sequencing, ELISA and flow cytometry, we used consecutive drug-naïve RA patients attending the early arthritis clinic for the first time (diagnosed RA at first visit). Each participant gave informed written consent. Supplementary Tables S1 provide demographic and clinical details of the patients and controls used in the different lines of work.

Experimental workflow (all details provided in supplementary material)

- PBMC were recovered from 30 ml of EDTA blood (using standard lymphoprep procedures) and naïve (CD45RA⁺ CD45RO⁻) and memory (CD45RA⁻ CD45RO⁺) CD4⁺ T-cell and monocytes (CD14⁺) were purified by cell sorting.
- Bisulfite converted DNA was hybridised to the Illumina Human Methylation 450K Beadchips.
- A standard CpG methylation data normalisation pipeline was applied.
- Different strategies to prioritise differentially methylated (DM) CpGs were applied to identify DM CpGs or DM regions and the associated genes.
- Commercial ELISAs were used to confirm the differential expression of 3 cytokines whose genes were DM (IL21, IL34 and RANKL).
- The TNF gene was chosen for confirmation using bisulfite DNA sequencing.
- Publically available gene expression datasets for CD4⁺ T-cells from early, drug naïve RA and HC were retrieved to examine associated differential expression of DM genes.
- Flow cytometry was used to assess levels of expression of 5 DM genes for cell surface proteins (CD4, CD25, CD62L, CXCR4, IL6R) towards identifying subpopulation(s) of cells affected.

Methylation array

PBMC were recovered from 30 mL of EDTA blood. Naïve (CD45RA⁺/CD45RO⁻) and memory (CD45RA⁻/CD45RO⁺)

CD4⁺ T-cell and monocytes (CD14⁺) were purified by cell sorting following antibody staining using standards protocols: anti-CD4 (Clone RPA-T4, BD), anti-CD3 (Clone RPA-T8, BD), CD45RA (Clone MEM55, Serotec), CD45RO (Clone UCHL1, Serotec), anti-CD14 (Clone M5E2, BD). DNA was extracted from purified cell subsets using the Nucleon extraction kit according to manufacturers' instructions. The concentration of genomic DNA was assessed by NanoDrop. Genomic DNA (650 ng) was bisulfite converted using the Zymo EZ DNA Methylation™ Kit. Bisulfite converted DNA was amplified using the Illumina Infinium Methylation Assay and hybridised to Illumina Human Methylation 450K Beadchips before scanning on the Illumina iScan microarray scanner [74]. All procedures were performed by Hologic Ltd. (Manchester, UK).

DNA methylation data analysis

A standard data analysis pipeline utilised a combination of R [75], bioconductor and custom scripts was designed (supplementary Figure S1). A total of 48 genome-wide DNA methylation profiles (from 3 cells subset of 10 HC and 6 early RA patient) were retrieved as idat files. Data quality control analysis and preprocessing were performed with the R package Minfi [76]. Plots of β values for density including all 48 samples, bean plots for each individual samples and strip plots for array control probes were generated using the same R package. Of the 48 samples, 2 failed quality control due to poor DNA quality or concentration (details in supplementary Figure S2). CpG probes which were identified to be common SNPs and cross-reactive probes that have been shown to hybridise to multiple locations in the genomes were filtered out [77]. Methylation levels for each CpG site were presented as β value or M value according to the analysis to be performed. β values are the ratio of the fluorescence intensity between the methylated and unmethylated probes, ranging from 0 (all copies of the CpG in the sample are unmethylated) to 1 (all copies of the CpG in the sample are fully methylated). M values are log-transformed β values preferably used for statistical testing [78].

Multidimensional scaling (MDS) for (i) each cell type, (ii) genders, and (iii) RA and HC was performed to examine the source of variation in the dataset and were plotted using the MDS plot function in the minfi package in R [76]. Two-sided t-tests (on M value) were performed on every CpG using the function rowttest in the genefilter package [79] for significance of the difference in methylation between HC and RA. The Log(p value) was calculated and data were presented using Manhattan plots generated using the qqman package [80]. Heatmaps were created using the heatmap2 function of the gplot package [81]. To identify clusters of DM CpG, a custom R scripts to score each individual CpG and prioritise them was developed (details in the results section,

full code available on request). Annotation related to each CpG (location in islands/open sea, associated gene symbols) were retrieved using getAnnotation in the minfi package [76] and the getNearestTSS function in FDb.InfiniumMethylation.hg19 package [82].

Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s13148-020-00837-1>.

Additional file 1: Supplementary Data 1-3: Full list of gene associated with DM-CpG-Cluster and DM-isolated-CpG in naïve, memory T-cells, and monocyte (3 Excel files).

Additional file 2: Supplementary LIST-1(1 Excel sheet), Supplementary LIST-2 (1 Excel sheet), Supplementary LIST-3 (1 Excel sheet), Supplementary LIST-4: Full list of gene included in the STRING analysis for naïve CD4+T-cells. (1 Excel sheet).

Additional file 3: Supplementary methods and results.

Abbreviations

AIDs: Autoimmune diseases; $\Delta\beta$ values: Difference in β values; DEG: Differentially expressed genes; DM: Differential methylation or differentially methylated; DMR: Differentially methylated regions; ELISA: Enzyme-linked immunosorbent assays; HC: Healthy controls; MDS: Multidimensional scaling; MFI: Mean fluorescence intensity; RA: Rheumatoid arthritis

Acknowledgements

The authors would like to thank Liz Straszynski and Adan Davisson for technical support with cell sorting. Two ERASMUS placement students also contributed to the sequencing data (Martin Leconte and Laura Dutto).

Authors' contributions

Conceptualization and study design: FP, IBM. Methodology and data acquisition: RP, RP, XX, KM, JSR. Software: RP, RH. Formal analysis: FP, RP, ANB, SA, PC, AD. Writing of original draft: RP, FP, ANB, SA. Writing, review and editing: RP, ANB, JSN, AD, IBM, PE, RH, FP. Funding acquisition: FP IBM, PE. Supervision: FP, ANB, RH. The authors read and approved the final manuscript.

Funding

This work has been directly funded by a European Union-funded FP7-integrated project EURO-TEAM No. 305549 and partly supported by the IMI-funded project BeTheCure No 115142-2.

Availability of data and materials

The raw DNA methylation data have been deposited to the Gene Expression Omnibus (GEO) under accession numbers GEO: GSE121192. Custom R code for prioritising DM CpG is available on request.

Ethics approval and consent to participate

Patients were recruited from an Early Inflammatory Arthritis clinic (ethically approved REC: 09/H1307/98) including HCs. Each participant gave informed written consent.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Leeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds, Leeds, UK. ²Leeds Institute for Data Analytics, University of Leeds, Leeds, UK. ³Institute of Infection Immunity and Inflammation, University of Glasgow, Glasgow, UK. ⁴Leeds University Business School, University of Leeds, Leeds, UK.

Received: 20 December 2019 Accepted: 8 March 2020

Published online: 07 April 2020

References

- Wolfe F, Mitchell DM, Sibley JT, Fries JF, Bloch DA, Williams CA, et al. The mortality of rheumatoid arthritis. *Arthritis Rheum.* 1994;37(4):481–94.
- Pincus T, Callahan LF. The 'side effects' of rheumatoid arthritis: joint destruction, disability and early mortality. *Br J Rheumatol.* 1993;32(Suppl 1): 28–37.
- Zheng SC, Widschwendter M, Teschendorff AE. Epigenetic drift, epigenetic clocks and cancer risk. *Epigenomics.* 2016;8(5):705–19.
- Firestein GS, McInnes IB. Immunopathogenesis of rheumatoid arthritis. *Immunity.* 2017;46(2):183–96.
- Neidhart M, Rethage J, Kuchen S, Kunzler P, Crowl RM, Billingham ME, et al. Retrotransposable L1 elements expressed in rheumatoid arthritis synovial tissue. *Arthritis Rheum.* 2000;43:2634–47.
- Rhead B, Hologue C, Cole M, Shao X, Quach HL, Quach D, et al. Rheumatoid arthritis naïve T cells share hypermethylation sites with synovial cells. *Arthritis Rheum.* 2017;69(3):550–9.
- Ballestar E, Li TL. New insights into the epigenetics of inflammatory rheumatic diseases. *Nat Rev Rheumatol.* 2017;13(10):593–605.
- Nair N, Plant D, Verstappen SM, Isaacs JD, Morgan AW, Hyrich KL, et al. Differential DNA methylation correlates with response to methotrexate in rheumatoid arthritis. *Rheumatology (Oxford).* 2019.
- Ponchel F, Burska AN. Epigenetic modifications: Are we closer to clinical applicability? *Journal of Pharmacogenomics & Pharmacoproteomics.* 2016;07(02).
- Straub RH, Paimela L, Peltomaa R, Scholmerich J, Leirisalo-Repo M. Inadequately low serum levels of steroid hormones in relation to interleukin-6 and tumor necrosis factor in untreated patients with early rheumatoid arthritis and reactive arthritis. *Arthritis Rheum.* 2002;46(3):654–62.
- Baillet A, Gossec L, Paternotte S, Etcheto A, Combe B, Meyer O, et al. Evaluation of serum interleukin-6 level as a surrogate marker of synovial inflammation and as a factor of structural progression in early rheumatoid arthritis: results from a French national multicenter cohort. *Arthritis Care Res.* 2015;67(7):905–12.
- Madhok R, Crilly A, Watson J, Capell HA. Serum interleukin-6 levels in rheumatoid-arthritis - correlations with clinical and laboratory indexes of disease-activity. *Ann Rheum Dis.* 1993;52(3):232–4.
- Nile CJ, Read RC, Akil M, Duff GW, Wilson AG. Methylation status of a single CpG site in the IL6 promoter is related to IL6 messenger RNA levels and rheumatoid arthritis. *Arthritis Rheum.* 2008;58(9):2686–93.
- Peters TJ, Buckley MJ, Statham AL, Pidsley R, Samarasinghe K, V Lord R, et al. De novo identification of differentially methylated regions in the human genome. *Epigenetics Chromatin.* 2015;8(1):6.
- Nakano K, Whitaker JW, Boyle DL, Wang W, Firestein GS. DNA methylome signature in rheumatoid arthritis. *Ann Rheum Dis.* 2013;72(1):110–7.
- Fu LH, Ma CL, Cong B, Li SJ, Chen HY, Zhang JG. Hypomethylation of proximal CpG motif of interleukin-10 promoter regulates its expression in human rheumatoid arthritis. *Acta Pharmacol Sin.* 2011;32(11):1373–80.
- Ishida K, Kobayashi T, Ito S, Komatsu Y, Yokoyama T, Okada M, et al. Interleukin-6 gene promoter methylation in rheumatoid arthritis and chronic periodontitis. *J Periodontol.* 2012;83(7):917–25.
- de la Rica L, Urquiza JM, Gomez-Cabrero D, Islam ABMMK, Lopez-Bigas N, Tegner J, et al. Identification of novel markers in rheumatoid arthritis through integrated analysis of DNA methylation and microRNA expression. *J Autoimmun.* 2013;41:6–16.
- Pratt AG, Swan DC, Richardson S, Wilson G, Hilken CM, Young DA, et al. A CD4 T cell gene signature for early rheumatoid arthritis implicates interleukin 6-mediated STAT3 signalling, particularly in anti-citrullinated peptide antibody-negative disease. *Ann Rheum Dis.* 2012.
- Gruden K, Hren M, Herman A, Blejec A, Albrecht T, Selbig J, et al. A "crossomics" study analysing variability of different components in peripheral blood of healthy caucasoid individuals. *PLoS One.* 2012;7(1):e28761.
- Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, et al. STRING v10: protein–protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* 2015;43(Database issue):D447–D52.
- Kandasamy K, Mohan SS, Raju R, Keerthikumar S, Kumar GS, Venugopal AK, et al. NetPath: a public resource of curated signal transduction pathways. *Genome Biol.* 2010;11(1):R3.
- Tian J, Chau C, Hales TG, Kaufman DL. GABAA receptors mediate inhibition of T cell responses. *J Neuroimmunol.* 1999;96(1):21–8.

24. Mendu SK, Bhandage A, Jin Z, Birnir B. Different subtypes of GABA-A receptors are expressed in human, mouse and rat T lymphocytes. *PLoS One*. 2012;7(8).
25. Tian JD, Yong J, Dang H, Kaufman DL. Oral GABA treatment downregulates inflammatory responses in a mouse model of rheumatoid arthritis. *Autoimmunity*. 2011;44(6):465–70.
26. Veldhoen M, Hocking RJ, Atkins CJ, Locksley RM, Stockinger B. TGF β in the context of an inflammatory cytokine milieu supports de novo differentiation of IL-17-producing T cells. *Immunity*. 2006;24(2):179–89.
27. Serada S, Fujimoto M, Mihara M, Koike N, Ohsugi Y, Nomura S, et al. IL-6 blockade inhibits the induction of myelin antigen-specific Th17 cells and Th1 cells in experimental autoimmune encephalomyelitis. *Proc Natl Acad Sci U S A*. 2008;105(26):9041–6.
28. Korn T, Bettelli E, Gao W, Awasthi A, Jager A, Strom TB, et al. IL-21 initiates an alternative pathway to induce proinflammatory T(H)17 cells. *Nature*. 2007;448(7152):484–7.
29. Langrish CL, Chen Y, Blumenschein WM, Mattson J, Basham B, Sedgwick JD, et al. IL-23 drives a pathogenic T cell population that induces autoimmune inflammation. *J Exp Med*. 2005;201(2):233–40.
30. Sutton CE, Lalor SJ, Sweeney CM, Brereton CF, Lavelle EC, Mills KH. Interleukin-1 and IL-23 induce innate IL-17 production from gammadelta T cells, amplifying Th17 responses and autoimmunity. *Immunity*. 2009;31(2):331–41.
31. Van Roon J, Glaudemans K, Bijlsma J, Lafeber F. Interleukin 7 stimulates tumour necrosis factor α and Th1 cytokine production in joints of patients with rheumatoid arthritis. *Ann Rheum Dis*. 2003;62(2):113–9.
32. Kawashima M, Miossec P. Defect of Th1 immune response of whole blood cells from active patients with rheumatoid arthritis (RA). *Arthritis Res Ther*. 2003;5:S14.
33. Kawashima M, Miossec P. Effect of treatment of rheumatoid arthritis with infliximab on IFN gamma, IL4, T-bet, and GATA-3 expression: link with improvement of systemic inflammation and disease activity. *Ann Rheum Dis*. 2005;64(3):415–8.
34. Muskardin TLW, Niewold TB. Type I interferon in rheumatic diseases. *Nat Rev Rheumatol*. 2018;14(4):214–28.
35. Rodriguez-Carrio J, Alperi-Lopez M, Lopez P, Ballina-Garcia FJ, Suarez A. Heterogeneity of the type I interferon signature in rheumatoid arthritis: a potential limitation for its use as a clinical biomarker. *Front Immunol*. 2017;8:2007.
36. Psarras A, Emery P, Vital EM. Type I interferon-mediated autoimmune diseases: pathogenesis, diagnosis and targeted therapy. *Rheumatology (Oxford)*. 2017;56(10):1662–75.
37. Ronnblom L. The importance of the type I interferon system in autoimmunity. *Clin Exp Rheumatol*. 2016;34(4 Suppl 98):21–4.
38. Lübbers J, Brink M, van de Stadt LA, Vosslander S, Wesseling JG, van Schaardenburg D, et al. The type I IFN signature as a biomarker of preclinical rheumatoid arthritis. *Ann Rheum Dis*. 2013;72(5):776–80.
39. de Jong TD, Lübbers J, Turk S, Vosslander S, Mantel E, Bontkes HJ, et al. The type I interferon signature in leukocyte subsets from peripheral blood of patients with early arthritis: a major contribution by granulocytes. *Arthritis Res Ther*. 2016;18:165.
40. van Baarsen LG, Bos WH, Rustenburg F, van der Pouw Kraan TC, Wolbink GJ, Dijkman BA, et al. Gene expression profiling in autoantibody-positive patients with arthralgia predicts development of arthritis. *Arthritis Rheum*. 2010;62(3):694–704.
41. de Jong TD, Blits M, de Ridder S, Vosslander S, Wolbink G, Nurmohamed MT, et al. Type I interferon response gene expression in established rheumatoid arthritis is not associated with clinical parameters. *Arthritis Res Ther*. 2016;18(1):290.
42. Castaneda-Delgado JE, Bastian-Hernandez Y, Macias-Segura N, Santiago-Algarra D, Castillo-Ortiz JD, Aleman-Navarro AL, et al. Type I interferon gene response is increased in early and established rheumatoid arthritis and correlates with autoantibody production. *Front Immunol*. 2017;8:285.
43. Bilgic H, Ytterberg SR, Amin S, McNallan KT, Wilson JC, Koeuth T, et al. Interleukin-6 and type I interferon-regulated genes and chemokines mark disease activity in dermatomyositis. *Arthritis Rheum*. 2009;60(11):3436–46.
44. Zimmermann M, Arruda-Silva F, Bianchetto-Aguilera F, Finotti G, Calzetti F, Scapini P, et al. IFN α enhances the production of IL-6 by human neutrophils activated via TLR8. *Sci Rep*. 2016;6:19674.
45. Yoshida Y, Tanaka T. Interleukin 6 and rheumatoid arthritis. *Biomed Res Int*. 2014;2014.
46. Tanaka Y, Martin ME. IL-6 targeting compared to TNF targeting in rheumatoid arthritis: studies of olokizumab, sarilumab and sirukumab. *Ann Rheum Dis*. 2014;73(9):1595–7.
47. Gasche JA, Hoffmann J, Boland CR, Goel A. Interleukin-6 promotes tumorigenesis by altering DNA methylation in oral cancer cells. *Int J Cancer*. 2011;129(5):1053–63.
48. Garaud S, Le Dantec C, Jousse-Joulin S, Hanrotel-Saliou C, Saraux A, Mageed RA, et al. IL-6 modulates CD5 expression in B cells from patients with lupus by regulating DNA methylation. *J Immunol*. 2009;182(9):5623–32.
49. Lal G, Zhang N, van der Touw W, Ding YZ, Ju WJ, Bottinger EP, et al. Epigenetic regulation of Foxp3 expression in regulatory T cells by DNA methylation. *J Immunol*. 2009;182(1):259–73.
50. Dienz O, Rincon M. The effects of IL-6 on CD4 T cell responses. *Clin Immunol*. 2009;130(1):27–33.
51. Teague TK, Marrack P, Kappler JW, Vella AT. IL-6 rescues resting mouse T cells from apoptosis. *J Immunol*. 1997;158(12):5791–6.
52. Lotz M, Jirik F, Kabouridis P, Tsoukas C, Hirano T, Kishimoto T, et al. B cell stimulating factor 2/interleukin 6 is a costimulant for human thymocytes and T lymphocytes. *J Exp Med*. 1988;167(3):1253–8.
53. Rochman I, Paul WE, Ben-Sasson S. IL-6 increases primed cell expansion and survival. *J Immunol*. 2005;174(8):4761–7.
54. Ivanov II, McKenzie BS, Zhou L, Tadokoro CE, Lepelley A, Lafaille JJ, et al. The orphan nuclear receptor ROR gamma t directs the differentiation program of proinflammatory IL-17(+) T helper cells. *Cell*. 2006;126(6):1121–33.
55. Veldhoen M, Hocking RJ, Atkins CJ, Locksley RM, Stockinger B. TGF beta in the context of an inflammatory cytokine milieu supports de novo differentiation of IL-17-producing T cells. *Immunity*. 2006;24(2):179–89.
56. Bettelli E, Carrier YJ, Gao WD, Korn T, Strom TB, Oukka M, et al. Reciprocal developmental pathways for the generation of pathogenic effector T(H)17 and regulatory T cells. *Nature*. 2006;441(7090):235–8.
57. Weissenbach M, Claahsen T, Weber C, Spitzer D, Wirth D, Vestweber D, et al. Interleukin-6 is a direct mediator of T cell migration. *Eur J Immunol*. 2004;34(10):2895–906.
58. Unutmaz D, Baldoni F, Abrignani S. Human naive T cells activated by cytokines differentiate into a split phenotype with functional features intermediate between naive and memory T cells. *Int Immunol*. 1995;7(9):1417–24.
59. Burgoyne CH, Field SL, Brown AK, Hensor EM, English A, Bingham SL, et al. Abnormal T cell differentiation persists in patients with rheumatoid arthritis in clinical remission and predicts relapse. *Ann Rheum Dis*. 2008;67(6):750–7.
60. Thiel S, Sommer U, Kortylewski M, Haan C, Behrmann I, Heinrich PC, et al. Termination of IL-6-induced STAT activation is independent of receptor internalization but requires de novo protein synthesis. *FEBS Lett*. 2000;470(1):15–9.
61. Thiel S, Dahmen H, Martens A, Muller-Newen G, Schaper F, Heinrich PC, et al. Constitutive internalization and association with adaptor protein-2 of the interleukin-6 signal transducer gp130. *FEBS Lett*. 1998;441(2):231–4.
62. Suwa T, Hogg JC, Quinlan KB, Van Eeden SF. The effect of interleukin-6 on L-selectin levels on polymorphonuclear leukocytes. *Am J Physiol Heart Circ Physiol*. 2002;283(3):H879–84.
63. Ponchel F, Morgan AW, Bingham SJ, Quinn M, Buch M, Verburg RJ, et al. Dysregulated lymphocyte proliferation and differentiation in patients with rheumatoid arthritis. *Blood*. 2002;100(13):4550–6.
64. Ponchel F, Goeb V, Parmar R, El-Sherbiny Y, Boissinot M, El Jawhari J, et al. An immunological biomarker to predict MTX response in early RA. *Ann Rheum Dis*. 2014;73(11):2047–53.
65. Hunt L, Hensor EM, Nam J, Burska AN, Parmar R, Emery P, et al. T cell subsets: an immunological biomarker to predict progression to clinical arthritis in ACPA-positive individuals. *Ann Rheum Dis*. 2016;75(10):1884–9.
66. Chtanova T, Tangye SG, Newton R, Frank N, Hodge MR, Rolph MS, et al. T follicular helper cells express a distinctive transcriptional profile, reflecting their role as non-Th1/Th2 effector cells that provide help for B cells. *J Immunol*. 2004;173(1):68–78.
67. Johnston RJ, Poholek AC, DiToro D, Yusuf I, Eto D, Barnett B, et al. Bcl6 and Blimp-1 are reciprocal and antagonistic regulators of T follicular helper cell differentiation. *Science*. 2009;325(5943):1006–10.
68. Nurieva RI, Chung Y, Hwang D, Yang XO, Kang HS, Ma L, et al. Generation of T follicular helper cells is mediated by interleukin-21 but independent of T helper 1, 2, or 17 cell lineages. *Immunity*. 2008;29(1):138–49.
69. Yu M, Cavero V, Lu Q, Li H. Follicular helper T cells in rheumatoid arthritis. *Clin Rheumatol*. 2015;34(9):1489–93.

70. Unutmaz D, Abrignani S. Cytokines can activate resting T-lymphocytes. *Chall Mod Med*. 1994;8:49–52.
71. Jones GW, McLoughlin RM, Hammond VJ, Parker CR, Williams JD, Malhotra R, et al. Loss of CD4(+) T cell IL-6R expression during inflammation underlines a role for IL-6 trans signaling in the local maintenance of Th17 Cells. *J Immunol*. 2010;184(4):2130–9.
72. Pilling D, Akbar AN, Girdlestone J, Orteu CH, Borthwick NJ, Amft N, et al. Interferon-beta mediates stromal cell rescue of T cells from apoptosis. *Eur J Immunol*. 1999;29(3):1041–50.
73. Firestein GS, Zvaifler NJ. How important are T cells in chronic rheumatoid synovitis? II. T cell-independent mechanisms from beginning to end. *Arthritis Rheum*. 2002;46(2):298–308.
74. Bibikova M, Barnes B, Tsan C, Ho V, Klotzle B, Le JM, et al. High density DNA methylation array with single CpG site resolution. *Genomics*. 2011;98(4):288–95.
75. R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2017.
76. Aryee MJ, Jaffe AE, Corrada-Bravo H, Ladd-Acosta C, Feinberg AP, Hansen KD, et al. Minfi: a flexible and comprehensive bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics*. 2014; 30(10):1363–9.
77. Chen YA, Lemire M, Choufani S, Butcher DT, Grafodatskaya D, Zanke BW, et al. Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. *Epigenetics*. 2013;8(2):203–9.
78. Du P, Zhang X, Huang CC, Jafari N, Kibbe WA, Hou L, et al. Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. *BMC bioinformatics*. 2010;11:587.
79. Gentleman R, Carey V, Huber W, Hahne F. Genefilter: genefilter: methods for filtering genes from high-throughput experiments. R package version 1.581. 2017.
80. Turner SD. qqman: an R package for visualizing GWAS results using Q-Q and Manhattan plots. *bioRxiv*. 2014.
81. Gregory RW, Ben B, Lodewijk B, Robert G, Wolfgang HAL, Thomas L, et al. gplots: Various R programming tools for plotting data. R package version 3.01. 2016.
82. Tim T, Jr. FDb.InfiniumMethylation.hg19: annotation package for Illumina Infinium DNA methylation probes. R package version 2.20. 2014.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

