View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Nottingham ePrints

The University of Nottingham

UNITED KINGDOM · CHINA · MALAYSIA

Glossop, John R. and Emes, Richard D. and Nixon, Nicola B. and Packham, Jon C. and Fryer, Anthony A. and Mattey, Derek L. and Farrell, William E. (2016) Genome-wide profiling in treatment-naive early rheumatoid arthritis reveals DNA methylome changes in T and B lymphocytes. Epigenomics, 8 (2). pp. 209-224. ISSN 1750-192X

Access from the University of Nottingham repository:
http://eprints.nottingham.ac.uk/30944/1/Epigenomics%202015-2.pdf

Copyright and reuse:

The Nottingham ePrints service makes this work by researchers of the University of Nottingham available open access under the following conditions.

This article is made available under the University of Nottingham End User licence and may be reused according to the conditions of the licence. For more details see: http://eprints.nottingham.ac.uk/end_user_agreement.pdf

A note on versions:

The version presented here may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the repository url above for details on accessing the published version and note that access may require a subscription.

For more information, please contact eprints@nottingham.ac.uk

Genome-wide profiling in treatment-naïve early rheumatoid arthritis reveals DNA methylome changes in T- and B-lymphocytes

John R. Glossop[1,2], Richard D. Emes[3,4], Nicola B. Nixon [2], Jon C. Packham[2], Anthony A. Fryer[1], Derek L. Mattey[1,2], William E. Farrell[1].

[1]Institute for Science and Technology in Medicine, Keele University, Guy Hilton Research Centre, Thornburrow Drive, Hartshill, Stoke-on-Trent, Staffordshire, ST4 7QB, UK.

[2]Haywood Rheumatology Centre, Haywood Hospital, High Lane, Burslem, Stoke-on-Trent, Staffordshire, ST6 7AG, UK.

[3]School of Veterinary Medicine and Science, University of Nottingham, Sutton Bonington Campus, Sutton Bonington, Leicestershire, LE12 5RD, UK.

[4]Advanced Data Analysis Centre, University of Nottingham, Sutton Bonington Campus, Sutton Bonington, Leicestershire, LE12 5RD, UK.

Author for correspondence/re-prints: Dr. John. R. Glossop, Institute for Science and Technology in Medicine, Keele University, Haywood Rheumatology Centre, Haywood Hospital, High Lane, Burslem, Stoke-on-Trent, Staffordshire, ST6 7AG, UK.

Tel: +44 (0) 1782 673812

Fax: +44 (0) 1782 673509

E-mail: j.r.glossop@keele.ac.uk

Key words: Early rheumatoid arthritis; Treatment-naïve; DNA methylation; T-lymphocyte; B-lymphocyte; CpG; Genome-wide; Illumina 450K array.

Funding support: This work was supported by funding from the Haywood Rheumatism Research and Development Foundation.

Running footline: DNA methylation in treatment-naïve RA

Abstract

Aims. Although aberrant DNA methylation has been described in rheumatoid arthritis (RA), no studies have interrogated this epigenetic modification in early disease. Following recent investigations of T- and B-lymphocytes in established disease, we now characterize in these cell populations genome-wide DNA methylation in treatment-naïve patients with early RA.

Patients & Methods. HumanMethylation450 BeadChips were used to examine genome-wide DNA methylation in lymphocyte populations from 23 early RA patients and 11 healthy individuals.

Results. Approximately 2,000 CpGs in each cell type were differentially methylated in early RA. A novel methylation signature in each cell type (150 sites in T-lymphocytes, 113 sites in B-lymphocytes) perfectly distinguished patients from controls. A subset of sites differentially methylated in early RA displayed similar changes in established disease.

Conclusions. Treatment-naïve early RA patients display novel disease-specific DNA methylation aberrations, supporting an important role for these changes in the pathogenesis of RA.

Introduction

Rheumatoid arthritis (RA) is a common autoimmune inflammatory joint disease that affects between 0.5 and 1% of the adult population [1, 2]. Estimates of disease incidence for the UK indicate in the region of 20,000 new RA cases each year [3], the majority of whom begin treatment with one or more disease modifying anti-rheumatic drugs. The disease is characterized by chronic inflammation and progressive destruction of synovial joint tissues, and a compelling literature now implicates alterations to the epigenome in the pathogenesis of RA. These alterations are multifactorial and include those that impact on DNA methylation, histone modifications and the expression of microRNAs [reviewed in 4-6].

DNA methylation describes the addition of a methyl group to the nucleobase cytosine, most frequently in the context of cytosine-guanine dinucleotides (CpGs). Inappropriate methylation of these sites in promoter-associated regions is frequently associated with gene silencing, whereas in transcriptionally competent genes these sites are commonly not methylated [7]. Changes to the DNA methylome have been described in multiple cell types in patients with RA. Global levels of methylation, for example, are altered in peripheral blood mononuclear cells and lymphocytes from patients with RA [8-11]. In addition, candidate gene studies also describe aberrant methylation across multiple, discrete, genes in these cell types [12-18]. Further to these studies, and in this case using genome-wide methylation arrays, we recently reported disease-associated methylation changes in multiple genes that were distinct to individual T- and B-lymphocyte populations [19]. Similar findings in the context of alterations to the epigenome are also

apparent in fibroblast-like synoviocytes derived from the joint [20-27], which further supports an important role for altered methylation in the pathogenesis of RA.

To date, studies of DNA methylation have been reported for patients with established RA, in many of these cases at the time of joint replacement. These patients have typically been treated over prolonged periods with a variety of disease-modifying agents, including methotrexate, a cornerstone of therapy in RA. However, a potential impact of methotrexate on DNA methylation has been suggested in RA and similar conditions [7, 28, 29]. Changes to the epigenome identified in established RA may therefore reflect, at least in part, the impact of treatment and the burden of chronic inflammation present during the evolution and progression of the disease. Thus, and as noted by others [30], such confounders have precluded identification of methylation changes that are important drivers of disease in RA. Indeed, to our knowledge, there has been no systematic examination of DNA methylation in patients with early disease, or in those at first diagnosis and who are naïve for treatment with disease-modifying anti-rheumatic drugs.

In this report, and in contrast to previous studies from our own and other laboratories, we have investigated genome-wide DNA methylation in T- and B-lymphocyte populations from newly diagnosed, treatment-naïve patients with RA. Our studies in these patients mitigate the likely impact of disease evolution and of different treatment modalities and their associated sequela apparent in established disease. In this way, the identified methylation changes are likely to reflect those associated with disease development and pathogenesis.

Patients & Methods

Patients

Genome-wide DNA methylation was investigated in 23 Caucasian patients recruited consecutively from the early synovitis clinic at the Haywood Rheumatology Centre in Stoke-on-Trent, UK, and also in a group of 11 healthy Caucasian females, the details of which we have previously reported [19]. All patients had symptomatic inflammatory arthritis suspected to be RA at presentation, and subsequently were classified with RA by a consultant rheumatologist and according to the 2010 ACR/EULAR classification criteria [31]. Demographic data and clinical characteristics for the patients are presented in **Table 1**. All patients were naïve for treatment with disease modifying anti-rheumatic drugs and biological therapies. With two exceptions, all patients had active disease at recruitment based on a disease activity score with 28-joint counts (DAS28) [32] ≥ 3.2, and approximately half (44%) had high disease activity (DAS28 > 5.1). The average age of the patients and healthy control group was 58.9 ± 15.8 years and 52.6 ± 4.9 years, respectively (mean ± SD). All study participants gave written informed consent. The study was approved by the East Midlands (Derby) Research Ethics Committee.

Purification of T- and B-lymphocyte populations

Whole blood samples (35 ml, EDTA) were collected from each patient at recruitment and prior to the start of treatment. Peripheral blood mononuclear cells were prepared using density-gradient centrifugation (Histopaque-1077, Sigma Aldrich). Thereafter, $CD19^+$ B-lymphocytes and $CD3^+$ T-lymphocytes were isolated from the mononuclear cell fraction by positive selection using anti-CD19 and anti-CD3 magnetic microbeads, respectively. All isolations were performed according to the manufacturer's instructions

(MiniMACS® Separation System; Miltenyi Biotec). Cells were pelleted, lysed (Buffer RLT Plus; Qiagen), and homogenized (QIAShredder spin columns; Qiagen) prior to storage at -80°C. We have previously demonstrated this method of cell isolation to yield high purity T-lymphocyte (≥99%) and B-lymphocyte (≥90%) populations from patients with RA and healthy individuals alike [19, 33].

Genome-wide DNA methylation profiling

HumanMethylation450 BeadChips (Illumina Inc.; hereafter referred to as 'array' or '450K array') were used to profile genome-wide DNA methylation at more than 480,000 CpG sites in lymphocyte populations from patients and healthy individuals. Genomic DNA was extracted from cell lysates according to the manufacturer's instructions using an AllPrep DNA/RNA/miRNA Universal kit (Qiagen), with further concentration by ethanol-based precipitation. DNA samples were sodium bisulfite treated and hybridized to 450K arrays according to manufacturer recommended protocols (performed by Hologic Tepnel Pharma Services, UK) that we have described previously [33]. Methylation at individual CpGs is reported as a β-value [34] on a linear scale, where 0 and 1 indicate unmethylated and completely methylated sites, respectively.

Sodium bisulfite pyrosequencing

Independent quantification of methylation at array-identified candidate CpG loci was conducted by sodium bisulfite pyrosequencing. For this purpose, amplicons containing CpGs of interest were prepared from whole-genome amplified, bisulfite-converted DNA using a touchdown PCR process, as we have described previously [33]. Oligonucleotide

primers for all PCR and pyrosequencing reactions were designed using PyroMark Assay Design software (Qiagen) and were specific for bisulfite-converted DNA. Primers were purchased from Biomers.net and assay details are provided in **Supplementary Table 1**. Pyrosequencing was performed according to the manufacturer's recommendations using a PyroMark Q24 instrument (Qiagen). Assays included one or more control dispensations to confirm successful bisulfite conversions. All data were analyzed using PyroMark Q24 software (v 2.0.6., build 20; Qiagen).

Data analysis

Processing and analysis of array data to determine methylation differences between treatment-naive early RA patients and healthy individuals was performed using idat files and the Bioconductor package Minfi [35]. Each array passed initial quality control assessment based on the performance of internal array controls. For each of the T- and B-lymphocyte array datasets, we excluded from analysis all probes located on the X and Y chromosomes and probes with a detection p-value > 0.01 in any one or more of the samples (n = 34 [23 patients and 11 healthy controls]). Array data were subsequently normalized using Subset-quantile Within Array Normalization (SWAN) [36], an approach which has been shown to reduce technical variation both within and between arrays, including the technical differences observed between the two different probe types on the array [34, 37]. Following these steps, a total of 470,366 and 473,105 CpGs in T-lymphocytes and B-lymphocytes, respectively, were retained and were used in subsequent analyses.

For each of the cell types separately, the 'dmpFinder' function in Minfi was used to identify CpGs that were differentially methylated between treatment-naïve early RA patients and healthy individuals (**step 1, Figure 1**). This function performs an F-test to test the equality of means and was used with logit transformed β-values (M-values), as recommended by Du *et al* [38]. The false discovery rate (FDR) [39] was used to adjust for multiple testing, with adjusted p-values < 0.05 considered statistically significant.

In addition to the analysis described above, sites identified as differentially methylated in treatment-naïve early RA patients relative to healthy individuals were further examined through comparison with methylation in patients with established disease. In these cases, we used the processing and normalization procedures described above to re-analyze the 450K array data derived from a cohort of established RA patients (n = 12) that we have previously described [19]. Criteria to identify sites (a) differentially methylated in both early and established RA relative to healthy individuals, and (b) showing altered methylation in established but not early RA, are described in the Results and **Figure 1** (**steps 2 and 3**).

Patterns of differential methylation within and between groups were examined by hierarchical clustering using Genesis software (v1.7.6) with Euclidian distance and complete linkage criteria [40]. Further statistical analyses were performed using Stata 8.0 (Intercooled; Stata Corporation, TX, USA). P-values < 0.05 were considered significant.

Results

Genome-wide DNA methylation is altered in treatment-naïve early RA

We investigated genome-wide DNA methylation in T- and B-lymphocytes from a cohort of RA patients recruited at diagnosis and who were naïve for treatment with disease-modifying anti-rheumatic therapies. On the basis of a difference in mean β-value ≥ 0.1, and as described in **Step 1 of Figure 1**, we identified 1,951 CpGs in T-lymphocytes and 2,238 CpGs in B-lymphocytes that showed methylation changes in early RA patients relative to healthy individuals. These sites, the majority of which (60%) were hypermethylated in each of the cell types, represented 1,216 and 1,362 unique genes in T- and B-lymphocytes, respectively. Summary characteristics for these sites are presented in **Supplementary Table 2** (a complete list of CpGs for each cell type are available in **Supplementary Tables 3 and 4**).

Further analysis across the differentially methylated sites revealed a subset of 150 CpGs in T-lymphocytes and 113 CpGs in B-lymphocytes that showed more substantial methylation differences (≥ 0.2) in early RA patients relative to their healthy counterparts (**step 2, Figure 1**). A summary of these sites is provided in **Table 2** and annotated lists of the sites are provided in **Supplementary Tables 5 and 6**. Hierarchical clustering methodology was used to determine whether methylation patterns across these sites could distinguish between patients with early RA and healthy individuals. As shown graphically by the heatmaps and dendrograms in **Figure 2**, analysis revealed that differentially methylated sites in T- and B-lymphocytes segregate without exception early RA patients from healthy individuals. Indeed, there were 8 CpGs in T-lymphocytes and 12 CpGs in B-lymphocytes that showed methylation changes in all 23

early RA patients, and over 40 sites in each cell type were differentially methylated in 80% or more of the patients (**Supplementary Figure 1**). Of particular note in this analysis, of the 150 and 113 CpGs identified in T- and B-lymphocytes respectively, 44 sites shared identify between the cell types; moreover, for each of these sites, the direction of change in methylation was the same in each of the cell types.

Array data for selected CpG sites were validated by bisulfite pyrosequencing analysis. In each case, we observed concordance between array β-values and methylation values determined by this technique, reinforcing our conclusions of differential methylation in treatment-naïve early RA patients relative to healthy controls. **Figure 3** shows examples of two sites in T-lymphocytes (*DUSP22* and *KCNJ5*) and one site in B-lymphocytes (*ADAMTS17*). In addition, and as previously reported by us [19, 33], a significant correlation between the two methods was apparent across the spectrum of methylation β-values (**Figure 3A**).

Discrete methylation changes in early RA are preserved in patients with established disease

We next examined the methylation status of the 150 and 113 CpGs identified in T- and B-lymphocytes from early RA patients in a group of 12 patients with established disease that we have previously described [19]. For each of these patients, disease was of at least 5-years duration and in some cases extended to 30 years. In that study, and using the same array platform, we observed significant heterogeneity in methylation β-values across the patient samples, which we interpret may reflect the impact of disease progression or treatment modalities in these patients. We therefore assessed methylation

at the CpGs identified in early disease in each of the 12 established RA patients on an individual basis. Employing the criteria shown in **Step 3 of Figure 1**, we identified 50 and 26 CpGs in T- and B- lymphocytes, respectively, that showed methylation changes in both established and early disease (**blue boxes, Figure 1**). Examples of gene-associated CpGs identified in T-lymphocytes (*DUSP22*, *KCNJ5* and *MAGI2*) and B-lymphocytes (*MAML3*) are shown in **Figure 4** (a complete list and annotation for these sites is provided in **Supplementary Tables 7 and 8**).

Discrete methylation changes in established, but not early, disease

Finally, we determined whether methylation changes specifically associated with disease progression were also apparent. In these cases, we reasoned that these changes would be present in established disease but not in early RA patients. To this end, and as described in **Step 1 of Figure 1**, we first identified those sites for which there we no statistically significant differences in methylation between treatment-naïve early RA patients and healthy individuals. We subsequently identified CpGs differentially methylated in established disease as those for which six or more patients displayed altered methylation relative to, in these cases, both the early RA and healthy control groups (**step 3, Figure 1**). Using this approach, we identified 218 CpGs in T-lymphocytes and 208 CpGs in B-lymphocytes that were aberrantly methylated and were unique to patients with established disease (**green boxes, Figure 1**; details for these sites are provided in **Supplementary Tables 9 and 10**). Examples of the genes identified in T-lymphocytes (*BANF2*) and B-lymphocytes (*PSORS1C3*, *PCDH12* and *JRK*) are presented in **Figure 5**. Validation of selected genes by pyrosequencing confirmed the methylation changes observed by array in all cases (data not shown).

Discussion

The pathogenic mechanisms that drive disease in RA are not completely understood but an expanding literature implicates an important role for aberrations to the epigenome, including those which impact on the DNA methylation [4-7]. However, and to our knowledge, studies focusing on DNA methylation in RA, including those from our own laboratory, have been restricted to patients with established disease and, by extension, subsequent to long-standing therapeutic intervention. These potential confounders, and others consequent to disease progression, have thus far precluded identification of aberrantly methylated loci that play an important role in the etiology and genesis of RA, as has recently been discussed in RA [30, 41] and related conditions [42, 43].

In an attempt to address these issues, we determined genome-wide DNA methylation profiles in a cohort of newly diagnosed treatment-naïve patients with RA. Using stringent analytical and statistical criteria, we have identified novel disease-associated methylation changes in enriched T- and B-lymphocyte populations from these patients. These changes to the epigenome in treatment-naive early disease (relative to healthy controls) target multiple genes, suggesting an important role for specific epigenetic modification in the genesis and etiology of RA. Through comparison of these new data in early RA with data from our earlier, similar work in patients with established disease [19], we further report methylation changes present in early RA that are preserved in established disease. Similarly, we have also determined additional changes to the methylome that appear to be confined to or specific to patients with established disease.

The initial finding of our work was the identification of 1,951 CpGs in T-lymphocytes and 2,238 CpGs in B-lymphocytes, representing 1,216 and 1,362 genes respectively, that were differentially methylated in patients with early RA relative to healthy controls. The overall number of CpGs we identified are similar to those reported in studies of genome-wide methylation in RA patients with established disease [25, 26]. Applying more stringent criteria to the dataset (specifically β-value differences ≥0.2), we identified 150 and 113 sites in T- and B-lymphocytes respectively that discriminated between and segregated all of the early RA patients from healthy controls. To our knowledge, we are the first to define discrete methylation signatures in RA using distinct T- and B-lymphocyte populations; interestingly, methylation patterns in T-lymphocytes have previously been shown to segregate patients from controls in other autoimmune conditions [44].

Importantly, we found that the methylation changes defining each signature in T- and B-lymphocytes were present in the majority of patients studied. Indeed, there were more than 30 sites in each cell type that showed aberrant methylation in 20 or more of the 23 patients; for some of these sites, differential methylation was apparent in all 23 of the patients investigated. Furthermore, for each individual patient, 70 or more of the 150 sites identified in T-lymphocytes, and a similar number of the 113 sites identified in B-lymphocytes, were inappropriately methylated. The identification of methylation changes at these sites, and in particular those that are present in all or nearly all patients, supports a potential role for these as “drivers” of disease, rather than as “passenger” methylation events, a concept that has been described in cancer epigenetics [45, 46]. We also identified in early RA a further subset of sites (n = 44) that were differentially

methylated in both the T- and B-lymphocyte populations. Moreover, for each of these sites the direction of the methylation change, hypermethylation or hypomethylation, was identical in each cell type. These sites therefore appear to reflect lymphocyte-specific, rather than cell subset-specific, methylation changes that are associated with RA.

Comparative analyses of our findings in early disease with methylation levels determined in patients with established RA led to the identification of a significant number of sites where methylation changes in treatment-naïve early RA were also apparent in established disease. Aberrant methylation at these sites was apparent in at least six of the 12 established RA patients. These findings in both early and established disease suggest that epigenetic dysregulation at these sites/genes may be important in the establishment and maintenance of the disease phenotype. Conversely, we also identified a subset of sites for which differential methylation was apparent in patients with established RA but not in those newly diagnosed with RA. These changes, again present in at least six of the 12 established RA patients, might reflect changes important in disease progression, or perhaps be consequent to disease progression, chronic inflammation [47] and responses to individual treatment modalities. They may also reflect, in part, the influence of inherent heterogeneity between patients, and may thus represent the "passenger" methylation changes discussed earlier.

Examination of the differentially methylated sites using gene ontology and pathway analysis did not identify any specific ontology/pathway terms that were significantly enriched in early RA (following adjustment for multiple testing; data not shown). Nevertheless, amongst the sites aberrantly methylated in early RA, and in this case also

in patients with established disease, we noted multiple sites within the gene for DUSP22, a protein tyrosine phosphatase for which aberrant methylation has also been described in autoimmune primary Sjögren's syndrome [42]. Sites within this gene were hypermethylated in 21 of the 23 early RA patients and in 10 of 12 patients with established disease. Amongst other actions, DUSP22 has been reported to negatively regulate STAT3 and the IL-6/STAT3 signaling pathway [48] that mediates the actions of IL-6, a key cytokine in RA [49]. A recent study has also demonstrated enhanced T-lymphocyte-mediated immune responses and spontaneous development of autoimmunity and inflammation in a DUSP22 knockout mouse model, suggesting an important suppressive role in the regulation of T-lymphocyte-mediated immunity and autoimmunity [50]. Together, these findings support a potential role for epigenetic-mediated dysregulation of DUSP22 that may be important in the development and progression of RA.

The potential impact on the DNA methylome of treatment, and as a potential confounder, has been recognized by us [19, 27] and other investigators [10, 28, 30]. Methotrexate, the first-line anti-rheumatic drug in RA [51], has been reported to be associated with global methylation levels in inflammatory arthritis [10], and discordance for treatment with methotrexate has been shown to significantly impact on identification of differentially methylated sites in juvenile idiopathic arthritis [28]. More recently, methotrexate has been shown to induce demethylation of the Foxp3 locus in regulatory T-lymphocytes [29], cells essential for immunological self-tolerance [52] and the development of which is known to be epigenetically controlled [53]. Given these findings, we considered it important to focus our investigation on a cohort of patients

who were naïve for treatment with disease-modifying anti-rheumatic drugs, including methotrexate. In doing so, we are able to specifically exclude the likely impact of these agents from our findings and the identification of candidate CpGs/genes in this study.

The patients studied in this work were recruited consecutively from the early synovitis clinic at the hospital, with a subsequent diagnosis of RA in all cases. Through this approach we were able to study DNA methylation at first diagnosis and prior to the start of treatment. A potential caveat associated with our findings is the impact of confounding. However, and in this context, we were able to mitigate the influence of race and age, previously associated with DNA methylation [54, 55], by studying Caucasian subjects in whom the mean ages of early RA patients and healthy individuals were closely matched. To address the inclusion of both male and female patients, we removed prior to analysis all CpGs located on the X and Y chromosomes, as we and others have previously described [36, 56]. Furthermore, clustering analysis did not define separate clusters within the early RA group contingent on gender (data not shown). An important next step will be for the methylation signatures reported here to be replicated and verified by us and other groups in independent cohorts of newly diagnosed treatment-naïve patients.

Conclusions

Our findings reveal, for the first time, genome-wide aberrant DNA methylation in treatment-naïve patients newly diagnosed with RA. These changes in methylation, identified using a stringent genome-wide array approach and independent of treatment, are present in individual T- and B-lymphocyte populations and define unique disease-associated methylation signatures. Altered methylation for a subset of the sites is also preserved in patients with established disease. These data support an important role for aberrant methylation within discrete lymphocyte populations in the development and pathogenesis of RA. This work provides an incentive for further investigations in patients with early RA, most notably, to explore the potential of these epigenetic profiles to be used to predict treatment response and inform the clinical management of this condition.

Executive Summary

Background

- Compelling evidence from studies of multiple cell types describes changes to the DNA methylome in patients with rheumatoid arthritis (RA).
- Studies to date have examined patients with long-term established disease, which has precluded the identification of aberrant methylation signals associated with the etiology of RA.
- This study examined genome-wide DNA methylation in a cohort of treatment-naive patients with newly diagnosed early RA.

Patients & Methods

- Genome-wide DNA methylation profiling using HumanMethylation450 BeadChips was performed in individual T- and B-lymphocyte populations from 23 treatment-naive patients with early RA and 11 healthy controls.
- Array data were processed and differentially methylated sites identified using Minfi. Sites identified in early RA were further examined in patients with established disease (n = 12). Validation was conducted by bisulfite pyrosequencing.

Results

- Differential methylation in early RA was identified at 1,951 CpGs in T-lymphocytes and 2,238 CpGs in B-lymphocytes, representing 1,216 and 1,362 genes respectively.
- Sites displaying more substantial changes in methylation (≥ 0.2 difference) perfectly segregated early RA patients and healthy individuals into two discrete groups.
- A subset of sites differentially methylated in early RA showed similar methylation changes in established disease also. Conversely, sites were also identified for which methylation was altered in established, but not early, RA.

Conclusions

- Our findings for the first time reveal genome-wide DNA methylation changes and a unique methylation signature in individual T- and B-lymphocyte populations that are present at first diagnosis in patients with RA.
- Aberrant methylation of the candidates identified is independent of confounding from disease-modifying treatments in this cohort of treatment-naïve patients.
- These data support an important role for aberrant methylation within discrete lymphocyte populations in the development of RA. Further work is needed to confirm our findings and explore DNA methylation in the context of treatment response and the clinical management of RA.

References

Articles of interest have been highlighted as:

* of interest

1. Symmons D, Turner G, Webb R, *et al*. The prevalence of rheumatoid arthritis in the United Kigndom: new estimates for a new century. *Rheumatology* 41(7), 793–800 (2002).
2. Myasoedova E, Crowson CS, Maradil Kremers H, Therneau TM, Gabriel SE. Is the incidence of rheumatoid arthritis rising? Results from Olmsted County, Minnesota, 1955–2007. *Arthritis Rheum.* 62(6), 1576-1582 (2010).
3. Wiles N, Symmons DPM, Harrison B, *et al*. Estimating the incidence of rheumatoid arthritis: trying to hit a moving target? *Arthritis Rheum.* 42(7), 1339–1346 (1999).
4. Bottini S, Firestein GS. Epigenetics in Rheumatoid Arthritis: A Primer for Rheumatologists. *Curr. Rheumatol. Rep.* 15(11), 372 (2013).
5. Glant TT, Mikecz K, Rauch TA. Epigenetics in the pathogenesis of rheumatoid arthritis. *BMC Med.* 12, 35 (2014).
6. Klein K, Gay S. Epigenetics in rheumatoid arthritis. *Curr. Opin. Rheumatol.* 27(1), 76–82 (2015).
7. Bird A. DNA methylation patterns and epigenetic memory. *Genes Dev.* 16(1), 6–21 (2002).
8. Richardson B, Scheinbart L, Strahler J, Gross L, Hanash S, Johnson M. Evidence for impaired T cell DNA methylation in systemic lupus erythematosus and rheumatoid arthritis. *Arthritis Rheum.* 33(11), 1665–1673 (1990).
9. Corvetta A, Della Bitta R, Luchetti MM, Pomponio G. 5-Methylcytosine content of DNA in blood, synovial mononuclear cells and synovial tissue from patients

affected by autoimmune rheumatic diseases. *J. Chromatogr.* 566(2), 481–491 (1991).

10. Kim YI, Logan JW, Mason JB, Roubenoff R. DNA hypomethylation in inflammatory arthritis: reversal with methotrexate. *J. Lab. Clin. Med.* 128(2), 165–172 (1996).
11. Liu CC, Fang TJ, Ou TT, *et al*. Global DNA methylation, DNMT1, and MBD2 in patients with rheumatoid arthritis. *Immunol. Lett.* 135(1-2), 96–99 (2011).
12. Fu LH, Cong B, Zhen YF, *et al*. [Methylation status of the IL-10 gene promoter in the peripheral blood mononuclear cells of rheumatoid arthritis patients]. *Yi Chuan* 29(11), 1357–1361 (2007).
13. 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.* 58(9), 2686–2693 (2008).
14. Kitamura T, Kabuyama Y, Kamataki A, *et al*. Enhancement of lymphocyte migration and cytokine production by ephrinB1 system in rheumatoid arthritis. *Am. J. Physiol. Cell. Physiol.* 294(1), C189-C196 (2008).
15. 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* 32(11), 1373–1380 (2011).
16. Ishida K, Kobayashi T, Ito S, *et al*. Interleukin-6 gene promoter methylation in rheumatoid arthritis and chronic periodontitis. *J. Periodontol.* 83(7), 917–925 (2012).
17. Liao J, Liang G, Xie S, *et al*. CD40L demethylation in CD4(+) T cells from women with rheumatoid arthritis. *Clin. Immunol.* 145(1), 13–18 (2012).

18. Cribbs AP, Kennedy A, Penn H, *et al*. Treg cell function in rheumatoid arthritis is compromised by ctla-4 promoter methylation resulting in a failure to activate the indoleamine 2,3-dioxygenase pathway. *Arthritis Rheumatol.* 66(9), 2344–2354 (2014).

19. Glossop JR, Emes RD, Nixon NB, *et al*. Genome-wide DNA methylation profiling in rheumatoid arthritis identifies disease-associated methylation changes that are distinct to individual T- and B-lymphocyte populations. *Epigenetics* 9(9), 1228–1237 (2014).

***First genome-wide study of DNA methylation in distinct lymphocyte populations in patients with RA.**

20. Karouzakis E, Gay RE, Michel BA, Gay S, Neidhart M. DNA hypomethylation in rheumatoid arthritis synovial fibroblasts. *Arthritis Rheum.* 60(12), 3613–3622 (2009).

21. Takami N, Osawa K, Miura Y, *et al*. Hypermethylated promoter region of DR3, the death receptor 3 gene, in rheumatoid arthritis synovial cells. *Arthritis Rheum.* 54(3), 779–787 (2006).

22. Zhou Q, Long L, Shi G, Zhang J, Wu T, Zhou B. Research of the methylation status of miR-124a gene promoter among rheumatoid arthritis patients. *Clin. Dev. Immunol.* 2013, 524204 (2013).

23. Karouzakis E, Rengel Y, Jungel A, *et al*. DNA methylation regulates the expression of CXCL12 in rheumatoid arthritis synovial fibroblasts. *Genes Immun.* 12(8), 643–652 (2011).

24. Park SH, Kim SK, Choe JY, *et al*. Hypermethylation of EBF3 and IRX1 genes in synovial fibroblasts of patients with rheumatoid arthritis. *Mol. Cells* 35(4), 298–304 (2013).

25. Nakano K, Whitaker JW, Boyle DL, Wang W, Firestein GS. DNA methylome signature in rheumatoid arthritis. *Ann. Rheum. Dis.* 72(1), 110–117 (2013).

26. de la Rica L, Urquiza JM, Gomez-Cabrero D, *et al*. Identification of novel markers in rheumatoid arthritis through the integrated analysis of DNA methylation and microRNA expression. *J. Autoimmun.* 41, 6–16 (2013).

27. Glossop JR, Haworth KE, Emes RD, *et al*. DNA methylation profiling of synovial fluid FLS in rheumatoid arthritis reveals changes common with tissue-derived FLS. *Epigenomics* 2015. In Press.

28. Ellis JA, Munro JE, Chavez RA, *et al*. Genome-scale case-control analysis of CD4+ T-cell DNA methylation in juvenile idiopathic arthritis reveals potential targets involved in disease. *Clin. Epigenetics* 4(1), 20 (2012).

29. Cribbs AP, Kennedy A, Penn H, *et al*. Methotrexate restores regulatory T cell function through demethylation of the foxp3 upstream enhancer in patients with rheumatoid arthritis. *Arthritis Rheumatol.* 2015. Jan 20. doi: 10.1002/art.39031. [Epub ahead of print].

***Study proposing demethylation of Foxp3 as a novel mechanism of action for methotrexate and restoration of regulatory T-cell function in RA.**

30. Plant D, Wilson AG, Barton A. Genetic and epigenetic predictors of responsiveness to treatment in RA. *Nat. Rev. Rheumatol.* 10(6), 329–337 (2014).

***Timely and well considered review article discussing epigenetic biomarkers for the prediction of treatment response in RA.**

31. Aletaha D, Neogi L, Silman AJ, *et al*. 2010 Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Arthritis Rheum.* 62(9), 2569–2581 (2010).
32. Prevoo ML, van 't Hof MA, Kuper HH, van Leeuwen MA, van de Putte LB, van Riel PL. Modified disease activity scores that include twenty-eight-joint counts. Development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. *Arthritis Rheum.* 38(1), 44–48 (1995).
33. Glossop JR, Nixon NB, Emes RD, *et al*. Epigenome-wide profiling identifies significant differences in DNA methylation between matched-pairs of T- and B-lymphocytes from healthy individuals. *Epigenetics* 8(11), 1188–1197 (2013).
34. Bibikova M, Barnes B, Tsan C, *et al*. High density DNA methylation array with single CpG resolution. *Genomics* 98(4), 288–295 (2011).
35. Aryee MJ, Jaffe AE, Corrada-Bravo H, *et al*. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics* 30(10), 1363–1369 (2014).
36. Maksimovic J, Gordon L, Oshlack A. SWAN: Subset-quantile within array normalization for illumina infinium HumanMethylation450 BeadChips. *Genome Biol.* 13(6), R44 (2012).
37. Dedeurwaerder S, Defrance M, Calonne E, Denis H, Sotiriou C, Fuks F. Evaluation of the Infinium Methylation 450K technology. *Epigenomics* 3(6), 771–784 (2011).
38. Du P, Zhang X, Huang CC, *et al*. Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. *BMC Bioinformatics* 11, 587 (2010).

39. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Statist. Soc. B* 57, 289–300 (1995).

40. Sturn A, Quackenbush J, Trajanoski Z. Genesis: cluster analysis of microarray data. *Bioinformatics* 18(1), 207–208 (2002).

41. Liu Y, Aryee MJ, Padyukov L, *et al*. Epigenome-wide association data implicate DNA methylation as an intermediary of genetic risk in rheumatoid arthritis. *Nat. Biotechnol.* 31(2), 142–147 (2013).

42. Altorok N, Coit P, Hughes T, *et al*. Genome-wide DNA methylation patterns in naïve CD4+ T cells from patients with primary Sjögren's syndrome. *Arthritis Rheumatol.* 66(3), 731–739 (2014).

43. Delgado-Calle J, Fernández AF, Sainz J, *et al*. Genome-wide profiling of bone reveals differentially methylated regions in osteoporosis and osteoarthritis. *Arthritis Rheum.* 65(1), 197–205 (2013).

44. Jeffries MA, Dozmorov M, Tang Y, Merrill JT, Wren JD, Sawalha AH. Genome-wide DNA methylation patterns in CD4+ T cells from patients with systemic lupus erythematosus. *Epigenetics* 6(5), 593–601 (2011).

45. De Carvalho DD, Sharma S, You JS, *et al*. DNA methylation screening identifies driver epigenetic events of cancer cell survival. *Cancer Cell* 21(5), 655–667 (2012).

46. Kalari S, Pfeifer GP. Identification of driver and passenger DNA methylation in cancer by epigenomic analysis. *Adv. Genet.* 70, 277–308 (2010).

47. Nakano K, Boyle DL, Firestein GS. Regulation of DNA methylation in rheumatoid arthritis synoviocytes. *J. Immunol.* 190(3), 1297–1303 (2013).

48. Sekine Y, Tsuji S, Ikeda O, *et al*. Regulation of STAT3-mediated signaling by LMW-DSP2. *Oncogene* 25(42), 5801–5806 (2006)

49. Fonseca JE, Santos MJ, Canhão H, Choy E. Interleukin-6 as a key player in systemic inflammation and joint destruction. *Autoimmun. Rev.* 8(7), 538–542 (2009).
50. Li JP, Yang CY, Chuang HC, *et al*. The phosphatase JKAP/DUPS22 inhibits T-cell receptor signalling and autoimmunity by inactivating Lck. *Nat. Commun.* 5, 3618 (2014).

    ***Knockout mouse model demonstrating the importance of DUSP22 in the suppression of T-cell-driven autoimmunity and inflammation.**
51. Cronstein BN. Low-Dose Methotrexate: A Mainstay in the Treatment of Rheumatoid Arthritis. *Pharmacol. Rev.* 57(2), 163–172 (2005).
52. Vignali DA, Collison LW, Workman CJ. How regulatory T cells work. *Nat. Rev. Immunol.* 8(7), 523–532 (2008).
53. Polansky JK, Kretschmer K, Freyer J, *et al*. DNA methylation controls Foxp3 gene expression. *Eur. J. Immunol.* 38(6), 1654–1663 (2008).
54. Zhang FF, Cardarelli R, Carroll J, *et al*. Significant differences in global genomic DNA methylation by gender and race/ethnicity in peripheral blood. *Epigenetics* 6(5), 623–629 (2011).
55. Boks MP, Derks EM, Weisenberger DJ, *et al*. The relationship of DNA methylation with age, gender and genotype in twins and healthy controls. *PLoS One* 4(8), e6767 (2009).
56. Duong CV, Emes RD, Wessely F, Yacqub-Usman K, Clayton RN, Farrell WE. Quantitative, genome-wide analysis of the DNA methylome in sporadic pituitary adenomas. *Endocr. Relat.Cancer* 19(6), 805–816 (2012).

Financial disclosure/Acknowledgements

The authors would like to Dr. Dimitra Blana and Dr. Ed Chadwick for analytical assistance and Janet Turner for assistance with the collection of clinical data. This work was supported by funding generously provided by the Haywood Rheumatism Research and Development Foundation. The authors have no conflict of interest to disclose and have no relevant affiliations with or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. The authors did not receive any writing assistance in the preparation of the manuscript

Figure legends

Figure 1. Filtering criteria for identification of CpGs differentially methylated in treatment-naïve early RA patients relative to healthy individuals. The starting number of CpGs indicated for T-lymphocytes and B-lymphocytes (470,366 and 473,105, respectively) are the sites remaining after processing and normalization, as described in the Methods. Numbers in the figure indicate the number of CpGs remaining at each successive step, and where: 1) red boxes indicates sites differentially methylated in early RA relative to healthy individuals; 2) blue boxes indicate sites differentially methylated in early RA and also in established RA (relative to healthy); and 3) green boxes indicate sites differentially methylated in established RA but not in early RA (relative to healthy).

Figure 2. Heatmap and clustering for the 150 and 113 differentially methylated CpGs in T- and B-lymphocytes, respectively, in treatment-naïve early RA patients. For T-lymphocytes (A) and B-lymphocytes (B), each row represents an individual CpG and each column an individual sample (listed beneath the heatmap). Color gradation from yellow to blue represents low to high DNA methylation respectively, with β-values ranging from 0 (no methylation; yellow) to 1 (complete methylation; blue). In both cell types, methylation patterns across these sites segregated RA patients (red bar) and healthy individuals (green bar) into two discrete groups.

Figure 3. Validation of differentially methylated CpGs in T- and B-lymphocytes in treatment-naive early RA patients. A. Scatter plot demonstrating excellent correlation between array methylation β-values and methylation as determined by bisulfite

pyrosequencing for T- and B-lymphocytes from treatment-naive early RA patients. One CpG in each of three separate genes was analyzed (*HMOX2*, cg14951292; *AMN*, cg09616556; and *PM20D1*, cg11965913). Spearman's r = 0.964, p < 0.00001. (Data for patients with established RA and healthy individuals was reported previously [16]). **B–D.** Validation using bisulfite pyrosequencing of two hypermethylated candidates in T-lymphocytes, *DUSP22* (**B**) and *KCNJ5* (**C**), and one hypomethylated candidate in B-lymphocytes, *ADAMTS17* (**D**), showing differential methylation in early RA relative to healthy individuals. In each plot, early RA patients (n = 23) and healthy individuals (n = 11) are depicted by circles and triangles respectively, and the short blue horizontal bar in each group indicates the mean value. Gene names are shown on the x-axis and where, for each gene, methylation values are shown for the array (filled symbols; left) and pyrosequencing (open symbols; right).

Figure 4. Identification of differentially methylated sites in early RA that also show altered methylation in patients with established disease. Shown are examples of three candidates in T-lymphocytes, *DUSP22* (**A**), *KCNJ5* (**B**) and *MAGI2* (**C**), and one candidate in B-lymphocytes, *MAML3* (**D**), that were identified by array analysis as differentially methylated in early RA, and which also showed similar methylation changes in patients with established disease (criteria for differential methylation in each RA group relative to healthy individuals are described in the main text). In each plot, RA patients (early and established) and healthy individuals are depicted by circles and triangles respectively. Filled circles in the early and established RA groups represent individual samples showing a methylation level at least 20% higher/lower than the mean of the healthy group for hypermethylated and hypomethylated sites, respectively

(indicated by the horizontal dashed line). The short blue horizontal bar in each group indicates the mean value. CpGs are presented by their associated gene in each case (shown on the x-axis).

Figure 5. Identification of sites differentially methylated in established RA that are not differentially methylated in patients with early RA. Shown are examples of one candidate in T-lymphocytes, *BANF2* (**A**), and three candidates in B-lymphocytes, *PSORS1C3* (**B**), *PCDH12* (**C**), and *JRK* (**D**), that displayed similar levels of methylation in early RA patients and healthy individuals but which were identified as differentially methylated in patients with established disease (criteria for differential methylation in established RA are described in the main text). In each plot, RA patients (early and established) and healthy individuals are depicted by circles and triangles respectively. Filled circles in the established RA group represent individual samples showing a methylation level at least 20% higher or lower than the mean of the healthy group for hypermethylated and hypomethylated sites, respectively (indicated by the horizontal dashed line). The short blue horizontal bar in each group indicates the mean value. CpGs are presented by their associated gene in each case (shown on the x-axis). For each gene, a minimum of six established RA patients showed a methylation difference ≥ 0.2 relative to the healthy and early RA groups. The differential methylation observed was also reflected by a mean difference ≥ 0.2 in the established RA group relative to the means of the healthy and early RA groups.

Supplementary Figure 1. Heatmap of methylation differences in individual early RA patients, relative to the healthy control group mean, for each of the 150 and

113 CpGs differentially methylated in T- and B-lymphocytes, respectively. For T-lymphocytes (A) and B-lymphocytes (B), each row represents an individual CpG and each column an individual treatment-naïve early RA patient (listed above the heatmap). For each patient, sites hypomethylated and hypermethylated relative to the mean of the healthy group are depicted by blue and red coloring, respectively (grey indicates sites not differentially methylated relative to healthy controls). The black horizontal line separates sites hypomethylated (above) and hypermethylated (below) in early RA. CpGs are ordered according to the total number of individual patients showing a methylation difference relative to the healthy mean (decreasing from top to bottom for hypomethylated sites, and increasing from top to bottom for hypermethylated sites).

Supplementary Table 1. Assay details for candidate CpGs/genes interrogated by bisulfite pyrosequencing.

*Further information that is not included here is available upon request. The horizontal dashed line between rows 7 and 8 separates candidate genes in T-lymphocytes (above) and B-lymphocytes (below).

†The prefix 'b-' denotes biotin labeling at the 5' end.

‡The sequence provided is post-bisulfite conversion. Letters 'Y' and 'R' denote the cytosine of the CpG site interrogated by the assay ('Y' and 'R' refer to sequencing along the upper and lower strands respectively).

Abbreviations: bp, base pairs

Supplementary Table 2. Summary characteristics for CpGs identified as differentially methylated in T-lymphocytes and B-lymphocytes in treatment-naïve early RA patients.*

Unless otherwise indicated, all figures are the number (%).

* Differentially methylated CpGs were identified according to the criteria described in the Methods/Results section and in Figure 1 (sites with a mean difference in β-value ≥ 0.1 that was statistically significant).

† Annotation describing CpG island, shore and shelf status, and genetic location relative to defined gene regions, is derived from the University of California at Santa Cruz (UCSC) database.

Abbreviations: TSS200, 200-bp block upstream of the transcription start site; TSS1500, 1500-bp block upstream of the transcription start site; 5' UTR, 5' untranslated region; 3' UTR, 3' untranslated region.

Supplementary Table 3. Complete list, β-value data and annotation for the 1,951 CpGs identified in T-lymphocytes as differentially methylated between treatment-naïve early RA patients and healthy individuals.*

*Unless otherwise stated, values are array methylation β-values. The dashed horizontal line between rows 1,239 and 1,240 separates CpGs that were hypermethylated (above) and hypomethylated (below) in T-lymphocytes from treatment-naive early RA patients.

†The 'dmpFinder' function in Minfi [35] was used to calculate F-test p-values adjusted for multiple testing using the false discovery rate (FDR) [39].

Supplementary Table 4. Complete list, β-value data and annotation for the 2,238 CpGs identified in B-lymphocytes as differentially methylated between treatment-naïve early RA patients and healthy individuals.*

*Unless otherwise stated, values are array methylation β-values. The dashed horizontal line between rows 1,295 and 1,296 separates CpGs that were hypermethylated (above) and hypomethylated (below) in B-lymphocytes from treatment-naive early RA patients.

†The 'dmpFinder' function in Minfi [35] was used to calculate F-test p-values adjusted for multiple testing using the false discovery rate (FDR) [39].

Supplementary Table 5. Complete list, β-value data and annotation for the 150 CpGs identified in T-lymphocytes as differentially methylated (≥ 0.2 mean difference) between treatment-naïve early RA patients and healthy individuals.*

*Unless otherwise stated, values are array methylation β-values. The dashed horizontal line between rows 92 and 93 separates CpGs that were hypermethylated (above) and hypomethylated (below) in T-lymphocytes from treatment-naive early RA patients.

†The 'dmpFinder' function in Minfi [35] was used to calculate F-test p-values adjusted for multiple testing using the false discovery rate (FDR) [39].

Supplementary Table 6. Complete list, β-value data and annotation for the 113 CpGs identified in B-lymphocytes as differentially methylated (≥ 0.2 mean difference) between treatment-naïve early RA patients and healthy individuals.*

*Unless otherwise stated, values are array methylation β-values. The dashed horizontal line between rows 95 and 96 separates CpGs that were hypermethylated (above) and hypomethylated (below) in B-lymphocytes from treatment-naive early RA patients.

†The 'dmpFinder' function in Minfi [35] was used to calculate F-test p-values adjusted for multiple testing using the false discovery rate (FDR) [39].

Supplementary Table 7. Complete list, β-value data and annotation for the 50 CpGs in T-lymphocytes that were differentially methylated in treatment-naïve early RA and which also were differentially methylated in patients with established RA.*

*Unless otherwise stated, values are array methylation β-values. The dashed horizontal line between rows 36 and 37 separates CpGs that were hypermethylated (above) and hypomethylated (below) in T-lymphocytes from both treatment-naive early RA patients and from patients with established disease.

Supplementary Table 8. Complete list, β-value data and annotation for the 26 CpGs in B-lymphocytes that were differentially methylated in treatment-naïve early RA and which also were differentially methylated in patients with established RA.*

*Unless otherwise stated, values are array methylation β-values. The dashed horizontal line between rows 25 and 26 separates CpGs that were hypermethylated (above) and hypomethylated (below) in B-lymphocytes from both treatment-naive early RA patients and from patients with established disease.

Supplementary Table 9. Complete list, β-value data and annotation for the 218 CpGs in T-lymphocytes that were not differentially methylated in treatment-naïve early RA but were differentially methylated in patients with established RA.*

*Unless otherwise stated, values are array methylation β-values. The dashed horizontal line between rows 121 and 122 separates CpGs that were hypermethylated (above) and hypomethylated (below) in T-lymphocytes from patients with established RA.

Supplementary Table 10. Complete list, β-value data and annotation for the 208 CpGs in B-lymphocytes that were not differentially methylated in treatment-naïve early RA but were differentially methylated in patients with established RA.*

*Unless otherwise stated, values are array methylation β-values. The dashed horizontal line between rows 103 and 104 separates CpGs that were hypermethylated (above) and hypomethylated (below) in B-lymphocytes from patients with established RA.