DNA methylation at diagnosis is associated with response to disease-modifying drugs in early rheumatoid arthritis

John R. Glossop[1,2], Nicola B. Nixon[2], Richard D. Emes[3,4], Julius Sim[5], Jon C. Packham[1,2], Derek L. Mattey[1,2], William E. Farrell[1], Anthony A. Fryer[1].

[1]Institute for Applied Clinical Sciences, 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.

[5]School of Health and Rehabilitation, Keele University, Staffordshire, ST5 5BG, UK.

Author for correspondence/re-prints: Professor Tony Fryer, Institute for Applied Clinical Sciences, Keele University, Guy Hilton Research Centre, Thornburrow Drive, Hartshill, Stoke-on-Trent, Staffordshire, ST4 7QB, UK.

Tel: +44 (0) 1782 674245

Fax: +44 (0) 1782 747319

E-mail: anthony.fryer@uhnms.nhs.uk

Key words: Disease-modifying antirheumatic drugs (DMARDs); disease activity score with 28 joint counts (DAS28); DNA methylation; Early rheumatoid arthritis; Illumina 450K array; T-lymphocyte; Treatment response.

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

Acknowledgements: The authors would like to thank the patients who participated in the study. We also thank Janet Turner, Cath Thwaites and Moira Dishman for assistance with the collection of clinical data.

Running footline: DNA methylation and treatment response in early RA

Abstract

Aims. A proof-of-concept study to explore whether DNA methylation at first diagnosis is associated with response to disease-modifying antirheumatic drugs (DMARDs) in patients with early rheumatoid arthritis (RA).

Patients & Methods. DNA methylation was quantified in T-lymphocytes from 46 treatment-naïve patients using HumanMethylation450 BeadChips. Treatment response was determined at six months using the EULAR response criteria.

Results. Initial filtering identified 21 CpGs that were differentially methylated between responders and non-responders. After conservative adjustment for multiple testing, six sites remained statistically significant, of which four showed high sensitivity and/or specificity (≥75%) for response to treatment. Moreover, methylation at two sites in combination was the strongest factor associated with response (80.0% sensitivity, 90.9% specificity, AUC 0.85).

Conclusions. DNA methylation at diagnosis is associated with DMARD treatment response in early RA.

Introduction

Rheumatoid arthritis (RA) is a chronic inflammatory joint disease of autoimmune origin that affects 0.5–1.0% of the adult population [1, 2]. Treatment of patients with centres on the use of a variety of synthetic disease-modifying antirheumatic drugs (DMARDs). Methotrexate is the first-line DMARD of choice for the treatment and management of RA, prescribed as monotherapy or in combination with other DMARDs. Although these agents are efficacious for the treatment of RA [3-5], clinically meaningful responses are not observed in all patients and a significant proportion remain refractory to treatment.

A substantial body of literature supports an important role for epigenetic dysregulation, including of DNA methylation, in the pathogenesis of RA [reviewed in 6-8]. Evidence also suggests that disease modifying agents such as methotrexate may influence DNA methylation [9, 10]. Moreover, methylation status as a potential biomarker associated with response to therapy has been demonstrated in other conditions [11] and proposed for use in RA by several investigators [12, 13]. DNA is methylated through enzymatic conversion of cytosine to methylcytosine; this occurring almost invariably at cytosine-phosphate-guanine sites (CpGs). In the context of promoter-associated sites, methylation is associated with transcriptional repression and gene silencing [14]. In RA, alterations to the DNA methylome are apparent in multiple cell types important in the disease process, including peripheral blood-derived mononuclear cells, lymphocytes and joint-derived fibroblasts. Recently, we were the first to define disease-associated methylation changes that were distinct to individual T- and B-lymphocyte populations [15]. Moreover, we reported methylation differences in these lymphocyte populations in treatment-naïve patients at first RA diagnosis [16]. Whilst providing evidence for a role

77 in the development of the disease, our findings support DNA methylation profiling at
78 diagnosis as a potential source of biomarkers for response to treatment in RA.
79
80 It is clear that the ability to identify which patients will respond to treatment offers
81 considerable benefits for the management of RA. For example, it would (i) facilitate
82 rapid dose-escalation and reduce time to effective response in those likely to be poor
83 responders to traditional regimens, and (ii) avoid unwanted side-effects in those likely
84 to show an effective response to lower doses or monotherapy. These benefits are all the
85 more important given evidence that response to first treatment with disease-modifying
86 agents is strongly associated with long-term outcome in these patients [17]. The search
87 for biomarkers associated with response has encompassed demographic and clinical
88 factors as well as genetic associations and expression profiling of proinflammatory and
89 other mediators [18-20]. However, no single factor or combination of factors have thus
90 far proven to be accurate and reliable in determining which patients will respond to
91 DMARD therapy.
92
93 Our aim therefore, in this proof-of-concept study, was to determine whether genome-
94 wide DNA methylation profiles at first diagnosis are associated with response to
95 treatment with conventional DMARDs (as determined by improvement in disease
96 activity using the validated European League Against Rheumatism (EULAR) response
97 criteria) in a typical population of newly-diagnosed, treatment-naïve patients with RA.
98 As in our previous work, we examined methylation in purified T-lymphocyte
99 populations, cells that are instrumental in the disease process and chronic inflammation

100 [21], and for which relationships with disease activity have recently been described [22-
101 24].

Patients and Methods

Study population

A prospective cohort of 46 Caucasian patients attending the early synovitis clinic at the Haywood Rheumatology Centre in Stoke-on-Trent, UK, and presenting with symptomatic inflammatory arthritis suspected to be RA was recruited. All patients were subsequently classified as having RA, according to the 2010 ACR/EULAR classification criteria, by a consultant rheumatologist [25]. No patients had been treated with DMARDs or biological agents at the time of recruitment. Clinical data collected at baseline included disease activity, erythrocyte sedimentation rate (ESR), rheumatoid factor (RF) and anti-citrullinated peptide antibodies (ACPA). Demographic and clinical characteristics are presented in **Table 1**. At diagnosis with RA, all patients began treatment with one or more DMARDs (methotrexate, hydroxychloroquine, and sulphasalazine) and the majority received parenteral corticosteroids, solely for the clinical management of RA and as directed by a consultant rheumatologist. Patients were followed for six-months and remained on treatment throughout. The study was approved by the East Midlands (Derby) Research Ethics Committee. All patients provided written informed consent.

Disease activity was determined at recruitment (prior to initiation of DMARD therapy) and after three and six months of treatment using the disease activity score with 28-joint counts (DAS28) [26], though data at three months was excluded from further analysis due to the known short-term effect of corticosteroid treatment on DAS28 scores. DAS28 scores range from 0-10: a score >5.1 indicates high disease activity while one of ≤3.2 denotes low disease activity. Response to treatment was determined at six months

according to the DAS28-based EULAR response criteria [26-28], which evaluate response in patients with RA based on a composite categorization incorporating both change in DAS28 from baseline (ΔDAS28) and final absolute DAS28 score. Specifically, these criteria classify response as 'good' (ΔDAS28 >1.2, current DAS28 ≤3.2), 'moderate' (ΔDAS28 >1.2, current DAS28 >3.2, or ΔDAS28 >0.6–1.2, current DAS28 ≤5.1) and 'no' (ΔDAS28 ≤0.6, or ΔDAS28 >0.6–1.2, current DAS28 >5.1) [28]. According to these criteria, responders were defined as patients with a 'good' or 'moderate' response to treatment, and non-responders as patients with 'no' response to treatment.

Isolation of T-lymphocytes

Fresh peripheral blood samples (35 ml, EDTA) were collected from each patient at baseline, prior to the initiation of treatment. $CD3^+$ T-lymphocytes were isolated from mononuclear cell preparations using positive selection with magnetic microbeads (MACS® Separation System; Miltenyi Biotec). We have previously shown this method to yield high-purity T-lymphocyte populations (mean ≥ 99%) in RA patients [15]. Genomic DNA was extracted using an AllPrep DNA/RNA/miRNA Universal kit (Qiagen) and stored at -20°C prior to use.

Genome-wide DNA methylation profiling

DNA methylation was quantified at >480,000 CpG sites using the HumanMethylation450 BeadChip (Illumina Inc.; hereafter referred to as 'array'). Details of array design and coverage have been described elsewhere [29]. Genomic DNA samples (n = 46) were treated with sodium bisulfite using an EZ DNA

Methylation Kit (Zymo Research) and subsequently were hybridized to arrays according to manufacturer recommended protocols, as previously described (performed by Hologic Tepnel Pharma Services, Manchester, UK) [30]. All samples passed stringent internal array quality control, including sample-independent (*e.g.* staining, hybridization) and sample-dependent (*e.g.* bisulfite conversion) controls. Methylation at individual CpG sites is reported as a β-value ranging from 0 to 1 (unmethylated to fully methylated, respectively) [29].

Sodium bisulfite Pyrosequencing

Array candidates were independently validated by bisulfite Pyrosequencing using a PyroMark Q24 instrument and analysis software (Qiagen), as we have previously described [15, 30]. Briefly, fresh genomic DNA aliquots were sodium bisulfite-converted and amplified using whole genome amplification [30, 31]. Thereafter, Touchdown PCR [32, 33] was used to prepare PCR amplicons containing CpGs of interest. Assay details are provided in **Supplementary Table 1**.

Data analysis

Array data (idat files) were processed and analyzed using the Bioconductor package Minfi [34]. We removed from analysis all CpGs with a detection p-value >0.01 in any one or more of the 46 samples and all probes targeting sites on the X and Y chromosomes (a total of 12,295 CpGs). Data were normalized by Subset-quantile Within Array Normalization (SWAN), as described by Maksimovic *et al*. [35], and multi-dimensional scaling plots were examined to confirm appropriate adjustment for potential confounding due to batch effects (processing date, array position and slide).

To identify methylation differences associated with treatment response, patients were stratified into responders and non-responders. CpGs showing altered methylation between the two groups were identified using the 'dmpFinder' function in Minfi. This function performs an F-test to compare groups and was used with logit-transformed β-values (M-values), as recommended by Du *et al*. [36]. P-values <0.05 were considered statistically significant and, together with a mean β-value difference ≥0.1 between the groups, were used as an initial screening tool to identify sites displaying differential methylation. Two further filtering steps were subsequently applied to identify differentially methylated CpGs as those sites where: 1) at least two-thirds of non-responders showed a β-value difference ≥0.1 relative to the responder mean; and 2) at least two-thirds of responders displayed a β-value equal to or in excess of the responder mean. Filtering criteria are summarized in **Figure 1**. We then applied a Bonferroni adjustment at stage 5, based on comparisons conducted using the final 21 CpGs identified.

The McNemar test was used to examine the incidence of patients with moderate/high disease activity between baseline and six-months. The association of baseline methylation status with treatment response was determined by calculating sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV), and by examining receiver operating characteristic (ROC) area under the curve (AUC) plots. ROC curves were constructed based on logistic regression analysis with response to treatment categorised as no response versus moderate/good response as described

above. Analyses were performed using Stata 12.0 (Intercooled; Stata Corporation, TX, USA) and considering p-values <0.05 as statistically significant.

Results

Characteristics of the patients

Table 1 summarizes the demographic and clinical characteristics for the RA patients at recruitment. Most patients (43/46, 93.5%) started treatment with MTX, either as monotherapy or in combination with other DMARDs. The majority of patients (33/46, 71.7%) remained on their indicated starting DMARD regimen throughout the course of the study. Of the remaining patients, all but two introduced or discontinued a single DMARD on one occasion during the six-month follow-up period.

Disease activity and treatment response

Moderate or high disease activity (DAS28 >3.2) was present in 43/46 (93.5%) patients at recruitment (three patients had low disease activity, with DAS28 scores of 2.27, 2.66 and 3.18). After six-months of treatment, 28/46 (60.9%) patients had moderate/high disease activity ($p < 0.001$ vs. baseline, McNemar test), with approximately two-thirds (63.0%) achieving an improvement in DAS28 ≥1.2. Classifying response by the EULAR response criteria, the number of patients achieving a good, moderate and no response to treatment at six-months was 16 (34.8%), 19 (41.3%), and 11 (23.9%), respectively. On this basis, 76.1% (35/46) of patients were classified as responders and the remainder as non-responders. Details of baseline characteristics and six-month treatment regimens for the two groups are presented in **Supplementary Table 2**.

Relationship between DNA methylation and treatment response

Use of the robust filtering steps described in the Methods section and shown in **Figure 1** identified 269 CpGs with a statistically significant difference in mean methylation β-

value ≥0.1 between responders and non-responders. Moreover, for a subset of 21 sites, methylation differences were present in at least two-thirds of the individual patients within each group (full annotation for these 21 sites is provided in **Supplementary Table 3**). The majority of these sites were hypermethylated in responders (16/21, 76.2%), were linked with a gene (15/21, 71.4%) and were associated with a CpG island and/or the surrounding shores/shelves (13/21, 61.9%).

To refine these sites further, we applied a conservative Bonferroni adjustment for multiple testing, based on the 21 comparisons undertaken. This revealed six CpGs for which the methylation differences between responders and non-responders remained statistically significant (p_{adj} <0.05; **Supplementary Table 3**). For each of these six CpGs, we plotted methylation against treatment response to determine a percentage methylation cut-off that in each case provided the greatest discrimination between patients that responded to treatment and those that did not. Examples of two differentially methylated CpGs are presented in **Figure 2**. We also calculated the corresponding sensitivity and specificity for each site to assess the association of methylation status with response. Using this approach, and as shown in **Table 2**, four sites were identified with a sensitivity and/or specificity ≥75% for discrimination between responders and non-responders. Most notably, hypermethylation of CpG-2 and hypomethylation of CpG-3 (shown in **Figure 2** and validated by Pyrosequencing in **Supplementary Figure 1**) each demonstrated a sensitivity and PPV of approximately 90%, although the corresponding specificity and NPV were lower (63.6% and 70.0%, and 63.6% and 63.6%, for CpG-2 and CpG-3, respectively). Using ROC curve analysis

to further evaluate the association with response, CpG-2 and CpG-3 also demonstrated the highest AUC values (0.78 and 0.76, respectively).

Combinations of CpGs associated with treatment response

Focusing on the four sites identified above, we next examined the ability to discriminate between responders and non-responders for each of the six possible pairs of sites. The combination of hypermethylation of CpG-2 and hypomethylation of CpG-3 demonstrated the best overall performance with a sensitivity of 80.0% and specificity of 90.9% (**Table 2**). As shown in **Figure 3**, 28 of 29 patients with this combination were responders (14 good and 14 moderate response; right chart, **Figure 3**). In contrast, all four patients failing to satisfy either cut-off were non-responders (left chart, **Figure 3**). The strength of the association of the CpG-2 + CpG-3 combination with response was also reflected in a ROC AUC of 0.85 (95% confidence interval [CI] 0.71, 0.94).

Discussion

This is the first study to examine the link between DNA methylation and first-line treatment response in RA. Using a prospective cohort of patients recruited at first diagnosis and prior to the initiation of treatment, our data indicate that baseline DNA methylation levels for a discrete subset of sites are significantly associated with response to treatment with disease-modifying agents. The methylation status at two specific sites assessed in combination, and which independently were associated with response, proved to be the strongest factor associated with treatment response.

Since early, effective intervention in RA reduces disease activity and inflammation, and improves long-term outcome [37-40], identification of baseline factors associated with treatment response has been a priority. However, examination of a broad range of clinical, molecular and genetic factors has not produced definitive biomarkers [18, 19]. Our findings now provide the first evidence that epigenetic profiling, in this case of DNA methylation, may have significant value in identifying which patients with RA may respond to first-line DMARD treatment. Furthermore, DNA methylation is an attractive biomarker since it is typically stable over time, is minimally affected by short-term stimuli and is readily measured [12]. The potential utility of methylation profiling is further supported by a very recently reported association between differential DNA methylation and response to second-line anti-TNF therapy in RA [41].

We were unable to formally examine the independence of the CpG-2 + CpG-3 association with treatment response in this proof-of-concept study. However, a preliminary assessment using our data suggested that it was independent of baseline

283 clinical variables including disease activity, autoantibodies and systemic inflammatory
284 markers, which individually did not appear to be associated with response. This would
285 be in agreement with the main body of literature, which indicates that ESR, RF and
286 ACPA are not independently associated with response to methotrexate and/or other
287 DMARDs [reviewed in 18]. Although not reported by all studies [42], evidence does
288 indicate that male sex is associated with a better response to methotrexate [43-45]. Our
289 data suggest a possible trend towards better response in males ($p < 0.1$), which may
290 reflect treatment with methotrexate for over 90% of the patients studied.
291
292 The CpG-2 + CpG-3 combination, which we identified as the strongest independent
293 factor associated with treatment response, comprises sites in *ADAMTSL2* (CpG-2), a
294 disintegrin and metalloproteinase with thrombospondin motif-like protein, and in
295 *BTN3A2* (CpG-3), a butyrophilin family member. Although the function of
296 ADAMTSL2 has not been fully determined, evidence supports a role in the regulation
297 of transforming growth factor-β (TGF-β) [46]. TGF-β is a pleiotropic cytokine with
298 important immunoregulatory functions [47, 48], which is implicated in RA synovial
299 pathology [49]. Butyrophilins are transmembrane proteins that share structural
300 similarities with B7 co-stimulatory molecules and are emerging as novel regulators of
301 T-lymphocyte function and immune responses [50, 51].
302
303 We focused on DNA methylation factors associated with response in the context of
304 DMARD treatment strategies that reflected standard clinical practice. Both responder
305 and non-responder groups included patients receiving methotrexate monotherapy and
306 patients receiving combination therapy, the proportions of which were not significantly

different either at baseline or at six-months follow-up (**Supplementary Table 2**). Importantly, methylation at two CpGs in combination was strongly associated with treatment response despite the limited variation in treatment regimens, supporting its potential utility as a marker of response at diagnosis in a real-world clinical setting. Furthermore, we purposefully used the EULAR criteria as the response measure in this study as these are universally accepted and encompass both improvement in disease activity over time and end-point disease activity. Reassuringly, the proportion of responders in this study is consistent with previous reports using these criteria [44, 52]. By quantifying methylation at baseline, we are also able to exclude potential confounding associated with DMARDs, including methotrexate, an impact of which on methylation has been suggested by several groups [9,10,53,54].

Although our proof-of-concept study is the first of its kind in RA, a limitation of our work was the relatively small number of patients that we were able to recruit. In an attempt to address this, we used a number of sequential filtering steps to identify sites differentially methylated between responders and non-responders to treatment. Furthermore, for the two CpGs comprising the strongest biomarker associated with response, we validated the array data by also quantifying methylation using an independent method (Pyrosequencing). This significantly reduces the risk of type I errors associated with genome-wide approaches. However, we recognise that an important next step will be to confirm our findings and determine the true predictive value of this biomarker in larger, independent patient cohorts.

Conclusions

In conclusion, we report the identification of a novel DNA methylation combination that is associated with response to treatment with conventional disease-modifying drugs in newly diagnosed patients with RA. Whilst our findings will require verification in larger, independent early RA cohorts, they provide the first evidence to support epigenetic profiling as a novel approach to identifying biomarkers associated with response to DMARD therapy. Ultimately, this has the potential to inform clinical management and patient care, towards the goal of a stratified, personalized medicine approach to treatment.

Executive Summary

Background

- Newly diagnosed patients with rheumatoid arthritis (RA) demonstrate variability of response to treatment with disease-modifying antirheumatic drugs (DMARDs).
- To date, no definitive biomarkers associated with response have been identified.
- This proof-of-concept study explored whether DNA methylation at first diagnosis is associated with response to treatment with DMARDs in patients with treatment-naïve early RA.

Patients & Methods

- HumanMethylation450 BeadChips were used to quantify genome-wide DNA methylation at diagnosis in T-lymphocytes from 46 treatment-naïve patients with early RA.
- Response to DMARD treatment was determined at six months using the DAS28-based EULAR response criteria. Sensitivity, specificity and receiver operating characteristic AUC data were used to assess associations of baseline methylation with treatment response.

Results

- At six-months, the numbers of patients achieving a good/moderate/no response to treatment were 16/19/11 (35/41/24%), respectively.
- Array analysis identified 21 CpGs displaying methylation differences between responders and non-responders, of which four statistically significant sites (p_{adj} <0.05, Bonferroni) showed high sensitivity and/or specificity ≥75% for treatment response.

- Methylation at two individual sites in combination (cg0301849 and cg14345882) was the strongest factor associated with response, with 80.0% sensitivity and 90.9% specificity (AUC 0.85). 28 of 29 patients with this combination were responders.

Conclusions

- DNA methylation of a novel CpG combination is associated with treatment response at first diagnosis in early RA patients prior to commencing treatment with DMARDs.
- These findings provide the first evidence to support epigenetic profiling as a novel approach to identifying biomarkers associated with DMARD treatment response in RA. This may ultimately have the potential to inform clinical management and patient care.

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 Kingdom: new estimates for a new century. *Rheumatology* 41(7), 793–800 (2002).
2. Myasoedova E, Crowson CS, Kremers HM, 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. O'Dell JR, Haire CE, Erikson N, *et al*. Treatment of rheumatoid arthritis with methotrexate alone, sulfasalazine and hydroxychloroquine, or a combination of all three medications. *N. Engl. J. Med*. 334(20), 1287–1291 (1996).
4. O'Dell JR, Leff R, Paulsen G, *et al*. Treatment of rheumatoid arthritis with methotrexate and hydroxychloroquine, methotrexate and sulfasalazine, or a combination of the three medications: results of a two-year, randomized, double-blind, placebo-controlled trial. *Arthritis Rheum*. 46(5), 1164–1170 (2002).
5. Nurmohamed MT, Dijkmans BA. Efficacy, tolerability and cost effectiveness of disease-modifying antirheumatic drugs and biologic agents in rheumatoid arthritis. *Drugs* 65(5), 661–694 (2005).
6. Bottini N, Firestein GS. Epigenetics in rheumatoid arthritis: a primer for rheumatologists. *Curr. Rheumatol. Rep*. 15(11), 372 (2013).
7. Glant TT, Mikecz K, Rauch TA. Epigenetics in the pathogenesis of rheumatoid arthritis. *BMC Med*. 12, 35 (2014).

8. Klein K, Gay S. Epigenetic modifications in rheumatoid arthritis, a review. *Curr. Opin. Pharmacol.* 13(3), 420–425 (2013).
9. 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).
10. de Andres MC, Perez-Pampin E, Calaza M, *et al*. Assessment of global DNA methylation in peripheral blood cell subpopulations of early rheumatoid arthritis before and after methotrexate. *Arthritis Res. Ther.* 17, 233 (2015).
***First RA study of global DNA methylation levels in distinct blood cell subpopulations and apparent reversal with methotrexate treatment**
11. Esteller M, Garcia-Foncillas J, Andion E, *et al*. Inactivation of the DNA-repair gene MGMT and the clinical response of gliomas to alkylating agents. *N. Engl. J. Med.* 343(19), 1350–1354 (2000).
12. Plant D, Wilson AG, Barton A. Genetic and epigenetic predictors of responsiveness to treatment in RA. *Nat. Rev. Rheumatol.* 10(6), 329–337 (2014).
13. Cribbs A, Feldmann M, Oppermann U. Towards an understanding of the role of DNA methylation in rheumatoid arthritis: therapeutic and diagnostic implications. *Ther. Adv. Musculoskelet. Dis.* 7(5), 206–219 (2015).
14. Bird A. DNA methylation patterns and epigenetic memory. *Genes Dev.* 16(1), 6–21 (2002).
15. 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* 2014;**9**(9):1228–37.

16. Glossop JR, Emes RD, Nixon NB, *et al*. Genome-wide profiling in treatment-naïve early rheumatoid arthritis reveals DNA methylome changes in T- and B-lymphocytes. *Epigenomics* 2016;**8**(2):209-24

***First DNA methylation array study in newly diagnosed RA patients naïve for treatment with disease-modifying antirheumatic drugs**

17. Farragher TM, Lunt M, Fu B, Bunn D, Symmons DP. Early treatment with, and time receiving, first disease-modifying antirheumatic drug predicts long-term function in patients with inflammatory polyarthritis. *Ann. Rheum. Dis*. 69(4), 689–695 (2010).

18. Romao VC, Canhao H, Fonseca JE. Old drugs, old problems: where do we stand in the prediction of rheumatoid arthritis responsiveness to methotrexate and other synthetic DMARDs? *BMC Med*. 11, 17 (2013).

***Comprehensive review article discussing current evidence for predicting response to disease modifying drugs in RA.**

19. Smith SL, Plant D, Eyre S, Barton A. The potential use of expression profiling: implications for predicting treatment response in rheumatoid arthritis. *Ann. Rheum. Dis*. 72(7), 1118–1124 (2013).

20. Grabiec AM, Reedquist KA. The ascent of acetylation in the epigenetics of rheumatoid arthritis. *Nat. Rev. Rheumatol*. 9(5), 311–318 (2013).

21. Cope AP. T cells in rheumatoid arthritis. *Arthritis Res. Ther*. 10(Suppl 1), S1 (2008).

22. Peres RS, Liew FY, Talbot J, *et al*. Low expression of CD39 on regulatory T cells as a biomarker for resistance to methotrexate therapy in rheumatoid arthritis. *Proc. Natl. Acad. Sci. U. S. A*. 112(8), 2509–2514 (2015).

23. Carvalheiro H, Duarte C, Silva-Cardoso S, da Silva JA, Souto-Carneiro MM. CD8+ T cell profiles in patients with rheumatoid arthritis and their relationship to disease activity. *Arthritis Rheumatol*. 67(2), 363–371 (2015).
24. Kuuliala K, Kuuliala A, Koivuniemi R, *et al*. Constitutive STAT3 phosphorylation in circulating CD4+ T lymphocytes associates with disease activity and treatment response in recent-onset rheumatoid arthritis. *PLoS One* 10(9), e0137385 (2015).
25. Aletaha D, Neogi T, Silman AJ, *et al*. 2010 Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Ann. Rheum. Dis*. 69(9), 1580–1588 (2010).
26. 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).
27. van Gestel AM, Prevoo ML, van 't Hof MA, van Rijswijk MH, van de Putte LB, van Riel PL. Development and validation of the European League Against Rheumatism response criteria for rheumatoid arthritis. Comparison with the preliminary American College of Rheumatology and the World Health Organization/International League Against Rheumatism Criteria. *Arthritis Rheum*. 39(1), 34–40 (1996).
28. van Gestel AM, Haagsma CJ, van Riel PL. Validation of rheumatoid arthritis improvement criteria that include simplified joint counts. *Arthritis Rheum*. 41(10), 1845–1850 (1998).
29. Bibikova M, Barnes B, Tsan C, *et al*. High density DNA methylation array with single CpG resolution. *Genomics* 98(4), 288–295 (2011).

30. 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* 2013;**8**(11):1188–97.
31. Mill J, Yazdanpanah S, Guckel E, Ziegler S, Kaminsky Z, Petronis A. Whole genome amplification of sodium bisulfite-treated DNA allows the accurate estimate of methylated cytosine density in limited DNA resources. *Biotechniques* 41(5), 603–607 (2006).
32. Don RH, Cox PT, Wainwright BJ, Baker K, Mattick JS. 'Touchdown' PCR to circumvent spurious priming during gene amplification. *Nucleic Acids Res*. 19(14), 4008 (1991).
33. Haworth KE, Farrell WE, Emes RD, *et al*. Combined influence of gene-specific cord blood methylation and maternal smoking habit on birth weight. *Epigenomics* 2013;**5**(1):37–49.
34. 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).
35. Maksimovic J, Gordon L, Oshlack A. SWAN: Subset-quantile within array normalization for illumina infinium HumanMethylation450 BeadChips. *Genome Biol*. 13(6), R44 (2012).
36. 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).

37. Finckh A, Liang MH, van Herckenrode CM, de Pablo P. Long-term impact of early treatment on radiographic progression in rheumatoid arthritis: a meta-analysis. *Arthritis Rheum*. 55(6), 864–872 (2006).

38. de Vries-Bouwstra JK, Goekoop-Ruiterman YP, Verpoort KN, *et al*. Progression of joint damage in early rheumatoid arthritis: association with HLA-DRB1, rheumatoid factor, and anti-citrullinated protein antibodies in relation to different treatment strategies. *Arthritis Rheum*. 58(5), 1293–1298 (2008).

39. Klareskog L, Catrina AI, Paget S. Rheumatoid arthritis. *Lancet* 373(9664), 659–672 (2009).

40. Scirè CA, Verstappen SM, Mirjafari H, *et al*. Reduction of long-term disability in inflammatory polyarthritis by early and persistent suppression of joint inflammation: results from the Norfolk Arthritis Register. *Arthritis Care Res. (Hoboken)* 63(7), 945–952 (2011).

41. Plant D, Webster A, Nair N, *et al*. Differential methylation as a biomarker of response to etanercept in patients with rheumatoid arthritis. *Arthritis Rheumatol*. 68(6), 1353–1360 (2016).

***Association between differential DNA methylation and response to second-line anti-TNF therapy in RA.**

42. Hider SL, Silman AJ, Thomson W, Lunt M, Bunn D, Symmons DP. Can clinical factors at presentation be used to predict outcome of treatment with methotrexate in patients with early inflammatory polyarthritis? *Ann. Rheum. Dis*. 68(1), 57–62 (2009).

511 43. Wessels JA, van der Kooij SM, le Cessie S, *et al*. A clinical pharmacogenetic
512 model to predict the efficacy of methotrexate monotherapy in recent-onset
513 rheumatoid arthritis. *Arthritis Rheum*. 56(6), 1765–1775 (2007).
514 44. Saevarsdottir S, Wallin H, Seddighzadeh M, *et al*. Predictors of response to
515 methotrexate in early DMARD naive rheumatoid arthritis: results from the initial
516 open-label phase of the SWEFOT trial. *Ann. Rheum. Dis*. 70(3), 469–475 (2011).
517 45. Ma MH, Ibrahim F, Walker D, *et al*. Remission in early rheumatoid arthritis:
518 predicting treatment response. *J. Rheumatol*. 39(3), 470–475 (2012).
519 46. Le Goff C, Morice-Picard F, Dagoneau N, *et al*. ADAMTSL2 mutations in
520 geleophysic dysplasia demonstrate a role for ADAMTS-like proteins in TGF-β
521 bioavailability regulation. *Nat. Genet*. 40(9), 1119–1123 (2008).
522 47. Schmidt-Weber CB, Blaser K. Regulation and role of transforming growth factor-
523 beta in immune tolerance induction and inflammation. *Curr. Opin. Immunol*. 16(6),
524 709–716 (2004).
525 48. Li MO, Flavell RA. TGF-β: A master of all T cell trades. *Cell* 134(3), 392–404
526 (2008).
527 49. Pohlers D, Beyer A, Koczan D, Wilhelm T, Thiesen HJ, Kinne RW. Constitutive
528 upregulation of the transforming growth factor-β pathway in rheumatoid arthritis
529 synovial fibroblasts. *Arthritis Res. Ther*. 9(3), R59 (2007).
530 50. Afrache H, Gouret P, Ainouche S, Pontarotti P, Olive D. The butyrophilin (BTN)
531 gene family: from milk fat to the regulation of the immune response.
532 *Immunogenetics* 64(11), 781–794 (2012).
533 51. Abeler-Dorner L, Swamy M, Williams G, Hayday AC, Bas A. Butyrophilins: an
534 emerging family of immune regulators. *Trends Immunol*. 33(1), 34–41 (2012).

52. Svensson B, Schaufelberger C, Teleman A, Theander J. Remission and response to early treatment of RA assessed by the Disease Activity Score. *Rheumatology (Oxford)* 39(9), 1031–1036 (2009).
53. 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).
54. 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*. 67(5), 1182–1192 (2015).

Financial disclosure/Acknowledgements

The authors would like to thank the patients who participated in the study. We also thank Janet Turner, Cath Thwaites and Moira Dishman for assistance with the collection of clinical data. This work was supported by funding provided by 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 at baseline (pre-treatment) between treatment responders and non-responders in patients with early RA. The starting number of CpGs indicated (482,421) is the total number of CpGs on the methylation array platform. Following initial processing (step 1), data were normalized using SWAN [35], implemented in the Bioconductor package Minfi [34]. Numbers in the figure indicate the number of CpGs remaining at each successive step.

Abbreviations: RA, rheumatoid arthritis; SWAN, subset-quantile within array normalization.

Figure 2. Pre-treatment methylation status discriminates responders and non-responders in patients with early RA. In **(A)** CpG-2 (cg03018489) and **(B)** CpG-3 (cg14345882), non-responders (n = 11) and responders (n = 35) are depicted by open circles and filled triangles, respectively, and where responders are divided into those showing a moderate (centre, n = 19) and good (right, n = 16) response to treatment. Good, moderate and no response categories are defined in the EULAR response criteria [23-25]. The horizontal dashed line indicates the methylation cut-off for distinguishing between responders and non-responders, and the short horizontal bar in each group indicates the mean value.

Abbreviations: RA, rheumatoid arthritis; EULAR, European League Against Rheumatism).

Figure 3. Pre-treatment methylation status at two CpG sites in combination is associated with response to treatment in patients with early RA patients. For CpG-2 (cg03018489) and CpG-3 (cg14345882) methylation status was defined as hypermethylated (above) or hypomethylated (below) relative to a cut-off of 60% and 20%, respectively. Shown on the x-axis are the four possible methylation combinations, with methylation status of CpG-2 given first and of CpG-3 given second, as indicated (the two combinations in which only one CpG satisfied the cut-off value are grouped together (centre chart)). Each chart depicts the proportion of patients achieving a good (white), moderate (striped) and no response (dark grey) to treatment, stratified by methylation status for the CpG-2/CpG-3 combination.

Abbreviations: RA, rheumatoid arthritis; Hypo, hypomethylated; Hyper, hypermethylated.

Supplementary Figure 1. Technical validation by bisulfite pyrosequencing of baseline methylation status for two CpGs differentially methylated between responders and non-responders in patients with early RA. In both **(A)** CpG-2 (cg03018489) and **(C)** CpG-3 (cg14345882), responders (n = 35) and non-responders (n = 11) are depicted by triangles and circles respectively. The short red horizontal bar shown in each group indicates the mean value. For each CpG, methylation values are shown for the array (filled symbols; left) and Pyrosequencing (open symbols; right). Bland-Altman plots in (B) CpG-2 (cg03018489) and (D) CpG-3 (cg14345882) show the agreement between % methylation levels as determined by 450K array and pyrosequencing analysis. Each point represents an individual patient. Shown by horizontal lines are the mean difference between the methods (bias) and the upper and

lower boundaries of the 95% limits of agreement (± 1.96 SD). The intraclass correlation coefficient between the methods is 0.963 for CpG-2, and 0.690 for CpG-3.

Abbreviations: RA, rheumatoid arthritis; 450K, HumanMethylation450 BeadChip

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

‡The sequence indicated 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 in the forward and reverse orientation, respectively).

Abbreviations: bp, base pairs.

Supplementary Table 2. Baseline demographic and clinical characteristics in early RA patients who responded and did not respond to DMARD treatment at 6-months follow-up.

* Mann-Whitney U test (continuous variables) or Fisher's exact test (categorical), as appropriate.

† data unavailable for two patients.

‡ data unavailable for one patient.

§ 26/45 (57.8%) patients were positive for ACPA/ RF (data unavailable for one patient).

¥ The total number of patients starting treatment with a given DMARD, whether received as monotherapy or in combination with other DMARDs.

* One further patient received monotherapy with hydroxychloroquine.

One patient was not receiving DMARD treatment.

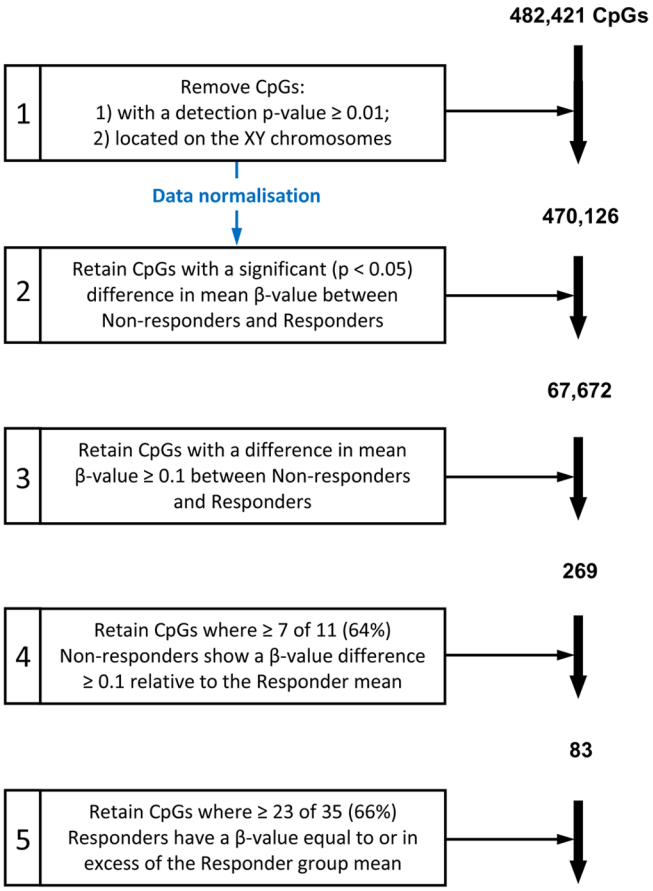
Abbreviations: RA, rheumatoid arthritis; DMARDs, disease-modifying anti-rheumatic drugs; RF, rheumatoid factor; ACPA, anti-citrullinated peptide antibodies; DAS28, disease activity score with 28-joint count; ESR, erythrocyte sedimentation rate.

Supplementary Table 3. Complete list and annotation for the 21 CpGs identified as differentially methylated at baseline (pre-treatment) between responders and non-responders in patients with early RA patients.*

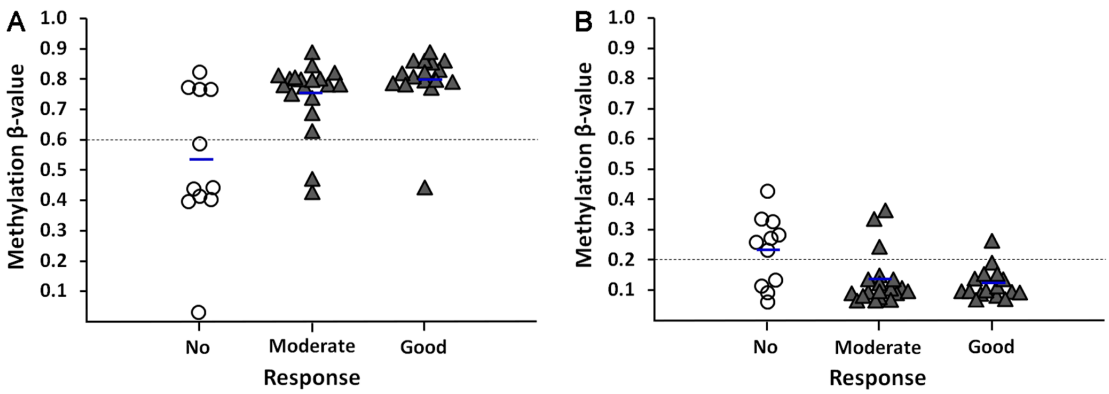
*Bold blue font indicates CpGs with statistically significant ($p < 0.05$, Bonferroni-adjusted) differences in methylation between responders and non-responders**.** The dashed horizontal line between rows 18 and 19 separates CpGs that were hypermethylated (above) and hypomethylated (below) in responders relative to non-responders.

†The 'dmpFinder' function in Minfi [34] was used to calculate F-test p-values.

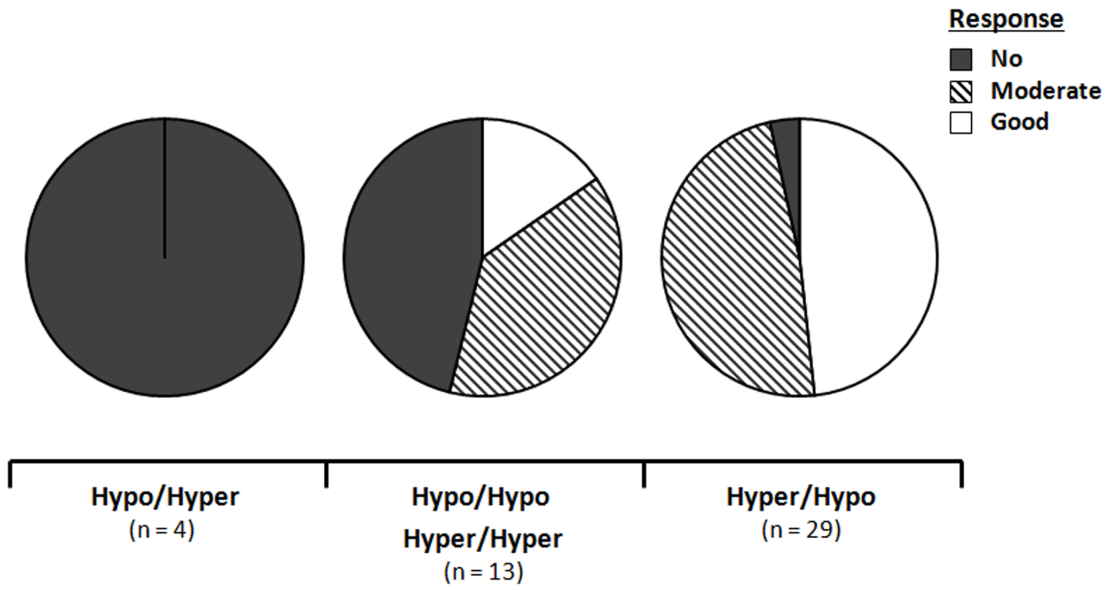
Abbreviations: RA, rheumatoid arthritis.


482,421 CpGs
1
Remove CpGs:
1) with a detection p-value ≥ 0.01;
2) located on the XY chromosomes
Data normalisation
470,126
2
Retain CpGs with a significant (p < 0.05) difference in mean β-value between Non-responders and Responders
67,672
3
Retain CpGs with a difference in mean β-value ≥ 0.1 between Non-responders and Responders
269
4
Retain CpGs where ≥ 7 of 11 (64%) Non-responders show a β-value difference ≥ 0.1 relative to the Responder mean
83
5
Retain CpGs where ≥ 23 of 35 (66%) Responders have a β-value equal to or in excess of the Responder group mean
21

641

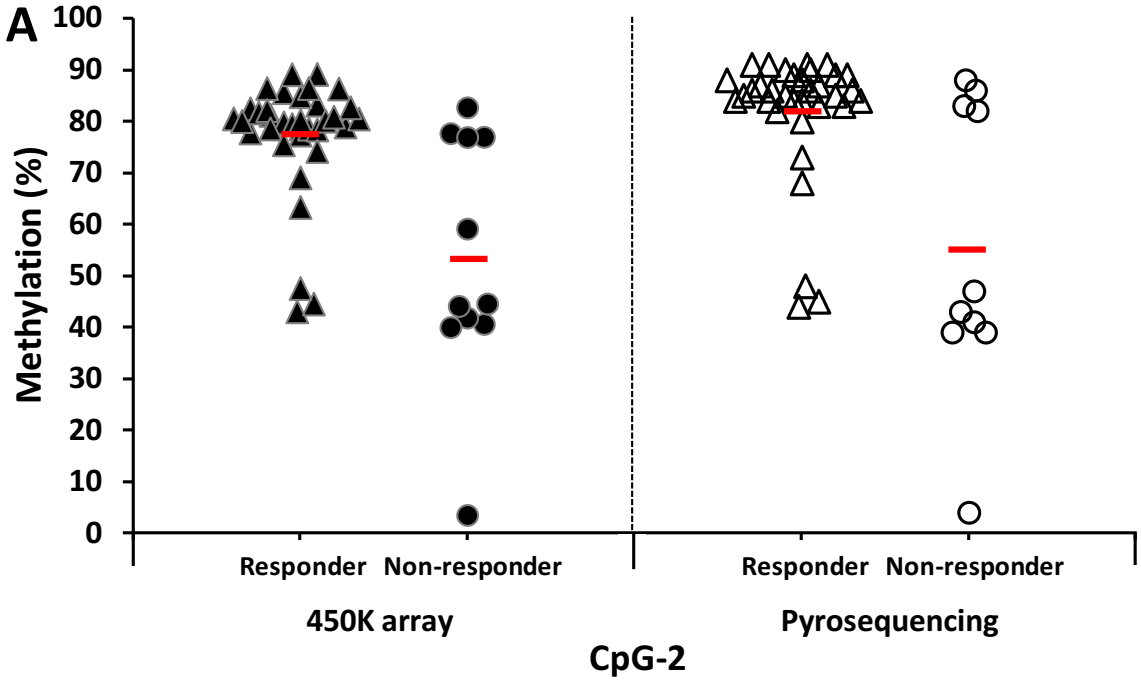

A
B
Methylation β-value
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
No
Moderate
Good
Response

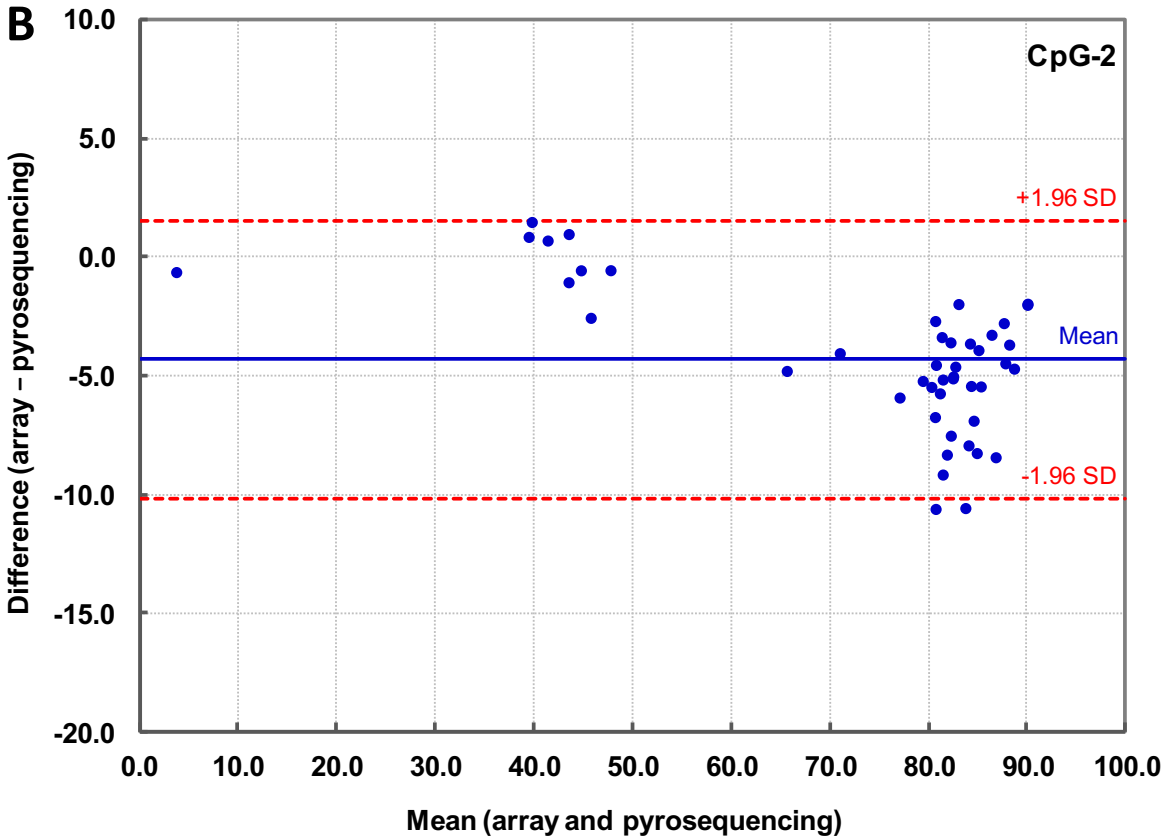
642


Response
No
Moderate
Good
Hypo/Hyper
(n = 4)
Hypo/Hypo
Hyper/Hyper
(n = 13)
Hyper/Hypo
(n = 29)
CpG-2 + CpG-3 methylation

643

644





continued over…

Table 1. Demographic and clinical characteristics at baseline for the cohort of 46 treatment-naïve patients with early RA.

Number	46
Male/female, No. (%)	16/30 (34.8/65.2)
Age, mean ± SD (years)	57.7 ± 13.9
RF positive, No. (%)[†§]	23 (52.3)
ACPA positive, No. (%)[‡§]	22 (48.9)
DAS28, mean ± SD	5.29 ± 1.4
ESR, mean ± SD	30.1 ± 23.7
Corticosteroids, No. (%)	45 (97.8)
Starting DMARD, No. (%)[¥]	
Methotrexate (MTX)	43 (93.5)
Hydroxychloroquine (HCQ)	29 (63.0)
Sulphasalazine (SSZ)	23 (50.0)
Starting treatment regimens, No. (%)	
*Monotherapy (MTX)**	15 (32.6)
Triple therapy (MTX+HCQ+SSZ)	20 (43.5)
Dual therapy (two of MTX, HCQ and SSZ)	10 (21.7)

† of 44 patients (data unavailable for two patients).

‡ of 45 patients (data unavailable for one patient).

§ 26/45 (57.8%) patients were positive for ACPA/ RF (data unavailable for one patient).

¥ The total number of patients starting treatment with a given DMARD, whether received as monotherapy or in combination with other DMARDs.

* One further patient started monotherapy with hydroxychloroquine.

646

Table 2. Association of baseline methylation status with treatment response in patients with early RA.*

pG ID	Methylation in responders: Hyper/Hypo	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	ROC AUC (95% CI)
ıdividual sites						
pG-1 *(cg07225509)*	Hyper	77.1	72.7	90.0	50.0	0.75 (0.59, 0.86)
pG-2 *(cg03018489)*	Hyper	91.4	63.6	88.9	70.0	0.78 (0.64, 0.89)
pG-3 *(cg14345882)*	Hypo	88.6	63.6	88.6	63.6	0.76 (0.61, 0.87)
pG-4 *(cg23974730)*	Hypo	82.9	63.6	87.9	53.9	0.73 (0.59, 0.86)
ombinations						
pG-1 + CpG-2	Hyper/Hyper	71.4	90.9	96.2	50.0	0.81 (0.66, 0.91)
pG-1 + CpG-3	Hyper/Hypo	65.7	81.8	92.0	42.9	0.74 (0.59, 0.86)
pG-1 + CpG-4	Hyper/Hypo	60.0	90.9	95.5	41.7	0.75 (0.61, 0.87)
pG-2 + CpG-3	**Hyper/Hypo**	**80.0**	**90.9**	**96.6**	**58.8**	**0.85 (0.71, 0.94)**
pG-2 + CpG-4	Hyper/Hypo	77.1	72.7	90.0	50.0	0.75 (0.59, 0.86)
pG-3 + CpG-4	Hypo/Hypo	74.3	90.9	96.3	52.6	0.83 (0.69, 0.92)

*Of the six CpGs identified as significantly differentially methylated between responders and non-responders (see main text), shown are the four CpGs with a sensitivity and/or specificity ≥75% and that showed most promise for discriminating between responders and non-responders. Also shown are the six possible CpG pairs derived from these four sites. All individual sites and combinations shown were significantly associated with treatment response ($p < 0.05$, Fisher's exact test). The CpG-2 + CpG-3 combination displayed the best overall performance ($p < 0.001$; bold font).