Supplementary Materials:

Supplementary Methods:

Immunophenotyping

Whole blood was collected from each patient at study enrollment and was processed within 3 hours by the Mount Sinai's Human Immune Monitoring Core. Cells were stained with a pre-optimized T-cell antibody cocktail that contained antiCD45, antiCD3, antiCD4, antiCD8, antiCCR4 and antiCCR6 antibodies (Supplementary Table 2) and analyzed on a BD LSR Fortessa (BD, San Jose, CA). FlowJo 9.4 software (Treestar Inc, San Carlos, CA) was used for post-acquisition analysis (Supplementary Figure S2). As the Human Immune Monitoring Core analysis was performed at the initial visit of the study, complete immunophenotyping data were available only for the discovery cohort. One GA-treated patient is missing from the analysis of CCR6 and CCR4 expression in CD8 T cells (total 15 GA-treated patients for the CCR6/CCR4 analysis of CD8 T cell).

DNA methylation analysis

CD4 T cells were positively isolated using magnetic beads (Miltenyi Biotec) and DNA was extracted with QIAamp DNA Blood Mini Kit (Qiagen). 500ng of genomic DNA from each sample were bisulfite-treated with Methylamp One-Step DNA Modification Kit (EpiGentek) and methylation levels were measured at ~850,000 CpG sites by using the Infinium MethylationEPIC BeadChip array at the New York Genome Center. Probes with detection p value more than 0.01 and bead count less than 3 in more than 5% of the samples were filtered out. Non-cg probes, probes that fall near SNPs as defined by Zhou et al (Zhou *et al.*, 2017) or align to multiple locations as defined by Nordlund et al (Nordlund *et al.*, 2013) were removed.

Finally, probes at the sex chromosomes X and Y were also removed. β-values were normalized with BMIQ v1.6 to correct for Type-I and Type-II probe bias. PCA analysis revealed 2 outliers in the treatment-naïve group of the discovery cohort which were removed from the rest of the analysis (Supplementary Figure S3).

As we have previously described (Huynh *et al.*, 2014; Watson *et al.*, 2016), our analysis was focused on genomic regions rather than individual CpG sites as regulatory DNA modifications generally involve multiple consecutive CpGs. To identify differentially methylated regions between treated and untreated patients, we utilized a linear regression model at each individual CpG site to identify the contribution of treatment status in DNA methylation changes after controlling for age, gender, race, disease duration, the total CD4 T cell percentage as well as the percentage of CCR6-CCR4+, CCR6+CCR4+, CCR6+CCR4- CD4 T cells in each sample for the discovery cohort. Since CCR4 and CCR6 are expressed mainly in memory CD4 T cells (Fagin *et al.*, 2012a) and highly correlate with the total number of CD4 memory cells (Pearson, n=5, r=0.9833, p=0.0026) (Supplementary Figure S4), our above analysis would be able to control for potential naïve/memory imbalances of our samples. We then used a 1 kb sliding window to define genomic regions with closely located CpG sites, with a maximum genomic length of 3.33kb. Based on these regions, we combined each CpG specific p value from our linear regression model within a single region with the Stouffer's method (Whitlock, 2005). This was followed by the Bonferroni correction for multiple hypothesis testing. Differentially methylated regions (DMRs) were defined as regions with more than 4 CpGs that have an absolute median β-value change greater than 0.02 and an adjusted combined p value of less than 0.01. Data analysis was performed in R Studio by utilizing the R packages ChAMP (Tian *et al.*, 2017) for data preprocessing and normalization, LIMMA for linear regression (Ritchie *et al.*,

2015), bumphunter for the 1kb sliding window function (Jaffe *et al.*, 2012) and DMRCATE for the p value combination with the Stouffer method (Peters *et al.*, 2015).

Controlling for cell-type heterogeneity in our DNA methylation analysis

DNA methylation profiles are directly linked with cell types and, for CD4 T cells in particular, with the frequency of naive/memory T cells. The DNA methylation profiles of these cells, however, follow a linear pattern, as shown in (Durek *et al.*, 2016) and, thus, cell-specific methylation patterns/biases can be accounted for within a linear model framework. Notably, the majority of the currently available methods to correct for cell-type heterogeneity in DNA methylation studies use the estimated proportions of different cell-types as covariates in a linear model. (Michels *et al.*, 2013; Lehne *et al.*, 2015; Rahmani *et al.*, 2016; Teschendorff and Zheng, 2017; Zou, 2017). Moreover, several publications have used a similar approach to our analysis (Huynh *et al.*, 2014; Watson *et al.*, 2016). Using a reference dataset to estimate proportions of naïve/memory CD4 T cells would introduce more bias, as these reference datasets are either from healthy controls or patients with other immune mediated diseases. It is not safe to assume that the methylation patterns of immune cells are similar between MS and other immune mediated diseases or healthy individuals, as shown in (Maltby *et al.*, 2017; Ruhrmann *et al.*, 2017), where CD4 DNA methylation levels differ between MS and healthy controls. Reference-free methods on the other hand may remove variation attributable to the endpoint of interest and overcorrect the dataset (Lehne *et al.*, 2015; Teschendorff and Zheng, 2017). Our study design has the advantage of having the cell proportions measured by FACS, which represent the true proportions of the cell-types in the same samples that we used for the DNA methylation analysis. Moreover, it is worth noting that the cell proportions we measured, CCR6+CCR4+, CCR6+CCR4- and CCR6-CCR4+, highly and significantly correlate with the memory T cell

proportion as measured by CD45RO+ and CCR7+/CCR7- from 5 healthy individuals (r=0.9833, p=0.0026) (Supplementary Figure S4). Therefore, by including these FACS measured cell-type proportions in our linear model we could effectively control for the cell-type heterogeneity that could be introduced by the difference in naïve/memory CD4 T cell percentages between treated and untreated patients. To further reinforce our findings, we conducted a separate DNA methylation decomposition analysis that investigated the DNA methylation changes within the different cell types from each sample, thus our analysis is controlling for cell type heterogeneity with two different methods.

Decomposition of DNA methylation values

To determine the contribution of each cell type to the DNA methylation changes observed at the *MIR-21* locus in our analysis, we decomposed the β-values of the CpG sites in that locus to each cell type. As it has already been suggested, (Michels *et al.*, 2013), the beta values we obtain with DNA methylation analysis in the context of Epigenome Wide Association Studies, are the percentage of cells in any given sample at any given CpG site that are methylated and thus the measured proportion of methylated DNA is assumed to increase as a linear mixture of the distinct cell-specific methylation profiles.

For example:

Betavalue = ((Proportion of cell type1) x (Methylation level of cell type 1)) +

((Proportion of cell type2) x (Methylation level of cell type 2)) +

((Proportion of cell type3) x (Methylation level of cell type 3)) +

((Proportion of cell type4) x (Methylation level of cell type 4)) +

In previous studies (Houseman *et al.*, 2012), this equation has been used to estimate the

proportions of cell types based on their pre-determined methylation levels from a reference dataset. This has been done by using least squares regression with the reference DNA profiles as independent variables and the measured beta value as the dependent variable. The coefficients of this regression model then would represent the cell proportions (Houseman *et al.*, 2012; Teschendorff and Zheng, 2017). To avoid negative values for these coefficients, they are constrained to be non-negative and/or to sum to 1 (Houseman *et al.*, 2012; Teschendorff and Zheng, 2017).

In our study, we already had the proportions of cell types measured by FACS, but we wanted to identify the cell type specific methylation levels. To this end we used the same equation as above with the difference that we used the cell proportions as the independent variables, and the coefficients of the model as the cell specific methylation levels. We also used a least squares regression model to obtain those coefficients, which we constrained to be between 0 and 1, as this is the biological range of beta values. We then fit the model by using the FACS-measured cell proportions and the measured beta values for a given CpG site from either all the treatment naïve patients or all the FAE treated patients.

For example:

Betavalue of the miR21 promoter =

((Proportion of Th2) x (Methylation level of Th2 at the miR21 promoter)) +

((Proportion of Th17) x (Methylation level of Th17 at the miR21 promoter)) +

((Proportion of Th1Th17) x (Methylation level of Th1Th17 at the miR21 promoter)) +

((Proportion of DN) x (Methylation level of DN at the miR21 promoter)) +

By fitting the above model, we can estimate the methylation level for each cell type at a

given CpG site that is the most representative for either the treatment naïve or FAE treated patient population. However, to compare these cell-type specific methylation levels between the two groups, we estimated their distribution by bootstrapping the residuals of the model. Bootstrapping is a widely used nonparametric statistical method that can estimate the distribution of almost any statistic by resampling with replacement from the original sample (Efron, 1979). For example, if there are 47 original samples (treatment naive patients), resampling with replacement will create a new set of 47 samples by randomly selecting one from the original set 47 times. This may result in some samples not being selected in the new dataset, whereas other samples could be selected more than once. By doing this 10000 times, we can obtain the distribution of the DNA methylation values from the treatment naïve patients for each different cell type. Of note, this is a similar approach to the method that was used by (*Houseman et al.*, 2012) to estimate the distribution of cell proportions from reference DNA methylation values in their method detailed above (Houseman *et al.*, 2012). This distribution now represents the "null distribution", which we can use to obtain a p-value for the cell specific DNA methylation values that we obtained from the FAE treated patients with our model. To this end we used the cumulative distribution function of the null distribution, which essentially tells us the probability of obtaining a methylation value from the null distribution that is less or equal to a given methylation value. By forming the null hypothesis that the cell specific DNA methylation values of the FAE treated patient population come from the same distribution as the cell specific DNA methylation values of the treatment naïve patient group (null distribution), we estimated the probability of obtaining a methylation level from the null distribution that is equal or higher than the one estimated from the FAE treated patients (p-value). Given that the obtained p-value was <0.05 only for the Th17 cell subtype, we rejected the null hypothesis for that cell type, in favor

of the alternative hypothesis, which suggests that FAE-treated patients have a hypermethylated miR21 promoter in their Th17 cells compared to treatment naïve patients. This analysis was done in each *MIR-21* DMR CpG site separately and in the average methylation value of the *MIR-21* locus. Data analysis was performed in python 2.7 by utilizing the packages scipy 0.19.0 (for constrained least squares: scipy.optimize.lsq_linear), numpy 1.13.1 and pandas 0.20.2

Power of the pair-wise DNA methylation analysis

Measurements regarding DNA from the same subject can have a very low variance and therefore, given a paired study design, smaller sample sizes can be sufficient. One example is this study by Baranzini et al published in Nature (Baranzini et al., 2010) that investigated genetic and epigenetic differences in three pairs of monozygotic twins that were only discordant for the diagnosis of Multiple Sclerosis. Indeed, the standard deviation of the methylation level at the miR21 DMR in the discovery cohort is 0.097. On the other hand, because of the paired design, the standard deviation of differences at the miR21 DMR from the longitudinal cohort is only 0.044. Based on this standard deviation of differences, the sample size required for 80% power for a paired t-test to detect an effect size of 0.05 (based on the discovery cohort) and alpha=0.05 is 6 (Sańchez, 1993). Thus, our study, based on the experimental design, has sufficient power to detect changes at the miR21 locus.

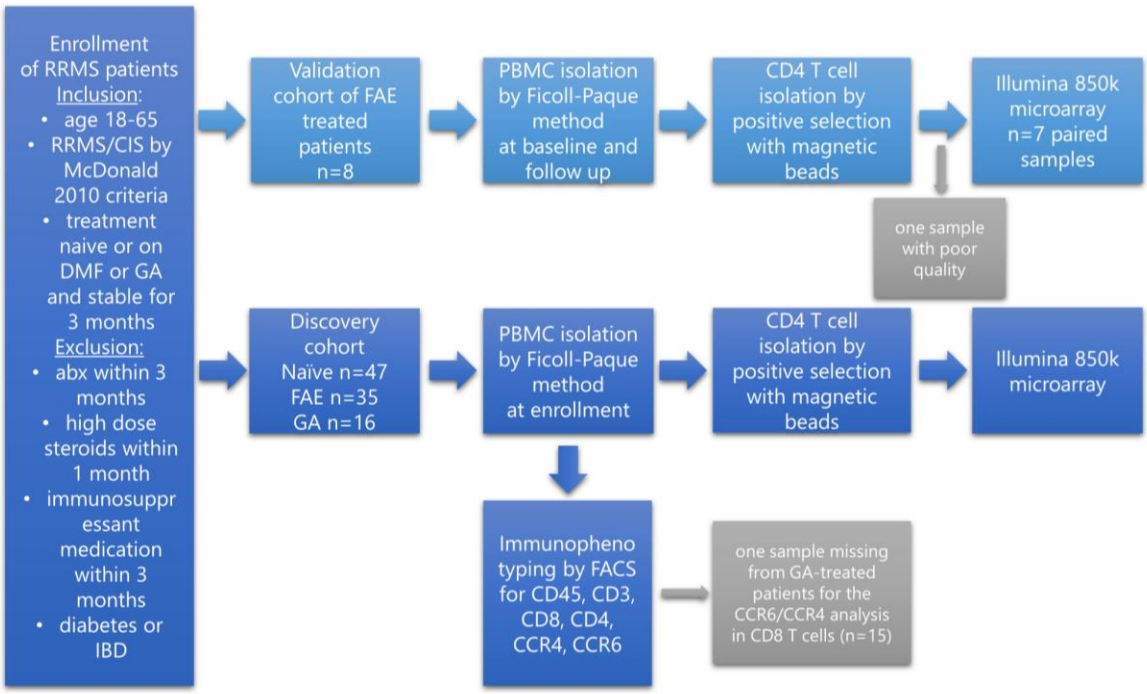
Isolation and in vitro culture of naïve and memory CD4 T cells

PBMCs were collected from healthy donors by the Mount Sinai's Human Immune Monitoring Core and stored in liquid nitrogen until further use. After thawing, PBMCs were stained with anti-CD4-FITC (BD), anti-CD8-PerCP (BD), anti CD45RO-BV421 (BD) and anti-CCR7-PE (BD) antibodies for 20 mins in PBS plus 3%FBS on ice. Naïve CD45RO-CCR7+ CD4+ or CD45RO-CCR7+ CD8+ T cells and memory CD45RO+ CD4+ T cells were then

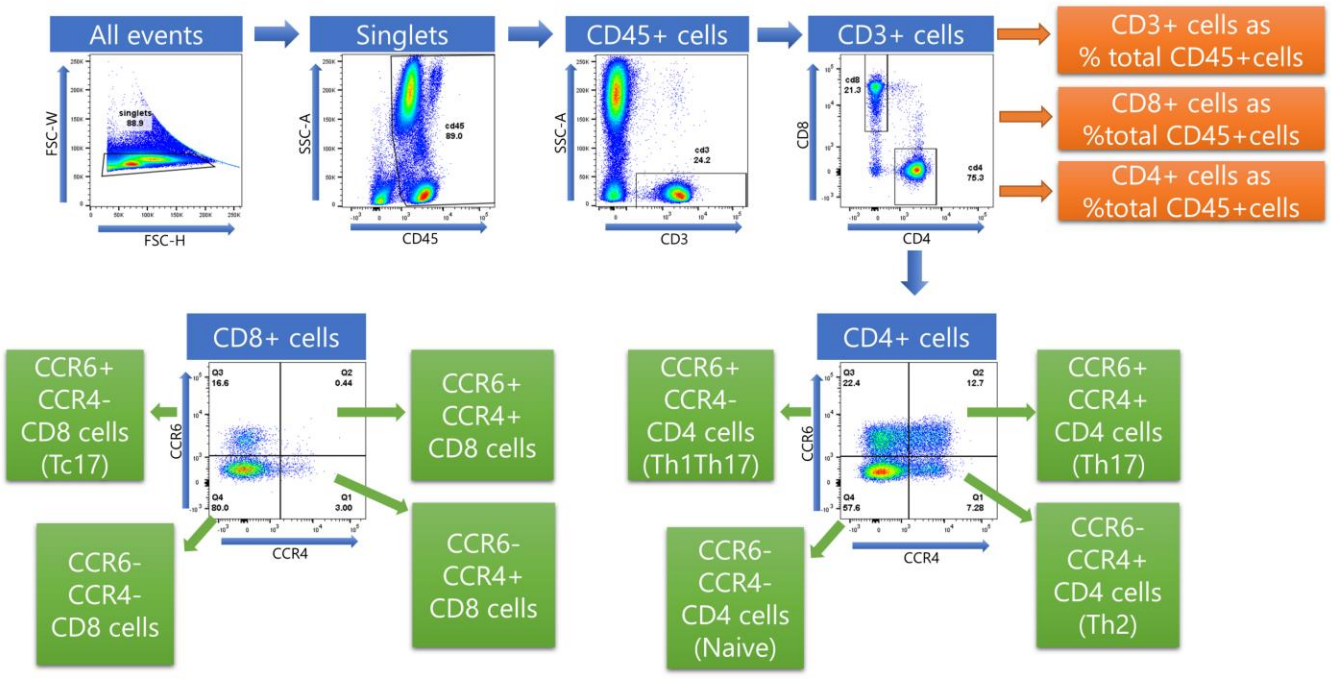
isolated by FACS on a BD FACS Aria Fusion (BD, San Jose, CA). CD4 T cells were then cultured for 3 days (for DNA methylation and RNA studies) or 6 days (for protein expression by flow cytometry) in X-VIVO 15 media (Lonza) and stimulated with antiCD3/CD28 coated beads (Dynabeads, ThermoFisher) at 1:1 ratio. Th17 polarization was performed with 12.5ng/mL IL-1b (Peprotech), 25ng/mL IL-6 (Peprotech), 25ng/mL IL-23 (Peprotech) and 1ng/mL TGFbeta (Peprotech) and 1ug/mL anti-IL4 antibody (Invitrogen). CD8 T cells were cultured for 3 days for all analyses as their viability can decrease in longer cultures. To promote their viability in vitro, we also supplemented the XVIVO-15 media with 1ng/mL IL7 (Peprotech) and 10ng/mL IL15 (Peprotech) (Montes *et al.*, 2005; Ghassemi *et al.*, 2016). T cytotoxic-17 (Tc17) polarization was performed with 12.5ng/mL IL-1b (Peprotech), 25ng/mL IL-6 (Peprotech), 25ng/mL IL-23 (Peprotech), 1ug/mL anti-IL4 antibody (Invitrogen) and 1ng/mL TGFbeta (Peprotech) in addition to the above IL7 and IL15. MMF (Sigma) was dissolved in DMSO and stored at -20oC. MMF at the specified concentration (either 20uM or 50uM) was added to the cultures twice daily for 3 days given its very short half-life and to mimic its in vivo administration schedule. CD4 T cells that were kept for 6 days in culture received MMF once daily along with half media changes every day from day 3 to day 5. Viability of cells was measured with eBioscience Fixable Viability Dye eFluor780. At the end of the third day, CD4 and CD8 cells were harvested and either genomic DNA or total RNA (including microRNA) was isolated with Allprep DNA/RNA micro kit (Qiagen) and miRNAeasy mini kit (Qiagen) respectively. On day 6 for CD4 and day 3 for CD8, cells were stained with anti-CCR6-BV421 (BD) and eBioscience Fixable Viability Dye eFluor780, fixed and permeabilized with the Foxp3/Transcription Factor Staining Buffer Set (ThermoFischer) and then stained with 1:100 anti-SMAD7-FITC (SantaCruz) and analyzed by flow cytometry.

EpiTYPER MassArray analysis

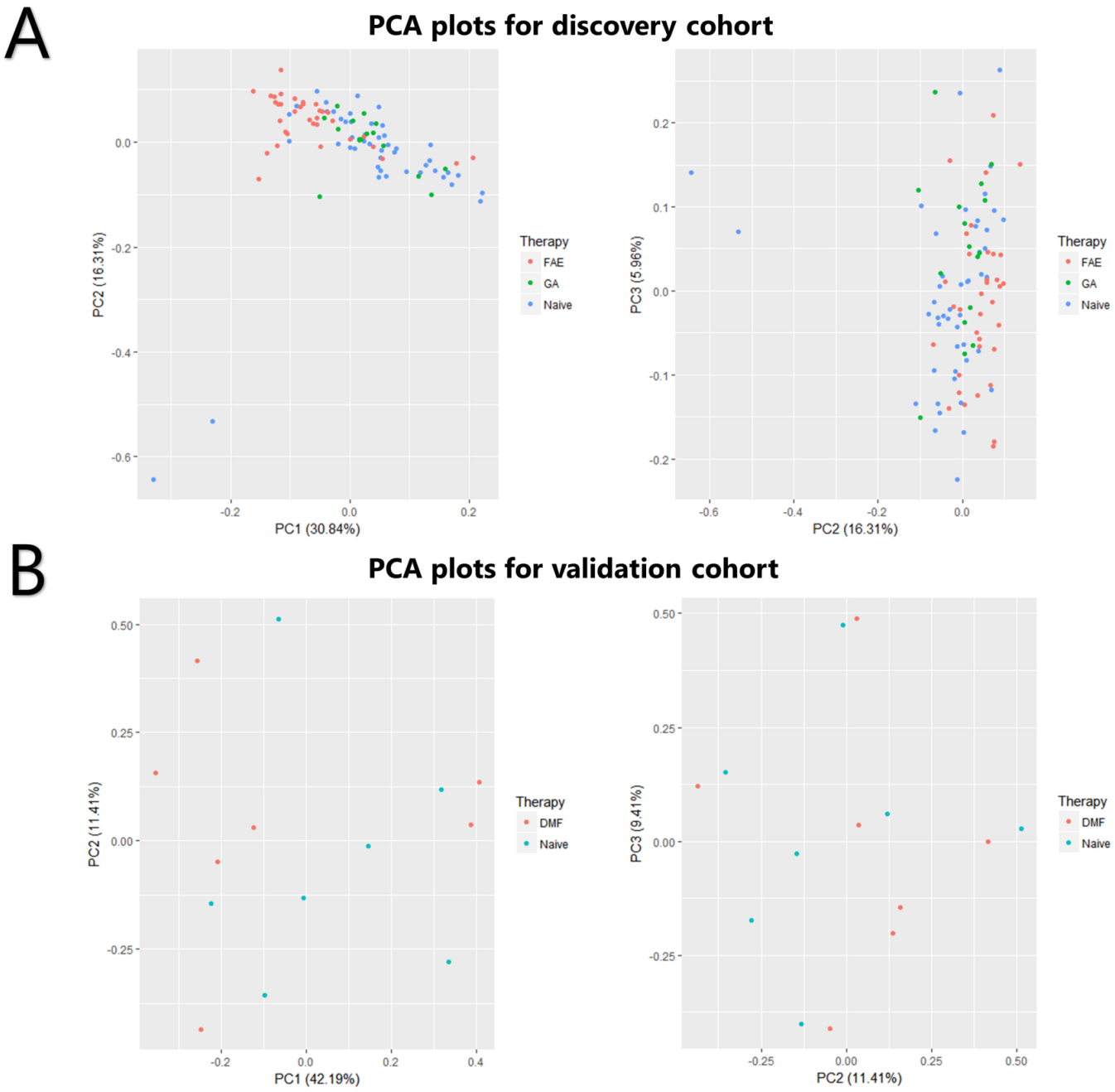
DNA methylation analysis of ex vivo and in vitro stimulated cells was performed with EpiTYPER MassArray system (Agena Bioscience) as previously described (Moyon *et al.*, 2016) at the Epigenetics Core facility at the CUNY Advanced Science Research Center (ASRC). Briefly, genomic DNA was bisulfite treated with Methylamp One-Step DNA Modification Kit (EpiGentek) and was used in a touchdown PCR reaction to amplify the *MIR-21* promoter and regular PCR reaction to amplify the TNF promoter (primers for *MIR-21* promoter with required tag (lower case); LEFT: aggaagagagGGATTGGTTTATTTGGGGATTT, RIGHT: cagtaatacgactcactatagggagaaggctACAAAATAATACAACCATAAAAATATCAC) (primers for TNF promoter with required tag (lower case); LEFT: aggaagagagGGGTAGTAGGGATAAGTTTGGGATA, RIGHT: cagtaatacgactcactatagggagaaggctCCTCCAAAACCTCCAAATATAAAAT). All PCR reactions were performed by using HotStarTaq DNA Polymerase kit (Qiagen) and included additional Mg for a total of 3mM concentration. After PCR amplification, all samples were run in agarose gels to confirm the presence of a single band before running them on EpiTYPER MassArray.



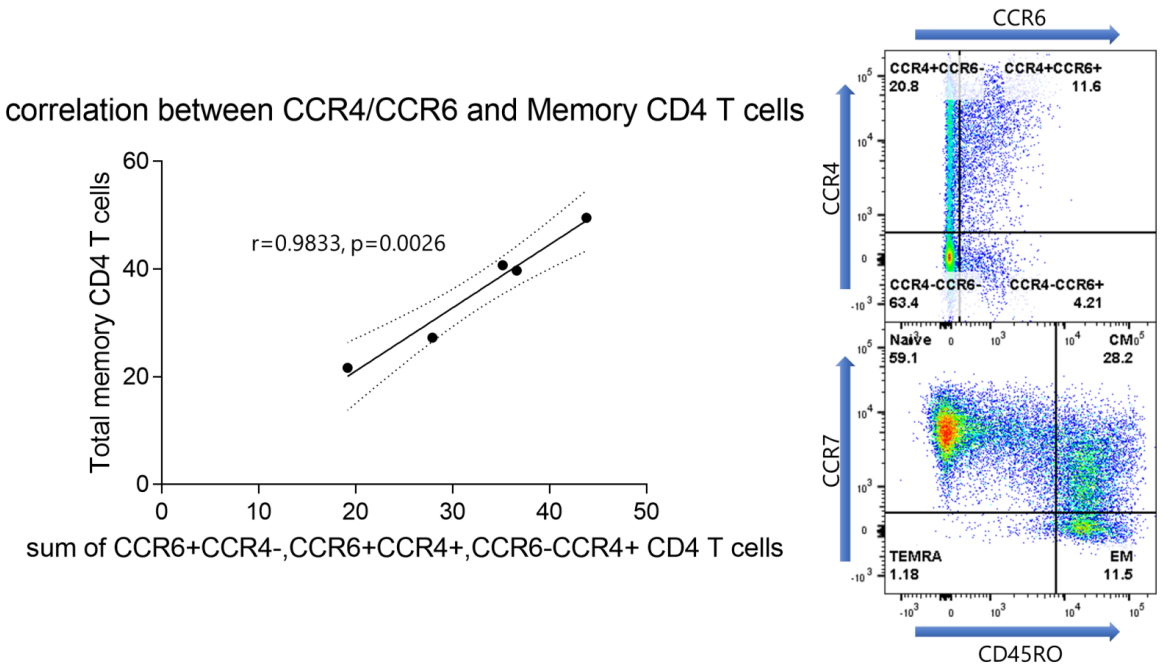
Supplementary Fig. S1. Study design flowchart: This research was approved by the Institutional Review Board (IRB) and informed consent was obtained for all subjects. Inclusion criteria for study enrollment were age 18-65, diagnosis of Relapsing Remitting MS (RRMS) or Clinically Isolated Syndrome by McDonald 2010 criteria (Polman *et al.*, 2011), being treatment naïve or being treated with either GA or FAEs and having stable disease for at least 3 months. Exclusion criteria were antibiotic use within 3 months, use of high dose corticosteroids within 1 month of enrollment, diagnosis of diabetes or inflammatory bowel disease, recent gastroenteritis, and treatment with an immunosuppressant medication for any condition in the 3 months preceding enrollment. PBMCs from 47 treatment naïve, 35 FAE and 16 GA-treated patients were collected for immunophenotyping and CD4 T cell DNA methylation analysis as a discovery cohort. CD4 T cells were also collected from 8 patients before and after FAE therapy as a validation cohort for our DNA methylation analysis, with one sample being excluded at quality control.



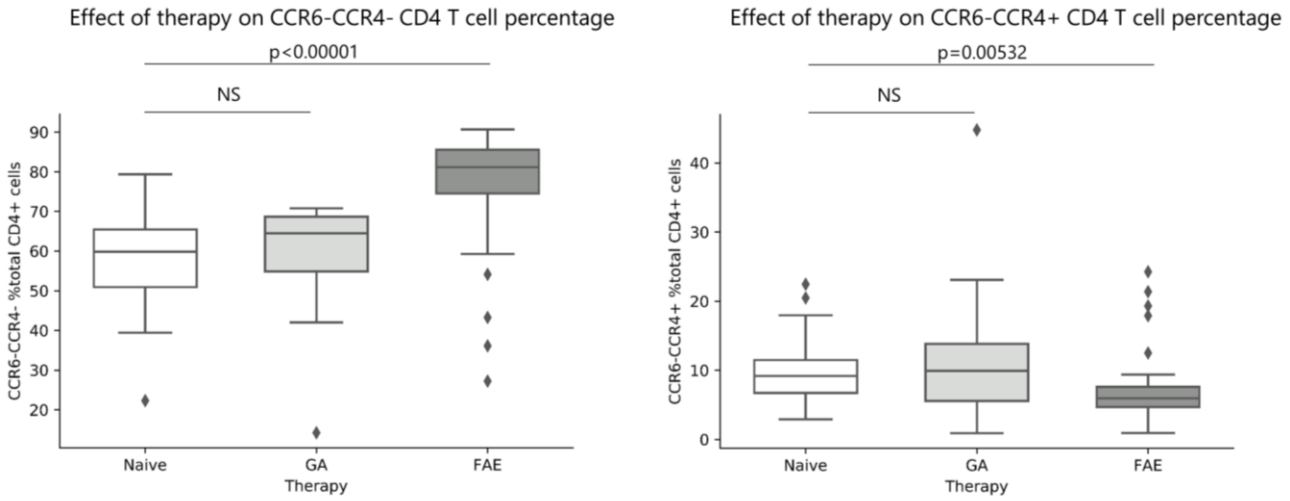
Supplementary Fig. S2. Gating strategy: Debris and doublets were excluded using light scatter measurements and major cell populations were identified based on their forward and side scatter properties. Subsequently, cells were gated using the corresponding cluster of differentiation or chemokine receptor targets and were represented as percentages of the parent population. Of note, SSC high CD3+ cells represent a minor subpopulation of neutrophils (Puellmann *et al.*, 2006) and, thus, were gated out from the analysis.



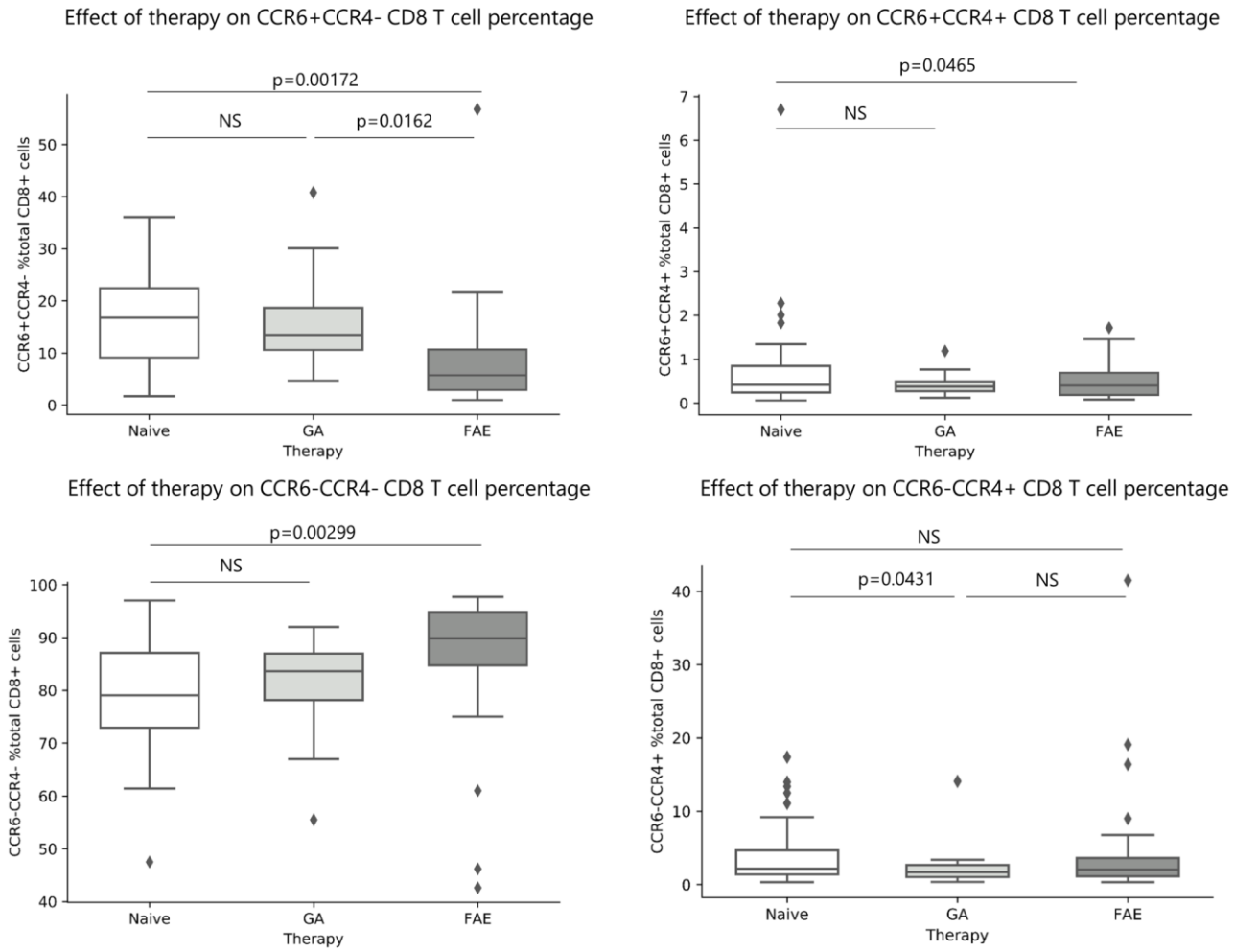
Supplementary Fig. S3. DNA methylation analysis PCA plots. (A): PCA analysis of our discovery cohort identified 2 outliers in the treatment-naïve group which were excluded from the rest of the DNA methylation analysis. **(B):** PCA analysis of our validation cohort did not identify any outliers.



Supplementary Fig. S4. Correlation between CCR6 and CCR4 stained CD4 T cells with memory T cell markers in healthy donors. Whole blood from 5 healthy donors was stained for CD3, CD4, CCR6, CCR4, CD45RO and CCR7 and analyzed by FACS. The sum of CCR6+CCR4+, CCR6+CCR4- and CCR6-CCR4+ CD4 T cell percentage was correlated to the total memory T cell population from each donor, defined as the sum of CD45RO+CCR7+ central memory (CM) and CD45RO+ CCR7- effector memory (EM). TEMRA: T cell effector memory CD45RA+

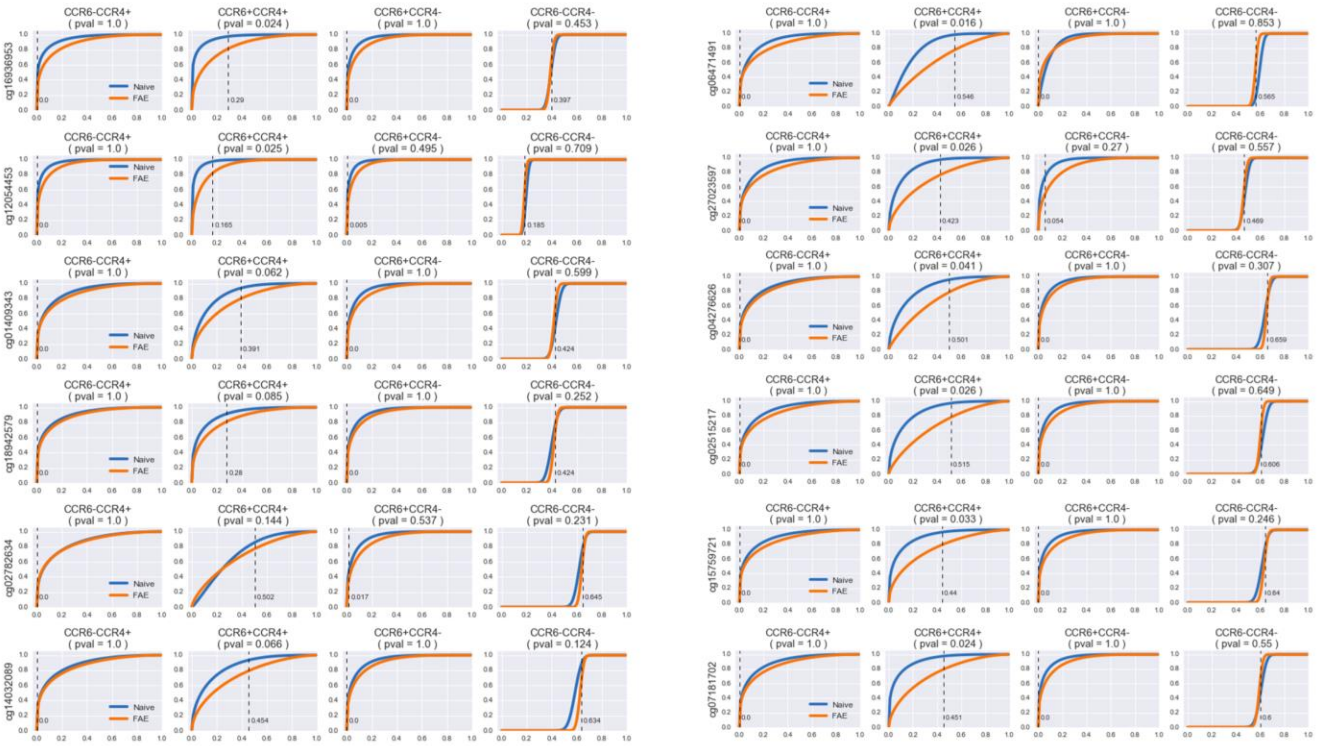


Supplementary Fig. S5 Effect of therapy on CCR6-CCR4- and CCR6-CCR4+ CD4 T cells. Whole blood from 47 treatment-naïve, 35 fumaric acid ester (FAE) and 16 glatiramer acetate (GA) treated multiple sclerosis (MS) patients was analyzed by FACS. FAE therapy was associated with mildly reduced CCR6-CCR4+ CD4 T cells (Th2) and increased CCR6-CCR4- (Naïve) percentage of CD4 T cells (linear regression controlling for age, gender, race and disease duration; Th2 t_{76}=-2.870, p=0.00532; Naïve t_{76}=6.171, p=0.0000000305). NS: not-significant (p>0.05). The lines in the box plot represent the quartiles of the dataset and the whiskers show the rest of the distribution, except for “outliers” that were determined by a function of the inter-quartile range based on the package seaborn in python.

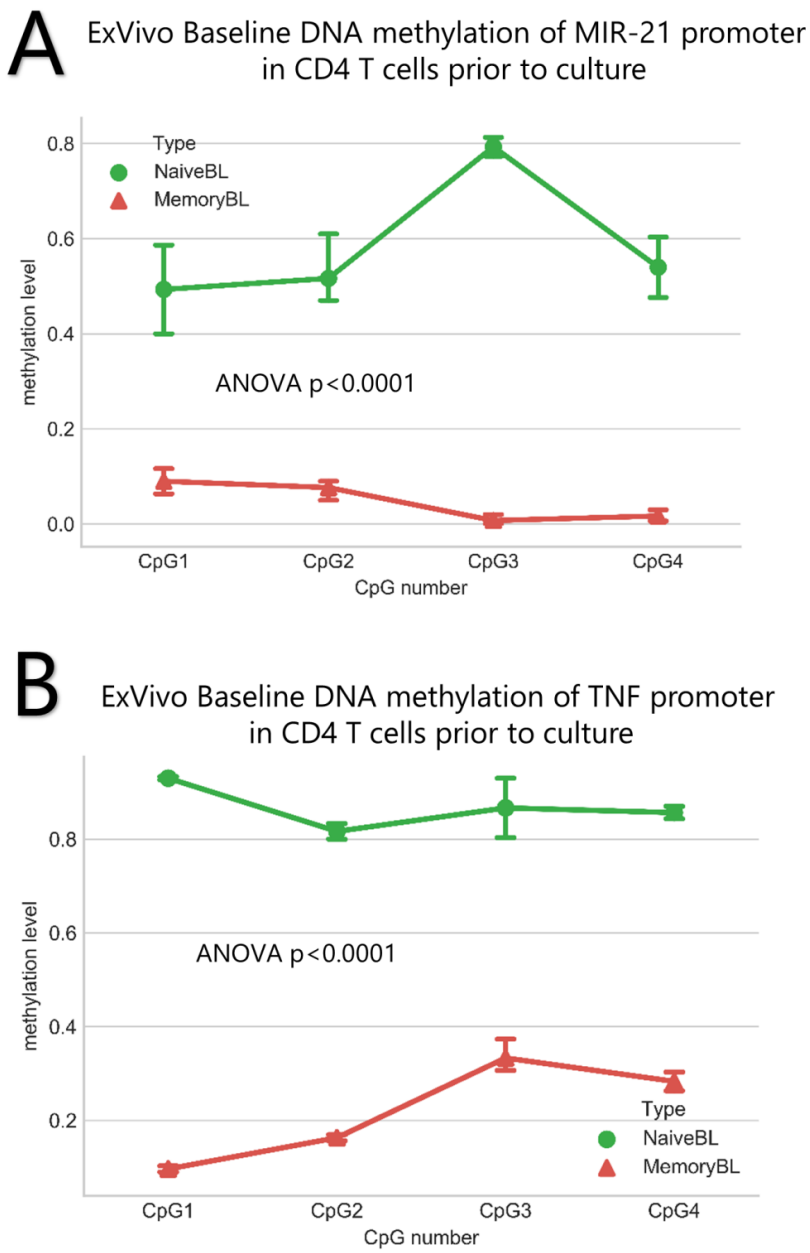


Supplementary Fig. S6 Effect of therapy on CCD6-CCR4+, CCR6+CCR4+, CCR6-CCR4- and CCR6+CCR4- CD8 T cells. Whole blood from 47 treatment-naïve, 35 fumaric acid ester (FAE) and 15 glatiramer acetate (GA) treated multiple sclerosis (MS) patients was analyzed by FACS. FAE therapy was associated with a significant reduction of CCR6+CCR4- CD8 T cells and an increased CCR6-CCR4- percentage of CD8 T cells (linear regression controlling for age, gender, race and disease duration; CCR6+CCR4- t_{76}=-3.249, p=0.00172; CCR6-CCR4- t_{76}=3.068, p=0.00299). CCR6+CCR4- CD8 T cells were lower in FAE-treated than GA-treated patients (linear regression controlling for age, gender, race and disease duration; CCR6+CCR4- t_{76}=-2.5, p=0.0162). Although CCR6+CCR4+ CD8 T cells were at very low abundance in both treatment naïve and treated patients, FAE treatment was associated with a mildly lower percentage of these cells (linear regression controlling for age, gender, race and disease duration; CCR6+CCR4+ t_{76}=-2.024,

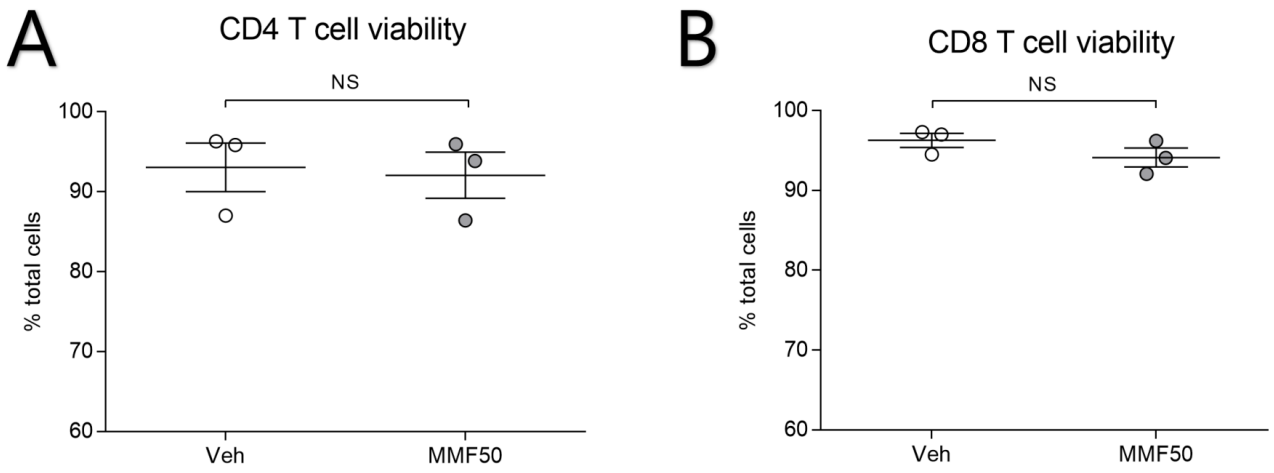
p=0.0465). Finally, CCR6-CCR4+ CD8 T cells did not change with FAE treatment (linear regression controlling for age, gender, race and disease duration; CCR6-CCR4+ t_{76}=-0.653, p=0.51579). GA had a minimal effect on CD8 T cells, with only a mild reduction of CCR6-CCR4+ CD8 T cells (linear regression controlling for age, gender, race and disease duration; CCR6-CCR4+ t_{56}=-2.070, p=0.0431). NS: not-significant (p>0.05). The lines in the box plot represent the quartiles of the dataset and the whiskers show the rest of the distribution, except for “outliers” that were determined by a function of the inter-quartile range based on the package seaborn in python.



Supplementary Fig. S7. Decomposition of DNA methylation values at the *MIR-21* locus. Decomposition of the β-values of the CpG sites in the *MIR-21* locus was done by utilizing a constrained least squares regression model, as the measured β-value of each sample is a linear combination of the β-values of each cell type and their corresponding proportions in the sample. Each row represents a different CpG site from the locus that was identified in our discovery and validation cohort. The cumulative distribution function of the null distribution (treatment naïve) of each CpG methylation value of the *MIR-21* DMR is shown in blue. The cumulative distribution function of the treatment distribution of each CpG methylation value of the *MIR-21* DMR is shown in orange. Each point on the cumulative distribution function represents the probability (y axis) that a given cell type has a methylation value below a certain β-value (x axis). The dotted line represents the methylation value of each CpG site of the *MIR-21* locus in each cell type that was obtained by fitting the model in all the FAE-treated patients. The p value was obtained from the cumulative distribution function of the null distribution from each cell type as the probability of obtaining a treatment naïve methylation value higher than the methylation value of each CpG site from FAE-treated patients. CCR6+CCR4+ CD4 T cells (Th17) exhibited significantly hypermethylated CpGs at the *MIR-21* promoter (right side) in FAE-treated patients compared to treatment naïve controls.



Supplementary Fig. S8. Ex vivo DNA methylation analysis from Naïve and Memory CD4 T cells prior to culture. Naïve (CD45RO-CCR7+) and memory (CD45RO+) CD4 T cells were isolated from human PBMCs by FACS. The baseline DNA methylation levels of naïve (NaïveBL) and memory CD4 T cells (MemoryBL) at the *MIR-21* and *TNF* promoters were measured ex vivo by EpiTYPER MassArray. **(A):** Naïve CD4 T cells had a hypermethylated *MIR-21* locus compared to memory CD4 T cells, whose methylation values were lower than 0.1 (two-way ANOVA; n=3, $F_{(1, 16)} = 225.1$; $p<0.0001$). **(B):** Similarly the TNF promoter was hypermethylated in naïve compared to memory CD4 T cells (two-way ANOVA; n=3, $F_{(1, 16)} = 812.5$, $p<0.0001$).



Supplementary Fig. S9. Naïve CD4 and CD8 T cell viability was not affected by FAE treatment in vitro. Naïve CD45RO-CCR7+ CD4 or CD8 T cells were sorted from human PBMCs and then activated in culture with antiCD3/CD28 coated beads under Th17 or Tc17 polarizing conditions with vehicle (Veh) or MMF 50uM (MMF50) added twice daily for 3 days. Viability was assessed by FACS after staining with eBioscience Fixable Viability Dye eFluor780. No difference in viable cells was detected in either CD4 or CD8 T cells (two-tailed t test; CD4: n=3, t_4=0.2346, p=0.8260; CD8: n=3, t_4=1.442, p=0.2228).