Basic and translational research

EXTENDED REPORT

DNA methylome signature in rheumatoid arthritis

Kazuhisa Nakano,[1] John W Whitaker,[2] David L Boyle,[1] Wei Wang,[2] Gary S Firestein[1]

[1]Department of Medicine, University of California San Diego School of Medicine, La Jolla, California, USA
[2]Department of Chemistry and Biochemistry, University of California San Diego School of Medicine, La Jolla, California, USA

Correspondence to
Dr Gary S Firestein, Division of Rheumatology, Allergy and Immunology, University of California San Diego School of Medicine, 9500 Gilman Drive, Mail code 0656, La Jolla, CA 92093-0656, USA; gfirestein@ucsd.edu

ABSTRACT

Objectives Epigenetics can influence disease susceptibility and severity. While DNA methylation of individual genes has been explored in autoimmunity, no unbiased systematic analyses have been reported. Therefore, a genome-wide evaluation of DNA methylation loci in fibroblast-like synoviocytes (FLS) isolated from the site of disease in rheumatoid arthritis (RA) was performed.

Methods Genomic DNA was isolated from six RA and five osteoarthritis (OA) FLS lines and evaluated using the Illumina HumanMethylation450 chip. Cluster analysis of data was performed and corrected using Benjamini–Hochberg adjustment for multiple comparisons. Methylation was confirmed by pyrosequencing and gene expression was determined by qPCR. Pathway analysis was performed using the Kyoto Encyclopedia of Genes and Genomes.

Results RA and control FLS segregated based on DNA methylation, with 1859 differentially methylated loci. Hypomethylated loci were identified in key genes relevant to RA, such as *CHI3L1*, *CASP1*, *STAT3*, *MAP3K5*, *MEFV* and *WISP3*. Hypermethylation was also observed, including *TGFBR2* and *FOXO1*. Hypomethylation of individual genes was associated with increased gene expression. Grouped analysis identified 207 hypermethylated or hypomethylated genes with multiple differentially methylated loci, including *COL1A1*, *MEFV* and *TNF*. Hypomethylation was increased in multiple pathways related to cell migration, including focal adhesion, cell adhesion, transendothelial migration and extracellular matrix interactions. Confirmatory studies with OA and normal FLS also demonstrated segregation of RA from control FLS based on methylation pattern.

Conclusions Differentially methylated genes could alter FLS gene expression and contribute to the pathogenesis of RA. DNA methylation of critical genes suggests that RA FLS are imprinted and implicate epigenetic contributions to inflammatory arthritis.

INTRODUCTION

Rheumatoid arthritis (RA) is an immune-mediated disease marked by symmetric inflammation in diarthrodial joints and destruction of the extracellular matrix. Genomics has rapidly advanced our understanding of susceptibility and severity of RA, and many associated polymorphisms in key genes have been described. However, identical twins have a concordance rate of 12%–15%[1][2] suggesting that epigenetic influences can affect either the onset or progression of disease. One of the most widely studied epigenetic mechanisms is DNA methylation, which plays a key role regulating gene expression and could potentially contribute to immune dysregulation.[3][4]

Fibroblast-like synoviocytes (FLS) form the synovial intimal lining of joints and play an integral role in RA by producing key cytokines, small molecule mediators and proteases.[5] Rheumatoid FLS exhibit a unique aggressive phenotype that could contribute to the cytokine milieu and joint destruction. Recent data in a severe combined immunodeficiency (SCID) mouse model show that RA synoviocytes can migrate between joints, thereby serving as a mechanism that spreads the RA phenotype and cartilage damage to distant sites.[6]

The mechanism of the aggressive rheumatoid phenotype is uncertain, although several studies implicate abnormal tumour suppressor gene structure and function.[7] An alternative explanation is that differential DNA methylation might alter gene expression and synoviocyte function. To address this question, we evaluated DNA methylation patterns in RA FLS and compared them with FLS derived from individuals with non-inflammatory joint disease. The data show that the RA FLS display a DNA methylome signature that distinguishes them from osteoarthritis (OA) and normal FLS, with differentially methylated (DM) genes that are critical to cell trafficking, inflammation and cell–extracellular matrix interactions.

METHODS
FLS and patient phenotype

FLS were isolated from synovial tissues obtained from female RA and OA patients at the time of joint replacement as described previously.[8] The mean ages of RA and OA patients were 53 ± 9 and 68 ± 16, respectively. Clinical information on four patients (two RA and two OA) was limited because the samples were de-identified. The erythrocyte sedimentation rates for the remaining RA and OA patients were 38 ± 15 and 19 ± 10, respectively. Of the four RA patients with clinical information, three were seropositive for serum rheumatoid factor and/or anti-cyclic citrullinated peptide antibody. Four were treated with low dose prednisone, two with methotrexate, two with a tumour necrosis factor (TNF) blocker and one with leflunomide. OA patients were treated with acetaminophin and narcotics for pain. Confirmatory studies were performed with an additional five RA, six OA and six normal FLS lines. Normal FLS were obtained from the San Diego Tissue Bank.

Isolation of genomic DNA and qPCR analysis

Fifth passage FLS were grown to 80% confluence and harvested. Genomic DNA of 10^6 FLS was isolated using the MagMAX DNA Multi-Sample Kit (Applied Biosystems, Foster City, California, USA). mRNA from cultured FLS was isolated using

RNA-STAT (Tel-test, Friendswood, Texas, USA). OD 260/280 was routinely ~2.0. cDNA was prepared from 1 μg of total RNA and qPCR was performed with TaqMan reagents (Life Technologies, Carlsbad, California, USA). Gene Ct values were normalised to glyceraldehyde 3-phosphate dehydrogenase expression using a standard curve methodology to obtain relative cell equivalents, as previously described.[9]

Infinium HumanMethylation450 BeadChip analysis

The Infinium HumanMethylation450 BeadChip was processed as described by the manufacturer (Illumina, San Diego, California, USA). This chip covers 96% of RefSeq genes and provides comprehensive gene region coverage, targeting multiple sites with promoter, 5′ UTR (untranslated region), 1st exon, gene body and 3′ UTR. The methylation level of a locus was measured as $\beta = M/(U+M+100)$ where M was the fluorescence level of the methylation probe and U was the methylation level of the unmethylated probe. A constant value of 100 was added as a pseudocount to prevent division by a small number (or 0) for background subtraction. The β values vary from 0 (no unmethylation) to 1 (100% methylation). To measure the difference in methylation at a locus between OA and RA, the average β levels were compared as described in the Results section.

Pyrosequencing

Sodium bisulphite modification of genomic DNA was performed with the Epitect bisulphite kit (Qiagen, Valencia, California, USA). The primers for PCR amplification and sequencing were designed using PyroMark (Qiagen) assay design SW 2.0. PCR was performed with biotinylated primers to convert the PCR product to single-stranded DNA templates. Pyrosequencing reactions and methylation quantification were performed by the Stanford Protein and Nucleic Acid Facility (http://pan.stanford.edu). The following loci were evaluated: TNF (chr6:31543300, 31543289), FERMT3 (chr11:63974123, 63974131, 63974162), ITGA4 (chr2:182321373, 182321354), MYEF2 (chr15:48470794, 48470754), SYNJ2 (chr6:158404121, 158404061) and CHI3L1 (chr1:203156625).

Enrichment of multiple methylated loci

DM loci were ranked by their individual corrected p values (see Results section for statistical methods). For each gene associated with loci where p was less than 0.05, relative methylation was calculated as its enrichment factor (EF).

$$EF = (\text{number of loci from gene A beneath cut-off}/ \\ \text{total loci from gene A})/(\text{total number of loci} \\ \text{beneath cut-off/total number of loci})$$

For enriched genes (EF>1), a p value for the level of enrichment was calculated using the hypergeometric distribution. The resulting p values were adjusted with the Benjamini–Hochberg correction. If a gene was enriched at multiple loci DM levels then only the level with the lowest enrichment p value was reported.

Pathway and GO analyses

Pathway enrichment analysis was carried out for genes with DM using the Kyoto Encyclopedia of Genes and Genomes (KEGG) human pathways and modules (www.genome.jp/kegg/download). The resulting p values were adjusted with the Benjamini–Hochberg correction. As the KEGG pathways represent groups of related bimolecular pathways, a p value cut-off of <0.1 was used as previously described.[10] [11] Gene ontology

(GO) analysis was carried using human GO term associations (www.geneontology.org). GO term enrichment analysis was carried out using model-based gene set analysis which uses probabilistic inference to identify the active GO terms.[12] Marginal probability values >0.50 were considered significant.

Evaluation of confirmatory cell lines

To assess the predicative capability of the original 11 RA/OA FLS lines on the new 11 RA/OA FLS lines, an support vector machine (SVM) model was trained on the original data using the 20 differentially methylated loci (DML) with the most significant p values and classified the new lines using R (http://www.R-project.org/). The SVM model used the default radial basis kernel with the following parameters: cost=8 and γ=1/64. The 1859 DML were then ranked based on the similarity of the two RA outliers to the average methylation frequency of the original six RA FLS lines and their dissimilarity to the average methylation frequency of the original five OA FLS lines. The DML were ranked and the top 20 DML were selected for hierarchical clustering. To assess the predictive capability of these 20 DML in all 28 RA/OA/normal FLS lines, an SVM model was trained using these 20 DML via twofold cross validation. Twofold cross validation was repeated 100 times to adequately sample various training and test dataset combinations. The SVM accuracy, sensitivity and specificity statistics were determined by counting the number of test cases where an RA sample was classified as RA/non-RA (ie, OA or normal) and the number of the cases where a non-RA sample was classified as RA/non-RA.

RESULTS

Initial probe and data filtering

The Infinium HumanMethylation450 BeadChip (Illumina, Inc.) was used to determine the methylation status of 485 512 loci from FLS isolated from 11 patients at the time of total joint replacement surgery (six RA; five OA). Loci were removed from subsequent analysis if any of the probes for a locus could not be disguised from background with a p value <0.01 or if not enough beads were present upon the chip for accurate measurement of their methylation level. After filtering, 476 331 loci were available for further analysis.

Global methylation status

To assess global methylation status of RA and OA FLS, the methylation scores over all filtered loci within a sample were summed. The difference between the two samples was assessed using Student t test. Initial analysis included all loci and was then repeated for only loci located in promoters. Overall, global hypermethylation or hypomethylation in RA FLS was not observed in the loci evaluated by the Illumina chip (p values 0.528 and 0.627, respectively). This observation was confirmed by evaluating LINE-1 methylation using pyrosequencing to quantify global methylation status.[13] LINE-1 methylation was 70.2±0.62% in the RA and 70.9±0.13% in OA FLS (p>0.10).

Analysis of DM individual loci

Although global methylation was not different for RA and OA in the array, DM loci clustered to the two populations. To identify the autosomal loci that are DM between OA and RA, two filters were used: (i) an average difference in methylation level was greater than 0.1 and (ii) a p value <0.05 calculated using the Student t test and corrected for multiple testing with the Benjamini–Hochberg adjustment. A total of 1859 loci in 1206 different genes (732 of which were hypomethylated and 1127 which were hypermethylated) were identified as significantly

Table 1 Number of hypermethylated and hypomethylated sites

	Hypomethylated in RA	Hypermethylated in RA
DM loci*	732	1127
DM genes†	63	144
DM KEGG pathways	6	10
DM GO terms	15	3

*Difference between OA and RA >0.1, with p<0.05 after Benjamini–Hochberg adjustment.
†Six additional genes had both hypomethylated and hypermethylated loci (ADAMTS2, C7orf50, DIP2C, FRMD4A, JARID2, SH3PXD2A).
DM, differentially methylated; GO, gene ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; OA, osteoarthritis; RA, rheumatoid arthritis.

DM in RA FLS (table 1). Examples of key individual loci are shown in table 2 and the complete list of 1859 loci is shown in supplementary online table S1. Many of the DM loci are associated with genes that have relevance to joint architecture and inflammation. For example, CHI3L1 is a cartilage specific antigen, also known as gp39, which is the target for autoimmunity in RA.

STAT3 is a key signalling protein that is activated in RA and is associated with interleukin (IL)-6 function. TNF signalling (TRAF2), IL-1 regulation (caspase 1) and purinoreceptors were also hypomethylated in RA.

Permutation analysis was carried out to assess the significance of loci identified as DM. The 11 samples were randomly assigned to OA and RA while maintaining the same number of OA and RA labels, that is, five OA and six RA. The permutation analysis was repeated 1000 times. The average number of significant loci during the permutation analysis was 4.9, compared with 1859 for the correct disease identification. The permutation analysis strongly supports these loci as DM and not a result of random chance.

To assess the ability of the 1859 loci to distinguish OA from RA, the methylation patterns at the loci across the 11 samples were hierarchically clustered by measuring the Euclidian distance between methylation levels across the loci (figure 1A). The clustering of the samples distinguishes OA from RA, which significantly segregate according to diagnosis (p<0.05). The clustering of loci also shows that groups of loci have

Table 2 Examples of genes with individual differentially methylated loci

	Description	CpG Chr coordinate	Gene region
Hypermethylated genes			
PHLPP1	PH domain and leucine rich repeat protein phosphatase 1	18:60386672	Body
RXRA	Retinoid X receptor, α	9:137243592	Body
CYFIP1	Cytoplasmic FMR1 interacting protein 1	5:22985092	Body
ADAMTS2	ADAM metallopeptidase with thrombospondin type 1 motif, 2	5:178712051	Body
CD55	CD55 molecule, decay accelerating factor for complement	1:207497486	Body
CABLES1	Cdk5 and Abl enzyme substrate 1	18:20717693	Body
HAS1	Hyaluronan synthase 1	19:52223620	Body
ITGB8	Integrin, β 8	7:20449718	3'UTR
HBEGF	Heparin-binding EGF-like growth factor	5:139722098	Body
PTPN14	Protein tyrosine phosphatase, non-receptor type 14	1:214588993	Body
COL4A2	Collagen, type IV, α 2	13:110979651	Body
COL4A1	Collagen, type IV, α 1	13:110892021	Body
TGFBR2	Transforming growth factor, β receptor II (70/80 kDa)	3:30684783	Body
EGF	Epidermal growth factor	4:110837358	Body
ITGBL1	Integrin, β-like 1 (with EGF-like repeat domains)	13:102109225	Body
TIMP2	TIMP metallopeptidase inhibitor 2	17:76871734	Body
MAP3K1	Mitogen-activated protein kinase kinase kinase 1	5:56115499	Body
FOXO1	Forkhead box O1	13:41223175	Body
Hypomethylated genes			
CHI3L1	Chitinase 3-like one (cartilage glycoprotein-39)	1:203156625	TSS
WISP3	WNT1 inducible signalling pathway protein 3	6:112375870	Body
STK24	Serine/threonine kinase 24	13:99201496	Body
MMP20	Matrix metallopeptidase 20	11:102473828	Body
EGFLAM	EGF-like, fibronectin type III and laminin G domain	5:38403758	Body
PTGIS	Prostaglandin I2 (prostacyclin) synthase	20:48183673	Body
DCBLD1	Discoidin, CUB and LCCL domain containing 1	6:117802710	TSS
MAP3K5	Mitogen-activated protein kinase kinase kinase 5 (ASK1)	6:136915088	Body
MGMT	O-6-methylguanine-DNA methyltransferase	10:131323903	Body
CDK14	Cyclin-dependent kinase 14	7:90794392	3'UTR
STAT3	Signal transducer and activator of transcription 3	17:40489721	Body
ADAM32	ADAM metallopeptidase domain 32	8:38965386	Body
RASGRF2	Ras protein-specific guanine nucleotide-releasing factor 2	5:80460495	Body
CASP1	Caspase 1, apoptosis-related cysteine peptidase	11:104916773	TSS
P2RX6	Purinergic receptor P2X, ligand-gated ion channel, 6	22:21381565	3'UTR
TRAF2	Tumour necrosis factor receptor-associated factor 2	9:139796499	Body
MEFV	Mediterranean fever	16:3307037	TSS

p value cut-off of <0.05 used for determining significance. If a gene had multiple significant loci then the coordinates and gene region information of the most significant locus given.
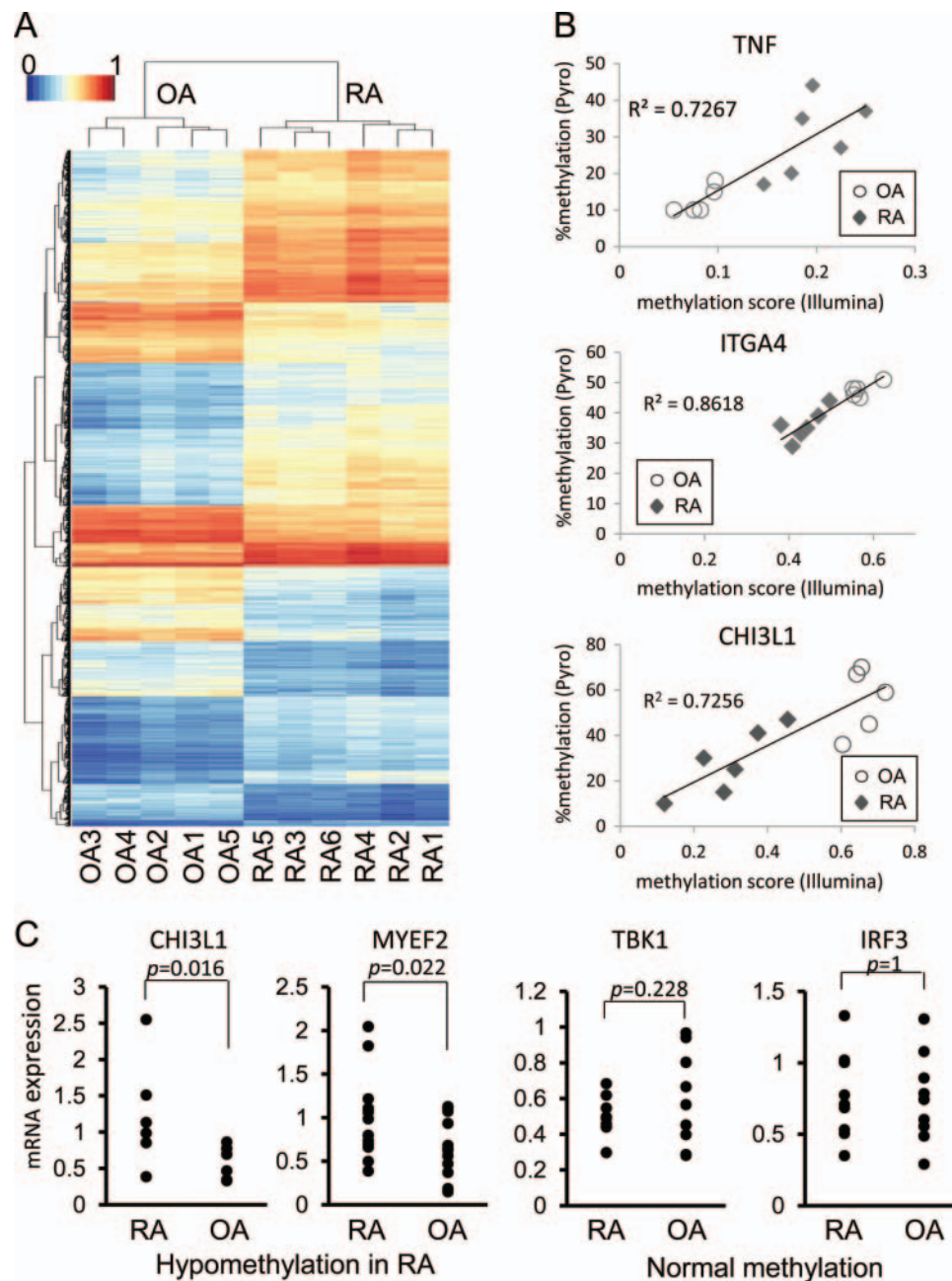
Figure 1 Hierarchical clustering and heatmap of differentially methylated (DM) loci. The methylation levels at the 1859 significantly DM loci were used for hierarchical clustering. (A) The clustering of the samples is shown by the dendrogram at the top and the clustering of the loci is shown by the dendrogram on the left. The methylation levels at the loci are shown in the heatmap. (B) Representative correlation in DM loci between the HumanMethylation450 BeadChip and pyrosequencing. DM loci of five genes (TNF; 2, FERMT3; 3, CHI3L1; 1, ITGA4; 2, and MYEF2; 2 loci) of the HumanMethylation450 BeadChip were validated by pyrosequencing (all 11 samples; six rheumatoid arthritis (RA), five osteoarthritis (OA)). The correlation is shown for a hypermethylated locus (*TNF*) and hypomethylated loci (*IGTA4 and CHI3L1*) (*TNF*, chr6:31543300; *ITGA4*, chr2:182321354; *CHI3L1*, chr1:203156625, GRCh37/hg19). The y-axis shows the percent methylated cytosines. The x-axis shows the mean methylation score from the Illumina chip. p Value <0.01 for all loci. (C) Gene expression of hypomethylated genes in RA fibroblast-like synoviocytes (FLS). Gene expression was determined by PCR in 6–13 separate OA and RA FLS lines for seven genes that were significantly hypomethylated (*CHI3L1, COL1A1, MYEF2, ITG4A, SYNJ2, STK24, MAP3K5*). Representative examples are shown for hypomethylated genes. Taken as a group, expression of hypomethylated genes in RA was significantly greater than OA (p<0.01). As a group, expression of genes that were normally methylated in RA was similar to OA (representative examples shown) (*AXIN, IKKE, TBK1, NANOG, POU5F1, MAP2K6, IRF3*).

similar patterns of DM across the samples. To confirm the Illumina chip data, we repeated the study using new aliquots of the original 11 FLS lines and obtained similar results (data not shown). In addition, pyrosequencing was used to confirm 10 DM loci of five genes (TNF, FERMT3, CHI3L1, ITGA4 and MYEF2) and validate the chip methodology. Results obtained

for all 10 loci were in accordance with results obtained by methylation array with average Spearman's rank correlation coefficient $r_s = 0.90 \pm 0.02$ (mean±SEM, p<0.01) (figure 1B for representative examples). Therefore, the chip results correlate highly with pyrosequencing and minimise the need for additional sequencing to confirm data in additional DM loci.

Table 3 Representative genes with more than one differentially methylated loci

RA status	Gene	Differentially methylated loci	Total loci on array	p Value for enrichment	Enrichment factor
Hypo	RHOJ	2	9	2.16E-05	525.33
Hypo	MEFV	2	11	9.64E-05	200.66
Hypo	CD93	3	15	2.44E-06	200.26
Hypo	MYEF2	2	13	0.0001139	169.79
Hypo	ITGA4	3	18	4.52E-06	157.79
Hypo	P2RY6	2	21	0.0001567	142.51
Hypo	AXIN2	2	37	0.0002877	102.35
Hypo	COL22A1	3	36	2.26E-05	83.44
Hypo	COL1A1	4	41	1.26E-06	79.55
Hypo	COL6A1	2	33	0.0006962	60.68
Hypo	ST14	2	43	0.0009217	51.33
Hypo	SYNJ2	2	68	0.0026518	27.84
Hyper	IGSF6	3	6	1.21E-06	275.53
Hyper	EGF	2	13	0.0003606	107.08
Hyper	ITGB8	2	13	0.0005041	84.78
Hyper	TNF	4	27	2.66E-06	71.05
Hyper	AKT3	2	30	0.0008412	63.26
Hyper	MAP3K1	2	24	0.0013940	45.92
Hyper	PLXNA2	3	45	0.0001585	43.66
Hyper	TRERF1	4	50	1.49E-05	41.35
Hyper	FOXO1	2	32	0.0020262	36.94
Hyper	TIMP2	4	64	5.67E-05	28.37
Hyper	COL4A1	2	71	0.0035030	26.73
Hyper	COL4A2	5	139	1.53E-05	22.99
Hyper	SRPK2	2	45	0.0056023	20.18
Hyper	ADAM12	2	52	0.0057208	19.88
Hyper	PTPN14	2	59	0.0080113	16.26
Hyper	DNMT3A	2	77	0.0088166	15.35
Hyper	CREB5	2	82	0.0142486	11.7
Hyper	PRKCZ	2	263	0.0332803	7.22

A p value cut-off of <0.05 was used for determining significance.
Hyper, hypermethylated; Hypo, hypomethylated; RA, rheumatoid arthritis.

Analysis of genes with multiple DM loci

We then identified the genes with multiple DM loci (see Methods section). In all, 207 genes were enriched for multiple hypermethylated or hypomethylated loci (table 1). Representative examples of DM genes are shown in table 3, and the complete lists of genes are provided in supplementary online table S2. Of interest, several genes associated with matrix regulation are affected, including several collagen proteins, proteases (ADAM12) and protease inhibitors (TIMP2).

Correlation of gene expression and methylation status

Functional analysis of the DM was then performed by evaluating expression of 14 separate genes with hypomethylated and normally methylated loci in RA FLS by qPCR (see figure 1C for representative examples). Although there was individual variation, expression of genes with hypomethylated loci was significantly greater in RA FLS than in OA FLS (n = 7 genes; p<0.01). On the other hand, expression of normally methylated genes was similar in RA and OA FLS (n=7 genes; p>0.10). Therefore, expression levels of a broad panel of genes correlated with methylation status.

Pathway and GO analyses

The biological pathways and gene ontologies that are enriched with DM loci were examined. Loci were mapped to pathways via their relationships to genes within KEGG pathways and for GO using all 1859 DM loci identified (see table 1). As shown in

table 4, pathway analysis showed that interactions between cells and the matrix and cell recruitment were especially prominent in the DM group, including focal adhesion, cell adhesion molecules, leukocyte transendothelial interactions and ECM–receptor interactions. Other pathways, such as mTOR signalling and adipokine signalling, were also DM in the RA cells. Examples of genes in the 'Focal adhesion' pathway, which are enriched with hypomethylated loci in RA, include ITGA, ITGB, Actinin, RTK, Parvin, DOCK1 and Bcl-2. Supplementary online table S3 shows a complete list of the DM pathways. The complete list of differential GO terms is shown in supplementary online table S4. Of interest, the number of GO terms associated with hypomethylated DNA in RA was greater than for hypermethylated DNA (14 vs 3) and included the PcG protein complex, which is associated with remodelling chromatin to permit epigenetic gene silencing. Thus, hypomethylated terms clustered in RA especially related to membrane and transcription factor biology.

Confirmatory cell lines

Methylation was then determined with 11 new FLS lines (five RA and six OA). All 22 RA/OA lines were hierarchically clustered using the 1859 DML from the original dataset to determine if the methylation patterns of the new lines matched the original dataset. Hierarchical clustering of three of five new RA and six of six new OA cell lines conformed to the RA/OA signatures found in the original data. There were no distinguishing

Table 4 Examples of gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis of hypomethylated and hypermethylated genes in rheumatoid arthritis (RA)

Analysis type	Marginal probability (GO)	p value (KEGG)	Name
RA hypomethylated			
GO	0.998		Transcription coactivator activity
GO	0.985		ATPase binding
GO	0.956		PcG protein complex
GO	0.922		Positive regulation of transcription from RNA polymerase II promoter
GO	0.820		Extrinsic to plasma membrane
GO	0.811		Transcription factor binding
GO	0.788		G-protein activated inward rectifier potassium channel activity
GO	0.762		Regulation of exocytosis
GO	0.710		Nuclear hormone receptor binding
GO	0.592		Ubiquitin thiolesterase activity
GO	0.554		Phospholipase C activity
GO	0.536		External side of plasma membrane
GO	0.523		Actin cytoskeleton
GO	0.511		Biological adhesion
KEGG		0.007	Focal adhesion
KEGG		0.007	Glycosphingolipid biosynthesis—lacto and neolacto series
KEGG		0.007	ECM–receptor interaction
KEGG		0.014	Leukocyte transendothelial migration
KEGG		0.034	Protein digestion and absorption
KEGG		0.044	Adipocytokine signalling pathway
KEGG		0.062	Cell adhesion molecules
KEGG		0.085	Nitrogen metabolism
RA hypermethylated			
GO	0.771		Golgi apparatus
GO	0.766		Cell surface receptor linked signalling pathway
GO	0.603		Binding
KEGG		0.058	ErbB signalling pathway
KEGG		0.095	Fc ε RI signalling pathway
KEGG		0.095	mTOR signalling pathway
KEGG		0.095	Adipocytokine signalling pathway

The significance of GO terms and KEGG pathways were measured using marginal probabilities and p values, respectively. GO terms and KEGG pathways were shown if they had a marginal probability >0.5 or p values <0.1, respectively. GO terms and KEGG pathways which are irrelevant to RA were omitted from the table but are included in the supplementary information.

clinical features in the two RA FLS that did not conform with the other nine. An SVM model trained on the initial 11 RA/OA FLS lines using the 20 most significant DM loci was able to predict the diagnosis of the 11 new RA/OA lines with 100% accuracy (100% sensitivity and 100% specificity). To identify a DM locus subset that distinguishes all 11 RA lines via hierarchical clustering, the original 1859 DM loci were then ranked based on the similarity of the two non-conforming RA FLS lines to the original six RA lines to expand the definition of the RA signature. The top 20 DM loci from this new ranking (see Supplementary online table S5, including loci in key genes like ANGPT1, TRAF2, MMP20, and STK24) distinctly clustered all 11 RA lines from all 11 OA, including the six new OA lines and three new RA lines that were not included during DM loci ranking (see figure 2). Further confirmation was obtained with six additional normal FLS lines, which clustered with OA and did not overlap with RA (see figure 2). Using these 20 DM loci and twofold cross validation on all 28 FLS lines (RA/OA/normal), the SVM model predicted RA with 100% accuracy (100% sensitivity and 100% specificity).

DISCUSSION

FLS contribute to the pathogenesis of RA by producing inflammatory mediators and contributing to cartilage damage.[14–16] Understanding the molecular mechanisms that regulate FLS

activation could provide insights into the pathogenesis of RA and lead to novel therapeutic strategies. In the present study, we evaluated the epigenetic profile of RA FLS by exploring DNA methylation patterns in an unbiased fashion. A novel methylome signature was identified, especially involving pathways relevant to cell movement, adhesion and trafficking. A simplified algorithm identifying a limited subset of DM loci that distinguishes RA from OA was developed and confirmed with new FLS lines from patients with arthritis as well as normal individuals. These data strongly support the notion that RA FLS methylation patterns are distinct from control and normal FLS.

Rheumatoid FLS exhibit a unique aggressive phenotype that contributes to the cytokine milieu and joint destruction.[17] Functional studies suggest that RA cells are imprinted in situ and maintain these features after many passages in tissue culture. For example, RA FLS, unlike OA or normal synoviocytes, adhere to and invade cartilage explants in SCID mice.[18] RA FLS can grow under anchorage-independent conditions, are less susceptible to contact inhibition and resistant to apoptosis.[19 20] RA synoviocytes can potentially 'metastasize' and transfer the invasive phenotype from one joint to another.[6]

Several mechanisms have been implicated in the aggressive phenotype, including abnormal sentrin or phosphatase and tensin homologue expression, preferential shunting of stressed cells to DNA repair rather than programmed cell death,[21–23]
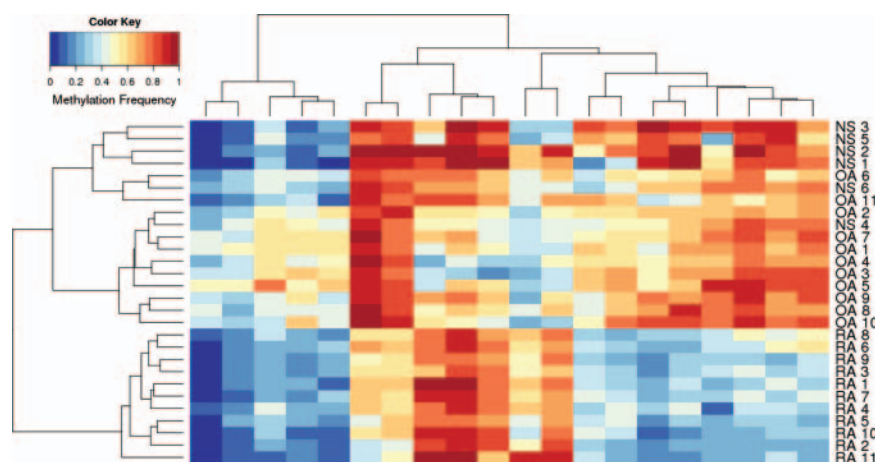
Figure 2 Hierarchical clustering and heatmap of differentially methylated (DM) loci. Hierarchical clustering was performed using the top 20 DM loci as described in the text. Note that the osteoarthritis (OA) fibroblast-like synoviocytes (FLS) lines were completely segregated from rheumatoid arthritis (RA) using these loci, including three new RA and six new OA FLS lines that were independent of the algorithm. In addition, six new normal FLS lines clustered with OA. See supplementary online table S5 for identification of the loci used in the model.

and somatic mutations of regulatory genes.[24–29] Micrcosatellite instability has also been identified in RA synovium, in part due to decreased DNA repair.[30] These genetic modifications potentially affect the natural history of disease.

Epigenetics can also profoundly influence cell activation and gene expression through a variety of mechanisms, including DNA methylation, histone modification and microRNA production.[31] Normal ontogeny relies on a carefully orchestrated sequence of DNA methylation to repress regulatory genes by methylating cytosine in CpG loci, either in promoters or in genes themselves.[32] Methylation abnormalities have been associated with many diseases, most notably cancer where expression or suppression of key genes allows cells to escape normal homeostatic controls.[33] Hypomethylation and hypermethylation are associated with many malignancies and can contribute to transformation.[34]

DNA methyltransferases (DNMTs) are responsible for initiating and maintaining CpG methylation in the human genome by converting cytosine to methylcytosine.[35] In mammalian cells, DNMT1, DNMT3a and DNMT3b are the primary enzymes responsible for CpG methylation. DNMT3a and DNMT3b mainly regulate de novo methylation while DNMT1 maintains methylation, especially during cell division. DNMT expression and DNA methylation are not immutable but are influenced by the environment and modify gene expression throughout life and even in progeny. For instance, pregnant mice fed a diet rich in methyl donors give birth to pups with increased levels of DNA methylation and increased airway reactivity in murine asthma for at least two subsequent generations.[36 37]

Global hypomethylation has been described in peripheral blood mononuclear cells of patients with RA, although the specific genes involved are not known.[38] Using chip technology to evaluate methylation of 809 CpG sites, no differences were observed in unfractionated peripheral blood mononuclear cells of twins disconcordant for RA.[39] We used an alternative approach by focusing on a homogeneous population of cells from the site of disease and chip technology that evaluates many more loci (>450 000 methylation sites). Modest global hypomethylation was also reported in cultured RA FLS when compared with OA cells, and treating RA FLS with a DNMT inhibitor decreased methylation further.[40] Our studies suggest

that global methylation levels are similar in OA and RA FLS for the specific loci assayed.

The genome-wide analysis identified distinct methylation profiles of OA and RA FLS involving 1859 loci located in 1206 genes. Cluster analysis showed that the two types of FLS could be distinguished based solely on methylation. Additional analysis identified 207 genes with multiple hypermethylated or hypomethylated loci. Many of these genes play a key role in inflammation, matrix regulation, leukocyte recruitment and immune responses and could represent novel therapeutic targets. Gene expression levels correlated with methylation status, with high expression in hypomethylated genes in RA FLS and normal expression in genes that were not DM. General concordance between methylation and mRNA levels was observed despite the large number of factors that can influence gene expression. The differences are especially interesting in light of the fact that most loci are in the body of genes rather than the promoter regions, as previously described.[41]

Analysis of DM KEGG pathways revealed a preponderance of DM genes related to cell adhesion, movement and trafficking. The traditional view of synoviocytes as relatively static cells that remain in the synovial intimal lining requires considerable revision. Instead, they are a dynamic population that move within the synovium to the surface of cartilage, disengage from the intimal lining to populate the synovial fluid compartment and mediate destruction. Synoviocytes can also release from the extracellular matrix and migrate to other joints. The combination of functional observations related to the 'spread' of RA via migrating FLS and imprinting cell movement via the methylome provides evidence that targeting genes in these KEGG pathways could be effective in RA.

These data provide evidence that epigenetic changes occur in RA synoviocytes and that they persist in culture. Imprinting could potentially occur before clinical disease and contribute to susceptibility. Alternatively, the changes could be induced after initiation of synovitis. The inflammatory milieu could potentially contribute to synoviocyte imprinting and affect their function. Thus, local inflammation could alter the expression of enzymes responsible for initiating and maintaining DNA methylation. Additional studies on cells from patients with high risk for developing RA would be needed to resolve this question. Experiments to evaluate the effects of the

inflammatory milieu on the methylome and the stability of differential methylation are also needed.

The ability to distinguish RA FLS based on DNA methylation could provide insights into how synoviocytes are altered and contribute to the pathogenesis of this disease. Evidence of RA phenotypic stratification based on the transcriptome also supports this concept,[42] and epigenetic imprinting is one potential mechanism. Therefore, defining the methylome in FLS offers a new way to explore these differences and identify novel therapeutic targets.

REFERENCES

1. **Silman AJ,** MacGregor AJ, Thomson W, et al. Twin concordance rates for rheumatoid arthritis: results from a nationwide study. *Br J Rheumatol* 1993;**32**:903–7.
2. **Jarvinen P,** Aho K. Twin studies in rheumatic diseases. *Semin Arthritis Rheum* 1994;**24**:19–28.
3. **Issa JP,** Ottaviano YL, Celano P, et al. Methylation of the oestrogen receptor CpG island links ageing and neoplasia in human colon. *Nat Genet* 1994;**7**:536–40.
4. **Esteller M.** CpG island hypermethylation and tumor suppressor genes: a booming present, a brighter future. *Oncogene* 2002;**21**:5427–40.
5. **Bartok B,** Firestein GS. Fibroblast-like synoviocytes: key effector cells in rheumatoid arthritis. *Immunol Rev* 2010;**233**:233–55.
6. **Lefevre S,** Knedla A, Tennie C, et al. Synovial fibroblasts spread rheumatoid arthritis to unaffected joints. *Nat Med* 2009;**15**:1414–20.
7. **Tak PP,** Zvaifler NJ, Green DR, et al. Rheumatoid arthritis and p53: how oxidative stress might alter the course of inflammatory diseases. *Immunol Today* 2000;**21**:78–82.
8. **Alvaro-Gracia JM,** Zvaifler NJ, Firestein GS. Cytokines in chronic inflammatory arthritis. V. Mutual antagonism between interferon-gamma and tumor necrosis factor-alpha on HLA-DR expression, proliferation, collagenase production, and granulocyte macrophage colony-stimulating factor production by rheumatoid arthritis synoviocytes. *J Clin Invest* 1990;**86**:1790–8.
9. **Boyle D,** Rosengren S, Bugbee W, et al. Quantitative biomarker analysis of synovial gene expression by real-time PCR. *Arthritis Res Ther* 2003;**5**:R352–60.
10. **Xu T,** Ping J, Yu Y, et al. Revealing parasite influence in metabolic pathways in Apicomplexa infected patients. *BMC Bioinform* 2010;**11**:S13.
11. **Shen Z,** Want EJ, Chen W, et al. Sepsis plasma protein profiling with immunodepletion, three-dimensional liquid chromatography tandem mass spectrometry, and spectrum counting. *J Proteome Res* 2006;**5**:3154–60.
12. **Bauer S,** Gagneur J, Robinson PN. GOing Bayesian: model-based gene set analysis of genome-scale data. *Nucleic Acids Res* 2010;**38**:3523–32.
13. **Choi I-S,** Estecio MRH, Nagano Y, et al. Hypomethylation of LINE-1 and Alu in well-differentiated neuroendocrine tumors (pancreatic endocrine tumors and carcinoid tumors). *Mod Pathol* 2007;**20**:802–10.
14. **Firestein GS.** Evolving concepts of rheumatoid arthritis. *Nature* 2003;**423**:356–61.
15. **Lee DM,** Kiener HP, Agarwal SK, et al. Cadherin-11 in synovial lining formation and pathology in arthritis. *Science* 2007;**315**:1006–10.
16. **Kiener HP,** Niederreiter B, Lee DM, et al. Cadherin 11 promotes invasive behavior of fibroblast-like synoviocytes. *Arthritis Rheum* 2009;**60**:1305–10.
17. **Firestein GS.** Invasive fibroblast-like synoviocytes in rheumatoid arthritis. Passive responders or transformed aggressors? *Arthritis Rheum* 1996;**39**:1781–90.
18. **Muller-Ladner U,** Kriegsmann J, Franklin BN, et al. Synovial fibroblasts of patients with rheumatoid arthritis attach to and invade normal human cartilage when engrafted into SCID mice. *Am J Pathol* 1996;**149**:1607–15.
19. **Lafyatis R,** Remmers EF, Roberts AB, et al. Anchorage-independent growth of synoviocytes from arthritic and normal joints. Stimulation by exogenous platelet-derived growth factor and inhibition by transforming growth factor-beta and retinoids. *J Clin Invest* 1989;**83**:1267–76.
20. **Baier A,** Meineckel I, Gay S, et al. Apoptosis in rheumatoid arthritis. *Curr Opin Rheumatol* 2003;**15**:274–9.
21. **Pap T,** Franz JK, Hummel KM, et al. Activation of synovial fibroblasts in rheumatoid arthritis: lack of expression of the tumour suppressor PTEN at sites of invasive growth and destruction. *Arthritis Res* 2000;**2**:59–64.
22. **Franz JK,** Pap T, Hummel KM, et al. Expression of sentrin, a novel antiapoptotic molecule, at sites of synovial invasion in rheumatoid arthritis. *Arthritis Rheum* 2000;**43**:599–607.
23. **You X,** Boyle D, Hammaker D, et al. PUMA-mediated apoptosis in fibroblast-like synoviocytes does not require p53. *Arthritis Res Ther* 2006;**8**:R157.
24. **Firestein GS,** Echeverri F, Yeo M, et al. Somatic mutations in the p53 tumor suppressor gene in rheumatoid arthritis synovium. *Proc Natl Acad Sci* 1997;**94**:10895–900.
25. **Igarashi H,** Hashimoto J, Tomita T, et al. TP53 mutations coincide with the ectopic expression of activation-induced cytidine deaminase in the fibroblast-like synoviocytes derived from a fraction of patients with rheumatoid arthritis. *Clin Exp Immunol* 2010;**161**:71–80.
26. **Inazuka M,** Tahira T, Horiuchi T, et al. Analysis of p53 tumour suppressor gene somatic mutations in rheumatoid arthritis synovium. *Rheumatology* 2000;**39**:262–6.
27. **RÈme T,** Travaglio A, Gueydon E, et al. Mutations of the p53 tumour suppressor gene in erosive rheumatoid synovial tissue. *Clin Exp Immunol* 1998;**111**:353–8.
28. **Bang H,** Egerer K, Gauliard A, et al. Mutation and citrullination modifies vimentin to a novel autoantigen for rheumatoid arthritis. *Arthritis Rheum* 2007;**56**:2503–11.
29. **Da Sylva T,** Connor A, Mburu Y, et al. Somatic mutations in the mitochondria of rheumatoid arthritis synoviocytes. *Arthritis Res Ther* 2005;**7**:R844–51.
30. **Lee S-H,** Chang DK, Goel A, et al. Microsatellite instability and suppressed DNA repair enzyme expression in rheumatoid arthritis. *J Immunol* 2003;**170**:2214–20.
31. **Karouzakis E,** Gay RE, Gay S, et al. Epigenetic control in rheumatoid arthritis synovial fibroblasts. *Nat Rev Rheumatol* 2009;**5**:266–72.
32. **Christophersen NSy,** Helin K. Epigenetic control of embryonic stem cell fate. *J Exp Med* 2010;**207**:2287–95.
33. **Kulis M,** Esteller M. DNA methylation and cancer. *Adv Genet* 2010;**70**:27–56.
34. **Ren J,** Singh BN, Huang Q, et al. DNA hypermethylation as a chemotherapy target. *Cell Signal* 2011;**23**:1082–93.
35. **Turek-Plewa J,** Jagodzinski PP. The role of mammalian DNA methyltransferases in the regulation of gene expression. *Cell Mol Biol Lett* 2005;**10**:631–47.
36. **Hollingsworth JW,** Maruoka S, Boon K, et al. In utero supplementation with methyl donors enhances allergic airway disease in mice. *J Clin Invest* 2008;**118**:3462–9.
37. **Miller RL.** Prenatal maternal diet affects asthma risk in offspring. *J Clin Invest* 2008;**118**:3265–8.
38. **Liu CC,** Fang TJ, Ou TT, et al. Global DNA methylation, DNMT1, and MBD2 in patients with rheumatoid arthritis. *Immunol Lett* 2011;**135**:96–9.
39. **Javierre BM,** Fernandez AF, Richter J, et al. Changes in the pattern of DNA methylation associate with twin discordance in systemic lupus erythematosus. *Genome Res* 2010;**20**:170–9.
40. **Karouzakis E,** Gay RE, Michel BA, et al. DNA hypomethylation in rheumatoid arthritis synovial fibroblasts. *Arthritis Rheum* 2009;**60**:3613–22.
41. **Lister R,** Pelizzola M, Dowen RH, et al. Human DNA methylome at base resolution show widespread epigenomic differences. *Nature* 2009;**462**:315–22.
42. **van Baarsen L,** Bos C, Pouw Kraan T, et al. Transcription profiling of rheumatic diseases. *Arthritis Res Ther* 2009;**11**:207.

ARD

DNA methylome signature in rheumatoid arthritis

Kazuhisa Nakano, John W Whitaker, David L Boyle, Wei Wang and Gary S Firestein

Updated information and services can be found at:

http://ard.bmj.com/content/72/1/110

These include:

References

This article cites 42 articles, 8 of which you can access for free at:

http://ard.bmj.com/content/72/1/110#BIBL

Email alerting service

Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Topic Collections

Articles on similar topics can be found in the following collections

Notes