Mecanismos de desregulación de la metilación del DNA y de micrornas en células implicadas en la patogénesis de la artritis reumatoide

Lorenzo de la Rica Lázaro

MECANISMOS DE DESREGULACIÓN DE LA METILACIÓN DEL DNA Y DE MICRORNAS EN CÉLULAS IMPLICADAS EN LA PATOGÉNESIS DE LA ARTRITIS REUMATOIDE

Memoria presentada por Lorenzo de la Rica Lázaro para optar al grado de doctor por la Universitat de Barcelona.

UNIVERSITAT DE BARCELONA - FACULTAT DE BIOLOGÍA

PROGRAMA DE DOCTORAT EN BIOMEDICINA 2013

Este trabajo ha sido realizado en el Grupo de Cromatina y Enfermedad del Programa de Epigenética y Biología de Cáncer (PEBC) del Institut d'Investigació Biomèdica de Bellvitge (IDIBELL)

Director: Dr. Esteban Ballestar Doctorando: Lorenzo de la Rica Lázaro

AGRADECIMIENTOS

ÍNDICE

RESUMEN

La presente tesis doctoral se ha centrado en investigar los procesos de desregulación en la metilación del DNA y de microRNAs en dos tipos celulares implicados en la patogenia de la artritis reumatoide, cuya función se encuentra exacerbada en el sinovio de las personas afectadas: fibroblastos sinoviales y osteoclastos. En primer lugar, se han analizado los fibroblastos sinoviales obtenidos de articulaciones procedentes de pacientes con artritis reumatoide, comparando los niveles de metilación del DNA y de expresión de microRNAs entre dos series obtenidas de pacientes y controles. En segundo lugar se ha estudiado el proceso de osteoclastogénesis, por el cual los monocitos, se fusionan y diferencian en osteoclastos multinucleados, capaces de degradar el hueso.

Los fibroblastos sinoviales y los osteoclastos son dos tipos celulares que en artritis reumatoide tienen una función aberrante, dado que se encuentran hiperactivos. En la articulación afectada, los fibroblastos sinoviales hiperactivos degradan el cartílago, mientras que los osteoclastos, contribuyen a la destrucción del hueso. Tras un proceso continuado de degradación, la articulación puede perder su función y la persona que padece la AR quedar seriamente afectada.

En el caso de los fibroblastos sinoviales, se han extraído de rodillas de personas con artritis reumatoide, y los hemos comparado con controles examinando sus perfiles de metilación de DNA y miRNAs, para investigar las diferencias entre los fibroblastos de pacientes y controles. Las variaciones entre los dos grupos encontradas indican potenciales vías de señalización y genes desregulados en esta enfermedad. Ente los ejemplos más reseñables cabe destacar la hipometilación del gen *IL6R*, *TNFAIP8*, *HOXA11y CD74*. Además, los datos de metilación, expresión de microRNAs así como expresión de mRNA, se integraron para conocer en profundidad las interconexiones entre las diferentes capas de regulación que están detrás del comportamiento aberrante de este tipo celular.

Por otro lado, hemos analizado *in vitro* qué cambios epigenéticos suceden durante la diferenciación de monocitos a osteoclastos. Para ello, hemos realizado un análisis de metilación de DNA durante el proceso de diferenciación de monocitos y osteoclastos derivados de los anteriores, así como de varias series temporales para conocer las dinámicas de metilación. Como resultado hemos descubierto que en este proceso de diferenciación, el perfil de metilación cambia drásticamente. Genes clave en la función del osteoclasto, como la *CTSK*, *ACP5* o *TM4SF7*, se hipometilan de forma activa, rápidamente, y se sobreexpresan. También genes específicos de monocito, como el *CX3CR1* se hipermetilan y silencian específicamente. Además, hemos determinado que el factor de transcripción PU.1, tiene un papel clave en este cambio

epigenético, dado que actúa como conector dual en el reclutamiento de la maquinaria epigenética (TET2 y DNMT3b) que va a modificar, en un sentido o en otro (hipometilación e hipermetilación) el epigenoma. La caracterización del mecanismo a nivel molecular en este proceso de diferenciación es clave para posteriormente encontrar vías de señalización que potencialmente puedan ser inhibidas farmacológicamente.

Por último, se ha analizado el perfil de expresión de microRNAs en este proceso de diferenciación mieloide. Se ha realizado inicialmente un screening de expresión de 380 miRNAs a varios tiempos (0d, 2d y 21d), centrándonos posteriormente en dos clusters de miRNAs, el miR-212/132 y el miR- 99b/let-7e/125a. Se ha demostrado el papel clave que estos microRNAs desempeñan en el proceso de diferenciación, dado que su inhibición funcional, retrasó la formación de los osteoclastos, así como la expresión de marcadores de osteoclasto, y la represión de genes de linaje alternativo.

ABREVIATURAS

5UTR	Región no traducida 5’
5meC	5-metilcitosina
5hmeC	5-hidroximetilcitosina
5caC	5-carboxilcitosina
5fC	5-formilcitosina
5-azadC	5-aza-2-desoxicitidina
AR	Artritis reumatoide
BER	Mecanismo de reparación del DNA por escisión de bases
DNA	Ácido desoxirribonucléico
DNMTs	DNA metiltransferasas
MCSF	Factor estimulante de colonias de macrófagos
miRNA	micro RNA
MMPs	Metaloproteinasas de matriz
MOs	Monocitos
OCs	Osteoclastos
RASF	Fibroblastos sinoviales de artritis reumatoide
RANKL	Ligando del receptor activador del factor nuclear kappa-B
RNA	Ácido ribonucléico
TET	Proteína “Translocación 10-11”
TDG	Timina-DNA-Glicosilasa

INTRODUCCIÓN

1. REGULACIÓN DE LA FUNCIÓN GÉNICA

La función de los genes depende no sólo de la secuencia de nucleótidos del DNA donde están codificados, sino también de una diverdidad de mecanismos reguladores a los que se ve sometida. La regulación de la actividad final de un gen puede controlarse a nivel transcripcional, posttranscripcional, por modificaciones postraduccionales, etc

Entre los mecanismos de regulación transcripcional se incluye el control ejercido por los factores de transcripción, que se complementa con la metilación del DNA, las modificaciones de las histonas, y otros mecanismos estudiados desde la perspectiva de la Epigenética. Dependiendo de los niveles de cada marca epigenética en un determinado promotor, se obtendrá un efecto sobre la actividad transcripcional, bien sea de activación, o por el contrario, de represión. A nivel de regulación postranscripcional de la función de los genes, existen diversos mecanismos, entre los cuales cabe destacar los mediados por los microRNAs. Son pequeños RNAs complementarios al RNA mensajero, que o bien median su degradación endo/exonucleolítica, o bien impiden la traducción del mismo a proteína.

1.1. EPIGENÉTICA

La Epigenética se define como el estudio de los cambios heredables en la expresión de genes que no son debidos a modificaciones de la secuencia de DNA[1, 2]. En general, dichos cambios de expresión están determinados por distintos tipos de modificaciones covalentes del DNA y las proteínas que lo empaquetan, las histonas.

Las principales marcas epigenéticas son la metilación del DNA y la modificación de aminoácidos conservados de los extremos N-terminales de las histonas (proteínas alrededor de las cuales se posiciona el DNA, en unidades repetitivas denominadas nucleosomas). Entre ambas modificaciones, hay una clara conexión mecanística, de manera que se pueden acoplar por medio de proteínas DNA metiltransferasas (DNMTs, capaces de reclutar desacetilasas y metiltransferasas de histonas[3, 4]) y proteínas MBD (Methyl CpG Binding Domain Proteins, que se unen a DNA metilado y reclutan enzimas modificadoras de histonas[5]).

Los eventos epigenéticos tienen implicaciones a nivel de transcripción, reparación y replicación de DNA, biología del cáncer[6], arquitectura nuclear[7], estabilidad cromosómica[8], impronta genética[9, 10], inactivación del Cromosoma X en mujeres[11],etc.

Los mecanismos epigenéticos en individuos normales son responsables de numerosos procesos relacionados con la diferenciación celular y el desarrollo (impronta genética, inactivación del cromosoma X en mujeres, expresión de genes

específicos de tejido durante el desarrollo…) y la desregulación de dichos mecanismos está relacionada con el desarrollo de enfermedades como el cáncer o diversos síndromes (síndrome de Rett, síndrome de inmunodeficiencia, inestabilidad centromérica y anormalidades faciales (ICF), síndrome de Prader-Willy, etc) [Revisado en[12]], por tanto no es de extrañar que haya sido uno de los campos de investigación con mayor expansión en la última década.

1.2. Metilación del DNA

La metilación del DNA es un proceso muy bien caracterizado química y enzimáticamente. Sucede sobre la posición 5' del anillo de citosina en el contexto de dinucleótidos CpG, desde un donador S-adenosil metionina (SAM).

Las enzimas encargadas de realizar este proceso son las DNA metiltransferasas (DNMTs): DNMT3A y DNMT3B se encargan de la metilación del DNA de novo durante el desarrollo embrionario, y DNMT1 se encarga de mantener el patrón de metilación de la hebra patrón en el DNA hemimetilado resultante tras la replicación. Hay otra DNMT, la DNMT3L, sin actividad catalítica pero con actividad reguladora sobre DNMT 3A y 3B[13, 14].

Los dinucleótidos CpG, a lo largo del genoma, se encuentran representados por debajo de lo esperable estadísticamente, y tienden a agruparse en lo que se conoce como "islas CpG" (región con al menos 200 pb y con un porcentaje de GC mayor de 50 y con un promedio de CpG observado/esperado mayor de 0.6). Las islas CpG, en general se encuentran en regiones repetitivas del genoma, y en regiones promotoras de genes codificantes de proteínas. De hecho, alrededor del 50% de los genes posee una isla CpG en su región promotora. Los dinucleótidos CpG dispersos suelen encontrarse metilados, mientras que los situados en islas CpG, no suelen estarlo en células normales[15].

Tradicionalmente, se acepta que la metilación del promotor de un gen, tiene implicaciones en su expresión, provocando una represión transcripcional[16]. Se ha visto que este mecanismo de regulación tiene tiene implicaciones en Impronta genética[9, 10], inactivación del cromosoma X en mujeres[11], etc, así como en determinados procesos patológicos. Sin embargo, ahora se conoce más sobre el efecto de la metilación fuera del contexto del promotor, siendo igualmente importantes el 5UTR y el primer exón. Por otra parte, la metilación del cuerpo del gen, se ha visto positivamente correlacionado con la expresión[17, 18], evidenciando la complejidad de mecanismos regulatorios en torno a la metilación del DNA.

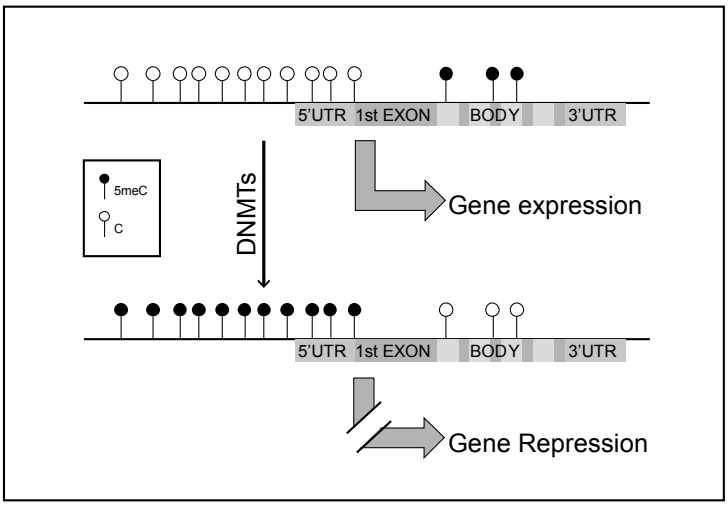


Figura 1. Regulación de la expresión génica por niveles de metilación del promotor. Promotores de genes con sus dinucleótidos CpG sin grupos metilo, están asociados a genes transcripcionalemnte activos. Estos dinucleótidos CpG pueden ser metilados por la familia de enzimas DNMTs, provocando un silenciamiento de la expresión de estos genes. Como se indica en la figura, el grado de metilación existente en el cuerpo del gen ("Body") está directamente relacionado con el nivel de expresión del mismo.

La metilación del DNA es una marca epigenética estable, que puede ser empleada por la célula como mecanismo de represión de genes en el largo plazo. Sin embargo, se ha observado que es una marca más dinámica de lo que originariamente se había considerado, pues existen mecanismos para revertir la 5meC a su estado original de citosina desmetilada.

1.3. Desmetilación activa del DNA

Uno de los mecanismos por los cuales las células pierden los niveles de metilación en determinadas regiones (promotores, etc…) es debido a un mantenimiento ineficiente del patrón de metilación por parte de la DNMT1 tras la división del DNA. Para que se de este proceso de desmetilación pasiva, la célula ha de dividirse, y si se divide con una velocidad superior a la actividad de la DNMT1, se observará hipometilación en determinados genes. La funciinalidad de estos genes que anteriormente estaban metilados, y ahora no, puede cambiar, de manera que los genes hipometilados pueden sobreexpresarse en ausencia de marcas silenciadoras como los grupos metilo. Sin embargo, también se han observado procesos de desmetilación en ausencia de división celular. Ejemplos de estas evidencias son:

-Desmetilación global del pronúcleo paterno[19]: tras la sustitución de la protamina por histonas tras la fertilización, el genoma del pronúcleo paterno sufre un proceso de desmetilación global.

-Desmetilación de genes específicos en células somáticas:

-Linfocitos T: 20 minutos después de ser estimulados, el enhancer del promotor de la interlekina 2 se desmetila, en ausencia de replicación[20].

-Neuronas: el gen del factor neurotrífico derivado de cerebro (*Brain-Derived Neurotrophic Factor* BDNF), gen clave para la plasticidad neuronal en adultos, se hipometila en neuronas despolarizadas por KCl, evitando ser reprimido por la proteína MeCP2. Las neuronas son células post-mitóticas, por lo que procesos de desmetilación activa son necesarios[21].

-Células dendríticas: el proceso de diferenciación de monocito a célula dendrítica sucede en ausencia de división, y se ha observado una desmetilación en genes clave para la identidad de estas células inmunes[22, 23].

1.4. Mecanismos de desmetilación activa

El estudio de la maquinaria implicada en los procesos de hipometilación activa de regiones genómicas ha sido un asunto controvertido durante los últimos años. Varios mecanismos han sido propuestos: eliminación catalítica del grupo metilo de la 5meC, desaminación de 5meC a Timina, y reparación de la MISSMATCH por el mecanismo de "Reparación por Escisión de Bases", BER (del inglés , Base Excision Repair), etc.

Uno de los mecanismos mejor documentados, es el de la oxidación de la 5meC a derivados como 5hidroximetilcitosina (5hmeC), 5carboxilcitosina (5caC) o 5 formilcitosina (5fC), por parte de enzimas de la familia TET, y la posterior acción de la proteína Timina-DNA-Glicosilasa (TDG), la cual escindiría la base oxidada creando un sitio de reparación[24].

1.4.1. Oxidación de la 5meC catalizada por la familia de enzimas TET

La 5meC puede ser oxidada sucesivamente a 5-hidroximetilcitosina (5hmC), 5-formilcitosina (5fC) y 5-carboxilcitosina (5caC) por la familia de proteínas TET (Ten eleven translocation family)[25, 26]. Esta familia de enzimas son unas dioxigenasas de DNA dependientes de Fe(II) y de 2-oxo-glutarato, y está compuesta por tres miembros: TET1, TET2 y TET3. TET1 está implicada, entre otras funciones, en la reprogramación epigenética que sufren las células germinales primordiales durante su desarrollo, y que conlleva la pérdida de metilación[27]. TET2 actúa principalmente a nivel del desarrollo mieloide, dado que mutaciones de pérdida de función en pacientes, conllevan el desarrollo de diferentes enfermedades mieloides, como síndromes mielodisplásicos, neoplasmas mieloproliferativas o leucemias monomielocíticas crónicas[28]. TET3, por su parte, ha sido relacionada con la pérdida de metilación que sufre el pronucleolo paterno tras la fecundación[29].

Dado que en esta tesis se ha analizado el proceso de diferenciación de monocito a osteoclasto, es decir, un proceso de diferenciación mieloide, nos centraremos en el papel de TET2, dado que parece tener más importancia en este linaje. El papel de TET2 en un proceso de diferenciación mieloide ha sido diseccionado en el proceso de transdiferenciación de linfocito a macrófago, donde se ha visto el papel fundamental de TET2 en la adquisición de los cambios epigenéticos necesarios para mantener el linaje mieloide al que pertenecen los macrófagos[30].

La hebra de DNA con nucleótidos oxidados derivados de la 5meC es reconocida por la enzima Timina-DNA-Glicosilasa (TDG), y la base de 5caC, por la cual tiene una gran afinidad, es escindida[24]. La implicación de estas enzimas en procesos de desmetilación activa ha sido confirmada gracias al uso de Knock-Outs (KOs) para ambas encimas. Se ha visto que el nivel de 5hmeC se ve reducido en ratones KO para TET1 y TET2, mientras que los niveles de 5meC se ven incrementados[31, 32]. Por otro lado, se ha analizado la acumulación de 5-formilcitosina (5fC) y 5-carboxilcitosina (5caC) en células KOs para TDG. Se ha visto un incremento de estos derivados oxidados de la 5meC, confirmando el papel de esta enzima en la escisión de los nucleótidos intermediarios de la desmetilación oxidativa[33, 34]. En la figura 2 se muestra un esquema con los diferentes intermediarios que intervienen en la desmetilación activa.

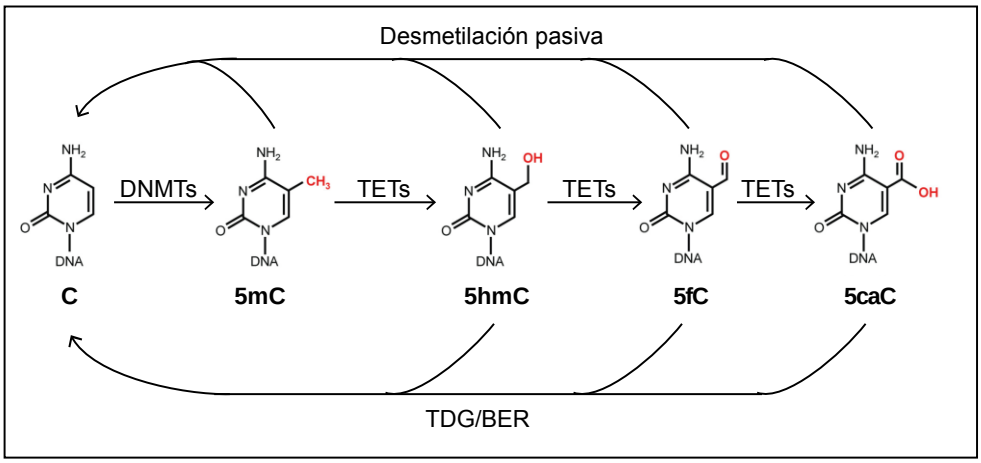


Figura 2. Desmetilación activa mediada por derivados oxidados de la 5meC. El nucleótido citosina, puede ser metilado en su carbono 5' por la familia de enzimas DNTMs. Este proceso puede ser revertido de forma pasiva, por la acción ineficiente de las DNMTs tras el proceso d división celular, o de forma activa. Para que este proceso se realice de forma activa se requiere la actividad de la familia de enzimas TET, que oxidan a la 5meC a 5hmC, 5fC o 5caC. Estos derivados oxidados son reconocidos por la Timina DNA glicosilasa, que elimina la base nitrogenada oxidada, dejando sólo el azúcar, en lo que se denomina un sitio abásico. Esto provoca un desapareamiento en el DNA, que es

reconocido por la maquinaria de reparación del DNA por escisión de bases, BER, por sus siglas en inglés (*Base Excision Repair*), resultando en una nueva Citosina sin el carbono 5 metilado.

1.5. MicroRNAs

Los microRNAs (miRNAs) son son una extensa familia de RNAs con ~21 nucleótidos de longitud que regulan la expresión de genes a nivel postranscripcional. Fueron descubiertos a principios de la década de los 2000[35] y su importancia ha sido estudiada en la práctica totalidad de procesos celulares, desde diferenciación, respuesta inmune, estrés, etc.

1.5.1 Síntesis de microRNAs

Los miRNAs se procesan desde precursores inmaduros (pri-miRNAs) que son transcritos desde genes independientes, o a partir de intrones de genes que codifican proteínas. Los pri-miRNAs se doblan adquiriendo estructura secundaria *hairpin*, que sirven de sustrato para Drosha y Dicer, ambas pertenecientes a la familia III de las RNAsas. El pri-miRNA es procesado por Drosha[36] dando lugar a un oligonucleótido de alrededor de 70 pares de bases (pre-miRNA) que es exportado del núcleo al citoplasma. Dicer, procesa el pre-miRNA a un dúplex miRNA/miRNA* de 20 pares de bases[37, 38], el miRNA maduro. Una de las hebras del dúplex miRNA/miRNA*es integrada al complejo miRISC (miRNA-induced silencing complex – complejo de silenciamiento inducido por miRNAs)[39, 40]. En función de la complementariedad interaccionará con un determinado RNA mensajero para reprimir su función. Puede evitar el que sea traducido y provocar la degradación exonucleolítica de la cola de poliA del RNA mensajero cuando la complementariedad es imperfecta con la región 3'UTR[41]. Cuando hay complementariedad perfecta, puede provocar la degradación endonuclolítica del mRNA por parte de AGO2 (Argonaute proteins)[42].

De la misma manera que cualquier gen, los genes que codifican miRNAs, pueden sufrir procesos de regulación transcripcional tales como metilación de sus promotores[43], etc.

La funcionalidad de los microRNAs puede ser estudiada gracia a algoritmos bioinformáticos que predicen los RNAs mensajeros (mRNA) dianas de los microRNAs estudiados. Algunos de los más populares son miRanda[44], TargetScan[45], PicTar[46], PITA[47], etc. En ellos, se analiza la probabilidad de interacción teniendo en cuenta diferentes parámetros, tales como la complementariedad, la estructura secundaria formada, la temperatura de fusión, etc... Gracias a estas herramientas se pueden predecir dianas potenciales que son silenciadas por el efecto de los microRNAs.

2. ENFERMEDAD AUTOINMUNE

El sistema inmune es un mecanismo muy evolucionado de protección frente a innumerables ataques. Para defendernos, el sistema inmune debe aprender antes qué elementos son propios del organismo, contra los que no hace falta actuar, y cuáles pueden suponer una amenaza externa, que sea necesario erradicar. La tolerancia inmunológica es el mecanismo que se encarga de ello[48]. Sucede a dos niveles: central y periférico. Dentro de los órganos implicados en la tolerancia inmunológica central, destaca el timo (para la maduración de los linfocitos T) y la médula ósea (para los linfocitos B), con papeles centrales en inmunología. Respecto a la tolerancia inmunológica periférica, los principales órganos encargados son los nódulos linfáticos, el bazo, etc… Gracias a estos mecanismos, se desarrollan bien un número bajo de células inmunes auto-reactivas, bien células auto-reactivas que son inactivas[49].

El hecho de que en todos los individuos considerados "sanos" haya linfocitos y anticuerpos auto-reactivos[50] y que estos no desarrollen ningún tipo de patología, depende de varios factores como la susceptibilidad genética, la exposición a factores contaminantes, tabaco, el nivel de estrés, y un largo etcétera. En algunos individuos, estos linfocitos y anticuerpos pueden atacar al propio organismo, causando patologías de tipo autoinmune.

Las enfermedades autoinmunes son dolencias multifactoriales causadas por la activación aberrante de linfocitos Ts y/o Bs, en ausencia de una infección activa u otra causa discernible[51]. Cuando el sistema inmune actúa contra un órgano concreto, se denomina autoinmunidad órgano-específica, y si afecta a todo el organismo, autoinmunidad sistémica. Entre las órgano-específicas, destacan la diabetes tipo I (afecta a los islotes pancreáticos), la enfermedad celiaca o la enfermedad de Crohn's (ambos con el tracto gastro-intestinal afectado) o la esclerosis múltiple (afecta al sistema nervioso). Las enfermedades autoinmunes sistémicas más comunes son la artritis reumatoide y el lupus eritematoso sistémico, en las cuales el sistema inmune ataca a varias partes del cuerpo, como las articulaciones (enfermedades reumáticas), piel, pulmones, corazón, etc[52].

2.1. Las enfermedades autoinmunes como modelo de enfermedades complejas

Las enfermedades autoinmunes son enfermedades complejas, cuya aparición depende de factores diversos, es decir, son enfermedades multifactoriales. Este grupo de enfermedades comenzó a ser estudiado desde el punto de vista conceptual y metodológico de la genética para investigar locus asociados a las patologías. De esta manera, usando una aproximación de genes candidatos se descubrió la

implicación de los genes del complejo mayor de histocompatibilidad (HLA - *Human leukocyte antigen*) en artritis reumatoide[53], lupus eritematoso sistémico[54] y escleroderma[55]. Más recientemente, gracias al avance de la tecnología, se han podido realizar estudios de asociación a nivel genómico global, GWAS (*Genome-wide Asociation Studies*), donde se ha descubierto la asociación de genes con la susceptibilidad a sufrir determinadas enfermedades autoinmunes, como son *IRF5*, *BLK*, *CD40*, *STAT4*, *PTPN22*, etc[56-58]. La mayor parte de estos genes están involucrados en la respuesta inmune, tanto innata como adaptativa.

Pese a los grandes esfuerzos invertidos por grandes consorcios internacionales[59, 60], ha quedado patente que el componente genético no es el único involucrado en el desarrollo de este tipo de enfermedades. Esta idea se ve reforzado por la existencia de una baja tasa de concordancia entre gemelos monocigóticos para estas enfermedades[61, 62]. Por tanto, existen evidencias que llevan a pensar que, por encima del componente genético, otros factores están actuando en las personas afectadas para desarrollar estas enfermedades. Entre los factores que posiblemente influencian el desarrollo de autoinmunidad, encontramos factores ambientales tales como agentes infecciosos como el virus del Epstein-Barr[63], el humo del tabaco[64], etc.

Muchos de los factores ambientales mencionados anteriormente, son capaces de modificar el perfil epigenético de las personas que se ven expuestos a ellos. La Epigenética está adquiriendo gran relevancia en el estudio de las causas subyacentes a las enfermedades autoinmunes. Como ejemplo, vemos que proteínas codificadas en el genoma del EBV (virus del Epstein-Barr) como LMP1 pueden activar a la DNMT1[65], modificando el estado de metilación del DNA, o que algunos fármacos hipometilantes como la 5-azadC causan enfermedades como el lupus inducido por fármacos[66]. También se ha visto que el humo del tabaco, que es un agente modificador del programa epigenético en cáncer[67, 68], promueve la aparición de artritis reumatoide el gemelo fumador de parejas de gemelos monocigóticos discordantes para la enfermedad, explicando potencialmente por qué es precisamente el gemelo que fuma el que padece la enfermedad, en un mismo entorno genético[69].

Por todo lo comentado anteriormente, es evidente que el interés del estudio de este tipo de enfermedades desde el punto de vista epigenético se ha convertido en un campo en continuo crecimiento durante la última década.

2.2. Artritis reumatoide

La artritis reumatoide (AR) es una enfermedad autoinmune que afecta principalmente a las articulaciones. En ellas, se produce hiperplasia de la membrana

sinovial, e inflamación de la articulación en general, que culmina con la destrucción de cartílago y hueso, dejándola sin función.

La artritis reumátide tiene una prevalencia de alrededor del 1%, y una incidencia de 3 casos nuevos al año por cada 10.000 habitantes. En la patogenia de la artritis reumatoide los linocitos T CD4 hiperactivados segregan citoquinas proinflamatorias que atraen a células inmunes circulantes en sangre periférica. Esta confluencia de células en un entorno de inflamación, rovoca la proliferación y activación de los sinoviocitos, o células residentes en el sinovio de la articulación, entre los que destacan los fibroblastos sinoviales de artritis reumatoide (RASF) y los osteoclastos. El sistema inmune está desregulado a varios niveles, ya que son varios los tipos celulares inmunes (Linfocitos B, T, macrófagos, etc) que intervienen, directa o indirectamente, en la patogénesis de esta enfermedad[70].

Como se ha comentado en el párrafo anterior, estas células inmunes, segregan grandes cantidades de citoquinas proinflamatorias en la articulación afectada, de manera que se crea un microentorno proinflamatorio con consecuencias muy negativas para la funcionalidad de la articulación[71]. En este entorno, células sinoviales como los fibroblastos sinoviales (conocidas como RASF – *Rheumatoid Arthritis Synovial Fibroblasts*), y los osteoclastos, están hiperactivados y son hiper-reactivos, y son los que finalmente ejecutarán la destrucción del cartílago y del hueso, respectivamente[72].

La AR se ha estudiado desde el punto de vista genético, y se han descubierto algunos locus de susceptibilidad como HHLA-DRB1, PTPN22, PADI4, STAT4, IL6ST, SPRED2, RBPJ, CCR6 y IRF5[56]. Sin embargo, la baja tasa de concordancia para esta enfermedad en gemelos monocigóticos (inferior al 15%[62, 73]), evidencia la necesidad de otros mecanismos aparte de los genéticos para que los pacientes acaben desarrollando la enfermedad.

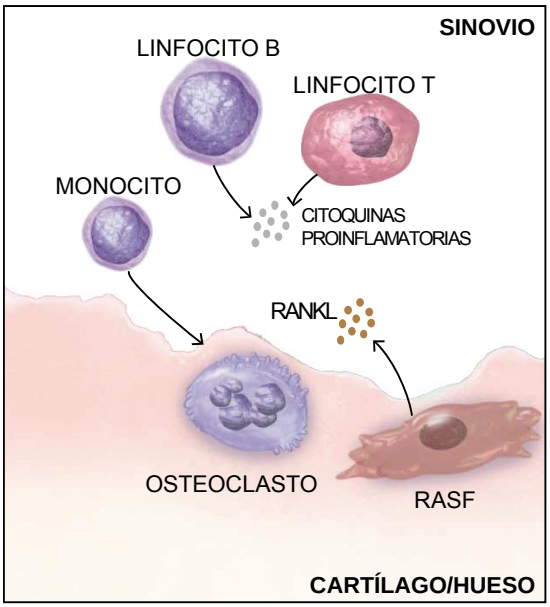


Figura 3. Tipos celulares implicados en la creación del entorno pro-inflamatorio en las articulaciones afectadas. Células inmunes extravasadas de sangre periférica al sinovio, segregan citoquinas proinflamatorias. Las citoquinas proinflamatorias activan y hacen proliferar a las células residentes en lamembrana sinovial, los sinoviocitos. Uno de los tipos de sinoviocitos son los RASF, que aparte de degradar el cartílago, segregan RANKL. El RANKL presente en la articulación, es capaz de activar la diferenciación de los monocitos a osteoclastos, en un proceso llamado osteoclastogénesis. Los osteoclastos activados de forma aberrante degradarán el hueso de la articulación. La acción conjunta y mantenida en el tiempo de RASFs y osteoclastos termina por degradar la articulación, provocando un pérdida de funcionalidad de la misma. Figura adaptada de E.H.S. Choy, and G.S. Panayi N Engl J Med 2001[71].

Como se puede apreciar en la figura 3, aparte de la importancia de las células efectoras de la destrucción del cartílago y el hueso (RASF y osteoclastos), las desregulación de las células inmunes circulantes también tiene gran importancia en el desarrollo de la enfermedad.

3. RASF

Los fibroblastos sinoviales procedentes de pacientes con artritis reumatoide son más agresivos que los de personas sanas por varios motivos. En primer lugar, expresan mayores niveles de metaloproteinasas (MMPs) y citoquinas que aquellos extraídos de donantes sanos[74, 75]. Además, muestran un comportamiento de célula tumoral ya que: son más invasivos en el cartílago[76], tienen más resistencia a la apoptosis[77], y son capaces de crecer sin estar adheridos a un sustrato[78].

Este tipo celular ha sido uno de los más ampliamente estudiado dentro de la AR, también, desde el punto de vista epigenético, donde sean encontrado bastantes particularidades que diferencian estos fibroblastos de los existentes en las articulaciones de personas sanas.

3.1. Desregulación de la metilación del DNA en rasf

Los RASF tienen unas características epigenéticas únicas, que explican en parte lo agresivo de su comportamiento en el sinovio de los pacientes afectados. En AR, se ha comprobado que el tejido sinovial obtenido de pacientes con AR está hipometilado en relación al de pacientes sanos[79]. Además, se ha comprobado que debido a esta hipometilación existente en los RASF, éstos sobreexpresan el retrotransposon *Line 1*[80], elemento silenciado en condiciones fisiológicas. Esta sobreexpresión de *Line 1* provoca la sobreexpresión de genes que probablemente contribuyan al fenotipo agresivo como son el proto-oncogén met (*MET*), p38delta MAP kinasa (*MAPK13*) y la proteína de unión a galectina 3 (*LGALS3BP*)[81].

Los RASF expresan menores niveles de la enzima metilante DNMT1, y tienen menores niveles de metilación global, es decir, muestran hipometilación global. Se ha logrado que fibroblastos sinoviales extraídos de controles sanos se comporten de forma agresiva e invasiva similar a los RASF tratando a estos fibroblastos sanos con agentes hipometilantes como la 5-azadC[82], lo cual evidencia la importancia de este entorno hipometilado en la patogénesis de la AR.

Hay un microRNA cuyua expresión parece ser regulada por metilación en RASF. El promotor del gen que codifica el el microRNA *hsa-miR-203* está hipometilado y como consecuencia el miR-203 sobreexpresado[83]. Indirectamente, esta sobreexpresión provoca una mayor producción de *MMP-1* e *IL-6*, importantes en la invasividad y en la inflamación respectivamente.

Curiosamente, en RASF también se ha demostrado hipermetilación en algunos genes específicos, como es el caso del receptor de apoptosis 3 (death receptor 3, *DR3* o *TNFRSF25*), que permite una mayor resistencia a la apoptosis[84].

Hasta aquí, se han mostrado algunos ejemplos de trabajos donde se han analizado genes candidatos, sin embargo también se han hecho análisis de metilación globales,

como el realizado por el grupo de Firestein en 2013[85]. En él, se analizaba la metilación global de 6 RASF y 6 OASF, descubriéndose 1859 genes con cambios de metilación entre los dos grupos. Entre los genes hipometilados, surgieron varios relevantes para la AR, como *CHI3L1*, *CASP1*, *STAT3*, *MAP3K5*, *MEFV*. También algunos hipermetilados como *TGFBR2* o *FOXO1*. Todos estos trabajos evidencian la importancia de la epigenética en la AR. Algunos de los cambios de metilación más importantes se muestran en la Figura 4, y su potencial relación con la activación de osteoclastos.

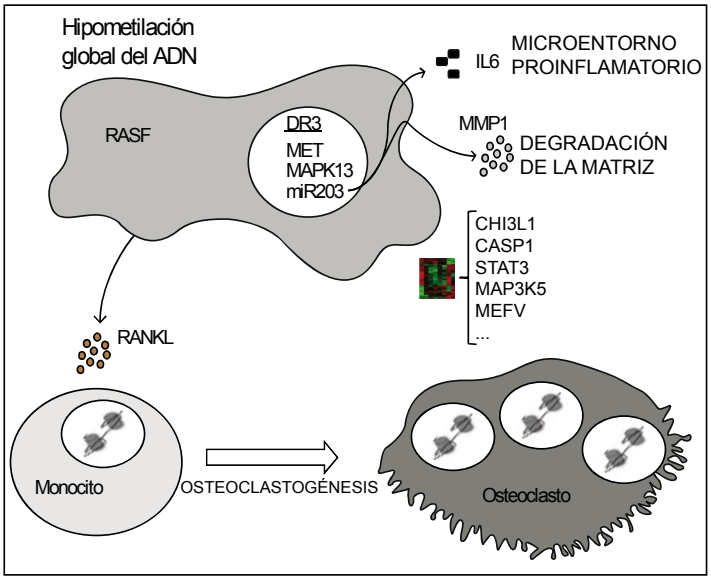


Figura 4. Desregulación de los niveles de metilación del DNA en RASF. Resumen de los hallazgos realizados en RASF desde el punto de vista de la metilación del DNA. Aparte de la hipometilación global detectada, varios estudios con genes candidatos han demostrado la importancia de algunos genes desregulados por metilación en el comportamiento agresivo de los RASF. Además, se muestra la relación entre el comportamiento agresivo de los RASF y su capacidad para segregar RANKL, el cual activa la osteoclastogénesis.

3.2. Desregulación de los niveles de microRNAs en RASF

Desde el año 2008, en que se descubrió la primera evidencia de la implicación de los miRNAs en la patogénesis de la AR, se ha avanzado mucho en el conocimiento sobre las rutas afectadas por la desregulación de microRNAs. Los microRNAs podrían ser una potencial diana para el tratamiento farmacológico de la AR, o bien ayudar a su diagnóstico temprano como biomarcadores.

En AR se ha comprobado que miR-155 se sobreexpresa en el tejido sinovial y en RASF, y que esto provoca una reducción de los niveles de MMP-3, indicando su posible papel en la modulación de las propiedades destructivas de este miRNA[86]. miR-146a también está sobreexpresado en RASF[87], así como miR-203,el cual, como ya se ha comentado anteriormente, está regulado por los niveles de metilación de su promotor[83].

Hay un microRNA, el miR-18a, que forma parte de un loop de retroalimentación positiva con NK-kB. Este microRNA contribuye a la destrucción y a la inflamación crónica de la articulación dado que se provoca la sobreexpresión de enzimas que degradan matriz y moléculas proinflamatorias[88]. Otro miembro del mismo cluster (cluster miR-17-92) también se ha visto implicado en AR. miR-20a está silenciado en RASF, y esto provoca la sobreexpresión de ASK1 apoptosis signal-regulating kinase 1, de manera que se producen mayores cantidades de IL6, CXCL-10 y TNF-a, contribuyendo a la patogénesis de la AR a través de mediadiores de inflamación[89].

Por otro lado, miR-124 está a niveles más bajos en RASF, y esto tiene efectos a nivel de proliferación y de reclutamiento de monocitos, dado que miR-124 tiene como diana a CDK-2 y a la proteína quimioatrayente de monocitos 1 (MCP-1)[90].

Otro microRNA silenciado es el miR-34a*, un microRNA proapoptítico que actúa a través de XIAP. Su silenciamiento, en parte explicaría la mayor resistencia de los RASF a la apoptosis[91]. Además, miR-19b está silenciado en RASF y tiene como diana TLR2, un gen muy importante en la inflamación existente en AR. TLR2, que está sobreexpresado en RASF, a su vez regula la expresión de IL-6 y MMP3[92]. Por lo que hay varios ejemplos de microRNAs cuya regulación está afectada en RASF, y que potencialmente podrían darnos pistas sobre los mecanismos de acción de la AR, así como potenciales terapias.

4.OSTEOCLASTOS

Los osteoclastos son células gigantes, multinucleadas que se encargan de degradar hueso[93]. Se diferencian a partir de progenitores monocito/macrófago[94] tras su estimulación con MCSF[95] (*Macrophage colony-stimulating factor*) y RANKL[96] (*Receptor activator of nuclear factor kappa-B ligand*). El proceso de diferenciación que da lugar a osteoclastos a partir de sus progenitores es conocido como osteoclastogénesis. Este complejo proceso de diferenciación requiere la fusión de células progenitoras, la reorganización del citoesqueleto[97] y la activación de un programa genético específico para cumplir la función de degradar el hueso.

Se puede emular la diferenciación que tiene lugar en el hueso *in vitro*, tratando progenitores de osteoclastos procedentes de PBMCs, o CD14 con MCSF y RANKL[98]. Estos osteoclastos generados in vitro son capaces de degradar hueso, y expresan marcadores de osteoclastos[99].

4.1. Características generales de los osteoclastos y enfermedades en las que están involucrados

4.1.1. Proteínas importantes en la degradación ósea. Marcadores de osteoclastos

Para degradar el hueso, los osteoclastos expresan una serie de enzimas que se encargan del catabolismo óseo, y que nos permiten cuantificar sus niveles para determinar la presencia y calidad de los osteoclastos. La catepsina K CTSK es una cisteín proteasa activa a pHs ácidos (a partir de 4.5) que segregan los osteoclastos a la laguna de resorción para que degrade el colágeno y otras proteínas de la matriz ósea[100]. La metaloproteinasa de matriz 9 (MMP9) es una enzima que también degrada matriz extracelular cuya expresión en osteoclastos[101] es inducida tras la estimulación con RANKL[102]. La anhidrasa carbónica II (CA2) por otro lado, es la encargada de generar los iones H^+ necesarios (con CO2 y H2O genera H2CO3) para acidificar el medio óseo que va a ser degradado[103, 104].

Por último, la fosfatasa ácida resistente a tartrato(TRAP) es la encargada de degradar varias proteínas de la matriz ósea como la sialoproteína ósea, fosfotroteinas de matriz ósea, así como la osteopontina[105]. Es usado como marcador por excelencia para determinar la presencia de osteoclastos[106, 107].

4.1.2. Enfermedades relacionadas con el funcionamiento aberrante de los osteoclastos

La función correcta de los osteoclastos es de vital importancia para la correcta homeostasis del sistema óseo. Esta importancia se hace patente cuando la función de los osteoclastos es deficiente, y se produce una osteopetrosis[108] generalizada. Por el

contrario, la hiperactivación aberrante de los osteoclastos produce una pérdida de masa y densidad ósea, conocida como osteoporosis[109], tal como se muestra en la figura 5.

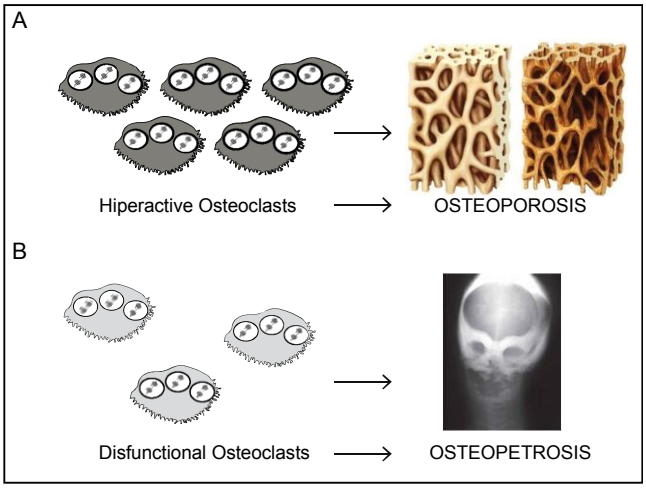


Figura 5. Enfermedades relacionadas con el funcionamiento aberrante de los osteoclastos. Cuando hay una excesiva diferenciación de osteoclastos, y estos están hiperactivados, la homeostasis ósea se desregula. Hay más degradación del hueso que síntesis, de manera que la densidad ósea es inferior, y la resistencia de los mismos se ve comprometida. Esto provoca osteoporosis, y las personas que la padecen, tienen un riesgo grande de sufrir fracturas óseas. Por otro lado, cuando el proceso de diferenciación está menos activo de lo normal, o la función de los osteoclastos no es la correcta, la renovación de los huesos, necesaria para el desarrollo, el crecimiento, y la curaciónd e fracturas, se ve comprometida. Se provoca un fenómeno conocido como osteopetrosis, en el cual, por ejemplo a nivel craneal, no se permite el crecimiento del sistema nervioso de acuerdo a la edad del individuo. Imagen superior extraída de Medical picture. Imagen inferior extraída de Jaypee Brothers Medical Publishers.

Precisamente la hiperactivación aberrante de los osteoclastos es una circunstancia que se da en varios tipos de tumores, como el mieloma múltiple[110], el cáncer de próstata y el cáncer de mama[111]. En estos tumores, los blastos del mieloma, o las células metastáticas, envían señales pro-osteoclásticas que finalmente provocan un aumento en la degradación del hueso y por tanto un incrementado riesgo de fracturas.

Además de tumores en los que se promueve la diferenciación de osteoclastos, hay un tipo de tumor propio del linaje osteclástico, que sucede cuando el proceso de diferenciación de monocito a osteoclasto se ve afectado. Es el tumor de células gigantes, también llamado osteoclastoma, en donde hay una sobreactivación de los osteoclastos y un crecimiento descontrolado de estas células

gigantes, que causa daños severos a la estructura ósea de los pacientes afectados[112, 113].

4.1.3. Importancia de los osteoclastos en la patogénesis de la artritis reumatoide

En la artritis reumatoide hay dos eventos determinantes que provocan la pérdida de función de la articulación. La degradación del cartílago en la articulación, realizado por parte de los RASF, tal como se ha mencionado anteriormente, y la degradación del hueso, que es llevada a cabo por osteoclastos hiperactivados de manera aberrante.

La presencia de osteoclastos en la superficie del pannus sinovial demuestra su implicación en la destrucción ósea en artritis reumatoide[114, 115]. Esta implicación además se ve reforzada con el uso de modelos animales con artritis experimental, en los cuales si se bloquea la osteoclastogénesis, se ve que las articulaciones quedan protegidas de destrucción focal del hueso[116, 117]. La causa más probable del incremento en la presencia y actividad de los osteoclastos en las articulaciones afectadas, es que en estos sinovios, hay una mayor concentración de RANKL. Este RANKL es producido por los Linfocitos T y los RASF del sinovio[118, 119] y estimula la diferenciación de los osteoclastos.

Se ha demostrado la importancia de la osteoclastogénesis por vías alternativas, como por el hecho de que una de las terapias empleadas en el tratamiento de la AR, el denosumab (un anticuerpo anti-RANKL), atenúa el desarrollo y la progresión de la degradación ósea en pacientes con AR[120, 121]. Este fármaco es un inhibidor competitivo del ligando de RANK, RANKL. De esta manera se consigue bloquear una de las señales principales que desencadena la diferenciación de los osteoclastos.

4.2. Diferenciación de osteoclastos

Las rutas de señalización activadas en los progenitores tras la estimulación con MCSF y RANKL han sido extensamente estudiadas. Tras la estimulación con MCSF y RANKL se producen varios eventos que en primer lugar mostraremos de manera descriptiva, para después analizarlos en mayor detalle. En primer lugar se activa TRAF-6[122, 123], así como los adaptadores de ITAM[124] (*immunoreceptor tyrosine-based activation motif*). Los adaptadores de ITAM son por un lado DAP12[125], cuyo receptor es TREM-2[125] y por otro FcRg[126], que actúa a través de OSCAR y de oscilaciones en las concentraciones de calcio[127]. Todas estas señales culminan con la activación de NF-kB, MAPK y c-Jun[128], los cuales actúan sobre NFATc1, uno de los reguladores transcripcionales principales de la osteoclastogénesis[129]. NFATc1 actúa en conjunción con otros dos fatores de transcripción que ya estaban presentes en el progenitor, como son PU.1 y MITF[130].

El receptor de MCSF es Csf1R (*Colony-stimulating factor 1 Receptor*), un receptor tirosina kinasa que se autofosforila en siete de los residuos tirosínicos de su cola citoplasmática. Esta cola fosforilada sirve para reclutar varias moléculas entre las que se encuentran PI3 kinasas, Src o Grb2[131]. De esta manera se estimula la expresión de RANK, entre otras moléculas[132].

Por su parte, RANKL interacciona con su receptor RANK[133], y le permite reclutar factores asociados al receptor de TNF, conocidos como proteínas TRAF[134, 135]. A partir de aquí se activan al menos seis cascadas de señalización: NF-kB, AP-1 y p38 MAP Kinasa, ERK (extracelular signal regulated kinase) , Src y PI3-kinasa. La principal molécula adaptadora en este paso es TRAF6[136, 137], la cual es capaz de fosforilar IkB, que deja de inhibir a NF-kB[138] ya que es degradada por la ruta del proteasoma. Los heterodímeros p50-RelA ó p52-RelB se transportan de forma activa al núcleo y estimulan la expresión de genes osteolásticos[139].

La activación de la ruta de la p38 MAP Kinasa por fosforilación activa al regulador transcripcional MITF, importante en la sobreexpresión de genes importantes para el osteoclasto como la CTSK y TRAP[140]. Además, como resultado de esta compleja ruta de señalización el factor de transcripción CREB (cAMP response element-binding) es fosforilado, e induce la expresión de c-Fos, el cual es clave en la sobreexpresión de NFATc1[141].

Por tanto, como resultado de esta estimulación aumenta la expresión del factor de transcripción NFATc1. La actividad de este factor de transcripción está regulado por calcineurina y por oscilaciones en las concentraciones de Ca^{2+}, mediadas por la señalización por ITAMs (a través de sus moléculas adaptadoras, DAP12 y FcRg). Las oscilaciones en la concentración de Ca provocadas por ITAMs activan a la CaMKIV[141] (Calcium/calmodulin-dependent protein kinase type IV) e inhiben a la calcineurina (calcium/calmoduline-dependent protein phosphatase) y esto resulta en la fosforilación de NFATc1 y consiguiente activación, permitiéndole ser translocado al núcleo y realizar su función controlando la diferenciación de los pre-osteoclastos y la multinucleación.

El esquema de las rutas de señalización activadas en la diferenciación de osteoclastos puede encontrarse en la Figura 6.

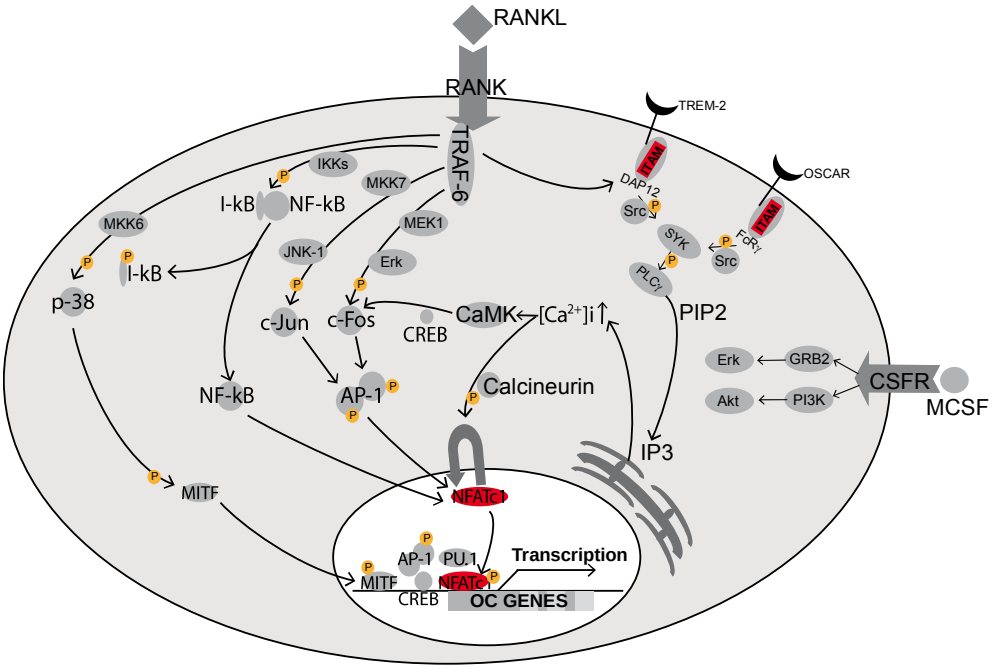


Figura 6. Esquema de las rutas de señalización activadas en la diferenciación de osteoclastos. Tras la unión de RANKL a su receptor, RANK, se activa TRAF-6, el principal adaptador de la señal inducida por RANKL. TRAF-6 es capaz de reclutar y activar a los adaptadores ITAMS DAP12 y FcRg,los cuales reclutan kinasas que fosforilan a la PLCg, activando la ruta de las IP3 y la liberación de calcio desde el retículo endoplasmático. Las oscilaciones de calcio provocan la activación de la CaMK, la cual a través de CREB fosforila y activa a Fos por un lado, y a la Calcineurina por otro, la cual fosforila a NFATc1 y lo activa. Por otro lado, TRAF-6 también es capaz de reclutar y activar otra serie de kinasas, cuyo resultado final es la activación de varias rutas de señalización:NF-kB, AP-1, p-38, etc. Estas moléculas activadas, pueden pasar al núcleo o actuar en el citoplasma, pero de cualquier forma acaban activando y reclutando factores de transcripción (MITF, PU.1, NFATc1, etc), o bien unirse directamente a promotores de genes de osteoclasto que deben ser activados. El resultado final es que NFATc1, el principal factor de trasncripción de la osteoclastogénesis, sufe un proceso de autoamplificación, y de activación redundante por varias vías diferentes, que permite la activación transcripcional de los genes encargados de la osteoclastogénesis y de la degradación ósea.

4.3. Factores de transcripción implicados en osteoclastogénesis

Tal y como se ha mencionado en el epígrafe anterior, hay varios factores de transcripción que intervienen en el proceso de diferenciación de monocitos a osteoclatos. Los más importantes son PU.1, MITF y NFATc1.

PU.1 es importante en la producción de células B, de macrófagos, células dendríticas y osteoclastos. Referente a este último tipo celular, el defecto de PU.1 en ratones knock out provoca osteopetrosis aguda, debida a un defecto en la generación de progenitores monocíticos que den lugar a osteoclastos[142].

MITF (Microphthalmia-associated transcription factor) es un factor de transcripción que es fosforilado y activado por ERK, el cual es inducido por la estimulación de los monocitos con MCSF. Los ratones con MITF mutado, también sufren osteopetrosis severa debido a la incapacidad de los progenitores osteoclásticos de fusionarse[143].

NFATc1 es un factor de transcripción regulado por Ca^{2+}-Calcineurina cuya expresión se activa tras la estimulación con RANKL. Se le considera el principal regulador de la osteoclastogénesis. NFATc1 induce la expresión de muchos genes responsables de la diferenciación de los OCs, así como de su función degradando hueso. Algunos ejemplos incluyen *OSCAR*, *TRAP*, *CTSK*, etc[144].

Se ha descrito con anterioridad que varios de estos factores de transcripción interaccionan para activar o reprimir la expresión de genes. Mientras que MITF y NFATc1 tienen una acción más específica y exclusiva en los osteoclatos, PU.1 tiene una expresión en más tipos celulares. Además, debido a su versatilidad y a su capacidad de interaccionar con maquinaria epigenética, a continuación ampliaremos el conocimiento actual sobre este factor de transcripción.

4.4. El factor de transcripción PU.1

El factor de transcripción PU.1 es una pieza clave en el desarrollo de osteoclastos. La proteína tiene 264 aminoácidos, está codificada en la región p11.22[145] del cromosoma 11 y es el producto del gen Spi-1[146]. Pertenece a la familia de factores de transcripción ETS[147] (*E-twenty six*), ya que tiene un dominio de unión al DNA ETS[148].

4.4.1. Dominios de PU.1

Este factor de transcripción consta de tres dominios principales: el dominio de transactivación, el dominio PEST, y el dominio de unión del DNA ETS[149]. Se muestra un esquema de la estructura de PU.1 en la Figura 7.

-Dominio de Transactivación: en el extremo amino terminal. Gracias a este dominio es capaz de interaccionar con otras proteínas, modulando su función.

-Dominio PEST: está en la parte central, es rico en prolina (P), ácido glutámico (E), serinas (S) y treoninas (T), y es necesario para algunas interacciones proteína-proteína.

-Dominio de unión del DNA ETS: está en el extremo carboxi terminal. Reconoce cajas PU (rico en purinas), con la secuencia consenso GAGGAA. PU.1 se une al DNA como monómero.

Figura 7. Esquema de las estructura del factor de transcripción PU.1. Motivos que conforman su estructura. Dominio de transactivación, dominio PEST, dominio de unión del DNA ETS.

4.4.2. Función de PU.1

La forma en que se regula la unión de PU.1 a sus sitios de unión es extremadamente compleja, lo que da pistas de la importancia de este factor de transcripción para el correcto funcionamiento de las células inmunes. La unión de PU.1 a un determiando sitio de unión, depende de muchos factores, entre los que cabe destacar la secuencia, la accesibilidad de esa región, la concentración de PU.1 en la célula, así como la presencia de otros factores de transcripción que cooperen en su reclutamiento[150].

PU.1 es clave en la hematopoyesis[151]. Regula la diferenciación de linfocitos B, así como de todo el linaje mieloide y eritroide, dependiendo de en qué concentraciones se encuentre[152, 153].

En células mieloides, PU.1 está presente a altas concentraciones, y controla directamente la expresión de genes críticos para la diferenciación y el funcionamiento de macrófagos. Entre estos genes se encuentra la integrina CD11b, y los receptores M-CSF y GM-CSF. PU.1 es necesario también para la diferenciación de osteoclastos, dado que su ausencia en ratones KO para PU.1 provoca osteopetrosis[142].

4.4.3. PU.1 en la diferenciación de osteoclastos

Las primeras evidencias que mostraron la importancia de PU.1 en la diferenciación de los osteoclastos surgieron en 1997, donde se observó que tras la inducción de la osteoclastogénesis la concentración de PU.1 aumentaba. En ratones KO para PU.1, sin embargo, se observaba una depleción en el número de osteoclastos y macrófagos, indicando el origen común de ambos tipos celulares, así como el papel en la etapa inicial de la diferenciación de PU.1[142]. El motivo por el cual PU.1 es tan importante en este proceso de diferenciación radica en que controla la expresión de muchos genes clave para los osteoclastos. Por ejemplo, la expresión de RANK, el receptor de RANKL, está regulada por la unión de PU.1 a su promotor[154]. PU.1 no actúa aislado, sino que interacciona con diferentes "partners" para activar o reprimir

la expresión de genes. Uno de los factores de transcripción asocaiados a PU.1 en osteoclastogénesis más importantes es MITF, mencionado anteriormente. MITF interacciona con PU.1 y c-Fos y determina que su localización subcelular sea nuclear[155]. Además, la interacción entre PU.1 y MITF es necesaria para activar la expresión de uno de los marcadores más importantes del osteoclasto, el gen ACP5 (TRAP)[156, 157], así como de algunos elementos clave en la ruta de señalización activada por RANKL, el receptor OSCAR (*osteoclast-associated receptor*)[158]. Aparte de con MITF, se ha descrito la interacción de PU.1 con NFATc1, uno de los reguladores principales de este proceso. La interacción PU.1-NFATc1 contribuye a la expresión del receptor anteriormente citado, OSCAR[159].

PU.1 es un factor de transcripción clave para la activación de otros genes necesarios en la osteoclastogénesis. Por ejemplo, la expresión del gen de la CTSK, o de la integrina beta 3 es posible gracias a la interacción y a las sinergias entre PU.1 y NFATc1 (activado por fosforilación por p38 MAP kinasa)[160, 161].

La red de factores de transcripción MITF, PU.1, NFATc1 es clave en estas células, y se cree que si bien PU.1 y MITF actúan al principio, activando la expresión de genes diana, es NFATc1 el que a posteriori, mantiene la expresión de estos genes en las células diferenciadas[130].

4.4.4. PU.1 interacciona con maquinaria modificadora de la cromatina

Se ha visto que PU.1 interacciona con el remodelador de la cromatina SWI/SNF en genes de osteoclastos tales como TRAP o CTSK[130]. Pero además, hay otras evidencias que otorgan a PU.1 un papel importante en la modificación de la estructura de la cromatina de los genes a los que se une, reclutando otros miembros de la maquinaria epigenética.

Por ejemplo, PU.1 interacciona con la histona desacetilasa 1, HDAC1[162]. HDAC1 desacetila histonas, y contribuye a que la estructura de los nucleosomas sea más compacta y el DNA esté menos accesible, reprimiéndose la expresión de los genes a los que se une.

En la diferenciación de macrófagos, se ha visto que PU.1 silencia la expresión del cluster de microRNAs miR-17-92 reclutando a Egr-2/Jarid1b, los cuales desmetilan lalisina 4 de la histona 3 (H3K4)[163].

Por otro lado, en la diferenciación de monocitos, se ha visto que PU.1 reprime la expresión de Bim, un factor proapoptótico, cambiando el estado de la cromatina de su promotor. PU.1 es capaz de reclutar a SUZ12 y a EZH2, los cuales trimetilan a la lisina 27 de la histona 3 (H3K27), marcando de forma represiva ese promotor[164].

En la diferenciación de células de leucemia, PU.1 es capaz de activar la expresión de Cebpa y Cbfb. Para ello incrementa la cantidad e la marca de cromatina transcripcionalmente activa acetilación de H3K9 a través de mecanismos todavía desconocidos[165]

PU.1 no solo es capaz de modificar el estado de la cromatina reclutando modificadores de marcas de histonas. Se ha descrito que también es capaz de interaccionar con la DNMT3a/b para alterar también el estado de metilación de determinados promotores[166].

Por lo tanto, se pone en evidencia la importancia de PU.1 no sólo para reclutar a otros factores de transcripción para activar o reprimir la expresión de determiandos genes, sino que también es capaz de modificar el estado de la cromatina allá donde se une. Esto lo consigue interaccionando con diferentes tipos de modificadores de la cromatina.

4.5. Cambios en microRNAs durante la osteoclastogénesis

La implicación de los microRNAs en el proceso de osteoclastogénesis ha sido de reciente descubrimiento. Desde 2007 se ha incrementado el conocimiento en este campo, empleando como modelos la osteoclastogénesis humana, la de ratón, o la diferenciación desde células RAW264.7.

Gracias al uso del modelo de RAW264.7, se descubrió la implicación del miRNA-223, el cual se expresa en la etapa inicial, con un papel "antiosteoclástico", probablemente implicado en la modulación de los niveles de osteoclastos, ya que cuando se sobreexpresaba miR-223, se inhibe la formación de osteoclastos[167]. Para demostrar la importancia de los microRNAs en este proceso, se ha silenciado la maquinaria responsable del procesamiento de los microRNAs, como son DGCR8, Dicer, y Ago2. En todos los casos, se inhibió la formación de osteoclastos funcionales, demostrando que un correcto metabolismo de los microRNAs es necesario. Concretamente, el silenciamiento de Dicer redujo la expresión de TRAP y de NFATc1, así como el número de osteoclastos TRAP positivas[168]. Además, se comprobó que en cultivos de médula ósea de ratón, miR-223 es necesario para la formación de osteoclastos, al contrario que lo que sucede en RAW264.7. Este microRNa actúa silenciando el gen NFI-A. El silenciamiento de este gen es necesario para la expresión del receptor de MCSF[169]. Se ha comprobado su importancia en un modelo expreimental de artritis (artritis inducida por colágeno), en el que se ha visto que inhibiendo la expresión de este miRNA a través del uso de lentivirus, se reduce la severidad de la artritis[170, 171].

También en RAW se ha comprobado la importancia de otro microRNA, miR-155, el cual reprime la expresión de MITF, y evita la diferenciación de estas células a osteoclastos[172]. Además, en un modelo experimental en ratón de artritis autoinmune, se ha visto que depleccionando la expresión de miR-155, las articulaciones de estos ratones sufrían menor degradación ósea[173, 174], confirmando la importancia de miR-155, así como su posible uso como diana terapéutica en pacientes con AR.

El perfil de expresión de microRNAs cambia drásticamente durante el proceso de diferenciación de macrófagos de médula ósea (los progenitores de los osteoclastos en médula ósea de ratón) a osteoclastos[175]. Además, miR-21 se sobreexpresa de forma notable. miR-21 regula la expresión de PDCD4 (*programmed cell death*), silenciándolo. De esta manera, PDCD4 deja de reprimir c-Fos, y se permite la normal diferenciación de BMM a osteoclastos en ratón.

miR-146 es un microRNA que está presente en altas concentraciones el sinovio de la articulación de la AR, así como en los PBMCs. Se ha visto que si se silencia este miRNA en PBMCs, se disminuye drásticamente el número de osteoclastos TRAP positivas. Además, si se suminsitra este miRNA en ratones con artritis inducida por colágeno, se evita la destrucción de la articulación[176].

Otros ejemplos de microRNAs importantes para el proceso de osteoclastogénesis son el miR-29b[177] o mir-124[178] (deben silenciarse durante el proceso, dado que son reguladores negativos).

OBJETIVOS

Dada la prevalencia de las enfermedades autoinmunes, y más concretamente, de la artritis reumatoide, es necesario profundizar en el conocimiento de los mecanismos que pueden conducir al desarrollo de la autoinmunidad, y las consecuencias que ello conlleva para los pacientes. La presente tesis doctoral pretende generar nuevo conocimiento sobre los cambios epigenéticos y de expresión de microRNAs asociados a los dos tipos celulares más importantes implicados en la destrucción de las articulaciones de los pacientes con artritis reumatoide. En el caso de los RASF, se planteó investigar las alteraciones en la metilación del DNA y expresión de microRNAs con respecto a fibroblastos control, así como realizar un análisis integrado de metilación y expresión. Con respecto a los osteoclastos, el estudio se diseño con el objeto de caracetrizar la dinámica y mecanismso de cambios en la metilación de DNA y microRNAs durante la diferenciación de monocitos a osteoclastos, proceso altamente exacerbado en los pacientes de artritis reumatoide.

De forma específica los objetivos de este trabajo son:

1. El estudio de los cambios de metilación del DNA asociados a la diferenciación de monocito a osteoclasto y sus bases moleculares. Dado que este proceso de diferenciación está hiperactivado en pacientes con AR, es fundamental caracterizarlo molecularmente. Además, los osteoclastos maduros tienen hasta 50 núcleos y albergan gran cantidad de información genética redundante, la cual es necesario regular. Comparando los perfiles de metilación entre los precursores y las células diferenciadas, se analizará la importancia de los cambios de metilación para el proceso, cómo suceden estos cambios y qué maquinaria está implicada en los mismos. (Artículo 1).

2. El estudio de los cambios en los niveles de microRNAs que suceden durante la diferenciación de monocito a osteoclasto, así como de su funcionalidad. Complementando al objetivo número 1, dada la gran cantidad de información genética existente en el osteoclasto, aparte de la regulación transcripcional ejercida por la metilación del DNA, es interesante conocer los mecanismos de regulación postranscripcionales que permiten a esta célula realizar su función. Se estudiará el efecto funcional de modicifar los niveles de microRNAs clave en la expresión de genes clave para la osteoclastogénesis, así como su efecto en la diferenciación de osteoclastos. Finalmente se intentará elucidar el mecanismo de acción de los microRNAs, estudiando a través de qué dianas actúan. (Artículo 2).

3. El estudio de los cambios de metilación del DNA y en los niveles de microRNAs que diferencian a los RASF de sus homólogos sanos. A partir del estudio de las características que diferencian a los fibroblastos obtenidos a partir de pacientes con AR, de los de OA, se profundizará en los posibles genes, microRNAs y rutas implicadas en el comportamiento agresivo de este tipo celular. Además, se estudiarán los cambios en estas dos marcas de manera integrada, dado que en una célula nada sucede de forma aislada, y esta aproximación aportará información clave sobre las múltiples capas que se desregulan en las células sinoviales de los pacientes con artritis reumatoide. (Artículo 3).

RESULTADOS

Por la presente certifico que el doctorando **LORENZO DE LA RICA LÁZARO** presentará la Tesis Doctoral en forma de compendio de tres artículos, dos de los cuales ya han sido publicados. Su contribución a los artículos se indica a continuación:

ARTÍCULO 1:

Lorenzo de la Rica, José M. Urquiza, David Gómez-Cabrero, Abul B.M.M.K Islam, Nuria López-Bigas, Jesper Tegnér, René E.M. Toes and Esteban Ballestar

TÍTULO: Identification of novel markers in rheumatoid arthritis through integrated analysis of DNA methylation and microRNA expression.

REVISTA: Journal of Autoimmunity 2013 Jan; (41) ; 6-16

doi:10.1016/j.jaut.2012.12.005

(Factor de impacto: 8.15)

En este artículo Lorenzo de la Rica fue el responsable de concebir, en colaboración conmigo la mayor parte de los experimentos, así como de su realización. Se encargó de cultivar, mantener y amplificar las muestras de fibroblastos así como de aislar todas las muestras empleadas en este estudio. Llevó a cabo la mayor parte de los experimentos (Figuras 2, 3 y 4) y análisis de datos del estudio, a partir de los análisis bioinformáticos realizados por José M. Urquiza. Lorenzo también analizó en interpretó todos los resultados obtenidos en colaboración conmigo. Fue el encargado de la realización de todas las figuras y participó en la escritura y revisión del artículo.

ARTÍCULO 2:

Lorenzo de la Rica, Javier Rodríguez-Ubreva, Mireia García, Abul B. M. M. K. Islam, José M. Urquiza, Henar Hernando, Jesper Christensen, Kristian Helin, Carmen Gómez-Vaquero, and Esteban Ballestar

TÍTULO: PU.1 target genes undergo Tet2-coupled demethylation and DNMT3b-mediated methylation in monocyte-to-osteoclast differentiation

REVISTA: Genome Biology 2013 Sep 12;14(9):R99 doi: 10.1186/gb-2013-14-9-r99

(Factor de impacto: 10.3)

En este artículo Lorenzo de la Rica fue el responsable de concebir, en colaboración conmigo la mayor parte de los experimentos, así como de su realización. Fue el responsable de la puesta a punto del modelo de osteoclastogénesis en el laboratorio, y de todos los experimentos necesarios para la determinación de la presencia de osteoclastos (qPCRs de marcadores, tinciones TRAP, inmunofluorescencias, etc). Llevó a cabo la mayor parte de los experimentos del estudio (Figuras 1, 2, 3A,B, 4A,B,E,F, 5A,B,C,D y 6 ; Figuras suplementarias 1 , 2A,B, 3, 4, 5C) y generó todas las muestras de osteoclastos empleadas en el mismo. Asistiendo a Javier Rodriguez-Ubreva, colaboró en los experimentos de ChIP e Inmunoprecipitaciones (Fig 3C, 4C,D y 5E y Supp Fig 5 A y B). También analizó los datos generados a partir del análisis bioinformático del array de metilación y los arrays de expresión realizados por José M. Urquiza, y se encargó de la representación de los datos y la búsqueda de genes relevantes para el proceso. Finalmente, Lorenzo también analizó en interpretó todos los resultados obtenidos en colaboración conmigo. Fue el encargado de la realización de todas las figuras y participó en la escritura y revisión del artículo.

ARTÍCULO 3:

Lorenzo de la Rica, Natalia Ramírez-Comet, Laura Ciudad, Mireia García, José M. Urquiza, Carmen Gómez-Vaquero and Esteban Ballestar

TÍTULO: MicroRNA profiling reveals key role of miR-212/132 and miR- 99b/let-7e/125a clusters in monocyte to osteoclast differentiation

REVISTA: In preparation

En este artículo Lorenzo de la Rica fue el responsable de concebir, en colaboración conmigo la mayor parte de los experimentos, así como de su realización. Fue el responsable de realizar el screening de microRNAs. Llevó a cabo la mayor parte de los experimentos del estudio y se encargó de la representación de los datos y la búsqueda de miRNAs relevantes para el proceso. Finalmente, Lorenzo también analizó en interpretó todos los resultados obtenidos en colaboración conmigo. Fue el encargado de la realización de todas las figuras y participó en la escritura y revisión del artículo.

Y para que así conste a todos los efectos firmo la presente en L'Hospitalet de Llobregat,
Barcelona a 27 de Septiembre de 2013

Esteban Ballestar, Ph.D.
Chromatin and Disease Group, Leader
Cancer Epigenetics and Biology Programme (PEBC)
Bellvitge Biomedical Research Institute (IDIBELL)
Avda. Gran Via 199-203
08908 L'Hospitalet de Llobregat, Barcelona, Spain
Tel: +34 932607133
Fax: +34 932607219
e-mail: eballestar@idibell.org

ARTÍCULO 1:

Revista:

Journal of Autoimmunity, [41 (2013) 6-16 doi:10.1016/j.jaut.2012.12.005 ISSN: 0896-8411]

Título:

Identification of novel markers in rheumatoid arthritis through integrated analysis of DNA methylation and microRNA expression

Autores:

Lorenzo de la Rica[a], José M. Urquiza[a], David Gómez-Cabrero[b], Abul B. M. M. K. Islam[c,d], Nuria López-Bigas[c,e], Jesper Tegnér[b], René E. M. Toes[f], and Esteban Ballestar[a]

Afiliaciones:

a - Chromatin and Disease Group, Cancer Epigenetics and Biology Programme (PEBC), Bellvitge Biomedical Research Institute (IDIBELL), 08907 L'Hospitalet de Llobregat, Barcelona, Spain
b - Department of Medicine, Karolinska Institutet, Computational Medicine Unit, Centre for Molecular Medicine, and Swedish e-science Research Centre (SeRC), Solna, Stockholm, Sweden
c - Department of Experimental and Health Sciences, Barcelona Biomedical Research Park, Universitat Pompeu Fabra (UPF), 08003 Barcelona, Spain
d - Department of Genetic Engineering and Biotechnology, University of Dhaka, Dhaka 1000, Bangladesh
e - Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain
f - Department of Rheumatology, Leiden University Medical Center, Leiden, The Netherlands

RESUMEN EN CASTELLANO

Las enfermedades autoinmunes reumáticas son trastornos complejos cuya etiopatogenia se atribuye a una interacción entre la predisposición genética y los factores ambientales. Ambos factores (los genes de susceptibilidad a la autoinmunidad, y los factores ambientales), están involucrados en la generación de perfiles epigenéticos aberrantes en tipos celulares específicos que, en última instancia, provocan cambios en la expresión normal de los genes. Además, los cambios en los perfiles de expresión de los miRNAs, también causan la desregualción de genes asociados a fenotipos aberrantes. En la artritis reumatoide, varios tipos de células son las que finalmente destruyen la articulación, siendo los fibroblastos sinoviales uno de los más importantes. En esteartículo, se ha realizado un estudio de los niveles de metilación del DNA y de los niveles de expresión de microRNAs en un conjunto de muestras de fibroblastos sinoviales procedentes de pacientes con artritis reumatoide, y se ha comparado con fibroblastos extraídos a pacientes con osteoartritis, con un fenotipo normal. El análisis de los cambios de metilación del DNA permitió identificar cambios en nuevos genes diana clave como IL6R, CAPN8 y DPP-4, así como en varios genes HOX. Los cambios de metilación de una porción significativa de genes tienen consecuencias a nivel de la expresión de los mismos, mostrando una relación inversa. También el análisis de los niveles de expresión de los microRNAs mostró varios grupos de microRNAs con cambios específicos en los RASF. Además, se realizó un análisis integrado de los datos de metilación, expresión de microRNAs y expresión génica, observando microRNAs controlados por los niveles de metilación de sus promotores, así como genes que son regulados simultáneamente por los niveles de metilación de sus promotores, así como por microRNAs que se les unen. Muchos de los genes descubiertos podrían ser usados como marcadores clínicos en artritis reumatoide.

En este estudio se han identificado nuevas diana desreguladas en los fibroblastos sinoviales de artritis reumatoide. Gracias al análisis integrativo de microRNAs y cambios epigenéticos, se establece una nueva forma de análisis de los datos provenientes de pacientes, en la que se tienen en cuenta varias capas regulatorias de manera simultanea, y no de manera aislada.

ABSTRACT

Autoimmune rheumatic diseases are complex disorders, whose etiopathology is attributed to a crosstalk between genetic predisposition and environmental factors. Both variants of autoimmune susceptibility genes and environment are involved in the generation of aberrant epigenetic profiles in a cell-specific manner, which ultimately result in dysregulation of expression. Furthermore, changes in miRNA expression profiles also cause gene dysregulation associated with aberrant phenotypes. In rheumatoid arthritis, several cell types are involved in the destruction of the joints, synovial fibroblasts being among the most important. In this study we performed DNA methylation and miRNA expression screening of a set of rheumatoid arthritis synovial fibroblasts and compared the results with those obtained from osteoarthritis patients with a normal phenotype. DNA methylation screening allowed us to identify changes in novel key target genes like IL6R, CAPN8 and DPP4, as well as several HOX genes. A significant proportion of genes undergoing DNA methylation changes were inversely correlated with expression. miRNA screening revealed the existence of subsets of miRNAs that underwent changes in expression. Integrated analysis highlighted sets of miRNAs that are controlled by DNA methylation, and genes that are regulated by DNA methylation and are targeted by miRNAs with a potential use as clinical markers. Our study enabled the identification of novel dysregulated targets in rheumatoid arthritis synovial fibroblasts and generated a new workflow for the integrated analysis of miRNA and epigenetic control.

Keywords: rheumatoid arthritis, rheumatoid arthritis synovial fibroblasts, DNA methylation, epigenetic, microRNAs, integration

1. INTRODUCTION

Rheumatoid arthritis (RA) is a chronic autoimmune inflammatory disease characterized by the progressive destruction of the joints. RA pathogenesis involves a variety of cell types, including several lymphocyte subsets, dendritic cells, osteoclasts and synovial fibroblasts (SFs). In healthy individuals, SFs are essential to keep the joints in shape, doing so by providing nutrients, facilitating matrix remodeling and contributing to tissue repair [1]. In contrast to normal SFs or those isolated from patients with osteoarthritis (osteoarthritis synovial fibroblasts, OASFs), rheumatoid arthritis synovial fibroblasts (RASFs) show activities associated with an aggressive phenotype, like upregulated expression of protooncogenes, specific matrix-degrading enzymes, adhesion molecules, and cytokines [2]. Differences in phenotype and gene expression between RASFs and their normal counterparts reflect a profound change in processes involved in gene regulation at the transcriptional and post-transcriptional levels. The first group comprises epigenetic mechanisms, like DNA methylation, whilst miRNA control constitutes one of the best studied mechanisms of the second.

DNA methylation takes place in cytosine bases followed by guanines. In relation with transcription, the repressive role of methylation at CpG sites located at or near the transcription start sites of genes, especially when those CpGs are clustered as CpG islands, is well established [3]. Methylation of CpGs located in other regions like gene bodies is also involved in gene regulation [4, 5]. At the other side of gene regulation lie microRNAs (miRNAs), a class of endogenous, small, non-coding regulatory RNA molecules that modulate the expression of multiple target genes at the post-transcriptional level and that are implicated in a wide variety of cellular processes and disease pathogenesis [6].

The study of epigenetic- and miRNA-mediated alterations in association with disease is becoming increasingly important as these processes directly participate in the generation of aberrant profiles of gene expression ultimately determining cell function and are pharmacologically reversible. Epigenetics is particularly relevant in autoimmune rheumatic diseases as it is highly dependent on environmental effects. As indicated above, both genetics and environmental factors contribute to ethiopathology of autoimmune rheumatic disorders. This double contribution is typically exemplified by the partial concordance in monozygotic twins (MZ) [7, 8]. It is of inherent interest to identify autoimmune disease phenotypes for which the environment plays a critical role [9]. Many environmental factors, including exposure to chemicals, tobacco smoke, radiation, ultraviolet (UV) light and infectious agents among other external factors, are associated with the development of autoimmune

rheumatic disorders [10]. Most of these environmental factors are now known to directly or indirectly induce epigenetic changes, which modulate gene expression and therefore associate with changes in cell function. For this reason, epigenetics provides a source of molecular mechanisms that can explain the environmental effects on the development of autoimmune disorders [11]. The close relationship between environment and epigenetic status and autoimmune rheumatic disease is also exemplified by using animal models [12]. This type of studies is also essential for the identification of novel clinical markers for disease onset, progression and response to treatments.

In this line, initial reports demonstrated hypomethylation-associated reactivation of endogenous retroviral element L1 in the RA synovial lining at joints [13]. Additional sequences have since been found to undergo hypomethylation in RASFs, like IL-6 [14] and CXCL12 [15]. Candidate gene analysis has also enabled genes to be identified that are hypermethylated in RASFs [16]. More recently, DNA methylation profiling of RASFs versus OASFs has led to the identification of a number of hypomethylated and hypermethylated genes [17]. With respect to miRNAs, reduced miR-34a levels have been linked with increased resistance of RASFs to apoptosis [18], and lower miR-124a levels in RASFs impact its targets, CDK-2 and MCP-1 [19]. Conversely, miR-203 shows increased expression in RASFs [20]. Interestingly, overexpression of this miR-203 is demethylation-dependent, highlighting the importance of investigating multiple levels of regulation and the need to use integrated strategies that consider interconnected mechanisms.

In this study, we have performed the first integrated comparison of DNA methylation and miRNA expression data, together with mRNA expression data from RASFs versus OASFs (Figure 1) in order to investigate the relevance of these changes in these cells and to overcome the limitations of using a small number of samples. Our analysis identifies novel targets of DNA methylation- and miRNA-associated dysregulation in RA. Integration of the analysis of these two datasets suggests the existence of several genes for which the two mechanisms could act in the same or in opposite directions.

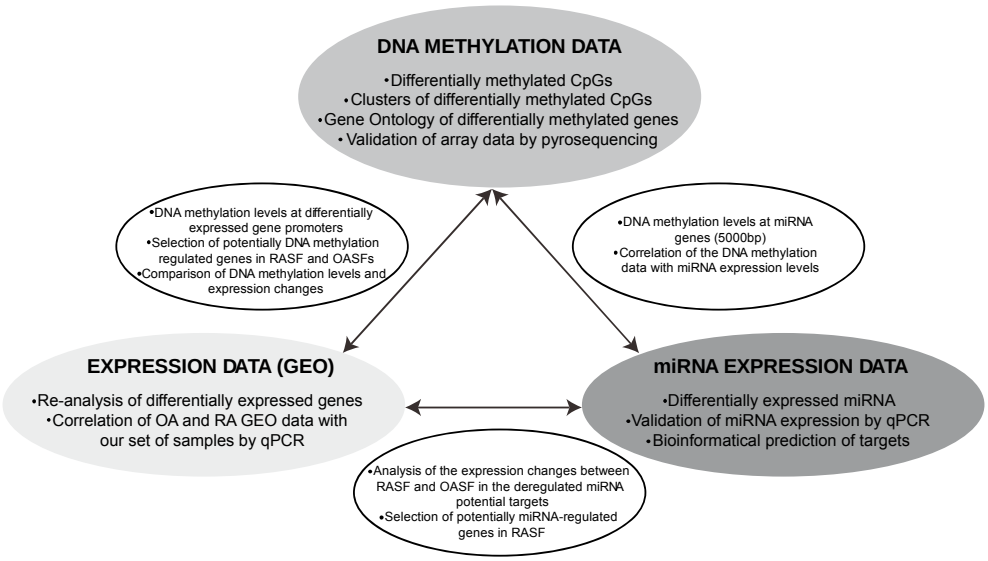


Figure 1. Scheme depicting the strategy designed in this study where DNA methylation and miRNA data are integrated with expression array data. The grey oval areas show the type of information that individual analysis of DNA methylation, miRNA expression and expression datasets can provide. This is listed within these grey oval areas and are described in detail in the Results section. Between these grey oval areas, smaller elliptical panels show the type of analysis that can provide the combined information between DNA methylation and expression datasets (left), DNA methylation and miRNA expression data sets (right), or expression and miRNA expression datasets (bottom).

2. MATERIAL AND METHODS

2.1. Subjects and sample preparation

Fibroblast-like synoviocytes (FLSs) were isolated from synovial tissues extracted from RA and OA patients at the time of joint replacement in the Department of Rheumatology of Leiden University Medical Center. All RA patients met the 1987 criteria of the American College of Rheumatology. Before tissue collection, permission consistent with the protocol of the Helsinki International Conference on Harmonisation Good Clinical Practice was obtained. All individuals gave informed consent. Synovial tissues were collected during the arthroscopy, frozen in Tissue-Tek OCT compound (Sakura Finetek, Zoeterwoude, Netherlands) and cut into 5-µm slices using a cryotome (Leica CM 1900). Fibroblast cultures were maintained in Dulbecco's modified Eagle's medium supplemented with 10% fetal calf serum.

2.2. DNA methylation profiling using universal bead arrays

Infinium HumanMethylation450 BeadChips (Illumina, Inc.) were used to analyze DNA methylation. With this analysis it is possible to cover > 485,000 methylation sites per sample at single-nucleotide resolution. This panel covers 99% of RefSeq genes, with an average of 17 CpG sites per gene region distributed across the promoter, 5'UTR, first exon, gene body, and 3'UTR. It covers 96% of CpG islands, with additional coverage in island shores and the regions flanking them. Bisulfite conversion of DNA samples was done using the EZ DNA methylation kit (Zymo Research, Orange, CA). After bisulfite treatment, the remaining assay steps were identical to those of the Infinium Methylation Assay, using reagents and conditions supplied and recommended by the manufacturer. Two technical replicates of each bisulfite-converted sample were run. The results were all in close agreement and were averaged for subsequent analysis. The array hybridization was conducted under a temperature gradient program, and arrays were imaged using a BeadArray Reader (Illumina Inc.). The image processing and intensity data extraction software and procedures were those described by Bibikova and colleagues [21]. Each methylation datum point was represented as a combination of the Cy3 and Cy5 fluorescent intensities from the M (methylated) and U (unmethylated) alleles. Background intensity, computed from a set of negative controls, was subtracted from each datum point.

2.3. Detection of differentially methylated CpGs

Differentially methylated CpGs were selected using an algorithm in the statistical computing language R [22], version 2.14.0. In order to process Illumina Infinium HumanMethylation450 methylation data, we used the methods available in the LIMMA and LUMI packages [23] from the Bioconductor repository [24]. Before statistical analysis, a pre-process stage was applied, whose main steps were: 1) Adjusting color balance, i.e., normalizing between two color channels; 2) Quantile normalizing based on color balance-adjusted data; 3) Removing probes with a detection p-value > 0.01; 4) Filtering probes located in sex chromosomes; 5) Filtering probes considered to be SNPs (single nucleotide polymorphisms). Specifically, the probes were filtered out using Illumina identifiers for SNPs, i.e. those probes with an "rs" prefix in their name; 6) Non-specific filtering based on the IQR (interquartile range) [25], using 0.20 as the threshold value. Subsequently, a Bayes-moderated t-test was carried out using LIMMA [26]. Several criteria have been proposed to indentify significant differences in methylated CpGs. In this study, we adopted the median-difference beta-value between the two sample groups for each CpG [27, 28]. Specifically we considered a probe as differentially methylated if (1) the absolute

value of the median-difference between b-values is higher than 0.1 and the statistical test was significant (p-value<0.05).

2.4 Identification of genomic clusters of differentially methylated CpGs

A clustering method available in Charm package [29] was applied to the differentially methylated CpGs. Although Charm is a package specific for analyzing DNA methylation data from two-color Nimblegen microarrays, we reimplemented the code to invoke the main clustering function using genomic CpG localization. By using this approach, we identified Differentially Methylated Regions (DMR) by grouping differentially methylated probes closer than 500 pbs. In this analysis, the considered lists of CpGs were those associated with a value of $p < 0.01$.

2.5 Bisulfite pyrosequencing

CpGs were selected for technical validation of Infinium Methylation 450K by the bisulfite pyrosequencing technique in the RASF and OASF samples. CpG island DNA methylation status was determined by sequencing bisulfite-modified genomic DNA. Bisulfite modification of genomic DNA was carried out as described by Herman and colleagues [30]. 2 µl of the converted DNA (corresponding to approximately 20 – 30 ng) were then used as a template in each subsequent PCR. Primers for PCR amplification and sequencing were designed with the PyroMark® Assay Design 2.0 software (Qiagen). PCRs were performed with the HotStart Taq DNA polymerase PCR kit (Qiagen) and the success of amplification was assessed by agarose gel electrophoresis. Pyrosequencing of the PCR products was performed with the PyromarkTM Q24 system (Qiagen). All primer sequences are listed in Supplementary Table 1.

2.6 Gene expression data analysis and comparison of DNA expression and DNA methylation data

To compare expression and methylation data, we used RASF and OASF expression data from the Gene Expression Omnibus (GEO) under the accession number (GSE29746) [31]. Agilent one-color expression data were examined using LIMMA [24]. The pre-process stage consisted of background correction, followed by normalization. Thus, the applied background correction is a convolution of normal and exponential distributions that are fitted to the foreground intensities using the background intensities as a covariate, as explained in the LIMMA manual. Next, a well-known quantile method was performed to normalize the green channel between the arrays and then the green channel intensity values were log2-

transformed. Values of average replicate spots were analyzed with a Bayes-moderated t-test. Expression genes matching methylated genes were then studied. Genes differentially expressed between RASF and OASF groups were selected if they met the criteria of having values of p and FDR (False Discovery Rate) lower than 0.05 as calculated by Benjamini-Hochberg and a greater than two-fold or less than 0.5-fold change in expression. Expression data were validated by quantitative RT-PCR. Primer sequences are listed in Supplementary Table 1.

2.7 microRNA expression screening, target prediction and integration with DNA methylation data

Total RNA was extracted with TriPure (Roche, Switzerland) following the manufacturer's instructions. Ready-to-use microRNA PCR Human Panel I and II V2.R from Exiqon (Reference 203608) were used according to the instruction manual (Exiqon). For each RT-PCR reaction 30 ng of total RNA was used. Samples from OASF and RASF patients were pooled and two replicates of each group were analyzed on a Roche LightCycler® 480 real-time PCR system. Results were converted to relative values using the inter-plate calibrators included in the panels (log 2 ratios). RASF and OASF average expression values were normalized with respect to reference gene miR-103. Differentially expressed microRNAs (FC > 2 or < 0.5) were selected.

To predict the potential targets of the dysregulated microRNAs, we used the algorithms of several databases, specifically TargetScan [32], PicTar [33], PITA [34], miRBase [35], microRNA.org [36], miRDB/MirTarget2 [37], TarBase [38] , and miRecords [39], StarBase/CLIPseq [40]. Only targets predicted in at least four of these databases and differentially expressed between RASFs and OASFs were included in the heatmaps.

To compare the DNA methylation bead array data with the miRNA expression levels, miRNAs were mapped to Illumina 450k probes. For each differentially expressed miRNA we studied the CpGs within a 5000 bp window around the transcription start site. Using the GRCh37 assembly annotation for Illumina, the genomic localization of probes was extracted in order to match them with miRNA loci. Genomic features of miRNAs were taken from the miRBase [41] and Illumina annotation was obtained from IlluminaHumanMethylation450K.db Bioconductor Package [41].

2.8 Gene Ontology Analysis

Gene Ontology analysis was done with the FatiGO tool [42], which uses Fisher's exact test to detect significant over-representation of GO terms in one of the sets (list of selected genes) with respect to the other one (the rest of the genome). Multiple test correction to account for the multiple hypothesis tested (one for each GO term) is applied to reduce false positives. GO terms with adjusted P-value < 0.05 are considered significant.

2.9 Graphics and Heatmaps

All graphs were created using Prism5 Graphpad. Heatmaps were generated from the expression or methylation data using the Genesis program from Graz University of Technology [43].

3. RESULTS

3.1. Comparison of DNA methylation patterns between RASF and OASF reveals both hypomethylation and hypermethylation of key genes

We performed high-throughput DNA methylation screening to compare SF samples from six RA and six OA patients. To this end, we used a methylation bead array that allows the interrogation of > 450,000 CpG sites across the entire genome covering 99% of RefSeq genes. Statistical analysis of the combined data from the 12 samples showed that 2571 CpG sites, associated with 1240 different genes, had significant differences in DNA methylation between RASFs and OASFs (median β differences > 0.10, $p < 0.05$) (Figure 2A and Supplementary Table 2). Specifically, we found 1091 hypomethylated CpG sites (in 575 genes) and 1479 hypermethylated CpG sites (in 714 genes).

The list of genes differentially methylated between RASFs and OASFs includes a number with known implications for RA pathogenesis and some potentially interesting novel genes (Table 1). One of the best examples is IL6R. Our results indicated that IL6R is hypomethylated in RASFs with respect to OASFs, and that hypomethylation is probably associated with IL6R overexpression in RASFs. IL6 and IL6R are factors well known to be associated with RA pathogenesis and progression. IL6R overexpression plays a key role in acute and chronic inflammation and increases the risk of joint destruction in RA. Also, IL6R antibodies have recently been approved for the treatment of RA [44]. Another interesting example in the hypomethylated gene list is TNFAIP8, or TIPE2, a negative mediator of apoptosis that plays a role in inflammation [45]. We also identified CAPN8 as the gene with the greatest difference

between RASFs and OASFs. This gene has not previously been associated with RA, although it is involved in other inflammatory processes such as irritable bowel syndrome [46].

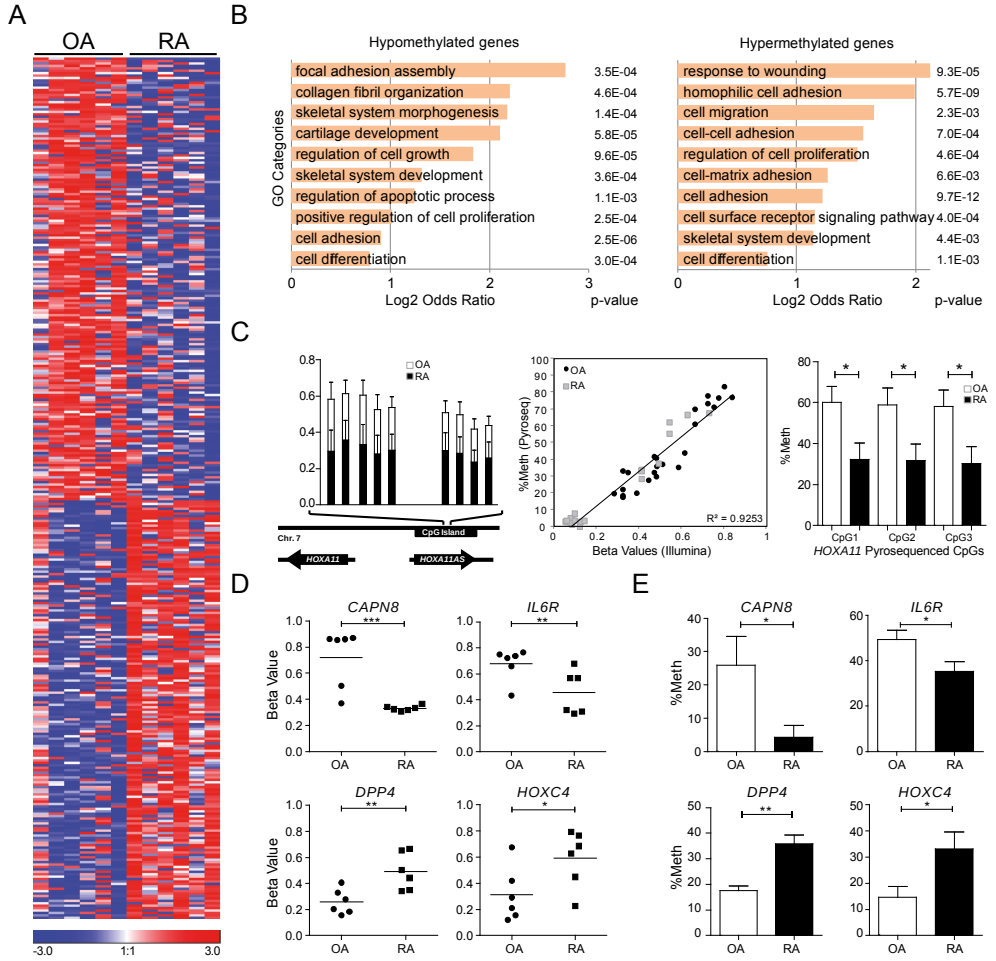


Figure 2. Comparison of the DNA methylation profiles between RASFs and OASFs samples. (A) Heatmap including the methylation data for the six RASF and OASF samples shows significant differential methylation. There are both significant hypermethylated and hypomethylated genes. In this heatmap, all the genes with a value of $p < 0.05$ and a difference in median □ > 0.2 are shown. The scale at the bottom distinguishes hypermethylated (red) and hypomethylated (blue) genes. (B) Summary of the gene ontology (GO) analysis for the category "biological process' among hypomethylated and hypermethylated genes. P-values are shown on the right (C) Methylation data from the array analysis corresponding to HOXA11 gene in which 9 consecutive CpGs are hypomethylated in RASFs relative to OASFs (left), comparison of the array data and pyrosequencing, where the excellent correlation between the two sources of data is shown by a regression line (center), methylation values as obtained through

pyrosequencing corresponding to three selected CpGs comparing RASF and OASF samples. (D) Comparison of the array data (left) and pyrosequencing data (right) of four selected hypomethylated (CAPN8, IL6R) and hypermethylated (DPP4, HOXC4) genes.

Conversely, hypermethylated genes include factors like DPP4 and CCR6. DPP4 encodes a serine protease, which cleaves a number of regulatory factors, including chemokines and growth factors. DPP4 inhibitors have recently emerged as novel pharmacological agents for inflammatory disease [47]. Several lines of evidence have also shown a role for CCR6 in RA [48].

We then set out to determine whether our differentially methylated genes could be involved in biological functions relevant to RA pathogenesis. We therefore performed Gene Ontology analysis to test whether some molecular functions or biological processes were significantly associated with the genes with the greatest difference in DNA methylation status between RASFs and OASFs. The analysis was performed independently for gene lists in the hypomethylated and hypermethylated group. We observed significantly enriched functional processes that are potentially relevant in the biology of SFs (Figure 2B), including the following categories: focal adhesion assembly (GO:0048041), cartilage development (GO:0051216) and regulation of cell growth (GO:0001558) for hypomethylated genes. For hypermethylated genes, we observed enrichment in categories such as response to wounding (GO:0009611), cell migration (GO:0016477) and cell adhesion (GO:0007155). Hypermethylated and hypomethylated genes shared several functional categories, such as cell differentiation (GO:0030154), cell adhesion(GO:0007155) and skeletal system development(GO:0001501) characteristic of this cell type.

We also compared our data with those reported in a recent study by Nakano and colleagues [17]. We found a significant overlap of genes that were hypomethylated and hypermethylated in both sets of samples (Supp. Fig 1). These included genes like MMP20, RASGRF2 and TRAF2 from the list of hypomethylated genes, and ADAMTS2, EGF and TIMP2 from among the hypermethylated genes (see Supp. Fig 1 and Table 2 in [17]). The use of a limited set of samples in the identification of genes introduces a bias associated with each particular sample cohort, which would explain the partial overlap between different experiments. However in this case, we observed an excellent overlap between both experiments.

Gene Name	Δmeth CpG	Region	Description	Δ Beta (RA-OA)	FC Express (RA/OA)	Previously reported RA implication (ref)
CAPN8	1	Body	calpain 8	-0.52	N/A	
SERPINA5	1	TSS1500	serpin peptidase inhibitor, clade A member 5	-0.40	N/A	
FCGBP	1	Body	Fc fragment of IgG binding protein	-0.35	0.34	Detected in plasma sera related with autoimmunity [11600203]
HOXA11	13	TSS1500	homeobox A11	-0.30	0.40	
IL6R	1	Body	interleukin 6 receptor	-0.29	N/A	Its ligand (IL6) is overexpressed in RA [11053081]
S100A14	3	TSS1500	S100 calcium binding protein A14	-0.27	N/A	Involved in invasion through MMP2 (elevated in RA plasma) [22451655]
TMEM51	2	5'UTR	transmembrane protein 51	-0.27	4.21	
CSGALNACT1	3	TSS200	chondroitin sulfate N-acetylgalactosaminyltransferase 1	-0.22	0.48	Involved in cartilage development and endocondral ossification [20812917] and [21148564]
COL14A1	2	Body	collagen, type XIV, alpha 1	-0.22	3.76	
CD74	8	TSS1500	CD74 molecule	-0.22	N/A	Initiates MIF signal transduction (levels related with RA course) [12782713]
TNFAIP8	3	Body	tumor necrosis factor, alpha-induced protein 8	-0.20	3.75	Negative regulator of innate and adaptative immunity [18455983]
TNFRSF8	1	Body|TSS 1500	tumor necrosis factor receptor superfamily, member 8	-0.19	2.93	Its overexpression contributes to proinflammatory immune responses [22628304]
KCNJ15	2	5'UTR|TSS 200	potassium inwardly-rectifying channel, subfamily J, member 15	-0.15	5.51	
CCR6	1	TSS1500	chemokine (C-C motif) receptor 6	0.23	0.67	Migration, proliferation, and MMPs production [11472439] and [15593223]
DPP4	1	TSS200	dipeptidyl-peptidase 4	0.23	N/A	Its Inhibition increases cartilage invasion by RASF [20155839]
PRKCZ	17	Body|TSS 1500	protein kinase C, zeta	0.25	N/A	Inactivates syndecan-4 (integrin co-receptor), reducing DC motility.[20607801]
HLA-DRB5	3	Body	major histocompatibility complex, class II, DR beta 5	0.26	3.16	SNP associated with cutaneous manifestations rheumatoid vasculitis [22641591]
ALOX5AP	1	TSS1500	arachidonate 5-lipoxygenase-activating protein	0.29	N/A	Deficit of this molecule ameliorates symptoms in CIA [9091585]
BCL6	2	Body	B-cell CLL/lymphoma 6	0.30	N/A	RA synovial T cells express BCL6, potent B cell regulator [18975336]
SPTBN1	2	TSS1500	spectrin, beta, non-erythrocytic 1	0.27	0.23	Associated with CD43 abrogates T cell activation [12354383]
HOXC4	13	5'UTR|TSS 1500	homeobox C4	0.4	N/A	

Table 1. Selection of genes differentially methylated and/or expressed in RASF vs. OASF, and previously described implications in RA.

We also performed an analysis to identify genomic clusters of differentially methylated CpGs, which highlighted several regions of consecutive CpGs that are hypomethylated or hypermethylated in RASFs compared with OASFs. Among hypermethylated CpG clusters in RASFs we identified TMEM51 and PTPRN2. With respect to hypomethylated genes, up to nine clustered CpGs were hypomethylated around the transcription start sites of HOXA11 (Figure 2C, left) and nine in CD74, the major histocompatibility complex, class II invariant chain-encoding gene. CD74 levels have been reported to be higher in synovial tissue samples from patients with RA than in tissue from patients with osteoarthritis [49]. HOXA11 was considered another interesting gene, as HOX genes are a direct target of EZH2, a Polycomb group protein involved in differentiation and in establishing repressive marks, including histone H3K27me3 and DNA methylation, under normal and pathological conditions. In fact, additional HOXA genes were identified as being differentially methylated between

RASFs and OASFs (Table 1), suggesting that the Polycomb group differentiation pathway may be responsible for these differences.

To validate our analysis, we used bisulfite pyrosequencing of selected genes (Supp. Fig 2). In all cases, pyrosequencing of individual genes confirmed the results of the analysis. In fact, comparison of the bead array and pyrosequencing methylation data (Figure 2C, center and right) showed an excellent correlation, supporting the validity of our analysis. Additional genes that were subjected to pyrosequencing analysis included CAPN8 and IL6R, both of which were hypomethylated in RASFs, and DPP4 and HOXC4 in the hypermethylated group (Figure 2D and 2E). In all cases, the analysis was validated by pyrosequencing in a larger cohort of samples.

3.2. Integration of DNA methylation data with expression data from RASFs and OASFs

DNA methylation is generally associated with gene repression, particularly when it occurs at promoter CpG islands. However, DNA methylation changes at promoters with low CpG density can also regulate transcription, and changes in gene bodies also affect transcriptional activity [4], although they do not necessarily repress it. We therefore integrated our DNA methylation data with a high-throughput expression analysis of RASFs and OASFs from a recent study (GSE29746) [31]. To integrate expression data with our methylation results, we first reanalyzed the expression data as described in the Materials and Methods. Applying the threshold criterion of a value of $p < 0.01$, we identified 3470 probes differentially expressed between RASFs and OASFs (for which $FC > 2$ or < 0.6) (Figure 3A). We then compared the results from the analysis of the expression arrays with the DNA methylation data. Our analysis showed that 208 annotated CpGs displayed an inverse correlation between expression and methylation levels (Figure 3B and Supplementary Table 3). To examine the relationship between methylation and gene expression further, we also performed an analysis focusing on the relative position of the CpG site that undergoes a significant change in methylation. We found that genes with a methylation change at the TSS or the 5'UTR generally exhibited an inverse correlation between DNA methylation and gene expression (Figure 3C), whereby an increase in methylation tended to be accompanied by a decrease in expression. Curiously, this relationship is positive when looking at CpGs containing probes located at gene bodies with a significant methylation change (Figure 3C). Figure 3D shows two examples of an inverse correlation between DNA methylation and expression data .

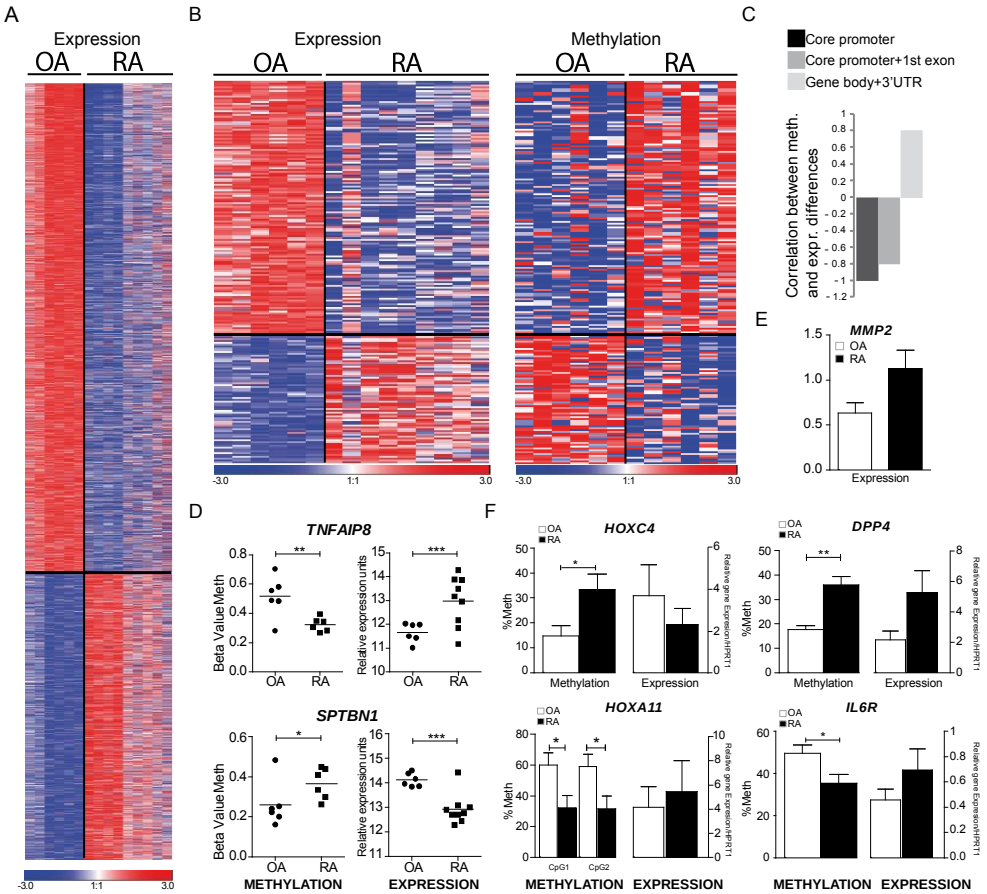


Figure 3. Integration of DNA methylation with expression data. (A) Heatmaps including the reanalysis of expression data (GSE29746) for a collection of RASF and OASF samples. The scale at the bottom distinguishes upregulated (red) and downregulated (blue) genes. (B) Heatmap comparison of inversely correlated expression and methylation. (C) Correlation between differences in DNA methylation and expression of associated genes, focusing on different regions where the CpG sites are located in the probe, core promoter, core + 1st exon and gene body. (D) Illustrative examples of genes featuring an inverse correlation between methylation and expression. (E) Validation by quantitative RT-PCR of MMP2, previously described as overexpressed in RASFs. (F) Examples of genes whose expression data were validated by quantitative RT-PCR.

We performed quantitative RT-PCR to investigate the expression status of several of the genes displaying a change in DNA methylation in the set of samples used in this study. This analysis included several of the genes mentioned above, as well as others, like MMP2, for which increased expression in RASFs has been described [50]. Our analysis confirmed the elevated levels for this gene in our collection of RASFs (Figure 3E). We also observed an inverse correlation between

DNA methylation and expression for genes like HOXC4, HOXA11, CAPN8 and IL6R (Figure 3F), although genes like DPP4 did show a direct relationship. Specifically, we found that hypermethylated DPP4 had higher levels of expression in RASFs than in OASFs (Figure 3F). Elevated levels for DPP4 are compatible with the data obtained by other researchers [51]. However, it also indicates that for some genes, other mechanisms contribute more to their expression levels than do DNA methylation changes.

3.3. miRNA screening in RASF and OASF

Changes in expression levels can certainly be due to transcriptional control, like that determined by epigenetic changes at gene promoters, DNA methylation, or differences in transcription factor binding. At the post-transcriptional level, miRNAs are recognized as being major players in gene expression regulation. We compared the expression levels of miRNAs in pooled RASF and OASF RNA samples. The screening led to the identification of a number of miRNAs that are overexpressed in RASFs with respect to OASFs, as well as downregulated miRNAs (Figure 4A). Among the most upregulated and downregulated miRNAs, we identified several that have been previously associated with relevant or related events like miR-203 [20], which is upregulated in RASFs with respect to OASFs, and miR-124, which is downregulated in RASFs [52] (Figure 4B). Other additional miRNAs identified in previous work in relation to RA include miR-146a and miR-34a (Figure 4A). As indicated above, the overlap with other studies highlights the robustness of our data. However, it is likely that the analysis of a limited number of samples introduces a bias associated with specific characteristics of the samples in the studied cohort. Integrated analysis can help to identify relevant targets. We then performed quantitative PCR to validate a selection of the miRNAs in the entire cohort. Examples of miR-625*, downregulated in RASF, and miR-551b, upregulated in RASF, are shown in Figure 4C.

As explained, miRNA-dependent control is associated with the expression control of a number of targets either by inducing direct mRNA degradation or through translational inhibition [53]. Accumulated evidence has shown that most miRNA targets are affected at the mRNA levels, and therefore comparison of mRNA expression array and miRNA expression data is useful for identifying and evaluating the impact of miRNA misregulation at the mRNA levels.

To explore this aspect, we investigated the relationship between miRNA expression differences between RASFs and OASFs and their involvement in gene control by looking at levels of their potential targets. To this end, we obtained a matrix with the potential targets for each of the five miRNAs most strongly

upregulated and downregulated in RASFs relative to OASFs. We considered bona fide putative targets those predicted by at least four databases. As before, we used the expression microarrays data for RASFs and OASFs generated in another study (GSE29746) [31].

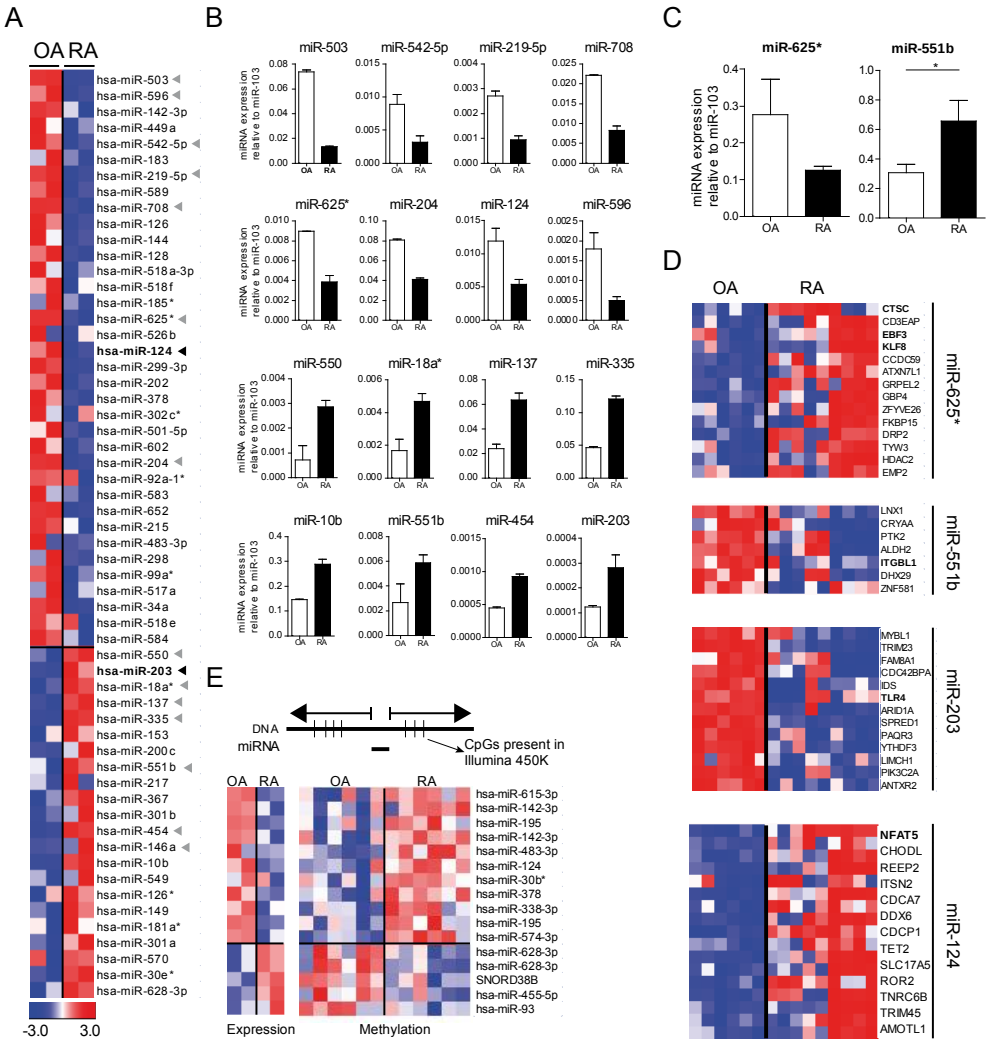


Figure 4. miRNA dysregulation in RASF. (A) Heatmaps showing the miRNA expression data for pooled RASFs and OASFs. miRNAs previously described as dysregulated in RASFs are highlighted with a black arrow. Those represented in the adjacent section are highlighted with a grey arrow. The scale at the bottom distinguishes distinguishing upregulated (red) and downregulated (blue) genes. (B) Examples of the most upregulated and downregulated miRNAs in RASFs with respect to OASFs. (C) Validation of the miRNA data by quantitative RT-PCR. (D) Heatmaps showing the expression levels of putative targets (at least four hits for prediction in miRNA databases) for selected

miRNAs in RASFs and OASFs. (E) Correlation between DNA methylation and miRNA expression data.

When looking at the expression levels of putative targets of selected miRNAs we found more genes with potential effects on the RASF phenotype. These included genes like CTSC, KLF8 or EBF3, which are upregulated in RASFs concomitant with downregulation of miR625* and ITGBL1, which is downregulated in RASF concomitant with upregulation of miR551b. Additional putative targets included TLR4 for miR-203 and NFAT5 for miR-124 (Figure 4D). TLR4 is upregulated in RA and plays a key role in the disease, whereas NFAT5 is a critical regulator of inflammatory arthritis.

3.4. Integrated analysis of both miRNAs and DNA methylation reveals multiple layers of regulation in genes relevant to RA pathogenesis

We performed two separate analyses to explore the potential connection between miRNA and DNA methylation control for genes associated with the RASF phenotype. The first analysis focused on the potential regulation of miRNAs by DNA methylation. DNA methylation can also repress the expression of miRNAs, since miRNA-associated promoters are subjected to similar mechanisms of transcriptional control as protein-coding genes. We compared the data from the bead array analysis with miRNA expression data. Our analysis showed 11 downregulated miRNAs, like miR-124, that were located near CpG sites and were hypermethylated in RASFs. Only four upregulated miRNAs were located near a CpG site hypomethylated in RASFs (Figure 4E).

The second analysis investigated the potential influence of DNA methylation and miRNA control on specific targets. As explained above, differences in expression patterns between RASFs and OASFs could be due to altered mechanisms of control at the epigenetic level, like DNA methylation, or at the post-transcriptional level. We generated a list of selected genes whose expression patterns differed significantly between RASFs and OASFs. Then we matched the expression data with our DNA methylation data from bead arrays and with a selection of miRNAs that might target those genes (as predicted at least by four databases) and that have significant differences in expression between RASFs and OASFs. This yielded a list of genes potentially regulated by DNA methylation, at the transcriptional level, and targeted by miRNAs, at the post-transcriptional level. The list of genes comprises six groups (Supplementary Table 4): i) downregulated genes in which hypermethylation concurs with overexpression of a miRNA that targets them. Methylation and miRNA regulate in the same repressive direction in this group; ii) downregulated genes in which

hypomethylation is potentially overcome by co-occurrence of upregulation of a miRNA that targets them. In this group, miRNA control is potentially the predominant mechanism; iii) downregulated genes in which hypermethylation predominates over downregulation of miRNAs that potentially target them; iv) upregulated genes in which hypomethylation concurs with downregulation of a miRNA that targets them; v) upregulated genes in which hypermethylation is overcome by downregulation of a miRNA that targets them, and, vi) upregulated genes in which hypermethylation predominates over upregulation of a miRNA that targets them. Integrated analysis would require further validation to provide bona fide targets determined by both regulatory mechanisms. However, this novel approach to integrating miRNA and DNA methylation analysis provides a new workflow for exploring the multiple layers of gene dysregulation in RA in greater depth.

4. DISCUSSION

In this study we have identified novel dysregulated targets in rheumatoid arthritis (RA) synovial fibroblasts at the DNA methylation and miRNA expression levels. By using a double approach and integrated analysis of the DNA methylation, miRNA expression and mRNA expression data we have established a new pipeline for investigating the complexity of gene dysregulation in the context of this disease when using primary samples. As indicated above, dysregulation of gene expression arises from a combination of factors, including genetic polymorphisms in genes associated with regulatory roles and miRNAs, environmental factors and their combined effect on transcription factor function and epigenetic profiles, like DNA methylation and histone modification profiles. Understanding the relationship between different elements of regulation is key not only for understanding their intricate connections within the disease but also in the higher propensity to associated disorders [54]. DNA methylation-associated regulation and miRNA control are major regulatory elements and provide useful targets and markers of gene dysregulation in disease. In the context of RASFs, a few studies have previously shown the existence of genes with DNA methylation alterations in RASFs. Most of these have involved examining candidate genes. Examples include the identification of the TNFRSF25 gene (encoding DDR9), which is hypermethylated at its CpG island in synovial cells of RA patients [16], and CXCL12 upregulated and hypomethylated in RASFs [15]. More recently, Firestein and colleagues [17] took an array-based approach to identify hypomethylated and hypermethylated genes in RASFs. Regarding miRNA profiling in RASFs, several studies have demonstrated specific roles

for miRNAs that are dysregulated in RA synovial tissues [18-20, 55]. However, there were no previous systematic efforts to combine analyses of these two types of mechanisms in the context of RA.

To the best of our knowledge, our study constitutes the first attempt to integrate high-throughput omics data from primary samples in the context of RA. The need of integrating several levels of regulation is relevant for several reasons: first, from a biological point of view, it is essential to understand the molecular mechanisms underlying aberrant changes in gene expression associated with the acquisition of the aggressive phenotype of RASFs; second, from a more translational point of view, understanding multiple levels of regulation of target genes that undergo dysregulation in RA, could potentially allow to predict their behavior following the use of specific therapeutic compounds. It can also serve to make a better use of them as clinical markers of disease onset, progression or response to therapy.

Our analysis of individual datasets not only has allowed us to confirm changes described by others but also to determine novel genes with altered DNA methylation patterns, including MMP20, RASGRF2, EGF, TIMP2 and others. Most importantly we have identified new genes that are relevant to the RA phenotype. This includes IL6R, which is well known as an overexpressed gene in RASFs and a target for antibody-based therapy [44]. Additional targets include CAPN8, TNFAIP8, CD74 and CCR6. Methylation alterations in RASFs occur at promoter CpG islands in genes like DPP4 or HOXC4, and downstream of the TSS in genes like CAPN8 and IL6R. This last observation is in agreement with recent reports showing that gene expression can be also affected by methylation changes at gene bodies [4, 5]. In any case, we have found a canonical inverse relationship between DNA methylation and expression status for a subset of more than 200 genes. At the miRNA level, analysis of the expression dataset has allowed us to validate previously described miRNAs, like miR-203 and miR-124, as well as identifying novel miRNAs, like miR-503, miR-625*, miR-551b, and miR-550, that are potentially associated with dysregulated targets in RASFs.

Integrative analysis has been carried out at different levels. Firstly, the combined analysis of DNA methylation and expression data generated a list of genes in which methylation changes were inversely correlated with expression changes. This list of genes potentially contains those regulated through DNA methylation in a canonical manner, where DNA methylation associates with gene repression (Supplementary Table 3). Secondly, we also studied the potential relationship

between expression changes and miRNA expression changes that potentially target them (as defined by the cumulative use of miRNA target prediction databases). In this case, we identified a number of genes undergoing expression changes in RASFs that are potentially targeted by concomitantly dysregulated miRNAs.

Another level of integration is achieved by looking at genes that may be targeted or regulated by the combined action of miRNA and DNA methylation. Thus, we explored the potential combined effect of miRNAs and DNA methylation in genes undergoing expression changes in RASFs (Supplementary Table 2). Our analysis revealed gene targets in which methylation and miRNA control possibly concur in direction or have antagonistic effects. This classification of genes in different groups is important because pharmacological compounds or other experimental approaches influencing one of the mechanism (DNA methylation) but not the other (miRNA expression) or viceversa, would have to consider the existence of multiple levels of regulation for interpreting the outcome of such treatment. Finally, by looking at the potential control of miRNA expression by DNA methylation, we identified a further regulatory mechanism for several miRNAs, including miR-124. In this case, hypermethylation of a specific miRNA promoter, would have a positive effect on the expression levels of its targets, and, for instance, pharmacological demethylation of that miRNA would result in overexpression of the miRNA and downregulation of its targets.

As indicated above, epigenetic profiles and miRNA expression patterns are cell type-specific. The need to use primary samples for the target tissue or cell type of a particular disease is usually a limitation to performing epigenetic or miRNA analysis, given the access to small amounts of tissue or cells that can be obtained in most cases. The reduced number of laboratories with access to RASFs, OASFs or SF from normal individuals is a good reflection of such limitation. Genetic analysis of genetically complex diseases does not have such a limitation, since in most cases can be done with peripheral blood. In this sense, the use of integrative approaches to investigate epigenetic and miRNA-mediated control of a limited set of samples overcomes partially this obstacle by providing extra sets of data for internal validation within a small cohort of samples and an increase of the robustness of the analysis.

In conclusion, our study highlights the need of investigating the multiple layers of regulation at the transcriptional and post-transcriptional levels as well as integrating the datasets during the analysis. As targets for therapy, it is important to understand the intricate connections between the various control mechanisms and

to consider the existence of both processes that operate in the same direction or have antagonistic effects. The use of integrative approaches will also be necessary for the rational design of targeted therapies as well as for the use of different clinical markers for the classification. In this sense, the workflow designed in this study has allowed us to identify novel targets and their regulatory mechanism in RASF and opens up a number of possibilities for future research on epigenetics aspects on RA.

Acknowledgements

We would like to thank Dr. Gary Firestein for sharing the raw data of his DNA methylation study with us. We would also like to thank José Luis Pablos for his valuable feedback on his expression dataset. This work was supported by grant SAF2011-29635 from the Spanish Ministry of Science and Innovation, grant from Fundación Ramón Areces and grant 2009SGR184 from AGAUR (Catalan Government). LR is supported by a PFIS predoctoral fellowship and AI was supported by a AGAUR predoctoral fellowship. NL-B acknowledges funding from the Spanish Ministry of Science and Technology (grant number SAF2009-06954)

REFERENCES

[1] Lefevre S, Knedla A, Tennie C, Kampmann A, Wunrau C, Dinser R et al. Synovial fibroblasts spread rheumatoid arthritis to unaffected joints. Nat Med, 2009;15:1414-20.

[2] Tolboom TC, van der Helm-Van Mil AH, Nelissen RG, Breedveld FC, Toes RE, Huizinga TW. Invasiveness of fibroblast-like synoviocytes is an individual patient characteristic associated with the rate of joint destruction in patients with rheumatoid arthritis. Arthritis Rheum, 2005;52:1999-2002.

[3] Deaton AM, Bird A. CpG islands and the regulation of transcription. Genes Dev, 2011;25:1010-22.

[4] Ball MP, Li JB, Gao Y, Lee JH, LeProust EM, Park IH et al. Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells. Nat Biotechnol, 2009;27:361-8.

[5] Rauch TA, Wu X, Zhong X, Riggs AD, Pfeifer GP. A human B cell methylome at 100-base pair resolution. Proc Natl Acad Sci U S A, 2009;106:671-8.

[6] Tili E, Michaille JJ, Costinean S, Croce CM. MicroRNAs, the immune system and rheumatic disease. Nat Clin Pract Rheumatol, 2008;4:534-41.

[7] Selmi C, Lu Q, Humble MC. Heritability versus the role of the environment in autoimmunity. J Autoimmun, 2012;39:249-52.

[8] Ballestar E. Epigenetics lessons from twins: prospects for autoimmune disease. Clin Rev Allergy Immunol, 2010;39:30-41.

[9] Miller FW, Pollard KM, Parks CG, Germolec DR, Leung PS, Selmi C et al. Criteria for environmentally associated autoimmune diseases. J Autoimmun, 2012;39:253-8.

[10] Miller FW, Alfredsson L, Costenbader KH, Kamen DL, Nelson LM, Norris JM et al. Epidemiology of environmental exposures and human autoimmune diseases: Findings from a National Institute of Environmental Health Sciences Expert Panel Workshop. J Autoimmun, 2012;39:259-71.

[11] Selmi C, Leung PS, Sherr DH, Diaz M, Nyland JF, Monestier M et al. Mechanisms of environmental influence on human autoimmunity: A national institute of environmental health sciences expert panel workshop. J Autoimmun, 2012;39:272-84.

[12] Germolec D, Kono DH, Pfau JC, Pollard KM. Animal models used to examine the role of the environment in the development of autoimmune disease: Findings from an NIEHS Expert Panel Workshop. J Autoimmun, 2012;39:285-93.

[13] Neidhart M, Rethage J, Kuchen S, Kunzler P, Crowl RM, Billingham ME et al. Retrotransposable L1 elements expressed in rheumatoid arthritis synovial tissue:

association with genomic DNA hypomethylation and influence on gene expression. Arthritis Rheum, 2000;43:2634-47.

[14] Nile CJ, Read RC, Akil M, Duff GW, Wilson AG. Methylation status of a single CpG site in the IL6 promoter is related to IL6 messenger RNA levels and rheumatoid arthritis. Arthritis Rheum, 2008;58:2686-93.

[15] Karouzakis E, Rengel Y, Jungel A, Kolling C, Gay RE, Michel BA et al. DNA methylation regulates the expression of CXCL12 in rheumatoid arthritis synovial fibroblasts. Genes Immun, 2011;12:643-52.

[16] Takami N, Osawa K, Miura Y, Komai K, Taniguchi M, Shiraishi M et al. Hypermethylated promoter region of DR3, the death receptor 3 gene, in rheumatoid arthritis synovial cells. Arthritis Rheum, 2006;54:779-87.

[17] Nakano K, Whitaker JW, Boyle DL, Wang W, Firestein GS. DNA methylome signature in rheumatoid arthritis. Ann Rheum Dis, 2012.

[18] Niederer F, Trenkmann M, Ospelt C, Karouzakis E, Neidhart M, Stanczyk J et al. Down-regulation of microRNA-34a* in rheumatoid arthritis synovial fibroblasts promotes apoptosis resistance. Arthritis Rheum, 2012;64:1771-9.

[19] Nakamachi Y, Kawano S, Takenokuchi M, Nishimura K, Sakai Y, Chin T et al. MicroRNA-124a is a key regulator of proliferation and monocyte chemoattractant protein 1 secretion in fibroblast-like synoviocytes from patients with rheumatoid arthritis. Arthritis Rheum, 2009;60:1294-304.

[20] Stanczyk J, Ospelt C, Karouzakis E, Filer A, Raza K, Kolling C et al. Altered expression of microRNA-203 in rheumatoid arthritis synovial fibroblasts and its role in fibroblast activation. Arthritis Rheum, 2011;63:373-81.

[21] Bibikova M, Lin Z, Zhou L, Chudin E, Garcia EW, Wu B et al. High-throughput DNA methylation profiling using universal bead arrays. Genome Res, 2006;16:383-93.

[22] Dessau RB, Pipper CB. ["R"--project for statistical computing]. Ugeskr Laeger, 2008;170:328-30.

[23] Du P, Kibbe WA, Lin SM. lumi: a pipeline for processing Illumina microarray. Bioinformatics, 2008;24:1547-8.

[24] Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S et al. Bioconductor: open software development for computational biology and bioinformatics. Genome Biol, 2004;5:R80.

[25] Falcon S, Gentleman R. Using GOstats to test gene lists for GO term association. Bioinformatics, 2007;23:257-8.

[26] Smyth GK. Limma: linear models for microarray data. Bioinformatics and Computational Biology Solutions using R and Bioconductor, 2005:397-420.

[27] Maksimovic J, Gordon L, Oshlack A. SWAN: Subset-quantile Within Array Normalization for Illumina Infinium HumanMethylation450 BeadChips. Genome Biol, 2012;13:R44.

[28] Touleimat N, Tost J. Complete pipeline for Infinium((R)) Human Methylation 450K BeadChip data processing using subset quantile normalization for accurate DNA methylation estimation. Epigenomics, 2012;4:325-41.

[29] Aryee MJ, Wu Z, Ladd-Acosta C, Herb B, Feinberg AP, Yegnasubramanian S et al. Accurate genome-scale percentage DNA methylation estimates from microarray data. Biostatistics, 2011;12:197-210.

[30] Herman JG, Graff JR, Myohanen S, Nelkin BD, Baylin SB. Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands. Proc Natl Acad Sci U S A, 1996;93:9821-6.

[31] Del Rey MJ, Usategui A, Izquierdo E, Canete JD, Blanco FJ, Criado G et al. Transcriptome analysis reveals specific changes in osteoarthritis synovial fibroblasts. Ann Rheum Dis, 2012;71:275-80.

[32] Lewis BP, Burge CB, Bartel DP. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell, 2005;120:15-20.

[33] Krek A, Grun D, Poy MN, Wolf R, Rosenberg L, Epstein EJ et al. Combinatorial microRNA target predictions. Nat Genet, 2005;37:495-500.

[34] Kertesz M, Iovino N, Unnerstall U, Gaul U, Segal E. The role of site accessibility in microRNA target recognition. Nat Genet, 2007;39:1278-84.

[35] Griffiths-Jones S, Grocock RJ, van Dongen S, Bateman A, Enright AJ. miRBase: microRNA sequences, targets and gene nomenclature. Nucleic Acids Res, 2006;34:D140-4.

[36] Betel D, Wilson M, Gabow A, Marks DS, Sander C. The microRNA.org resource: targets and expression. Nucleic Acids Res, 2008;36:D149-53.

[37] Wang X, El Naqa IM. Prediction of both conserved and nonconserved microRNA targets in animals. Bioinformatics, 2008;24:325-32.

[38] Vergoulis T, Vlachos IS, Alexiou P, Georgakilas G, Maragkakis M, Reczko M et al. TarBase 6.0: capturing the exponential growth of miRNA targets with experimental support. Nucleic Acids Res, 2011;40:D222-9.

[39] Xiao F, Zuo Z, Cai G, Kang S, Gao X, Li T. miRecords: an integrated resource for microRNA-target interactions. Nucleic Acids Res, 2009;37:D105-10.

[40] Yang JH, Li JH, Shao P, Zhou H, Chen YQ, Qu LH. starBase: a database for exploring microRNA-mRNA interaction maps from Argonaute CLIP-Seq and Degradome-Seq data. Nucleic Acids Res, 2011;39:D202-9.

[41] Kozomara A, Griffiths-Jones S. miRBase: integrating microRNA annotation and deep-sequencing data. Nucleic Acids Res, 2011;39:D152-7.

[42] Al-Shahrour F, Diaz-Uriarte R, Dopazo J. FatiGO: a web tool for finding significant associations of Gene Ontology terms with groups of genes. Bioinformatics, 2004;20:578-80.

[43] Sturn A, Quackenbush J, Trajanoski Z. Genesis: cluster analysis of microarray data. Bioinformatics, 2002;18:207-8.

[44] Thompson CA. FDA approves tocilizumab to treat rheumatoid arthritis. Am J Health Syst Pharm, 2010;67:254.

[45] Sun H, Gong S, Carmody RJ, Hilliard A, Li L, Sun J et al. TIPE2, a negative regulator of innate and adaptive immunity that maintains immune homeostasis. Cell, 2008;133:415-26.

[46] Swan C, Duroudier NP, Campbell E, Zaitoun A, Hastings M, Dukes GE et al. Identifying and testing candidate genetic polymorphisms in the irritable bowel syndrome (IBS): association with TNFSF15 and TNFalpha. Gut, 2012.

[47] Yazbeck R, Howarth GS, Abbott CA. Dipeptidyl peptidase inhibitors, an emerging drug class for inflammatory disease? Trends Pharmacol Sci, 2009;30:600-7.

[48] Ruth JH, Shahrara S, Park CC, Morel JC, Kumar P, Qin S et al. Role of macrophage inflammatory protein-3alpha and its ligand CCR6 in rheumatoid arthritis. Lab Invest, 2003;83:579-88.

[49] Waldburger JM, Palmer G, Seemayer C, Lamacchia C, Finckh A, Christofilopoulos P et al. Autoimmunity and inflammation are independent of class II transactivator type PIV-dependent class II major histocompatibility complex expression in peripheral tissues during collagen-induced arthritis. Arthritis Rheum, 2011;63:3354-63.

[50] Li G, Zhang Y, Qian Y, Zhang H, Guo S, Sunagawa M et al. Interleukin-17A promotes rheumatoid arthritis synoviocytes migration and invasion under hypoxia by increasing MMP2 and MMP9 expression through NF-kappaB/HIF-1alpha pathway. Mol Immunol, 2012;53:227-36.

[51] Solau-Gervais E, Zerimech F, Lemaire R, Fontaine C, Huet G, Flipo RM. Cysteine and serine proteases of synovial tissue in rheumatoid arthritis and osteoarthritis. Scand J Rheumatol, 2007;36:373-7.

[52] Kawano S, Nakamachi Y. miR-124a as a key regulator of proliferation and MCP-1 secretion in synoviocytes from patients with rheumatoid arthritis. Ann Rheum Dis, 2011;70 Suppl 1:i88-91.

[53] He L, Hannon GJ. MicroRNAs: small RNAs with a big role in gene regulation. Nat Rev Genet, 2004;5:522-31.

[54] Ngalamika O, Zhang Y, Yin H, Zhao M, Gershwin ME, Lu Q. Epigenetics, autoimmunity and hematologic malignancies: A comprehensive review. J Autoimmun, 2012;39:451-65.

[55] Nakasa T, Miyaki S, Okubo A, Hashimoto M, Nishida K, Ochi M et al. Expression of microRNA-146 in rheumatoid arthritis synovial tissue. Arthritis Rheum, 2008;58:1284-92.

[56] Kobayashi K, Yagasaki M, Harada N, Chichibu K, Hibi T, Yoshida T et al. Detection of Fcgamma binding protein antigen in human sera and its relation with autoimmune diseases. Immunol Lett, 2001;79:229-35.

[57] Nishimoto N, Kishimoto T, Yoshizaki K. Anti-interleukin 6 receptor antibody treatment in rheumatic disease. Ann Rheum Dis, 2000;59 Suppl 1:i21-7.

[58] Watanabe Y, Takeuchi K, Higa Onaga S, Sato M, Tsujita M, Abe M et al. Chondroitin sulfate N-acetylgalactosaminyltransferase-1 is required for normal cartilage development. Biochem J, 2010;432:47-55.

[59] Sato T, Kudo T, Ikehara Y, Ogawa H, Hirano T, Kiyohara K et al. Chondroitin sulfate N-acetylgalactosaminyltransferase 1 is necessary for normal endochondral ossification and aggrecan metabolism. J Biol Chem, 2011;286:5803-12.

[60] Leng L, Metz CN, Fang Y, Xu J, Donnelly S, Baugh J et al. MIF signal transduction initiated by binding to CD74. J Exp Med, 2003;197:1467-76.

[61] Simhadri VL, Hansen HP, Simhadri VR, Reiners KS, Bessler M, Engert A et al. A novel role for reciprocal CD30-CD30L signaling in the cross-talk between natural killer and dendritic cells. Biol Chem, 2012;393:101-6.

[62] Matsui T, Akahoshi T, Namai R, Hashimoto A, Kurihara Y, Rana M et al. Selective recruitment of CCR6-expressing cells by increased production of MIP-3 alpha in rheumatoid arthritis. Clin Exp Immunol, 2001;125:155-61.

[63] Ospelt C, Mertens JC, Jungel A, Brentano F, Maciejewska-Rodriguez H, Huber LC et al. Inhibition of fibroblast activation protein and dipeptidylpeptidase 4 increases cartilage invasion by rheumatoid arthritis synovial fibroblasts. Arthritis Rheum, 2010;62:1224-35.

[64] Buhligen J, Himmel M, Gebhardt C, Simon JC, Ziegler W, Averbeck M. Lysophosphatidylcholine-mediated functional inactivation of syndecan-4 results in decreased adhesion and motility of dendritic cells. J Cell Physiol, 2010;225:905-14.

[65] Nishimura WE, Costallat LT, Fernandes SR, Conde RA, Bertolo MB. Association of HLA-DRB5*01 with protection against cutaneous manifestations of rheumatoid vasculitis in Brazilian patients. Rev Bras Reumatol, 2012;52:366-74.

[66] Griffiths RJ, Smith MA, Roach ML, Stock JL, Stam EJ, Milici AJ et al. Collagen-induced arthritis is reduced in 5-lipoxygenase-activating protein-deficient mice. J Exp Med, 1997;185:1123-9.

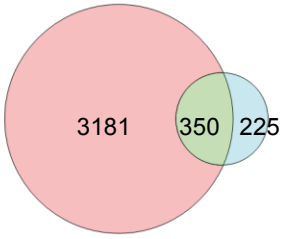
[67] Manzo A, Vitolo B, Humby F, Caporali R, Jarrossay D, Dell'accio F et al. Mature antigen-experienced T helper cells synthesize and secrete the B cell chemoattractant CXCL13 in the inflammatory environment of the rheumatoid joint. Arthritis Rheum, 2008;58:3377-87.

[68] Pradhan D, Morrow J. The spectrin-ankyrin skeleton controls CD45 surface display and interleukin-2 production. Immunity, 2002;17:303-15.

Supplementary Figure 1.

Data comparison between Firenstein dataset and Ballestar's.

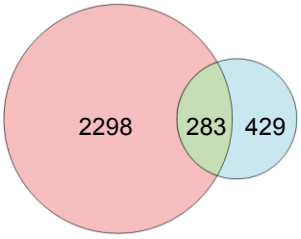
Hypomethylated genes



Nakano K *et al* [16] de la Rica L *et al*

Gene	Gene	Gene	Gene	Gene	Gene	Gene	Gene
ABI1	CD74	FGFRL1	KATNAL2	MMP20	PPP1R10	SPOP	WARS
ACACB	CDC42EP3	FGGY	KCNJ15	MOBKL1A	PPP3CA	SPTLC2	WIPF1
ACAN	CDH13	FLJ12825	KDELC2	MPP7	PPT2	SRGAP1	WWP2
ACTN1	CFDP1	FLJ22536	KDM4B	MRPS27	PRDM8	SRGAP3	WWTR1
ADAMTS16	CHST10	FLT1	KIAA0196	MSC	PRKCA	SSH1	YPEL2
ADARB2	CHST11	FLVCR2	KIAA0317	MYO10	PRKCH	STEAP1	ZBTB20
AKAP2	CHST15	FNIP2	KIAA1409	MYO18A	PRKCZ	STK10	ZNF83
ALDH7A1	CHSY1	FOXP1	KIAA1549	MYOM2	PRR4	SV2C	
ALX4	CLDN14	FRMD4B	KIRREL3	NALCN	PRRT1	SYNJ2	
AMOTL1	CLDN20	FRMD6	KLHL23	NAV1	PTPRK	TANC2	
ANKRD11	CLUAP1	FSIP1	KLHL25	NAV2	PTPRN2	TBCD	
ANTXR1	CLYBL	FUT8	KLHL6	NECAB2	RAB11FIP5	TCERG1L	
APCDD1L	CMTM4	FZD6	KLRAQ1	NFATC1	RADIL	TCF7L1	
ARHGEF10	COL11A2	GAP43	KRTAP1-1	NGEF	RAGE	TCIRG1	
ARNT2	COL14A1	GAS7	LAMB3	NOS1AP	RAPGEF4	TEAD1	
ATG7	COL27A1	GFRA1	LATS2	NPAS2	RASA3	TFB1M	
ATP11A	COL5A1	GJB2	LAYN	NR4A3	RASGRF2	TK2	
ATP6V0E1	CRABP1	GPC6	LCLAT1	NRBP2	RASSF4	TLE2	
ATXN1	CRIM1	GPD2	LCP1	NRXN1	RBPMS	TMED10	
BAHCC1	CSGALNA CT1	GPR133	LDLRAD3	NTM	RFPL2	TMEM145	
BCAT1	CUX1	GPR171	LHFPL2	NUDT4	RGL1	TMEM204	
BICC1	CYR61	GPR87	LIF	NUDT4P1	RHOT1	TMEM51	
BOC	DHRS3	GPRC5B	LMF1	OBSCN	RNF220	TNFAIP8	
BTG2	DIP2C	H2AFY	LMO4	ODZ2	RSPO2	TNFRSF10 B	
C10orf11	DISP1	H2AFY2	LOC145845	ODZ4	SBK2	TNFRSF8	
C11orf17	DLEU2	HDAC9	LOC146880	OLFM1	SDK1	TNNI3K	
C11orf30	DNAH7	HEPHL1	LOC340357	PALLD	SDK2	TNNT3	
C18orf45	DNM3	HERC2	LOC728613	PALM2-AKAP2	SERPINA5	TNS3	
C1orf198	DOCK5	HLA-DRB1	LOXL3	PARK7	SETBP1	TPD52L1	
C21orf33	DOK1	HOXA11AS	LRIG1	PAX7	SH3BP2	TPM1	
C22orf15	DYNC1H1	HOXA3	LRRC27	PCCA	SH3BP4	TPO	
C2orf39	EBF2	HOXD8	LRRC33	PCSK6	SH3TC2	TPST1	
C4orf38	EFNA5	HTRA4	LRRFIP2	PDE6A	SHANK2	TRAF2	
C5orf13	EFR3A	ID3	LTC4S	PDGFC	SIM2	TRAF3IP2	
C5orf27	EGFLAM	IFT140	MACF1	PDLIM4	SKI	TRIM26	
C7orf10	EIF2AK4	IGF1R	MACROD1	PDXK	SLAMF8	TRIM40	
C8orf34	EIF2C2	IL12RB1	MAGI1	PEBP4	SLC12A8	TRIO	
C9orf3	ELF5	IL1RL1	MAML2	PEMT	SLC14A1	TRIP4	
C9orf45	EMILIN2	IL31RA	MAP3K5	PFKP	SLC25A37	TSHZ1	
CA10	EPAS1	ILDR1	MAST4	PHC2	SLC2A5	TSHZ2	
CACNA1C	EPHB3	INO80C	MBNL1	PHF11	SLC41A3	TSNARE1	
CAPN10	ERICH1	IRF1	MBNL2	PKNOX2	SLC8A1	TSPAN9	
CAPN13	FAF1	IRF7	MBP	PLCH2	SLIT3	TTLL6	
CBFA2T3	FAM107B	IRS1	MED12L	PLEKHA7	SND1	UBAC2	
CCDC102A	FAM176B	ITGBL1	MED13L	PNMAL2	SNORA74B	UBE2H	
CCDC111	FCGBP	JAKMIP2	MIA	POU2F2	SORBS1	UTRN	
CCDC88C	FGD6	JARID2	MICAL2	PPAN	SORCS2	VOPP1	
CD109	FGF1	JAZF1	MIR2117	PPAN-P2RY11	SOX9	VWA2	
CD28	FGFR2	JUB	MIR365-1	PPM1H	SPIN1	VWCE	

Hypermethylated genes

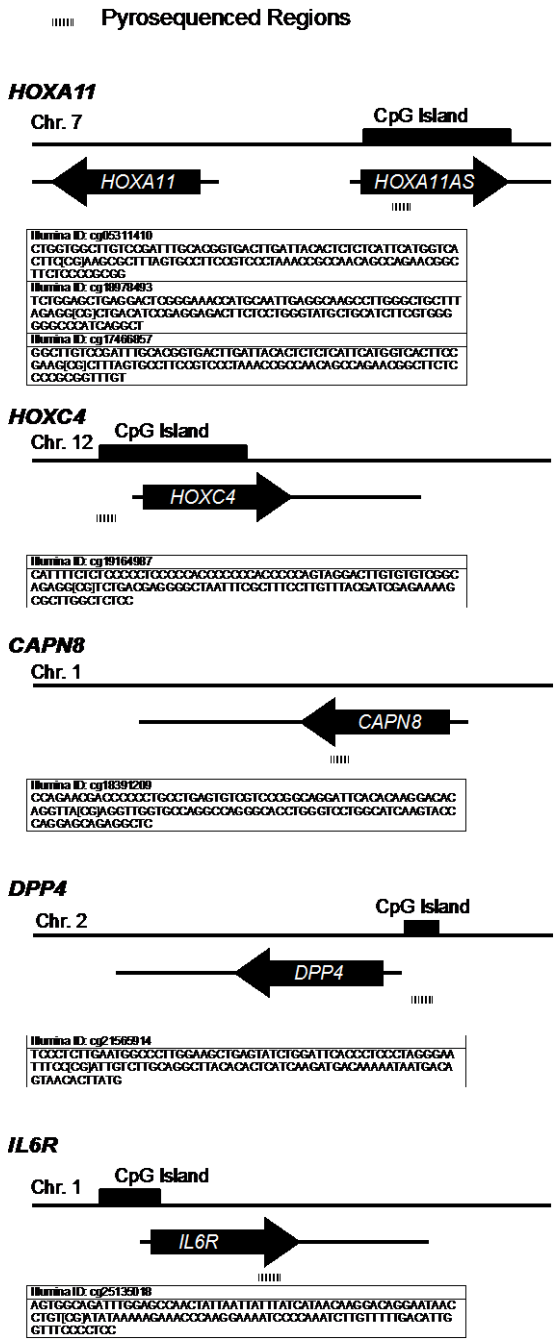


Nakano K *et al* [16] de la Rica L *et al*

Gene	Gene	Gene	Gene	Gene	Gene
AACS	CCDC85C	FMN1	LOH12CR1	PPAP2B	SPTBN1
ABR	CCNY	FOXN3	LRRC16A	PPFIBP2	SREBF1
ACOX2	CD248	FOXP2	LRRC66	PRDM16	SRPK2
ACTN2	CD81	FRMD4A	MAD1L1	PRKAG2	ST3GAL2
ACVR1	CDC42EP4	FYN	MAML2	PRKAR1B	STK32A
ADAMTS2	CDSN	GAB2	MARK3	PRKCZ	STK32C
ADAMTSL2	CELSR1	GALNT9	MC1R	PROCR	SUV420H2
ADARB2	CHRNA7	GAS7	MCF2L	PSORS1C1	SYN3
ADCY2	CLIC5	GATA2	MEIS2	PTGER4	SYNE2
ADPRHL1	CMTM1	GDNF	MGMT	PTPRN2	TBC1D1
AFAP1	COL18A1	GIMAP4	MIR548F5	RALYL	TBCD
AGAP1	CORIN	GLP2R	MRPL42P5	RAP1GAP	TEX2
AGXT2	CPT1A	GNG2	MSI2	RAP1GAP2	TIMP2
AKAP12	CRIP1	GP5	MXRA7	RAPGEF5	TIMP3
AKAP13	CRYGN	GPD1	MYL2	RASGEF1C	TJP1
ALOX5AP	CTBP1	GPR123	MYLK2	RBMS3	TMCO3
ANK1	CUGBP2	GPR133	MYO1D	RBPMS	TMEM120B
ANK2	CUL1	GRIN2A	MYOM1	RERE	TMEM51
ANKRD11	CYP4F12	GSG1L	NAV2	RGMA	TMOD4
ANKRD37	DAB1	HCCA2	NCALD	RNF220	TNFAIP8L3
ANKS1B	DACT2	HDAC4	NCOA7	RNF39	TNIK
ARC	DAPK1	HECW1	NCOR2	RNF44	TNNT3
ARHGEF10 L	DCLK1	HEG1	NCRNA00188	ROR2	TNS1
ARID5B	DIP2C	HHEX	NDFIP2	RP1L1	TNXB
ARPP-21	DLG1	HLA-DRB1	NKAIN2	RPH3AL	TRERF1
ASAM	DLX6AS	HOXC4	NLGN1	RTN4RL1	TRIO
ASB2	DNAJC17	HOXC5	NOVA1	S100A4	TSPAN12
ASCC2	DNAJC22	HOXC6	NPEPL1	SAMD13	TTF1
ATP11A	DPYSL2	HOXC8	NPHS2	SDCBP2	UBE2QL1
ATP2A3	DUSP10	HSPA12A	NT5E	SH3GL1	UMODL1
ATRNL1	DUSP8	IQCE	NXN	SH3PXD2A	UNC84A
AUTS2	EBF1	JARID2	NXPH2	SLC12A8	VAC14
B3GNTL1	EBF2	KCNE1	ODZ2	SLC1A2	WWOX
BAIAP2	EBF3	KCNJ12	OGDH	SLC20A2	ZBTB16
BCAS3	ECE1	KCNMA1	PAPPA	SLC22A18	ZFHX3
BRPF3	ECT2L	KCNMB3	PARD3	SLC25A37	ZFHX4
C10orf11	EGF	KCNQ1DN	PARK2	SLC45A1	ZFPM2
C11orf53	EGFR	KIAA1671	PBX1	SLC45A4	ZMAT4
C12orf34	EMID2	KIAA1949	PCDH9	SLCO3A1	
C14orf179	ESD	KIF1A	PDE4D	SMARCD3	
C3orf21	ESRRA	KIF26A	PDE4DIP	SMOC2	
C3orf50	EXT1	KLB	PEAR1	SND1	
C5orf62	FAM170B	KLHDC8A	PELI2	SNED1	
C7orf50	FAM53B	LEPR	PET112L	SNORD65	
CACNA1H	FDX1	LHX2	PHACTR1	SNTG2	
CACNA2D3	FERD3L	LIMA1	PHLDB2	SNX29	
CACNG1	FHAD1	LIMS2	PIGV	SORCS2	
CCDC14	FLJ42875	LIPC	PLEC1	SOX7	
	FLNC	LOC284837	POLR1A	SPRED3	

Supplementary Figure 2.

Scheme showing the pyrosequenced regions



Supplementary Table 1

Primer sequences used in this study

Name	Type	Sequence
PYRO_HOXA11_PCR_F	Pyrosequencing	AGAGGTAGGTAGGGAAGAT
PYRO_HOXA11_PCR_R	Pyrosequencing	[Btn]ATCCCCCTCCCATAAACT
PYRO_HOXA11_SEQ1	Pyrosequencing	AAGATGAGGGGAGAG
PYRO_HOXA11_SEQ2	Pyrosequencing	GGGTTTTGYGGATTAGTGATAAA
PYRO_HOXA11_SEQ3	Pyrosequencing	GTTTAGGGAYGGAAGGTATTAA
PYRO_HOXA11_SEQ4	Pyrosequencing	GGATAAGTTATTAGGTAGGTATA
PYRO_HOXC4_PCR_F	Pyrosequencing	GGATGGGGGAAGAAATTTGT
PYRO_HOXC4_PCR_R	Pyrosequencing	[Btn]CACTCCCTTTACCCTTTTCAAAT
PYRO_HOXC4_SEQ	Pyrosequencing	TTTATTTTTTTATTTTTAGTAGGAT
PYRO_CAPN8_PCR_F	Pyrosequencing	AGGGGTTTTGTATTTGTTAGAGT
PYRO_CAPN8_PCR_R	Pyrosequencing	[Btn]TCTCCAAAACRACCCCCCTACCT
PYRO_CAPN8_SEQ	Pyrosequencing	GTGTTTTGGTTTGGTATTAAT
PYRO_miR628_PCR_F	Pyrosequencing	AAGGTGATTTTTTTTAGGGAATTTG
PYRO_miR628_PCR_R	Pyrosequencing	[Btn]CTTCCCTTCCACTACCACTCTTACTAA
PYRO_miR628_SEQ	Pyrosequencing	TTATAAAGGAGTAGTATTAGAATAG
PYRO_DPP4_PCR_F	Pyrosequencing	[Btn]AATGTTTAGAGTAGTATTTGGGAAAAAGT
PYRO_DPP4_PCR_R	Pyrosequencing	ACCCTTAAAAACTAAATATCTAAATTCACC
PYRO_DPP4_Seq	Pyrosequencing	AATTCACCCTCCCTA
PYRO_IL6R_PCR_F	Pyrosequencing	TGGATATTAAGTTAAAAGTTAGTGGTAGAT
PYRO_IL6R_PCR_R	Pyrosequencing	[Btn]TTACTTCCCACTTTATAACTAACATACC
PYRO_IL6R_SEQ	Pyrosequencing	ATAAGGATAGGAATAATTTGT
Hs_IL6R_RT_F	RT-PCR	GCTCCACGACTCTGGAAACT
Hs_IL6R_RT_R	RT-PCR	GGACCCCACTCACAAACAAC
Hs_HOXA11_RT_F	RT-PCR	GGCCACACTGAGGACAAGG
Hs_HOXA11_RT_R	RT-PCR	AGAACTCCCGTTCCAGCTCT
Hs_HOXC4_RT_F	RT-PCR	CCAGCAAGCAACCCATAGTC
Hs_HOXC4_RT_R	RT-PCR	ACTTGCTGCCGGGTATAGG
Hs_CAPN8_RT_F	RT-PCR	GTCCATCAGCTTTGGGCTAC
Hs_CAPN8_RT_R	RT-PCR	TAAACTGAGGGCTGGGACAC
Hs_DPP4_RT_F	RT-PCR	CACCGTGGAAGGTTCTTCTG
Hs_DPP4_RT_R	RT-PCR	GCGACTGTCAGCTGTAGCAT
Hs_MMP2_RT_F	RT-PCR	CCCAAAACGGACAAAGAGTT
Hs_MMP2_RT_R	RT-PCR	TGTCCTTCAGCACAAACAGG
Hs_CXCL12_RT_F	RT-PCR	GTGGTCGTGCTGGTCCTC
Hs_CXCL12_RT_R	RT-PCR	AGATGCTTGACGTTGGCTCT

Supplementary Table 4.

Coordinated regulation of genes at the transcriptional (DNA methylation) and/or post-transcriptional (miRNAs) levels in RASF

				DNA Methylation			miRNA		
	Gene Target		mRNA Gene	Diff Meth	Grup	CpGs	miRNA	miR Expression	Regulatory Mechanism
UPREGULATED GENES	**EPHA4**	EPH receptor A4			Body	1	miR10b		**Both**
	KLF11	Kruppel-like factor 11			Body	1			**Both**
	NCOA2	nuclear receptor coactivator 2			Body	1	miR137		**Both**
	EPHA4	EPH receptor A4			Body	1	miR335		**Both**
	PFKFB3	6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 3			Body	4	miR454		**Both**
	RNF145	ring finger protein 145			Body	1			**Both**
	HOXA3	homeobox A3			TSS1500	1	miR10b		**miRNA**
	FURIN	furin (paired basic amino acid cleaving enzyme)			Body	1	miR137		**miRNA**
	CRIM1	cysteine rich transmembrane BMP regulator 1 (chordin-like)			Body	2	miR335		**miRNA**
	SRPR	signal recognition particle receptor (docking protein)			3'UTR	1			**miRNA**
	FRMD6	FERM domain containing 6			5'UTR	2			**miRNA**
	NPTN	neuroplastin			Body	1	miR454		**miRNA**
	AP2A2	adaptor-related protein complex 2, alpha 2 subunit			Body/3'UTR	7			**METH**
	CPD	carboxypeptidase D			Body	1			**METH**
	FBXO9	F-box protein 9			3'UTR	1	miR204		**METH**
	MAPRE2	microtubule-associated protein, RP/EB family, member 2			5'UTR|Body	1			**METH**
	NCOA7	nuclear receptor coactivator 7			5'UTR	1			**METH**
	PTPRJ	protein tyrosine phosphatase, receptor type, J			Body	1			**METH**
	CCND1	cyclin D1			TSS1500	1	miR503		**METH**
	GPM6A	glycoprotein M6A			TSS1500	1			**METH**
	SESN1	sestrin 1			Body	1	miR708		**METH**
DOWNREGULATED GENES	**BNC2**	basonuclin 2			Body	1	miR204		**Both**
	ELOVL6	ELOVL fatty acid elongase 6			Body	1			**miRNA**
	HMGA2	high mobility group AT-hook 2			Body	1	miR204		**miRNA**
	NOVA1	neuro-oncological ventral antigen 1			Body	1			**miRNA**
	HMGA2	high mobility group AT-hook 2			Body	1			**miRNA**
	ZFHX4	zinc finger homeobox 4			5'UTR	1	miR503		**miRNA**
	PSD3	pleckstrin and Sec7 domain containing 3			Body	1	miR335		**METH**

Color	Mean
	Downregulated (miRNA and mRNA expression)/ Hypomethylated
	Upregulated (miRNA and mRNA expression)/ Hypermethylated

ARTÍCULO 2:

Revista:

Genome Biology

Título:

PU.1 targets genes undergo TET2-coupled demethylation and DNMT3b-mediated methylation in monocyte-to-osteoclast differentiation

Autores:

Lorenzo de la Rica[1], Javier Rodríguez-Ubreva[1], Mireia García[2], Abul B. M. M. K. Islam[3,4], José M. Urquiza[1], Henar Hernando[1], Jesper Christensen[5], Kristian Helin[5], Carmen Gómez-Vaquero[2], and Esteban Ballestar[1]

Afiliaciones:

[1] Chromatin and Disease Group, Cancer Epigenetics and Biology Programme (PEBC), Bellvitge Biomedical Research Institute (IDIBELL), 08908 L'Hospitalet de Llobregat, Barcelona, Spain

[2] Rheumatology Service, Bellvitge University Hospital (HUB), 08908 L'Hospitalet de Llobregat, Barcelona, Spain

[3] Department of Experimental and Health Sciences, Barcelona Biomedical Research Park, Universitat Pompeu Fabra (UPF), 08003 Barcelona, Spain

[4] Department of Genetic Engineering and Biotechnology, University of Dhaka, Dhaka 1000, Bangladesh

[5] Biotech Research and Innovation Center (BRIC), and Center for Epigenetics University of Copenhagen, Ole Maaløes Vej 5, 2200 Copenhagen, Denmark

RESUMEN EN CASTELLANO

La metilación del DNA es un mecanismo epigenético clave para la dirigir y estabilizar los cambios que sufre una célula durante sus procesos de diferenciación y desarrollo. El aumento o la disminución de la metilación en una región específica, depende en gran medida de los factores de transcripción unidos a esa determinada región, y de la maquinaria epigenética que éstos recluten. Hay un proceso de diferenciación dentro del sistema hematopoyético con varias características únicas: la conversión de monocitos a osteoclastos. Este proceso de diferenciación se ve alterado en ciertas enfermedades autoinmunes, así como en el cáncer, y se sabe qué factores de transcripción están involucrados en él. En el presente trabajo, nos centramos en el análisis de los cambios de metilación del DNA durante la osteoclastogénesis, y observamos hipometilación e hipermetilación en varios miles de genes, incluyendo algunos importantes para la diferenciación y la función de los osteoclatos. Además, en varios de los genes que se hipometilan, se ha detectado 5-hidroximetilcitosina, un intermediario implicado en las vías de desmetilación activa. Tras analizar los motivos de unión a factores de transcripción en las secuecnias aledañas a los CpGs que se hipo o hipermetilan, se encontró un enriquecimiento específico en motivos para PU.1, NF-kB y AP-1 (Jun / Fos). Específicamente los motivos de unión de PU.1 se encontraban tanto en las regiones que se hipermetilaban como en las que se hipometilan, y gracias a los datos de ChIP-Seq, se comprobó que efectivamente estaba unido a ambos tipos de genes. Por otro lado, PU.1 interacciona con DNMT3b y TET2, actuando de forma dual en lo que respecta a la hipermetilación o hipometilación de determinadas regiones genómicas. Gracias a experimentos de silenciamiento de PU.1 mediado por siRNAs, se comprobó la importancia de PU.1 para la adquisición del estado epigenético correcto. El silenciamiento de PU.1 impidió que los cambios de metilación se realizaran correctamente, ya que TET2 y DNMT3b eran reclutados en menor medida a los promotores estudiados durante el proceso de desarrollo de monocito a osteoclastos. Como conclusión, el presente trabajo ha permitido la identificación de los cambios en la metilación del DNA durante la osteoclastogénesis, y además, ha revelado una nueva función de PU.1 dirigiendo los cambios de metilación durante la diferenciación.

ABSTRACT

Background: DNA methylation is a key epigenetic mechanism for driving and stabilizing cell-fate decisions. Local deposition and removal of DNA methylation are tightly coupled with transcription factor binding, although the relationship varies with the specific differentiation process. Conversion of monocytes to osteoclasts is a unique terminal differentiation process within the hematopoietic system. This differentiation model is relevant to autoimmune disease and cancer, and there is abundant knowledge on the sets of transcription factors involved.

Results: Here we focused on DNA methylation changes during osteoclastogenesis. Hypermethylation and hypomethylation changes took place in several thousand genes, including all relevant osteoclast differentiation and function categories. Hypomethylation occurred in association with changes in 5-hydroxymethylcytosine, a proposed intermediate toward demethylation. Transcription factor binding motif analysis revealed an overrepresentation of PU.1, NF-kB and AP-1 (Jun/Fos) binding motifs in genes undergoing DNA methylation changes. Among these, only PU.1 motifs were significantly enriched in both hypermethylated and hypomethylated genes; ChIP-seq data analysis confirmed its association to both gene sets. Moreover, PU.1 interacts with both DNMT3b and TET2, suggesting its participation in driving hypermethylation and hydroxymethylation-mediated hypomethylation. Consistent with this, siRNA-mediated PU.1 knockdown in primary monocytes impaired the acquisition of DNA methylation and expression changes, and reduced the association of TET2 and DNMT3b at PU.1 targets during osteoclast differentiation.

Conclusions: The work described here identifies key changes in DNA methylation during monocyte-to-osteoclast differentiation and reveals novel roles for PU.1 in this process.

INTRODUCTION

DNA methylation plays a fundamental role in differentiation as it drives and stabilizes gene activity states during cell-fate decisions. Recent reports have shown a close relationship between the participation of transcription factors during differentiation and the generation of cell type-specific epigenetic signatures [1-3]. Several mechanisms explain the co-occurrence of DNA methylation changes and transcription factor binding, including the active recruitment of enzymes involved in DNA methylation deposition, interference or alternative use of the same genomic regions. One of the best models for investigating these mechanisms is the hematopoietic differentiation system given the profound knowledge on the transcription factors implicated at different stages. Many studies have focused on hematopoiesis in order to learn about the type, distribution and role of epigenetic changes, particularly DNA methylation during differentiation. However, the role of DNA methylation changes and the mechanisms participating in their acquisition in terminal differentiation processes remain elusive, even though these are amongst the most important since they produce functional cell types with very specific roles.

A singular differentiation process within the hematopoietic system is represented by differentiation from monocytes (MOs) to osteoclasts (OCs), which are giant, multinucleated cells that are specialized in degrading bone [4]. OCs differentiate from monocyte/macrophage progenitors following M-CSF [5] and RANKL [6] stimulation. Osteoclastogenesis requires cell fusion, cytoskeleton re-organization [7] and the activation of the specific gene sets necessary for bone catabolism. The signaling pathways activated after M-CSF and RANKL induction have been extensively described, and act through TRAF-6 [8, 9], immunoreceptor tyrosine-based activation motif (ITAM) [10] adaptors DAP12 [11] and FcRg [12] associated with their respective receptors, TREM-2 [13] and OSCAR, as well as calcium oscillations [14]. Signals end in the activation of NF¬kB, MAPK and c-Jun, leading to the activation of NFATc1 [15], the master transcription factor of osteoclastogenesis, together with PU.1 and MITF [16], which is already present in the progenitors. These transcription factors bind to the promoter and help up-regulating OC markers such as dendritic cell-specific transmembrane protein (DC-STAMP/TM7SF4) [17], tartrate-resistant acid phosphatase (TRACP/ACP5) [18], cathepsin K (CTSK) [19], matrix metalloproteinase 9 (MMP9) [20] and carbonic anhydrase 2 (CA2).

OC deregulation is involved in several pathological contexts, either in the form of deficient function, as is the case in osteopetrosis [21], or aberrant hyperactivation, as in osteoporosis [22]. These cells are also involved in autoimmune rheumatic disease. For instance, in rheumatoid arthritis aberrantly activated OCs are major effectors of joint destruction [23]. Moreover, OCs cause bone complications in several diseases, such as multiple myeloma [24], prostate cancer and breast cancer [25], and there is also a specific tumor with OC origin, the giant cell tumor of bone [26].

In vitro generation of OCs allows this cell type to be investigated, whereas isolating primary bone OCs for this purpose is very difficult. MOs stimulated with RANKL and M-CSF generate functional OCs [27], which degrade bone and express OC markers [28]. As indicated, the involvement of transcription factors in this model has been well studied, however very few reports have analyzed the role of epigenetic changes during osteoclastogenesis, and these focus mainly on histone modifications [29, 30]. Given the relationship between transcription factors and DNA methylation, we hypothesized that examining DNA methylation changes would provide clues about the involvement of specific factors in the dynamics and hierarchy of these changes in terminal differentiation.

In this study, we compared the DNA methylation profiles of MOs and derived OCs following M-CSF and RANKL stimulation. We found that osteoclastogenesis was associated with the drastic reshaping of the DNA methylation landscape. Hypermethylation and hypomethylation occur in many relevant functional categories and key genes, including those whose functions are crucial to OC biology, like CTSK, ACP5 and DC-STAMP. Hypomethylation occurred early, concomitantly with transcription changes, was DNA replication-independent and associated with a change in 5¬hydroxymethylcytosine, which has been proposed as an intermediate in the process of demethylation. Inspection of transcription factor binding motif overrepresentation in genes undergoing DNA methylation changes revealed the enrichment of the PU.1 binding motif in hypermethylated genes and AP-1, NF-kB and also PU.1 motifs among hypomethylated genes. In fact, analysis of PU.1 ChIPseq data showed its general association to a high number of both hypo and hypermethylated sites. Chromatin immunoprecipitation assays and immunoprecipitation experiments, suggested a potential novel role for PU.1 recruiting DNMT3B to hypermethylated promoters, and TET2, which converts 5-methylcytosine to 5-hydroxymethylcytosine, to genes that become demethylated. This has been demonstrated by performing siRNA-mediated downregulation of PU.1 which partially impaired DNA methylation, expression and recruitment of TET2 and DNMT3B to PU.1 targets, supporting the participation of PU.1 in the acquisition of DNA methylation changes at their target sites.

RESULTS

Cell differentiation and fusion in osteoclastogenesis are accompanied by hypomethylation and hypermethylation of key functional pathways and genes

To investigate the acquisition of DNA methylation changes during monocyte-to-osteoclast differentiation we first obtained three sets of matching samples corresponding to MOs (CD14+ cells) from peripheral blood and OCs derived from the same CD14+ cells, 21 days after the addition of M-CSF and RANKL. The quality of mature, bone-resorbing OCs obtained under these conditions was confirmed by several methods, including the presence of more than three nuclei in TRAP-positive cells (in some cases, up to 40 nuclei per cell were counted), the upregulation of OC markers, such as *CA2*, *CTSK*, *ACP5*/*TRACP* and *MMP9*, and downregulation of the monocytic gene *CX3CR1* (Supplementary Figure 1). At 21 days, over 84% of the nuclei detected in these preparations could be considered to be osteoclastic nuclei (in polykaryons, nuclei and not cells were counted) (Supplementary Figure 1).

We then performed DNA methylation profiling using bead arrays that interrogate the DNA methylation status of > 450,000 CpG sites across the entire genome covering 99% of RefSeq genes. Statistical analysis of the combined data from the three pairs of samples revealed that 3515 genes (8028 CpGs) displayed differential methylation (FC ≥ 2 or FC ≤ 0.5; FDR ≤ 0.05). Specifically, we identified 1895 hypomethylated genes (3597 CpG sites) and 2054 hypermethylated genes (4429 CpGs) (Figure 1A and Supplementary Table I). Changes corresponding to the average three pairs of monocytes/osteoclasts (Figure 1B) were almost identical to the pattern obtained for each individual pair of samples (Supplementary Figure 2A), highlighting the specificity of the differences observed.

Over a third of the differentially methylated CpG-containing probes (33% for hypomethylated CpGs, 45% for hypermethylated CpGs) mapped to gene promoters, the best-described regulatory region for DNA methylation, although DNA methylation changes also occurred at a similar scale in gene bodies (51% for hypomethylated CpGs, 40 % for hypermethylated CpGs) (Figure 1C). Gene ontology analysis of hypomethylated CpGs revealed significant enrichment (FDR ≤ 0.05) for a variety of functional categories of relevance in OC differentiation and function (Figure 1D). We observed very high significance for terms like immune response (FDR = 4.25E-25) and signal transduction (FDR = 1.45E-21), but also more specific categories such as ruffle organization (FDR = 9.91E-2), calcium ion transport (FDR = 4.6E-2) and OC differentiation (FDR = 1.94E-1). In the case of hypermethylated genes, we also found highly significant enrichment of signal transduction (FDR = 4.09E-17),

and enrichment of categories related to other hematopoietic cell types, suggesting that hypermethylation and associated silencing take place in gene sets that become silent in differentiated OCs (Figure 1D). Together, these data indicate that DNA hypomethylation is targeted to genomic regions that are activated during osteoclastogenesis, and hypermethylation silences alternative lineage genes that are not expressed in OCs.

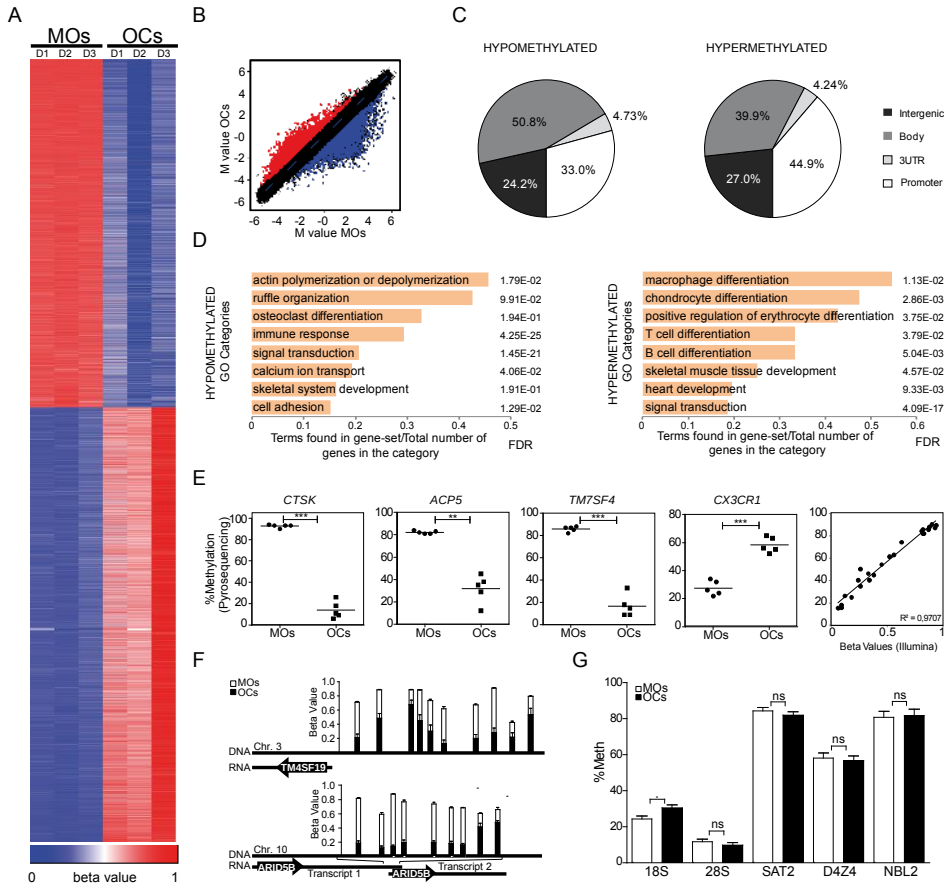


Figure 1. High-throughput methylation comparison between monocytes (MOs) and derived osteoclasts (OCs). (A) Heatmap including the data for three paired samples of MOs (MO D1, D2 and D3) and their derived OCs (OC D1, D2, D3) harvested on day 21. The heatmap includes all CpG-containing probes displaying significant methylation changes (a total of 8028 with FC ≥ 2 or FC ≤ 0.5; p ≤ 0.01 and FDR ≤ 0.05) (Data in Supp. Table I). A scale is shown at the bottom, whereby beta values (i.e. the ratio of the methylated probe intensity to the overall intensity, where overall intensity is the sum of methylated and unmethylated probe intensities) ranging from 0 (unmethylated, blue) to 1 (completely methylated, red) are shown. (B) Scatterplot showing the mean methylation profile of three matching MO/OC pairs. Genes with significant differences

(FC > 2, FDR < 0.05) in averaged results from the three pairs of samples are highlighted in red (hypermethylated) or blue (hypomethylated). (C) Distribution of differentially methylated CpGs among genomic regions (promoter, gene bodies, 3'UTR and intergenic) in different subsets of CpGs (hypomethylated, hypermethylated). (D) Gene ontology enrichment analysis of hypomethylated and hypermethylated CpGs showing the most important categories related with OCs and/or bone biology. (E) Technical validation of the array data by bisulfite pyrosequencing of modified DNA. Three representative hypomethylated genes (*ACP5*, *CTSK* and *TM7SF4*) and one hypermethylated gene (*CX3CR1*) from the array data were validated by BS pyrosequencing. A representation showing the excellent correlation between array data (beta values) and pyrosequencing data (% methylation) including the data for the four genes is also presented (right panel). (F) Cluster analysis of contiguous differentially methylated regions (< 500 bp). Two examples of regions with more than nine consecutive CpGs differentially methylated are shown. (G) Analysis of methylation levels in repetitive elements (Sat2, D4Z4, NBL2) and ribosomal RNA genes (18S and 28S regions) as obtained from bisulfite sequencing analysis.

Remarkably, among the group of hypomethylated genes (Supplementary Table I), we identified changes in several of the archetypal OC genes near their transcription start sites. For example, *CTSK*, the lysosomal cysteine proteinase involved in bone remodeling and resorption, is hypomethylated more than 60%. The *ACP5/TRACP* gene is hypomethylated around 47%. Finally, *TM7SF4*, which encodes for DC-STAMP, a seven-pass transmembrane protein involved in signal transduction in OCs and dendritic cells, undergoes 59% hypomethylation. We also observed significant hypomethylation at the osteoclast-specific transcription factor gene *NFATC1*, although in this case hypomethylation occurred at CpG sites located in its gene body region. Conversely, *CX3CR1*, an important factor for MO adhesion to blood vessels that is downregulated during osteoclastogenesis, displayed an increase in methylation of over 28% (Supplementary Table I).

To confirm that differences in DNA methylation identified between MOs and OCs were robust, we carried out bisulfite genomic pyrosequencing of the aforementioned selection of genes, looking at CpG sites corresponding to the oligonucleotide probe represented in the methylation array. In all cases, bisulfite pyrosequencing confirmed the results of the beadchip array (Figure 1E and Supplementary Table II). This analysis showed a very close correlation between the array and the pyrosequencing data ($R^2 = 0.9707$) (Figure 1E).

We also investigated the coordinated hypomethylation or hypermethylation of adjacent CpGs by analyzing the different sequence window lengths (from 500 bp to 1,000,000 bp). With the largest sequence windows we were able to observe the

coordinated hypermethylation of multiple CpGs across several genes, like those in the HOXA gene cluster. However, the majority of CpGs undergoing coordinated methylation changes were identified within the single gene level. By analyzing CpGs that are concomitantly deregulated within a 500-bp window, we identified several genes displaying coordinated hypomethylation or hypermethylation of many CpG sites (Supplementary Table III). Among these, we identified several CpG clusters in genes potentially involved in OC function and/or differentiation, including 10 CpGs at the promoter of the *TM4SF19* gene, also known as OC maturation-associated gene 4 protein, and 9 CpGs in the gene body of *ARID5B*, the AT-rich interactive domain 5B (MRF1-like) (Figure 1F).

To examine the specificity of the DNA methylation changes further we performed bisulfite sequencing of repetitive elements (Sat2, D4Z4 and NBL2 repeats) and ribosomal RNA genes (Figure 1G and Supplementary Figure 2B). We also performed genome-wide amplification of unmethylated DNA Alu repeats (AUMA), the most common family of repetitive elements that are present in tandem or interspersed in the genome [31]. These experiments showed no significant DNA methylation changes in any of these repetitive elements (Supplementary Figure 2C), reinforcing the notion of the high specificity of hypomethylation and hypermethylation of the identified gene sets.

Hypomethylation is replication-independent and involves changes in 5-hydroxymethylcytosine

To investigate the dynamics of DNA methylation in relation to gene expression changes we first examined how DNA methylation changes are associated with expression changes and then compared the dynamics of DNA methylation and expression changes.

We used osteoclastogenesis expression data (available from the ArrayExpress database under accession number E-MEXP-2019) on 0, 5 and 20 days [32]. Our analysis showed that most changes occurred within the first 5 days, since the expression changes between 0 and 5 days were very similar to those observed between 0 and 20 days, and very few genes changed between 5 and 20 days (Figure 2A and Additional file 6). The 0-to-20-day comparison showed that 2895 genes were upregulated (FC>2; FDR<0.05) and 1858 were downregulated (FC<0.5; FDR<0.05). We found different relationships between DNA methylation changes and gene expression (Figure 2B). An inverse relationship between DNA methylation and gene

expression was mainly observed for changes occurring in CpGs in the proximity of the TSS and within the first exon (Figure 2C) and it was less frequent in those at gene bodies and 3'UTR (Figure 2C). Comparing DNA methylation and expression data revealed that 452 genes were both hypomethylated and overexpressed and 280 genes were both hypermethylated and repressed at the selected thresholds (Additional file 7). We selected a panel of 10 genes from those undergoing hypomethylation and hypermethylation to investigate the dynamics of DNA methylation and expression changes, and performed bisulfite pyrosequencing and quantitative RT-PCR over the entire osteoclastogenesis for three sets of samples (Figure 2D). We found that the promoters of genes like *ACP5*, *CTSK*, *TM7SF4* and TM4SF19 rapidly became hypomethylated following RANKL and M-CSF stimulation (Figure 2D, top). In fact, around 60% of the entire range of hypomethylation occurred between days 0 and 4. Changes in mRNA levels occurred at a similar pace or, in some cases, in an even more gradual manner and were slightly delayed with respect to changes in DNA methylation. In contrast, hypermethylated genes like *PPP1R16B*, *CD6* and *NR4A2* (Figure 2D, bottom) displayed loss of expression before experiencing an increase in DNA methylation, highlighting the different dynamics and mechanisms involved in hypomethylation and hypermethylation events.

It is well established that osteoclastogenesis occurs in the absence of cell division. We tested the levels of cell division in our monocyte-to-osteoclast differentiation experiments by treating cells with BrdU pulses. Consistent with previous observations, fewer than 9.8% were found to be BrdU-positive between 1 and 4 days, confirming the virtual absence of replication (Figure 2E and Supplementary Figure 3). This implies that the large DNA methylation changes observed in this time period are independent of DNA replication. This conclusion is also supported by the fact that treatment with 5-Aza-2'-deoxycytidine (5azadC), a pharmacological compound that results in replication-coupled DNA demethylation [33], had no significant effect on osteoclastogenesis (Figure 2F).

The existence of DNA methylation changes in the absence of replication is particularly significant for genes undergoing demethylation, given the controversy around active DNA demethylation mechanisms. In this context, recent studies have drawn attention towards a family of enzymes, the Tet proteins, which convert 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5hmC) [34, 35] and other modified forms of cytosine, 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) [36]. 5hmC, 5fC and 5caC may represent intermediates in an active demethylation pathway that ultimately replaces 5mC with cytosine in non-dividing cells [37, 38].

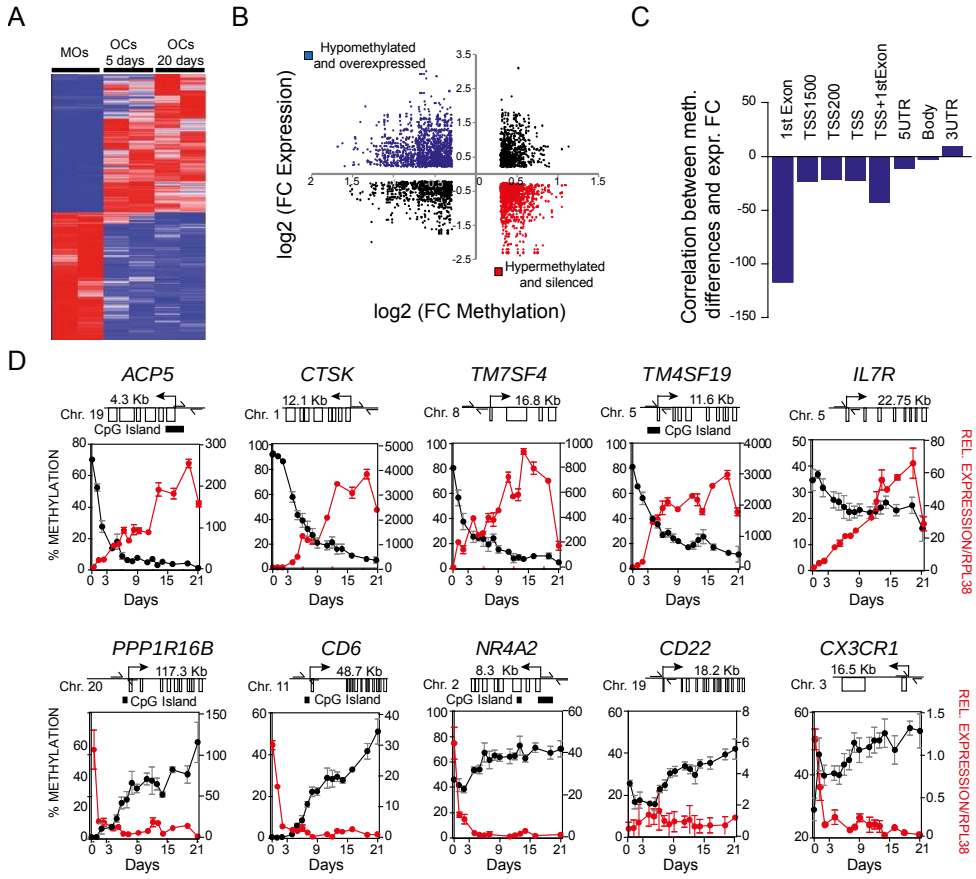


Figure 2 (Part 1). (A) Heatmap showing expression levels on 0, 5 and 20 days. The heatmap includes all the genes displaying significant expression changes (a total of 4753 with FC ≥ 2 or FC ≤ 0.5; p ≤ 0.01 and FDR ≤ 0.05). (Data in Supp. Table IV). (B) Scatterplots showing the relationship between the log2-transformed FC in expression and the log2-transformed FC in DNA methylation. 62% of the hypomethylated genes are overexpressed (in blue); on the contrary, 55% of the hypermethylated genes are repressed (in red). (C) Correlation between methylation and expression data (slope from the linear regression between DNA methylation differences versus expression differences) for all the differentially methylated genes organized by genomic location (first exon, transcription start site, 5'UTR, gene body, 3'UTR). (D) DNA methylation and expression dynamics of selected loci during monocyte-to-osteoclast differentiation. Methylation percentage as determine by bisulfite pyrosequencing. Quantitative RT-PCR data is relative to RPL38. DNA methylation and expression data are represented with a black and a red line respectively.

To establish the potential involvement of hydroxymethylation we focused on the 5hmC levels at early time points in several of the genes that are hypomethylated during osteoclastogenesis, using a method that cleaves DNA that has C, 5mC, but not

the glucosyl-5hmC produced as a result of treatment with the 5-hydroxymethylcytosine specific glucosyltransferase enzyme (Figure 2G). For several genes that become hypomethylated, like *ACP5* and *TM4SF19*, we observed an initial increase in 5hmC levels followed by a slight but significant decrease (Figure 2H), indicating the involvement of hydroxymethylation, and therefore the activity of Tet proteins, in genes that undergo a reduction in DNA methylation.

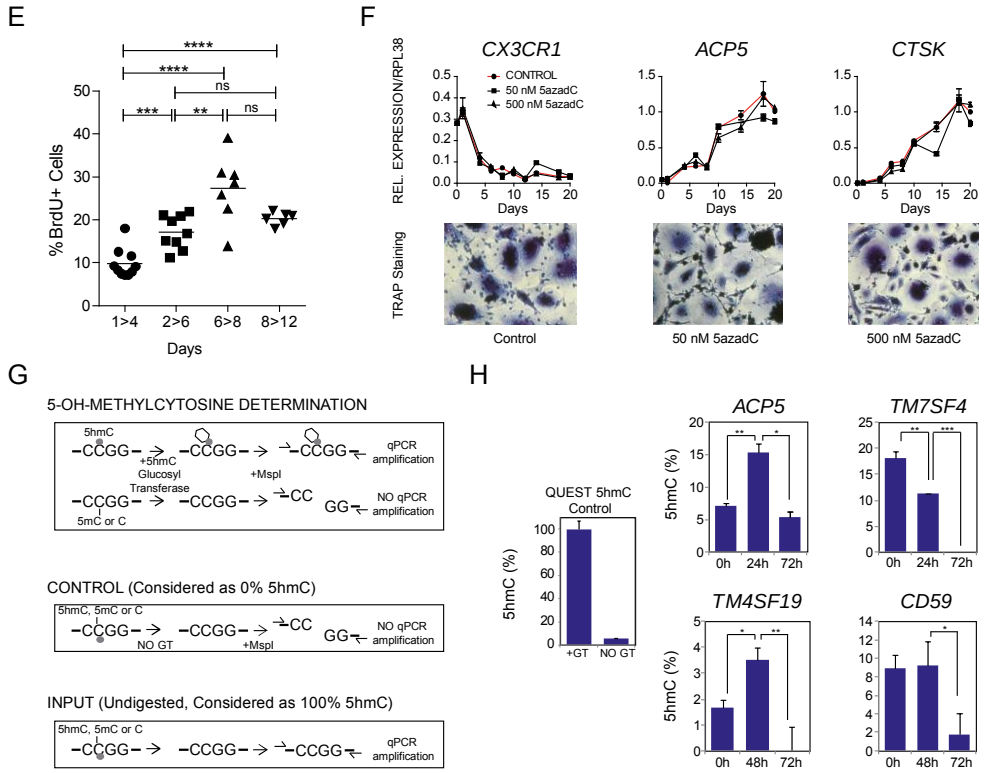


Figure 2 (Part 2). (E) BrdU assay showing the percentage of replicating cells at different times. From day 1 to day 4, only 9.46% of cells divide. (F) Effects of 5azadC treatment (50 nM, 500 nM) on osteoclastogenesis monitoring ACP5 and CTSK updownregulation and CX3CR1 downregulation and TRAP staining over time. (G) Workflow of the technique used to check the presence of 5 hydroxymethylcytosine in hypomethylated genes. DNA was treated with a 5hmC-specific glucosyltransferase. Cytosines bearing a 5-hydroxymethyl are protected against digestion by the MspI restriction enzyme, making possible the detection of amplification in the region by qPCR. When no 5hmC is present, glucose is not transferred to the cytosine and the DNA is cleaved at the restriction enzyme target region (CCGG for MspI), and less amplification is detected by qPCR. Several controls are used to set the 0% and 100% content of 5hmC. (H) 5hmC content in

several of the CpGs that are rapidly demethylated after RANKL and MCSF stimulation of OC precursors.

Sequences undergoing DNA methylation changes are enriched for binding motifs for AP-1, NF-kB and PU.1, key transcription factors in osteoclastogenesis

Different studies have recently shown that transcription factor binding events are associated with changes in the DNA methylation profiles and the response to different situations [2, 39, 40]. To address this further, we first investigated the potential overrepresentation of transcription factor binding motifs among the sequences undergoing DNA methylation changes during OC differentiation using the TRANSFAC database and focusing on a region of 500 bp around the CpG sites identified as undergoing hypomethylation or hypermethylation. We noted highly significant overrepresentation of a small selection of transcription factor binding motifs for genes that undergo hypomethylation or hypermethylation (Figure 3A). We observed that the overrepresentation of binding motifs was very specific to the direction of the DNA methylation change (hypomethylation or hypermethylation).

In the case of hypomethylated genes, we observed highly significant enrichment of binding motifs of the AP-1 family and NF-kB subunits (Figure 3A). We also observed enrichment of PU.1 (FDR 1.07E-12). In fact, 39% of all hypomethylated genes had binding motifs for AP-1, 15% genes had NF-kB binding motifs and another 15% genes had binding motifs for PU.1 or other ETS-related factors (PU.1 alone, 9%) (Figure 3B). As aforementioned, these three groups of TFs play critical roles in osteoclastogenesis [41]. For instance, c-Fos, a component of the dimeric TF AP-1, regulate the switch between monocytes/macrophages and OC differentiation. Fra-1 is downstream to c-Fos, whereas PU.1 and NF-kB are upstream. NF-kB is critical in the expression of a variety of cytokines involved in OC differentiation. In the case of hypermethylated genes, we identified even greater enrichment of the binding motifs of ETS-related transcription factors, especially PU.1 (Figure 3A). In fact, the PU.1 binding motif is present in 15% of all hypermethylated genes (Figure 3B). Other motifs of ETS-related transcription factors from our list of hypermethylated genes included SPIB, ESE1, ETS1, ETS2 and others (Figure 3A). Much lower or insignificant levels of enrichment were obtained for AP-1 family members and NF-kB subunits among the hypermethylated genes. Previous studies have shown that genes that become methylated during hematopoietic differentiation are characterized by ETS transcription factors [2]. This appears to be particularly relevant in monocytic differentiation [1]. Interestingly, most of the reports about the role of PU.1 in

osteoclastogenesis are associated with the activation of osteoclast-specific genes. However, in relation with methylation changes, PU.1 appears to be better correlated with those changes in the direction of repression. Overall, the analysis of transcription factor motifs showed that several of the factors associated with osteoclastogenesis had a significant overrepresentation of their binding motifs among the sets of hypo-and hypermethylated genes (Additional file 9).

To confirm the association of some of these factors with genes that become hypo-and hypermethylated we performed chromatin immunoprecipitation (ChIP) assays with a selection of transcription factors including PU.1, the NFkB subunit p65 and c-Fos, given their known role in osteoclastogenesis as well as the presence of binding motifs for them around hypo-and hypermethylated genes (in the case of c-Fos, it was chosen as a component of the dimeric TF AP-1). To select candidate genes we considered genes with motifs for these three factors among the list of hypo-and hypermethylated genes. For instance, genes that become hypomethylated and have binding sites for p65 include *CCL5*, *IL1R* and *TNFR5SF*. In the case of transcription factor c-Fos, we looked at genes that become hypomethylated like *IL7R*, *CD59* and *IL1R*. In the case of PU.1, we chose key genes with PU.1 binding near the differentially methylated CpG, including *ACP5* and *TM4SF7* (hypomethylated) and *CX3CR1* (hypermethylated). ChIP assays demonstrated the interaction of these factors with most of the aforementioned promoters, even before the stimulation with M-CSF and RANKL (Figure 3C), as if these genes were primed by these factors in monocytes. No binding was observed in control sequences like *Sat2* repeats and the *MYOD1* and *TDRD1* promoters. Interestingly, in the case of PU.1, with both hypo-and hypermethylated genes displaying binding at 0 days, genes becoming demethylated showed a slight increase in PU.1 binding, whereas hypermethylated genes showed a slight decrease in PU.1 association (Figure 3C).

A

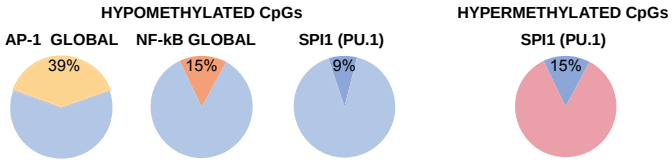
HYPOMETHYL ATED CpGs

TRANSFAC binding motif	p-value	TF family
V$AP1_C	1.03E-211	AP-1
V$BACH1_01	4.41E-190	
V$AP1_Q6_01	4.69E-183	AP-1
V$BACH2_01	5.02E-167	
V$NFKB_Q6_01	1.36E-40	NF-kB
V$NFKAPPAB_01	6.11E-39	NF-kB
V$NFKAPPAB65_01	1.05E-33	NF-kB
V$NFKB_C	6.83E-29	NF-kB
V$CMAF_01	1.42E-15	
V$VMAF_01	4.61E-15	
V$PU1_Q4	4.25E-14	ETS
V$ISRE_01	6.13E-14	

HYPERMETHYL ATED CpGs

TRANSFAC binding motif	p-value	TF family
V$PU1_Q4	5.05E-67	ETS
V$SPIB_01	5.93E-50	ETS
V$PU1_01	4.84E-46	ETS
V$ESE1_Q3	1.84E-43	
V$ETS2_B	5.38E-35	ETS
V$ETS1_B	4.33E-26	ETS
V$CEBPB_02	6.51E-22	CEBP
V$ETS_Q4	1.10E-21	ETS
V$ELF1_Q6	2.71E-15	
V$CEBPB_01	1.35E-14	CEBP
V$CEBPA_01	1.36E-13	CEBP
V$CEBP_Q2_01	2.64E-12	CEBP

B



C

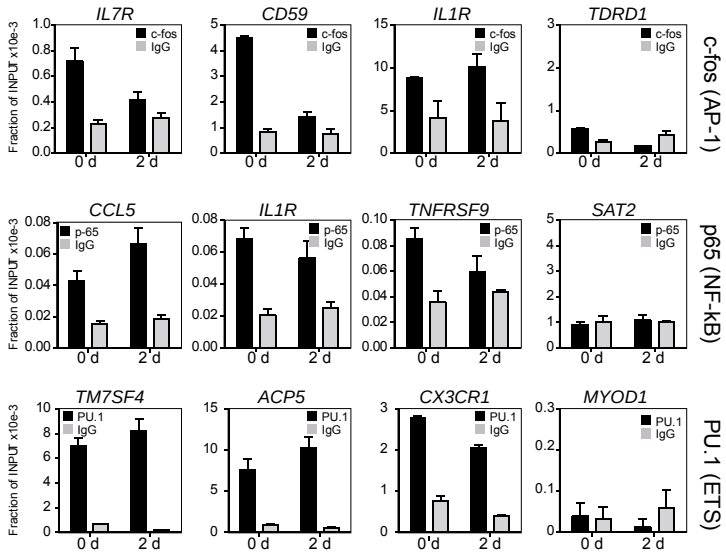


Figure 3. Association of transcription factors with DNA methylation changes during monocyte to OC differentiation. (A) Significant enrichment of predicted TF (TRANSFAC motif) in hypo/hypermethylated CpG sites regions. A 500-bp window centered around the hypo-/hypermethylated CpG sites was tested. The name of the transcription factor binding motif, the p-value and the TF family are provided. Below we show three of the

motifs that have a higher representation in this analysis (B) Diagrams showing the percentage of hypo-/hypermethylated CpGs with AP-1, NF-kB and PU.1 binding sites relative to the total number of hypo-/hypermethylated CpGs. (C) Quantitative ChIP assays showing the binding of three selected transcription factors (p65 NF-kB subunit, Fos and PU.1 to target genes selected by the presence of the putative binding motifs according to the TRANSFAC analysis). Samples were analyzed at 0 ad 2 days after RANKL/M-CSF stimulation. We used Sat2 repeats and the TRDR1 MyoD1 promoter as negative control sequences.

PU.1 recruits DNMT3b and TET2 to hypermethylated and hypomethylated genes

To investigate the potential role of the aforementioned transcription factors in the acquisition of DNA methylation changes we chose PU.1 and NF-kB (p65 subunit) as two representative examples. We first checked their expression levels during osteoclastogenesis, by carrying out qRT-PCR and western blot assays. mRNA and protein analysis (Figure 4A and 4B) both confirmed the expression of these factors. PU.1 revealed an increase at the mRNA levels, although there was no change at the protein level. In the case of p65 NF-kB we only observed a clear increase at the protein level (Figure 4B). In parallel, we also confirmed the presence of DNMT3b, a de novo DNA methyltransferase, and the Ten eleven translocation (TET) protein TET2, as enzymatic activities potentially related with DNA demethylation (Figure 4A and 4B). TET proteins are responsible for conversion of 5mC in 5hmC [34], 5fC and 5cac [36]. Recent evidences support a role for TET-dependent active DNA demethylation process [44, 45]. We focused on TET2 given their high levels in hematopoietic cells of myeloid origin [46, 47]. Also, we have recently reported that TET2 plays a role in derepressing genes in pre-B cell to macrophage differentiation [46], and recent data shows that TET2 is required for active DNA demethylation in primary human MOs [47]. In fact TET1 and TET3 were undetectable in western blot (not shown) and qRT-PCR evidenced their low levels in this cell type (Figure 4A, only shown for TET1).

The confirmed binding of factors like PU.1 and the p65 subunit of NF-kB to hypo-and hypermethylated genes (Figure 3C) raised the possibility of their potential direct interaction with factors involved in maintaining the DNA methylation homeostasis. Some of these interactions have already been explored. For instance, PU.1 physically interacts with the de novo DNA methyltransferases DNMT3A and DNMT3B [48]. Such an interaction, if it occurred in osteoclastogenesis, could provide a potential mechanism to explain how PU.1 target genes become hypermethylated.

One would expect that these transcription factors could also interact with factors participating in demethylation processes. Our previous results suggested the existence of 5hmC enrichment in genes that become hypomethylated, and therefore it is reasonable to test whether NF-kB p65 and PU.1 interact with Tet proteins, the enzymes catalyzing hydroxylation of 5mC.

We therefore tested the recruitment by NF-kB p65 and PU.1 of both DNMT3b and TET2 by carrying out immunoprecipitation assays with osteoclastogenesis samples 0, 2 and 4 days after stimulation with M-CSF and RANKL. Our results showed that PU.1 directly interacted with both DNMT3b and TET2 (Figure 4C). It is plausible that these two interactions may involve different subpopulations of PU.1, for instance with specific post-translational modifications like Ser phosphorylation. However, we did not address this aspect at this point. In the case of NF-kB, we did not observe binding with either of these factors (Figure 4C). This could perhaps be explained by the fact that p65 is shuttling back to the cytoplasm much of the time.

To confirm the interaction between PU.1 and DNMT3b and TET2, we performed reciprocal immunoprepitation experiments with anti-DNMT3b and anti-TET2. These confirmed the direct interaction with PU.1 (Figure 4C). Our results suggested that PU.1 may play a dual coupling transcription factor that can interact with the DNA methyltransferases and enzymes perhaps participating in, or leading to, demethylation. It is likely that other factors participate in the recruitment of these enzymes, however at this stage we focused on PU.1 because of its ability to bind both DNMT3b and Tet2 and its association with both hyper-and hypomethylated sequences.

We then investigated the dual role of PU.1 in recruiting TET2 and DNMT3b to promoters. To this end we performed chromatin immunoprecipitation assays with PU.1, TET2 and DNMT3b in MOs at 0 and 2 days following stimulation with M-CSF/RANKL. We amplified gene promoters with predicted binding sites for PU.1 that become both demethylated (*ACP5*, *TMS7SF4* and *TM4SF19*) as well as hypermethylated (*CX3CR1* and *NR4A2*) and used a non-target of PU.1 (*MYOD1*) as negative control. Our analysis showed binding of PU.1 at both 0 and 2 days (Figure 4D). For genes that become hypomethylated, we observed an increased recruitment of TET2 at these promoters after 2 days, whereas DNMT3b was initially enriched but its association with these promoters was lost after M-CSF and RANKL stimulation (Figure 4D). In genes that become hypermethylated (*CX3CR1* and *NR4A2*), we also observed association of PU.1 at both 0 and 2 days. However we again observed a slight decrease at 2 days together. We also observed increased recruitment of

DNMT3b at 2 days after M-CSF/ RANKL stimulation. We did not observe association of PU.1, DNMT3b and TET2 in the negative control for PU.1 binding, the MyoD promoter.

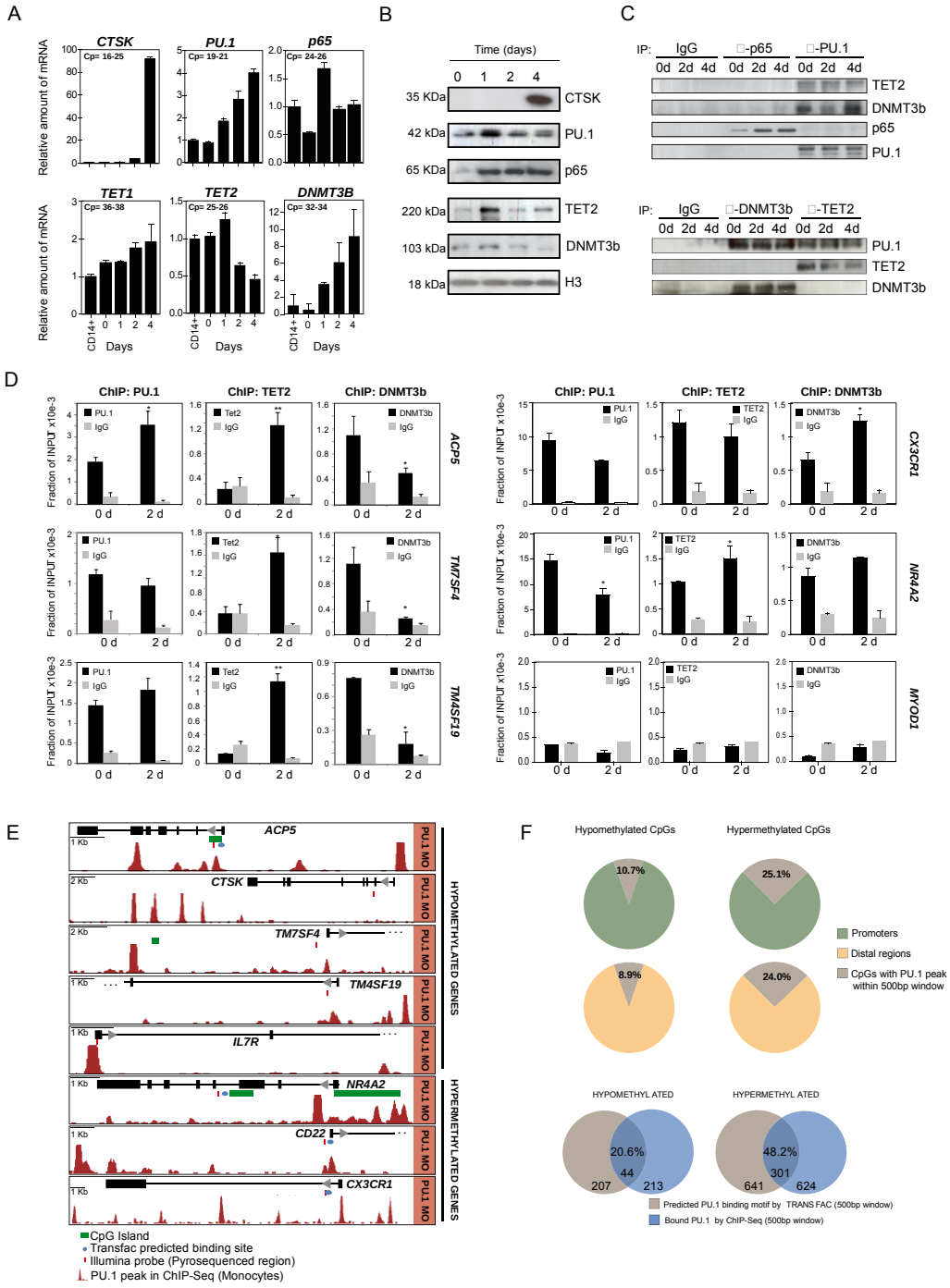


Figure 4. (A) Quantitative RT-PCR analysis showing the mRNA levels of factors and elements of the DNA methylation machinery during osteoclastogenesis. (B) Western blot showing the protein levels of factors and elements of the DNA methylation machinery during osteoclastogenesis. (C) Immunoprecipitation of p65 and PU.1 showing the interaction with DNMT3B and TET2 on 0, 2 and 4 days after RANKL and M-CSF stimulation. IgG was used as a negative control. Reciprocal immunoprecipitation experiments are shown in the bottom panel. (D) Quantitative ChIP assays showing the binding of PU.1, TET2 and DNMT3b binding to hypomethylated genes (*ACP5*, *TM7SF4* and *TM4SF19*) and hypermethylated genes (*CX3CR1*, *NR4A2*), all direct targets of PU.1, and a negative control *(MYOD1* promoter) without PU.1 target sites. The experiment was performed with three biological triplicates but only one experiment is shown. T-student test comparing binding of each antibody between 0 d vs 2 d was performed: * corresponds to p-value<0.05; ** means pvalue<0.01; *** means p-value<0.001. (E) Examples showing PU.1 binding (from ChIPseq data, GSE31621) to the region neighbouring hypo-and hypermethylated CpGs. The PU.1 binding motif location is presented as a horizontal blue dot and the CpG displaying differential methylation (Illumina probe) between MO and OC is marked with a red bar. (F) Analysis of ChIPseq analysis for PU.1 and comparison to TRANSFAC predictions. Top panel: proportion of the CpG-containing probes displaying DNA methylation changes that have peaks for PU.1 binding in the same 500 bp window. Diagrams are separated in the hypo-and hypermethylated sets and in promoter and distal regions (gene bodies, 3'UTR and intergenic regions). Bottom panel, Venn diagrams showing the overlap of PU.1 targets from ChIPseq data (GEO accession number: GSE31621) in MOs and TRANSFAC prediction for PU.1, both using a window of 500 pb with the CpG that displays significant changes in methylation centered.

To evaluate the extent to which hypo-and hypermethylated genes correlate with PU.1 occupancy, we used our DNA methylation data and PU.1 ChIPseq data (GSE31621) obtained in MOs [1]. Most of the individual example genes previously analyzed displayed PU.1 binding overlapping or in the proximity of the CpG sites undergoing a methylation change (Figure 4E). To systematize the analysis we used a window of 500 bp centered around the CpG displaying DNA methylation changes. Under these conditions we found that 10.7% of all hypomethylated CpGs located in promoter regions genes and 25.1% of all hypermethylated CpGs located in promoter regions had PU.1 peaks within this 500 bp window (Figure 4F). These numbers were similar when focusing on CpGs located in distal regions (Figure 4F). We also compared the ChIPseq data to the prediction by TRANSFAC analysis and observed that the overlap between the two sets of list was around 20.6% for hypomethylated genes and 46.9% for hypermethylated genes, again using the same 500 bp for both datasets (Figure 4F). These analyses reinforced the notion of PU.1 associated with a high number of genes undergoing DNA methylation changes, however it also reveals

the weakness in the predictive power of TRANSFAC motif searches and the need of experimental validation of its results.

Dowregulation of PU.1 in MOs impairs activation of OC markers, hypomethylation and recruitment of DNMT3b and TET2

To investigate a potential causal relationship between PU.1 and DNA methylation changes in monocyte-to-osteoclast differentiation we investigated the effects of ablating PU.1 expression in MOs. We therefore downregulated PU.1 levels in MOs using transient transfection experiments with a mix of two siRNAs targeting exon2 and the 3'UTR of PU.1 (Figure 5A). In parallel, we used a control siRNA. Following transfection we stimulated differentiation using RANKL/M-CSF. In these conditions, we checked by qRT-PCR and western blot the effects on PU.1 levels at 1, 2 , 4 and 6 days following RANKL/M-CSF stimulation of MOs and confirmed the PU.1 downregulation close to 60% (Figure 5B and 5C). We then observed that the upregulation of genes like *ACP5* and *CTSK* (both PU.1-direct targets) was partially impaired (Figure 5D). In the case of genes like *CX3CR1* and *NR4A2* we determined that downregulation was also impaired in PU.1-siRNA treated MOs. Interestingly, we also analyzed two genes that are not direct PU.1 targets, one upregulated and hypomethylated during osteoclastogenesis (*PLA2G4E*) and a second one, highly methylated, that does not experience DNA methylation changes during OC differentiation (*FSCN3*). PU.1-siRNA treatment had only small effects on gene expression changes during osteoclast differentiation (perhaps due to indirect effects) when compared to control siRNA, confirming the specificity of the changes observed for the other genes (Figure 5D, bottom). We then tested the effects of PU.1 downregulation in DNA methylation changes. We looked at both PU.1-target genes that become hypomethylated and hypermethylated. In both cases, we observed that downregulation of PU.1 impaired the acquisition of DNA methylation changes, in contrast with the changes observed for control siRNA-treated MOs (Figure 5D). In the case of *TM7SF4*, one of the key genes undergoing hypomethylation, we did not detect an effect of PU.1 downregulation on its DNA methylation dynamics (Additional file 10). However, this could perhaps be explained because this gene undergoes changes before downregulation of PU.1 by siRNA is effective, within day 1 (Additional file 10) and suggests the participation of other factors in this process. At any rate, the observed effects only occurred in PU.1 targets. It did not affect genes that are not targeted by PU.1 (*PLA2G4E*, *FSCN3*). In the case of *PLA2G4E*, PU.1-siRNA treatment did not impair the loss of methylation that occurred in the control experiment. For *FSCN3*, we observed no loss of methylation in any case (Figure 5D).

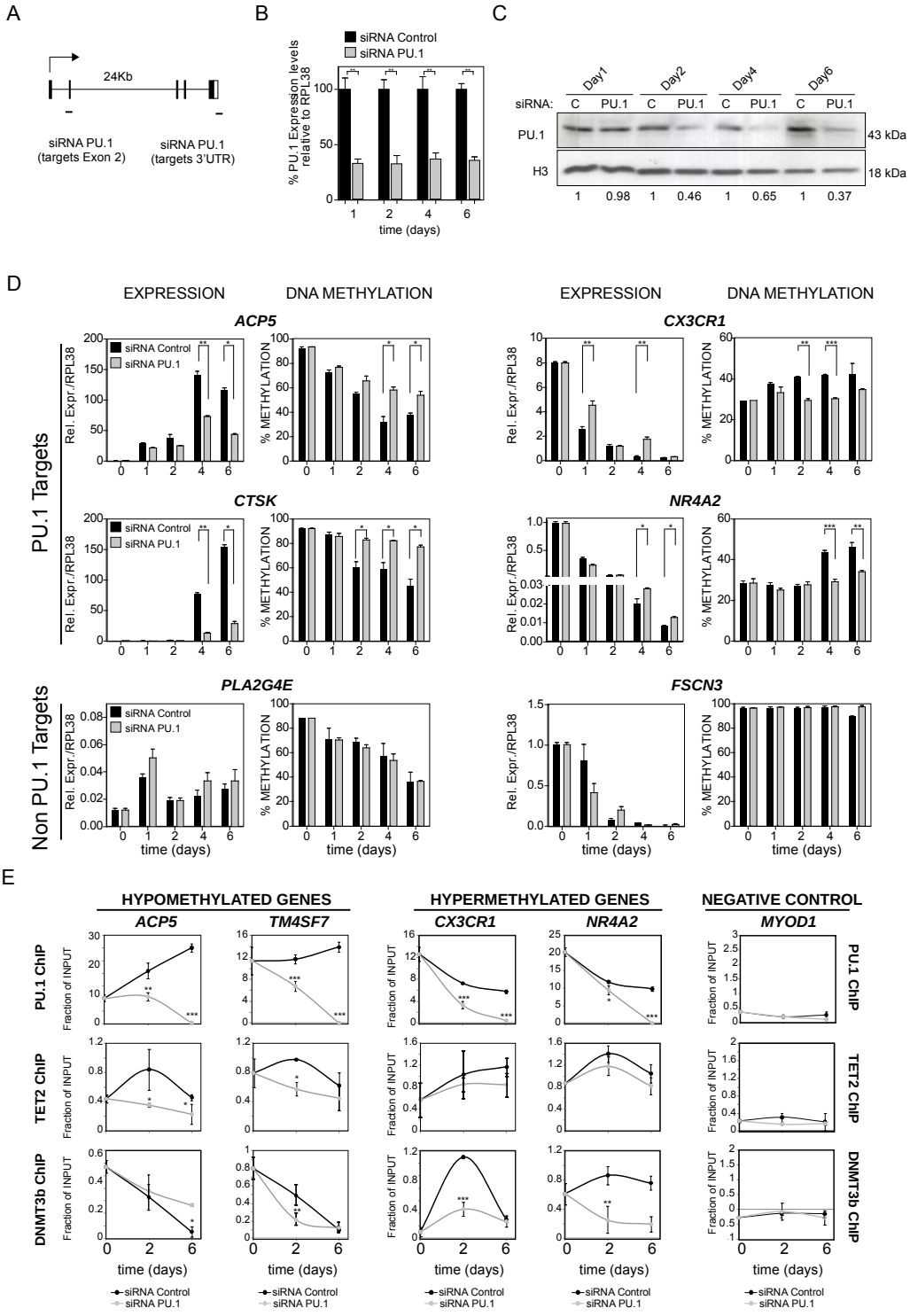

A
24Kb
siRNA PU.1 (targets Exon 2)
siRNA PU.1 (targets 3'UTR)
B
siRNA Control
siRNA PU.1
% PU.1 Expression levels relative to RPL38
time (days)
C
Day1
Day2
Day4
Day6
siRNA:
C
PU.1
PU.1
H3
43 kDa
18 kDa
1 0.98 1 0.46 1 0.65 1 0.37
D
EXPRESSION
DNA METHYLATION
PU.1 Targets
Non PU.1 Targets
ACP5
CX3CR1
CTSK
NR4A2
PLA2G4E
FSCN3
Rel. Expr./RPL38
% METHYLATION
time (days)
E
HYPOMETHYLATED GENES
HYPERMETHYLATED GENES
NEGATIVE CONTROL
ACP5
TM4SF7
CX3CR1
NR4A2
MYOD1
PU.1 ChIP
TET2 ChIP
DNMT3b ChIP
Fraction of INPUT
time (days)
siRNA Control
siRNA PU.1

Figure 5. Effect of the downregulation of PU.1 in the recruitment and function of epigenetic machinery to OC promoters. PU.1 has a direct role in leading DNA methylation changes at their targets. (A) Scheme depicting the two regions of the SPI1 gene (PU.1) (exon 2 and 3'UTR) targeted by the two siRNAs used in this study. (B) Effects of siRNA experiments on PU.1 levels at 1, 2, 4 and 6 days as analyzed by qRT-pCR (C) Effects of siRNA experiments on PU.1 levels at 1, 2, 4 and 6 days as analyzed by western blot (D) Effects of PU.1 downregulation on expression and methylation of PU.1-target genes that become demethylated (*ACP5*, *CTSK*), genes that become hypermethylated (*CX3CR1*, *NR4A2*) and non pU.1 target genes, *PLA2G4E*, which becomes also overexpressed and demethylated, and *FSCN3*, which is hypermethylated and does not undergo loss of methylation during osteoclastogenesis (E) ChIP assays showing the effects of PU.1 downregulation in its recruitment, together with TET2 and DNMT3b binding to the same genes. Data were obtained at 0, 2 and 6 days after M-CSF /RANL stimulation. To simplify the representation negative control assays with IgG for each time point have been substracted to the experiments with each antibody. We have used the *MYOD1* promoter as a negative control (data without substracting the background is presented in Additional file 10). The experiment was performed with three biological triplicates but only one experiment is shown. Error bars correspond to technical replicates. Some of them are smaller than the data point icon. T-student test was performed: * corresponds to p-value<0.05; ** means p-value<0.01; *** means p-value<0.001.

Finally, we compared the effect of PU.1 downregulation in the recruitment of DNMT3b and TET2 to hyper-and hypomethylated promoters (Figure 5E and Additional file 10). As expected, we observed that PU.1 dowregulation resulted in a decrease of the levels of PU.1 associated with the promoters of both hypo-and hypermethylated genes. Most importantly, it also reduced the association of DNMT3b and TET2 reinforcing the notion of the role of these factors and their association with PU.1 in the DNA methylation changes occurring at these CpG sites (Figure 5E). The time course analysis (at 2 and 6 days) of these results also revealed a complex dynamics for the PU.1, TET2 and DNMT3b interactions with their target genes, particularly in the case of hypermethylated genes. It is possible that perturbation of PU.1 levels could be compensated by additional factors that participate in the acquisition of DNA methylation changes of these genes. These aspects will need to be further investigated.

DISCUSSION

Our results provide evidence of the participation of transcription factors, focusing on PU.1, in determining changes in DNA methylation during monocyte-to-osteoclast differentiation. First, a detailed analysis of the sequences undergoing DNA methylation changes produced evidences of the participation of several transcription factors, given the specific overrepresentation of certain motifs in hypo-and hypermethylated genes. This initial analysis was validated in several candidate genes and using ChIPseq data for human primary monocytes [1]. Second, further analyses on one these candidate transcription factors, PU.1, and manipulation of its levels revealed a novel role for this factor in mediating DNA methylation changes during osteoclastogenesis, by direct binding of both DNMT3B and TET2.

In general, DNA methylation changes in differentiation or any other dynamic process are of interest for two reasons: 1) these changes are generally associated with gene expression changes, particularly when associated with promoters or gene bodies, and reveal aspects intrinsic to identity and function of the corresponding cell types. 2) they can be considered as epigenetic footprints that, despite not necessarily being associated with an expression or organizational change, reveal a change in the milieu of a particular CpG and therefore can be used to trace the participation of specific transcription factors or other nuclear elements in that environment/neighborhood. This information can then be used to reconstruct cell signaling events, transcription factors involved and mechanisms participating in differentiation. In this sense, our data show that DNA methylation changes are involved in the differentiation dynamics and stabilization of the OC phenotype since they are concomitant with, or even precede, expression changes. These data are closely correlated with gene expression changes, and a majority of genes that undergo hypomethylation or hypermethylation at their promoters or gene bodies also experience a change in expression, although the relationship varies between different gene sets. Finally, gene ontology analysis reveals that all relevant functional categories and the majority of key genes for differentiation or the activity of functional OCs undergo DNA methylation changes and that genes within all relevant functional categories undergo DNA methylation changes..

Our study suggests that both hypomethylation and hypermethylation events are equally important. Hypomethylation events, in many cases associated with gene activation, affect genes that are specific to this differentiation process or are related with the function of differentiated OCs. In contrast, the identity of genes affected by hypermethylation events is less closely correlated with OC function, given that most of them are related with gene repression. In fact, we found that hypermethylation

affects genes that are specific to other cellular types. Given that osteoclastogenesis involves cell fusion and the generation of highly polyploid cells, we had speculated whether the existence of redundant copies of genetic material could lead to massive gene repression, and the silencing of extra copies. However, hypermethylation does not seem to be predominant over hypomethylation. The two activities are very specific to particular gene sets and there are no indications of changes in repetitive elements.

A number of transcription factors are essential for OC formation. Some of these factors are involved in various differentiation processes. Among these, PU.1, c-Fos, NF-kB and other factors are essential for osteoclastogenesis. In fact, NF-kB and PU.1-deficient mice show a macrophage differentiation failure, and osteoclastogenesis is inhibited at an early stage of differentiation. c-Fos is a component of the dimeric TF AP-1, which also includes FosB, Fra-1, Fra-2, and Jun proteins such as c-Jun, JunB, and JunD. Other key factors involved in OC differentiation include CEBPalpha [42] and Bach1[49]. Osteoclastogenesis also depends on the activity of more specific transcription factors like NFATc1 and MITF. Interestingly, the analysis of the presence of transcription factor binding sites in sequences that undergo DNA methylation changes shows a significant enrichment in binding motifs of transcription factors that are key in OC differentiation, some of which we have validated for a selection of putative target genes.

One of the most interesting factors in this process is the ETS factor PU.1. In fact, PU.1 is the earliest molecule known to influence the differentiation and commitment of precursor myeloid cells to the OC lineage. PU.1 functions in concert with other transcription factors, including c-Myb, C/EBPa, cJun and others, to activate osteoclast-specific genes.

Our results reveal two hitherto undescribed roles for PU.1 in the context of monocyte-to-OC differentiation. First, we have identified the association of PU.1 with genes that become repressed through hypermethylation and describe its direct interaction with DNMT3b in the context of osteoclastogenesis. Second, we identify a novel interaction between PU.1 and TET2 and their association with genes that become demethylated. Our study shows that PU.1 may act as a dual adaptor during osteoclastogenesis, in the directions of hypomethylation and hypermethylation. This is compatible with previous data on genome wide DNA methylation profiling comparing cell types across the hematopietic differentiation system where an overrepresentation of ETS transcription factor binding sites was found [2]. In monocyte-to-osteoclast differentiation, PU.1 is best known for its role in the

activation of osteoclast-specific genes. However, studies in other models have previously shown that PU.1 can participate in the repression of genes in concert with elements of the epigenetic machinery. For instance, PU.1 is known to generate a repressive chromatin structure characterized by H3K9me3 in myeloid and erythroid differentiation [50]. Also, PU.1 has been shown to act in concert with MITF to recruit corepressors to osteoclast-specific in committed myeloid precursors capable of forming either macrophages or OC [51]. Moreover, previous studies have shown that PU.1 can form a complex with DNMT3a and DNMT3b [48]. However, this is the first report where the association between PU.1 and DNMTs in association with gene repression is shown in this context.

Moreover, our findings constitute the first report where the binding of PU.1 to TET2 has been described. Several recent reports have pointed at TET2-mediated hydroxylation of 5-methylcytosine as an intermediate step towards demethylation [52] and our data shows changes in 5hmC at genes that become demethylated in osteoclastogenesis, reinforcing the possibility that PU.1-mediated recruitment of TET2 is leading to 5hmC-mediated demethylation. However the detailed mechanisms that couple hydroxylation of 5mC and demethylation are still object of debate.

The manipulation of PU.1 levels by using siRNAs has shown that PU.1 has a direct role in recruiting DNMT3b and TET2 to its target promoters, as well as showing how impaired association of PU.1 results in defective acquisition of DNA methylation changes in both directions as well as reduced effect on gene expression changes. Therefore, our data reveal a novel role of PU.1 as a dual adaptor with the ability to bind both epigenetically repressive and epigenetically activating events and targeting DNA methylation changes in both directions (Figure 6). The incomplete impairment of DNA methylation and expression changes, as well as partial loss of Tet2 and

DNMT3b following PU.1 knock-down indicates that additional transcription factors are also participating in this process. In future studies, it will also be interesting to identify the mechanisms that operate in the specific recruitment of PU.1-TET2 to genes that become demethylated, and to determine how PU.1-DNMT3b is recruited to genes that become hypermethylated. It is likely that specific transcription factors play a role, and specific post-translational modifications in PU.1 may participate in the coupling of its associated complexes to specific factors. In this context, Ser phosphorylation of PU.1 has already been shown to play a role in its recruitment to promoters [53] and could also participate in discriminating interaction with epigenetic modifiers.

Our study has allowed us to identify key DNA methylation changes during OC differentiation and has revealed an implication of PU.1 in the acquisition of DNA methylation and expression changes as well as identifying novel interactions with DNMT3b and TET2.

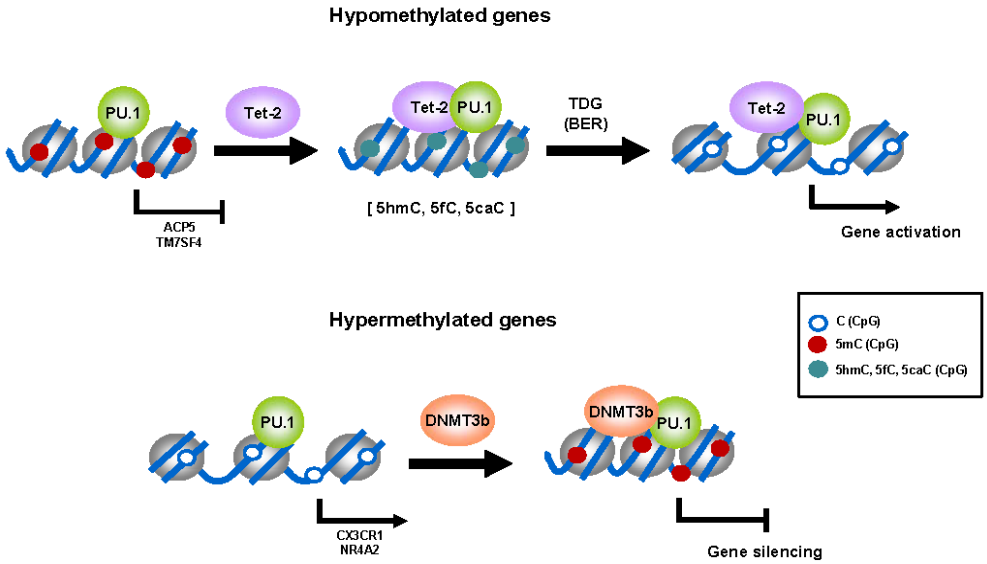


Figure 6. Model showing the proposed recruitment of TET2 and DNMT3b by PU.1 to its target genes that become hypo- or hypermethylated respectively during osteoclastogenesis. Genes that become hypomethylated exchange PU.1-DNMT3b by PU.1-TET2 (although whether pre-existing subpopulations of these association exist or alternatively if a post-translational or another event mediates exchange of TET2 and DNMT3b is not elucidated in the present stay). TDG is likely to mediate conversion of 5hmC/5fmC/5caC to demethylated cytosine. Hypermethylated genes do not display binding of TET2, and recruit DNMt3b as differentiation to OCs is triggered.

CONCLUSIONS

Our study of the DNA methylation changes in monocyte-to-osteoclast differentiation reveals the occurrence of both hypomethylation and hypermethylation changes. These changes occur in the virtual absence of DNA replication suggesting the participation of active mechanisms, particularly relevant for hypomethylation events, for which the mechanisms are still subject of debate. Also, when comparing the dynamics of DNA methylation and expression changes, hypomethylation occurs concomitant or even earlier than expression changes. In contrast, for the majority of genes becoming hypermethylated, hypermethylation follows expression changes. Hypomethylation takes place in relevant functional categories related with OC differentiation and most of the genes that are necessary for OC function undergo hypomethylation including *ACP5*, *CTSK* and *TM7SF4* among others. The analysis of overrepresentation of transcription factor binding motifs reveals the enrichment of specific motifs for hypomethylated and hypermethylated genes. Among these, PU.1 and other ETS-related binding motifs are highly enriched in both hypomethylated and hypermethylated genes. We have demonstrated that PU.1 is bound to both hypo- and hypermethylated promoters and that it is able to recruit both DNMT3b and TET2. Most importantly, downregulation of PU.1 with siRNAs not only shows a reduction in the recruitment of these enzymes to their corresponding target sites but also a specific reduction in the acquisition of DNA methylation and expression changes at PU.1 target genes. Our results demonstrate a key role of PU.1 in driving DNA methylation changes during OC differentiation.

MATERIALS AND METHODS

Differentiation of OCs from peripheral blood mononuclear cells

Human samples (blood) used in this study came from anonymous blood donors and were obtained from the Catalan Blood and Tissue Bank (Banc de Sang i Teixits) in Barcelona as thrombocyte concentrates (buffy coats). The anonymous blood donors received oral and written information about the possibility that their blood would be used for research purposes, and any questions that arose were then answered. Prior to obtaining the first blood sample the donors signed a consent form at the Banc de Teixits. The Banc de Teixits follows the principles set out in the WMA Declaration of Helsinki. The blood was carefully layered on a Ficoll–Paque gradient (Amersham, Buckinghamshire, UK) and centrifuged at 2000 rpm for 30 min without braking. After centrifugation, peripheral blood mononuclear cells (PBMCs), in the interface between the plasma and the Ficoll– Paque gradient, were collected and washed twice with ice-cold PBS, followed by centrifugation at 2000 rpm for 5 min. Pure CD14+ cells were isolated from PBMCs using positive selection with MACS magnetic CD14 antibody (Miltenyi Biotec). Cells were then resuspended in α-minimal essential medium (α-MEM, Glutamax no nucleosides) (Invitrogen, Carlsbad, CA, USA) containing 10% fetal bovine serum, 100 units/ml penicillin, 100μg/ml streptomycin and antimycotic and supplemented with 25 ng/mL human M-CSF and 50 ng/ml hRANKL soluble (PeproTech EC, London, UK). Depending on the amount needed, cells were seeded at a density of $3 \cdot 10^5$ cells/well in 96-well plates, $5 \cdot 10^6$ cells/well in 6-well plates or $40 \cdot 10^6$ cells in 10 mm plates and cultured for 21 days (unless otherwise noted); medium and cytokines were changed twice a week. The presence of OCs was checked by tartrate-resistant acid phosphatase (TRAP) staining using the Leukocyte Acid Phosphatase Assay Kit (Sigma–Aldrich) according to the manufacturer's instructions. A phalloidin/DAPI stain allowed us to confirm that the populations were highly enriched in multinuclear cells, some of them containing more than 40 nuclei. We used several methods to determine that on day 21 almost 85% of the nuclei detected were “osteoclastic nuclei” (in polykaryons, nuclei and not cells were quantified). OCs (TRAP-positive cells with more than three nuclei) were also analyzed at the mRNA level: upregulation of key OC markers (*TRAP/ACP5*, *CA2*, *MMP9* and *CTSK*) and the downregulation of the MO marker *CX3CR1* were confirmed.

Treatment of MOs with 5-aza-2-deoxycytidine

In some cases we performed monocyte-to-osteoclast differentiation experiments in the presence of different subtoxic concentrations of the DNA replication-coupled demethylating drug 5-aza-2deoxycytidine (at 50 nM, 500 mM) for 72 h.

Visualization of OCs with phalloidin and DAPI staining

PBMCs or pure isolated CD14+ cells were seeded and cultured in glass Lab-Tek Chamber Slides (Thermo Fisher Scientific) for 21 days in the presence of hM-CSF and hRANKL. OCs were then washed twice with PBS and fixed (3.7% paraformaldehyde, 15 min). Cells were permeabilized with 0.1% (V/V) Triton X-100 for 5 min and stained for F-actin with 5 U/mL Alexa Fluor® 647-Phalloidin (Invitrogen). Cells were then mounted in Mowiol-DAPI mounting medium. Cultures were visualized by CLSM (Leica TCP SP2 AOBS confocal microscope).

DNA methylation profiling using universal bead arrays

Infinium HumanMethylation450 BeadChips (Illumina, Inc.) were used to analyze DNA methylation. This array allows interrogating > 485,000 methylation sites per sample at single-nucleotide resolution, covering 99% of RefSeq genes, with an average of 17 CpG sites per gene region distributed across the promoter, 5'UTR, first exon, gene body and 3'UTR. It covers 96% of CpG islands, with additional coverage in CpG island shores and the regions flanking them. DNA samples were bisulfite converted using the EZ DNA methylation kit (Zymo Research, Orange, CA). After bisulfite treatment, the remaining assay steps were performed following the specifications and using the reagents supplied and recommended by the manufacturer. The array was hybridized using a temperature gradient program, and arrays were imaged using a BeadArray Reader (Illumina Inc.). The image processing and intensity data extraction software and procedures were those previously described [54]. Each methylation data point is obtained from a combination of the Cy3 and Cy5 fluorescent intensities from the M (methylated) and U (unmethylated) alleles. Background intensity computed from a set of negative controls was subtracted from each data point. For representation and further analysis we used both Beta values and M values. The Beta-value is the ratio of the methylated probe intensity and the overall intensity (sum of methylated and unmethylated probe intensities). The M-value is calculated as the log2 ratio of the intensities of methylated probe versus unmethylated probe. The Beta-value ranges from 0 to 1 and is more intuitive and was used in heatmaps and in comparisons with DNA methylation percentages from bisulfite pyrosequencing experiments, however for statistic purposes it is more adequate the use of M values [55].

Detection of differentially methylated CpGs

The approach to select differentially methylated CpGs was implemented in R [56], a well-known language in statistical computing. In order to process Illumina Infinium HumanMethylation450 methylation data, we used the methods supplied in limma

[57], genefilter, and lumi [58] packages from Bioconductor repository. Previous to statistical analysis, a pre-process stage is applied, the main steps are: 1) Color balance adjustment, i.e., normalization between two color channels; 2) Performing quantile normalization based on color balance adjusted data, and 3) variance filtering by IQR (Interquartile range) using 0.50 for threshold value. Subsequently, for statistical analysis, eBayes moderated t-statistics test was carried out from limma package [57]. Specifically, a paired limma was performed as designed in IMA package [59]. To choose significant differences in methylated CpGs several criteria have been proposed. In this study, we considered a probe as differentially methylated if (1) has a fold-change >2 for hypermethylated and <0.5 hypomethylated) and (2) the statistical test was significant (p-value<0.01 and FDR<0.05).

Identification of genomic clusters of differentially methylated CpGs

A clustering method was applied to the differenced methylated CpGs from charm package [60]. We re-implemented the code to invoke the main clustering function using genomic CpG localisation: identify Differentially Methylated Regions (DMRs) by grouping differentially methylated probes (DMPs). The maximum allowable gap between probe positions for probes to be clustered into the same region was set to 500 bp. It has been shown that in many cases methylation changes are observed over a range of CpGs, which may be identified for instance at shores close to Transcription Starting Sites. We considered that DMR are more robust signals than DMPs. In this analysis, the considered list of CpGs attains a p-value below 0.01 and FDR < 0.05.

Bisulfite sequencing and pyrosequencing

We used bisulfite pyrosequencing to validate CpG methylation changes resulting from the analysis with the Infinium HumanMethylation450 BeadChips. Bisulfite modification of genomic DNA isolated from MOs, OCs, and samples from time course or PU.1-knockdown experiments was carried out as described by Herman et al. [61]. 2 µl of the converted DNA (corresponding to approximately 20-30 ng) were then used as a template in each subsequent PCR. Primers for PCR amplification and sequencing were designed with the PyroMark® Assay Design 2.0 software (Qiagen). PCRs were performed with the HotStart Taq DNA polymerase PCR kit (Qiagen), and the success of amplification was assessed by agarose gel electrophoresis. PCR products were pyrosequenced with the Pyromark™ Q24 system (Qiagen). In the case of repetitive elements (Sat2, D4Z4, NBL2 and 18S rRNA and 28 rRNA) we performed standard bisulfite sequencing of a minimum of 10 clones. Results from bisulfite pyrosequencing and sequencing of multiple clones are presented as a percentage of

methylation. All primer sequences are listed in Additional file 11. Raw data for bisulfite sequencing of all samples is presented in Additional file 4.

Gene expression data analysis and comparison of DNA expression data versus DNA methylation data

In order to compare expression data versus methylation data, we used CD14+ and OC expression data from ArrayExpress database www.ebi.ac.uk/arrayexpress) under the accession name (EMEXP-2019) from a previous publication [32]. Affymetrix GeneChip Human Genome U133 Plus 2.0 expression data was processed using limma [57] and affy [62] packages from bioconductor. The pre-processing stage is divided in three major steps: 1) background correction, 2) normalization, and 3). reporter summarization. Here, the expresso function in affy package was chosen for preprocessing. Thus, the RMA method [63] was applied for background correction. Then, a quantile normalization was performed. In addition, we introduced a specific step for PM (perfect match probes) adjustment, utilizing the PM-only model based expression index (option 'pmonly'). And finally, for summarization step, the median polish method was taken. Next, as previously in the methylation analysis, a variance filtering by IQR (Interquartile range) using 0.50 for threshold value was executed. After preprocessing, a statistical analysis was applied, using eBayes moderated t-statistics test from limma package. Subsequently, expression genes matching to methylated genes were studied. Genes differentially expressed between MOs and Mo-OCs groups were selected with a criteria of p-value lower than 0.01 and False Discovery Rate (FDR) lower than 0.05 as calculated by Benjamini-Hochberg and a fold-change of expression higher than 2 or lower than 0.5. Validation of expression data was performed by quantitative RT-PCR. All primer sequences are listed in Additional file 11.

Gene ontology analysis

Gene ontology (GO) was analyzed with the FatiGO tool [64], which uses Fisher's exact test to detect significant over-representation of GO terms in one of the sets (list of selected genes) with respect to the other (the rest of the genome). Multiple test correction to account for the multiple hypotheses tested (one for each GO term) was applied to reduce false positive results. GO terms with adjusted values of $p < 0.05$ were considered significant.

Analysis of transcription factor binding

We used the STORM algorithm [65] to identify potential overrepresentation of transcription factor motifs in the 500 bp region around the center of the hypomethylated/hypermethylated CpG sites (as well as for all other CpGs-containing probes contained in the array) assuming cutoff values of p = 0.00002 (for hypo-/hypermethylated probes) and 0.00001 (for all other probes), using position frequency matrices (PFMs) from the TRANSFAC database (Professional version, release 2009.4) [66]. Enrichment analysis of predicted TF in the probes of significant hypomethylated probes (n = 421) was conducted using GiTools ([67]; www.gitools.org). We calculated two-tailed probabilities, and a final adjusted FDR p-value (with 0.25 cutoff) was considered statistically significant. We downloaded PU.1 ChIPseq data for CD14+ MOs generated by Michael Rehli's laboratory [1] from the Gene Expression Omnibus (GSE31621). The genomic locations of the calculated peaks were mapped to GRCh37.p10 human alignment obtained from Biomart [68], by using bedtools (intersect function) in order to obtain the PU.1 occupied genes. To determine whether a given CpG (from the Illumina bead array) was positive for PU.1 binding, we used the same 500 bp window used for TRANSFAC analysis.

Graphics and heatmaps

All graphs were created using Prism5 Graphpad. Heatmaps were generated from the expression or methylation data using the Genesis program (Graz University of Technology) [43].

BrdU proliferation assays

BrdU was used at a final concentration of 300 μm, as previously described. On the days specified, BrdU pulsing solution was added to each well for 2 to 4 days. For confocal microscopy of monocyteto-osteoclast differentiation samples, CD14+ cells were seeded on Millicell EZ 8-well glass slides (Millipore) and cultured in differentiation media. At different times BrdU was added to the medium and after 2-4 days cells were fixed (4% paraformaldehyde, 30 minutes, RT), permeabilized (PBS-BSA-Triton X-100 0,8%, 10 minutes, RT) and treated with HCl 2N for 30 minutes. After DNA opening, HCl was neutralized by two 5-minute washes with NaBo (0.1M, pH 8.5) and two 5-minutes washes with PBT. Cells were incubated with anti-BrdU antibody (18 h at 4ºC, 1:1000 dilution) and an anti-mouse Alexa-568 conjugated antibody was added to visualize the BrdU-positive nuclei. A phalloidin incubation step and Mowiol-DAPI mounting media were used.

Transfection of primary human MOs

We used two different Silencer® select pre-designed siRNAs against human PU.1 (one targeting exon 2 and another targeting the 3'UTR) and a Silencer® select negative control to perform PU.1 knockdown experiments in peripheral blood MOs. We used Lipofectamine RNAiMAX Transfection Reagent (Invitrogen) for efficient siRNA transfection. mRNA and protein levels were examined by quantitative RT-PCR and western blot at 1, 2, 4 and 6 days after siRNA transfection. In this case MO samples were prepared by incubating PBMCs in plates in α-MEM without serum for 30 min and washing out the unattached cells. Under these conditions over 80% are MOs. This alternative protocol was used for increased viability following transfection. These experiments were performed with three biological replicates.

Chromatin immunoprecipitation (ChIP) assays and immunoprecipitation experiments

Immunoprecipitation was performed by standard procedures in CD14+ cells at 0, 2 and 4 days after treatment with M-CSF and RANKL. Cell extracts were prepared in 50 mM Tris-HCl, pH 7.5, 1 mM EDTA, 150 mM NaCl, 1% Triton-X-100 and protease cocktail inhibitors (Complete, Roche Molecular Biochemicals). Cellular extracts and samples from immunoprecipitation experiments were electrophoresed and western blotted following standard procedures. For chromatin immunoprecipitation (ChIP) assays, CD14+ at 0, 2 and 4 days after treatment with MCSF and RANKL were crosslinked with 1% formaldehyde and subjected to immunoprecipitation after sonication. ChIP experiments were performed as described [46]. Analysis was performed by real-time quantitative PCR. Data are represented as the ratio of the bound fraction over the input for each specific factor. We used a mouse monoclonal antibody against the TET2 N-t for ChIPs and a rabbit polyclonal antibody against TET2 for western blot. For DNMT3b we used a rabbit polyclonal against amino acids 1-230 of human DNMT3b (Santa Cruz Biotechnology). We also used a rabbit polyclonal against the C-t of PU.1 (sc-352, Santa Cruz Biotechnology), a rabbit polyclonal against the N-t of c-Fos (sc-52, Santa Cruz Biotechnology) and a rabbit polyclonal against the C-t of NF-kB p65 (sc-372, Santa Cruz Biotechnology). IgG was used as a negative control. Primer sequences were designed to contain either predicted or known TF binding (from TRANSFAC or ChIPseq data) as close as possible from the CpG undergoing methylation changes. Primer sequences are shown in Additional file 11. These experiments were performed with three biological replicates.

5hmC detection

5hmC was analyzed using the Quest 5-hmC Detection system (Zymo). Genomic DNA was treated with a specific 5hmC glucosyltransferase (GT) or left untreated (No GT, 0% 5hmC). DNA was then digested with MspI (100U) at 37°C overnight, followed by column purification. The MspI-resistant fraction (bearing the glucosile group, and therefore the original 5hmC) was quantitated by qPCR using primers designed around at least one MspI site (CCGG), and normalized to the amplification of the same region in the original DNA input. The amplification obtained in the untreated (no GT, MspI sensible) was then substracted to the samples in order to calculate the level of 0% 5hmC. The resulting values were the percentage of 5hmC present in each of the samples. Primer sequences are shown in Additional file 11.

Amplification of UnMethylated Alus (AUMA)

This method, aiming at amplifying unmethylated Alus, was performed as described [31, 39]. Products were resolved on denaturing sequencing gels. Bands were visualized by silver staining the gels. AUMA fingerprints were visually checked for methylation differences between bands in different samples.

List of abbreviations

5azadC, 5-aza-2´-deoxycytidine

5mC, 5-methylcytosine

5hmC, 5-hydroxymethylcytosine

AUMA, amplification of unmethylated Alu repeats

Competing interests

The authors declare that they have no competing interests

Author's contributions

LR and EB conceived experiments; LR, MG, JR-U and HH performed experiments; AI and JMU performed biocomputing analysis; LR, JMU, JC, KH, CGV and EB analyzed the data; EB wrote the paper.

Data Access

Methylation array data for this publication has been deposited in NCBI's Gene Expression Omnibus and is accessible through GEO Series accession number GSE46648.

Acknowledgements

We thank Dr Fátima Al-Shahrour and Dr Núria López-Bigas for helpful suggestions about the bioinformatic analyses, and Dr Mercedes Garayoa and Dr Antonio Garcia-Gomez for their helpful suggestions about protocols and providing antibodies. This work was supported by grant SAF2011-29635 from the Spanish Ministry of Science and Innovation, grant CIVP16A1834 from Fundación Ramón Areces and grant 2009SGR184 from AGAUR (Catalan Government). LR is supported by a PFIS predoctoral fellowship.

REFERENCES

1. Pham TH, Benner C, Lichtinger M, Schwarzfischer L, Hu Y, Andreesen R, Chen W, Rehli M: **Dynamic epigenetic enhancer signatures reveal key transcription factors associated with monocytic differentiation states**. *Blood* 2012, **119**(24):e161-171.
2. Hogart A, Lichtenberg J, Ajay SS, Anderson S, Margulies EH, Bodine DM: **Genome-wide DNA methylation profiles in hematopoietic stem and progenitor cells reveal overrepresentation of ETS transcription factor binding sites**. *Genome Res* 2012, **22**(8):1407-1418.
3. Zhang JA, Mortazavi A, Williams BA, Wold BJ, Rothenberg EV: **Dynamic transformations of genome-wide epigenetic marking and transcriptional control establish T cell identity**. *Cell* 2012, **149**(2):467-482.
4. Blair HC, Teitelbaum SL, Ghiselli R, Gluck S: **Osteoclastic bone resorption by a polarized vacuolar proton pump**. *Science* 1989, **245**(4920):855-857.
5. Wiktor-Jedrzejczak W, Bartocci A, Ferrante AW, Jr., Ahmed-Ansari A, Sell KW, Pollard JW, Stanley ER: **Total absence of colony-stimulating factor 1 in the macrophage-deficient osteopetrotic (op/op) mouse**. *Proc Natl Acad Sci U S A* 1990, **87**(12):4828-4832.
6. Lacey DL, Timms E, Tan HL, Kelley MJ, Dunstan CR, Burgess T, Elliott R, Colombero A, Elliott G, Scully S *et al*: **Osteoprotegerin ligand is a cytokine that regulates osteoclast differentiation and activation**. *Cell* 1998, **93**(2):165-176.
7. Saltel F, Chabadel A, Bonnelye E, Jurdic P: **Actin cytoskeletal organisation in osteoclasts: a model to decipher transmigration and matrix degradation**. *Eur J Cell Biol* 2008, **87**(8-9):459-468.
8. Wong BR, Besser D, Kim N, Arron JR, Vologodskaia M, Hanafusa H, Choi Y: **TRANCE, a TNF family member, activates Akt/PKB through a signaling complex involving TRAF6 and c-Src**. *Mol Cell* 1999, **4**(6):1041-1049.
9. Kobayashi N, Kadono Y, Naito A, Matsumoto K, Yamamoto T, Tanaka S, Inoue J: **Segregation of TRAF6-mediated signaling pathways clarifies its role in osteoclastogenesis**. *Embo J* 2001, **20**(6):1271-1280.
10. Blank U, Launay P, Benhamou M, Monteiro RC: **Inhibitory ITAMs as novel regulators of immunity**. *Immunol Rev* 2009, **232**(1):59-71.
11. Humphrey MB, Ogasawara K, Yao W, Spusta SC, Daws MR, Lane NE, Lanier LL, Nakamura MC: **The signaling adapter protein DAP12 regulates multinucleation during osteoclast development**. *J Bone Miner Res* 2004, **19**(2):224-234.

12. Koga T, Inui M, Inoue K, Kim S, Suematsu A, Kobayashi E, Iwata T, Ohnishi H, Matozaki T, Kodama T *et al*: **Costimulatory signals mediated by the ITAM motif cooperate with RANKL for bone homeostasis**. *Nature* 2004, **428**(6984):758-763.
13. Humphrey MB, Daws MR, Spusta SC, Niemi EC, Torchia JA, Lanier LL, Seaman WE, Nakamura MC: **TREM2, a DAP12-associated receptor, regulates osteoclast differentiation and function**. *J Bone Miner Res* 2006, **21**(2):237-245.
14. Negishi-Koga T, Takayanagi H: **Ca2+-NFATc1 signaling is an essential axis of osteoclast differentiation**. *Immunol Rev* 2009, **231**(1):241-256.
15. Takayanagi H, Kim S, Koga T, Nishina H, Isshiki M, Yoshida H, Saiura A, Isobe M, Yokochi T, Inoue J *et al*: **Induction and activation of the transcription factor NFATc1 (NFAT2) integrate RANKL signaling in terminal differentiation of osteoclasts**. *Dev Cell* 2002, **3**(6):889-901.
16. Sharma SM, Bronisz A, Hu R, Patel K, Mansky KC, Sif S, Ostrowski MC: **MITF and PU.1 recruit p38 MAPK and NFATc1 to target genes during osteoclast differentiation**. *J Biol Chem* 2007, **282**(21):15921-15929.
17. Kim K, Lee SH, Ha Kim J, Choi Y, Kim N: **NFATc1 induces osteoclast fusion via up-regulation of Atp6v0d2 and the dendritic cell-specific transmembrane protein (DC-STAMP)**. *Mol Endocrinol* 2008, **22**(1):176-185.
18. Yu M, Moreno JL, Stains JP, Keegan AD: **Complex regulation of tartrate-resistant acid phosphatase (TRAP) expression by interleukin 4 (IL-4): IL-4 indirectly suppresses receptor activator of NF-kappaB ligand (RANKL)-mediated TRAP expression but modestly induces its expression directly**. *J Biol Chem* 2009, **284**(47):32968-32979.
19. Matsumoto M, Kogawa M, Wada S, Takayanagi H, Tsujimoto M, Katayama S, Hisatake K, Nogi Y: **Essential role of p38 mitogen-activated protein kinase in cathepsin K gene expression during osteoclastogenesis through association of NFATc1 and PU.1**. *J Biol Chem* 2004, **279**(44):45969-45979.
20. Sundaram K, Nishimura R, Senn J, Youssef RF, London SD, Reddy SV: **RANK ligand signaling modulates the matrix metalloproteinase-9 gene expression during osteoclast differentiation**. *Exp Cell Res* 2007, **313**(1):168-178.
21. Tolar J, Teitelbaum SL, Orchard PJ: **Osteopetrosis**. *N Engl J Med* 2004, **351**(27):2839-2849.
22. Rachner TD, Khosla S, Hofbauer LC: **Osteoporosis: now and the future**. *Lancet* 2011, **377**(9773):1276-1287.
23. Scott DL, Wolfe F, Huizinga TW: **Rheumatoid arthritis**. *Lancet* 2010, **376**(9746):1094-1108.
24. Mundy GR, Raisz LG, Cooper RA, Schechter GP, Salmon SE: **Evidence for the secretion of an osteoclast stimulating factor in myeloma**. *N Engl J Med* 1974, **291**(20):1041-1046.

25. Yoneda T: **Cellular and molecular mechanisms of breast and prostate cancer metastasis to bone**. *Eur J Cancer* 1998, **34**(2):240-245.
26. Mii Y, Miyauchi Y, Morishita T, Miura S, Honoki K, Aoki M, Tamai S: **Osteoclast origin of giant cells in giant cell tumors of bone: ultrastructural and cytochemical study of six cases**. *Ultrastruct Pathol* 1991, **15**(6):623-629.
27. Nicholson GC, Malakellis M, Collier FM, Cameron PU, Holloway WR, Gough TJ, Gregorio-King C, Kirkland MA, Myers DE: **Induction of osteoclasts from CD14-positive human peripheral blood mononuclear cells by receptor activator of nuclear factor kappaB ligand (RANKL)**. *Clin Sci (Lond)* 2000, **99**(2):133-140.
28. Sorensen MG, Henriksen K, Schaller S, Henriksen DB, Nielsen FC, Dziegiel MH, Karsdal MA: **Characterization of osteoclasts derived from CD14+ monocytes isolated from peripheral blood**. *J Bone Miner Metab* 2007, **25**(1):36-45.
29. Youn MY, Takada I, Imai Y, Yasuda H, Kato S: **Transcriptionally active nuclei are selective in mature multinucleated osteoclasts**. *Genes Cells* 2010, **15**(10):1025-1035.
30. Saltman LH, Javed A, Ribadeneyra J, Hussain S, Young DW, Osdoby P, Amcheslavsky A, van Wijnen AJ, Stein JL, Stein GS *et al*: **Organization of transcriptional regulatory machinery in osteoclast nuclei: compartmentalization of Runx1**. *J Cell Physiol* 2005, **204**(3):871-880.
31. Rodriguez J, Vives L, Jorda M, Morales C, Munoz M, Vendrell E, Peinado MA: **Genome-wide tracking of unmethylated DNA Alu repeats in normal and cancer cells**. *Nucleic Acids Res* 2008, **36**(3):770-784.
32. Gallois A, Lachuer J, Yvert G, Wierinckx A, Brunet F, Rabourdin-Combe C, Delprat C, Jurdic P, Mazzorana M: **Genome-wide expression analyses establish dendritic cells as a new osteoclast precursor able to generate bone-resorbing cells more efficiently than monocytes**. *J Bone Miner Res* 2010, **25**(3):661-672.
33. Momparler RL: **Pharmacology of 5-Aza-2'-deoxycytidine (decitabine)**. *Semin Hematol* 2005, **42**(3 Suppl 2):S9-16.
34. Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, Agarwal S, Iyer LM, Liu DR, Aravind L *et al*: **Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1**. *Science* 2009, **324**(5929):930-935.
35. Kriaucionis S, Heintz N: **The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain**. *Science* 2009, **324**(5929):929-930.
36. Ito S, Shen L, Dai Q, Wu SC, Collins LB, Swenberg JA, He C, Zhang Y: **Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine**. *Science* 2011, **333**(6047):1300-1303.
37. Wu H, Zhang Y: **Mechanisms and functions of Tet protein-mediated 5-methylcytosine oxidation**. *Genes Dev* 2011, **25**(23):2436-2452.

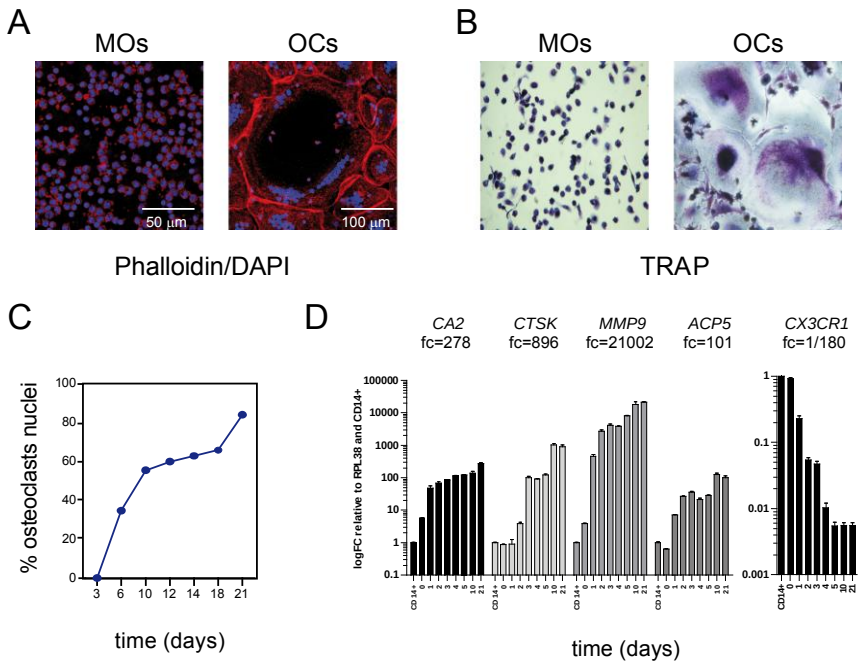
38. Williams K, Christensen J, Helin K: **DNA methylation: TET proteins-guardians of CpG islands?** *EMBO Rep* 2012, **13**(1):28-35.
39. Hernando H, Shannon-Lowe C, Islam AB, Al-Shahrour F, Rodriguez-Ubreva J, Rodriguez-Cortez VC, Javierre BM, Mangas C, Fernandez AF, Parra M *et al*: **The B cell transcription program mediates hypomethylation and overexpression of key genes in Epstein-Barr virus-associated proliferative conversion**. *Genome Biol* 2013, **14**(1):R3.
40. Stadler MB, Murr R, Burger L, Ivanek R, Lienert F, Scholer A, van Nimwegen E, Wirbelauer C, Oakeley EJ, Gaidatzis D *et al*: **DNA-binding factors shape the mouse methylome at distal regulatory regions**. *Nature* 2012, **480**(7378):490-495.
41. Edwards JR, Mundy GR: **Advances in osteoclast biology: old findings and new insights from mouse models**. *Nat Rev Rheumatol* 2011, **7**(4):235-243.
42. Chen W, Zhu G, Hao L, Wu M, Ci H, Li YP: **C/EBPalpha regulates osteoclast lineage commitment**. *Proc Natl Acad Sci U S A* 2013.
43. Smink JJ, Begay V, Schoenmaker T, Sterneck E, de Vries TJ, Leutz A: **Transcription factor C/EBPbeta isoform ratio regulates osteoclastogenesis through MafB**. *Embo J* 2009, **28**(12):1769-1781.
44. Shen L, Wu H, Diep D, Yamaguchi S, D'Alessio AC, Fung HL, Zhang K, Zhang Y: **Genome-wide analysis reveals TET- and TDG-dependent 5-methylcytosine oxidation dynamics**. *Cell* 2013, **153**(3):692-706.
45. Cortellino S, Xu J, Sannai M, Moore R, Caretti E, Cigliano A, Le Coz M, Devarajan K, Wessels A, Soprano D *et al*: **Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair**. *Cell* 2011, **146**(1):67-79.
46. Kallin EM, Rodriguez-Ubreva J, Christensen J, Cimmino L, Aifantis I, Helin K, Ballestar E, Graf T: **Tet2 facilitates the derepression of myeloid target genes during CEBPalpha-induced transdifferentiation of pre-B cells**. *Mol Cell* 2012, **48**(2):266-276.
47. Klug M, Schmidhofer S, Gebhard C, Andreesen R, Rehli M: **5-Hydroxymethylcytosine is an essential intermediate of active DNA demethylation processes in primary human monocytes**. *Genome Biol* 2013, **14**(5):R46.
48. Suzuki M, Yamada T, Kihara-Negishi F, Sakurai T, Hara E, Tenen DG, Hozumi N, Oikawa T: **Site-specific DNA methylation by a complex of PU.1 and Dnmt3a/b**. *Oncogene* 2006, **25**(17):2477-2488.
49. Hama M, Kirino Y, Takeno M, Takase K, Miyazaki T, Yoshimi R, Ueda A, Itoh-Nakadai A, Muto A, Igarashi K *et al*: **Bach1 regulates osteoclastogenesis in a mouse model via both heme oxygenase 1-dependent and heme oxygenase 1-independent pathways**. *Arthritis Rheum* 2012, **64**(5):1518-1528.

50. Stopka T, Amanatullah DF, Papetti M, Skoultchi AI: **PU.1 inhibits the erythroid program by binding to GATA-1 on DNA and creating a repressive chromatin structure**. *Embo J* 2005, **24**(21):3712-3723.
51. Hu R, Sharma SM, Bronisz A, Srinivasan R, Sankar U, Ostrowski MC: **Eos, MITF, and PU.1 recruit corepressors to osteoclast-specific genes in committed myeloid progenitors**. *Mol Cell Biol* 2007, **27**(11):4018-4027.
52. Vincent JJ, Huang Y, Chen PY, Feng S, Calvopina JH, Nee K, Lee SA, Le T, Yoon AJ, Faull K *et al*: **Stage-specific roles for tet1 and tet2 in DNA demethylation in primordial germ cells**. *Cell Stem Cell* 2013, **12**(4):470-478.
53. Azim AC, Wang X, Park GY, Sadikot RT, Cao H, Mathew B, Atchison M, van Breemen RB, Joo M, Christman JW: **NF-kappaB-inducing kinase regulates cyclooxygenase 2 gene expression in macrophages by phosphorylation of PU.1**. *J Immunol* 2007, **179**(11):7868-7875.
54. Bibikova M, Lin Z, Zhou L, Chudin E, Garcia EW, Wu B, Doucet D, Thomas NJ, Wang Y, Vollmer E *et al*: **High-throughput DNA methylation profiling using universal bead arrays**. *Genome Res* 2006, **16**(3):383-393.
55. Du P, Zhang X, Huang CC, Jafari N, Kibbe WA, Hou L, Lin SM: **Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis**. *BMC Bioinformatics* 2010, **11**:587.
56. The-R-Development-Core-Team: **R: A language and environment for statistical computing. R Foundation for Statistical Computing**. 2012.
57. Smyth GK: **Limma: linear models for microarray data. In: 'Bioinformatics and Computational Biology Solutions using R and Bioconductor'. R. Gentleman, V. Carey, S. Dudoit, R. Irizarry, W. Huber (eds)**. *Springer, New York* 2005:397-420.
58. Du P, Kibbe WA, Lin SM: **lumi: a pipeline for processing Illumina microarray**. *Bioinformatics* 2008, **24**(13):1547-1548.
59. Wang D, Yan L, Hu Q, Sucheston LE, Higgins MJ, Ambrosone CB, Johnson CS, Smiraglia DJ, Liu S: **IMA: an R package for high-throughput analysis of Illumina's 450K Infinium methylation data**. *Bioinformatics* 2012, **28**(5):729-730.
60. Aryee MJ, Wu Z, Ladd-Acosta C, Herb B, Feinberg AP, Yegnasubramanian S, Irizarry RA: **Accurate genome-scale percentage DNA methylation estimates from microarray data**. *Biostatistics* 2011, **12**(2):197-210.
61. Herman JG, Graff JR, Myohanen S, Nelkin BD, Baylin SB: **Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands**. *Proc Natl Acad Sci U S A* 1996, **93**(18):9821-9826.
62. Gautier L, Cope L, Bolstad BM, Irizarry RA: **affy--analysis of Affymetrix GeneChip data at the probe level**. *Bioinformatics* 2004, **20**(3):307-315.

63. Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B, Speed TP: **Summaries of Affymetrix GeneChip probe level data**. *Nucleic Acids Res* 2003, **31**(4):e15.
64. Al-Shahrour F, Diaz-Uriarte R, Dopazo J: **FatiGO: a web tool for finding significant associations of Gene Ontology terms with groups of genes**. *Bioinformatics* 2004, **20**(4):578-580.
65. Schones DE, Smith AD, Zhang MQ: **Statistical significance of cis-regulatory modules**. *BMC Bioinformatics* 2007, **8**:19.
66. Matys V, Fricke E, Geffers R, Gossling E, Haubrock M, Hehl R, Hornischer K, Karas D, Kel AE, Kel-Margoulis OV *et al*: **TRANSFAC: transcriptional regulation, from patterns to profiles**. *Nucleic Acids Res* 2003, **31**(1):374-378.
67. Perez-Llamas C, Lopez-Bigas N: **Gitools: analysis and visualisation of genomic data using interactive heat-maps**. *PLoS One* 2011, **6**(5):e19541.
68. Flicek P, Amode MR, Barrell D, Beal K, Brent S, Carvalho-Silva D, Clapham P, Coates G, Fairley S, Fitzgerald S *et al*: **Ensembl 2012**. *Nucleic Acids Res* 2012, **40**(Database issue):D84-90.

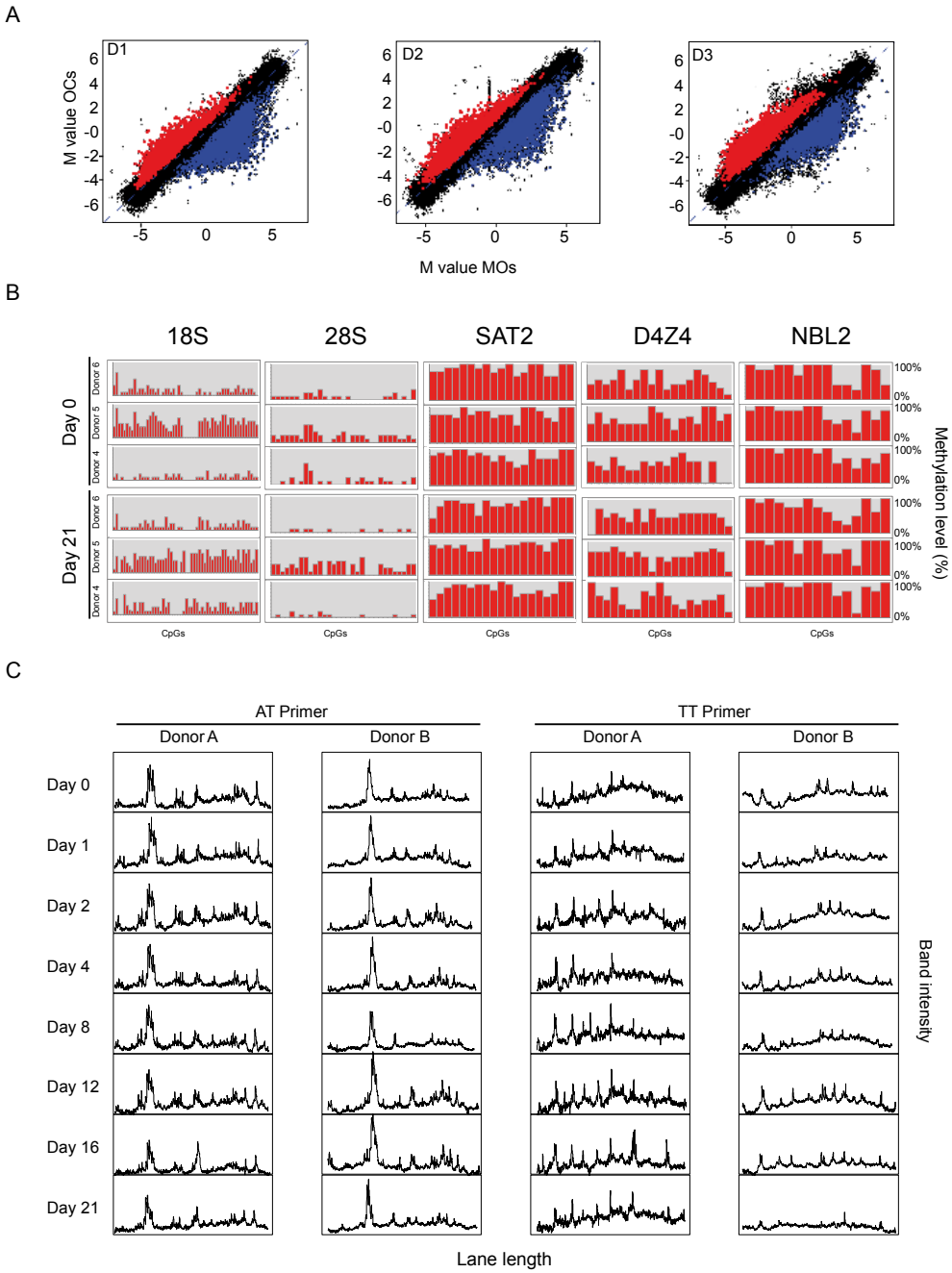
ADDITIONAL FIGURES

ADDITIONAL FILE 1



Add. Figure 1. (A) Visualization of the formation of the actin ring and the generation of polykaryons in monocyte (MO) to osteoclast (OC) differentiation with phalloidin and DAPI staining. (B) TRAP (Tartrate resistant acid phosphatase-OC marker) staining in MO and OC preparations, showing this activity only in OCs. (B) Determination of the typical percentage of osteoclastic nuclei present in the preparations used for the experiments; over 84% efficiency was achieved at 21 days. (C) Upregulation of OC specific markers (*CA2*, *CTSK*, *MMP9*, *ACP5*) was checked by qPCR; downregulation of a monocyte specificgene (*CX3CR1*) was also monitored.

ADDITIONAL FILE 3

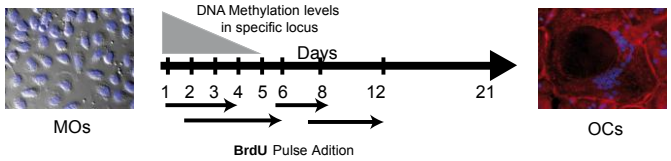


Add. Figure 2. (A) Scatterplots showing DNA methylation profiles of matching MO/OC pairs. Genes with significant differences (FC > 2, FDR < 0.05) in averaged results from three samples are highlighted in red (hypermethylated) or blue (hypomethylated). Three panels corresponding for each of the three individual comparisons of MO/OC pairs (D1, D2 and D3) are shown. (B) Bisulphite sequencing analysis of repetitive

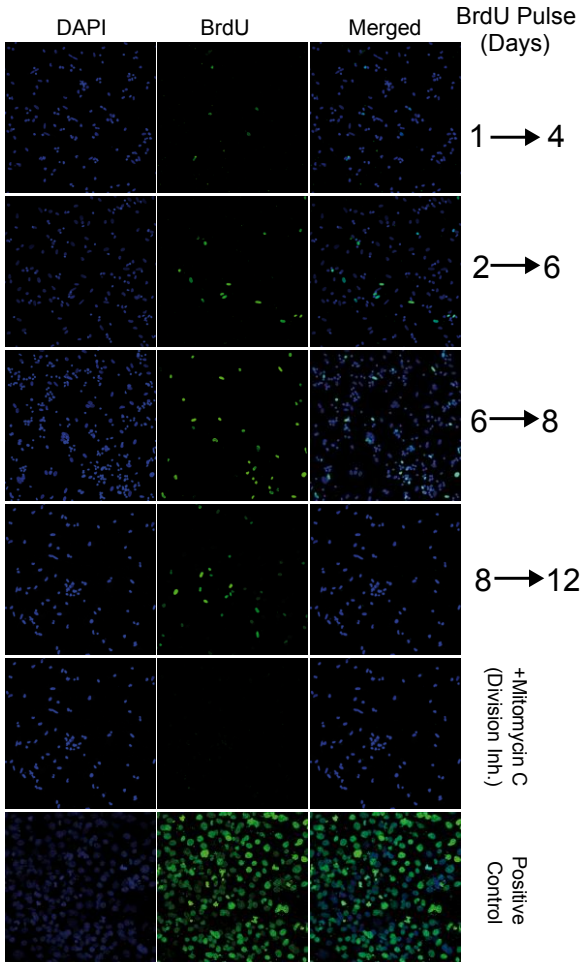
sequences performed on monocytes (day 0) and osteoclasts (day 21) from three different donors (donor A, donor B and donor C), showing no relevant differences in the DNA methylation levels. (C) AUMA (Amplification of Unmethylated Alus) analysis of two independent monocyte-to-osteoclast differentiation experiments. Graphs correspond to the scanned intensities of the bands obtained with two different sets of primers. No significant differences are observed.

ADDITIONAL FILE 8

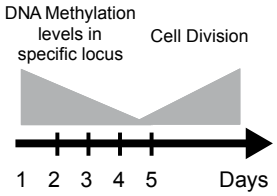
A



B

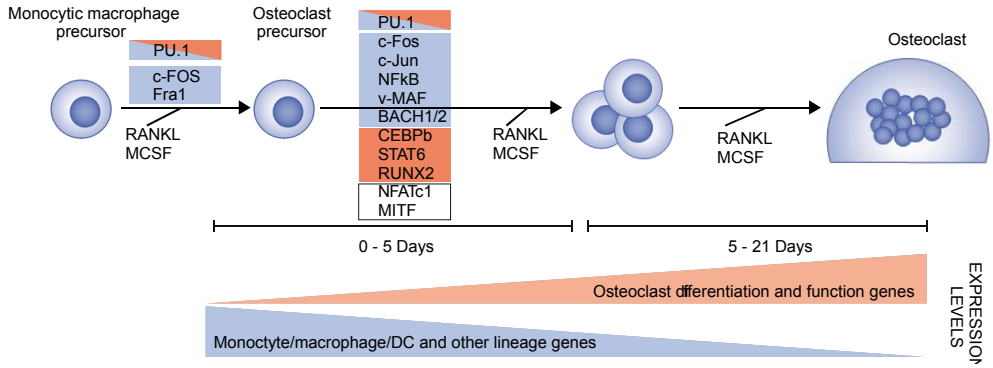


C



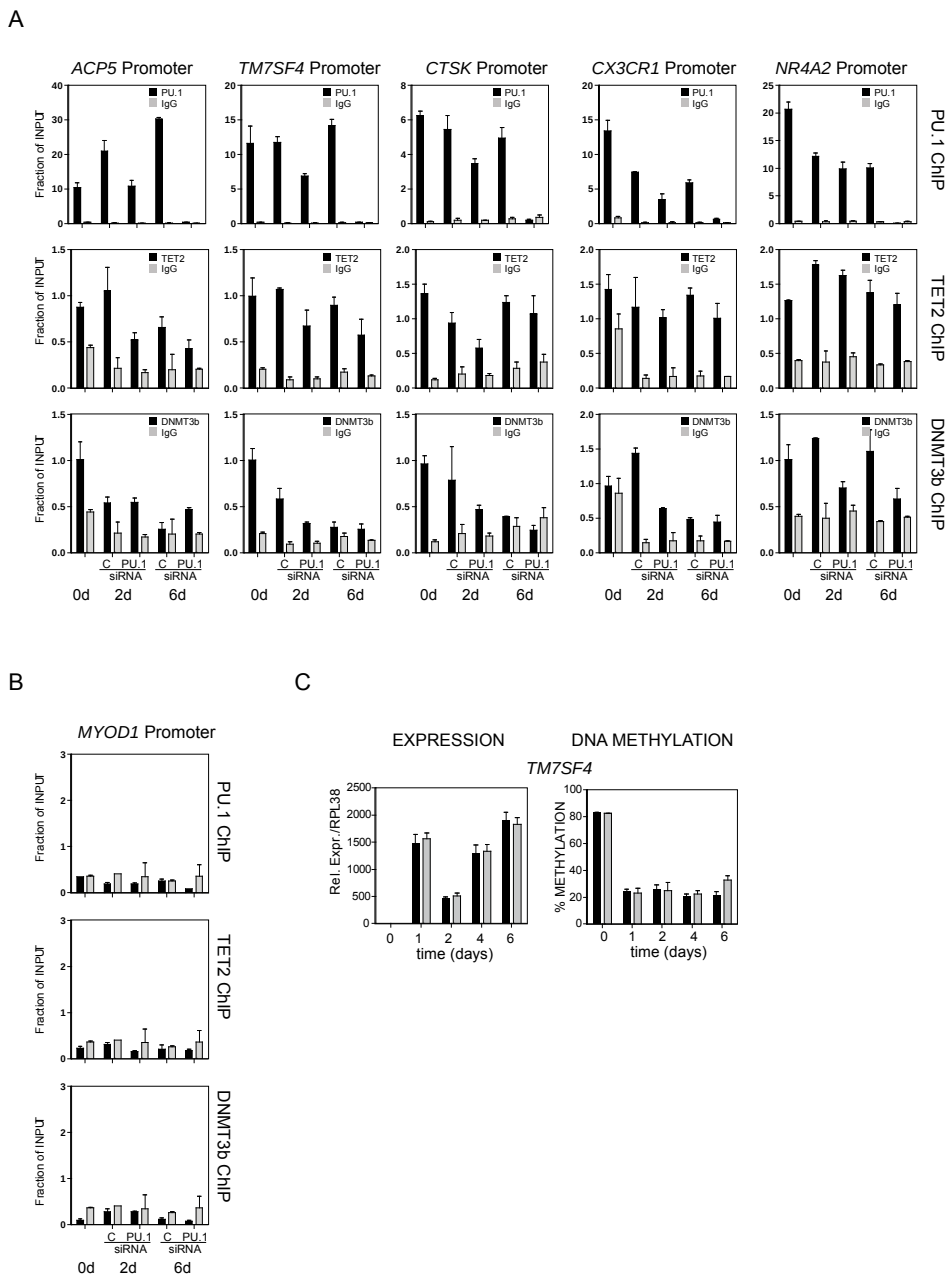
Supp. Figure 3. (A) Scheme showing the BrdU pulses added to monocytes differentiating into osteoclasts. (B) Representative immunofluorescence images at the selected time points showing BrdU positive cells. (C) Representation of the time scale where DNA demethylation occurs during osteoclast differentiation, together with the cell division observed at later time points.

ADDITIONAL FILE 9



Supp. Figure 4. Osteoclast differentiation scheme showing transcription factors that are known to be involved in monocyte-to-osteoclast differentiation. We have in red or blue the presence of binding motifs for those factors (according to 2 TRANSFAC analysis) among the sequences surrounding the CpGs that become hypo- or hypermethylated. Those arising from our analysis are highlighted in red and blue (associated with hypermethylation and hypomethylation, respectively).

ADDITIONAL FILE 10



Supp. Figure 5. (A) ChIP assays showing the effects of PU.1 downregulation in its recruitment, together with TET2 and DNMT3b binding to the same genes. Data were obtained at 0, 2 and 6 days after M-CSF/RANL stimulation. (B) We have used the MYOD1 promoter as a negative control. (C) Effects of PU.1 downregulation on expression and methylation of PU.1-target gene TM7SF4.

ADDITIONAL TABLES

Additional File 2. List of hypomethylated and hypermethylated genes during monocyte to osteoclast differentiation (FC<0.5 (hypomethylated, sheet 1) or FC>2

Additional File 4. Individual raw data corresponding to bisulfite pyrosequencing and standard bisulfite sequencing of individual MO and OC samples (Figure 1E), time course methylation data (Figure 2D, E) and PU.1 siRNA experiments (Figure 5D). Data are presented as supplied by PyroMark® Assay Design Software 2.0 for PyroMark Q96 MD (Qiagen), which automatically generates methylation percentages in a datasheet format.

Bisulfite pyrosequencing of individual MO and OC samples (Figure 1E):

CTSK	MOs	OCs
D1	94	6
D2	90	9
D3	93	26
D4	93	18
D5	93	11

Illumina ID: cg11946165

CpG Sequenced:
TCTTTAAAAGTAACCAAAAACAGCAGTCCTGGTTATTTATGACAGCACTTGAATCAATGC[CG]TAAGTT
CTGATGGACTCACATGTGACTCTGTTGCTAAAACTCTCAGGTGGTGGGATGCCC

ACP5	MOs	OCs	MOs	OCs	MOs	OCs
D1	82	45	85	9	100	6
D2	83	12	87	7	100	5
D3	81	38	84	22	76	23
D4	81	35	89	15	74	14
D5	84	29	82	10	72	8

Illumina ID: cg21207418

CpG Sequenced:
CTCGGCCCACACAGCCTCCGGGTGGACCTGCAGGGGCCTGTTTGTGCTGTAGGCTTGACA[CG]TCCAG
GTATCTCTGTGTGTCTGTGTATCTCAGTGTGAGTGTGTGTGTGTGTGCACACTTG

TM7SF4/DC STAMP	MOs	OCs
D1	87	9
D2	88	18
D3	82	33
D4	85	15
D5	87	9

Illumina ID: cg01136183

CpG Sequenced:
AAGGAACCCCATCCAGTCCAGCCGGTTGGCTTGCTCCTCCCCTCCTCCCACTCCAGTTCA[CG]CTCCAG
CCCACTGAAGAGTGGTGCCCACCCCTAGGCCCCTGCCTAAATGGCTCTTCCAGA

CX3CR1	MOs	OCs
D1	24	52
D2	26	56
D3	34	63
D4	32	65
D5	22	55

Illumina ID: cg04569233

CpG Sequenced:

TAGCTGTCCACTGCTCCACCCCACCCACAGGTACCCAACTAGTCCTGTGCACCTACCTGG[CG]TGGACT
GCCAAGGGAACCTCTGGATCTGCCAGTCAGCCACCCTGTCCTGCTCAGACTTTA

FOXP1	MOs	OCs
D1	17	41
D2	7	25
D3	6	53
D4	10	46
D5	8	31

Illumina ID: cg02520804

CpG Sequenced:

TTGGGTAGCATTCTCCTCATAAAGAAGGATACATTAAAAAAAATAACTTGTTTCGCGACT[CG]GCATCC
ATAAGGAACTCAAATGCTGCCCAGAGAGGGGCTGAGTATTTCCTTCCAAGTGAG

Correlation of Array data and pyrosequencing data (Data of Figure 1E, last right):

		CTSK		ACP5		TM7SF4		CX3CR1	
		PYROSEQ	M Value (Illumina)	PYROSEQ (%)	M Value (Illumina)	PYROSEQ	M Value (Illumina)	PYROSEQ	M Value (Illumina)
Mos	**D1**	94	0,89	82	0,82	87	0,85	24	0,40
Mos	**D2**	90	0,90	83	0,82	88	0,88	26	0,35
Mos	**D3**	93	0,87	81	0,84	82	0,85	34	0,40
Ocs	**D1**	6	0,16	45	0,54	9	0,18	52	0,61
Ocs	**D2**	9	0,16	12	0,26	18	0,25	56	0,63
Ocs	**D3**	26	0,50	38	0,45	33	0,46	63	0,74

Bisulfite sequencing of Repetitive regions in individual MO and OC samples (Figure 1G):

18S	**MONOCYTES**			**OSTEOLASTS**		
%Meth	MOs_D4	MOs_D5	MOs_D6	OCs_D4	OCs_D5	OCs_D6
CpG1	30.00	33.33	10.00	10.00	10.00	10.00
CpG2	70.00	88.89	20.00	50.00	50.00	50.00
CpG3	0.00	44.44				
CpG4	10.00	44.44	10.00	10.00	40.00	
CpG5	10.00	22.22	0.00	10.00	50.00	62.50
CpG6	10.00	44.44	0.00	10.00	30.00	28.57
CpG7	20.00	33.33	0.00	0.00	20.00	14.29
CpG8	20.00	22.22	0.00	10.00	10.00	12.50
CpG9	50.00	77.78	10.00	10.00	60.00	37.50
CpG10	10.00	44.44	0.00	10.00	40.00	12.50
CpG11	20.00	33.33	20.00	20.00	50.00	37.50
CpG12	20.00	33.33	10.00	10.00	50.00	37.50
CpG13	10.00	33.33	10.00	30.00	50.00	12.50
CpG14	30.00	55.56	10.00	20.00	40.00	12.50
CpG15	20.00	66.67	0.00	10.00	50.00	12.50
CpG16	20.00	77.78	10.00	30.00	50.00	25.00
CpG17	10.00	66.67	10.00	10.00	30.00	25.00
CpG18	20.00	50.00	10.00	10.00	40.00	25.00
CpG19	0.00	50.00	0.00	10.00	40.00	12.50
CpG20	10.00	40.00	0.00	10.00	40.00	50.00
CpG21	20.00	30.00	20.00	40.00	44.44	25.00
CpG22	20.00	20.00	10.00	10.00	75.00	37.50
CpG23	10.00	20.00	10.00	20.00	62.50	
CpG24	20.00	37.50	20.00	20.00	37.50	
CpG25	0.00	40.00	10.00	0.00	37.50	
CpG26	30.00	60.00	20.00	10.00		
CpG27	10.00	37.50	10.00	10.00	33.33	
CpG28					66.67	12.50
CpG29						50.00
CpG30						25.00
CpG31					60.00	25.00
CpG32					60.00	50.00
CpG33	11.11		10.00	10.00	60.00	37.50
CpG34	20.00	50.00	10.00	40.00	70.00	12.50
CpG35	0.00	50.00	10.00	30.00	60.00	25.00

CpG36	10.00	60.00	0.00	0.00	30.00	12.50
CpG37	0.00	50.00	30.00	30.00	70.00	37.50
CpG38	10.00	30.00	0.00	10.00	50.00	25.00
CpG39	0.00	60.00	0.00	0.00	40.00	12.50
CpG40	0.00	50.00	10.00	10.00	50.00	25.00
CpG41	10.00	40.00	10.00	10.00	50.00	12.50
CpG42	10.00	50.00	20.00	20.00	50.00	37.50
CpG43	20.00	70.00	10.00	10.00	70.00	37.50
CpG44	20.00	70.00	20.00	20.00	70.00	37.50
CpG45	20.00	40.00	0.00	10.00	50.00	25.00
CpG46	10.00	60.00	10.00	10.00	40.00	12.50
CpG47	10.00	60.00	10.00	20.00	50.00	25.00
CpG48	20.00	40.00	20.00	10.00	40.00	12.50
CpG49	30.00	50.00	20.00	30.00	70.00	12.50
CpG50	10.00	40.00	0.00	10.00	30.00	37.50
CpG51	20.00	60.00	10.00	30.00	70.00	25.00
CpG52	10.00	40.00	10.00	20.00	60.00	37.50
CpG53	20.00	50.00	30.00	20.00	70.00	37.50
CpG54	20.00	50.00	20.00	10.00	30.00	37.50
CpG55	10.00	50.00	10.00	10.00	40.00	12.50
CpG56	0.00	40.00	10.00	10.00	70.00	37.50

28S	**MONOCYTES**			**OSTEOLASTS**		
%Meth	MOs_D4	MOs_D5	MOs_D6	OCs_D4	OCs_D5	OCs_D6
CpG1	10.00	20.00			30.00	
CpG2	10.00	10.00			30.00	8.33
CpG3	10.00	20.00	10.00		10.00	
CpG4	10.00	20.00			20.00	
CpG5	10.00	20.00	20.00	10.00	30.00	16.67
CpG6	10.00	20.00		10.00	20.00	
CpG7		10.00	10.00		50.00	8.33
CpG8	20.00	50.00	60.00	10.00	50.00	
CpG9	20.00	50.00	40.00	10.00	10.00	8.33
CpG10	10.00	30.00			30.00	
CpG11	30.00	20.00	10.00		40.00	16.67
CpG12	10.00					8.33
CpG13			10.00	10.00	40.00	8.33
CpG14	10.00	10.00			30.00	
CpG15	10.00	20.00			40.00	
CpG16		30.00			40.00	
CpG17	10.00		20.00			
CpG18		20.00			30.00	
CpG19		20.00	10.00		10.00	0.00
CpG20		20.00	20.00	10.00	50.00	8.33
CpG21		10.00	10.00			
CpG22			10.00			
CpG23						
CpG24		20.00	20.00	10.00	30.00	
CpG25	10.00	20.00	10.00		20.00	
CpG26	10.00	20.00	10.00		20.00	

CpG27	20.00	20.00			10.00	8.33
CpG28		10.00		10.00	10.00	
CpG29	10.00	10.00	10.00		10.00	
CpG30		20.00		10.00	30.00	
CpG31	30.00	10.00	20.00		30.00	8.33

SAT2	**MONOCYTES**			**OSTEOLASTS**		
%Meth	MOs_D4	MOs_D5	MOs_D6	OCs_D4	OCs_D5	OCs_D6
CpG1	80.00	80.00	83.33	41.67	83.33	50.00
CpG2	80.00	80.00	91.67	75.00	75.00	66.67
CpG3	90.00	80.00	83.33	91.67	83.33	83.33
CpG4	90.00	70.00	100.00	91.67	75.00	91.67
CpG5	100.00	100.00	100.00	83.33	83.33	91.67
CpG6	100.00	80.00	83.33	83.33	66.67	83.33
CpG7	88.89	90.00	91.67	50.00	66.67	91.67
CpG8	100.00	90.00	83.33	91.67	100.00	91.67
CpG9	77.78	70.00	66.67	75.00	75.00	75.00
CpG10	88.89	80.00	83.33	75.00	75.00	100.00
CpG11	100.00	80.00	66.67	83.33	83.33	91.67
CpG12	66.67	50.00	50.00	75.00	83.33	58.33
CpG13	77.78	100.00	58.33	91.67	91.67	66.67
CpG14	100.00	100.00	100.00	100.00	91.67	66.67
CpG15	100.00	90.00	75.00	100.00	83.33	75.00
CpG16	66.67	70.00	75.00	75.00	66.67	58.33
CpG17	66.67	60.00	75.00	100.00	91.67	66.67
CpG18	100.00	100.00	100.00	100.00	100.00	100.00
CpG19	100.00	100.00	100.00	100.00	100.00	100.00

D4Z4	**MONOCYTES**			**OSTEOLASTS**		
%Meth	MOs_D4	MOs_D5	MOs_D6	OCs_D4	OCs_D5	OCs_D6
CpG1	42.86	50.00	62.50		71.43	100.00
CpG2	57.14	83.33	50.00	75.00	71.43	62.50
CpG3	42.86	50.00	37.50	50.00	71.43	50.00
CpG4	57.14	66.67	75.00	62.50	85.71	87.50
CpG5	85.71	50.00	50.00	50.00	71.43	37.50
CpG6	28.57	50.00	37.50	37.50	57.14	25.00
CpG7	57.14	50.00	37.50	75.00	71.43	25.00
CpG8	85.71	50.00	62.50	37.50	57.14	75.00
CpG9	42.86	100.00	50.00	75.00	14.29	62.50
CpG10	85.71	83.33	62.50	50.00	57.14	100.00
CpG11	28.57	66.67	50.00	50.00	42.86	62.50
CpG12	42.86	33.33	75.00	50.00	57.14	37.50
CpG13	42.86	66.67	87.50	62.50	57.14	25.00
CpG14	57.14	71.43	62.50	62.50	57.14	50.00
CpG15	85.71	50.00	62.50	62.50	71.43	37.50
CpG16	71.43	100.00		62.50	71.43	50.00
CpG17	50.00	100.00	62.50	62.50	80.00	50.00
CpG18	33.33	60.00		50.00	71.43	66.67
CpG19	14.29	80.00		25.00	14.29	16.67

NBL2	MONOCYTEs			OSTEOLASTS		
%Meth	MOs_D4	MOs_D5	MOs_D6	OCs_D4	OCs_D5	OCs_D6
CpG1	100.00	87.50	100.00	100.00	100.00	87.50
CpG2	85.71	100.00	100.00	87.50	100.00	87.50
CpG3	85.71	100.00	100.00	100.00	87.50	100.00
CpG4	100.00	87.50	85.71	87.50	100.00	100.00
CpG5	100.00	87.50	100.00	75.00	100.00	87.50
CpG6	71.43	87.50	100.00	75.00	62.50	87.50
CpG7	100.00	100.00	100.00	100.00	100.00	100.00
CpG8	100.00	100.00	85.71	75.00	100.00	100.00
CpG9	100.00	62.50	100.00	62.50	62.50	100.00
CpG10	42.86	50.00	57.14	37.50	62.50	62.50
CpG11	42.86	62.50	71.43	25.00	75.00	75.00
CpG12	28.57	25.00	42.86	50.00	28.57	12.50
CpG13	100.00	87.50	71.43	100.00	100.00	87.50
CpG14	85.71	62.50	57.14	62.50	100.00	75.00
CpG15	42.86	87.50	85.71	100.00	100.00	100.00

Summary Table time course: HYPOMETHYLATED GENES (Figure 2D)

ACP5_CPG1		ACP5_CPG2		ACP5_CPG3		CTSK		TM7SF4/DC STAMP		TM4SF19								IL7R		CD59	
%METH MEAN	SEM	%METH MEAN	SEM	%METH MEAN	SEM	%METH MEAN	SEM	%METH MEAN	SEM	%METH MEAN	SEM	%METH MEAN	SEM	%METH MEAN	SEM	%METH MEAN	SEM	%METH MEAN	SEM	%METH MEAN	SEM
80.83	2.14	85.50	7.82	70.33	2.34	92.17	0.54	80.50	0.96	81.46	1.36	87.73	5.59	84.21	2.92	63.26	2.24	34.64	4.06	61.75	6.18
80.67	0.52	68.17	8.01	52.50	5.79	90.00	1.06	56.67	6.17	65.62	1.63	78.72	6.41	64.93	8.16	42.26	9.28	36.86	1.53	38.13	8.80
75.00	10.84	32.50	7.40	27.67	4.56	85.83	1.58	37.67	6.38	56.11	5.07	70.47	12.07	54.40	12.70	28.39	12.47	31.72	3.36	31.39	10.00
60.17	5.27	16.67	3.14	13.83	2.48	57.17	4.21	25.50	4.58	39.75	1.94	53.94	3.78	37.51	3.80	19.71	2.08	27.06	3.12	16.45	3.54
58.67	8.19	17.50	5.05	18.17	5.00	42.67	3.79	24.83	2.52	35.38	2.92	45.99	8.93	30.36	8.66	16.53	7.02	26.35	3.17	18.90	2.39
52.40	13.63	10.67	4.13	9.00	4.98	38.50	5.53	24.00	5.87	27.19	2.17	34.97	4.53	22.88	2.14	10.57	2.99	24.35	4.47	16.41	8.93
47.17	4.07	8.67	1.86	6.50	2.81	31.67	6.15	19.67	5.57	28.87	3.46	47.63	8.21	30.59	9.02	14.61	4.97	22.52	3.27	9.91	3.91
46.20	7.01	7.33	2.66	5.50	2.26	26.50	5.21	24.50	3.32	24.34	3.22	40.55	5.07	23.51	2.93	10.41	4.45	22.40	2.27	9.96	4.62
43.17	5.53	11.00	3.46	7.83	3.31	20.33	4.00	15.50	2.67	22.05	1.20	33.73	4.67	19.54	4.15	8.46	3.26	23.28	2.23	9.38	2.80
40.33	4.63	7.83	1.60	4.83	1.17	18.33	3.53	13.50	1.84	17.09	0.61	30.15	3.79	14.84	3.30	5.34	2.44	22.18	2.21	6.24	2.32
38.50	5.51	9.00	1.83	7.00	0.82	20.75	4.66	8.50	3.62	18.92	3.06	28.82	5.50	13.32	4.70	6.39	3.20	22.70	2.82	10.50	3.92
42.00	7.29	5.67	2.80	3.17	2.40	15.33	3.85	9.17	2.23	22.91	4.85	33.89	11.20	13.77	3.70	5.64	2.73	24.14	2.53	6.69	2.54
42.50	1.73	6.75	1.26	5.25	2.50	15.50	3.75	7.75	1.03	25.80	5.49	26.27	5.89	18.15	12.79	8.76	7.27	26.08	2.35	7.80	3.26
36.50	4.64	6.83	2.64	3.83	1.17	10.33	1.71	10.33	1.50	17.31	2.54	27.05	5.66	16.23	8.96	7.85	7.10	23.09	2.15	4.60	2.03
36.00	4.38	4.33	1.51	4.33	3.14	7.67	1.87	10.33	2.32	12.74	1.05	24.02	4.75	9.65	3.13	2.69	2.09	24.97	3.07	4.93	4.43
33.60	5.98	4.60	1.95	1.50	1.00	6.50	2.62	5.00	1.52	11.59	6.12	21.60	9.45	13.15	9.72	3.61	2.18	16.17	4.93	4.75	6.21

Summary Table time course: HYPERMETHYLATED GENES (Figure 2D)

PPP1R16B		CD6						NR4A2								CD22		SNCG		CX3CR1	
%METH MEAN	SEM	%METH MEAN	SEM	%METH MEAN	SEM	%METH MEAN	SEM	%METH MEAN	SEM	%METH MEAN	SEM	%METH MEAN	SEM	%METH MEAN	SEM	%METH MEAN	SEM	%METH MEAN	SEM	%METH MEAN	SEM
0.94	0.20	0.58	0.21	0.60	0.36	0.58	0.50	46.24	1.43	46.99	9.99	53.46	6.70	48.18	3.03	25.47	1.37	32.64	6.38	29.01	7.14
0.97	0.31	0.42	0.18	0.56	0.14	0.42	0.43	41.74	4.96	44.09	12.04	50.11	8.99	44.63	11.37	16.50	3.23	26.41	13.27	38.87	3.36
4.82	2.48	0.50	0.20	0.59	0.17	0.50	0.49	38.54	2.77	38.24	9.46	46.92	8.38	37.96	10.16	17.37	3.85	17.54	6.42	37.04	4.04
5.59	1.26	1.85	0.64	1.87	1.44	1.85	1.56	53.64	2.22	47.27	9.19	54.44	5.98	46.93	5.21	15.85	0.54	25.14	3.65	40.59	5.65
9.64	1.83	5.95	1.28	5.35	2.18	5.95	3.13	54.17	3.31	53.24	8.15	56.16	11.60	52.21	7.00	15.55	3.03	24.42	4.50	39.94	7.15
17.12	2.79	8.68	2.15	4.38	2.86	8.68	5.27	67.46	7.51	50.31	8.29	67.66	16.70	53.26	4.76	22.76	1.85	29.39	9.51	38.14	2.87
18.70	3.78	16.22	1.45	9.08	3.78	16.22	3.55	61.48	6.82	56.95	11.75	68.45	10.81	58.71	9.15	24.59	2.57	38.17	11.53	40.36	1.80
26.57	5.34	22.23	2.34	12.19	2.26	22.23	5.23	65.15	3.98	63.07	12.10	70.08	7.00	64.30	11.14	30.37	2.12	47.10	4.93	45.73	3.71
23.94	2.07	22.36	1.84	13.70	2.14	22.36	4.51	63.64	3.13	62.89	8.38	70.85	8.88	60.86	8.24	31.26	2.23	37.56	9.06	52.67	7.88
28.38	2.80	28.82	4.77	16.82	8.99	28.82	11.68	64.34	5.71	68.02	8.38	73.13	10.45	63.15	8.20	33.82	1.95	45.51	8.84	50.74	6.66
27.69	6.61	28.43	4.01	20.15	6.82	28.43	8.01	64.55	2.68	61.22	4.36	66.57	4.69	58.02	5.00	32.27	2.21	41.74	6.04	51.07	8.04
26.06	2.81	29.44	2.37	20.34	8.68	29.44	5.80	72.86	7.60	64.32	9.73	76.99	8.79	64.26	9.17	29.56	4.35	42.95	19.34	52.90	4.26
21.08	1.20	27.80	1.33	19.57	4.81	27.80	2.65	63.00	2.98	63.77	4.49	70.71	7.94	58.17	2.26	34.72	1.95	45.07	7.20	56.50	7.73
32.86	1.44	32.92	1.04	23.19	7.17	32.92	2.32	71.25	3.68	71.55	8.96	69.82	13.53	65.44	10.25	35.27	3.02	48.82	14.30	51.79	6.64
30.72	4.12	41.72	3.77	22.05	6.47	41.72	9.23	67.31	5.02	69.73	6.74	72.84	6.76	68.91	9.14	39.33	2.52	55.04	11.57	54.97	5.42
45.90	9.65	51.17	5.96	26.22	7.33	51.17	13.33	70.33	6.33	65.96	15.90	66.09	11.37	63.41	6.80	41.97	4.88	59.97	9.01	61.50	7.97

Pyroseq. PU.1 siRNA (Fig 5D)

	siRNA control		siRNA PU.1		
ACP5	**Mean**	**SEM**	**Mean**	**SEM**	**P-VALUE**
Day 0	92.46	1.07			
Day 1	72.34	2.22	76.87	1.10	0.069
Day 2	55.17	1.27	65.81	3.97	0.026
Day 4	31.90	4.52	58.33	2.50	0.006
Day 6	37.45	1.78	54.06	3.21	0.033

	siRNA control		siRNA PU.1		
CTSK	**Mean**	**SEM**	**Mean**	**SEM**	**P-VALUE**
Day 0	92.32	2.34			
Day 1	87.14	2.19	85.80	2.39	0.21241
Day 2	60.98	4.23	82.96	1.43	0.03224
Day 4	59.11	5.37	82.33	0.30	0.01315
Day 6	45.33	5.55	77.18	1.52	0.01145

	siRNA control		siRNA PU.1		
CX3CR1	**Mean**	**SEM**	**Mean**	**SEM**	**P-VALUE**
Day 0	29.42	2.45			
Day 1	37.40	0.90	33.43	2.62	0.20493
Day 2	40.98	0.28	29.50	0.89	0.00328
Day 4	42.18	0.03	30.38	0.35	0.00006
Day 6	42.36	5.29	34.89	0.35	0.17717

	siRNA control		siRNA PU.1		
NR4A2	**Mean**	**SEM**	**Mean**	**SEM**	**P-VALUE**
Day 0	28.46	1.24			
Day 1	27.46	1.31	25.05	1.03	0.0925
Day 2	27.32	0.49	27.71	1.47	0.4123
Day 4	43.46	1.10	29.14	1.37	0.0009
Day 6	46.61	1.81	34.21	0.58	0.0053

	siRNA control		siRNA PU.1		
PLA2G4E	**Mean**	**SEM**	**Mean**	**SEM**	**P-VALUE**
Day 0	88.30	0.13			
Day 1	70.97	9.22	70.42	1.93	0.4722
Day 2	68.82	3.03	64.05	2.42	0.1518
Day 4	57.25	10.39	53.64	5.59	0.3772
Day 6	36.12	8.06	36.65	0.78	0.4838

Additional file 5. List of CpG clusters that undergo hyper or hypomethylation coordinately

Additional file 6. Differentially expressed genes between Mos, OC samples at 5 days and OC samples at 20 days after RANKL/M-CSF stimulation (FC>2, FC<0.5; FDR<0.05)

Additional file 7. List of genes whose expressiois potentially regulated by DNA methylation

Additional file 11. List of primers

Primer Type	Primer ID	Primer Sequence
BS Pyrosequencing	PYRO_CTSK_F	TGGGGATTTTAATTGAGATAGATGA
BS Pyrosequencing	PYRO_CTSK_R	[Btn]ACCCAAAAAATCCAAATAAAACTATCTT
BS Pyrosequencing	PYRO_CTSK_Seq	GAGTTATATGTGAGTTTATTAGAA
BS Pyrosequencing	PYRO_ACP5_F	AGGAGATGTGTTTTGGGTAATT
BS Pyrosequencing	PYRO_ACP5_R	[Btn]CCCTCCCCTTAAATCATATAAACC
BS Pyrosequencing	PYRO_ACP5_Seq	ATATATAGATATATAGAGATATTTG
BS Pyrosequencing	PYRO_DCSTAMP_TM7SF4_F	TGTTTGGGGTTATGAGTGTAG
BS Pyrosequencing	PYRO_DCSTAMP_TM7SF4_R	[Btn]TTACCCTCACTCCCATACT
BS Pyrosequencing	PYRO_DCSTAMP_TM7SF4_Seq	GGTTATGAGTGTAGAGG
BS Pyrosequencing	PYRO_CX3CR1_F	ATATTYGTTTTTGGTAAAGTTTGAGTAGGA
BS Pyrosequencing	PYRO_CX3CR1_R	[Btn]ACATTATTAACCTATTAACTATCCACTACT
BS Pyrosequencing	PYRO_CX3CR1_Seq	AGAGGTTTTTTTGGTAGT
BS Pyrosequencing	PYRO_TM4SF19_F	ATATGAAAATGAGTAGAGGGTGGGTTATTA
BS Pyrosequencing	PYRO_TM4SF19_R	[Btn]ACCTAAAAATTATCTTTCCAAAACTCTT
BS Pyrosequencing	PYRO_TM4SF19_Seq	AGGTTTGTGGGTAGG
BS Pyrosequencing	PYRO_CD22_F	TAGAGAAGTAGGGGGTGTGGTTATG
BS Pyrosequencing	PYRO_CD22_R	[Btn]TCCCAACTCTAAAAAATATACCTAACC
BS Pyrosequencing	PYRO_CD22_Seq	ATTTTGTATTTAATAAGTAAGTT
BS Pyrosequencing	PYRO_SNCG_F	GGTGGGGTAGGTTTAGTTTATATTT
BS Pyrosequencing	PYRO_SNCG_R	[Btn]ACCCCATATCTACCACATTCC
BS Pyrosequencing	PYRO_SNCG_Seq	GGTAAGTAGTTTTAGAAATTGT
BS Pyrosequencing	PYRO_CD6_F	TTTTTGGGGTGTAGTTTGGATGGG
BS Pyrosequencing	PYRO_CD6_R	[Btn]ACTCTACCCTTTACTATTCTTATTCCTAT
BS Pyrosequencing	PYRO_CD6_Seq	ATAGGTTGGGTTTGAT
BS Pyrosequencing	PYRO_NR4A2_F	GTGGGGAGGAATYGTAGATTTTAGTTATA
BS Pyrosequencing	PYRO_NR4A2_R	[Btn]ATCCCAACAACCAAACACTTC
BS Pyrosequencing	PYRO_NR4A2_Seq	AGTAAATAAAAATTGTTTAGTGGA
BS Pyrosequencing	PYRO_PPP1R16B_F	TAGAAAGAGGTTTGAATGAGGTGATAGA
BS Pyrosequencing	PYRO_PPP1R16B_R	[Btn]AAACCAAAAACTCAAAATTCCTAACTT
BS Pyrosequencing	PYRO_PPP1R16B_Seq	AGAGGTAYGATTATTTT
BS Pyrosequencing	Hs_PYRO_IL7R_F2	ATGATAATTTTAGGTATAATTTTTGGTATG
BS Pyrosequencing	Hs_PYRO_IL7R_R2	[Btn]TCACCATTTTAAACATAACCACTTTC
BS Pyrosequencing	Hs_PYRO_IL7R_Seq2	GGTATGGTTTTTTTTTATTTTAAG
BS Pyrosequencing	Hs_PYRO_CD59_F	TGAGGAGTTAGAGTTTTTAGGTATGT
BS Pyrosequencing	Hs_PYRO_CD59_R	[Btn]TCATATAACCACTATAATTCCCACTCT
BS Pyrosequencing	Hs_PYRO_CD59_Seq	GATTTATTTAGTGTTGTGGT
BS Pyrosequencing	PYRO_FSCN3_F	AGTTTATTTTGGTGTTTAAAGGGGAATAG
BS Pyrosequencing	PYRO_FSCN3_R	[Btn]AAAAAACCCTAAC**R**CAACAACATTAATCC

RT-PCR	Hs_CTSK_RT_F	ACAACAAGGTGGATGAAATCTC
RT-PCR	Hs_CTSK_RT_R	GCCAGTTCATATGTATGGACAC
RT-PCR	Hs_MMP9_RT_F	TCACTTTCCTGGGTAAGGAG
RT-PCR	Hs_MMP9_RT_R	GAACAAACTGTATCCTTGGTC
RT-PCR	Hs_CA2_RT_F	CTTGGATTACTGGACCTACCC
RT-PCR	Hs_CA2_RT_R	GGAATTTCAACACCTGCTCG
RT-PCR	Hs_TRAcP_RT_F	GATCACAATCTGCAGTACCTG
RT-PCR	Hs_TRAcP_RT_R	TTCAGTCCCATAGTGGAAGC
RT-PCR	Hs_CX3CR1_RT2_F	CACAAAGGAGCAGGCATGGAAG
RT-PCR	Hs_CX3CR1_RT2_R	CAGGTTCTCTGTAGACACAAGGC
RT-PCR	RT_TM4SF19_F	GCCTTTCTCCAGGTTCTGTCT
RT-PCR	RT_TM4SF19_R	CAAGGCTCAGTCCCAGGATA
RT-PCR	RT_CD22_F	CTCCTTTTGCTCTCAGATGCT
RT-PCR	RT_CD22_R	GAGATGCATGGTGTCGTGTC
RT-PCR	RT_SNCG_F	ACGGAAGCAGCTGAGAAGAC
RT-PCR	RT_SNCG_R	CTGAGGTCACGCTCTGTACAAC
RT-PCR	RT_CD6_F	ACCAGCTCAACACCAGCAGT
RT-PCR	RT_CD6_R	GCTGCTCCCGTTTGTCAG
RT-PCR	RT_NR4A2_F	cgctatcaattttcttctgttaaatg
RT-PCR	RT_NR4A2_R	gattcctccccacaaacaaa
RT-PCR	RT_PPP1R16B_F	GCCGCAAGAAAGTGTCCTT
RT-PCR	RT_PPP1R16B_R	GGGCTGACCTTATTCTTCAGG
RT-PCR	Hs_RT_IL7R_F	CGTTTCTGGAGAAAGTGGCTA
RT-PCR	Hs_RT_IL7R_R	GCGATCCATTCACTTCCAAC
RT-PCR	Hs_CTSK_ChIP_F	TAATTCCCTACCCTGGCACA
RT-PCR	Hs_CTSK_ChIP_R	CGGGGATAGAAATGCTGAGA
RT-PCR	Hs_RT_CD59_F	CACAATGGGAATCCAAGGAG
RT-PCR	Hs_RT_CD59_R	TGCAGTCAGCAGTTGGGTTA
RT-PCR	RT_FSCN3_F	CCTTTGAGGCATGCAAGAAT
RT-PCR	RT_FSCN3_R	TCTCATGCTCATTGCTCACC
RT-PCR	Hs_RT_PLA2G4E_F	ACACAGGACCTGGACACTCC
RT-PCR	Hs_RT_PLA2G4E_R	AACAGGTGGCATGGAGACA
RT-PCR	Hs_RT_HPRT1_F	TGACACTGGCAAAACAATGCA
RT-PCR	Hs_RT_HPRT1_R	GGTCCTTTTCACCAGCAAGCT
RT-PCR	Hs_RT_RPL38_F	TGGGTGAGAAAGGTCCTGGTC
RT-PCR	Hs_RT_RPL38_R	CGTCGGGCTGTGAGCAGGAA

ChIP	OHmeDIP_ACP5_F	acagccacagcctctcaagt
ChIP	OHmeDIP_ACP5_R	ggcttgacacgtccaggtat
ChIP+OH_Zymo	OHmeDIP_TM7SF4_F	ctccatttccatccctagca
ChIP+OH_Zymo	OHmeDIP_TM7SF4_R	gtgaactggagtgggaggag
ChIP	ChIP_TM4SF19_F	ggagggtgaggtgtttctca
ChIP	ChIP_TM4SF19_R	acccaccctctgctcatttt
ChIP	ChIP_CD22_F	atggaggggaaacctctgtc
ChIP	ChIP_CD22_R	ttttacCTGTTTCCGCGTGT
ChIP	ChIP_SNCG_F	ccagaaactgctagtgacgttg
ChIP	ChIP_SNCG_R	gacaagacccaccgggta
ChIP	ChIP_CD6_F	ttttcttctgggccacca
ChIP	ChIP_CD6_R	gcatccatgcagacacactt
ChIP	ChIP_NR4A2_F	cgctatcaattttcttctgttaaatg
ChIP	ChIP_NR4A2_R	gattcctccccacaaacaaa
ChIP	ChIP_PPP1R16B_F	Gcttcacctcccttctttcc
ChIP	ChIP_PPP1R16B_R	agatctgtggggccaggt
ChIP	Hs_ChIP_IL7R_F	CGTTTCTGGAGAAAGTGGCTA
ChIP	Hs_ChIP_IL7R_R	gggaactgaataacctgaaacc
ChIP	Hs_ChIP_FSCN3_F	cctttcctcccacttttcct
ChIP	Hs_ChIP_FSCN3_R	TCTGTGCCCAGAGAGCAAG
ChIP	Hs_ChIP_TDRD6_F	aacgtcccaacaagctcct
ChIP	Hs_ChIP_TDRD6_R	TTCCGGGCACGCATTTAC
ChIP	Hs_ChIP_CSF1R_FRA1_F	gcgcatcctagacctcactt
ChIP	Hs_ChIP_CSF1R_FRA1_R	CAGGGACACTGGGCTCTATC
ChIP	Hs_ChIP_CD59_FrCre_F	tgagaaggaagggacaggaa
ChIP	Hs_ChIP_CD59_FrCre_R	cggacagacagatgggttct
ChIP	Hs_ChIP_CCl5_NFkB_F	gccaatgcttggttgctatt
ChIP	Hs_ChIP_CCl5_NFkB_R	CGTGCTGTCTTGATCCTCTG
ChIP	Hs_ChIP_IL1R_NFkB_F	tcctaggtccctcaaaagca
ChIP	Hs_ChIP_IL1R_NFkB_R	gtccccaacgctctaacaaa
ChIP	Hs_ChIP_IL1R_CREB_F	gacctcccatcttacgcaga
ChIP	Hs_ChIP_IL1R_CREB_R	ggctgatggctgacttgatg
ChIP	Hs_ChIP_IL32_NFkB_F	ggacagggtccaaattcctt
ChIP	Hs_ChIP_IL32_NFkB_R	GGCAGAGGGAAAGTCCAGA
ChIP	Hs_ChIP_TM4SF19_CREB_F	ggagggtgaggtgtttctca
ChIP	Hs_ChIP_TM4SF19_CREB_R	acccaccctctgctcatttt
ChIP	Hs_ChIP_TNFRSF9_NFkB_F	gatttcggggtcagcagata
ChIP	Hs_ChIP_TNFRSF9_NFkB_R	CAGGTCAAACACAGGAGTGC
ChIP	Hs_ChIP_IL7R_AP1_F	gcaatctcggctcactgc
ChIP	Hs_ChIP_IL7R_AP1_R	gtggtgggcgcctgtaat

ChIP	Hs_qChIP_PLA2G4E_F	gaccaagccctctgtcactc
ChIP	Hs_qChIP_PLA2G4E_R	ccagcaatgaatggttaggc
ChIP	Hs_qChIP_SAT2_F	tgaatggaatcgtcatcgaa
ChIP	Hs_qChIP_SAT2_R	ccattcgataattccgcttg
ChIP	Hs_qChIP_TDRD1_F	tgcaaaggaacttttgagc
ChIP	Hs_qChIP_TDRD1_R	gtccacgtgcaactcaatgt
ChIP	Hs_qChIP_MYOD1_F	gttcctattggcctcggact
ChIP	Hs_qChIP_MYOD1_R	gcttcctcacccctagcttc
ChIP+OH_Zymo	Hs_qChIP_C1S_F	atttcccctggtaccaatcc
ChIP+OH_Zymo	Hs_qChIP_C1S_R	cagccaaagggtgtgtttct
ChIP+OH_Zymo	Hs_ChIP_CD59_F	ccagggaactctgcatttct
ChIP+OH_Zymo	Hs_ChIP_CD59_R	ctattctccagagccccaca
OH_Zymo	OHZymo_ACP5_CpG1_F	CGtgtcaagcctacagcaca
OH_Zymo	OHZymo_ACP5_CpG1_R	gagcaggacaCGggattg
OH_Zymo	OHZymo_CTSK_F	TGAGAATGATGGCTGTGGAG
OH_Zymo	OHZymo_CTSK_R	CCCACATATGGGTAGGCATC
OH_Zymo	OHZymo_TM4SF19_F	cccgtttcctgccttagaat
OH_Zymo	OHZymo_TM4SF19_R	ccacaggcctgcattgaa
OH_Zymo	OHZymo_NR4A2_F	ACTGCCCAGTGGACAAGC
OH_Zymo	OHZymo_NR4A2_R	ctcccctcagcctacCTTCT
OH_Zymo	Hs_Zymo_PLA2G4E_F	gccaagtgcattacagaacc
OH_Zymo	Hs_Zymo_PLA2G4E_R	ttggcttttccagaacactg
BS Sequencing	BS_Repetitive_18S_F	GGTTTGTGGYGYGGGGTT
BS Sequencing	BS_Repetitive_18S_R	AACTCTAAAATTACCACAATTA
BS Sequencing	BS_Repetitive_28S_F	GAGTGAATAGGGAAGAGTTTA
BS Sequencing	BS_Repetitive_28S_R	AAAATTCTTTTCAACTTTCCCTTAC
BS Sequencing	BS_Repetitive_D4Z4_F	GTTTGTTGTTGGATGAGTTTTTGG
BS Sequencing	BS_Repetitive_D4Z4_R	AAATCTCTCACCRAACCTAAACC
BS Sequencing	BS_Repetitive_SAT2_F	ATGGAAATGAAAGGGGTTATTATT
BS Sequencing	BS_Repetitive_SAT2_R	AAATTATTCCATTCCATTCCATTAA
BS Sequencing	BS_Repetitive_NBL2_F	GTAGTTGGTGTTAATGTGTGTTAT
BS Sequencing	BS_Repetitive_NBL2_R	CACTCTCTATATATTTCTTTCCC

ARTÍCULO 3:

Revista:

En preparación

Título:

Brief Report: MicroRNA profiling reveals key role of miR-212/132 and miR- 99b/let-7e/125a clusters in monocyte to osteoclast differentiation

Autores:

Lorenzo de la Rica[1], Natalia Ramírez-Comet[1], Laura Ciudad[1], Mireia García[2], José M. Urquiza[1], Carmen Gómez-Vaquero[2] and Esteban Ballestar[1]

Afiliaciones:

[1] Chromatin and Disease Group, Cancer Epigenetics and Biology Programme (PEBC), Bellvitge Biomedical Research Institute (IDIBELL), 08908 L'Hospitalet de Llobregat, Barcelona, Spain

[2] Rheumatology Service, Bellvitge University Hospital (HUB), 08908 L'Hospitalet de Llobregat, Barcelona, Spain

RESUMEN EN CASTELLANO

Los microRNAs (miRNAs) ejercen un efecto inhibitorio sobre la expresión de genes, e influencian la elección de linaje celular durante la hematopoyesis. La diferenciación de monocitos en osteoclastos es un proceso de diferenciación terminal único en el sistema hematopoyético. Esta diferenciación es relevante y está desregulada en procesos autoinmunes, así como en algunos tipos de cáncer. En la actualidad, está extensamente caracterizada la implicación de factores de transcripción y cambios de metilación en este proceso, sin embargo, el conocimiento sobre los mecanismos de regulación post-transcripcional es muy limitado. En este trabajo se han investigado los cambios de expresión en miRNAs durante la osteocaltogénesis. Un análisis de los perfiles de expresión de microRNAs en monocitos a día 0, y 2 y 20 días después de ser estimulados con MCSF y RANKL mostró cambios globales en el perfil de expresión de los mismos. Este perfil de expresión revela la participación de los miRNAs durante la diferenciación de osteoclastos, así como en la función de los osteoclastos maduros. En el presente estudio, por un lado se confirman cambios en la expresión de miRNAs previamente implicados en diferenciación mieloide, pero también se describen nuevos miRNAs implicados en este proceso de diferenciación mieloide. De manera especídica, los clusters de miRNAs miR-212/132 y miR- 99b/let-7e/125a aumentaban su expresión de manera rápida tras la estimulación con MCSF y RANKL. El aumento de la expresión sucede con diferentes diana (subida rápida y bajada, o subida rápida y mantenimiento de los niveles altos hasta el final del proceso de diferenciación). Los miRNAs anteriormente citados tienen dianas tan diversas como los genes *KDM6B*, *TNFS4*, *ARID3B* and *NR4A2* entre otros. La inhibición de la expresión de los miRNAs tenía efecto sobre la eficiencia de diferenciación, así como en la actividad de los osteoclastos, y confirmó el efecto en múltiples de las dianas predichas, tal como confirman los ensayos de luciferasa. Nuestros resultads revelan un papel clave de los cambios de expresión de miRNAs durante el proceso de diferenciación de los osteoclastos, e identifica nuevas dianas potenciales para manipular o inhibir el proceso.

ABSTRACT

MicroRNAs (miRNAs) exert negative effects on gene expression and influence cell lineage choice during hematopoiesis. Conversion of monocytes to osteoclasts is a unique terminal differentiation process within the hematopoietic system. This differentiation model is relevant to autoimmune disease and cancer. Currently, there is abundant knowledge on the expression changes, involvement of transcription factors and DNA methylation changes involved in this process but little is known about post-transcriptional regulatory mechanisms. In this study, we focused on miRNA expression changes during osteoclastogenesis. Analysis of the miRNA expression profiles in monocytes at 0, 2 and 20 days following M-CSF and RANKL stimulation revealed broad changes accompanying early stages of osteoclast differentiation and in mature osteclasts. These profiles reveal the participation of miRNAs during differentiation as well as in function of differentiated osteoclasts. We observed changes in expression of miRNAs previously described for their involvement in myeloid differentiation but also novel miRNAs processes. Specifically, miR-212/132 and miR- 99b/let-7e/125a clusters became rapidly upregulated, with different dynamics and target key genes like KDM6B, TNFS4, ARID3B and NR4A2 among others. Blocking of miRNA expression impacted the efficiency of differentiation as well as the activity of differenciated osteoclasts and confirmed the effects on several putative targets, along with luciferase assays. Our results reveal a key role of miRNA expression changes during osteoclast differentiation and identify novel targets for potential intervention.

INTRODUCTION

The successful generation of differentiated cell types from their progenitors depends on the highly coordinated regulation of gene expression by regulators including transcription factors (TFs), epigenetic modifications, and small non-coding RNAs. Despite their importance, the interplay between these regulators is not completely understood. For instance, TFs regulate microRNA (miRNA) expression and are themselves regulated by miRNAs, thereby establishing complex loops of regulation. However, the net contribution of TFs and miRNAs in different terminal differentiation processes is likely to be specific to each of them and is yet to be determined. MiRNAs regulate gene expression through sequence complementarity with their target mRNAs by mediating their decay or interfering with their translation. Many studies have focused on hematopoiesis in order to learn about the type, distribution and role of epigenetic and miRNA expression changes. However, the specific relationships in terminal differentiation processes remain elusive, even though these are amongst the most important since they produce functional cell types with very specific roles.

Osteoclasts are giant, multinucleated cells that degrade bone[1]. They differentiate from monocyte/macrophages progenitors[2] after M-CSF[3] and RANKL[4] induction. During osteoclastogenesis, progenitor cells fuse, re-organize[5] their cytoskeleton and activate the gene expression profile necessary for bone destruction. Several signaling pathways are activated after M-CSF and RANKL induction, and involve TRAF-6[6, 7], immunoreceptor tyrosine-based activation motif (ITAM)[8] adaptors DAP12[9] and FcRg[10] associated with their receptors, TREM-2[11] and OSCAR, together with calcium oscillations[12]. These signals activate NFkB, MAPK and c-Jun[13] that will coordinately turn NFATc1[14] on. NFATc1 is the master transcription factor of osteoclastogenesis, and works together with PU.1 and MITF[15], that were already present in the progenitors. The aforementioned transcription factors bind to osteoclast markers promoters and mediate their over-expression. Examples of such are tartrate-resistant acid phosphatase (TRAcP or ACP5)[16], cathepsin K (CTSK)[17], dendritic cell-specific transmembrane protein (DC-STAMP or TM7SF4)[18], matrix metalloproteinase 9 (MMP9)[19] or carbonic anhidrase 2 (CA2).

Osteoclast deregulation is involved in or causes several diseases, either by deficient function such as in osteopetrosis[20] or by aberrant hyperactivation, where decreased bone mass is detected (osteoporosis[21]). They are also involved in autoimmune disease, specifically, in rheumatoid arthritis aberrantly activated osteoclasts are one of the main effectors of joint destruction[22]. Moreover, osteoclast cause bone complications of several diseases, such as multiple

myeloma[23], prostate cancer and breast cancer[24], in which myeloma blasts or metastatic cells of the aforementioned conditions send pro-osteoclastic signals that end up in increased bone degradation and bone fractures. Indeed, there is a specific tumor with osteoclast origin, the giant cell tumor of bone (GCTB) also known as osteoclastoma where osteoclast upregulation and uncontrolled growth causes several bone injuries[25, 26].

In vitro generation of osteoclasts is an excellent tool to work with osteoclasts, given the great complexity of isolating primary bone osteoclasts. Among the PBMC population osteoclast precursor, CD14+ monocytic cells are able to give rise to osteoclasts following RANKL and M-CSF stimulation[27]. The osteoclasts generated are able to degrade bone and express osteoclast markers[28]. During the maturation of the osteoclast, progenitor cells fuse, giving rise to a multinuclear polykaryons where genetic information has to be tightly regulated. In this regard, few reports have analyzed the epigenetic status of this unusual cell type, and they have mainly focused on histone modifications with opposed results[29, 30], or in the DNA methylation landscape[31].However, the regulation of the functionality of the genes from a microRNA perspective, has been studied, using mouse bone marrow, RAW cell line, but less commonly, human samples.

The importance of miRNAs in osteoclast differentiation was confirmed when the miRNA processing machinery was silenced in mouse models. Knock-out mice for DGCR8, Dicer and Ago2 showed an impairment in osteoclast formation, and a reduced expression of TRAP and NFATc1[32]. The first specific miRNA that was found to be implicated in osteoclastogenesis was miR-223, which acts as a negative regulator when RAW264.7 differentiates to osteoclasts[33]. On the other hand, this miRNA had a pro-osteoclastic function when the differentiation was studied from bone marrow. miR-223 silences NFI-A, therefore allowing the expression of M-CSF receptor[34]. The importance of this miRNA in arthritis was demonstrated in mice by inhibiting it in vivo with lentiviral vectors in collagen induced arthritis mouse. The ablation of miR-223 expression ameliorated arthritis severity in the joints[35, 36]. miR-155 was also found to be a negative regulator of OC differentiation from RAW264.7, as it targets MITF mRNA, and prevents them to differentiate into osteoclasts[37]. Its function is especially relevant when translating this into in vivo CIA mouse models. Silencing of miR-155 in CIA mice, protected their joints of bone degradation, thereby confirming its potential as a therapeutic target in RA patients[38, 39]. miRNa expression profile changes drastically during the differentiation of bone marrow macrophages (BMMs) to osteoclasts. miR-21 is upregulated after RANKL and M-CSF induction, targeting the expression of PDCD4

(programmed cell death). The silencing of PDCD4 prevents its repression over c-Fos, which is a key mediator in the differentiation of BMMs to osteoclasts in mice[40]. miR-146 is at high concentrations in RA synovium as well as in PBMCs. When it is silenced, TRAP-positive multinucleated cell levels drops, and joint destruction is prevented in CIA mice joints, indicateing the potential therapeutic use of it in RA[41]. Other examples of important miRNAs for OC differentiation are miR-29b[42] and miR-124[43], that should be downregulated during the procces, as they act as a negative regulators.

In this study, we provide the first systematic high throughput analysis of the miRNA landscape variation upon RANKL and M-CSF stimulation in human osteoclastogenesis. Moreover, when studying the dynamics of the process, two miRNA clusters, miR-212/132 and miR- 99b/let-7e/125a were found responsible for the silencing of monocytic and inappropriate alternative lineage genes. Functionally inhibition of these, caused osteoclast differentiation delay/impairment through aberrant expression of their targets, mainly monocytic genes. Moreover, we have demonstrated that NR4A2 and CX3CR1 are directly regulated by miR212 and 132, and miR99b. Together, our data suggests a novel role for the aforementioned miRNA clusters in osteoclast differentiation from monocytic myeloid precursors, that could be potentially therapeutically inhibited to treat bone related diases such as rheumatoid arthritis and GBCT.

METHODS

Differentiation of OCs from peripheral blood mononuclear cells

Human samples (blood) used in this study came from anonymous blood donors and were obtained from the Catalan Blood and Tissue Bank (Banc de Sang i Teixits) in Barcelona as thrombocyte concentrates (buffy coats). The anonymous blood donors received oral and written information about the possibility that their blood would be used for research purposes, and any questions that arose were then answered. Prior to obtaining the first blood sample the donors signed a consent form at the Banc de Teixits. The Banc de Teixits follows the principles set out in the WMA Declaration of Helsinki. The blood was carefully layered on a Ficoll–Paque gradient (Amersham, Buckinghamshire, UK) and centrifuged at 2000 rpm for 30 min without braking. After centrifugation, peripheral blood mononuclear cells (PBMCs), in the interface between the plasma and the Ficoll– Paque gradient, were collected and washed twice with ice-cold PBS, followed by centrifugation at 2000 rpm for 5 min. Pure CD14+ cells were isolated from PBMCs using positive selection with MACS magnetic CD14 antibody (Miltenyi Biotec). Cells were then resuspended in g-minimal essential medium (g-MEM, Glutamax no nucleosides) (Invitrogen, Carlsbad, CA, USA) containing 10% fetal bovine serum, 100 units/ml penicillin, 100 ug/ml streptomycin and antimycotic and supplemented with 25 ng/mL human M-CSF and 50"ng/ml hRANKL soluble (PeproTech EC, London, UK). Depending on the amount needed, cells were seeded at a density of 3•105 cells/well in 96-well plates, 5•106 cells/well in 6-well plates or 40•106 cells in 10 mm plates and cultured for 21 days (unless otherwise noted); medium and cytokines were changed twice a week. The presence of OCs was checked by tartrate-resistant acid phosphatase (TRAP) staining using the Leukocyte Acid Phosphatase Assay Kit (Sigma–Aldrich) according to the manufacturer's instructions. A phalloidin/DAPI stain allowed us to confirm that the populations were highly enriched in multinuclear cells, some of them containing more than 40 nuclei. We used several methods to determine that on day 21 almost 85% of the nuclei detected were “osteoclastic nuclei” (in polykaryons, nuclei and not cells were quantified). OCs (TRAP-positive cells with more than three nuclei) were also analyzed at the mRNA level: upregulation of key OC markers (TRAP/ACP5, CA2, MMP9 and CTSK) and the downregulation of the MO marker CX3CR1 were confirmed. RESULTS

Visualization of OCs with phalloidin and DAPI staining

PBMCs or pure isolated CD14+ cells were seeded and cultured in glass Lab-Tek Chamber Slides (Thermo Fisher Scientific) for 21 days in the presence of hM-CSF and hRANKL. OCs were then washed twice with PBS and fixed (3.7% paraformaldehyde, 15 min). Cells were permeabilized with 0.1% (V/V) Triton X-100 for 5 min and stained for F-actin with 5 U/mL Alexa Fluor® 647-Phalloidin (Invitrogen). Cells were then mounted in Mowiol-DAPI mounting medium. Cultures were visualized by CLSM (Leica TCP SP2 AOBS confocal microscope).

microRNA expression screening, target prediction and integration with DNA methylation data

Total RNA was extracted with TriPure (Roche, Switzerland) following the manufacturer instructions. Ready-to-use microRNA PCR Human Panel I V2.R from Exiqon (Reference 203608) were used according to the instruction manual (Exiqon). For each RT-PCR reaction 30 ng total RNA was used. Paired samples of MOs at day 0 (MOs), 2 (OC 48h) and 21 (OCs) days after M-CSF and RANKL stimulation were obtained from three female healthy donors (Age 25-28), and were analyzed on a Roche LightCycler® 480 real-time PCR system. Results were converted to relative values using the inter-plate calibrators included on the panels (log 2 ratios). MOs, OCs 48h and OCs average expression values were normalized to reference gene miR-103. A t-test was then performed and microRNAs differentially expressed (FC >2 or <0.5), with a significant p-value (p-val <0.05) in at least one of the comparisons were selected and represented on the heatmap. The complete list of the raw expression data can be found in the supplementary material. Validation of the array expression data was performed in the samples used (validation set), as well as in a larger cohort of samples obtained from independent donors (replication set) using Exiqon microRNA LNA™ PCR primer sets (hsa-miR-99b-5p Ref. 204367; hsa-miR-125a-5p Ref. 204339; hsa-miR-132-3p Ref. 204129; hsa-miR-212-3p Re. 204170; hsa-miR-103a-3p Ref. 204063).

In order to predict the potential targets of the de-regulated microRNAs, we used the algorithms of several databases, specifically TargetScan, PicTar, PITA, miRBase, microRNA.org, miRDB/MirTarget2, TarBase, miRecords, StarBase/CLIPseq . Only Targets predicted in at least four of those databases were selected to further analysis.

To compare the DNA methylation bead array data with the expression levels of the miRNAs, a mapping between miRNAs and Illumina 450k probes was performed. For each differentially expressed miRNA we studied the CpGs within 2500 pb upstream and downstream of the Transcription Starting Site. We considered the CpGs included in the Illumina 450K array. MiRNAs genomic features were obtained from miRBase. Illumina annotation was obtained from IlluminaHumanMethylation450K.db Biocondcutor Package.

Transfection of primary human MOs

To perform the miRNa inhibitors experiments, we used unlabelled miRCURY LNA™ microRNA Power inhibitors to inhibit miR-99b (Ref. 427491-00), miR-125a (Ref. 426713-00), miR-132 (Ref. 426779-00), miR-212 (Ref. 426953-00) or as a control (Negative Control A Re.199020-00). 5 or 10 nM of Power inhibitors were transfected in CD14+ MOs using HapyFect Transfection Reagent (Tecran, UK). The efficiency of transfection was quantified by flow citometry using the Negative Control A, 5`-fluorescein labeled (Ref. 199020-04). In order to silence PU.1, we used two different Silencer® select pre-designed siRNAs against human PU.1 (one targeting exon 2 and another targeting the 3'UTR) and a Silencer® select negative control to perform PU.1 knockdown experiments in peripheral blood MOs. We used Lipofectamine RNAiMAX Transfection Reagent (Invitrogen) for efficient siRNA transfection. mRNA and protein levels were examined by quantitative RT-PCR and western blot at 1, 2, 4 and 6 days after siRNA transfection. These experiments were performed with more than three biological replicates.

Gene expression data analysis of RNA expression

In order to analyze expression data versus methylation data, we used CD14+ and OC expression data from ArrayExpress database (www.ebi.ac.uk/arrayexpress) under the accession name (EMEXP-2019) from a previous publication[44]. Affymetrix GeneChip Human Genome U133 Plus 2.0 expression data was processed using limma and affy packages from bioconductor. The pre-processing stage is divided in three major steps: 1) background correction, 2) normalization, and 3). reporter summarization. Here, the expresso function in affy package was chosen for preprocessing. Thus, the RMA method was applied for background correction. Then, a quantile normalization was performed. In addition, we introduced a specific step for PM (perfect

matchprobes) adjustment, utilizing the PM-only model based expression index (option 'pmonly'). And finally, for summarization step, the median polish method was taken. Next, a variance filtering by IQR (Interquartile range) using 0.50 for threshold value was executed. After preprocessing, a statistical analysis was applied, using eBayes moderated tstatistics test from limma package. Validation of expression data was performed by quantitative RT-PCR.

Graphics and heatmaps

All graphs were created using Prism5 Graphpad. Heatmaps were generated from the expression or methylation data using the Genesis program (Graz University of Technology)

RESULTS

MicroRNA expression profile changes drastically during osteoclastogenesis

To analyze the dynamics of microRNA expression during human osteoclastogenesis, we first generated three sets of matching samples of MOs (CD14+ cells) from peripheral blood, OCs 48 hours post RANKL and M-CSF induction, and mature OCs derived from the same CD14+ cells. The quality of the osteoclasts was confirmed microscopically by the presence of more than three nuclei in TRAP-positive cells (TRAP staining) and the formation of the actine ring (phalloidin/DAPI staining), as shown in Figure 1A. At the molecular level the upregulation of osteoclastic markers (CA2, CTSK, ACP5/TRACP and MMP9,) and the silencing of monocytic genes (CX3CR1) was confirmed (Figure 1B). We then performed a microRNA profiling to analyze the dynamics on the microRNA expression during the differentiation of monocytes to osteoclasts. Statistical analysis of the combined expression data from three biological replicates showed 115 microRNAs differentially expressed in at least one of the time points analyzed (Figure 1C). We organized the microRNAs according to their expression profile into 8 groups (the result of combining 3 comparison groups two-by-two) and focused in those microRNA whose expression increased at 48h. microRNAs that rapidly become upregulated after M-CSF and RANKL stimulation are potentially more important for the differentiation process than for the function of the osteoclasts. Two microRNA clusters ranked top in terms of fold change and relative expression levels: miR-212/132 and miR- 99b/let-7e/125a (Figure 1D). We confirmed the overexpression of the microRNAs by qPCR in the samples used for the Array (Figure 1E), as well as in a replication cohort (data not shown). To further analyze the expression dynamics of these microRNAs during the differentiation process we generated a time course of osteoclastogenesis from three different healthy donors, and checked the miRNA levels at several time points (Days 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 11, 12, 13, 14, 16, 19 and 21 days). These two clusters showed different dynamics when we analyzed their expression levels in a timely manner. For example, after RANKL/M-CSF stimulation, miR- 99b/let-7e/125a cluster members expression increased rapidly during the first four days; after day 4, its levels remained stably high until day 21 (Figure 1F, top). On the other hand, miR-212/132 cluster members expression peaked at day 2 showing an increase over day 0 of 47 fold (miR132) to 170 fold (miR-212). Strikingly, after day 2 peak their expression levels drop by 5 fold (Figure 1F, bottom). It appears that the function of miR-132 and miR-212 is involved in the early events of osteoclastogenesis as their expression levels are tightly regulated, and constrained to the first four days of differentiation.

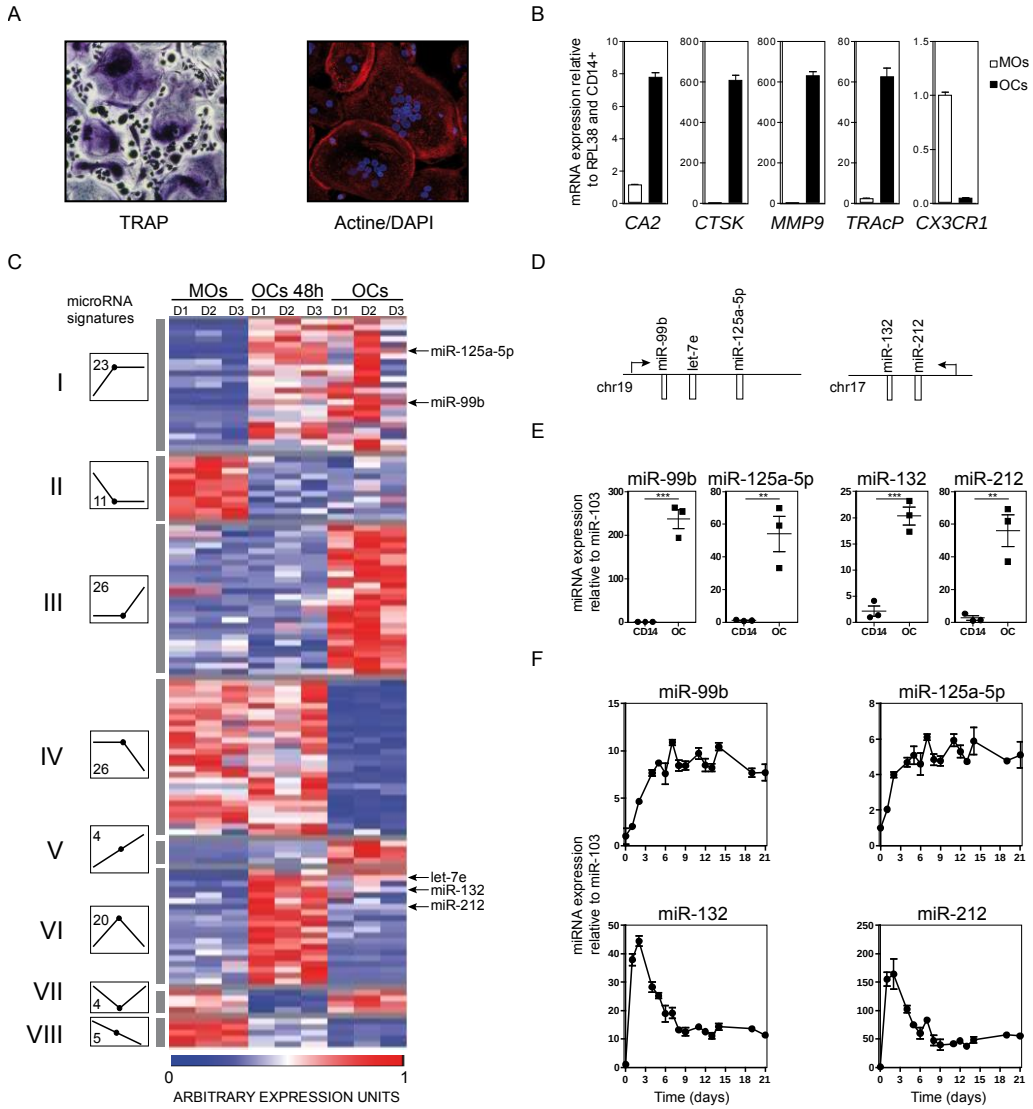


Figure 1. A. Validation of the presence of osteoclast by TRAP staining and by phalloidine staining to se the actin ring. B. Molecular characterization of the differentiation process. Several OC markers are upregulated(CA2, CTSK, MMP9 and TRAcP), and monocyte markers are silenced (CX3CR1). C. Heatmap showing the expression array data from the miRNA expression screening. miRNAs were subdivided in 8 groups (from I t VIII) according to their expression profiling (scheme); the number of miRNAs in each group is indicated inside the expression dynamics scheme. D. Representation of the genomic distribution of miR132/212 and miR 99b/125a/let7e clusters. E. Validation of array data by qPCR in a independent biological replicas. F. Expression dynamics of the indicated miRNAs during several time-points.

Several other miRNAs undergo expression changes after RNAKL and M-CSF stimulation, and many of them are concordant with previously described data in mice. For example, miR-223 is silenced, in accordance with its previously described function as a negative regulator of OC in RAW33. miR155 is upregulated, as it happens in BMM deriver osteoclasts 38, 39 (but contrary to what happens in RAW37). miR-21 is upregulated, concordant with previously described information40. miR-146 is upregulated at later steps41. On the other hand, miR-12443 is upregulated at 48 hours, but silenced right after.

DNA demethylation and PU.1 binding are implicated in the upregulation of miR-212/132 and miR- 99b/let-7e/125a clusters

The expression of miRNA genes can be regulated by several cellular mechanisms. We wondered whether the transcriptional activation of these two miRNA clusters may involve epigenetic modifications of their promoters, as well as the binding of the key myeloid transcription factor PU.1. We have previously described that both transcriptional regulators have a key role in modulating the expression of key genes related with the osteoclastic differentiation process31. So we decided to translate our previous findings to the regulation of the miRNA transcriptional program.

We analyzed the features of the miRNa promoters in order to see the CG composition, as well as the PU.1 ocupancy. As it can be seen in Figure 2A, both miRNa promoters are surrounded by both CpG islands as well as PU.1 bound molecules. To analyze the presence of PU.1 in miRNa promoters,we used PU.1 ChIP-Seq data available on human monocytes, from Gene Expression Omnibus (GSE31621) from other publication[45]. As it can be seen on Figure 2A, PU.1 was targeting the promoter of miR-212/132 cluster, but not miR- 99b/let-7e/125a. To properly demonstrate the importance of PU.1 in regulating the expression of miRNAs, we knocked down PU.1 by using a siRNA approach. We validated the extent of the silencing by qPCR (achieving a 60% of downregulation) (Figure 2B) as well as by western blot (around 35 to 70 % of protein levels downregulation) (Figure 2C). As it can be seen in Figure 2D, when PU.1 is absent, an increase in the expression of several miRNAs was detected, indicating the potential inhibitory effect of PU.1 in this miRNA promoters.

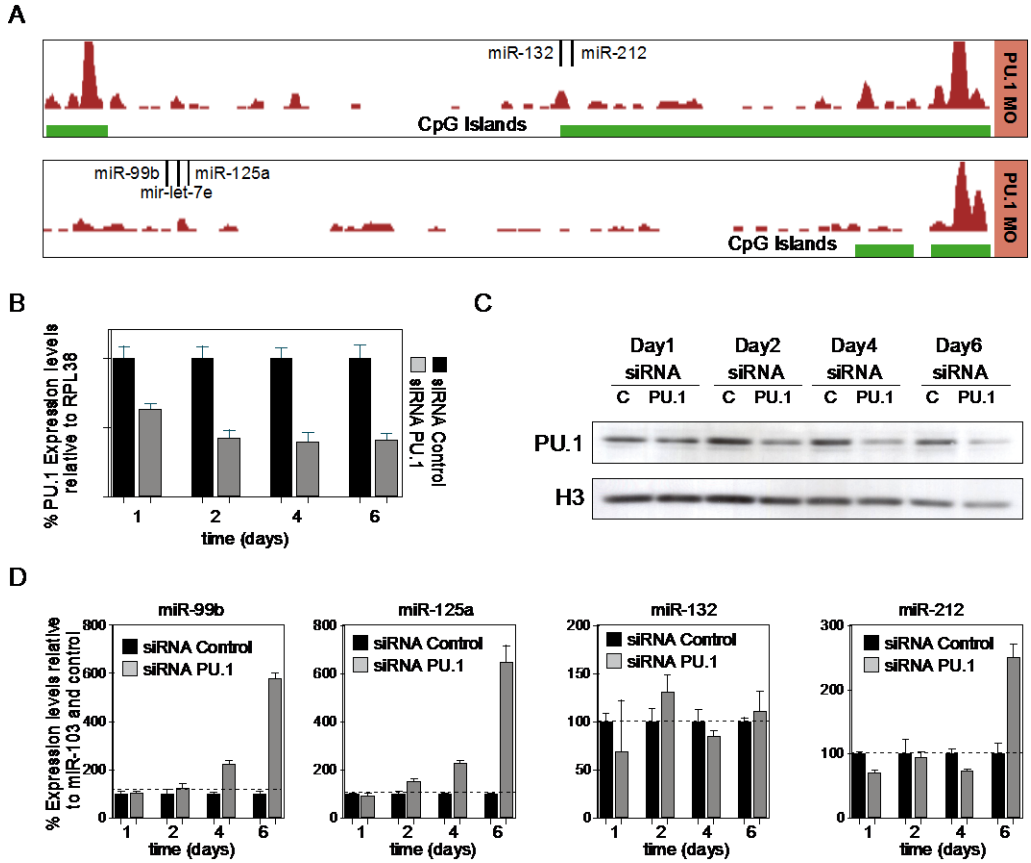


Figure 2. A. Genomic features of the miR132/212 and miR 99b/125a/let7e clusters. Peaks show the occupancy of PU.1 in monocytes. Green bars indicate CpG island presence. B. Validation of PU.1 knock down by qPCR. C. Validation of the of PU.1 protein knock down by western blot. D. Expression levels of the indicated miRNAs on PU.1 defficient samples at several time points.

Inhibition of microRNA function delays osteoclastogenesis

We also researched the putative mRNA targets for each of the selected miRNAs by the combination of several miRNA target prediction algorithms. We extracted the expression data of the putative targets from a mRNA expression array already published44 to look for genes that become downregulated after MCSF and RANKL stimulation, that is, that could be potentially regulated by miRNA posttranslational targeting (Figure 3A).

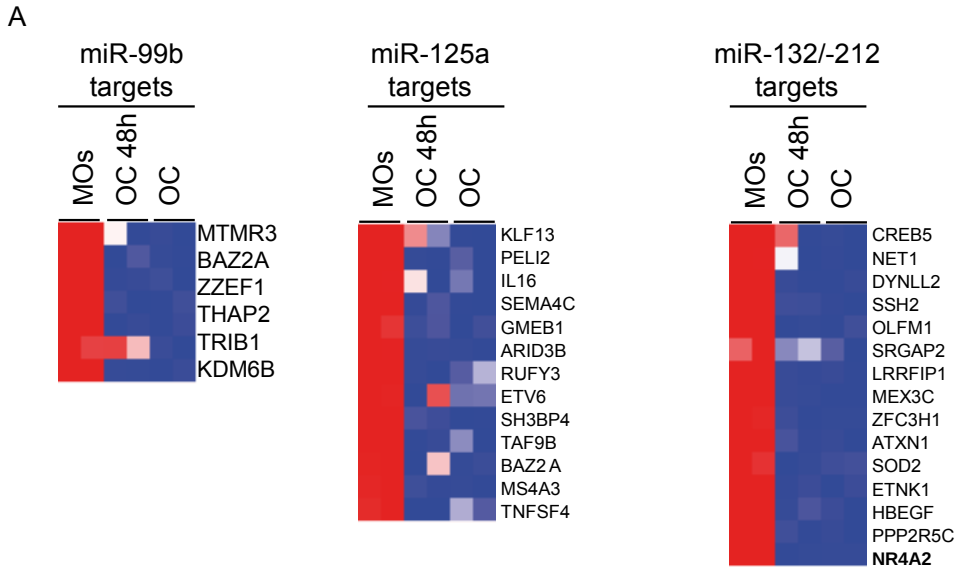


Figure 3. A. Heatmaps showing array expression levels of several miRNA target genes in MOs, OCs 48h and mature OCs. The targets shown were predicted by at least 4 target prediction algorithms.

One of the candidate targets that arose from this analysis as a target for miR-212 was the nuclear receptor subfamily 4, group A, member 2 (NR4A2). This gene was already described to have an implication in osteoclast, as it is rapidly silenced after MCSF/RANKL stimulation, and moreover, its promoter is hypermethylated31.

To further characterize the role of the selected microRNAs in osteoclastogenesis, we transfected the precursor monocytes with antagomiRs. We checked the efficiency of transfection by flow citometry of cells transfected with a control antagomiR fluorescein conjugated. We assessed 72.8% efficiency (Figure 4A). The antagomiRs bind specifically to the complementary miRNA and inhibit their function, therefore, preventing them to silence their target mRNAs. Then, we measured the expression levels of key osteoclast marker genes as well as several target genes for the miRNA, in order to asses if there is a delay in the differentiation process, and through which mechanisms they are working. Finally, we analyzed by TRAP staining if there were any impairment or delay in the formation of mature, bone resorbing osteoclasts.

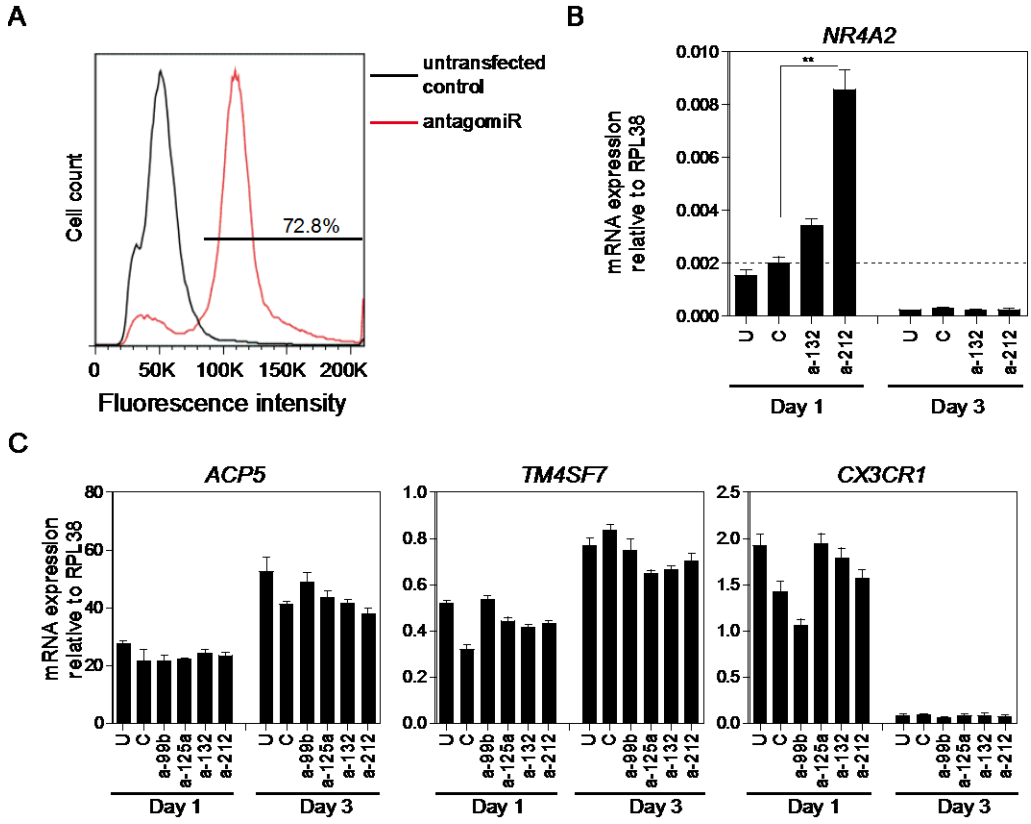


Figure 4 A. Quantification of the transfection efficiency by flow citometry. B. Functional effect of miRNA inhibition on NR4A2 expression levels at several time points. C. Functional effect of miRNA inhibition on the expression of several osteoclast and monocytemarkers.

When we transfected with miR-212 antagomiR, we saw that NR4A2 silencing was impaired by 5-fold at day 1 (Figure 2B). On the other hand, when transfecting individual miRNA inhibitors, we found little effect in regards to OC marker genes. ACP5, CTSK, TM4SF7 or CX3CR1 expression levels did not change after inhibiting individual miRNAs (Figure 4C).

microRNA function in osteoclastogenesis occur through repressing alternative lineage genes

After assessing the effect of miRNA inhibition in osteoclastogenesis, we decided to further investigate the molecular mechanisms through which miRNAs were exerting their inhibitory effect. We detected potential binding sites for miR-212 in the 3'UTR

of NR4A2, therefore, we decided to demonstrate the direct inhibition of the latter by miR-212, through a luciferase reporter assay of the 3'UTR region.

DISCUSSION

Our results provide evidence of the implication of miRNA regulation of key genes during monocyte-to-osteoclast. First a high throughput screening of miRNA expression identified 115 miRNAs that undergo expression changes. We were able to further validate these expression changes by qPCR in independent biological replicas, indicating the robustness of our screening. Moreover, we have further characterized the expression dynamics of two of the miRNA clusters that become more strongly regulated during the process: miR-212/132 and miR- 99b/let-7e/125a clusters. After a comprehensive analysis of several time points, the expression profile of each of the miRNAs showed different dynamics: miR-212/132 cluster peaked at day 2 and was rapidly silenced, while miR- 99b/let-7e/125a was upregulated and its expression was stably maintained through differentiation.

We hypothesize that the expression dynamics that miR-212/132 cluster shows are more related with miRNAs important for the differentiation process, while those shown by miR- 99b/let-7e/125a cluster seem to be more implicated in osteoclast function.

REFERENCES

1. Blair, H. C., Teitelbaum, S. L., Ghiselli, R. & Gluck, S. Osteoclastic bone resorption by a polarized vacuolar proton pump. Science 245, 855-7 (1989).

2. Yasuda, H. et al. Osteoclast differentiation factor is a ligand for osteoprotegerin/osteoclastogenesis-inhibitory factor and is identical to TRANCE/RANKL. Proc Natl Acad Sci U S A 95, 3597-602 (1998).

3. Wiktor-Jedrzejczak, W. et al. Total absence of colony-stimulating factor 1 in the macrophage-deficient osteopetrotic (op/op) mouse. Proc Natl Acad Sci U S A 87, 4828-32 (1990).

4. Lacey, D. L. et al. Osteoprotegerin ligand is a cytokine that regulates osteoclast differentiation and activation. Cell 93, 165-76 (1998).

5. Saltel, F., Chabadel, A., Bonnelye, E. & Jurdic, P. Actin cytoskeletal organisation in osteoclasts: a model to decipher transmigration and matrix degradation. Eur J Cell Biol 87, 459-68 (2008).

6. Wong, B. R. et al. TRANCE, a TNF family member, activates Akt/PKB through a signaling complex involving TRAF6 and c-Src. Mol Cell 4, 1041-9 (1999).

7. Kobayashi, N. et al. Segregation of TRAF6-mediated signaling pathways clarifies its role in osteoclastogenesis. Embo J 20, 1271-80 (2001).

8. Blank, U., Launay, P., Benhamou, M. & Monteiro, R. C. Inhibitory ITAMs as novel regulators of immunity. Immunol Rev 232, 59-71 (2009).

9. Humphrey, M. B. et al. The signaling adapter protein DAP12 regulates multinucleation during osteoclast development. J Bone Miner Res 19, 224-34 (2004).

10. Koga, T. et al. Costimulatory signals mediated by the ITAM motif cooperate with RANKL for bone homeostasis. Nature 428, 758-63 (2004).

11. Humphrey, M. B. et al. TREM2, a DAP12-associated receptor, regulates osteoclast differentiation and function. J Bone Miner Res 21, 237-45 (2006).

12. Negishi-Koga, T. & Takayanagi, H. Ca2+-NFATc1 signaling is an essential axis of osteoclast differentiation. Immunol Rev 231, 241-56 (2009).

13. Ikeda, F. et al. Critical roles of c-Jun signaling in regulation of NFAT family and RANKL-regulated osteoclast differentiation. J Clin Invest 114, 475-84 (2004).

14. Takayanagi, H. et al. Induction and activation of the transcription factor NFATc1 (NFAT2) integrate RANKL signaling in terminal differentiation of osteoclasts. Dev Cell 3, 889-901 (2002).

15. Sharma, S. M. et al. MITF and PU.1 recruit p38 MAPK and NFATc1 to target genes during osteoclast differentiation. J Biol Chem 282, 15921-9 (2007).

16. Yu, M., Moreno, J. L., Stains, J. P. & Keegan, A. D. Complex regulation of tartrate-resistant acid phosphatase (TRAP) expression by interleukin 4 (IL-4): IL-4 indirectly suppresses receptor activator of NF-kappaB ligand (RANKL)-mediated TRAP expression but modestly induces its expression directly. J Biol Chem 284, 32968-79 (2009).

17. Matsumoto, M. et al. Essential role of p38 mitogen-activated protein kinase in cathepsin K gene expression during osteoclastogenesis through association of NFATc1 and PU.1. J Biol Chem 279, 45969-79 (2004).

18. Kim, K., Lee, S. H., Ha Kim, J., Choi, Y. & Kim, N. NFATc1 induces osteoclast fusion via up-regulation of Atp6v0d2 and the dendritic cell-specific transmembrane protein (DC-STAMP). Mol Endocrinol 22, 176-85 (2008).

19. Sundaram, K. et al. RANK ligand signaling modulates the matrix metalloproteinase-9 gene expression during osteoclast differentiation. Exp Cell Res 313, 168-78 (2007).

20. Tolar, J., Teitelbaum, S. L. & Orchard, P. J. Osteopetrosis. N Engl J Med 351, 2839-49 (2004).

21. Rachner, T. D., Khosla, S. & Hofbauer, L. C. Osteoporosis: now and the future. Lancet 377, 1276-87 (2011).

22. Scott, D. L., Wolfe, F. & Huizinga, T. W. Rheumatoid arthritis. Lancet 376, 1094-108 (2010).

23. Mundy, G. R., Raisz, L. G., Cooper, R. A., Schechter, G. P. & Salmon, S. E. Evidence for the secretion of an osteoclast stimulating factor in myeloma. N Engl J Med 291, 1041-6 (1974).

24. Yoneda, T. Cellular and molecular mechanisms of breast and prostate cancer metastasis to bone. Eur J Cancer 34, 240-5 (1998).

25. Mii, Y. et al. Osteoclast origin of giant cells in giant cell tumors of bone: ultrastructural and cytochemical study of six cases. Ultrastruct Pathol 15, 623-9 (1991).

26. Joyner, C. J., Quinn, J. M., Triffitt, J. T., Owen, M. E. & Athanasou, N. A. Phenotypic characterisation of mononuclear and multinucleated cells of giant cell tumour of bone. Bone Miner 16, 37-48 (1992).

27. Nicholson, G. C. et al. Induction of osteoclasts from CD14-positive human peripheral blood mononuclear cells by receptor activator of nuclear factor kappaB ligand (RANKL). Clin Sci (Lond) 99, 133-40 (2000).

28. Sorensen, M. G. et al. Characterization of osteoclasts derived from CD14+ monocytes isolated from peripheral blood. J Bone Miner Metab 25, 36-45 (2007).

29. Youn, M. Y., Takada, I., Imai, Y., Yasuda, H. & Kato, S. Transcriptionally active nuclei are selective in mature multinucleated osteoclasts. Genes Cells 15, 1025-35 (2010).

30. Saltman, L. H. et al. Organization of transcriptional regulatory machinery in osteoclast nuclei: compartmentalization of Runx1. J Cell Physiol 204, 871-80 (2005).

31. de la Rica, L. et al. PU.1 target genes undergo Tet2-coupled demethylation and DNMT3b-mediated methylation in monocyte-to-osteoclast differentiation. Genome Biol 14, R99 (2013).

32. Mizoguchi, F. et al. Osteoclast-specific Dicer gene deficiency suppresses osteoclastic bone resorption. J Cell Biochem 109, 866-75 (2010).

33. Sugatani, T. & Hruska, K. A. MicroRNA-223 is a key factor in osteoclast differentiation. J Cell Biochem 101, 996-9 (2007).

34. Sugatani, T. & Hruska, K. A. Impaired micro-RNA pathways diminish osteoclast differentiation and function. J Biol Chem 284, 4667-78 (2009).

35. Li, Y. T. et al. Brief report: amelioration of collagen-induced arthritis in mice by lentivirus-mediated silencing of microRNA-223. Arthritis Rheum 64, 3240-5 (2012).

36. Shibuya, H. et al. Overexpression of microRNA-223 in rheumatoid arthritis synovium controls osteoclast differentiation. Mod Rheumatol 23, 674-85 (2013).

37. Mann, M., Barad, O., Agami, R., Geiger, B. & Hornstein, E. miRNA-based mechanism for the commitment of multipotent progenitors to a single cellular fate. Proc Natl Acad Sci U S A 107, 15804-9 (2010).

38. Bluml, S. et al. Essential role of microRNA-155 in the pathogenesis of autoimmune arthritis in mice. Arthritis Rheum 63, 1281-8 (2011).

39. Zhang, J. et al. Interferon-beta-induced miR-155 inhibits osteoclast differentiation by targeting SOCS1 and MITF. FEBS Lett 586, 3255-62 (2012).

40. Sugatani, T., Vacher, J. & Hruska, K. A. A microRNA expression signature of osteoclastogenesis. Blood 117, 3648-57 (2011).

41. Nakasa, T., Shibuya, H., Nagata, Y., Niimoto, T. & Ochi, M. The inhibitory effect of microRNA-146a expression on bone destruction in collagen-induced arthritis. Arthritis Rheum 63, 1582-90 (2011).

42. Rossi, M. et al. miR-29b negatively regulates human osteoclastic cell differentiation and function: Implications for the treatment of multiple myeloma-related bone disease. J Cell Physiol 228, 1506-15 (2013).

43. Lee, Y. et al. MicroRNA-124 regulates osteoclast differentiation. Bone 56, 383-389 (2013).

44. Gallois, A. et al. Genome-wide expression analyses establish dendritic cells as a new osteoclast precursor able to generate bone-resorbing cells more efficiently than monocytes. J Bone Miner Res 25, 661-72 (2009).

45. Pham, T. H. et al. Mechanisms of in vivo binding site selection of the hematopoietic master transcription factor PU.1. Nucleic Acids Res 41, 6391-6402 (2013).

RESUMEN GLOBAL DE RESULTADOS Y DISCUSIÓN

En la presente tesis doctoral se ha analizado la desregulación epigenética en dos tipos celulares relevantes para la artritis reumatoide, los fibroblastos sinoviales y los osteoclastos. La actividad de ambos tipos celulares se encuentra exacerbada en la articulación de las personas artríticas, y se han empleado dos aproximaciones diferentes para su estudio.

Por un lado hemos analizado los fibroblastos sinoviales, responsables de la degradación del cartílago en estas personas. Hemos trabajado con los fibroblastos sinoviales de rodillas de personas con artritis reumatoide, y los hemos comparado con fibroblastos control analizando la desregulación de sus perfiles de metilación de DNA y microRNAs, para investigar las diferencias entre los fibroblastos de pacientes y controles. Las diferencias en los perfiles de metilación de DNA y de expresión de microRNAs entre los dos grupos indican potenciales vías de señalización y genes desregulados en esta enfermedad.

Por otro lado, hemos estudiado un proceso de diferenciación relevante para la articulación artrítica, como es la osteoclastogénesis, puesto que su sobreactivación es responsable de la degradación del hueso en la rodilla de estos pacientes. Hemos analizado *in vitro* qué cambios epigenéticos suceden durante la diferenciación de monocitos a osteoclastos. Como resultado hemos descubierto el papel dual clave del factor de transcripción PU.1 en el reclutamiento de la maquinaria epigenética (TET2 y DNMT3b) que va a modificar, en un sentido o en otro (hipometilación e hipermetilación) el epigenoma de los monocitos, para que se diferencien a osteoclastos. Conocer y caracterizar a nivel molecular qué sucede en este proceso de diferenciación es clave para posteriormente encontrar vías de señalización que potencialmente puedan ser inhibidas farmacológicamente.

En este apartado se presentan los resultados globales de esta tesis doctoral de manera comentada.

1. Análisis de metilación y del perfil de microRNAs en fibroblastos sinoviales aislados de pacientes con artritis reumatoide y osteoartritis

1.1. La comparación de los patrones de metilación del DNA entre fibroblastos sinoviales de pacientes de AR e individuos control revela cambios: hipometilación e hipermetilación en genes clave para el desarrollo de la artritis

Hemos comparado los perfiles de metilación del DNA de fibroblastos sinoviales extraídos de 6 pacientes con artritis reumatoide con los de 6 obtenidos de individuos con osteoartritis. Este tipo celular es habitualmente utilizado como control respecto

a artritis reumatoide aparte de por su disponibilidad, debido a que carece de componente inflamatorio. Se han comparado los perfiles de metilación de ambos grupos por medio del array de metilación 450K de Illumina. Entre los más de 1200 genes con cambios significativos de metilación detectados, hay algunos cuya función en AR ha sido extensamente estudiada, como el caso de *IL6R*, que se encuentra hipometilado en RASFesta enfermedad. La hipometilación y sobreexpresión de *IL6R* juega un papel principal en la inflamación aguda y crónica de la articulación, e incrementa el riesgo de rotura[179]. Además, un fármaco empleado para el tratamiento de la artritis reumatoide, como es el tocilizumab (agente biológico), tiene como diana esta proteína[180]. Otro ejemplo sería el del regulador negativo de apoptosis *TNFAIP8*, que también se hipometila, y podría explicar la mayor resistencia a la apoptosis que este tipo celular muestra. En dirección opuesta, hemos encontrado entre otros, dos genes previamente implicados en la patogénesis de la RA, como son *CCR6*[181] y *DPP4*[182]. Pero quizá lo más interesante de este tipo de análisis globales de metilación, es la posibilidad de identificar nuevos marcadores potenciales, cuya metilación se ve desregulada. Este es el caso del gen *CAPN8*, cuya implicación nunca se había descrito en RA, y su estado de metilación se mantiene firmemente entre todas las muestras analizadas. A nivel funcional, agrupando los genes por categorías GO (Gene Ontology), en los genes hipometilados y potencialmente sobreexpresados observamos enriquecimiento en "desarrollo de cartílago", "regulación del crecimiento celular" así como "ensamblaje de adhesión focal". Categorías comunes para genes hipo e hipermetilados son "diferenciación celular", "adhesión celular" y "desarrollo del sistema esquelético". Las categorías mencionadas son de especial relevancia dado que son precisamente los fibroblastos sinoviales los que en AR destruyen el cartílago, y por tanto, empeoran el sistema esquelético de las personas afectadas. Asimismo, hemos analizado los cambios de metilación por regiones o "clusters" en el genoma. Hemos visto varios genes con hasta 9 CpGs representadas en el array cuya metilación cambia en el mismo sentido (hipo o hipermetilados). Los genes *HOXA11* y *CD74*, por ejemplo, tienen en su promotor hasta 9 CpGs hipometiladas cada uno. Curiosamente, los niveles de CD74 son más elevados en el tejido sinovial procedente de pacientes con AR[183]. Por lo tanto, el estado epigenético de su promotor podría explicar por qué se encuentra a mayores niveles en el tejido sinovial de estos pacientes. Los genes cuya metilación cambia de manera más importante y relevante para AR en el array de metilación, han sido validados por técnicas alternativas como la pirosecuenciación de DNA modificado con bisulfito, observando en todos ellos la misma tendencia observada en el estudio global con array.

1.2. Integración de los datos del array de metilación del DNA con datos de expresión génica de RASF y OASF

Hemos analizado las consecuencias funcionales de los cambios de metilación, analizando cómo influyen en la expresión génica. Para ello hemos cruzado los datos generados en el array de metilación, con el de un array de expresión realizado en el mismo tipo de muestras de AR y controles[184]. Hemos detectado 208 genes cuya expresión potencialmente está regulada por los niveles de metilación en su DNA, entre los que se encuentran *HOXC4*, *HOXA11*, *CAPN8* and IL6R, demostrando la importancia de este tipo de regulación epigenética para la expresión de genes clave en el desarrollo de la enfermedad.

1.3. Análisis de microRNAs en RASF y OASF

Aparte de la regulación transcripcional regulada por los niveles de metilación del DNA, otros mecanismos influyen en la funcionalidad final de estos genes. Un mecanismo postranscripcional de regulación de suma importancia son los microRNAs. Hemos analizado los cambios en sus niveles de expresión en algunas de las muestras de fibroblastos obtenidos de pacientes de AR y controles usadas en el array de metilación. Hemos encontrado tanto microRNAs que se sobreexpresan (por ejemplo: miR-203, miR-550) como que se silencian (miR-124, miR-503) en los pacientes con artritis. Dada que la principal función de los micorRNAs es regular la expresión de sus dianas, hemos realizado un análisis de las dianas más robustas de cada microRNA, y lo hemos cruzado con los datos de expresión disponibles. De esta manera hemos identificado varios genes cuyos niveles de expresión y por tanto, de funcionalidad, podrían estar regulados por microRNAs. Entre los genes cuya regulación podría ser llevada a cabo por microRNAs, encontramos genes como CTSC, KLF8 o EBF3, regualdos todos por miR625*, o ITGBL1, que se ve regualdo por miR-551b.

1.4 Análisis integrativo de microRNAs y metilación de DNA muestra múltiples capas regulatorias en genes importantes para la patogénesis de la AR:

La regulación de la expresión génica en condiciones patológicas, así como en condiciones fisiológicas, depende de diversos mecanismos: factores de transcripción, metilación del DNA, marcas en histonas, niveles de microRNAs, etc. Además, unas capas pueden regular a las otras, y es el equilibrio entre unas y otras lo que provoca finalmente una respuesta de activación o represión. Por ello es importante analizar las relaciones entre diferentes capas regulatorias, como son los niveles de metilación, y los niveles de los microRNAs, especialmente en enfermedades complejas como es

la AR. Hemos realizado un análisis integrado de microRNAs y metilación de DNA, y hemos encontrado varios microRNAs cuya expresión parece estar controlada por los niveles de metilación de su promotor, dado que muestran una relación inversa entre su metilación y su expresión. Entre estos, se encuentra el miR-124, miR-378, etc. El segundo análisis integrativo que hemos realizado está basado en la posible sinergia o antagonismo entre estos dos niveles de regulación, y su efecto en la expresión de genes concretos. Hemos encontrado varios genes cuya expresión está reprimida por los dos mecanismos estudiados, de manera que su promotor está hipermetilado, y su mRNA es degradado por la mediación de algunos microRNAs, que se sobreexpresan. Este es el ejemplo de genes como *KLF11*, o *EPHA4*, cuya importancia en el desarrollo de esta patología debe ser estudiado en profundidad, dada la existencia de mecanismos redundantes empleados para su silenciamiento.

1.5. Discusión de la importancia de los hallazgos presentados en el análisis de metilación y del perfil de microRNAs en fibroblastos sinoviales aislados de pacientes con artritis reumatoide y osteoartritis

El estudio de enfermedades multigénicas complejas requiere analizar de manera conjunta varias capas de regulación de la expresión génica que potencialmente pueden desregularse de manera coordinada. Nuestro estudio supera la dificultad de extraer datos de un número limitado de muestras gracias a la integración de información de metilación, microRNAs y expresión en muestras de AR y controles. El análisis integrado nos ha permitido descubrir nuevos genes, cuya desregulación está involucrada en la patogénesis de la AR, como por ejemplo *IL6R*, *CAPN8*, *CD74*, *CCR6*, *DPP4* o *HOXC4*. Hemos comprobado que la expresión aberrante de estos genes puede estar determinado por un patrón irregular en la metilación del DNA de sus genes. Aparte de a nivel transcripcional, la regulación de la expresión puede darse postranscripcionalmente, como es el caso de la ejercida por microRNAs. Hemos analizado los cambios en los niveles que suceden en fibroblastos sinoviales de AR,y hemos observado un gran número de microRNAs sobreexpresados (miR-550, etc.) o silenciados (miR-503, etc.) aberrantemente. De la misma manera tras cruzar los datos de expresión de microRNAs y mRNAs, hemos identificado una serie de genes potencialmente regulados por microRNAs. Un mismo gen puede verse afectado por diferentes capas de regulación. En el presente estudio hemos decidido hacer un análisis integrativo multifactorial en el cual en primer lugar, analizamos la regulación de los niveles de los microRNAs por estudiando la metilación de sus promotores. En segundo lugar hemos combinado los datos de metilación, microRNAs con los de expresión, identificando genes cuyo silenciamiento o sobreexpresión son clave para la patología, dado que hay varios mecanismos (metilación o microRNAs) que se

encargan de controlar sus niveles de expresión. En conclusión, nuestro estudio demuestra la importancia de investigar múltiples capas de regulación no sólo a nivel transcripcional, sino también a nivel postranscripcional. El uso de herramientas genómicas aplicadas al análisis de tipos celulares específicos provenientes de pacientes, nos ofrece una cantidad de información muy elevada. Es por ello que se hace necesario ir un paso más allá, e integrar diferentes tipos de información para obtener aquellos genes más relevantes involucrados en enfermedades complejas como es la AR.

2. Estudio epigenético de la diferenciación Monocito a Osteoclasto: El factor de transcripción mieloide PU.1 dirige la desmetilación mediada por TET2, así como la metilación por parte de DNMT3b en el proceso de diferenciación de monocito a Osteoclasto

2.1. La diferenciación celular y la fusión en osteoclastogénesis sucede con la hipometilación e hipermetilación de genes y rutas clave para la función de resorción ósea

Hemos analizado los cambios de metilación que suceden en el proceso de diferenciación de monocito a osteoclasto, a partir de muestras pareadas procedentes de tres pacientes. Para ello, los niveles de metilación en el DNA genómico de estas muestras ha sido analizado en el array 450K de Illumina, donde se analiza el estado de metilación de más de 480.000 CpGs en el contexto de genes (promotor, cuerpo, 3'UTR), islas CpG, regiones intergénicas, etc. Hemos encontrado 3515 genes cuya metilación cambia durante el proceso (1895 hipometilados y 2054 hipermetilados). Algunos de los genes hipometilados y sobreexpresados, son genes clave en la función del osteoclasto, y son usados de rutina como marcadores, por ejemplo la *Catepsina K* (*CTSK*)[100] , *Fosfatasa Ácida Resistente a Tartrato* (*TRAP* o *ACP5*)[107], *Transmembrane 4 Superfamily, member 7* (*TM4SF7* o *DC-STAMP*)[185], etc. Los cambios observados en el array pudieron confirmarse pirosecuenciando DNA modificado con bisulfito en algunos genes seleccionados, viendo una correlación excelente entre los valores obtenidos por ambas técnicas (R^2=0.97). Hemos confirmado que los cambios de metilación observados no se deben a ningún mecanismo inespecífico que provoque una hipo o hipermetilación general del genoma, estudiando el grado de metilación de varias regiones repetitivas (18S, 28S, SAT2, D4Z4 y NBL2) mediante secuenciación de DNA modificado con bisulfito. No apreciamos cambio alguno en la metilación de estas regiones, por lo que concluimos que los cambios de metilación que suceden

durante la osteoclastogénesis son específicos de algunos promotores, claves para la función del osteoclasto.

2.2. La hipometilación es independiente de replicación y conlleva cambios en 5-hidroximetilcitosina

Para estudiar la dinámica de los cambios de metilación en relación con los cambios de expresión, decidimos estudiar el proceso de diferenciación en una serie temporal. Para ello diferenciamos monocitos a osteoclastos de tres donantes y recogimos muestras de DNA y RNA durante 21 días. Al analizar la dinámica de hipometilación algunos genes (*ACP5*, *CTSK*, *TM7SF4*, *TM4SF19*, *IL7R*), comprobamos que el 60% de los cambios de metilación que sucedían durante el proceso, ocurrían antes del cuarto día. Los cambios de expresión, sucedían a la misma velocidad, o ligeramente retrasados a los cambios de metilación. Posteriormente realizamos un estudio de proliferación de estos progenitores por medio de BrdU, comprobando que en estas condiciones, apenas un 10% de los progenitores de osteoclastos proliferan. Por tanto, la hipometilación observada en ausencia de división celular ha de efectuarse por medio de mecanismos de desmetilación activa. Algunos de los mecanismos propuestos conllevan la presencia de derivados oxidados de la 5-metilcitosina, como la 5-hidroximetilcitosina, 5carboxilcitosina o 5-formilcitosina[24, 25]. Decidimos analizar la presencia de 5-hidroximetilcitosina en promotores que se hipometilan, y comprobamos que en algunos hay un incremento de este intermediario, o bien el intermediario ya estaba presente, predisponiendo al gen para ser hipometilado. Por tanto las evidencias mostradas indican que hay un proceso de desmetilación activa en este proceso de diferenciación, y que este se lleva a cabo a través de la oxidación de la metilcitosina.

2.3. Las regiones que sufren cambios en su metilación están enriquecidas en motivos de unión para AP-1, NF-kB y PU.1, factores de transcripción muy importantes en osteoclastogénesis

Para estudiar el enriquecimiento en motivos de unión de factores de transcripción alrededor de los CpGs con cambios de metilación, creamos una ventana de 500 pares de bases alrededor de los mismos (+- 250pb) y analizamos las secuencias por medio de la herramienta TRANSFAC. Este análisis nos mostró el enriquecimiento específico, y altamente significativo de sitios de unión para AP-1, NF-kB y PU.1 en CpGs hipometilados. Curiosamente, los sitios de unión para PU.1 también están específicamente enriquecidos en CpGs hipermetiladas, indicando el potencial papel bimodal de PU.1 en este proceso de diferenciación. Para comprobar si los sitios de

unión potenciales de los factores de transcripción mencionados anteriormente estaban realmente ocupados por AP-1, NF-kB o PU.1 en genes seleccionados, realizamos un ChIP con c-fos (miembro de AP-1), p65 (miembro de NF-kb) y PU.1 (miembro de la familia de factores de transcripción ETS). En todos los casos pudimos comprobar la presencia del factor de transcripción en varios genes, no así en regiones control negativas como *TRDR1* (un factor testicular), *SAT2* (región repetitiva) o *MYOD1* (factor expresado en músculo).

2.4. PU.1 recluta a DNMT3b y TET2 a genes hipermetilados e hipometilados respectivamente

Para investigar el potencial papel de los factores de transcripción anteriormente mencionados en los cambios de metilación observados, analizamos la implicación de PU.1 y p-65 en el proceso. También analizamos la implicación de la enzima TET2, responsable de la conversión de 5metilcitosina a 5hidroximetilcitosina, 5carboxilcitosina o 5-formilcitosina[24, 30] (intermediarios en el proceso de desmetilación activa), así como el papel de DNMT3b, responsable de transferir grupos metilo *de novo* a la citosina. Tras comprobar por qPCR y western blot la expresión de los citados genes, investigamos si estas proteínas interaccionaban entre ellas. Para ello inmunoprecipitamos p-65 y PU.1 en muestras de osteoclastos a 0, 2 y 4 días tras ser estimulados con MCSF y RANKL. Nuestros resultados muestran un reclutamiento de DNMT3b y TET2 por parte de PU.1, dotándole de un potencial papel bivalente en modificar el entorno epigenético del núcleo de los osteoclastos. Estas interacciones pudieron ser confirmadas por medio de la inmunoprecipitación inversa. Inmunoprecipitamos TET2 y DNMT3b y confirmamos la presencia de PU.1 en el inmunoprecipitado de ambas proteínas. Una vez confirmada la interacción de PU.1 con estas dos proteínas, investigamos en profundidad el papel dual de PU.1. Mediante una inmunoprecipitación de cromatina (ChIP) de monocitos a 0 y 2 días después de ser estimulados con MCSF y RANKL, analizamos el reclutamiento de TET2 y DNMT3b por parte de PU.1 a promotores de genes que sufren cambios en sus niveles de metilación. En el caso de genes que se hipometilan, observamos la presencia de PU.1 desde el principio, y un ligero aumento a día 2. TET2, sin embargo, es reclutado al promotor de los genes hipometilados a día 2. Por otro lado, en genes que se hipermetilan, se aprecia un descenso en la ocupación de PU.1, y sin embargo un aumento en el reclutamiento de DNMT3b a día 2.

2.5. El silenciamiento de PU.1 en monocitos impide la activación de marcadores de osteoclasto, y el reclutamiento de DNMT3b y TET2

Para confirmar el papel de PU.1 en el reclutamiento de maquinaria modificadora de la cromatina, investigamos los efectos de silenciar PU.1 en monocitos humanos primarios por medio de siRNAs (RNA pequeño de interferencia). Una vez confirmado el silenciamiento de PU.1 a nivel de mRNA mediante qPCR (alrededor de un 60% de silenciamiento) y proteína por medio de western blot (alrededor de un 50-60%), comprobamos el efecto de la ausencia de PU.1 en el proceso de diferenciación. En primer lugar comprobamos que la sobreexpresión de marcadores osteoclásticos como son la *Catepsina K* y la enzima *TRAP*, no solo se ve retrasada, sino que se reduce entre un 30 y un 70% a días 4 y 6. También se han analizado los cambios de metilación en los promotores de estos genes, observando un retraso en la hipometilación, que probablemente dificulte la sobreexpresión de sus tránscritos, y muestra la importancia de PU.1 en la consecución del epigenoma adecuado para el proceso de diferenciación. No sólo los genes hipometilados se ven afectados. Los monocitos con PU.1 silenciado, tienen dificultades a la hora de silenciar genes no necesarios para el osteoclasto, como es el caso de los genes *CX3CR1* ó NR4A2. Su expresión es más elevada que en los controles, y sus niveles de metilación son inferiores, puesto que tardan más en ser hipermetilados. No se observaron cambios en dos genes control, uno que se hipometila sin PU.1 pegado a su promotor, y otro que permanece siempre hipermetilado. Finalmente, comprobamos el efecto del silenciamiento de PU.1 en el reclutamiento de DNMT3b y TET2 a los promotores de genes que se hipo (*ACP5*, *TM4SF7*) e hipermetilan (*CX3CR1*, *NR4A2*). Hemos comprobado la reducción en la ocupación de los promotores por parte de PU.1 cuando éste se encuentra silenciado. Éste hecho trae como consecuencia una reducción en el reclutamiento de TET2 a los promotores de genes hipometilados, y una disminución de la presencia de DNMT3b a los genes hipermetilados, demostrando de forma inequívoca el papel central de PU.1 en el reclutamiento de remodeladores cromatínicos, necesarios para la correcta diferenciación de progenitores mieloides a osteoclastos.

2.6. Discusión de la importancia del estudio epigenético de la diferenciación Monocito a Osteoclasto

En este trabajo hemos evidenciado la importancia de la regulación a nivel epigenético de este proceso, tanto en el sentido de la hipometilación, como en el sentido de la hipermetilación. Estos cambios afectan a genes específicos importantes para la función de los osteoclastos maduros, y no suceden de forma inespecífica a nivel de todo el genoma. Gracias a los estudios de enriquecimiento de secuencias, hemos podido comprobar que los cambios de metilación observados en este proceso de diferenciación podrían estar relacionados con factores de transcripción

importantes para el mismo, entre los que cabe destacar el papel de PU.1. Hemos observado un gran número de motivos de unión a este factor de transcripción específicamente en regiones que se hipermetilan o que se hipometilan, indicando su potencial conexión con los cambios epigenéticos observados. Por este motivo, se analizó y confirmó la relación entre PU.1, con la maquinaria epigenética que modifica el estado de metilación en este proceso de diferenciación: DNMT3b y TET2. Estos resultados evidencian el potencial papel dual de adaptador de la maquinaria epigenética, usada no solo para permitir la expresión de genes osteoclásticos, sino también para reprimir la expresión de genes no necesarios para el tipo celular estudiado. Asimismo supone la primera evidencia de interacción entre PU.1 y TET2, y añade más conocimiento al papel de PU.1 no solo como activador transcripcional, sino también como represor en procesos de diferenciación complejos como la osteoclastogénesis.

3. Estudio de los niveles de microRNAs revela el importante papel de los clusters miR-212/132 and miR- 99b/let-7e/125a en la diferenciación de monocito a osteoclasto

3.1. La diferenciación celular y la fusión en osteoclastogénesis sucede con el cambio del perfil de expresión de microRNAs

En este trabajo hemos evidenciado la importancia de la regulación de los microRNAs durante el proceso de diferenciación de osteoclastos. Cuando se analizan los niveles de expresión de microRNAs en diferentes puntos temporales, se aprecia un gran cambio en el perfil de microRNAs a tiempos tempranos, indicando que los cambios en microRNAs han de suceder rápidemante para que el proceso de diferenciación pueda llevarse a cabo de forma correcta. Además, se aprecia la presencia de ocho grupos de microRNAs en función de su perfil de expresión a tiempos tempranos (importantes para la osteoclastogénesis) y tiempos tardíos (importantes para la función del osteoclasto).

3.2. La inhibición de algunos microRNAs demuestra la importancia de los mismos en la diferenciación de monocitos a osteoclastos

Se ha inhibido el efecto de los microRNAs por medio de antagomiRs, y se ha caracterizado el efecto que provocaban en el proceso de diferenciación. Se ha visto que no se retrasa la expresión de los marcadores de osteoclastos. Sin embargo, la expresión de algunas dianas predichas por algoritmos sí que se retrasó.

CONCLUSIONES

De los resultados obtenidos de esta tesis doctoral se puede concluir:

1. Los fibroblastos sinoviales de artritis reumatoide tienen un perfil de metilación del DNA y expresión de microRNAs que los diferencia de los fibroblastos control, y que contribuye a explicar su comportamiento agresivo e invasivo.

2. La existencia de procesos de desregulación epigenética afecta a genes importantes para el sistema inmune o la agresividad de los RASF, entre los que se incluyen: MMP20, RASGRF2, EGF, TIMP2, IL6R, CAPN8, TNFAIP8, CD74, CCR6, DPP4, HOXC4. Estos genes pueden ser usados como marcadores en la identificación de RASF.

3. Los cambios de metilación observados en los RASF tienen relevancia a nivel de la expresión de los genes a los que afectan.

4. Se ha descrito la desregulación a nivel de microRNAs, encontrando algunos ejemplos de expresión aberrante de los mismos: miR-203, miR-124, miR-503, miR-625*, miR-551b y miR-550.

5. Existe una desregulación coordinada a varios niveles, tal como muestra el análisis integrado de los datos de metilación, expresión y microRNAs realizado por primera vez sobre las muestras de muestras de RASF .

6. Durante el proceso de diferenciación de monocito a osteoclasto hay grandes cambios en el metiloma. Se ha identificado hipometilación en 1895 genes e hipermetilación en 2054.

7. Durante la osteoclastogénesis se observa la hipometilación de genes clave para la función del osteoclasto (*CTSK*, *ACP5*, *TM7SF4*) en los cuales, se produce una ganancia de expresión

8. En el proceso de diferenciación de osteoclastos, también se observa hipermetilación de genes propios de otros linajes, como el gen *CX3CR1*, o receptor de la fractalina, que además, se silencia.

9. Los cambios de metilación sufridos en el DNA de estas células durante la osteoclastogénesis son específicos, y no están acompañados de pérdidas de metilación globales en secuencias repetititivas de distintos tipos

10. La pérdida de metilación sucede de manera rápida al inicio de la diferenciación de los monocitos, y ocurre en ausencia de división celular.

11. Durante el proceso de osteoclastogénesis varios genes se hipometilan de forma activa, estando implicados en este proceso derivados oxidados de la 5meC como la 5hmeC.

12. Existe un enriquecimientos en sitios de unión para factores de transcripción en los genes que se hipo- e hipermetilan. En concreto, se ha observado un enriquecimiento en sitios de unión para AP-1, NF-kB y PU.1 en regiones que se hipometilan, y de sitios de unión para PU.1 en las que se hipermetilan. Además, los sitios de unión predichos, están efectivamente ocupados por los factores de transcripción que los reconocen.

13. Se ha descrito la interacción durante la osteoclastogénesis del factor de transcripción PU.1 con enzimas modificadoras de la cromatina tales como DNMT3b y, por primera vez con TET2.

14. PU.1 y TET2 están ocupando promotores de genes que se hipometilan. Por otro lado, PU.1 y DNMT3b ocupan promotores de genes que se hipermetilan.

15. PU.1 es clave para el reclutamiento de la maquinaria epigenética que hipometila (TET2) o hipermetila (DNMT3B) genes clave para la función del osteoclasto, dado que en su ausencia, los cambios observados suceden en menor grado.

16. La expresión de microRNAs varía drásticamente durante el proceso de diferenciación de los osteoclastos.

17. La inhibición del efecto de los microRNAs provoca alteraciones funcionales en los osteocalstos.

BIBLIOGRAFÍA

1. Holliday, R. The inheritance of epigenetic defects. Science 238, 163-70 (1987).
2. Wolffe, A. P. & Matzke, M. A. Epigenetics: regulation through repression. Science 286, 481-6 (1999).
3. Rountree, M. R., Bachman, K. E. & Baylin, S. B. DNMT1 binds HDAC2 and a new co-repressor, DMAP1, to form a complex at replication foci. Nat Genet 25, 269-77 (2000).
4. Robertson, K. D. et al. DNMT1 forms a complex with Rb, E2F1 and HDAC1 and represses transcription from E2F-responsive promoters. Nat Genet 25, 338-42 (2000).
5. Nan, X. et al. Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. Nature 393, 386-9 (1998).
6. Esteller, M. Epigenetics in cancer. N Engl J Med 358, 1148-59 (2008).
7. Espada, J. & Esteller, M. Epigenetic control of nuclear architecture. Cell Mol Life Sci 64, 449-57 (2007).
8. Bestor, T. H. Transposons reanimated in mice. Cell 122, 322-5 (2005).
9. Feinberg, A. P., Cui, H. & Ohlsson, R. DNA methylation and genomic imprinting: insights from cancer into epigenetic mechanisms. Semin Cancer Biol 12, 389-98 (2002).
10. Bird, A. DNA methylation patterns and epigenetic memory. Genes Dev 16, 6-21 (2002).
11. Reik, W. & Lewis, A. Co-evolution of X-chromosome inactivation and imprinting in mammals. Nat Rev Genet 6, 403-10 (2005).
12. van der Maarel, S. M. Epigenetic mechanisms in health and disease. Ann Rheum Dis 67 Suppl 3, iii97-100 (2008).
13. Miller, O. J., Schnedl, W., Allen, J. & Erlanger, B. F. 5-Methylcytosine localised in mammalian constitutive heterochromatin. Nature 251, 636-7 (1974).
14. Bestor, T. H. The DNA methyltransferases of mammals. Hum Mol Genet 9, 2395-402 (2000).
15. Weber, M. et al. Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. Nat Genet 39, 457-66 (2007).
16. Herman, J. G. & Baylin, S. B. Gene silencing in cancer in association with promoter hypermethylation. N Engl J Med 349, 2042-54 (2003).
17. Hellman, A. & Chess, A. Gene body-specific methylation on the active X chromosome. Science 315, 1141-3 (2007).
18. Aran, D., Toperoff, G., Rosenberg, M. & Hellman, A. Replication timing-related and gene body-specific methylation of active human genes. Hum Mol Genet 20, 670-80 (2011).
19. Mayer, W., Niveleau, A., Walter, J., Fundele, R. & Haaf, T. Demethylation of the zygotic paternal genome. Nature 403, 501-2 (2000).
20. Bruniquel, D. & Schwartz, R. H. Selective, stable demethylation of the interleukin-2 gene enhances transcription by an active process. Nat Immunol 4, 235-40 (2003).

21. Martinowich, K. et al. DNA methylation-related chromatin remodeling in activity-dependent BDNF gene regulation. Science 302, 890-3 (2003).
22. Klug, M. et al. Active DNA demethylation in human postmitotic cells correlates with activating histone modifications, but not transcription levels. Genome Biol 11, R63 (2010).
23. Klug, M., Schmidhofer, S., Gebhard, C., Andreesen, R. & Rehli, M. 5-Hydroxymethylcytosine is an essential intermediate of active DNA demethylation processes in primary human monocytes. Genome Biol 14, R46 (2013).
24. He, Y. F. et al. Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. Science 333, 1303-7 (2011).
25. Tahiliani, M. et al. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science 324, 930-5 (2009).
26. Ito, S. et al. Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. Nature 466, 1129-33 (2010).
27. Dawlaty, M. M. et al. Combined deficiency of Tet1 and Tet2 causes epigenetic abnormalities but is compatible with postnatal development. Dev Cell 24, 310-23 (2013).
28. Quivoron, C. et al. TET2 inactivation results in pleiotropic hematopoietic abnormalities in mouse and is a recurrent event during human lymphomagenesis. Cancer Cell 20, 25-38 (2011).
29. Gu, T. P. et al. The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. Nature 477, 606-10 (2011).
30. Kallin, E. M. et al. Tet2 facilitates the derepression of myeloid target genes during CEBPalpha-induced transdifferentiation of pre-B cells. Mol Cell 48, 266-76 (2012).
31. Dawlaty, M. M. et al. Tet1 is dispensable for maintaining pluripotency and its loss is compatible with embryonic and postnatal development. Cell Stem Cell 9, 166-75 (2011).
32. Dawlaty, M. M. et al. Combined deficiency of Tet1 and Tet2 causes epigenetic abnormalities but is compatible with postnatal development. Dev Cell 24, 310-23 (2011).
33. Song, C. X. et al. Genome-wide profiling of 5-formylcytosine reveals its roles in epigenetic priming. Cell 153, 678-91 (2013).
34. Shen, L. et al. Genome-wide analysis reveals TET- and TDG-dependent 5-methylcytosine oxidation dynamics. Cell 153, 692-706 (2013).
35. Lagos-Quintana, M., Rauhut, R., Lendeckel, W. & Tuschl, T. Identification of novel genes coding for small expressed RNAs. Science 294, 853-8 (2001).
36. Lee, Y. et al. The nuclear RNase III Drosha initiates microRNA processing. Nature 425, 415-9 (2003).
37. Hutvagner, G. et al. A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. Science 293, 834-8 (2001).

38. Bernstein, E., Caudy, A. A., Hammond, S. M. & Hannon, G. J. Role for a bidentate ribonuclease in the initiation step of RNA interference. Nature 409, 363-6 (2001).
39. Lee, Y. S. et al. Distinct roles for Drosophila Dicer-1 and Dicer-2 in the siRNA/miRNA silencing pathways. Cell 117, 69-81 (2004).
40. Miyoshi, K., Okada, T. N., Siomi, H. & Siomi, M. C. Characterization of the miRNA-RISC loading complex and miRNA-RISC formed in the Drosophila miRNA pathway. Rna 15, 1282-91 (2009).
41. Zeng, Y., Yi, R. & Cullen, B. R. MicroRNAs and small interfering RNAs can inhibit mRNA expression by similar mechanisms. Proc Natl Acad Sci U S A 100, 9779-84 (2003).
42. Hutvagner, G. & Zamore, P. D. A microRNA in a multiple-turnover RNAi enzyme complex. Science 297, 2056-60 (2002).
43. Lujambio, A. et al. Genetic unmasking of an epigenetically silenced microRNA in human cancer cells. Cancer Res 67, 1424-9 (2007).
44. Betel, D., Wilson, M., Gabow, A., Marks, D. S. & Sander, C. The microRNA.org resource: targets and expression. Nucleic Acids Res 36, D149-53 (2008).
45. Lewis, B. P., Burge, C. B. & Bartel, D. P. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell 120, 15-20 (2005).
46. Lall, S. et al. A genome-wide map of conserved microRNA targets in C. elegans. Curr Biol 16, 460-71 (2006).
47. Kertesz, M., Iovino, N., Unnerstall, U., Gaul, U. & Segal, E. The role of site accessibility in microRNA target recognition. Nat Genet 39, 1278-84 (2007).
48. Klinman, N. R. The "clonal selection hypothesis" and current concepts of B cell tolerance. Immunity 5, 189-95 (1996).
49. Goodnow, C. C., Sprent, J., Fazekas de St Groth, B. & Vinuesa, C. G. Cellular and genetic mechanisms of self tolerance and autoimmunity. Nature 435, 590-7 (2005).
50. Wardemann, H. et al. Predominant autoantibody production by early human B cell precursors. Science 301, 1374-7 (2003).
51. Davidson, A. & Diamond, B. Autoimmune diseases. N Engl J Med 345, 340-50 (2001).
52. de la Rica, L. & Ballestar, E. in Patho-Epigenetics of Disease (eds. Minarovits, J. & Niller, H. H.) (Springer New York, New York, 2012).
53. Stastny, P. Association of the B-cell alloantigen DRw4 with rheumatoid arthritis. N Engl J Med 298, 869-71 (1978).
54. Grumet, F. C., Coukell, A., Bodmer, J. G., Bodmer, W. F. & McDevitt, H. O. Histocompatibility (HL-A) antigens associated with systemic lupus erythematosus. A possible genetic predisposition to disease. N Engl J Med 285, 193-6 (1971).
55. Gilchrist, F. C. et al. Class II HLA associations with autoantibodies in scleroderma: a highly significant role for HLA-DP. Genes Immun 2, 76-81 (2001).

56. Delgado-Vega, A., Sanchez, E., Lofgren, S., Castillejo-Lopez, C. & Alarcon-Riquelme, M. E. Recent findings on genetics of systemic autoimmune diseases. Curr Opin Immunol 22, 698-705 (2010).
57. Gregersen, P. K. Susceptibility genes for rheumatoid arthritis - a rapidly expanding harvest. Bull NYU Hosp Jt Dis 68, 179-82 (2010).
58. Sawcer, S. et al. Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. Nature 476, 214-9 (2011).
59. Eyre, S. et al. High-density genetic mapping identifies new susceptibility loci for rheumatoid arthritis. Nat Genet 44, 1336-40 (2012).
60. Stahl, E. A. et al. Genome-wide association study meta-analysis identifies seven new rheumatoid arthritis risk loci. Nat Genet 42, 508-14 (2010).
61. Grennan, D. M. et al. Family and twin studies in systemic lupus erythematosus. Dis Markers 13, 93-8 (1997).
62. Silman, A. J. et al. Twin concordance rates for rheumatoid arthritis: results from a nationwide study. Br J Rheumatol 32, 903-7 (1993).
63. Pender, M. P. Infection of autoreactive B lymphocytes with EBV, causing chronic autoimmune diseases. Trends Immunol 24, 584-8 (2003).
64. Pollard, K. M., Hultman, P. & Kono, D. H. Toxicology of autoimmune diseases. Chem Res Toxicol 23, 455-66 (2010).
65. Tsai, C. L. et al. Activation of DNA methyltransferase 1 by EBV LMP1 Involves c-Jun NH(2)-terminal kinase signaling. Cancer Res 66, 11668-76 (2006).
66. Strickland, F. M. & Richardson, B. C. Epigenetics in human autoimmunity. Epigenetics in autoimmunity - DNA methylation in systemic lupus erythematosus and beyond. Autoimmunity 41, 278-86 (2008).
67. Hussain, M. et al. Tobacco smoke induces polycomb-mediated repression of Dickkopf-1 in lung cancer cells. Cancer Res 69, 3570-8 (2009).
68. Liu, H., Zhou, Y., Boggs, S. E., Belinsky, S. A. & Liu, J. Cigarette smoke induces demethylation of prometastatic oncogene synuclein-gamma in lung cancer cells by downregulation of DNMT3B. Oncogene 26, 5900-10 (2007).
69. Silman, A. J., Newman, J. & MacGregor, A. J. Cigarette smoking increases the risk of rheumatoid arthritis. Results from a nationwide study of disease-discordant twins. Arthritis Rheum 39, 732-5 (1996).
70. Duke, O., Panayi, G. S., Janossy, G. & Poulter, L. W. An immunohistological analysis of lymphocyte subpopulations and their microenvironment in the synovial membranes of patients with rheumatoid arthritis using monoclonal antibodies. Clin Exp Immunol 49, 22-30 (1982).
71. Choy, E. H. & Panayi, G. S. Cytokine pathways and joint inflammation in rheumatoid arthritis. N Engl J Med 344, 907-16 (2001).
72. Scott, D. L., Wolfe, F. & Huizinga, T. W. Rheumatoid arthritis. Lancet 376, 1094-108 (2010).
73. Aho, K., Koskenvuo, M., Tuominen, J. & Kaprio, J. Occurrence of rheumatoid arthritis in a nationwide series of twins. J Rheumatol 13, 899-902 (1986).

74. Distler, J. H. et al. The induction of matrix metalloproteinase and cytokine expression in synovial fibroblasts stimulated with immune cell microparticles. Proc Natl Acad Sci U S A 102, 2892-7 (2005).
75. Tolboom, T. C. et al. Invasiveness of fibroblast-like synoviocytes is an individual patient characteristic associated with the rate of joint destruction in patients with rheumatoid arthritis. Arthritis Rheum 52, 1999-2002 (2005).
76. Muller-Ladner, U. et al. Synovial fibroblasts of patients with rheumatoid arthritis attach to and invade normal human cartilage when engrafted into SCID mice. Am J Pathol 149, 1607-15 (1996).
77. Baier, A., Meineckel, I., Gay, S. & Pap, T. Apoptosis in rheumatoid arthritis. Curr Opin Rheumatol 15, 274-9 (2003).
78. Lafyatis, R. 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 83, 1267-76 (1989).
79. Corvetta, A., Della Bitta, R., Luchetti, M. M. & Pomponio, G. 5-Methylcytosine content of DNA in blood, synovial mononuclear cells and synovial tissue from patients affected by autoimmune rheumatic diseases. J Chromatogr 566, 481-91 (1991).
80. Neidhart, M. et al. Retrotransposable L1 elements expressed in rheumatoid arthritis synovial tissue: association with genomic DNA hypomethylation and influence on gene expression. Arthritis Rheum 43, 2634-47 (2000).
81. Kuchen, S. et al. The L1 retroelement-related p40 protein induces p38delta MAP kinase. Autoimmunity 37, 57-65 (2004).
82. Karouzakis, E., Gay, R. E., Michel, B. A., Gay, S. & Neidhart, M. DNA hypomethylation in rheumatoid arthritis synovial fibroblasts. Arthritis Rheum 60, 3613-22 (2009).
83. Stanczyk, J. et al. Altered expression of microRNA-203 in rheumatoid arthritis synovial fibroblasts and its role in fibroblast activation. Arthritis Rheum 63, 373-81 (2011).
84. Takami, N. et al. Hypermethylated promoter region of DR3, the death receptor 3 gene, in rheumatoid arthritis synovial cells. Arthritis Rheum 54, 779-87 (2006).
85. Nakano, K., Whitaker, J. W., Boyle, D. L., Wang, W. & Firestein, G. S. DNA methylome signature in rheumatoid arthritis. Ann Rheum Dis 72, 110-7 (2013).
86. Stanczyk, J. et al. Altered expression of MicroRNA in synovial fibroblasts and synovial tissue in rheumatoid arthritis. Arthritis Rheum 58, 1001-9 (2008).
87. Nakasa, T. et al. Expression of microRNA-146 in rheumatoid arthritis synovial tissue. Arthritis Rheum 58, 1284-92 (2008).
88. Trenkmann, M. et al. Tumor necrosis factor alpha-induced microRNA-18a activates rheumatoid arthritis synovial fibroblasts through a feedback loop in NF-kappaB signaling. Arthritis Rheum 65, 916-27 (2013).

89. Philippe, L. et al. MiR-20a regulates ASK1 expression and TLR4-dependent cytokine release in rheumatoid fibroblast-like synoviocytes. Ann Rheum Dis 72, 1071-9 (2013).
90. Nakamachi, Y. et al. MicroRNA-124a is a key regulator of proliferation and monocyte chemoattractant protein 1 secretion in fibroblast-like synoviocytes from patients with rheumatoid arthritis. Arthritis Rheum 60, 1294-304 (2009).
91. Niederer, F. et al. Down-regulation of microRNA-34a* in rheumatoid arthritis synovial fibroblasts promotes apoptosis resistance. Arthritis Rheum 64, 1771-9 (2013).
92. Philippe, L. et al. TLR2 expression is regulated by microRNA miR-19 in rheumatoid fibroblast-like synoviocytes. J Immunol 188, 454-61 (2012).
93. Blair, H. C., Teitelbaum, S. L., Ghiselli, R. & Gluck, S. Osteoclastic bone resorption by a polarized vacuolar proton pump. Science 245, 855-7 (1989).
94. Yasuda, H. et al. Osteoclast differentiation factor is a ligand for osteoprotegerin/osteoclastogenesis-inhibitory factor and is identical to TRANCE/RANKL. Proc Natl Acad Sci U S A 95, 3597-602 (1998).
95. Wiktor-Jedrzejczak, W. et al. Total absence of colony-stimulating factor 1 in the macrophage-deficient osteopetrotic (op/op) mouse. Proc Natl Acad Sci U S A 87, 4828-32 (1990).
96. Lacey, D. L. et al. Osteoprotegerin ligand is a cytokine that regulates osteoclast differentiation and activation. Cell 93, 165-76 (1998).
97. Saltel, F., Chabadel, A., Bonnelye, E. & Jurdic, P. Actin cytoskeletal organisation in osteoclasts: a model to decipher transmigration and matrix degradation. Eur J Cell Biol 87, 459-68 (2008).
98. Nicholson, G. C. et al. Induction of osteoclasts from CD14-positive human peripheral blood mononuclear cells by receptor activator of nuclear factor kappaB ligand (RANKL). Clin Sci (Lond) 99, 133-40 (2000).
99. Sorensen, M. G. et al. Characterization of osteoclasts derived from CD14+ monocytes isolated from peripheral blood. J Bone Miner Metab 25, 36-45 (2007).
100. Drake, F. H. et al. Cathepsin K, but not cathepsins B, L, or S, is abundantly expressed in human osteoclasts. J Biol Chem 271, 12511-6 (1996).
101. Wucherpfennig, A. L., Li, Y. P., Stetler-Stevenson, W. G., Rosenberg, A. E. & Stashenko, P. Expression of 92 kD type IV collagenase/gelatinase B in human osteoclasts. J Bone Miner Res 9, 549-56 (1994).
102. Sundaram, K. et al. RANK ligand signaling modulates the matrix metalloproteinase-9 gene expression during osteoclast differentiation. Exp Cell Res 313, 168-78 (2007).
103. Vaananen, H. K. Immunohistochemical localization of carbonic anhydrase isoenzymes I and II in human bone, cartilage and giant cell tumor. Histochemistry 81, 485-7 (1984).
104. Sly, W. S., Hewett-Emmett, D., Whyte, M. P., Yu, Y. S. & Tashian, R. E. Carbonic anhydrase II deficiency identified as the primary defect in the

autosomal recessive syndrome of osteopetrosis with renal tubular acidosis and cerebral calcification. Proc Natl Acad Sci U S A 80, 2752-6 (1983).
105. Ek-Rylander, B., Flores, M., Wendel, M., Heinegard, D. & Andersson, G. Dephosphorylation of osteopontin and bone sialoprotein by osteoclastic tartrate-resistant acid phosphatase. Modulation of osteoclast adhesion in vitro. J Biol Chem 269, 14853-6 (1994).
106. Hammarstrom, L. E., Anderson, T. R., Marks, S. C., Jr. & Toverud, S. U. Inhibition by dithionite and reactivation by iron of the tartrate-resistant acid phosphatase in bone of osteopetrotic (ia) rats. J Histochem Cytochem 31, 1167-74 (1983).
107. Minkin, C. Bone acid phosphatase: tartrate-resistant acid phosphatase as a marker of osteoclast function. Calcif Tissue Int 34, 285-90 (1982).
108. Tolar, J., Teitelbaum, S. L. & Orchard, P. J. Osteopetrosis. N Engl J Med 351, 2839-49 (2004).
109. Rachner, T. D., Khosla, S. & Hofbauer, L. C. Osteoporosis: now and the future. Lancet 377, 1276-87 (2011).
110. Mundy, G. R., Raisz, L. G., Cooper, R. A., Schechter, G. P. & Salmon, S. E. Evidence for the secretion of an osteoclast stimulating factor in myeloma. N Engl J Med 291, 1041-6 (1974).
111. Yoneda, T. Cellular and molecular mechanisms of breast and prostate cancer metastasis to bone. Eur J Cancer 34, 240-5 (1998).
112. Mii, Y. et al. Osteoclast origin of giant cells in giant cell tumors of bone: ultrastructural and cytochemical study of six cases. Ultrastruct Pathol 15, 623-9 (1991).
113. Joyner, C. J., Quinn, J. M., Triffitt, J. T., Owen, M. E. & Athanasou, N. A. Phenotypic characterisation of mononuclear and multinucleated cells of giant cell tumour of bone. Bone Miner 16, 37-48 (1992).
114. Bromley, M. & Woolley, D. E. Chondroclasts and osteoclasts at subchondral sites of erosion in the rheumatoid joint. Arthritis Rheum 27, 968-75 (1984).
115. Gravallese, E. M. et al. Identification of cell types responsible for bone resorption in rheumatoid arthritis and juvenile rheumatoid arthritis. Am J Pathol 152, 943-51 (1998).
116. Pettit, A. R. et al. TRANCE/RANKL knockout mice are protected from bone erosion in a serum transfer model of arthritis. Am J Pathol 159, 1689-99 (2001).
117. Redlich, K. et al. Osteoclasts are essential for TNF-alpha-mediated joint destruction. J Clin Invest 110, 1419-27 (2002).
118. Anderson, D. M. et al. A homologue of the TNF receptor and its ligand enhance T-cell growth and dendritic-cell function. Nature 390, 175-9 (1997).
119. Takayanagi, H. et al. Involvement of receptor activator of nuclear factor kappaB ligand/osteoclast differentiation factor in osteoclastogenesis from synoviocytes in rheumatoid arthritis. Arthritis Rheum 43, 259-69 (2000).
120. Cohen, S. B. et al. Denosumab treatment effects on structural damage, bone mineral density, and bone turnover in rheumatoid arthritis: a twelve-month,

multicenter, randomized, double-blind, placebo-controlled, phase II clinical trial. Arthritis Rheum 58, 1299-309 (2008).
121. Deodhar, A. et al. Denosumab-mediated increase in hand bone mineral density associated with decreased progression of bone erosion in rheumatoid arthritis patients. Arthritis Care Res (Hoboken) 62, 569-74 (2010).
122. Wong, B. R. et al. TRANCE, a TNF family member, activates Akt/PKB through a signaling complex involving TRAF6 and c-Src. Mol Cell 4, 1041-9 (1999).
123. Kobayashi, N. et al. Segregation of TRAF6-mediated signaling pathways clarifies its role in osteoclastogenesis. Embo J 20, 1271-80 (2001).
124. Blank, U., Launay, P., Benhamou, M. & Monteiro, R. C. Inhibitory ITAMs as novel regulators of immunity. Immunol Rev 232, 59-71 (2009).
125. Humphrey, M. B. et al. TREM2, a DAP12-associated receptor, regulates osteoclast differentiation and function. J Bone Miner Res 21, 237-45 (2006).
126. Koga, T. et al. Costimulatory signals mediated by the ITAM motif cooperate with RANKL for bone homeostasis. Nature 428, 758-63 (2004).
127. Negishi-Koga, T. & Takayanagi, H. Ca2+-NFATc1 signaling is an essential axis of osteoclast differentiation. Immunol Rev 231, 241-56 (2009).
128. Ikeda, F. et al. Critical roles of c-Jun signaling in regulation of NFAT family and RANKL-regulated osteoclast differentiation. J Clin Invest 114, 475-84 (2004).
129. Takayanagi, H. et al. Induction and activation of the transcription factor NFATc1 (NFAT2) integrate RANKL signaling in terminal differentiation of osteoclasts. Dev Cell 3, 889-901 (2002).
130. Sharma, S. M. et al. MITF and PU.1 recruit p38 MAPK and NFATc1 to target genes during osteoclast differentiation. J Biol Chem 282, 15921-9 (2007).
131. Husson, H., Mograbi, B., Schmid-Antomarchi, H., Fischer, S. & Rossi, B. CSF-1 stimulation induces the formation of a multiprotein complex including CSF-1 receptor, c-Cbl, PI 3-kinase, Crk-II and Grb2. Oncogene 14, 2331-8 (1997).
132. Arai, F. et al. Commitment and differentiation of osteoclast precursor cells by the sequential expression of c-Fms and receptor activator of nuclear factor kappaB (RANK) receptors. J Exp Med 190, 1741-54 (1999).
133. Hsu, H. et al. Tumor necrosis factor receptor family member RANK mediates osteoclast differentiation and activation induced by osteoprotegerin ligand. Proc Natl Acad Sci U S A 96, 3540-5 (1999).
134. Rothe, M., Sarma, V., Dixit, V. M. & Goeddel, D. V. TRAF2-mediated activation of NF-kappa B by TNF receptor 2 and CD40. Science 269, 1424-7 (1995).
135. Galibert, L., Tometsko, M. E., Anderson, D. M., Cosman, D. & Dougall, W. C. The involvement of multiple tumor necrosis factor receptor (TNFR)-associated factors in the signaling mechanisms of receptor activator of NF-kappaB, a member of the TNFR superfamily. J Biol Chem 273, 34120-7 (1998).
136. Ye, H. et al. Distinct molecular mechanism for initiating TRAF6 signalling. Nature 418, 443-7 (2002).
137. Armstrong, A. P. et al. A RANK/TRAF6-dependent signal transduction pathway is essential for osteoclast cytoskeletal organization and resorptive function. J Biol Chem 277, 44347-56 (2002).

138. Darnay, B. G., Haridas, V., Ni, J., Moore, P. A. & Aggarwal, B. B. Characterization of the intracellular domain of receptor activator of NF-kappaB (RANK). Interaction with tumor necrosis factor receptor-associated factors and activation of NF-kappab and c-Jun N-terminal kinase. J Biol Chem 273, 20551-5 (1998).
139. Yamashita, T. et al. NF-kappaB p50 and p52 regulate receptor activator of NF-kappaB ligand (RANKL) and tumor necrosis factor-induced osteoclast precursor differentiation by activating c-Fos and NFATc1. J Biol Chem 282, 18245-53 (2007).
140. Mansky, K. C., Sankar, U., Han, J. & Ostrowski, M. C. Microphthalmia transcription factor is a target of the p38 MAPK pathway in response to receptor activator of NF-kappa B ligand signaling. J Biol Chem 277, 11077-83 (2002).
141. Sato, K. et al. Regulation of osteoclast differentiation and function by the CaMK-CREB pathway. Nat Med 12, 1410-6 (2006).
142. Tondravi, M. M. et al. Osteopetrosis in mice lacking haematopoietic transcription factor PU.1. Nature 386, 81-4 (1997).
143. Weilbaecher, K. N. et al. Linkage of M-CSF signaling to Mitf, TFE3, and the osteoclast defect in Mitf(mi/mi) mice. Mol Cell 8, 749-58 (2001).
144. Matsuo, K. et al. Nuclear factor of activated T-cells (NFAT) rescues osteoclastogenesis in precursors lacking c-Fos. J Biol Chem 279, 26475-80 (2004).
145. Nguyen, V. C. et al. Localization of the human oncogene SPI1 on chromosome 11, region p11.22. Hum Genet 84, 542-6 (1990).
146. Goebl, M. K. The PU.1 transcription factor is the product of the putative oncogene Spi-1. Cell 61, 1165-6 (1990).
147. Klemsz, M. J., McKercher, S. R., Celada, A., Van Beveren, C. & Maki, R. A. The macrophage and B cell-specific transcription factor PU.1 is related to the ets oncogene. Cell 61, 113-24 (1990).
148. Kodandapani, R. et al. A new pattern for helix-turn-helix recognition revealed by the PU.1 ETS-domain-DNA complex. Nature 380, 456-60 (1996).
149. Gupta, P., Gurudutta, G. U., Saluja, D. & Tripathi, R. P. PU.1 and partners: regulation of haematopoietic stem cell fate in normal and malignant haematopoiesis. J Cell Mol Med 13, 4349-63 (2009).
150. Pham, T. H. et al. Mechanisms of in vivo binding site selection of the hematopoietic master transcription factor PU.1. Nucleic Acids Res 41, 6391-6402 (2013).
151. Scott, E. W., Simon, M. C., Anastasi, J. & Singh, H. Requirement of transcription factor PU.1 in the development of multiple hematopoietic lineages. Science 265, 1573-7 (1994).
152. Back, J., Dierich, A., Bronn, C., Kastner, P. & Chan, S. PU.1 determines the self-renewal capacity of erythroid progenitor cells. Blood 103, 3615-23 (2004).

153. Nutt, S. L., Metcalf, D., D'Amico, A., Polli, M. & Wu, L. Dynamic regulation of PU.1 expression in multipotent hematopoietic progenitors. J Exp Med 201, 221-31 (2005).
154. Kwon, O. H., Lee, C. K., Lee, Y. I., Paik, S. G. & Lee, H. J. The hematopoietic transcription factor PU.1 regulates RANK gene expression in myeloid progenitors. Biochem Biophys Res Commun 335, 437-46 (2005).
155. Sato, M. et al. Microphthalmia-associated transcription factor interacts with PU.1 and c-Fos: determination of their subcellular localization. Biochem Biophys Res Commun 254, 384-7 (1999).
156. Luchin, A. et al. Genetic and physical interactions between Microphthalmia transcription factor and PU.1 are necessary for osteoclast gene expression and differentiation. J Biol Chem 276, 36703-10 (2001).
157. Partington, G. A., Fuller, K., Chambers, T. J. & Pondel, M. Mitf-PU.1 interactions with the tartrate-resistant acid phosphatase gene promoter during osteoclast differentiation. Bone 34, 237-45 (2004).
158. So, H. et al. Microphthalmia transcription factor and PU.1 synergistically induce the leukocyte receptor osteoclast-associated receptor gene expression. J Biol Chem 278, 24209-16 (2003).
159. Kim, K. et al. Nuclear factor of activated T cells c1 induces osteoclast-associated receptor gene expression during tumor necrosis factor-related activation-induced cytokine-mediated osteoclastogenesis. J Biol Chem 280, 35209-16 (2005).
160. Matsumoto, M. et al. Essential role of p38 mitogen-activated protein kinase in cathepsin K gene expression during osteoclastogenesis through association of NFATc1 and PU.1. J Biol Chem 279, 45969-79 (2004).
161. Crotti, T. N. et al. PU.1 and NFATc1 mediate osteoclastic induction of the mouse beta3 integrin promoter. J Cell Physiol 215, 636-44 (2008).
162. Kihara-Negishi, F. et al. In vivo complex formation of PU.1 with HDAC1 associated with PU.1-mediated transcriptional repression. Oncogene 20, 6039-47 (2001).
163. Pospisil, V. et al. Epigenetic silencing of the oncogenic miR-17-92 cluster during PU.1-directed macrophage differentiation. Embo J 30, 4450-64 (2011).
164. Ridinger-Saison, M. et al. Epigenetic silencing of Bim transcription by Spi-1/PU.1 promotes apoptosis resistance in leukaemia. Cell Death Differ 20, 1268-78 (2013).
165. Burda, P. et al. PU.1 activation relieves GATA-1-mediated repression of Cebpa and Cbfb during leukemia differentiation. Mol Cancer Res 7, 1693-703 (2009).
166. Suzuki, M. et al. Site-specific DNA methylation by a complex of PU.1 and Dnmt3a/b. Oncogene 25, 2477-88 (2006).
167. Sugatani, T. & Hruska, K. A. MicroRNA-223 is a key factor in osteoclast differentiation. J Cell Biochem 101, 996-9 (2007).
168. Mizoguchi, F. et al. Osteoclast-specific Dicer gene deficiency suppresses osteoclastic bone resorption. J Cell Biochem 109, 866-75 (2010).

169. Sugatani, T. & Hruska, K. A. Impaired micro-RNA pathways diminish osteoclast differentiation and function. J Biol Chem 284, 4667-78 (2009).
170. Li, Y. T. et al. Brief report: amelioration of collagen-induced arthritis in mice by lentivirus-mediated silencing of microRNA-223. Arthritis Rheum 64, 3240-5 (2012).
171. Shibuya, H. et al. Overexpression of microRNA-223 in rheumatoid arthritis synovium controls osteoclast differentiation. Mod Rheumatol 23, 674-85 (2013).
172. Mann, M., Barad, O., Agami, R., Geiger, B. & Hornstein, E. miRNA-based mechanism for the commitment of multipotent progenitors to a single cellular fate. Proc Natl Acad Sci U S A 107, 15804-9 (2010).
173. Bluml, S. et al. Essential role of microRNA-155 in the pathogenesis of autoimmune arthritis in mice. Arthritis Rheum 63, 1281-8 (2011).
174. Zhang, J. et al. Interferon-beta-induced miR-155 inhibits osteoclast differentiation by targeting SOCS1 and MITF. FEBS Lett 586, 3255-62 (2012).
175. Sugatani, T., Vacher, J. & Hruska, K. A. A microRNA expression signature of osteoclastogenesis. Blood 117, 3648-57 (2011).
176. Nakasa, T., Shibuya, H., Nagata, Y., Niimoto, T. & Ochi, M. The inhibitory effect of microRNA-146a expression on bone destruction in collagen-induced arthritis. Arthritis Rheum 63, 1582-90 (2011).
177. Rossi, M. et al. miR-29b negatively regulates human osteoclastic cell differentiation and function: Implications for the treatment of multiple myeloma-related bone disease. J Cell Physiol 228, 1506-15 (2013).
178. Lee, Y. et al. MicroRNA-124 regulates osteoclast differentiation. Bone 56, 383-389 (2013).
179. Nishimoto, N., Kishimoto, T. & Yoshizaki, K. Anti-interleukin 6 receptor antibody treatment in rheumatic disease. Ann Rheum Dis 59 Suppl 1, i21-7 (2000).
180. Thompson, C. A. FDA approves tocilizumab to treat rheumatoid arthritis. Am J Health Syst Pharm 67, 254 (2010).
181. Ruth, J. H. et al. Role of macrophage inflammatory protein-3alpha and its ligand CCR6 in rheumatoid arthritis. Lab Invest 83, 579-88 (2003).
182. Yazbeck, R., Howarth, G. S. & Abbott, C. A. Dipeptidyl peptidase inhibitors, an emerging drug class for inflammatory disease? Trends Pharmacol Sci 30, 600-7 (2009).
183. Waldburger, J. M. et al. Autoimmunity and inflammation are independent of class II transactivator type PIV-dependent class II major histocompatibility complex expression in peripheral tissues during collagen-induced arthritis. Arthritis Rheum 63, 3354-63 (2011).
184. Del Rey, M. J. et al. Transcriptome analysis reveals specific changes in osteoarthritis synovial fibroblasts. Ann Rheum Dis 71, 275-80 (2012).
185. Kim, K., Lee, S. H., Ha Kim, J., Choi, Y. & Kim, N. NFATc1 induces osteoclast fusion via up-regulation of Atp6v0d2 and the dendritic cell-specific transmembrane protein (DC-STAMP). Mol Endocrinol 22, 176-85 (2008).

ANEXOS

PU.1 targets TET2-coupled demethylation and DNMT3b-mediated methylation in monocyte-to-osteoclast differentiation
de la Rica L, Rodríguez-Ubreva, J, García M, Islam, A.B.M.M.K., Urquiza JM, Hernando H, Christensen J, Helin K, Gómez-Vaquero C, Ballestar E. Genome Biol. 2013,Sep 12;14(9):R99.

Identification of novel markers in rheumatoid arthritis through integrated analysis of DNA methylation and microRNA expression
de la Rica L, Urquiza JM, Gómez-Cabrero D, Islam A.B.M.M.K, López-Bigas N, Tegnér J, Toes R E.M., Ballestar E. J Autoimmun. 2013 Jan; (41) ; 6-16 .

Epigenetic regulation of PRAME in acute myeloid leukemia is different compared to CD34+ cells from healthy donors: Effect of 5-AZA treatment
Gutierrez-Cosío S, de la Rica L, Ballestar E, Santamaría C, Sánchez-Abarca LI, Caballero-Velazquez T, Blanco B, Calderón C, Herrero-Sánchez C, Carrancio S, Ciudad L, Cañizo C, Miguel JF, Pérez-Simón JA. Leuk Res 2012;36(7):895-9.

Journal of Autoimmunity 41 (2013) 6–16

Contents lists available at SciVerse ScienceDirect

Journal of Autoimmunity

journal homepage: www.elsevier.com/locate/jautimm

Identification of novel markers in rheumatoid arthritis through integrated analysis of DNA methylation and microRNA expression

Lorenzo de la Rica [a], José M. Urquiza [a], David Gómez-Cabrero [b], Abul B.M.M.K. Islam [c,d], Nuria López-Bigas [c,e], Jesper Tegnér [b], René E.M. Toes [f], Esteban Ballestar [a,*]

[a] Chromatin and Disease Group, Cancer Epigenetics and Biology Programme (PEBC), Bellvitge Biomedical Research Institute (IDIBELL), 08907 L'Hospitalet de Llobregat, Barcelona, Spain
[b] Department of Medicine, Karolinska Institutet, Computational Medicine Unit, Centre for Molecular Medicine, Swedish e-science Research Centre (SeRC), Solna, Stockholm, Sweden
[c] Department of Experimental and Health Sciences, Barcelona Biomedical Research Park, Universitat Pompeu Fabra (UPF), 08003 Barcelona, Spain
[d] Department of Genetic Engineering and Biotechnology, University of Dhaka, Dhaka 1000, Bangladesh
[e] Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain
[f] Department of Rheumatology, Leiden University Medical Center, Leiden, The Netherlands

ARTICLE INFO

Article history:
Received 14 December 2012
Accepted 16 December 2012

Keywords:
Rheumatoid arthritis
Rheumatoid arthritis synovial fibroblasts
DNA methylation
Epigenetic
MicroRNAs
Integration

ABSTRACT

Autoimmune rheumatic diseases are complex disorders, whose etiopathology is attributed to a crosstalk between genetic predisposition and environmental factors. Both variants of autoimmune susceptibility genes and environment are involved in the generation of aberrant epigenetic profiles in a cell-specific manner, which ultimately result in dysregulation of expression. Furthermore, changes in miRNA expression profiles also cause gene dysregulation associated with aberrant phenotypes. In rheumatoid arthritis, several cell types are involved in the destruction of the joints, synovial fibroblasts being among the most important. In this study we performed DNA methylation and miRNA expression screening of a set of rheumatoid arthritis synovial fibroblasts and compared the results with those obtained from osteoarthritis patients with a normal phenotype. DNA methylation screening allowed us to identify changes in novel key target genes like *IL6R*, *CAPN8* and *DPP4*, as well as several *HOX* genes. A significant proportion of genes undergoing DNA methylation changes were inversely correlated with expression. miRNA screening revealed the existence of subsets of miRNAs that underwent changes in expression. Integrated analysis highlighted sets of miRNAs that are controlled by DNA methylation, and genes that are regulated by DNA methylation and are targeted by miRNAs with a potential use as clinical markers. Our study enabled the identification of novel dysregulated targets in rheumatoid arthritis synovial fibroblasts and generated a new workflow for the integrated analysis of miRNA and epigenetic control.

1. Introduction

Rheumatoid arthritis (RA) is a chronic autoimmune inflammatory disease characterized by the progressive destruction of the joints. RA pathogenesis involves a variety of cell types, including several lymphocyte subsets, dendritic cells, osteoclasts and synovial fibroblasts (SFs). In healthy individuals, SFs are essential to keep the joints in shape, doing so by providing nutrients, facilitating matrix remodeling and contributing to tissue repair [1]. In contrast to normal SFs or those isolated from patients with osteoarthritis (osteoarthritis synovial fibroblasts, OASFs), rheumatoid arthritis synovial fibroblasts (RASFs) show activities associated with an aggressive phenotype, like upregulated expression of protooncogenes, specific matrix-degrading enzymes, adhesion molecules, and cytokines [2]. Differences in phenotype and gene expression between RASFs and their normal counterparts reflect a profound change in processes involved in gene regulation at the transcriptional and post-transcriptional levels. The first group comprises epigenetic mechanisms, like DNA methylation, whilst miRNA control constitutes one of the best studied mechanisms of the second.

DNA methylation takes place in cytosine bases followed by guanines. In relation with transcription, the repressive role of methylation at CpG sites located at or near the transcription start sites of genes, especially when those CpGs are clustered as CpG islands, is well established [3]. Methylation of CpGs located in other

* Corresponding author. Tel.: +34 932607133; fax: +34 932607219.
E-mail address: eballestar@idibell.org (E. Ballestar).

0896-8411/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.jaut.2012.12.005

regions like gene bodies is also involved in gene regulation [4,5]. At the other side of gene regulation lie microRNAs (miRNAs), a class of endogenous, small, non-coding regulatory RNA molecules that modulate the expression of multiple target genes at the post-transcriptional level and that are implicated in a wide variety of cellular processes and disease pathogenesis [6].

The study of epigenetic- and miRNA-mediated alterations in association with disease is becoming increasingly important as these processes directly participate in the generation of aberrant profiles of gene expression ultimately determining cell function and are pharmacologically reversible. Epigenetics is particularly relevant in autoimmune rheumatic diseases as it is highly dependent on environmental effects. As indicated above, both genetics and environmental factors contribute to ethiopathology of autoimmune rheumatic disorders. This double contribution is typically exemplified by the partial concordance in monozygotic twins (MZ) [7,8]. It is of inherent interest to identify autoimmune disease phenotypes for which the environment plays a critical role [9]. Many environmental factors, including exposure to chemicals, tobacco smoke, radiation, ultraviolet (UV) light and infectious agents among other external factors, are associated with the development of autoimmune rheumatic disorders [10]. Most of these environmental factors are now known to directly or indirectly induce epigenetic changes, which modulate gene expression and therefore associate with changes in cell function. For this reason, epigenetics provides a source of molecular mechanisms that can explain the environmental effects on the development of autoimmune disorders [11]. The close relationship between environment and epigenetic status and autoimmune rheumatic disease is also exemplified by using animal models [12]. This type of studies is also essential for the identification of novel clinical markers for disease onset, progression and response to treatments.

In this line, initial reports demonstrated hypomethylation-associated reactivation of endogenous retroviral element L1 in the RA synovial lining at joints [13]. Additional sequences have since been found to undergo hypomethylation in RASFs, like *IL-6* [14] and *CXCL12* [15]. Candidate gene analysis has also enabled genes to be identified that are hypermethylated in RASFs [16]. More recently, DNA methylation profiling of RASFs versus OASFs has led to the identification of a number of hypomethylated and hypermethylated genes [17]. With respect to miRNAs, reduced miR-34a levels have been linked with increased resistance of RASFs to apoptosis [18], and lower miR-124a levels in RASFs impact its targets, CDK-2 and MCP-1 [19]. Conversely, miR-203 shows increased expression in RASFs [20]. Interestingly, overexpression of this miR-203 is demethylation-dependent, highlighting the importance of investigating multiple levels of regulation and the need to use integrated strategies that consider interconnected mechanisms.

In this study, we have performed the first integrated comparison of DNA methylation and miRNA expression data, together with mRNA expression data from RASFs versus OASFs (Fig. 1) in order to investigate the relevance of these changes in these cells and to overcome the limitations of using a small number of samples. Our analysis identifies novel targets of DNA methylation- and miRNA-associated dysregulation in RA. Integration of the analysis of these two datasets suggests the existence of several genes for which the two mechanisms could act in the same or in opposite directions.

2. Material and methods

2.1. Subjects and sample preparation

Fibroblast-like synoviocytes (FLSs) were isolated from synovial tissues extracted from RA and OA patients at the time of joint replacement in the Department of Rheumatology of Leiden University Medical Center. All RA patients met the 1987 criteria of the American College of Rheumatology. Before tissue collection, permission consistent with the protocol of the Helsinki International Conference on Harmonisation Good Clinical Practice was obtained. All individuals gave informed consent. Synovial tissues

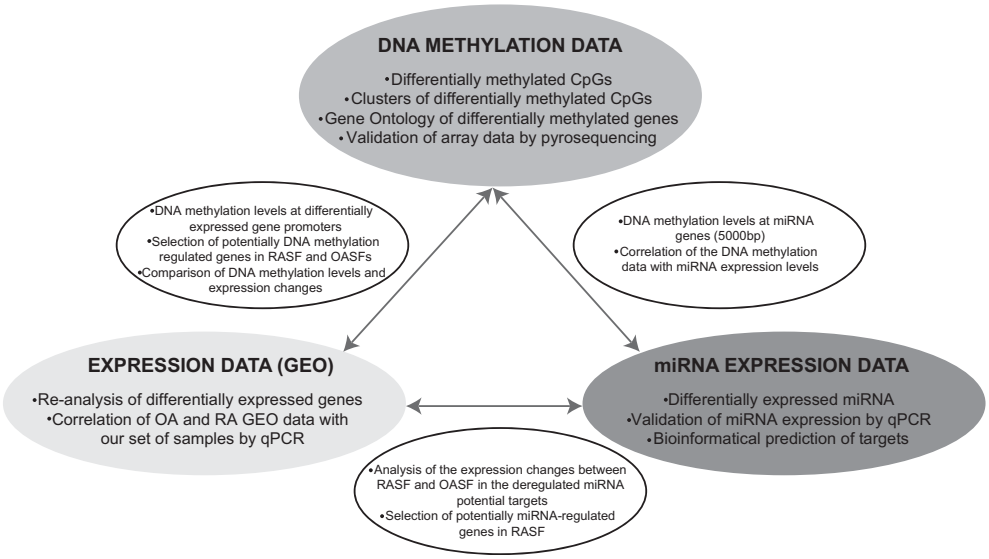


Fig. 1. Scheme depicting the strategy designed in this study where DNA methylation and miRNA data are integrated with expression array data. The grey oval areas show the type of information that individual analysis of DNA methylation, miRNA expression and expression datasets can provide. This is listed within these grey oval areas and are described in detail in the Results section. Between these grey oval areas, smaller elliptical panels show the type of analysis that can provide the combined information between DNA methylation and expression datasets (left), DNA methylation and miRNA expression datasets (right), or expression and miRNA expression datasets (bottom).

were collected during the arthroscopy, frozen in Tissue-Tek OCT compound (Sakura Finetek, Zoeterwoude, Netherlands) and cut into 5-μm slices using a cryotome (Leica CM 1900). Fibroblast cultures were maintained in Dulbecco's modified Eagle's medium supplemented with 10% fetal calf serum.

2.2. DNA methylation profiling using universal bead arrays

Infinium HumanMethylation450 BeadChips (Illumina, Inc.) were used to analyze DNA methylation. With this analysis it is possible to cover >485,000 methylation sites per sample at single-nucleotide resolution. This panel covers 99% of RefSeq genes, with an average of 17 CpG sites per gene region distributed across the promoter, 5′UTR, first exon, gene body, and 3′UTR. It covers 96% of CpG islands, with additional coverage in island shores and the regions flanking them. Bisulfite conversion of DNA samples was done using the EZ DNA methylation kit (Zymo Research, Orange, CA). After bisulfite treatment, the remaining assay steps were identical to those of the Infinium Methylation Assay, using reagents and conditions supplied and recommended by the manufacturer. Two technical replicates of each bisulfite-converted sample were run. The results were all in close agreement and were averaged for subsequent analysis. The array hybridization was conducted under a temperature gradient program, and arrays were imaged using a BeadArray Reader (Illumina Inc.). The image processing and intensity data extraction software and procedures were those described by Bibikova and colleagues [21]. Each methylation datum point was represented as a combination of the Cy3 and Cy5 fluorescent intensities from the M (methylated) and U (unmethylated) alleles. Background intensity, computed from a set of negative controls, was subtracted from each datum point.

2.3. Detection of differentially methylated CpGs

Differentially methylated CpGs were selected using an algorithm in the statistical computing language R [22], version 2.14.0. In order to process Illumina Infinium HumanMethylation450 methylation data, we used the methods available in the LIMMA and LUMI packages [23] from the Bioconductor repository [24]. Before statistical analysis, a pre-process stage was applied, whose main steps were: 1) Adjusting color balance, i.e., normalizing between two color channels; 2) Quantile normalizing based on color balance-adjusted data; 3) Removing probes with a detection p-value > 0.01; 4) Filtering probes located in sex chromosomes; 5) Filtering probes considered to be SNPs (single nucleotide polymorphisms). Specifically, the probes were filtered out using Illumina identifiers for SNPs, i.e. those probes with an "rs" prefix in their name; 6) Non-specific filtering based on the IQR (interquartile range) [25], using 0.20 as the threshold value.

Subsequently, a Bayes-moderated t-test was carried out using LIMMA [26]. Several criteria have been proposed to identify significant differences in methylated CpGs. In this study, we adopted the median-difference beta-value between the two sample groups for each CpG [27,28]. Specifically we considered a probe as differentially methylated if (1) the absolute value of the median-difference between b-values is higher than 0.1 and the statistical test was significant (p-value < 0.05).

2.4. Identification of genomic clusters of differentially methylated CpGs

A clustering method available in Charm package [29] was applied to the differentially methylated CpGs. Although Charm is a package specific for analyzing DNA methylation data from two-color Nimblegen microarrays, we reimplemented the code to invoke the main clustering function using genomic CpG localization. By using this approach, we identified Differentially Methylated Regions (DMR) by grouping differentially methylated probes closer than 500 pbs. In this analysis, the considered lists of CpGs were those associated with a value of $p < 0.01$.

2.5. Bisulfite pyrosequencing

CpGs were selected for technical validation of Infinium Methylation 450K by the bisulfite pyrosequencing technique in the RASF and OASF samples. CpG island DNA methylation status was determined by sequencing bisulfite-modified genomic DNA. Bisulfite modification of genomic DNA was carried out as described by Herman and colleagues [30]. 2 μl of the converted DNA (corresponding to approximately 20–30 ng) were then used as a template in each subsequent PCR. Primers for PCR amplification and sequencing were designed with the PyroMark® Assay Design 2.0 software (Qiagen). PCRs were performed with the HotStart Taq DNA polymerase PCR kit (Qiagen) and the success of amplification was assessed by agarose gel electrophoresis. Pyrosequencing of the PCR products was performed with the Pyromark™ Q24 system (Qiagen). All primer sequences are listed in Supplementary Table 1.

2.6. Gene expression data analysis and comparison of DNA expression and DNA methylation data

To compare expression and methylation data, we used RASF and OASF expression data from the Gene Expression Omnibus (GEO) under the accession number (GSE29746) [31]. Agilent one-color expression data were examined using LIMMA [24]. The pre-process stage consisted of background correction, followed by normalization. Thus, the applied background correction is a convolution of normal and exponential distributions that are fitted to the foreground intensities using the background intensities as a covariate, as explained in the LIMMA manual. Next, a well-known quantile method was performed to normalize the green channel between the arrays and then the green channel intensity values were log2-transformed. Values of average replicate spots were analyzed with a Bayes-moderated t-test. Expression genes matching methylated genes were then studied. Genes differentially expressed between RASF and OASF groups were selected if they met the criteria of having values of p and FDR (False Discovery Rate) lower than 0.05 as calculated by Benjamini-Hochberg and a greater than two-fold or less than 0.5-fold change in expression. Expression data were validated by quantitative RT-PCR. Primer sequences are listed in Supplementary Table 1.

2.7. microRNA expression screening, target prediction and integration with DNA methylation data

Total RNA was extracted with TriPure (Roche, Switzerland) following the manufacturer's instructions. Ready-to-use microRNA PCR Human Panel I and II V2.R from Exiqon (Reference 203608) were used according to the instruction manual (Exiqon). For each RT-PCR reaction 30 ng of total RNA was used. Samples from OASF and RASF patients were pooled and two replicates of each group were analyzed on a Roche LightCycler® 480 real-time PCR system. Results were converted to relative values using the inter-plate calibrators included in the panels (log 2 ratios). RASF and OASF average expression values were normalized with respect to reference gene miR-103. Differentially expressed microRNAs (FC > 2 or < 0.5) were selected.

To predict the potential targets of the dysregulated microRNAs, we used the algorithms of several databases, specifically TargetScan

[32], PicTar [33], PITA [34], miRBase [35], microRNA.org [36], miRDB/MirTarget2 [37], TarBase [38], and miRecords [39], StarBase/CLIPseq [40]. Only targets predicted in at least four of these databases and differentially expressed between RASFs and OASFs were included in the heatmaps.

To compare the DNA methylation bead array data with the miRNA expression levels, miRNAs were mapped to Illumina 450k probes. For each differentially expressed miRNA we studied the CpGs within a 5000 bp window around the transcription start site. Using the GRCh37 assembly annotation for Illumina, the genomic localization of probes was extracted in order to match them with miRNA loci. Genomic features of miRNAs were taken from the miRBase [41] and Illumina annotation was obtained from IlluminaHumanMethylation450K.db Bioconductor Package [41].

2.8. Gene ontology analysis

Gene Ontology analysis was done with the FatiGO tool [42], which uses Fisher's exact test to detect significant over-representation of GO terms in one of the sets (list of selected genes) with respect to the other one (the rest of the genome). Multiple test correction to account for the multiple hypothesis tested (one for each GO term) is applied to reduce false positives. GO terms with adjusted *P*-value < 0.05 are considered significant.

2.9. Graphics and heatmaps

All graphs were created using Prism5 Graphpad. Heatmaps were generated from the expression or methylation data using the Genesis program from Graz University of Technology [43].

3. Results

3.1. Comparison of DNA methylation patterns between RASF and OASF reveals both hypomethylation and hypermethylation of key genes

We performed high-throughput DNA methylation screening to compare SF samples from six RA and six OA patients. To this end, we used a methylation bead array that allows the interrogation of >450,000 CpG sites across the entire genome covering 99% of RefSeq genes. Statistical analysis of the combined data from the 12 samples showed that 2571 CpG sites, associated with 1240 different genes, had significant differences in DNA methylation between RASFs and OASFs (median β differences > 0.10, $p < 0.05$) (Fig. 2A and Supplementary Table 2). Specifically, we found 1091 hypomethylated CpG sites (in 575 genes) and 1479 hypermethylated CpG sites (in 714 genes).

The list of genes differentially methylated between RASFs and OASFs includes a number with known implications for RA pathogenesis and some potentially interesting novel genes (Table 1). One of the best examples is *IL6R*. Our results indicated that *IL6R* is hypomethylated in RASFs with respect to OASFs, and that hypomethylation is probably associated with *IL6R* overexpression in RASFs. *IL6* and *IL6R* are factors well known to be associated with RA pathogenesis and progression. *IL6R* overexpression plays a key role in acute and chronic inflammation and increases the risk of joint destruction in RA. Also, IL6R antibodies have recently been approved for the treatment of RA [44]. Another interesting example in the hypomethylated gene list is *TNFAIP8*, or *TIPE2*, a negative mediator of apoptosis that plays a role in inflammation [45]. We also identified *CAPN8* as the gene with the greatest difference between RASFs and OASFs. This gene has not previously been associated with RA, although it is involved in other inflammatory processes such as irritable bowel syndrome [46]. Conversely, hypermethylated genes include factors like *DPP4* and *CCR6*. *DPP4* encodes a serine protease, which cleaves a number of regulatory factors, including chemokines and growth factors. DPP4 inhibitors have recently emerged as novel pharmacological agents for inflammatory disease [47]. Several lines of evidence have also shown a role for *CCR6* in RA [48].

We then set out to determine whether our differentially methylated genes could be involved in biological functions relevant to RA pathogenesis. We therefore performed Gene Ontology analysis to test whether some molecular functions or biological processes were significantly associated with the genes with the greatest difference in DNA methylation status between RASFs and OASFs. The analysis was performed independently for gene lists in the hypomethylated and hypermethylated group. We observed significantly enriched functional processes that are potentially relevant in the biology of SFs (Fig. 2B), including the following categories: focal adhesion assembly (GO:0048041), cartilage development (GO:0051216) and regulation of cell growth (GO:0001558) for hypomethylated genes. For hypermethylated genes, we observed enrichment in categories such as response to wounding (GO:0009611), cell migration (GO:0016477) and cell adhesion (GO:0007155). Hypermethylated and hypomethylated genes shared several functional categories, such as cell differentiation (GO:0030154), cell adhesion (GO:0007155) and skeletal system development (GO:0001501) characteristic of this cell type.

We also compared our data with those reported in a recent study by Nakano and colleagues [17]. We found a significant overlap of genes that were hypomethylated and hypermethylated in both sets of samples (Supp. Fig. 1). These included genes like *MMP20*, *RASGRF2* and *TRAF2* from the list of hypomethylated genes, and *ADAMTS2*, *EGF* and *TIMP2* from among the hypermethylated genes (see Supp. Fig. 1 and Table 2 in [17]). The use of a limited set of samples in the identification of genes introduces a bias associated with each particular sample cohort, which would explain the partial overlap between different experiments. However in this case, we observed an excellent overlap between both experiments.

We also performed an analysis to identify genomic clusters of differentially methylated CpGs, which highlighted several regions of consecutive CpGs that are hypomethylated or hypermethylated in RASFs compared with OASFs. Among hypermethylated CpG clusters in RASFs we identified *TMEM51* and *PTPRN2*. With respect to hypomethylated genes, up to nine clustered CpGs were hypomethylated around the transcription start sites of *HOXA11* (Fig. 2C, left) and nine in *CD74*, the major histocompatibility complex, class II invariant chain-encoding gene. CD74 levels have been reported to be higher in synovial tissue samples from patients with RA than in tissue from patients with osteoarthritis [49]. *HOXA11* was considered another interesting gene, as HOX genes are a direct target of EZH2, a Polycomb group protein involved in differentiation and in establishing repressive marks, including histone H3K27me3 and DNA methylation, under normal and pathological conditions. In fact, additional HOXA genes were identified as being differentially methylated between RASFs and OASFs (Table 1), suggesting that the Polycomb group differentiation pathway may be responsible for these differences.

To validate our analysis, we used bisulfite pyrosequencing of selected genes (Supp. Fig. 2). In all cases, pyrosequencing of individual genes confirmed the results of the analysis. In fact, comparison of the bead array and pyrosequencing methylation data (Fig. 2C, center and right) showed an excellent correlation, supporting the validity of our analysis. Additional genes that were subjected to pyrosequencing analysis included *CAPN8* and *IL6R*, both of which were hypomethylated in RASFs, and *DPP4* and *HOXC4*

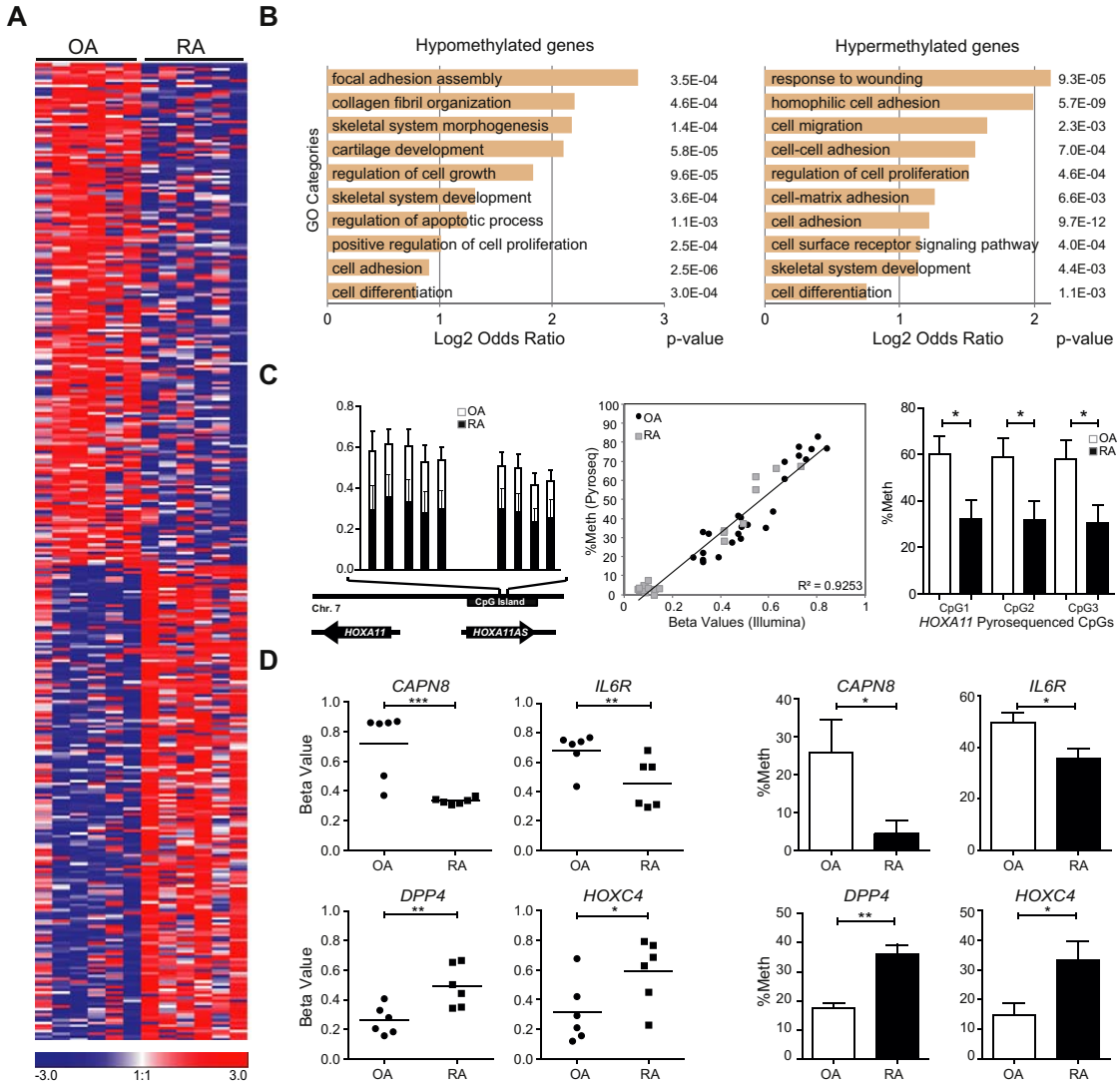


Fig. 2. Comparison of the DNA methylation profiles between RASFs and OASFs samples. (A) Heatmap including the methylation data for the six RASF and OASF samples shows significant differential methylation. There are both significant hypermethylated and hypomethylated genes. In this heatmap, all the genes with a value of $p < 0.05$ and a difference in median $\beta > 0.2$ are shown. The scale at the bottom distinguishes hypermethylated (red) and hypomethylated (blue) genes. (B) Summary of the gene ontology (GO) analysis for the category "biological process" among hypomethylated and hypermethylated genes. *P*-values are shown on the right (C) Methylation data from the array analysis corresponding to HOXA11 gene in which 9 consecutive CpGs are hypomethylated in RASFs relative to OASFs (left), comparison of the array data and pyrosequencing, where the excellent correlation between the two sources of data is shown by a regression line (center), methylation values as obtained through pyrosequencing corresponding to three selected CpGs comparing RASF and OASF samples. (D) Comparison of the array data (left) and pyrosequencing data (right) of four selected hypomethylated (*CAPN8*, *IL6R*) and hypermethylated (*DPP4*, *HOXC4*) genes.

in the hypermethylated group (Fig. 2D and E). In all cases, the analysis was validated by pyrosequencing in a larger cohort of samples.

3.2. Integration of DNA methylation data with expression data from RASFs and OASFs

DNA methylation is generally associated with gene repression, particularly when it occurs at promoter CpG islands. However, DNA methylation changes at promoters with low CpG density can also regulate transcription, and changes in gene bodies also affect transcriptional activity [4], although they do not necessarily repress it. We therefore integrated our DNA methylation data with a high-throughput expression analysis of RASFs and OASFs from a recent study (GSE29746) [31]. To integrate expression data with our methylation results, we first reanalyzed the expression data as described in the Materials and Methods. Applying the threshold criterion of a value of $p < 0.01$, we identified 3470 probes differentially expressed between RASFs and OASFs (for which FC > 2 or <0.6) (Fig. 3A). We then compared the results from the analysis of the expression arrays with the DNA methylation data. Our analysis showed that 208 annotated CpGs displayed an inverse correlation between expression and methylation levels (Fig. 3B and Supplementary Table 3).

Table 1
Selection of genes differentially methylated and/or expressed in RASF vs. OASF, and previously described implications in RA.

Gene name	Δmeth CpG	Region	Description	Δ beta (RA-OA)	FC express (RA/OA)	Previously reported RA implication (ref)
CAPN8	1	Body	Calpain 8	−0.52	N/A	
SERPINA5	1	TSS1500	Serpin peptidase inhibitor, clade A member 5	−0.40	N/A	
FCGBP	1	Body	Fc fragment of IgG binding protein	−0.35	0.34	Detected in plasma sera related with autoimmunity [56]
HOXA11	13	TSS1500	Homeobox A11	−0.30	0.40	
IL6R	1	Body	Interleukin 6 receptor	−0.29	N/A	Its ligand (IL6) is overexpressed in RA [57]
S100A14	3	TSS1500	S100 calcium binding protein A14	−0.27	N/A	Involved in invasion through MMP2 (elevated in RA plasma) [50]
TMEM51	2	5′UTR	Transmembrane protein 51	−0.27	4.21	
CSGALNACT1	3	TSS200	Chondroitin sulfate N-acetylgalactosaminyltransferase 1	−0.22	0.48	Involved in cartilage development and endocondral ossification [58] and [59]
COL14A1	2	Body	Collagen, type XIV, alpha 1	−0.22	3.76	
CD74	8	TSS1500	CD74 molecule	−0.22	N/A	Initiates MIF signal transduction (levels related with RA course) [60]
TNFAIP8	3	Body	Tumor necrosis factor, alpha-induced protein 8	−0.20	3.75	Negative regulator of innate and adaptative immunity [45]
TNFRSF8	1	Body|TSS 1500	Tumor necrosis factor receptor superfamily, member 8	−0.19	2.93	Its overexpression contributes to proinflammatory immune responses [61]
KCNJ15	2	5′UTR|TSS 200	Potassium inwardly-rectifying channel, subfamily J, member 15	−0.15	5.51	
CCR6	1	TSS1500	Chemokine (C-C motif) receptor 6	0.23	0.67	Migration, proliferation, and MMPs production [62]
DPP4	1	TSS200	Dipeptidyl-peptidase 4	0.23	N/A	Its inhibition increases cartilage invasion by RASF [63]
PRKCZ	17	Body|TSS 1500	Protein kinase C, zeta	0.25	N/A	Inactivates syndecan-4 (integrin co-receptor), reducing DC motility [64]
HLA-DRB5	3	Body	Major histocompatibility complex, class II, DR beta 5	0.26	3.16	SNP associated with cutaneous manifestations rheumatoid vasculitis [65]
ALOX5AP	1	TSS1500	Arachidonate 5-lipoxygenase-activating protein	0.29	N/A	Deficit of this molecule ameliorates symptoms in CIA [66]
BCL6	2	Body	B-cell CLL/lymphoma 6	0.30	N/A	RA synovial T cells express BCL6, potent B cell regulator [67]
SPTBN1	2	TSS1500	Spectrin, beta, non-erythrocytic 1	0.27	0.23	Associated with CD43 abrogates T cell activation [68]
HOXC4	13	5′UTR|TSS 1500	Homeobox C4	0.4	N/A	

To examine the relationship between methylation and gene expression further, we also performed an analysis focusing on the relative position of the CpG site that undergoes a significant change in methylation. We found that genes with a methylation change at the TSS or the 5′UTR generally exhibited an inverse correlation between DNA methylation and gene expression (Fig. 3C), whereby an increase in methylation tended to be accompanied by a decrease in expression. Curiously, this relationship is positive when looking at CpGs containing probes located at gene bodies with a significant methylation change (Fig. 3C). Fig. 3D shows two examples of an inverse correlation between DNA methylation and expression data.

We performed quantitative RT-PCR to investigate the expression status of several of the genes displaying a change in DNA methylation in the set of samples used in this study. This analysis included several of the genes mentioned above, as well as others, like *MMP2*, for which increased expression in RASFs has been described [50]. Our analysis confirmed the elevated levels for this gene in our collection of RASFs (Fig. 3E). We also observed an inverse correlation between DNA methylation and expression for genes like *HOXC4*, *HOXA11*, *CAPN8* and *IL6R* (Fig. 3F), although genes like *DPP4* did show a direct relationship. Specifically, we found that hypermethylated *DPP4* had higher levels of expression in RASFs than in OASFs (Fig. 3F). Elevated levels for *DPP4* are compatible with the data obtained by other researchers [51]. However, it also indicates that for some genes, other mechanisms contribute more to their expression levels than do DNA methylation changes.

3.3. miRNA screening in RASF and OASF

Changes in expression levels can certainly be due to transcriptional control, like that determined by epigenetic changes at gene promoters, DNA methylation, or differences in transcription factor binding. At the post-transcriptional level, miRNAs are recognized as being major players in gene expression regulation. We compared the expression levels of miRNAs in pooled RASF and OASF RNA samples. The screening led to the identification of a number of miRNAs that are overexpressed in RASFs with respect to OASFs, as well as downregulated miRNAs (Fig. 4A). Among the most upregulated and downregulated miRNAs, we identified several that have been previously associated with relevant or related events like miR-203 [20], which is upregulated in RASFs with respect to OASFs, and miR-124, which is downregulated in RASFs [52] (Fig. 4B). Other additional miRNAs identified in previous work in relation to RA include miR-146a and miR-34a (Fig. 4A). As indicated above, the overlap with other studies highlights the robustness of our data. However, it is likely that the analysis of a limited number of samples introduces a bias associated with specific characteristics of the samples in the studied cohort. Integrated analysis can help to identify relevant targets. We then performed quantitative PCR to validate a selection of the miRNAs in the entire cohort. Examples of miR-625*, downregulated in RASF, and miR-551b, upregulated in RASF, are shown in Fig. 4C.

As explained, miRNA-dependent control is associated with the expression control of a number of targets either by inducing direct mRNA degradation or through translational inhibition [53]. Accumulated evidence has shown that most miRNA targets are affected at the mRNA levels, and therefore comparison of mRNA expression array and

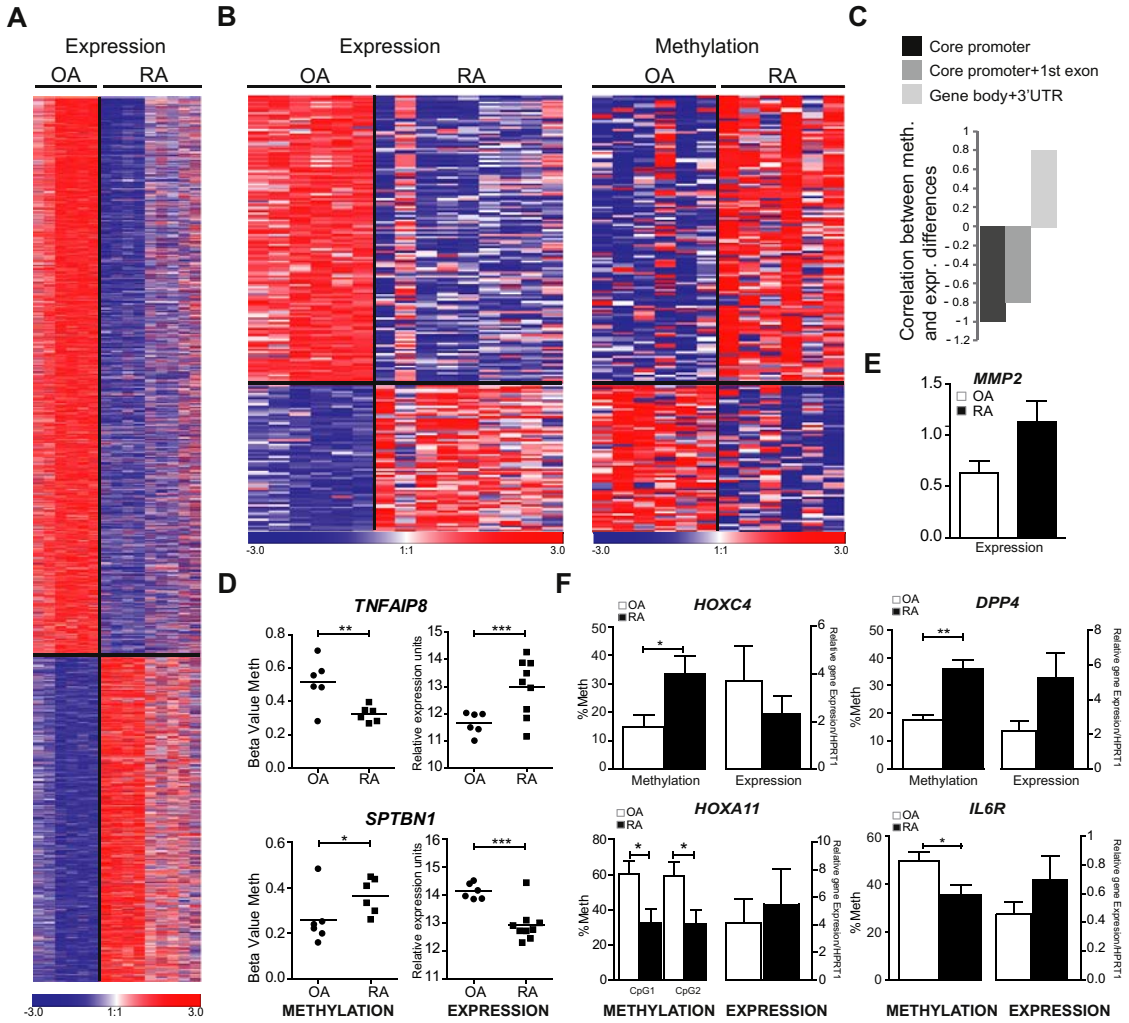


Fig. 3. Integration of DNA methylation with expression data. (A) Heatmaps including the reanalysis of expression data (GSE29746) for a collection of RASF and OASF samples. The scale at the bottom distinguishes upregulated (red) and downregulated (blue) genes. (B) Heatmap comparison of inversely correlated expression and methylation. (C) Correlation between differences in DNA methylation and expression of associated genes, focusing on different regions where the CpG sites are located in the probe, core promoter, core + 1st exon and gene body. (D) Illustrative examples of genes featuring an inverse correlation between methylation and expression. (E) Validation by quantitative RT-PCR of MMP2, previously described as overexpressed in RASFs. (F) Examples of genes whose expression data were validated by quantitative RT-PCR.

miRNA expression data is useful for identifying and evaluating the impact of miRNA misregulation at the mRNA levels.

To explore this aspect, we investigated the relationship between miRNA expression differences between RASFs and OASFs and their involvement in gene control by looking at levels of their potential targets. To this end, we obtained a matrix with the potential targets for each of the five miRNAs most strongly upregulated and downregulated in RASFs relative to OASFs. We considered *bona fide* putative targets those predicted by at least four databases. As before, we used the expression microarrays data for RASFs and OASFs generated in another study (GSE29746) [31].

When looking at the expression levels of putative targets of selected miRNAs we found more genes with potential effects on the RASF phenotype. These included genes like *CTSC*, *KLF8* or *EBF3*, which are upregulated in RASFs concomitant with downregulation of miR625* and *ITGBL1*, which is downregulated in RASF concomitant with upregulation of miR551b. Additional putative targets included TLR4 for miR-203 and NFAT5 for miR-124 (Fig. 4D). *TLR4* is upregulated in RA and plays a key role in the disease, whereas NFAT5 is a critical regulator of inflammatory arthritis.

3.4. Integrated analysis of both miRNAs and DNA methylation reveals multiple layers of regulation in genes relevant to RA pathogenesis

We performed two separate analyses to explore the potential connection between miRNA and DNA methylation control for genes associated with the RASF phenotype.

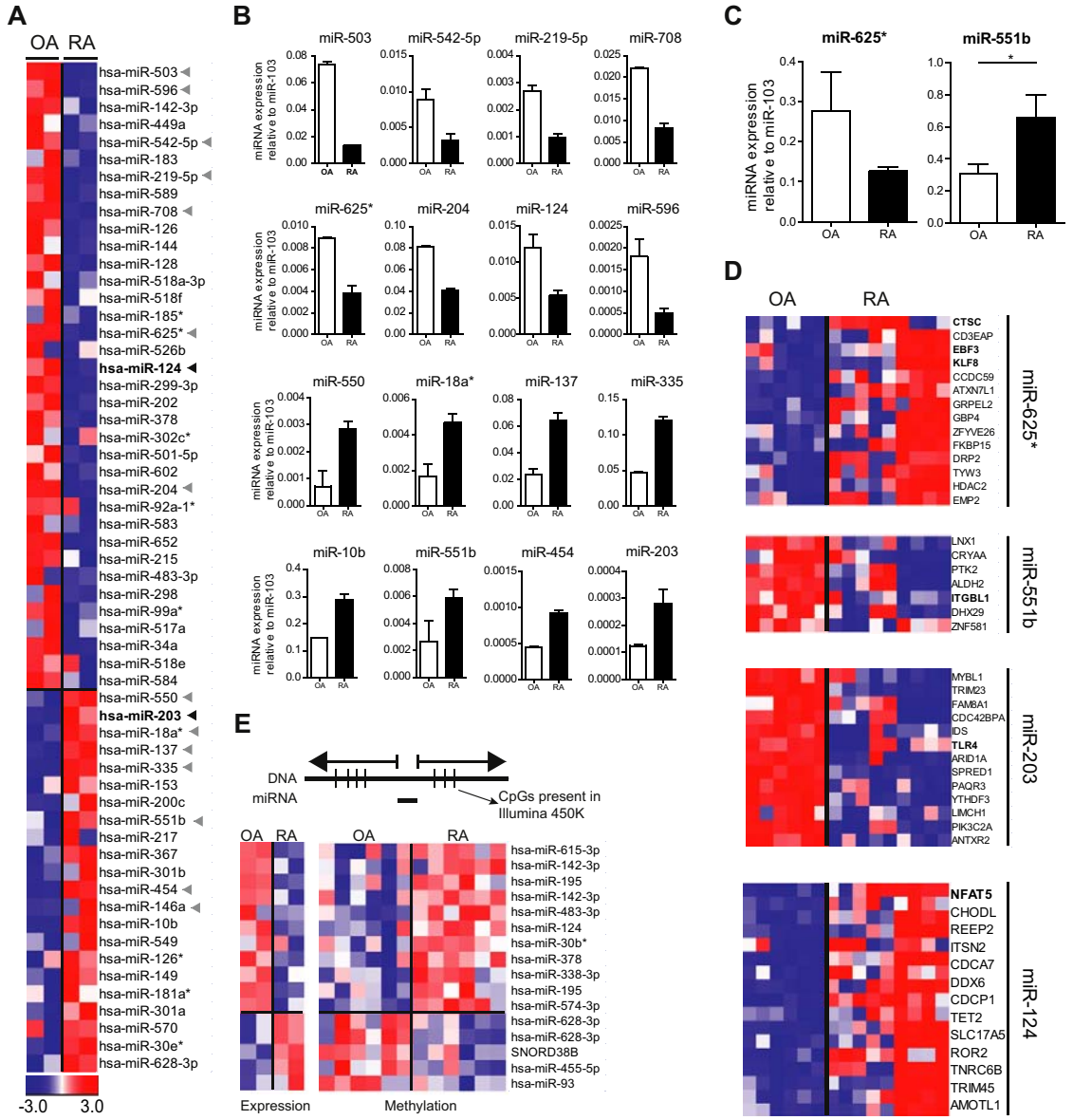


Fig. 4. miRNA dysregulation in RASF. (A) Heatmaps showing the miRNA expression data for pooled RASFs and OASFs. miRNAs previously described as dysregulated in RASFs are highlighted with a black arrow. Those represented in the adjacent section are highlighted with a grey arrow. The scale at the bottom distinguishes distinguishing upregulated (red) and downregulated (blue) genes. (B) Examples of the most upregulated and downregulated miRNAs in RASFs with respect to OASFs. (C) Validation of the miRNA data by quantitative RT-PCR. (D) Heatmaps showing the expression levels of putative targets (at least four hits for prediction in miRNA databases) for selected miRNAs in RASFs and OASFs. (E) Correlation between DNA methylation and miRNA expression data.

The first analysis focused on the potential regulation of miRNAs by DNA methylation. DNA methylation can also repress the expression of miRNAs, since miRNA-associated promoters are subjected to similar mechanisms of transcriptional control as protein-coding genes. We compared the data from the bead array analysis with miRNA expression data. Our analysis showed 11 downregulated miRNAs, like miR-124, that were located near CpG sites and were hypermethylated in RASFs. Only four upregulated miRNAs were located near a CpG site hypomethylated in RASFs (Fig. 4E).

The second analysis investigated the potential influence of DNA methylation and miRNA control on specific targets. As explained above, differences in expression patterns between RASFs and OASFs could be due to altered mechanisms of control at the epigenetic level, like DNA methylation, or at the post-transcriptional level. We generated a list of selected genes whose expression patterns differed significantly between RASFs and OASFs. Then we matched the expression data with our DNA methylation data from bead arrays and with a selection of miRNAs that might target those genes (as predicted at least by four databases) and that have significant

differences in expression between RASFs and OASFs. This yielded a list of genes potentially regulated by DNA methylation, at the transcriptional level, and targeted by miRNAs, at the post-transcriptional level. The list of genes comprises six groups (Supplementary Table 4): i) downregulated genes in which hypermethylation concurs with overexpression of a miRNA that targets them. Methylation and miRNA regulate in the same repressive direction in this group; ii) downregulated genes in which hypomethylation is potentially overcome by co-occurrence of upregulation of a miRNA that targets them. In this group, miRNA control is potentially the predominant mechanism; iii) downregulated genes in which hypermethylation predominates over downregulation of miRNAs that potentially target them; iv) upregulated genes in which hypomethylation concurs with downregulation of a miRNA that targets them; v) upregulated genes in which hypermethylation is overcome by downregulation of a miRNA that targets them, and, vi) upregulated genes in which hypermethylation predominates over upregulation of a miRNA that targets them. Integrated analysis would require further validation to provide *bona fide* targets determined by both regulatory mechanisms. However, this novel approach to integrating miRNA and DNA methylation analysis provides a new workflow for exploring the multiple layers of gene dysregulation in RA in greater depth.

4. Discussion

In this study we have identified novel dysregulated targets in rheumatoid arthritis (RA) synovial fibroblasts at the DNA methylation and miRNA expression levels. By using a double approach and integrated analysis of the DNA methylation, miRNA expression and mRNA expression data we have established a new pipeline for investigating the complexity of gene dysregulation in the context of this disease when using primary samples. As indicated above, dysregulation of gene expression arises from a combination of factors, including genetic polymorphisms in genes associated with regulatory roles and miRNAs, environmental factors and their combined effect on transcription factor function and epigenetic profiles, like DNA methylation and histone modification profiles. Understanding the relationship between different elements of regulation is key not only for understanding their intricate connections within the disease but also in the higher propensity to associated disorders [54]. DNA methylation-associated regulation and miRNA control are major regulatory elements and provide useful targets and markers of gene dysregulation in disease. In the context of RASFs, a few studies have previously shown the existence of genes with DNA methylation alterations in RASFs. Most of these have involved examining candidate genes. Examples include the identification of the *TNFRSF25* gene (encoding DDR9), which is hypermethylated at its CpG island in synovial cells of RA patients [16], and *CXCL12* upregulated and hypomethylated in RASFs [15]. More recently, Firestein and colleagues [17] took an array-based approach to identify hypomethylated and hypermethylated genes in RASFs. Regarding miRNA profiling in RASFs, several studies have demonstrated specific roles for miRNAs that are dysregulated in RA synovial tissues [18–20,55]. However, there were no previous systematic efforts to combine analyses of these two types of mechanisms in the context of RA.

To the best of our knowledge, our study constitutes the first attempt to integrate high-throughput omics data from primary samples in the context of RA. The need of integrating several levels of regulation is relevant for several reasons: first, from a biological point of view, it is essential to understand the molecular mechanisms underlying aberrant changes in gene expression associated with the acquisition of the aggressive phenotype of RASFs; second, from a more translational point of view, understanding multiple levels of regulation of target genes that undergo dysregulation in RA, could potentially allow to predict their behavior following the use of specific therapeutic compounds. It can also serve to make a better use of them as clinical markers of disease onset, progression or response to therapy.

Our analysis of individual datasets not only has allowed us to confirm changes described by others but also to determine novel genes with altered DNA methylation patterns, including *MMP20*, *RASGRF2*, *EGF*, *TIMP2* and others. Most importantly we have identified new genes that are relevant to the RA phenotype. This includes *IL6R*, which is well known as an overexpressed gene in RASFs and a target for antibody-based therapy [44]. Additional targets include *CAPN8*, *TNFAIP8*, *CD74* and *CCR6*. Methylation alterations in RASFs occur at promoter CpG islands in genes like *DPP4* or *HOXC4*, and downstream of the TSS in genes like *CAPN8* and *IL6R*. This last observation is in agreement with recent reports showing that gene expression can be also affected by methylation changes at gene bodies [4,5]. In any case, we have found a canonical inverse relationship between DNA methylation and expression status for a subset of more than 200 genes. At the miRNA level, analysis of the expression dataset has allowed us to validate previously described miRNAs, like miR-203 and miR-124, as well as identifying novel miRNAs, like miR-503, miR-625*, miR-551b, and miR-550, that are potentially associated with dysregulated targets in RASFs.

Integrative analysis has been carried out at different levels. Firstly, the combined analysis of DNA methylation and expression data generated a list of genes in which methylation changes were inversely correlated with expression changes. This list of genes potentially contains those regulated through DNA methylation in a canonical manner, where DNA methylation associates with gene repression (Supplementary Table 3). Secondly, we also studied the potential relationship between expression changes and miRNA expression changes that potentially target them (as defined by the cumulative use of miRNA target prediction databases). In this case, we identified a number of genes undergoing expression changes in RASFs that are potentially targeted by concomitantly dysregulated miRNAs.

Another level of integration is achieved by looking at genes that may be targeted or regulated by the combined action of miRNA and DNA methylation. Thus, we explored the potential combined effect of miRNAs and DNA methylation in genes undergoing expression changes in RASFs (Supplementary Table 2). Our analysis revealed gene targets in which methylation and miRNA control possibly concur in direction or have antagonistic effects. This classification of genes in different groups is important because pharmacological compounds or other experimental approaches influencing one of the mechanism (DNA methylation) but not the other (miRNA expression) or viceversa, would have to consider the existence of multiple levels of regulation for interpreting the outcome of such treatment. Finally, by looking at the potential control of miRNA expression by DNA methylation, we identified a further regulatory mechanism for several miRNAs, including miR-124. In this case, hypermethylation of a specific miRNA promoter, would have a positive effect on the expression levels of its targets, and, for instance, pharmacological demethylation of that miRNA would result in overexpression of the miRNA and downregulation of its targets.

As indicated above, epigenetic profiles and miRNA expression patterns are cell type-specific. The need to use primary samples for the target tissue or cell type of a particular disease is usually a limitation to performing epigenetic or miRNA analysis, given the access to small amounts of tissue or cells that can be obtained in most cases. The reduced number of laboratories with access to RASFs, OASFs or SF from normal individuals is a good reflection of

such limitation. Genetic analysis of genetically complex diseases does not have such a limitation, since in most cases can be done with peripheral blood. In this sense, the use of integrative approaches to investigate epigenetic and miRNA-mediated control of a limited set of samples overcomes partially this obstacle by providing extra sets of data for internal validation within a small cohort of samples and an increase of the robustness of the analysis.

In conclusion, our study highlights the need of investigating the multiple layers of regulation at the transcriptional and post-transcriptional levels as well as integrating the datasets during the analysis. As targets for therapy, it is important to understand the intricate connections between the various control mechanisms and to consider the existence of both processes that operate in the same direction or have antagonistic effects. The use of integrative approaches will also be necessary for the rational design of targeted therapies as well as for the use of different clinical markers for the classification. In this sense, the workflow designed in this study has allowed us to identify novel targets and their regulatory mechanism in RASF and opens up a number of possibilities for future research on epigenetics aspects on RA.

Acknowledgments

We would like to thank Dr. Gary Firestein for sharing the raw data of his DNA methylation study with us. We would also like to thank José Luis Pablos for his valuable feedback on his expression dataset. This work was supported by grant SAF2011-29635 from the Spanish Ministry of Science and Innovation, grant from Fundación Ramón Areces and grant 2009SGR184 from AGAUR (Catalan Government). LR is supported by a PFIS predoctoral fellowship and AI was supported by a AGAUR predoctoral fellowship. NL-B acknowledges funding from the Spanish Ministry of Science and Technology (grant number SAF2009-06954)

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.jaut.2012.12.005.

References

[1] Lefevre S, Knedla A, Tennie C, Kampmann A, Wunrau C, Dinser R, et al. Synovial fibroblasts spread rheumatoid arthritis to unaffected joints. Nat Med 2009;15:1414–20.
[2] Tolboom TC, van der Helm-Van Mil AH, Nelissen RG, Breedveld FC, Toes RE, Huizinga TW. Invasiveness of fibroblast-like synoviocytes is an individual patient characteristic associated with the rate of joint destruction in patients with rheumatoid arthritis. Arthritis Rheum 2005;52:1999–2002.
[3] Deaton AM, Bird A. CpG islands and the regulation of transcription. Genes Dev 2011;25:1010–22.
[4] Ball MP, Li JB, Gao Y, Lee JH, LeProust EM, Park IH, et al. Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells. Nat Biotechnol 2009;27:361–8.
[5] Rauch TA, Wu X, Zhong X, Riggs AD, Pfeifer GP. A human B cell methylome at 100-base pair resolution. Proc Natl Acad Sci U S A 2009;106:671–8.
[6] Tili E, Michaille JJ, Costinean S, Croce CM. MicroRNAs, the immune system and rheumatic disease. Nat Clin Pract Rheumatol 2008;4:534–41.
[7] Selmi C, Lu Q, Humble MC. Heritability versus the role of the environment in autoimmunity. J Autoimmun 2012;39:249–52.
[8] Ballestar E. Epigenetics lessons from twins: prospects for autoimmune disease. Clin Rev Allergy Immunol 2010;39:30–41.
[9] Miller FW, Pollard KM, Parks CG, Germolec DR, Leung PS, Selmi C, et al. Criteria for environmentally associated autoimmune diseases. J Autoimmun 2012;39:253–8.
[10] Miller FW, Alfredsson L, Costenbader KH, Kamen DL, Nelson LM, Norris JM, et al. Epidemiology of environmental exposures and human autoimmune diseases: findings from a National Institute of Environmental Health Sciences Expert Panel Workshop. J Autoimmun 2012;39:259–71.
[11] Selmi C, Leung PS, Sherr DH, Diaz M, Nyland JF, Monestier M, et al. Mechanisms of environmental influence on human autoimmunity: a national institute of environmental health sciences expert panel workshop. J Autoimmun 2012;39:272–84.
[12] Germolec D, Kono DH, Pfau JC, Pollard KM. Animal models used to examine the role of the environment in the development of autoimmune disease: findings from an NIEHS Expert Panel Workshop. J Autoimmun 2012;39:285–93.
[13] Neidhart M, Rethage J, Kuchen S, Kunzler P, Crowl RM, Billingham ME, et al. Retrotransposable L1 elements expressed in rheumatoid arthritis synovial tissue: association with genomic DNA hypomethylation and influence on gene expression. Arthritis Rheum 2000;43:2634–47.
[14] Nile CJ, Read RC, Akil M, Duff GW, Wilson AG. Methylation status of a single CpG site in the IL6 promoter is related to IL6 messenger RNA levels and rheumatoid arthritis. Arthritis Rheum 2008;58:2686–93.
[15] Karouzakis E, Rengel Y, Jungel A, Kolling C, Gay RE, Michel BA, et al. DNA methylation regulates the expression of CXCL12 in rheumatoid arthritis synovial fibroblasts. Genes Immun 2011;12:643–52.
[16] Takami N, Osawa K, Miura Y, Komai K, Taniguchi M, Shiraishi M, et al. Hypermethylated promoter region of DR3, the death receptor 3 gene, in rheumatoid arthritis synovial cells. Arthritis Rheum 2006;54:779–87.
[17] Nakano K, Whitaker JW, Boyle DL, Wang W, Firestein GS. DNA methylome signature in rheumatoid arthritis. Ann Rheum Dis 2012.
[18] Niederer F, Trenkmann M, Ospelt C, Karouzakis E, Neidhart M, Stanczyk J, et al. Down-regulation of microRNA-34a* in rheumatoid arthritis synovial fibroblasts promotes apoptosis resistance. Arthritis Rheum 2012;64:1771–9.
[19] Nakamachi Y, Kawano S, Takenokuchi M, Nishimura K, Sakai Y, Chin T, et al. MicroRNA-124a is a key regulator of proliferation and monocyte chemoattractant protein 1 secretion in fibroblast-like synoviocytes from patients with rheumatoid arthritis. Arthritis Rheum 2009;60:1294–304.
[20] Stanczyk J, Ospelt C, Karouzakis E, Filer A, Raza K, Kolling C, et al. Altered expression of microRNA-203 in rheumatoid arthritis synovial fibroblasts and its role in fibroblast activation. Arthritis Rheum 2011;63:373–81.
[21] Bibikova M, Lin Z, Zhou L, Chudin E, Garcia EW, Wu B, et al. High-throughput DNA methylation profiling using universal bead arrays. Genome Res 2006;16:383–93.
[22] Dessau RB, Pipper CB. "R"–project for statistical computing. Ugeskr Laeger 2008;170:328–30.
[23] Du P, Kibbe WA, Lin SM. Lumi: a pipeline for processing illumina microarray. Bioinformatics 2008;24:1547–8.
[24] Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, et al. Bioconductor: open software development for computational biology and bioinformatics. Genome Biol 2004;5:R80.
[25] Falcon S, Gentleman R. Using GOstats to test gene lists for GO term association. Bioinformatics 2007;23:257–8.
[26] Smyth GK. Limma: linear models for microarray data. In: Bioinformatics and computational biology solutions using R and bioconductor; 2005. p. 397–420.
[27] Maksimovic J, Gordon L, Oshlack A. SWAN: subset-quantile within array normalization for illumina infinium HumanMethylation450 BeadChips. Genome Biol 2012;13:R44.
[28] Touleimat N, Tost J. Complete pipeline for infinium((R)) Human Methylation 450K BeadChip data processing using subset quantile normalization for accurate DNA methylation estimation. Epigenomics 2012;4:325–41.
[29] Aryee MJ, Wu Z, Ladd-Acosta C, Herb B, Feinberg AP, Yegnasubramanian S, et al. Accurate genome-scale percentage DNA methylation estimates from microarray data. Biostatistics 2011;12:197–210.
[30] Herman JG, Graff JR, Myohanen S, Nelkin BD, Baylin SB. Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands. Proc Natl Acad Sci U S A 1996;93:9821–6.
[31] Del Rey MJ, Usategui A, Izquierdo E, Canete JD, Blanco FJ, Criado G, et al. Transcriptome analysis reveals specific changes in osteoarthritis synovial fibroblasts. Ann Rheum Dis 2012;71:275–80.
[32] Lewis BP, Burge CB, Bartel DP. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell 2005;120:15–20.
[33] Krek A, Grun D, Poy MN, Wolf R, Rosenberg L, Epstein EJ, et al. Combinatorial microRNA target predictions. Nat Genet 2005;37:495–500.
[34] Kertesz M, Iovino N, Unnerstall U, Gaul U, Segal E. The role of site accessibility in microRNA target recognition. Nat Genet 2007;39:1278–84.
[35] Griffiths-Jones S, Grocock RJ, van Dongen S, Bateman A, Enright AJ. miRBase: microRNA sequences, targets and gene nomenclature. Nucleic Acids Res 2006;34:D140–4.
[36] Betel D, Wilson M, Gabow A, Marks DS, Sander C. The microRNA.org resource: targets and expression. Nucleic Acids Res 2008;36:D149–53.
[37] Wang X, El Naqa IM. Prediction of both conserved and nonconserved microRNA targets in animals. Bioinformatics 2008;24:325–32.
[38] Vergoulis T, Vlachos IS, Alexiou P, Georgakilas G, Maragkakis M, Reczko M, et al. TarBase 6.0: capturing the exponential growth of miRNA targets with experimental support. Nucleic Acids Res 2011;40:D222–9.
[39] Xiao F, Zuo Z, Cai G, Kang S, Gao X, Li T. miRecords: an integrated resource for microRNA-target interactions. Nucleic Acids Res 2009;37:D105–10.
[40] Yang JH, Li JH, Shao P, Zhou H, Chen YQ, Qu LH. starBase: a database for exploring microRNA-mRNA interaction maps from Argonaute CLIP-Seq and Degradome-Seq data. Nucleic Acids Res 2011;39:D202–9.
[41] Kozomara A, Griffiths-Jones S. miRBase: integrating microRNA annotation and deep-sequencing data. Nucleic Acids Res 2011;39:D152–7.
[42] Al-Shahrour F, Diaz-Uriarte R, Dopazo J. FatiGO: a web tool for finding significant associations of Gene Ontology terms with groups of genes. Bioinformatics 2004;20:578–80.

[43] Sturn A, Quackenbush J, Trajanoski Z. Genesis: cluster analysis of microarray data. Bioinformatics 2002;18:207–8.
[44] Thompson CA. FDA approves tocilizumab to treat rheumatoid arthritis. Am J Health Syst Pharm 2010;67:254.
[45] Sun H, Gong S, Carmody RJ, Hilliard A, Li L, Sun J, et al. TIPE2, a negative regulator of innate and adaptive immunity that maintains immune homeostasis. Cell 2008;133:415–26.
[46] Swan C, Duroudier NP, Campbell E, Zaitoun A, Hastings M, Dukes GE, et al. Identifying and testing candidate genetic polymorphisms in the irritable bowel syndrome (IBS): association with TNFSF15 and TNFalpha. Gut 2012.
[47] Yazbeck R, Howarth GS, Abbott CA. Dipeptidyl peptidase inhibitors, an emerging drug class for inflammatory disease? Trends Pharmacol Sci 2009; 30:600–7.
[48] Ruth JH, Shahrara S, Park CC, Morel JC, Kumar P, Qin S, et al. Role of macrophage inflammatory protein-3alpha and its ligand CCR6 in rheumatoid arthritis. Lab Invest 2003;83:579–88.
[49] Waldburger JM, Palmer G, Seemayer C, Lamacchia C, Finckh A, Christofilopoulos P, et al. Autoimmunity and inflammation are independent of class II transactivator type PIV-dependent class II major histocompatibility complex expression in peripheral tissues during collagen-induced arthritis. Arthritis Rheum 2011;63:3354–63.
[50] Li G, Zhang Y, Qian Y, Zhang H, Guo S, Sunagawa M, et al. Interleukin-17A promotes rheumatoid arthritis synoviocytes migration and invasion under hypoxia by increasing MMP2 and MMP9 expression through NF-kappaB/HIF-1alpha pathway. Mol Immunol 2012;53:227–36.
[51] Solau-Gervais E, Zerimech F, Lemaire R, Fontaine C, Huet G, Flipo RM. Cysteine and serine proteases of synovial tissue in rheumatoid arthritis and osteoarthritis. Scand J Rheumatol 2007;36:373–7.
[52] Kawano S, Nakamachi Y. miR-124a as a key regulator of proliferation and MCP-1 secretion in synoviocytes from patients with rheumatoid arthritis. Ann Rheum Dis 2011;70(Suppl. 1):i88–91.
[53] He L, Hannon GJ. MicroRNAs: small RNAs with a big role in gene regulation. Nat Rev Genet 2004;5:522–31.
[54] Ngalamika O, Zhang Y, Yin H, Zhao M, Gershwin ME, Lu Q. Epigenetics, autoimmunity and hematologic malignancies: a comprehensive review. J Autoimmun 2012;39:451–65.
[55] Nakasa T, Miyaki S, Okubo A, Hashimoto M, Nishida K, Ochi M, et al. Expression of microRNA-146 in rheumatoid arthritis synovial tissue. Arthritis Rheum 2008;58:1284–92.
[56] Kobayashi K, Yagasaki M, Harada N, Chichibu K, Hibi T, Yoshida T, et al. Detection of Fcgamma binding protein antigen in human sera and its relation with autoimmune diseases. Immunol Lett 2001;79:229–35.
[57] Nishimoto N, Kishimoto T, Yoshizaki K. Anti-interleukin 6 receptor antibody treatment in rheumatic disease. Ann Rheum Dis 2000;59(Suppl. 1):i21–7.
[58] Watanabe Y, Takeuchi K, Higa Onaga S, Sato M, Tsujita M, Abe M, et al. Chondroitin sulfate N-acetylgalactosaminyltransferase-1 is required for normal cartilage development. Biochem J 2010;432:47–55.
[59] Sato T, Kudo T, Ikehara Y, Ogawa H, Hirano T, Kiyohara K, et al. Chondroitin sulfate N-acetylgalactosaminyltransferase 1 is necessary for normal endochondral ossification and aggrecan metabolism. J Biol Chem 2011;286: 5803–12.
[60] Leng L, Metz CN, Fang Y, Xu J, Donnelly S, Baugh J, et al. MIF signal transduction initiated by binding to CD74. J Exp Med 2003;197:1467–76.
[61] Simhadri VL, Hansen HP, Simhadri VR, Reiners KS, Bessler M, Engert A, et al. A novel role for reciprocal CD30-CD30L signaling in the cross-talk between natural killer and dendritic cells. Biol Chem 2012;393:101–6.
[62] Matsui T, Akahoshi T, Namai R, Hashimoto A, Kurihara Y, Rana M, et al. Selective recruitment of CCR6-expressing cells by increased production of MIP-3 alpha in rheumatoid arthritis. Clin Exp Immunol 2001;125:155–61.
[63] Ospelt C, Mertens JC, Jungel A, Brentano F, Maciejewska-Rodriguez H, Huber LC, et al. Inhibition of fibroblast activation protein and dipeptidylpeptidase 4 increases cartilage invasion by rheumatoid arthritis synovial fibroblasts. Arthritis Rheum 2010;62:1224–35.
[64] Buhligen J, Himmel M, Gebhardt C, Simon JC, Ziegler W, Averbeck M. Lysophosphatidylcholine-mediated functional inactivation of syndecan-4 results in decreased adhesion and motility of dendritic cells. J Cell Physiol 2010;225: 905–14.
[65] Nishimura WE, Costallat LT, Fernandes SR, Conde RA, Bertolo MB. Association of HLA-DRB5*01 with protection against cutaneous manifestations of rheumatoid vasculitis in Brazilian patients. Rev Bras Rheumatol 2012;52:366–74.
[66] Griffiths RJ, Smith MA, Roach ML, Stock JL, Stam EJ, Milici AJ, et al. Collagen-induced arthritis is reduced in 5-lipoxygenase-activating protein-deficient mice. J Exp Med 1997;185:1123–9.
[67] Manzo A, Vitolo B, Humby F, Caporali R, Jarrossay D, Dell'accio F, et al. Mature antigen-experienced T helper cells synthesize and secrete the B cell chemoattractant CXCL13 in the inflammatory environment of the rheumatoid joint. Arthritis Rheum 2008;58:3377–87.
[68] Pradhan D, Morrow J. The spectrin-ankyrin skeleton controls CD45 surface display and interleukin-2 production. Immunity 2002;17:303–15.

Genome Biology

Genome Biology

PU.1 target genes undergo Tet2-coupled demethylation and DNMT3b-mediated methylation in monocyte-to-osteoclast differentiation

Genome Biology 2013, **14**:R99 doi:10.1186/gb-2013-14-9-r99

Lorenzo de la Rica (ldelarica@idibell.org)
Javier Rodríguez-Ubreva (jrubreva@idibell.org)
Mireia García (mgperez@idibell.org)
Abul BMMK Islam (khademul.islam@gmail.com)
José M Urquiza (jurquiza@idibell.org)
Henar Hernando (hhernando@idibell.org)
Jesper Christensen (jesper.christensen@bric.ku.dk)
Kristian Helin (kristian.helin@bric.ku.dk)
Carmen Gómez-Vaquero (carmen.gomez@bellvitgehospital.cat)
Esteban Ballestar (eballestar@idibell.org)

ISSN 1465-6906

Article type Research

Submission date 7 May 2013

Acceptance date 9 September 2013

Publication date 12 September 2013

Article URL http://genomebiology.com/2013/14/9/R99

PU.1 target genes undergo Tet2-coupled demethylation and DNMT3b-mediated methylation in monocyte-to-osteoclast differentiation

Lorenzo de la Rica[1]
Email: ldelarica@idibell.org

Javier Rodríguez-Ubreva[1]
Email: jrubreva@idibell.org

Mireia García[2]
Email: mgperez@idibell.org

Abul BMMK Islam[3,4]
Email: khademul.islam@gmail.com

José M Urquiza[1]
Email: jurquiza@idibell.org

Henar Hernando[1]
Email: hhernando@idibell.org

Jesper Christensen[5]
Email: jesper.christensen@bric.ku.dk

Kristian Helin[5]
Email: kristian.helin@bric.ku.dk

Carmen Gómez-Vaquero[2]
Email: carmen.gomez@bellvitgehospital.cat

Esteban Ballestar[1*]
[*] Corresponding author
Email: eballestar@idibell.org

[1] Chromatin and Disease Group, Cancer Epigenetics and Biology Programme (PEBC), Bellvitge Biomedical Research Institute (IDIBELL), L'Hospitalet de Llobregat, Barcelona 08908, Spain

[2] Rheumatology Service, Bellvitge University Hospital (HUB), L'Hospitalet de Llobregat, Barcelona 08908, Spain

[3] Department of Experimental and Health Sciences, Barcelona Biomedical Research Park, Universitat Pompeu Fabra (UPF), Barcelona 08003, Spain

[4] Department of Genetic Engineering and Biotechnology, University of Dhaka, Dhaka 1000, Bangladesh

[5] Biotech Research and Innovation Center (BRIC), Center for Epigenetics
University of Copenhagen, Ole Maaløes Vej 5, Copenhagen 2200, Denmark

Abstract

Background

DNA methylation is a key epigenetic mechanism for driving and stabilizing cell-fate decisions. Local deposition and removal of DNA methylation are tightly coupled with transcription factor binding, although the relationship varies with the specific differentiation process. Conversion of monocytes to osteoclasts is a unique terminal differentiation process within the hematopoietic system. This differentiation model is relevant to autoimmune disease and cancer, and there is abundant knowledge on the sets of transcription factors involved.

Results

Here we focused on DNA methylation changes during osteoclastogenesis. Hypermethylation and hypomethylation changes took place in several thousand genes, including all relevant osteoclast differentiation and function categories. Hypomethylation occurred in association with changes in 5-hydroxymethylcytosine, a proposed intermediate toward demethylation. Transcription factor binding motif analysis revealed an overrepresentation of PU.1, NF-κB and AP-1 (Jun/Fos) binding motifs in genes undergoing DNA methylation changes. Among these, only PU.1 motifs were significantly enriched in both hypermethylated and hypomethylated genes; ChIP-seq data analysis confirmed its association to both gene sets. Moreover, PU.1 interacts with both DNMT3b and TET2, suggesting its participation in driving hypermethylation and hydroxymethylation-mediated hypomethylation. Consistent with this, siRNA-mediated PU.1 knockdown in primary monocytes impaired the acquisition of DNA methylation and expression changes, and reduced the association of TET2 and DNMT3b at PU.1 targets during osteoclast differentiation.

Conclusions

The work described here identifies key changes in DNA methylation during monocyte-to-osteoclast differentiation and reveals novel roles for PU.1 in this process.

Background

DNA methylation plays a fundamental role in differentiation as it drives and stabilizes gene activity states during cell-fate decisions. Recent reports have shown a close relationship between the participation of transcription factors during differentiation and the generation of cell type-specific epigenetic signatures [1-3]. Several mechanisms explain the co-occurrence of DNA methylation changes and transcription factor binding, including the active recruitment of enzymes involved in DNA methylation deposition, interference or alternative use of the same genomic regions. One of the best models for investigating these mechanisms is the hematopoietic differentiation system given the profound knowledge on the transcription factors implicated at different stages. Many studies have focused on hematopoiesis in order to learn about the type, distribution and role of epigenetic changes, particularly DNA

methylation during differentiation. However, the role of DNA methylation changes and the mechanisms participating in their acquisition in terminal differentiation processes remain elusive, even though these are amongst the most important since they produce functional cell types with very specific roles.

A singular differentiation process within the hematopoietic system is represented by differentiation from monocytes (MOs) to osteoclasts (OCs), which are giant, multinucleated cells that are specialized in degrading bone [4]. OCs differentiate from monocyte/macrophage progenitors following M-CSF [5] and RANKL [6] stimulation. Osteoclastogenesis requires cell fusion, cytoskeleton re-organization [7] and the activation of the specific gene sets necessary for bone catabolism. The signaling pathways activated after M-CSF and RANKL induction have been extensively described, and act through TRAF-6 [8,9], immunoreceptor tyrosine-based activation motif (ITAM) [10] adaptors DAP12 [11] and FcRg [12] associated with their respective receptors, TREM-2 [13] and OSCAR, as well as calcium oscillations [14]. Signals end in the activation of NF-kB, MAPK and c-Jun, leading to the activation of NFATc1 [15], the master transcription factor of osteoclastogenesis, together with PU.1 and MITF [16], which is already present in the progenitors. These transcription factors bind to the promoter and help up-regulating OC markers such as dendritic cell-specific transmembrane protein (*DC-STAMP*/*TM7SF4*) [17], tartrate-resistant acid phosphatase (*TRACP/ACP5*) [18], cathepsin K (*CTSK*) [19], matrix metalloproteinase 9 (*MMP9*) [20] and carbonic anhydrase 2 (*CA2*).

OC deregulation is involved in several pathological contexts, either in the form of deficient function, as is the case in osteopetrosis [21], or aberrant hyperactivation, as in osteoporosis [22]. These cells are also involved in autoimmune rheumatic disease. For instance, in rheumatoid arthritis aberrantly activated OCs are major effectors of joint destruction [23]. Moreover, OCs cause bone complications in several diseases, such as multiple myeloma [24], prostate cancer and breast cancer [25], and there is also a specific tumor with OC origin, the giant cell tumor of bone [26].

In vitro generation of OCs allows this cell type to be investigated, whereas isolating primary bone OCs for this purpose is very difficult. MOs stimulated with RANKL and M-CSF generate functional OCs [27], which degrade bone and express OC markers [28]. As indicated, the involvement of transcription factors in this model has been well studied, however very few reports have analyzed the role of epigenetic changes during osteoclastogenesis, and these focus mainly on histone modifications [29,30]. Given the relationship between transcription factors and DNA methylation, we hypothesized that examining DNA methylation changes would provide clues about the involvement of specific factors in the dynamics and hierarchy of these changes in terminal differentiation.

In this study, we compared the DNA methylation profiles of MOs and derived OCs following M-CSF and RANKL stimulation. We found that osteoclastogenesis was associated with the drastic reshaping of the DNA methylation landscape. Hypermethylation and hypomethylation occur in many relevant functional categories and key genes, including those whose functions are crucial to OC biology, like *CTSK*, *ACP5* and *DC-STAMP*. Hypomethylation occurred early, concomitantly with transcription changes, was DNA replication-independent and associated with a change in 5-hydroxymethylcytosine, which has been proposed as an intermediate in the process of demethylation. Inspection of transcription factor binding motif overrepresentation in genes undergoing DNA methylation changes revealed the enrichment of the PU.1 binding motif in hypermethylated genes and AP-1, NF-kB and also PU.1 motifs

among hypomethylated genes. In fact, analysis of PU.1 ChIPseq data showed its general association to a high number of both hypo- and hypermethylated sites. Chromatin immunoprecipitation assays and immunoprecipitation experiments, suggested a potential novel role for PU.1 recruiting DNMT3B to hypermethylated promoters, and TET2, which converts 5-methylcytosine to 5-hydroxymethylcytosine, to genes that become demethylated. This has been demonstrated by performing siRNA-mediated downregulation of PU.1 which partially impaired DNA methylation, expression and recruitment of TET2 and DNMT3B to PU.1 targets, supporting the participation of PU.1 in the acquisition of DNA methylation changes at their target sites.

Results

Cell differentiation and fusion in osteoclastogenesis are accompanied by hypomethylation and hypermethylation of key functional pathways and genes

To investigate the acquisition of DNA methylation changes during monocyte-to-osteoclast differentiation we first obtained three sets of matching samples corresponding to MOs (CD14+ cells) from peripheral blood and OCs derived from the same CD14+ cells, 21 days after the addition of M-CSF and RANKL. The quality of mature, bone-resorbing OCs obtained under these conditions was confirmed by several methods, including the presence of more than three nuclei in TRAP-positive cells (in some cases, up to 40 nuclei per cell were counted), the upregulation of OC markers, such as *CA2*, *CTSK*, *ACP5*/*TRACP* and *MMP9*, and downregulation of the monocytic gene *CX3CR1* (Additional file 1). At 21 days, over 84% of the nuclei detected in these preparations could be considered to be osteoclastic nuclei (in polykaryons, nuclei and not cells were counted) (Additional file 1). We then performed DNA methylation profiling using bead arrays that interrogate the DNA methylation status of > 450,000 CpG sites across the entire genome covering 99% of RefSeq genes. Statistical analysis of the combined data from the three pairs of samples revealed that 3515 genes (8028 CpGs) displayed differential methylation (FC $\geq$ 2 or FC $\leq$ 0.5; FDR $\leq$ 0.05). Specifically, we identified 1895 hypomethylated genes (3597 CpG sites) and 2054 hypermethylated genes (4429 CpGs) (Figure 1A and Additional file 2). Changes corresponding to the average three pairs of monocytes/osteoclasts (Figure 1B) were almost identical to the pattern obtained for each individual pair of samples (Additional file 3), highlighting the specificity of the differences observed.

Figure 1 High-throughput methylation comparison between monocytes (MOs) and derived osteoclasts (OCs). (A) Heatmap including the data for three paired samples of MOs (MO D1, D2, D3) and their derived OCs (OC D1, D2, D3) harvested on day 21. The heatmap includes all CpG-containing probes displaying significant methylation changes (8028 in total with FC ≥ 2 or FC ≤ 0.5; p ≤ 0.01 and FDR ≤ 0.05) (Additional file 2). Scale shown at the bottom, whereby beta values (i.e. the ratio of the methylated probe intensity to the overall intensity, where overall intensity is the sum of methylated and unmethylated probe intensities) ranging from 0 (unmethylated, blue) to 1 (completely methylated, red). **(B)** Scatterplot showing the mean methylation profile of three matching MO/OC pairs. Genes with significant differences (FC > 2, FDR < 0.05) in averaged results from the three pairs of samples are highlighted in red (hypermethylated) or blue (hypomethylated). **(C)** Distribution of differentially methylated CpGs among genomic regions (promoter, gene bodies, 3′UTR and intergenic) in different subsets of CpGs (hypomethylated, hypermethylated). **(D)** Gene ontology enrichment analysis of hypomethylated and hypermethylated CpGs showing the most important categories. **(E)** Technical validation of the array data by bisulfite pyrosequencing of modified DNA. BS pyrosequencing of three representative hypomethylated genes (*ACP5*, *CTSK* and *TM7SF4*) and one hypermethylated gene (*CX3CR1*) from the array data are shown. A representation showing the excellent correlation between array data (beta values) and pyrosequencing data (% methylation) including the data for the four genes (right panel). **(F)** Cluster analysis of contiguous differentially methylated regions (< 500 bp). Two examples of regions with more than nine consecutive CpGs differentially methylated are shown. **(G)** Analysis of methylation levels in repetitive elements (Sat2, D4Z4, NBL2) and ribosomal RNA genes (18S and 28S regions) as obtained from bisulfite sequencing analysis.

Over a third of the differentially methylated CpG-containing probes (33% for hypomethylated CpGs, 45% for hypermethylated CpGs) mapped to gene promoters, the best-described regulatory region for DNA methylation, although DNA methylation changes also occurred at a similar scale in gene bodies (51% for hypomethylated CpGs, 40% for hypermethylated CpGs) (Figure 1C). Gene ontology analysis of hypomethylated CpGs revealed significant enrichment (FDR ≤ 0.05) for a variety of functional categories of relevance in OC differentiation and function (Figure 1D). We observed very high significance for terms like immune response (FDR = 4.25E-25) and signal transduction (FDR = 1.45E-21), but also more specific categories such as ruffle organization (FDR = 9.91E-2), calcium ion transport (FDR = 4.6E-2) and OC differentiation (FDR = 1.94E-1). In the case of hypermethylated genes, we also found highly significant enrichment of signal transduction (FDR = 4.09E-17), and enrichment of categories related to other hematopoietic cell types, suggesting that hypermethylation and associated silencing take place in gene sets that become silent in differentiated OCs (Figure 1D). Together, these data indicate that DNA hypomethylation is targeted to genomic regions that are activated during osteoclastogenesis, and hypermethylation silences alternative lineage genes that are not expressed in OCs.

Remarkably, among the group of hypomethylated genes (Additional file 2), we identified changes in several of the archetypal OC genes near their transcription start sites. For example, *CTSK*, the lysosomal cysteine proteinase involved in bone remodeling and resorption, is hypomethylated more than 60%. The *ACP5/TRACP* gene is hypomethylated around 47%. Finally, *TM7SF4*, which encodes for DC-STAMP, a seven-pass transmembrane protein involved in signal transduction in OCs and dendritic cells, undergoes 59% hypomethylation. We also observed significant hypomethylation at the osteoclast-specific transcription factor gene *NFATC1*, although in this case hypomethylation occurred at CpG sites located in its

gene body region. Conversely, *CX3CR1*, an important factor for MO adhesion to blood vessels that is downregulated during osteoclastogenesis, displayed an increase in methylation of over 28% (Additional file 2).

To confirm that differences in DNA methylation identified between MOs and OCs were robust, we carried out bisulfite genomic pyrosequencing of the aforementioned selection of genes, looking at CpG sites corresponding to the oligonucleotide probe represented in the methylation array. In all cases, bisulfite pyrosequencing confirmed the results of the beadchip array (Figure 1E and Additional file 4). This analysis showed a very close correlation between the array and the pyrosequencing data ($R^2 = 0.9707$) (Figure 1E).

We also investigated the coordinated hypomethylation or hypermethylation of adjacent CpGs by analyzing the different sequence window lengths (from 500 bp to 1,000,000 bp). With the largest sequence windows we were able to observe the coordinated hypermethylation of multiple CpGs across several genes, like those in the HOXA gene cluster. However, the majority of CpGs undergoing coordinated methylation changes were identified within the single gene level. By analyzing CpGs that are concomitantly deregulated within a 500-bp window, we identified several genes displaying coordinated hypomethylation or hypermethylation of many CpG sites (Additional file 5). Among these, we identified several CpG clusters in genes potentially involved in OC function and/or differentiation, including 10 CpGs at the promoter of the *TM4SF19* gene, also known as OC maturation-associated gene 4 protein, and 9 CpGs in the gene body of *ARID5B*, the AT-rich interactive domain 5B (MRF1-like) (Figure 1F).

To examine the specificity of the DNA methylation changes further we performed bisulfite sequencing of repetitive elements (Sat2, D4Z4 and NBL2 repeats) and ribosomal RNA genes (Figure 1G and Additional file 3). We also performed genome-wide amplification of unmethylated DNA Alu repeats (AUMA), the most common family of repetitive elements that are present in tandem or interspersed in the genome [31]. These experiments showed no significant DNA methylation changes in any of these repetitive elements (Additional file 3), reinforcing the notion of the high specificity of hypomethylation and hypermethylation of the identified gene sets.

Hypomethylation is replication-independent and involves changes in 5-hydroxymethylcytosine

To investigate the dynamics of DNA methylation in relation to gene expression changes we first examined how DNA methylation changes are associated with expression changes and then compared the dynamics of DNA methylation and expression changes.

We used osteoclastogenesis expression data (available from the ArrayExpress database under accession number E-MEXP-2019) on 0, 5 and 20 days [32]. Our analysis showed that most changes occurred within the first 5 days, since the expression changes between 0 and 5 days were very similar to those observed between 0 and 20 days, and very few genes changed between 5 and 20 days (Figure 2A and Additional file 6). The 0-to-20-day comparison showed that 2895 genes were upregulated (FC > 2; FDR < 0.05) and 1858 were downregulated (FC < 0.5; FDR < 0.05). We found different relationships between DNA methylation changes and gene expression (Figure 2B). An inverse relationship between DNA methylation and gene expression was mainly observed for changes occurring in CpGs in the proximity of the TSS and within the first exon (Figure 2C) and it was less frequent in those at

gene bodies and 3′UTR (Figure 2C). Comparing DNA methylation and expression data revealed that 452 genes were both hypomethylated and overexpressed and 280 genes were both hypermethylated and repressed at the selected thresholds (Additional file 7). We selected a panel of 10 genes from those undergoing hypomethylation and hypermethylation to investigate the dynamics of DNA methylation and expression changes, and performed bisulfite pyrosequencing and quantitative RT-PCR over the entire osteoclastogenesis for three sets of samples (Figure 2D). We found that the promoters of genes like *ACP5*, *CTSK*, *TM7SF4* and *TM4SF19* rapidly became hypomethylated following RANKL and M-CSF stimulation (Figure 2D, top). In fact, around 60% of the entire range of hypomethylation occurred between days 0 and 4. Changes in mRNA levels occurred at a similar pace or, in some cases, in an even more gradual manner and were slightly delayed with respect to changes in DNA methylation. In contrast, hypermethylated genes like *PPP1R16B*, *CD6* and *NR4A2* (Figure 2D, bottom) displayed loss of expression before experiencing an increase in DNA methylation, highlighting the different dynamics and mechanisms involved in hypomethylation and hypermethylation events.

Figure 2 Dynamics of DNA methylation and its relationship with expression changes. **(A)** Heatmap showing expression levels on 0, 5 and 20 days for genes displaying significant methylation and expression changes (4753 in total with FC $\geq$ 2 or FC $\leq$ 0.5; p $\leq$ 0.01 and FDR $\leq$ 0.05). (Additional file 7). **(B)** Scatterplots showing the relationship between the $\log_2$-transformed FC in expression and the $\log_2$-transformed FC in DNA methylation. 62% of the hypomethylated genes are overexpressed (in blue); 55% of the hypermethylated genes are repressed (in red). **(C)** Correlation between methylation and expression data (slope from the linear regression between DNA methylation differences versus expression differences) for all differentially methylated genes organized by genomic location (first exon, TSS, 5′UTR, gene body, 3′UTR). **(D)** DNA methylation and expression dynamics of selected loci during monocyte-to-osteoclast differentiation. Methylation percentage determined by bisulfite pyrosequencing. Quantitative RT-PCR data relative to RPL38. DNA methylation and expression data are represented with a black and a red line respectively. **(E)** BrdU assay showing the percentage of replicating cells at different times. From days 1 to 4, only 9.46% of cells divide. **(F)** Effects of 5azadC treatment (50 nM, 500 nM) on osteoclastogenesis monitoring *ACP5*, *CTSK*, and *CX3CR1* levels and TRAP staining over time. **(G)** Workflow for testing the presence of 5 hydroxymethylcytosine in hypomethylated genes. DNA was treated with a 5hmC-specific glucosyltransferase. Cytosines bearing a 5-hydroxymethyl are protected against MspI digestion, and the surrounding region can be amplified by qPCR. When no 5hmC is present, glucose is not transferred to C, DNA is cleaved at CCGG sites, and there is less qPCR amplification. Several controls are used to set the 0% and 100% content of 5hmC. **(H)** 5hmC content in several of the CpGs that are rapidly demethylated after RANKL and M-CSF stimulation of OC precursors.

It is well established that osteoclastogenesis occurs in the absence of cell division. We tested the levels of cell division in our monocyte-to-osteoclast differentiation experiments by treating cells with BrdU pulses. Consistent with previous observations, fewer than 9.8% were found to be BrdU-positive between 1 and 4 days, confirming the virtual absence of replication (Figure 2E and Additional file 8). This implies that the large DNA methylation changes observed in this time period are independent of DNA replication. This conclusion is also supported by the fact that treatment with 5-Aza-2′-deoxycytidine (5azadC), a pharmacological compound that results in replication-coupled DNA demethylation [33], had no significant effect on osteoclastogenesis (Figure 2F).

The existence of DNA methylation changes in the absence of replication is particularly significant for genes undergoing demethylation, given the controversy around active DNA demethylation mechanisms. In this context, recent studies have drawn attention towards a family of enzymes, the Tet proteins, which convert 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5hmC) [34,35] and other modified forms of cytosine, 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) [36]. 5hmC, 5fC and 5caC may represent intermediates in an active demethylation pathway that ultimately replaces 5mC with cytosine in non-dividing cells [37,38]. To establish the potential involvement of these mechanisms, we here focused on the 5hmC levels at early time points in several of the genes that are hypomethylated during osteoclastogenesis, using a method that cleaves DNA that has C, 5mC, but not the glucosyl-5hmC produced as a result of treatment with the 5-hydroxymethylcytosine specific glucosyltransferase enzyme (Figure 2G). For several genes that become hypomethylated, like *ACP5* and *TM4SF19*, we observed an initial increase in 5hmC levels followed by a slight but significant decrease (Figure 2H). In other genes, like *TM7SF4* and *CD59*, of there were high levels of 5hmC before the addition of RANKL/M-CSF as if these genes were already primed for demethylation. In any case, our results

suggested the participation of hydroxymethylation, and therefore the activity of Tet proteins, in genes that undergo a reduction in DNA methylation.

Sequences undergoing DNA methylation changes are enriched for binding motifs for AP-1, NF-kB and PU.1, key transcription factors in osteoclastogenesis

Different studies have recently shown that transcription factor binding events are associated with changes in the DNA methylation profiles and the response to different situations [2,39,40]. To address this further, we first investigated the potential overrepresentation of transcription factor binding motifs among the sequences undergoing DNA methylation changes during OC differentiation using the TRANSFAC database and focusing on a region of 500 bp around the CpG sites identified as undergoing hypomethylation or hypermethylation. We noted highly significant overrepresentation of a small selection of transcription factor binding motifs for genes that undergo hypomethylation or hypermethylation (Figure 3A). We observed that the overrepresentation of binding motifs was very specific to the direction of the DNA methylation change (hypomethylation or hypermethylation).

Figure 3 Association of transcription factors with DNA methylation changes during moncoyte to OC differentiation. (A) Significant enrichment of predicted TF (TRANSFAC motif) in hypo-/hypermethylated CpG sites regions. A 500-bp window centered around the hypo-/hypermethylated CpG sites was tested. The name of the transcription factor binding motif, the p-value and the TF family are provided. Below we show three of the motifs that have a higher representation in this analysis **(B)** Diagrams showing the percentage of hypo-/hypermethylated CpGs with AP-1, NF-kB and PU.1 binding sites relative to the total number of hypo-/hypermethylated CpGs. **(C)** Quantitative ChIP assays showing the binding of three selected transcription factors (p65 NF-kB subunit, Fos and PU.1 to target genes selected by the presence of the putative binding motifs according to the TRANSFAC analysis). Samples were analyzed at 0 ad 2 days after RANKL/M-CSF stimulation. We used Sat2 repeats and the TRDR1 MyoD1 promoter as negative control sequences.

In the case of hypomethylated genes, we observed highly significant enrichment of binding motifs of the AP-1 family and NF-kB subunits (Figure 3A). We also observed enrichment of PU.1 (FDR 1.07E-12). In fact, 39% of all hypomethylated genes had binding motifs for AP-1, 15% genes had NF-kB binding motifs and another 15% genes had binding motifs for PU.1 or other ETS-related factors (PU.1 alone, 9%) (Figure 3B). As aforementioned, these three groups of TFs play critical roles in osteoclastogenesis [41]. For instance, c-Fos, a component of the dimeric TF AP-1, regulate the switch between monocytes/macrophages and OC differentiation. Fra-1 is downstream to c-Fos, whereas PU.1 and NF-kB are upstream. NF-kB is critical in the expression of a variety of cytokines involved in OC differentiation. In the case of hypermethylated genes, we identified even greater enrichment of the binding motifs of ETS-related transcription factors, especially PU.1 (Figure 3A). In fact, the PU.1 binding motif is present in 15% of all hypermethylated genes (Figure 3B). Other motifs of ETS-related transcription factors from our list of hypermethylated genes included SPIB, ESE1, ETS1, ETS2 and others (Figure 3A). Much lower or insignificant levels of enrichment were obtained for AP-1 family members and NF-kB subunits among the hypermethylated genes. Previous studies have shown that genes that become methylated during hematopoietic differentiation are characterized by ETS transcription factors [2]. This appears to be particularly relevant in monocytic differentiation [1]. Interestingly, most of the reports about

the role of PU.1 in osteoclastogenesis are associated with the activation of osteoclast-specific genes. However, in relation with methylation changes, PU.1 appears to be better correlated with those changes in the direction of repression. Overall, the analysis of transcription factor motifs showed that several of the factors associated with osteoclastogenesis had a significant overrepresentation of their binding motifs among the sets of hypo-and hypermethylated genes (Additional file 9).

To confirm the association of some of these factors with genes that become hypo- and hypermethylated we performed chromatin immunoprecipitation (ChIP) assays with a selection of transcription factors including PU.1, the NFkB subunit p65 and c-Fos, given their known role in osteoclastogenesis as well as the presence of binding motifs for them around hypo- and hypermethylated genes (in the case of c-Fos, it was chosen as a component of the dimeric TF AP-1). To select candidate genes we considered genes with motifs for these three factors among the list of hypo- and hypermethylated genes. For instance, genes that become hypomethylated and have binding sites for p65 include *CCL5*, *IL1R* and *TNFR5SF*. In the case of transcription factor c-Fos, we looked at genes that become hypomethylated like *IL7R*, *CD59* and *IL1R*. In the case of PU.1, we chose key genes with PU.1 binding near the differentially methylated CpG, including *ACP5* and *TM4SF7* (hypomethylated) and *CX3CR1* (hypermethylated). ChIP assays demonstrated the interaction of these factors with most of the aforementioned promoters, even before the stimulation with M-CSF and RANKL (Figure 3C), as if these genes were primed by these factors in monocytes. No binding was observed in control sequences like Sat2 repeats and the *MYOD1* and *TDRD1* promoters. Interestingly, in the case of PU.1, with both hypo- and hypermethylated genes displaying binding at 0 days, genes becoming demethylated showed a slight increase in PU.1 binding, whereas hypermethylated genes showed a slight decrease in PU.1 association (Figure 3C).

PU.1 recruits DNMT3b and TET2 to hypermethylated and hypomethylated genes

To investigate the potential role of the aforementioned transcription factors in the acquisition of DNA methylation changes we chose PU.1 and NF-kB (p65 subunit) as two representative examples. We first checked their expression levels during osteoclastogenesis, by carrying out qRT-PCR and western blot assays. mRNA and protein analysis (Figure 4A and 4B) both confirmed the expression of these factors. PU.1 revealed an increase at the mRNA levels, although there was no change at the protein level. In the case of p65 NF-kB we only observed a clear increase at the protein level (Figure 4B). In parallel, we also confirmed the presence of DNMT3b, a *de novo* DNA methyltransferase, and the Ten eleven translocation (TET) protein TET2, as enzymatic activities potentially related with DNA demethylation (Figure 4A and 4B). TET proteins are responsible for conversion of 5mC in 5hmC [34], 5fC and 5cac [36]. Recent evidences support a role for TET-dependent active DNA demethylation process [42,43]. We focused on TET2 given their high levels in hematopoietic cells of myeloid origin [44,45]. Also, we have recently reported that TET2 plays a role in derepressing genes in pre-B cell to macrophage differentiation [44], and recent data shows that TET2 is required for active DNA demethylation in primary human MOs [45]. In fact TET1 and TET3 were undetectable in western blot (not shown) and qRT-PCR evidenced their low levels in this cell type (Figure 4A, only shown for TET1).

Figure 4 Interactions between PU.1 and DNTM3b and TET2 and association with promoters undergoing DNA methylation changes. (A) Quantitative RT-PCR analysis for *CTSK*, *PU.1*, *p65 NF-kB*, *TET1*, *TET2* and *DNMT3B* during osteoclastogenesis. **(B)** Western blot for the same factors indicated above. **(C)** Immunoprecipitation experiment of p65 and PU.1 with DNMT3B and TET2 at 0, 2 and 4 days after RANKL and M-CSF stimulation. IgG used as a negative control. Reciprocal immunoprecipitation experiments in the bottom panel. **(D)** Quantitative ChIP assays showing PU.1, TET2 and DNMT3b binding to hypomethylated genes (*ACP5*, *TM7SF4*, *TM4SF19*) and hypermethylated genes (*CX3CR1*, *NR4A2*), all direct PU.1 targets, and a negative control *(MYOD1* promoter) without PU.1 target sites. The experiment was performed with three biological triplicates but only one experiment is shown. T-student test comparing binding of each antibody between 0 d vs 2 d was performed: * corresponds to p-value < 0.05; ** means p-value < 0.01; *** means p-value < 0.001. **(E)** Examples showing PU.1 binding (from ChIPseq data, GSE31621) to the region neighbouring hypo- and hypermethylated CpGs. The PU.1 binding motif location is presented as a horizontal blue dot and the CpG displaying differential methylation (Illumina probe) between MO and OC is marked with a red bar. **(F)** Analysis of ChIPseq analysis for PU.1 and comparison to TRANSFAC predictions. Top panel: proportion of the CpG-containing probes displaying DNA methylation changes that have peaks for PU.1 binding in the same 500 bp window. Diagrams are separated in the hypo- and hypermethylated sets and in promoter and distal regions (gene bodies, 3′UTR and intergenic regions). Bottom panel, Venn diagrams showing the overlap of PU.1 targets from ChIPseq data (GEO accession number: GSE31621) in MOs and TRANSFAC prediction for PU.1, both using a window of 500 pb centered by the CpG displaying significant methylation changes.

The confirmed binding of factors like PU.1 and the p65 subunit of NF-kB to hypo- and hypermethylated genes (Figure 3C) raised the possibility of their potential direct interaction with factors involved in maintaining the DNA methylation homeostasis. Some of these interactions have already been explored. For instance, PU.1 physically interacts with the *de novo* DNA methyltransferases DNMT3A and DNMT3B [46]. Such an interaction, if it occurred in osteoclastogenesis, could provide a potential mechanism to explain how PU.1 target genes become hypermethylated. One would expect that these transcription factors could also interact with factors participating in demethylation processes. Our previous results suggested the existence of 5hmC enrichment in genes that become hypomethylated, and therefore it is reasonable to test whether NF-kB p65 and PU.1 interact with Tet proteins, the enzymes catalyzing hydroxylation of 5mC.

We therefore tested the recruitment by NF-kB p65 and PU.1 of both DNMT3b and TET2 by carrying out immunoprecipitation assays with osteoclastogenesis samples 0, 2 and 4 days after stimulation with M-CSF and RANKL. Our results showed that PU.1 directly interacted with both DNMT3b and TET2 (Figure 4C). It is plausible that these two interactions may involve different subpopulations of PU.1, for instance with specific post-translational modifications like Ser phosphorylation. However, we did not address this aspect at this point. In the case of NF-kB, we did not observe binding with either of these factors (Figure 4C). This could perhaps be explained by the fact that p65 is shuttling back to the cytoplasm much of the time.

To confirm the interaction between PU.1 and DNMT3b and TET2, we performed reciprocal immunoprepitation experiments with anti-DNMT3b and anti-TET2. These confirmed the direct interaction with PU.1 (Figure 4C). Our results suggested that PU.1 may play a dual coupling transcription factor that can interact with the DNA methyltransferases and enzymes

perhaps participating in, or leading to, demethylation. It is likely that other factors participate in the recruitment of these enzymes, however at this stage we focused on PU.1 because of its ability to bind both DNMT3b and Tet2 and its association with both hyper- and hypomethylated sequences.

We then investigated the dual role of PU.1 in recruiting TET2 and DNMT3b to promoters. To this end we performed chromatin immunoprecipitation assays with PU.1, TET2 and DNMT3b in MOs at 0 and 2 days following stimulation with M-CSF/RANKL. We amplified gene promoters with predicted binding sites for PU.1 that become both demethylated (*ACP5*, *TMS7SF4* and *TM4SF19*) as well as hypermethylated (*CX3CR1* and *NR4A2*) and used a non-target of PU.1 (*MYOD1*) as negative control. Our analysis showed binding of PU.1 at both 0 and 2 days (Figure 4D). For genes that become hypomethylated, we observed an increased recruitment of TET2 at these promoters after 2 days, whereas DNMT3b was initially enriched but its association with these promoters was lost after M-CSF and RANKL stimulation (Figure 4D). In genes that become hypermethylated (*CX3CR1* and *NR4A2*), we also observed association of PU.1 at both 0 and 2 days. However we again observed a slight decrease at 2 days together. We also observed increased recruitment of DNMT3b at 2 days after M-CSF/ RANKL stimulation. We did not observe association of PU.1, DNMT3b and TET2 in the negative control for PU.1 binding, the MyoD promoter.

To evaluate the extent to which hypo- and hypermethylated genes correlate with PU.1 occupancy, we used our DNA methylation data and PU.1 ChIPseq data (GSE31621) obtained in MOs [1]. Most of the individual example genes previously analyzed displayed PU.1 binding overlapping or in the proximity of the CpG sites undergoing a methylation change (Figure 4E). To systematize the analysis we used a window of 500 bp centered around the CpG displaying DNA methylation changes. Under these conditions we found that 10.7% of all hypomethylated CpGs located in promoter regions genes and 25.1% of all hypermethylated CpGs located in promoter regions had PU.1 peaks within this 500 bp window (Figure 4F). These numbers were similar when focusing on CpGs located in distal regions (Figure 4F). We also compared the ChIPseq data to the prediction by TRANSFAC analysis and observed that the overlap between the two sets of list was around 20.6% for hypomethylated genes and 46.9% for hypermethylated genes, again using the same 500 bp for both datasets (Figure 4F). These analyses reinforced the notion of PU.1 associated with a high number of genes undergoing DNA methylation changes, however it also reveals the weakness in the predictive power of TRANSFAC motif searches and the need of experimental validation of its results.

Dowregulation of PU.1 in MOs impairs activation of OC markers, hypomethylation and recruitment of DNMT3b and TET2

To investigate a potential causal relationship between PU.1 and DNA methylation changes in monocyte-to-osteoclast differentiation we investigated the effects of ablating PU.1 expression in MOs. We therefore downregulated PU.1 levels in MOs using transient transfection experiments with a mix of two siRNAs targeting exon2 and the 3′UTR of PU.1 (Figure 5A). In parallel, we used a control siRNA. Following transfection we stimulated differentiation using RANKL/M-CSF. In these conditions, we checked by qRT-PCR and western blot the effects on PU.1 levels at 1, 2 , 4 and 6 days following RANKL/M-CSF stimulation of MOs and confirmed the PU.1 downregulation close to 60% (Figure 5B and 5C). We then observed that the upregulation of genes like *ACP5* and *CTSK* (both PU.1-direct targets) was partially impaired (Figure 5D). In the case of genes like *CX3CR1* and *NR4A2* we determined that

downregulation was also impaired in PU.1-siRNA treated MOs. Interestingly, we also analyzed two genes that are not direct PU.1 targets, one upregulated and hypomethylated during osteoclastogenesis (*PLA2G4E*) and a second one, highly methylated, that does not experience DNA methylation changes during OC differentiation (*FSCN3*). PU.1-siRNA treatment had only small effects on gene expression changes during osteoclast differentiation (perhaps due to indirect effects) when compared to control siRNA, confirming the specificity of the changes observed for the other genes (Figure 5D, bottom).

Figure 5 PU.1 has a direct role in leading DNA methylation changes at their targets. (A) Scheme depicting the two regions of the SPI1 gene (PU.1) (exon 2 and 3′UTR) targeted by the two siRNAs used in this study. **(B)** Effects of siRNA experiments on PU.1 levels at 1, 2, 4 and 6 days as analyzed by qRT-pCR **(C)** Effects of siRNA experiments on PU.1 levels at 1, 2, 4 and 6 days as analyzed by western blot **(D)** Effects of PU.1 downregulation on expression and methylation of PU.1-target genes that become demethylated (*ACP5*, *CTSK*), genes that become hypermethylated (*CX3CR1*, *NR4A2*) and non pU.1 target genes, *PLA2G4E*, which becomes also overexpressed and demethylated, and *FSCN3*, which is hypermethylated and does not undergo loss of methylation during osteoclastogenesis **(E)** ChIP assays showing the effects of PU.1 downregulation in its recruitment, together with TET2 and DNMT3b binding to the same genes. Data were obtained at 0, 2 and 6 days after M-CSF /RANL stimulation. To simplify the representation negative control assays with IgG for each time point have been substracted to the experiments with each antibody. We have used the *MYOD1* promoter as a negative control (data without substracting the background is presented in Additional file 10). The experiment was performed with three biological triplicates but only one experiment is shown. Error bars correspond to technical replicates. Some of them are smaller than the data point icon. T-student test was performed: * corresponds to p-value < 0.05; ** means p-value < 0.01; *** means p-value < 0.001.

We then tested the effects of PU.1 downregulation in DNA methylation changes. We looked at both PU.1-target genes that become hypomethylated and hypermethylated. In both cases, we observed that downregulation of PU.1 impaired the acquisition of DNA methylation changes, in contrast with the changes observed for control siRNA-treated MOs (Figure 5D). In the case of *TM7SF4*, one of the key genes undergoing hypomethylation, we did not detect an effect of PU.1 downregulation on its DNA methylation dynamics (Additional file 10). However, this could perhaps be explained because this gene undergoes changes before downregulation of PU.1 by siRNA is effective, within day 1 (Additional file 10) and suggests the participation of other factors in this process. At any rate, the observed effects only occurred in PU.1 targets. It did not affect genes that are not targeted by PU.1 (*PLA2G4E*, *FSCN3*). In the case of *PLA2G4E,* PU.1-siRNA treatment did not impair the loss of methylation that occurred in the control experiment. For *FSCN3*, we observed no loss of methylation in any case (Figure 5D).

Finally, we compared the effect of PU.1 downregulation in the recruitment of DNMT3b and TET2 to hyper- and hypomethylated promoters (Figure 5E and Additional file 10). As expected, we observed that PU.1 dowregulation resulted in a decrease of the levels of PU.1 associated with the promoters of both hypo- and hypermethylated genes. Most importantly, it also reduced the association of DNMT3b and TET2 reinforcing the notion of the role of these factors and their association with PU.1 in the DNA methylation changes occurring at these CpG sites (Figure 5E). The time course analysis (at 2 and 6 days) of these results also revealed a complex dynamics for the PU.1, TET2 and DNMT3b interactions with their target genes, particularly in the case of hypermethylated genes. It is possible that perturbation of

PU.1 levels could be compensated by additional factors that participate in the acquisition of DNA methylation changes of these genes. These aspects will need to be further investigated.

Discussion

Our results provide evidence of the participation of transcription factors, focusing on PU.1, in determining changes in DNA methylation during monocyte-to-osteoclast differentiation. First, a detailed analysis of the sequences undergoing DNA methylation changes produced evidences of the participation of several transcription factors, given the specific overrepresentation of certain motifs in hypo- and hypermethylated genes. This initial analysis was validated in several candidate genes and using ChIPseq data for human primary monocytes [1]. Second, further analyses on one these candidate transcription factors, PU.1, and manipulation of its levels revealed a novel role for this factor in mediating DNA methylation changes during osteoclastogenesis, by direct binding of both DNMT3B and TET2.

In general, DNA methylation changes in differentiation or any other dynamic process are of interest for two reasons: 1) these changes are generally associated with gene expression changes, particularly when associated with promoters or gene bodies, and reveal aspects intrinsic to identity and function of the corresponding cell types. 2) they can be considered as epigenetic footprints that, despite not necessarily being associated with an expression or organizational change, reveal a change in the milieu of a particular CpG and therefore can be used to trace the participation of specific transcription factors or other nuclear elements in that environment/neighborhood. This information can then be used to reconstruct cell signaling events, transcription factors involved and mechanisms participating in differentiation. In this sense, our data show that DNA methylation changes are involved in the differentiation dynamics and stabilization of the OC phenotype since they are concomitant with, or even precede, expression changes. These data are closely correlated with gene expression changes, and a majority of genes that undergo hypomethylation or hypermethylation at their promoters or gene bodies also experience a change in expression, although the relationship varies between different gene sets. Finally, gene ontology analysis reveals that all relevant functional categories and the majority of key genes for differentiation or the activity of functional OCs undergo DNA methylation changes and that genes within all relevant functional categories undergo DNA methylation changes..

Our study suggests that both hypomethylation and hypermethylation events are equally important. Hypomethylation events, in many cases associated with gene activation, affect genes that are specific to this differentiation process or are related with the function of differentiated OCs. In contrast, the identity of genes affected by hypermethylation events is less closely correlated with OC function, given that most of them are related with gene repression. In fact, we found that hypermethylation affects genes that are specific to other cellular types. Given that osteoclastogenesis involves cell fusion and the generation of highly polyploid cells, we had speculated whether the existence of redundant copies of genetic material could lead to massive gene repression, and the silencing of extra copies. However, hypermethylation does not seem to be predominant over hypomethylation. The two activities are very specific to particular gene sets and there are no indications of changes in repetitive elements.

A number of transcription factors are essential for OC formation. Some of these factors are involved in various differentiation processes. Among these, PU.1, c-Fos, NF-kB and other factors are essential for osteoclastogenesis. In fact, NF-kB and PU.1-deficient mice show a macrophage differentiation failure, and osteoclastogenesis is inhibited at an early stage of differentiation. c-Fos is a component of the dimeric TF AP-1, which also includes FosB, Fra-1, Fra-2, and Jun proteins such as c-Jun, JunB, and JunD. Other key factors involved in OC differentiation include CEBPalpha [47] and Bach1[48]. Osteoclastogenesis also depends on the activity of more specific transcription factors like NFATc1 and MITF. Interestingly, the analysis of the presence of transcription factor binding sites in sequences that undergo DNA methylation changes shows a significant enrichment in binding motifs of transcription factors that are key in OC differentiation, some of which we have validated for a selection of putative target genes.

One of the most interesting factors in this process is the ETS factor PU.1. In fact, PU.1 is the earliest molecule known to influence the differentiation and commitment of precursor myeloid cells to the OC lineage. PU.1 functions in concert with other transcription factors, including c-Myb, C/EBPa, cJun and others, to activate osteoclast-specific genes.

Our results reveal two hitherto undescribed roles for PU.1 in the context of monocyte-to-OC differentiation. First, we have identified the association of PU.1 with genes that become repressed through hypermethylation and describe its direct interaction with DNMT3b in the context of osteoclastogenesis. Second, we identify a novel interaction between PU.1 and TET2 and their association with genes that become demethylated. Our study shows that PU.1 may act as a dual adaptor during osteoclastogenesis, in the directions of hypomethylation and hypermethylation. This is compatible with previous data on genome wide DNA methylation profiling comparing cell types across the hematopietic differentiation system where an overrepresentation of ETS transcription factor binding sites was found [2]. In monocyte-to-osteoclast differentiation, PU.1 is best known for its role in the activation of osteoclast-specific genes. However, studies in other models have previously shown that PU.1 can participate in the repression of genes in concert with elements of the epigenetic machinery. For instance, PU.1 is known to generate a repressive chromatin structure characterized by H3K9me3 in myeloid and erythroid differentiation [49]. Also, PU.1 has been shown to act in concert with MITF to recruit corepressors to osteoclast-specific in committed myeloid precursors capable of forming either macrophages or OC [50]. Moreover, previous studies have shown that PU.1 can form a complex with DNMT3a and DNMT3b [46]. However, this is the first report where the association between PU.1 and DNMTs in association with gene repression is shown in this context.

Moreover, our findings constitute the first report where the binding of PU.1 to TET2 has been described. Several recent reports have pointed at TET2-mediated hydroxylation of 5-methylcytosine as an intermediate step towards demethylation [51] and our data shows changes in 5hmC at genes that become demethylated in osteoclastogenesis, reinforcing the possibility that PU.1-mediated recruitment of TET2 is leading to 5hmC-mediated demethylation. However the detailed mechanisms that couple hydroxylation of 5mC and demethylation are still object of debate.

The manipulation of PU.1 levels by using siRNAs has shown that PU.1 has a direct role in recruiting DNMT3b and TET2 to its target promoters, as well as showing how impaired association of PU.1 results in defective acquisition of DNA methylation changes in both directions as well as reduced effect on gene expression changes. Therefore, our data reveal a

novel role of PU.1 as a dual adaptor with the ability to bind both epigenetically repressive and epigenetically activating events and targeting DNA methylation changes in both directions (Figure 6). The incomplete impairment of DNA methylation and expression changes, as well as partial loss of Tet2 and DNMT3b following PU.1 knock-down indicates that additional transcription factors are also participating in this process. In future studies, it will also be interesting to identify the mechanisms that operate in the specific recruitment of PU.1-TET2 to genes that become demethylated, and to determine how PU.1-DNMT3b is recruited to genes that become hypermethylated. It is likely that specific transcription factors play a role, and specific post-translational modifications in PU.1 may participate in the coupling of its associated complexes to specific factors. In this context, Ser phosphorylation of PU.1 has already been shown to play a role in its recruitment to promoters [52] and could also participate in discriminating interaction with epigenetic modifiers.

Figure 6 Model showing a simplified diagram proposing the recruitment of TET2 and DNMT3b by PU.1 to its target genes that become hypo- or hypermethylated respectively during osteoclastogenesis. Genes that become hypomethylated exchange PU.1-DNMT3b by PU.1-TET2 (although whether pre-existing subpopulations of these associations may exist or, alternatively, post-translational or another mechanisms may mediate exchange of TET2 and DNMT3b. This is not elucidated at present). TDG is likely to mediate conversion of 5hmC/5fmC/5caC to demethylated cytosine. Hypermethylated genes experience an increase in the binding of DNMT3b as differentiation to OCs is triggered.

Our study has allowed us to identify key DNA methylation changes during OC differentiation and has revealed an implication of PU.1 in the acquisition of DNA methylation and expression changes as well as identifying novel interactions with DNMT3b and TET2.

Conclusions

Our study of the DNA methylation changes in monocyte-to-osteoclast differentiation reveals the occurrence of both hypomethylation and hypermethylation changes. These changes occur in the virtual absence of DNA replication suggesting the participation of active mechanisms, particularly relevant for hypomethylation events, for which the mechanisms are still subject of debate. Also, when comparing the dynamics of DNA methylation and expression changes, hypomethylation occurs concomitant or even earlier than expression changes. In contrast, for the majority of genes becoming hypermethylated, hypermethylation follows expression changes. Hypomethylation takes place in relevant functional categories related with OC differentiation and most of the genes that are necessary for OC function undergo hypomethylation including *ACP5*, *CTSK* and *TM7SF4* among others. The analysis of overrepresentation of transcription factor binding motifs reveals the enrichment of specific motifs for hypomethylated and hypermethylated genes. Among these, PU.1 and other ETS-related binding motifs are highly enriched in both hypomethylated and hypermethylated genes. We have demonstrated that PU.1 is bound to both hypo- and hypermethylated promoters and that it is able to recruit both DNMT3b and TET2. Most importantly, downregulation of PU.1 with siRNAs not only shows a reduction in the recruitment of these two enzymes to PU.1 target genes but also results in a specific reduction in the acquisition of DNA methylation and expression changes at those targets. Our results demonstrate a key role of PU.1 in driving DNA methylation changes during OC differentiation.

Materials and methods

Differentiation of OCs from peripheral blood mononuclear cells

Human samples (blood) used in this study came from anonymous blood donors and were obtained from the Catalan Blood and Tissue Bank (Banc de Sang i Teixits) in Barcelona as thrombocyte concentrates (buffy coats). The anonymous blood donors received oral and written information about the possibility that their blood would be used for research purposes, and any questions that arose were then answered. Prior to obtaining the first blood sample the donors signed a consent form at the Banc de Teixits. The Banc de Teixits follows the principles set out in the WMA Declaration of Helsinki. The blood was carefully layered on a Ficoll–Paque gradient (Amersham, Buckinghamshire, UK) and centrifuged at 2000 rpm for 30 min without braking. After centrifugation, peripheral blood mononuclear cells (PBMCs), in the interface between the plasma and the Ficoll–Paque gradient, were collected and washed twice with ice-cold PBS, followed by centrifugation at 2000 rpm for 5 min. Pure CD14+ cells were isolated from PBMCs using positive selection with MACS magnetic CD14 antibody (Miltenyi Biotec). Cells were then resuspended in α-minimal essential medium (α-MEM, Glutamax no nucleosides) (Invitrogen, Carlsbad, CA, USA) containing 10% fetal bovine serum, 100 units/ml penicillin, 100μg/ml streptomycin and antimycotic and supplemented with 25 ng/mL human M-CSF and 50ng/ml hRANKL soluble (PeproTech EC, London, UK). Depending on the amount needed, cells were seeded at a density of $3 \cdot 10^5$ cells/well in 96-well plates, $5 \cdot 10^6$ cells/well in 6-well plates or $40 \cdot 10^6$ cells in 10 mm plates and cultured for 21 days (unless otherwise noted); medium and cytokines were changed twice a week. The presence of OCs was checked by tartrate-resistant acid phosphatase (TRAP) staining using the Leukocyte Acid Phosphatase Assay Kit (Sigma–Aldrich) according to the manufacturer's instructions. A phalloidin/DAPI stain allowed us to confirm that the populations were highly enriched in multinuclear cells, some of them containing more than 40 nuclei. We used several methods to determine that on day 21 almost 85% of the nuclei detected were "osteoclastic nuclei" (in polykaryons, nuclei and not cells were quantified). OCs (TRAP-positive cells with more than three nuclei) were also analyzed at the mRNA level: upregulation of key OC markers (*TRAP/ACP5*, *CA2*, *MMP9* and *CTSK*) and the downregulation of the MO marker *CX3CR1* were confirmed.

Treatment of MOs with 5-aza-2-deoxycytidine

In some cases we performed monocyte-to-osteoclast differentiation experiments in the presence of different subtoxic concentrations of the DNA replication-coupled demethylating drug 5-aza-2-deoxycytidine (at 50 nM, 500 mM) for 72 h.

Visualization of OCs with phalloidin and DAPI staining

PBMCs or pure isolated CD14+ cells were seeded and cultured in glass Lab-Tek Chamber Slides (Thermo Fisher Scientific) for 21 days in the presence of hM-CSF and hRANKL. OCs were then washed twice with PBS and fixed (3.7% paraformaldehyde, 15 min). Cells were permeabilized with 0.1% (V/V) Triton X-100 for 5 min and stained for F-actin with 5 U/mL Alexa Fluor® 647-Phalloidin (Invitrogen). Cells were then mounted in Mowiol-DAPI mounting medium. Cultures were visualized by CLSM (Leica TCP SP2 AOBS confocal microscope).

DNA methylation profiling using universal bead arrays

Infinium HumanMethylation450 BeadChips (Illumina, Inc.) were used to analyze DNA methylation. This array allows interrogating > 485,000 methylation sites per sample at single-nucleotide resolution, covering 99% of RefSeq genes, with an average of 17 CpG sites per gene region distributed across the promoter, 5′UTR, first exon, gene body and 3′UTR. It covers 96% of CpG islands, with additional coverage in CpG island shores and the regions flanking them. DNA samples were bisulfite converted using the EZ DNA methylation kit (Zymo Research, Orange, CA). After bisulfite treatment, the remaining assay steps were performed following the specifications and using the reagents supplied and recommended by the manufacturer. The array was hybridized using a temperature gradient program, and arrays were imaged using a BeadArray Reader (Illumina Inc.). The image processing and intensity data extraction software and procedures were those previously described [53]. Each methylation data point is obtained from a combination of the Cy3 and Cy5 fluorescent intensities from the M (methylated) and U (unmethylated) alleles. Background intensity computed from a set of negative controls was subtracted from each data point. For representation and further analysis we used both Beta values and M values. The Beta-value is the ratio of the methylated probe intensity and the overall intensity (sum of methylated and unmethylated probe intensities). The M-value is calculated as the log2 ratio of the intensities of methylated probe versus unmethylated probe. The Beta-value ranges from 0 to 1 and is more intuitive and was used in heatmaps and in comparisons with DNA methylation percentages from bisulfite pyrosequencing experiments, however for statistic purposes it is more adequate the use of M values [54].

Detection of differentially methylated CpGs

The approach to select differentially methylated CpGs was implemented in R [55], a well-known language in statistical computing. In order to process Illumina Infinium HumanMethylation450 methylation data, we used the methods supplied in limma [56], genefilter, and lumi [57] packages from Bioconductor repository. Previous to statistical analysis, a pre-process stage is applied, the main steps are: 1) Color balance adjustment, i.e., normalization between two color channels; 2) Performing quantile normalization based on color balance adjusted data, and 3) variance filtering by IQR (Interquartile range) using 0.50 for threshold value. Subsequently, for statistical analysis, eBayes moderated t-statistics test was carried out from limma package [56]. Specifically, a paired limma was performed as designed in IMA package [58]. To choose significant differences in methylated CpGs several criteria have been proposed. In this study, we considered a probe as differentially methylated if (1) has a fold-change >2 for hypermethylated and <0.5 hypomethylated) and (2) the statistical test was significant (p-value < 0.01 and FDR < 0.05).

Identification of genomic clusters of differentially methylated CpGs

A clustering method was applied to the differenced methylated CpGs from charm package [59]. We re-implemented the code to invoke the main clustering function using genomic CpG localisation: identify Differentially Methylated Regions (DMRs) by grouping differentially methylated probes (DMPs). The maximum allowable gap between probe positions for probes to be clustered into the same region was set to 500 bp. It has been shown that in many cases methylation changes are observed over a range of CpGs, which may be identified for instance at shores close to Transcription Starting Sites. We considered that DMR are more robust

signals than DMPs. In this analysis, the considered list of CpGs attains a p-value below 0.01 and FDR < 0.05.

Bisulfite sequencing and pyrosequencing

We used bisulfite pyrosequencing to validate CpG methylation changes resulting from the analysis with the Infinium HumanMethylation450 BeadChips. Bisulfite modification of genomic DNA isolated from MOs, OCs, and samples from time course or PU.1-knockdown experiments was carried out as described by Herman et al. [60]. 2 μl of the converted DNA (corresponding to approximately 20–30 ng) were then used as a template in each subsequent PCR. Primers for PCR amplification and sequencing were designed with the PyroMark® Assay Design 2.0 software (Qiagen). PCRs were performed with the HotStart Taq DNA polymerase PCR kit (Qiagen), and the success of amplification was assessed by agarose gel electrophoresis. PCR products were pyrosequenced with the Pyromark™ Q24 system (Qiagen). In the case of repetitive elements (Sat2, D4Z4, NBL2 and 18S rRNA and 28 rRNA) we performed standard bisulfite sequencing of a minimum of 10 clones. Results from bisulfite pyrosequencing and sequencing of multiple clones are presented as a percentage of methylation. All primer sequences are listed in Additional file 11. Raw data for bisulfite sequencing of all samples is presented in Additional file 4.

Gene expression data analysis and comparison of DNA expression data versus DNA methylation data

In order to compare expression data versus methylation data, we used CD14+ and OC expression data from ArrayExpress database [61]) under the accession name (E-MEXP-2019) from a previous publication [32]. Affymetrix GeneChip Human Genome U133 Plus 2.0 expression data was processed using limma [56] and affy [62] packages from bioconductor. The pre-processing stage is divided in three major steps: 1) background correction, 2) normalization, and 3). reporter summarization. Here, the expresso function in affy package was chosen for preprocessing. Thus, the RMA method [63] was applied for background correction. Then, a quantile normalization was performed. In addition, we introduced a specific step for PM (perfect match probes) adjustment, utilizing the PM-only model based expression index (option 'pmonly'). And finally, for summarization step, the median polish method was taken. Next, as previously in the methylation analysis, a variance filtering by IQR (Interquartile range) using 0.50 for threshold value was executed. After preprocessing, a statistical analysis was applied, using eBayes moderated t-statistics test from limma package. Subsequently, expression genes matching to methylated genes were studied. Genes differentially expressed between MOs and Mo-OCs groups were selected with a criteria of p-value lower than 0.01 and False Discovery Rate (FDR) lower than 0.05 as calculated by Benjamini-Hochberg and a fold-change of expression higher than 2 or lower than 0.5. Validation of expression data was performed by quantitative RT-PCR. All primer sequences are listed in Additional file 11.

Gene ontology analysis

Gene ontology (GO) was analyzed with the FatiGO tool [64], which uses Fisher's exact test to detect significant over-representation of GO terms in one of the sets (list of selected genes) with respect to the other (the rest of the genome). Multiple test correction to account for the

multiple hypotheses tested (one for each GO term) was applied to reduce false positive results. GO terms with adjusted values of $p < 0.05$ were considered significant.

Analysis of transcription factor binding

We used the STORM algorithm [65] to identify potential overrepresentation of transcription factor motifs in the 500 bp region around the center of the hypomethylated/hypermethylated CpG sites (as well as for all other CpGs-containing probes contained in the array) assuming cutoff values of $p = 0.00002$ (for hypo-/hypermethylated probes) and 0.00001 (for all other probes), using position frequency matrices (PFMs) from the TRANSFAC database (Professional version, release 2009.4) [66]. Enrichment analysis of predicted TF in the probes of significant hypomethylated probes ($n = 421$) was conducted using GiTools ([67]; [68]). We calculated two-tailed probabilities, and a final adjusted FDR p-value (with 0.25 cutoff) was considered statistically significant.

We downloaded PU.1 ChIPseq data for CD14+ MOs generated by Michael Rehli's laboratory [1] from the Gene Expression Omnibus (GSE31621). The genomic locations of the calculated peaks were mapped to GRCh37.p10 human alignment obtained from Biomart [69], by using bedtools (intersect function) in order to obtain the PU.1 occupied genes. To determine whether a given CpG (from the Illumina bead array) was positive for PU.1 binding, we used the same 500 bp window used for TRANSFAC analysis.

Graphics and heatmaps

All graphs were created using Prism5 Graphpad. Heatmaps were generated from the expression or methylation data using the Genesis program (Graz University of Technology) [70].

BrdU proliferation assays

BrdU was used at a final concentration of 300 μm, as previously described. On the days specified, BrdU pulsing solution was added to each well for 2 to 4 days. For confocal microscopy of monocyte-to-osteoclast differentiation samples, CD14+ cells were seeded on Millicell EZ 8-well glass slides (Millipore) and cultured in differentiation media. At different times BrdU was added to the medium and after 2–4 days cells were fixed (4% paraformaldehyde, 30 minutes, RT), permeabilized (PBS-BSA-Triton X-100 0,8%, 10 minutes, RT) and treated with HCl 2N for 30 minutes. After DNA opening, HCl was neutralized by two 5- minute washes with NaBo (0.1M, pH 8.5) and two 5-minutes washes with PBT. Cells were incubated with anti-BrdU antibody (18 h at 4°C, 1:1000 dilution) and an anti-mouse Alexa-568 conjugated antibody was added to visualize the BrdU-positive nuclei. A phalloidin incubation step and Mowiol-DAPI mounting media were used.

Transfection of primary human MOs

We used two different Silencer® select pre-designed siRNAs against human PU.1 (one targeting exon 2 and another targeting the 3′UTR) and a Silencer® select negative control to perform PU.1 knockdown experiments in peripheral blood MOs. We used Lipofectamine RNAiMAX Transfection Reagent (Invitrogen) for efficient siRNA transfection. mRNA and protein levels were examined by quantitative RT-PCR and western blot at 1, 2, 4 and 6 days

after siRNA transfection. In this case MO samples were prepared by incubating PBMCs in plates in α-MEM without serum for 30 min and washing out the unattached cells. Under these conditions over 80% are MOs. This alternative protocol was used for increased viability following transfection. These experiments were performed with three biological replicates.

Chromatin immunoprecipitation (ChIP) assays and immunoprecipitation experiments

Immunoprecipitation was performed by standard procedures in CD14+ cells at 0, 2 and 4 days after treatment with M-CSF and RANKL. Cell extracts were prepared in 50 mM Tris–HCl, pH 7.5, 1 mM EDTA, 150 mM NaCl, 1% Triton-X-100 and protease cocktail inhibitors (Complete, Roche Molecular Biochemicals). Cellular extracts and samples from immunoprecipitation experiments were electrophoresed and western blotted following standard procedures.

For chromatin immunoprecipitation (ChIP) assays, CD14+ at 0, 2 and 4 days after treatment with M-CSF and RANKL were crosslinked with 1% formaldehyde and subjected to immunoprecipitation after sonication. ChIP experiments were performed as described [44]. Analysis was performed by real-time quantitative PCR. Data are represented as the ratio of the bound fraction over the input for each specific factor. We used a mouse monoclonal antibody against the TET2 N-t for ChIPs and a rabbit polyclonal antibody against TET2 for western blot. For DNMT3b we used a rabbit polyclonal against amino acids 1–230 of human DNMT3b (Santa Cruz Biotechnology). We also used a rabbit polyclonal against the C-t of PU.1 (sc-352, Santa Cruz Biotechnology), a rabbit polyclonal against the N-t of c-Fos (sc-52, Santa Cruz Biotechnology) and a rabbit polyclonal against the C-t of NF-kB p65 (sc-372, Santa Cruz Biotechnology). IgG was used as a negative control. Primer sequences were designed to contain either predicted or known TF binding (from TRANSFAC or ChIPseq data) as close as possible from the CpG undergoing methylation changes. Primer sequences are shown in Additional file 11. These experiments were performed with three biological replicates.

5hmC detection

5hmC was analyzed using the Quest 5-hmC Detection system (Zymo). Genomic DNA was treated with a specific 5hmC glucosyltransferase (GT) or left untreated (No GT, 0% 5hmC). DNA was then digested with MspI (100U) at 37°C overnight, followed by column purification. The MspI-resistant fraction (bearing the glucosile group, and therefore the original 5hmC) was quantitated by qPCR using primers designed around at least one MspI site (CCGG), and normalized to the amplification of the same region in the original DNA input. The amplification obtained in the untreated (no GT, MspI sensible) was then substracted to the samples in order to calculate the level of 0% 5hmC. The resulting values were the percentage of 5hmC present in each of the samples. Primer sequences are shown in Additional file 11.

Amplification of UnMethylated Alus (AUMA)

This method, aiming at amplifying unmethylated Alus, was performed as described [31,39]. Products were resolved on denaturing sequencing gels. Bands were visualized by silver

staining the gels. AUMA fingerprints were visually checked for methylation differences between bands in different samples.

Abbreviations

5azadC, 5-aza-2′-deoxycytidine; 5hmC, 5-hydroxymethylcytosine; 5mC, 5-methylcytosine; AUMA, Amplification of unmethylated Alu repeats

Competing interests

The authors declare that they have no competing interests

Author's contributions

LR and EB conceived experiments; LR, MG, JR-U and HH performed experiments; AI and JMU performed biocomputing analysis; LR, JMU, JC, KH, CGV and EB analyzed the data; EB wrote the paper. All authors read and approved the final manuscript.

Acknowledgements

We thank Dr Fátima Al-Shahrour and Dr Núria López-Bigas for helpful suggestions about the bioinformatic analyses, and Dr Mercedes Garayoa and Dr Antonio Garcia-Gomez for their helpful suggestions about protocols and providing antibodies. This work was supported by grant SAF2011-29635 from the Spanish Ministry of Science and Innovation, grant CIVP16A1834 from Fundación Ramón Areces and grant 2009SGR184 from AGAUR (Catalan Government). LR is supported by a PFIS predoctoral fellowship.

Data access

Methylation array data for this publication has been deposited in NCBI's Gene Expression Omnibus and is accessible through GEO Series accession number GSE46648.

References

1. Pham TH, Benner C, Lichtinger M, Schwarzfischer L, Hu Y, Andreesen R, Chen W, Rehli M: **Dynamic epigenetic enhancer signatures reveal key transcription factors associated with monocytic differentiation states.** *Blood* 2012, **119:**e161–e171.

2. Hogart A, Lichtenberg J, Ajay SS, Anderson S, Margulies EH, Bodine DM: **Genome-wide DNA methylation profiles in hematopoietic stem and progenitor cells reveal overrepresentation of ETS transcription factor binding sites.** *Genome Res* 2012, **22:**1407–1418.

3. Zhang JA, Mortazavi A, Williams BA, Wold BJ, Rothenberg EV: **Dynamic transformations of genome-wide epigenetic marking and transcriptional control establish T cell identity.** *Cell* 2012, **149:**467–482.

4. Blair HC, Teitelbaum SL, Ghiselli R, Gluck S: **Osteoclastic bone resorption by a polarized vacuolar proton pump.** *Science* 1989, **245:**855–857.

5. Wiktor-Jedrzejczak W, Bartocci A, Ferrante AW Jr, Ahmed-Ansari A, Sell KW, Pollard JW, Stanley ER: **Total absence of colony-stimulating factor 1 in the macrophage-deficient osteopetrotic (op/op) mouse.** *Proc Natl Acad Sci U S A* 1990, **87:**4828–4832.

6. Lacey DL, Timms E, Tan HL, Kelley MJ, Dunstan CR, Burgess T, Elliott R, Colombero A, Elliott G, Scully S, *et al*: **Osteoprotegerin ligand is a cytokine that regulates osteoclast differentiation and activation.** *Cell* 1998, **93:**165–176.

7. Saltel F, Chabadel A, Bonnelye E, Jurdic P: **Actin cytoskeletal organisation in osteoclasts: a model to decipher transmigration and matrix degradation.** *Eur J Cell Biol* 2008, **87:**459–468.

8. Wong BR, Besser D, Kim N, Arron JR, Vologodskaia M, Hanafusa H, Choi Y: **TRANCE, a TNF family member, activates Akt/PKB through a signaling complex involving TRAF6 and c-Src.** *Mol Cell* 1999, **4:**1041–1049.

9. Kobayashi N, Kadono Y, Naito A, Matsumoto K, Yamamoto T, Tanaka S, Inoue J: **Segregation of TRAF6-mediated signaling pathways clarifies its role in osteoclastogenesis.** *Embo J* 2001, **20:**1271–1280.

10. Blank U, Launay P, Benhamou M, Monteiro RC: **Inhibitory ITAMs as novel regulators of immunity.** *Immunol Rev* 2009, **232:**59–71.

11. Humphrey MB, Ogasawara K, Yao W, Spusta SC, Daws MR, Lane NE, Lanier LL, Nakamura MC: **The signaling adapter protein DAP12 regulates multinucleation during osteoclast development.** *J Bone Miner Res* 2004, **19:**224–234.

12. Koga T, Inui M, Inoue K, Kim S, Suematsu A, Kobayashi E, Iwata T, Ohnishi H, Matozaki T, Kodama T, *et al*: **Costimulatory signals mediated by the ITAM motif cooperate with RANKL for bone homeostasis.** *Nature* 2004, **428:**758–763.

13. Humphrey MB, Daws MR, Spusta SC, Niemi EC, Torchia JA, Lanier LL, Seaman WE, Nakamura MC: **TREM2, a DAP12-associated receptor, regulates osteoclast differentiation and function.** *J Bone Miner Res* 2006, **21:**237–245.

14. Negishi-Koga T, Takayanagi H: **Ca2 + −NFATc1 signaling is an essential axis of osteoclast differentiation.** *Immunol Rev* 2009, **231:**241–256.

15. Takayanagi H, Kim S, Koga T, Nishina H, Isshiki M, Yoshida H, Saiura A, Isobe M, Yokochi T, Inoue J, *et al*: **Induction and activation of the transcription factor NFATc1 (NFAT2) integrate RANKL signaling in terminal differentiation of osteoclasts.** *Dev Cell* 2002, **3:**889–901.

16. Sharma SM, Bronisz A, Hu R, Patel K, Mansky KC, Sif S, Ostrowski MC: **MITF and PU.1 recruit p38 MAPK and NFATc1 to target genes during osteoclast differentiation.** *J Biol Chem* 2007, **282:**15921–15929.

17. Kim K, Lee SH, Ha Kim J, Choi Y, Kim N: **NFATc1 induces osteoclast fusion via up-regulation of Atp6v0d2 and the dendritic cell-specific transmembrane protein (DC-STAMP).** *Mol Endocrinol* 2008, **22:**176–185.

18. Yu M, Moreno JL, Stains JP, Keegan AD: **Complex regulation of tartrate-resistant acid phosphatase (TRAP) expression by interleukin 4 (IL-4): IL-4 indirectly suppresses receptor activator of NF-kappaB ligand (RANKL)-mediated TRAP expression but modestly induces its expression directly.** *J Biol Chem* 2009, **284:**32968–32979.

19. Matsumoto M, Kogawa M, Wada S, Takayanagi H, Tsujimoto M, Katayama S, Hisatake K, Nogi Y: **Essential role of p38 mitogen-activated protein kinase in cathepsin K gene expression during osteoclastogenesis through association of NFATc1 and PU.1.** *J Biol Chem* 2004, **279:**45969–45979.

20. Sundaram K, Nishimura R, Senn J, Youssef RF, London SD, Reddy SV: **RANK ligand signaling modulates the matrix metalloproteinase-9 gene expression during osteoclast differentiation.** *Exp Cell Res* 2007, **313:**168–178.

21. Tolar J, Teitelbaum SL, Orchard PJ: **Osteopetrosis.** *N Engl J Med* 2004, **351:**2839–2849.

22. Rachner TD, Khosla S, Hofbauer LC: **Osteoporosis: now and the future.** *Lancet* 2011, **377:**1276–1287.

23. Scott DL, Wolfe F, Huizinga TW: **Rheumatoid arthritis.** *Lancet* 2010, **376:**1094–1108.

24. Mundy GR, Raisz LG, Cooper RA, Schechter GP, Salmon SE: **Evidence for the secretion of an osteoclast stimulating factor in myeloma.** *N Engl J Med* 1974, **291:**1041–1046.

25. Yoneda T: **Cellular and molecular mechanisms of breast and prostate cancer metastasis to bone.** *Eur J Cancer* 1998, **34:**240–245.

26. Mii Y, Miyauchi Y, Morishita T, Miura S, Honoki K, Aoki M, Tamai S: **Osteoclast origin of giant cells in giant cell tumors of bone: ultrastructural and cytochemical study of six cases.** *Ultrastruct Pathol* 1991, **15:**623–629.

27. Nicholson GC, Malakellis M, Collier FM, Cameron PU, Holloway WR, Gough TJ, Gregorio-King C, Kirkland MA, Myers DE: **Induction of osteoclasts from CD14-positive human peripheral blood mononuclear cells by receptor activator of nuclear factor kappaB ligand (RANKL).** *Clin Sci (Lond)* 2000, **99:**133–140.

28. Sorensen MG, Henriksen K, Schaller S, Henriksen DB, Nielsen FC, Dziegiel MH, Karsdal MA: **Characterization of osteoclasts derived from CD14+ monocytes isolated from peripheral blood.** *J Bone Miner Metab* 2007, **25:**36–45.

29. Youn MY, Takada I, Imai Y, Yasuda H, Kato S: **Transcriptionally active nuclei are selective in mature multinucleated osteoclasts.** *Genes Cells* 2010, **15:**1025–1035.

30. Saltman LH, Javed A, Ribadeneyra J, Hussain S, Young DW, Osdoby P, Amcheslavsky A, van Wijnen AJ, Stein JL, Stein GS, *et al*: **Organization of transcriptional regulatory**

machinery in osteoclast nuclei: compartmentalization of Runx1. *J Cell Physiol* 2005, **204:**871–880.

31. Rodriguez J, Vives L, Jorda M, Morales C, Munoz M, Vendrell E, Peinado MA: **Genome-wide tracking of unmethylated DNA Alu repeats in normal and cancer cells.** *Nucleic Acids Res* 2008, **36:**770–784.

32. Gallois A, Lachuer J, Yvert G, Wierinckx A, Brunet F, Rabourdin-Combe C, Delprat C, Jurdic P, Mazzorana M: **Genome-wide expression analyses establish dendritic cells as a new osteoclast precursor able to generate bone-resorbing cells more efficiently than monocytes.** *J Bone Miner Res* 2010, **25:**661–672.

33. Momparler RL: **Pharmacology of 5-Aza-2′-deoxycytidine (decitabine).** *Semin Hematol* 2005, **42:**S9–S16.

34. Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, Agarwal S, Iyer LM, Liu DR, Aravind L, *et al*: **Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1.** *Science* 2009, **324:**930–935.

35. Kriaucionis S, Heintz N: **The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain.** *Science* 2009, **324:**929–930.

36. Ito S, Shen L, Dai Q, Wu SC, Collins LB, Swenberg JA, He C, Zhang Y: **Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine.** *Science* 2011, **333:**1300–1303.

37. Wu H, Zhang Y: **Mechanisms and functions of Tet protein-mediated 5-methylcytosine oxidation.** *Genes Dev* 2011, **25:**2436–2452.

38. Williams K, Christensen J, Helin K: **DNA methylation: TET proteins-guardians of CpG islands?** *EMBO Rep* 2012, **13:**28–35.

39. Hernando H, Shannon-Lowe C, Islam AB, Al-Shahrour F, Rodriguez-Ubreva J, Rodriguez-Cortez VC, Javierre BM, Mangas C, Fernandez AF, Parra M, *et al*: **The B cell transcription program mediates hypomethylation and overexpression of key genes in Epstein-Barr virus-associated proliferative conversion.** *Genome Biol* 2013, **14:**R3.

40. Stadler MB, Murr R, Burger L, Ivanek R, Lienert F, Scholer A, van Nimwegen E, Wirbelauer C, Oakeley EJ, Gaidatzis D, *et al*: **DNA-binding factors shape the mouse methylome at distal regulatory regions.** *Nature* 2012, **480:**490–495.

41. Edwards JR, Mundy GR: **Advances in osteoclast biology: old findings and new insights from mouse models.** *Nat Rev Rheumatol* 2011, **7:**235–243.

42. Shen L, Wu H, Diep D, Yamaguchi S, D'Alessio AC, Fung HL, Zhang K, Zhang Y: **Genome-wide analysis reveals TET- and TDG-dependent 5-methylcytosine oxidation dynamics.** *Cell* 2013, **153:**692–706.

43. Cortellino S, Xu J, Sannai M, Moore R, Caretti E, Cigliano A, Le Coz M, Devarajan K, Wessels A, Soprano D, *et al*: **Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair.** *Cell* 2011, **146:**67–79.

44. Kallin EM, Rodriguez-Ubreva J, Christensen J, Cimmino L, Aifantis I, Helin K, Ballestar E, Graf T: **Tet2 facilitates the derepression of myeloid target genes during CEBPalpha-induced transdifferentiation of pre-B cells.** *Mol Cell* 2012, **48:**266–276.

45. Klug M, Schmidhofer S, Gebhard C, Andreesen R, Rehli M: **5-Hydroxymethylcytosine is an essential intermediate of active DNA demethylation processes in primary human monocytes.** *Genome Biol* 2013, **14:**R46.

46. Suzuki M, Yamada T, Kihara-Negishi F, Sakurai T, Hara E, Tenen DG, Hozumi N, Oikawa T: **Site-specific DNA methylation by a complex of PU.1 and Dnmt3a/b.** *Oncogene* 2006, **25:**2477–2488.

47. Chen W, Zhu G, Hao L, Wu M, Ci H, Li YP: **C/EBPalpha regulates osteoclast lineage commitment.** *Proc Natl Acad Sci U S A* 2013.

48. Hama M, Kirino Y, Takeno M, Takase K, Miyazaki T, Yoshimi R, Ueda A, Itoh-Nakadai A, Muto A, Igarashi K, *et al*: **Bach1 regulates osteoclastogenesis in a mouse model via both heme oxygenase 1-dependent and heme oxygenase 1-independent pathways.** *Arthritis Rheum* 2012, **64:**1518–1528.

49. Stopka T, Amanatullah DF, Papetti M, Skoultchi AI: **PU.1 inhibits the erythroid program by binding to GATA-1 on DNA and creating a repressive chromatin structure.** *Embo J* 2005, **24:**3712–3723.

50. Hu R, Sharma SM, Bronisz A, Srinivasan R, Sankar U, Ostrowski MC: **Eos, MITF, and PU.1 recruit corepressors to osteoclast-specific genes in committed myeloid progenitors.** *Mol Cell Biol* 2007, **27:**4018–4027.

51. Vincent JJ, Huang Y, Chen PY, Feng S, Calvopina JH, Nee K, Lee SA, Le T, Yoon AJ, Faull K, *et al*: **Stage-specific roles for tet1 and tet2 in DNA demethylation in primordial germ cells.** *Cell Stem Cell* 2013, **12:**470–478.

52. Azim AC, Wang X, Park GY, Sadikot RT, Cao H, Mathew B, Atchison M, van Breemen RB, Joo M, Christman JW: **NF-kappaB-inducing kinase regulates cyclooxygenase 2 gene expression in macrophages by phosphorylation of PU.1.** *J Immunol* 2007, **179:**7868–7875.

53. Bibikova M, Lin Z, Zhou L, Chudin E, Garcia EW, Wu B, Doucet D, Thomas NJ, Wang Y, Vollmer E, *et al*: **High-throughput DNA methylation profiling using universal bead arrays.** *Genome Res* 2006, **16:**383–393.

54. Du P, Zhang X, Huang CC, Jafari N, Kibbe WA, Hou L, Lin SM: **Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis.** *BMC Bioinformatics* 2010, **11:**587.

55. The-R-Development-Core-Team: *R: A language and environment for statistical computing.* R Foundation for Statistical Computing; 2012.

56. Smyth GK: **Limma: linear models for microarray data**. In *'Bioinformatics and Computational Biology Solutions using R and Bioconductor'.* Edited by Gentleman R, Carey V, Dudoit S, Irizarry R, Huber W. New York: Springer; 2005:397–420.

57. Du P, Kibbe WA, Lin SM: **Lumi: a pipeline for processing Illumina microarray.** *Bioinformatics* 2008, **24:**1547–1548.

58. Wang D, Yan L, Hu Q, Sucheston LE, Higgins MJ, Ambrosone CB, Johnson CS, Smiraglia DJ, Liu S: **IMA: an R package for high-throughput analysis of Illumina's 450K Infinium methylation data.** *Bioinformatics* 2012, **28:**729–730.

59. Aryee MJ, Wu Z, Ladd-Acosta C, Herb B, Feinberg AP, Yegnasubramanian S, Irizarry RA: **Accurate genome-scale percentage DNA methylation estimates from microarray data.** *Biostatistics* 2011, **12:**197–210.

60. Herman JG, Graff JR, Myohanen S, Nelkin BD, Baylin SB: **Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands.** *Proc Natl Acad Sci U S A* 1996, **93:**9821–9826.

61. www.ebi.ac.uk/arrayexpress.

62. Gautier L, Cope L, Irizarry RA, Bolstad BM: **Affy–analysis of Affymetrix GeneChip data at the probe level.** *Bioinformatics* 2004, **20:**307–315.

63. Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B, Speed TP: **Summaries of Affymetrix GeneChip probe level data.** *Nucleic Acids Res* 2003, **31:**e15.

64. Al-Shahrour F, Diaz-Uriarte R, Dopazo J: **FatiGO: a web tool for finding significant associations of Gene Ontology terms with groups of genes.** *Bioinformatics* 2004, **20:**578–580.

65. Schones DE, Smith AD, Zhang MQ: **Statistical significance of cis-regulatory modules.** *BMC Bioinformatics* 2007, **8:**19.

66. Matys V, Fricke E, Geffers R, Gossling E, Haubrock M, Hehl R, Hornischer K, Karas D, Kel AE, Kel-Margoulis OV, *et al*: **TRANSFAC: transcriptional regulation, from patterns to profiles.** *Nucleic Acids Res* 2003, **31:**374–378.

67. Perez-Llamas C, Lopez-Bigas N: **Gitools: analysis and visualisation of genomic data using interactive heat-maps.** *PLoS One* 2011, **6:**e19541.

68. www.gitools.org.

69. Flicek P, Amode MR, Barrell D, Beal K, Brent S, Carvalho-Silva D, Clapham P, Coates G, Fairley S, Fitzgerald S, *et al*: **Ensembl 2012.** *Nucleic Acids Res* 2012, **40:**D84–90.

70. Smink JJ, Begay V, Schoenmaker T, Sterneck E, de Vries TJ, Leutz A: **Transcription factor C/EBPbeta isoform ratio regulates osteoclastogenesis through MafB.** *Embo J* 2009, **28:**1769–1781.

Additional files

Additional_file_1 as PDF

Additional file 1 M-CSF and RANKL-induced monocyte-to-osteoclast differentiation. (A) Visualization of the formation of the actin ring and the generation of polykaryons in monocyte (MO) to osteoclast (OC) differentiation with phalloidin and DAPI staining. (B) TRAP (Tartrate resistant acid phosphatase-OC marker) staining in MO and OC preparations, showing this activity only in OCs. (B) Determination of the typical percentage of osteoclastic nuclei present in the preparations used for the experiments; over 84% efficiency was achieved at 21 days. (C) Upregulation of OC specific markers (*CA2*, *CTSK*, *MMP9*, *ACP5*) was checked by qPCR; downregulation of a monocyte specific gene (*CX3CR1*) was also monitored.

Additional_file_2 as XLSX

Additional file 2 List of hypomethylated and hypermethylated genes during monocyte to osteoclast differentiation (FC < 0.5 (hypomethylated, sheet 1) or FC > 2.

Additional_file_3 as PDF

Additional file 3 (A) Scatterplots showing DNA methylation profiles of matching MO/OC pairs. Genes with significant differences (FC > 2, FDR < 0.05) in averaged results from three samples are highlighted in red (hypermethylated) or blue (hypomethylated). Three panels corresponding for each of the three individual comparisons of MO/OC pairs (D1, D2 and D3) are shown. (B) Bisulphite sequencing analysis of repetitive sequences performed on monocytes (day 0) and osteoclasts (day 21) from three different donors (donor A, donor B and donor C), showing no relevant differences in the DNA methylation levels. (C) AUMA (Amplification of Unmethylated Alus) analysis of two independent monocyte-to-osteoclast differentiation experiments. Graphs correspond to the scanned intensities of the bands obtained with two different sets of primers. No significant differences are observed.

Additional_file_4 as DOCX

Additional file 4 Individual raw data corresponding to bisulfite pyrosequencing and standard bisulfite sequencing of individual MO and OC samples (Figure 1E), time course methylation data (Figure 2D, E) and PU.1 siRNA experiments (Figure 5D). Data are presented as supplied by PyroMark® Assay Design Software 2.0 for PyroMark Q96 MD (Qiagen), which automatically generates methylation percentages in a datasheet format.

Additional_file_5 as DOCX

Additional file 5 Clusters of consecutive CpGs hypomethylated (−) or hypermethylated (+) in OC vs MO.

Additional_file_6 as XLSX

Additional file 6 Differentially expressed genes between Mos, OC samples at 5 days and OC samples at 20 days after RANKL/M-CSF stimulation (FC > 2, FC < 0.5; FDR < 0.05).

Additional_file_7 as XLSX

Additional file 7 List of genes with an inverse relationship between DNA methylation and expression change (FC < 0.5 or FC > 2; FDR < 0.05 for both DNA methylation and expression data).

Additional_file_8 as PDF

Additional file 8 (A) Scheme showing the BrdU pulses added to monocytes differentiating into osteoclasts. (B) Representative immunofluorescence images at the selected time points showing BrdU positive cells. (C) Representation of the time scale where DNA demethylation occurs during osteoclast differentiation, together with the cell division observed at later time points.

Additional_file_9 as PDF

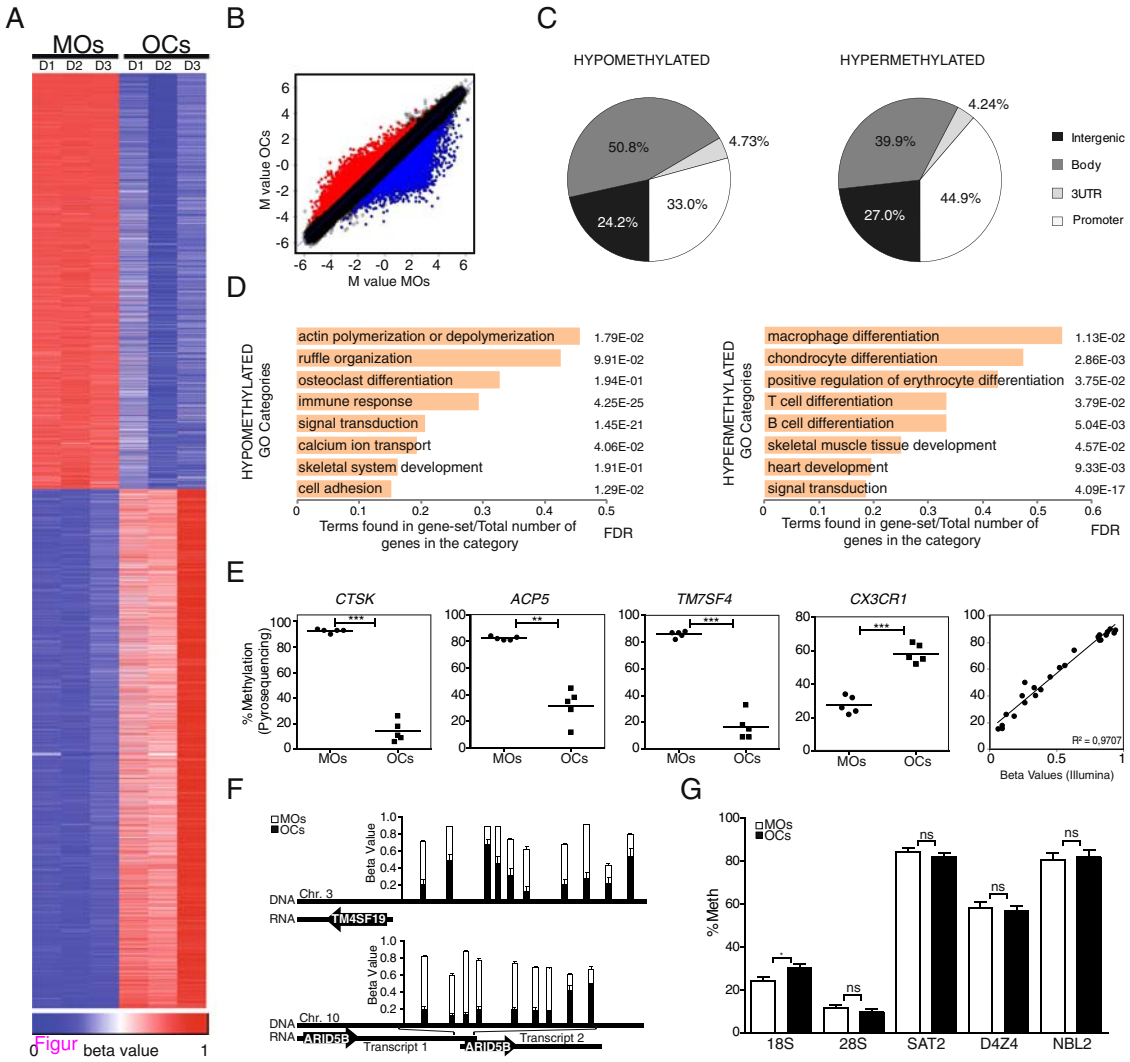
Additional file 9 Osteoclast differentiation scheme showing transcription factors that are known to be involved in monocyte-to-osteoclast differentiation. We have in red or blue the presence of binding motifs for those factors (according to TRANSFAC analysis) among the sequences surrounding the CpGs that become hypo- or hypermethylated. Those arising from our analysis are highlighted in red and blue (associated with hypermethylation and hypomethylation, respectively).

Additional_file_10 as PDF

Additional file 10 (A) ChIP assays showing the effects of PU.1 downregulation in its recruitment, together with TET2 and DNMT3b binding to the same genes. Data were obtained at 0, 2 and 6 days after M-CSF/RANL stimulation. (B) We have used the MYOD1 promoter as a negative control. (C) Effects of PU.1 downregulation on expression and methylation of PU.1-target gene TM7SF4.

Additional_file_11 as XLSX

Additional file 11 List of primers.


A
MOs
OCs
D1 D2 D3 D1 D2 D3
0
beta value
1
B
M value OCs
M value MOs
C
HYPOMETHYLATED
50.8%
4.73%
33.0%
24.2%
HYPERMETHYLATED
4.24%
39.9%
44.9%
27.0%
Intergenic
Body
3UTR
Promoter
D
HYPOMETHYLATED GO Categories
actin polymerization or depolymerization 1.79E-02
ruffle organization 9.91E-02
osteoclast differentiation 1.94E-01
immune response 4.25E-25
signal transduction 1.45E-21
calcium ion transport 4.06E-02
skeletal system development 1.91E-01
cell adhesion 1.29E-02
Terms found in gene-set/Total number of genes in the category
FDR
HYPERMETHYLATED GO Categories
macrophage differentiation 1.13E-02
chondrocyte differentiation 2.86E-03
positive regulation of erythrocyte differentiation 3.75E-02
T cell differentiation 3.79E-02
B cell differentiation 5.04E-03
skeletal muscle tissue development 4.57E-02
heart development 9.33E-03
signal transduction 4.09E-17
Terms found in gene-set/Total number of genes in the category
FDR
E
CTSK
ACP5
TM7SF4
CX3CR1
% Methylation (Pyrosequencing)
MOs
OCs
R² = 0,9707
Beta Values (Illumina)
F
MOs
OCs
Beta Value
DNA
RNA
Chr. 3
TM4SF19
Chr. 10
ARID5B
Transcript 1
Transcript 2
G
%Meth
18S
28S
SAT2
D4Z4
NBL2
ns

Figur

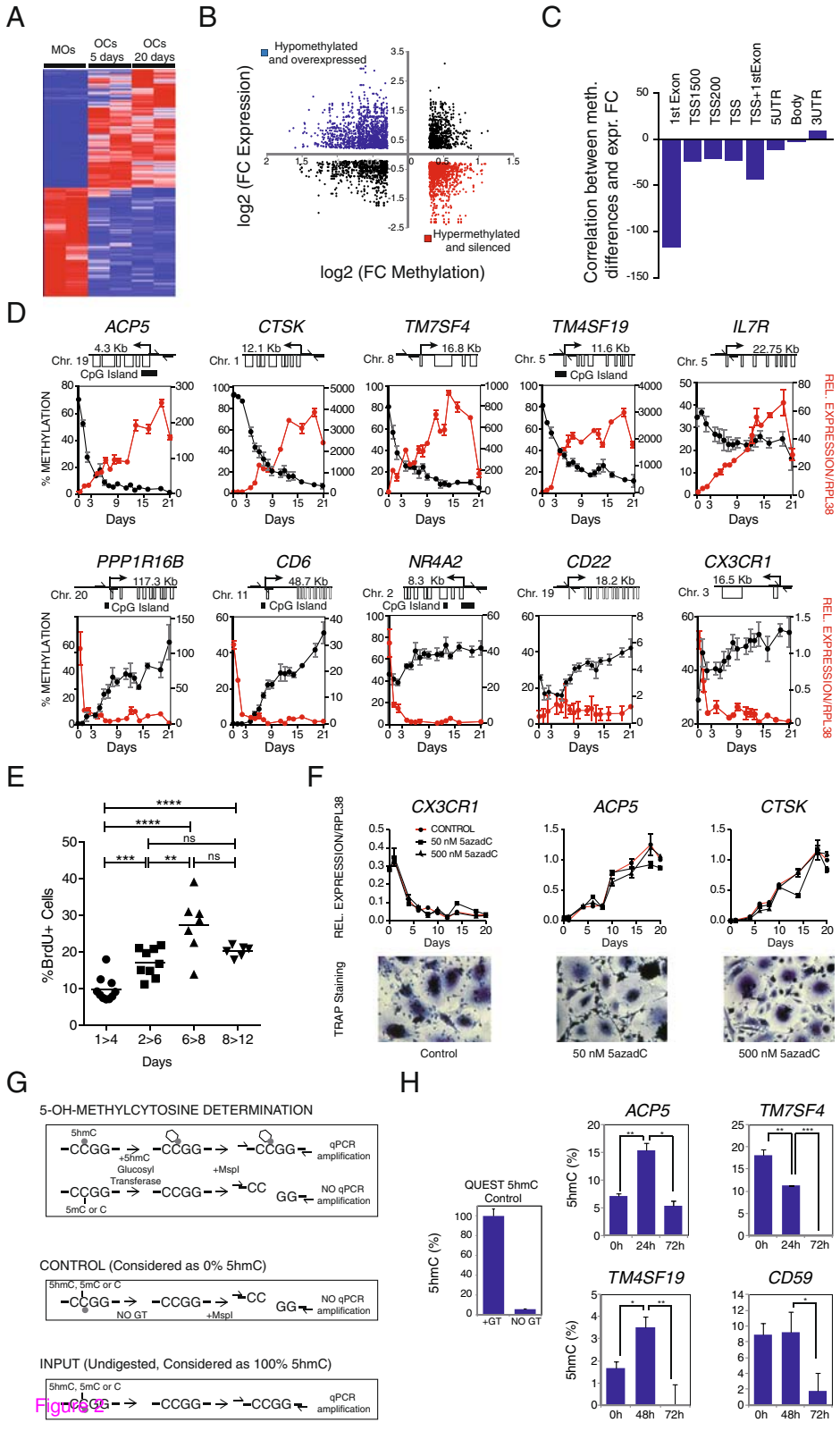

A
MOs
OCs 5 days
OCs 20 days
B
Hypomethylated and overexpressed
Hypermethylated and silenced
log2 (FC Expression)
log2 (FC Methylation)
C
Correlation between meth. differences and expr. FC
1st Exon
TSS1500
TSS200
TSS
TSS+1stExon
5UTR
Body
3UTR
D
ACP5
CTSK
TM7SF4
TM4SF19
IL7R
PPP1R16B
CD6
NR4A2
CD22
CX3CR1
% METHYLATION
REL. EXPRESSION/RPL38
CpG Island
Days
E
%BrdU+ Cells
1>4
2>6
6>8
8>12
Days
F
CX3CR1
ACP5
CTSK
CONTROL
50 nM 5azadC
500 nM 5azadC
REL. EXPRESSION/RPL38
TRAP Staining
Control
50 nM 5azadC
500 nM 5azadC
G
5-OH-METHYLCYTOSINE DETERMINATION
5hmC
+5hmC
Glucosyl Transferase
+MspI
qPCR amplification
NO qPCR amplification
5mC or C
CONTROL (Considered as 0% 5hmC)
NO GT
INPUT (Undigested, Considered as 100% 5hmC)
H
QUEST 5hmC Control
5hmC (%)
+GT
NO GT
ACP5
TM7SF4
TM4SF19
CD59
0h
24h
48h
72h

Figure 2

A

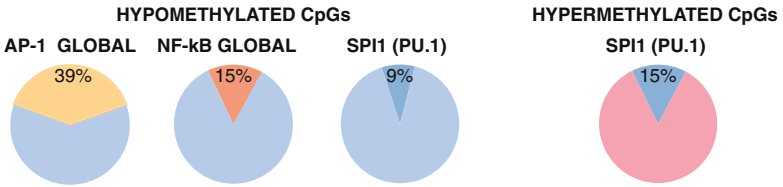
HYPOMETHYLATED CpGs

TRANSFAC binding motif	p-value	TF family
V$AP1_C	1.03E-211	AP-1
V$BACH1_01	4.41E-190	
V$AP1_Q6_01	4.69E-183	AP-1
V$BACH2_01	5.02E-167	
V$NFKB_Q6_01	1.36E-40	NF-kB
V$NFKAPPAB_01	6.11E-39	NF-kB
V$NFKAPPAB65_01	1.05E-33	NF-kB
V$NFKB_C	6.83E-29	NF-kB
V$CMAF_01	1.42E-15	
V$VMAF_01	4.61E-15	
V$PU1_Q4	4.25E-14	ETS
V$ISRE_01	6.13E-14	

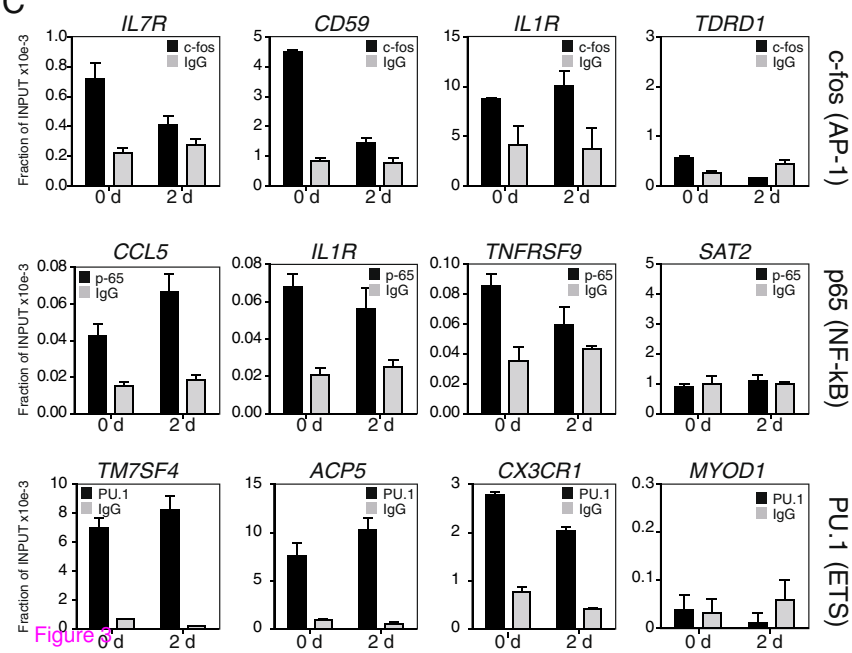
HYPERMETHYLATED CpGs

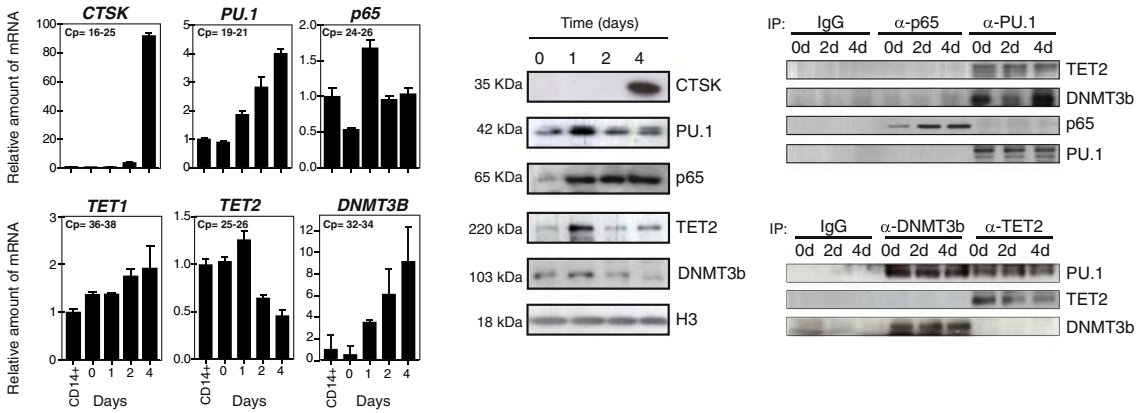
TRANSFAC binding motif	p-value	TF family
V$PU1_Q4	5.05E-67	ETS
V$SPIB_01	5.93E-50	ETS
V$PU1_01	4.84E-46	ETS
V$ESE1_Q3	1.84E-43	
V$ETS2_B	5.38E-35	ETS
V$ETS1_B	4.33E-26	ETS
V$CEBPB_02	6.51E-22	CEBP
V$ETS_Q4	1.10E-21	ETS
V$ELF1_Q6	2.71E-15	
V$CEBPB_01	1.35E-14	CEBP
V$CEBPA_01	1.36E-13	CEBP
V$CEBP_Q2_01	2.64E-12	CEBP

B

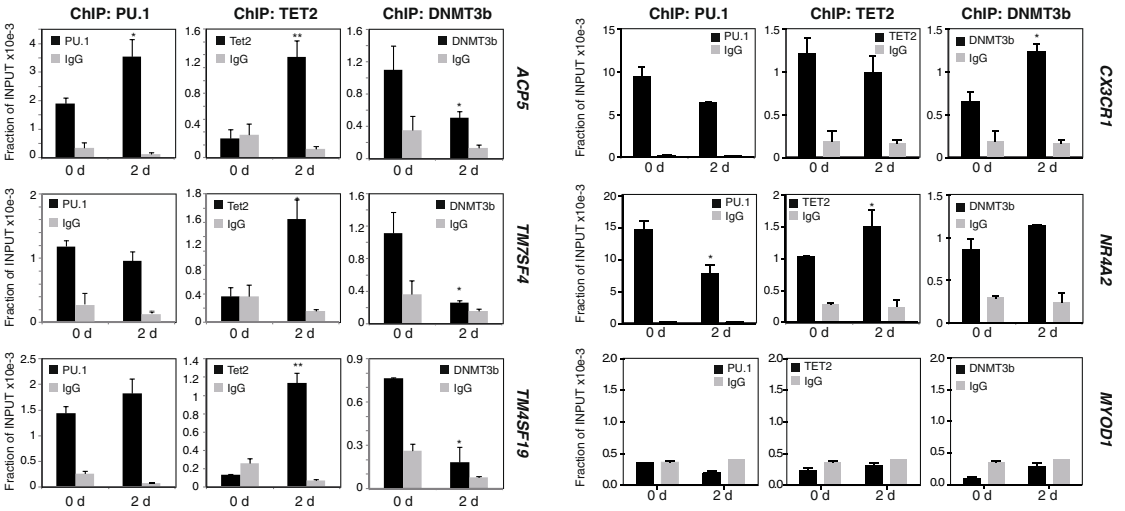


C

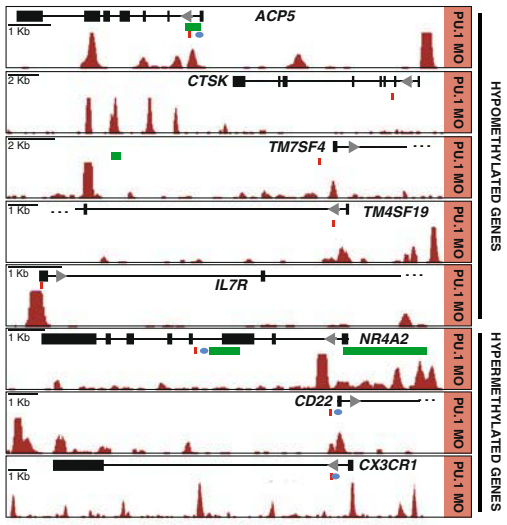



CTSK
PU.1
p65
TET1
TET2
DNMT3B
Relative amount of mRNA
Cp= 16-25
Cp= 19-21
Cp= 24-26
Cp= 36-38
Cp= 25-26
Cp= 32-34
CD14+ 0 1 2 4
Days
Time (days)
0 1 2 4
35 KDa
42 KDa
65 KDa
220 KDa
103 KDa
18 KDa
CTSK
PU.1
p65
TET2
DNMT3b
H3
IP: IgG α-p65 α-PU.1
0d 2d 4d 0d 2d 4d 0d 2d 4d
TET2
DNMT3b
p65
PU.1
IP: IgG α-DNMT3b α-TET2
0d 2d 4d 0d 2d 4d 0d 2d 4d
PU.1
TET2
DNMT3b

D


ChIP: PU.1
ChIP: TET2
ChIP: DNMT3b
Fraction of INPUT x10e-3
PU.1
Tet2
DNMT3b
IgG
0 d
2 d
ACP5
TM7SF4
TM4SF19
CX3CR1
NR4A2
MYOD1

E


ACP5
CTSK
TM7SF4
TM4SF19
IL7R
NR4A2
CD22
CX3CR1
1 Kb
2 Kb
PU.1 MO
HYPOMETHYLATED GENES
HYPERMETHYLATED GENES
CpG Island
Transfac predicted binding site
Illumina probe (Pyrosequenced region)
PU.1 peak in ChIP-Seq (Monocytes)

F

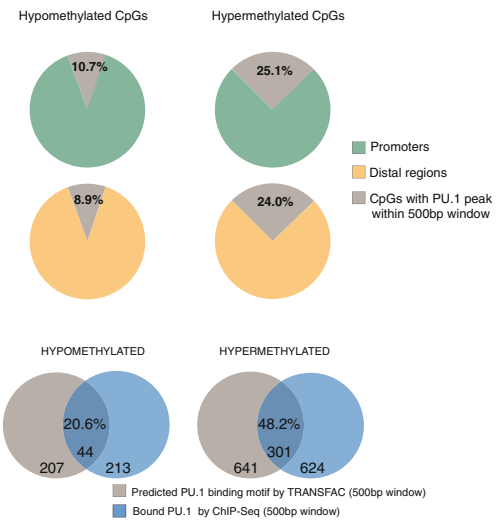

Hypomethylated CpGs
Hypermethylated CpGs
10.7%
25.1%
8.9%
24.0%
Promoters
Distal regions
CpGs with PU.1 peak within 500bp window
HYPOMETHYLATED
HYPERMETHYLATED
20.6%
44
207
213
48.2%
301
641
624
Predicted PU.1 binding motif by TRANSFAC (500bp window)
Bound PU.1 by ChIP-Seq (500bp window)

Figure 4

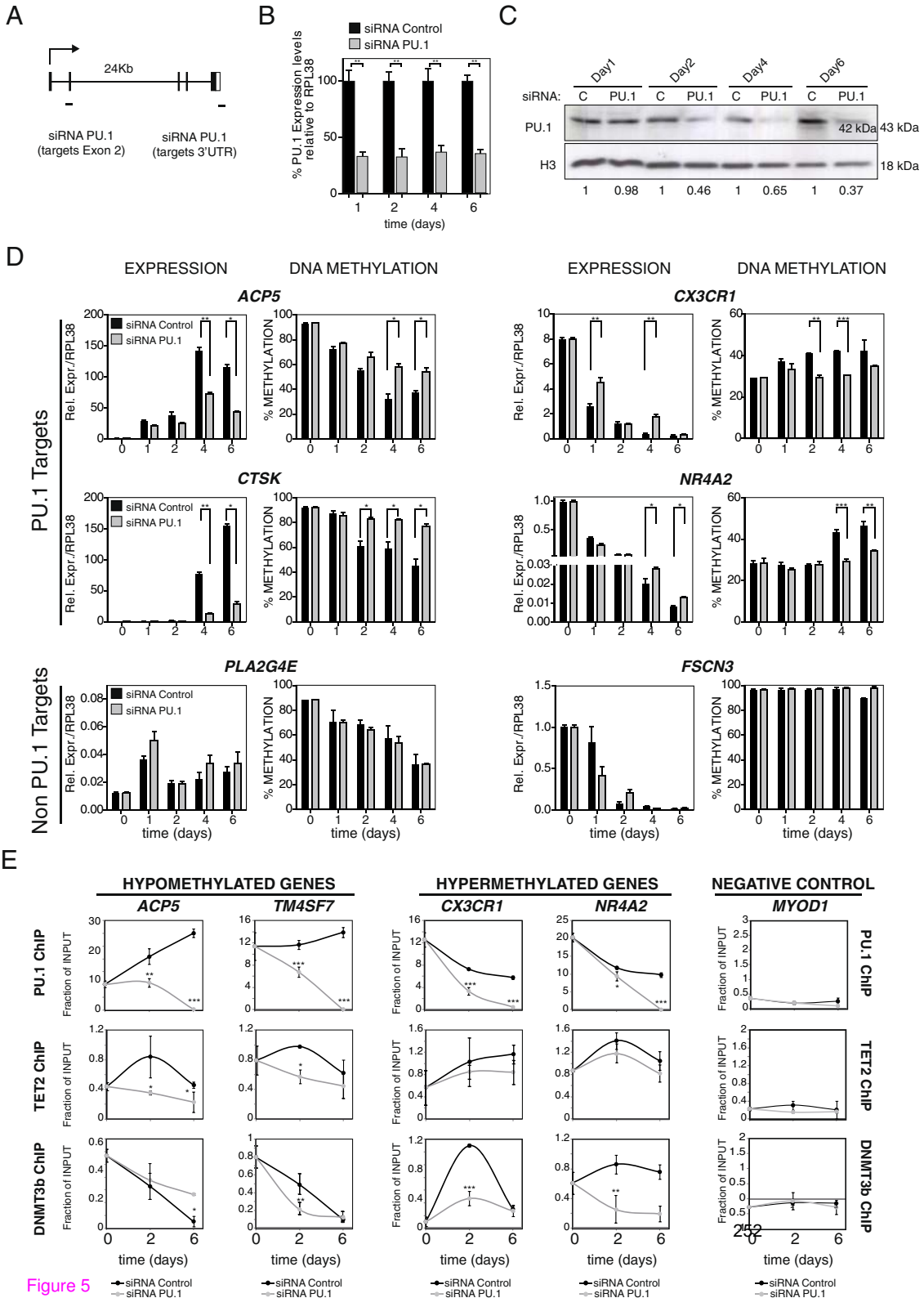

A
24Kb
siRNA PU.1 (targets Exon 2)
siRNA PU.1 (targets 3'UTR)
B
siRNA Control
siRNA PU.1
% PU.1 Expression levels relative to RPL38
time (days)
C
Day1
Day2
Day4
Day6
siRNA:
C
PU.1
PU.1
H3
42 kDa
43 kDa
18 kDa
1 0.98 1 0.46 1 0.65 1 0.37
D
EXPRESSION
DNA METHYLATION
PU.1 Targets
Non PU.1 Targets
ACP5
CX3CR1
CTSK
NR4A2
PLA2G4E
FSCN3
Rel. Expr./RPL38
% METHYLATION
time (days)
E
HYPOMETHYLATED GENES
HYPERMETHYLATED GENES
NEGATIVE CONTROL
ACP5
TM4SF7
CX3CR1
NR4A2
MYOD1
PU.1 ChIP
TET2 ChIP
DNMT3b ChIP
Fraction of INPUT
time (days)
siRNA Control
siRNA PU.1

Figure 5

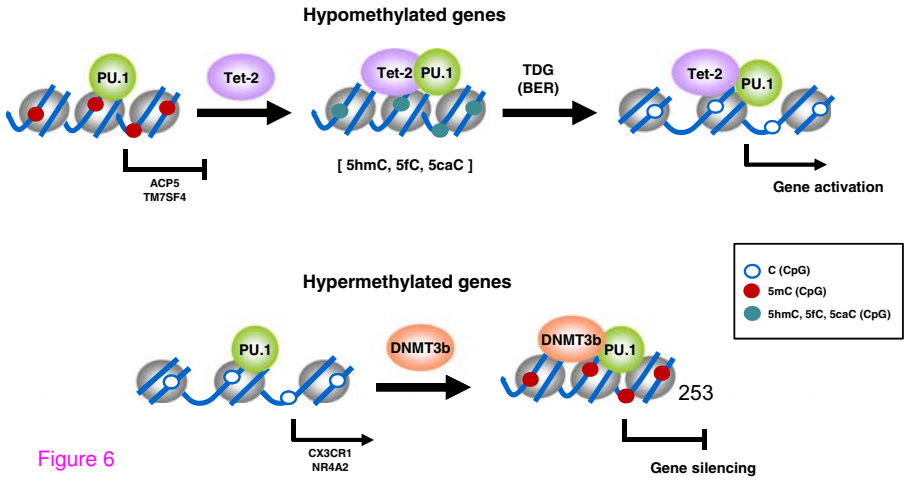

Hypomethylated genes
PU.1
Tet-2
Tet-2
PU.1
TDG
(BER)
Tet-2
PU.1
ACP5
TM7SF4
[5hmC, 5fC, 5caC]
Gene activation
C (CpG)
5mC (CpG)
5hmC, 5fC, 5caC (CpG)
Hypermethylated genes
PU.1
DNMT3b
DNMT3b
PU.1
CX3CR1
NR4A2
Gene silencing

Figure 6

Additional files provided with this submission:

Additional file 1: 1732670654987478_add1.pdf, 6632K
http://genomebiology.com/imedia/6049231811079634/supp1.pdf
Additional file 2: 1732670654987478_add2.xlsx, 1865K
http://genomebiology.com/imedia/1159981979107963/supp2.xlsx
Additional file 3: 1732670654987478_add3.pdf, 4235K
http://genomebiology.com/imedia/7107845610796341/supp3.pdf
Additional file 4: 1732670654987478_add4.docx, 75K
http://genomebiology.com/imedia/2061154675107963/supp4.docx
Additional file 5: 1732670654987478_add5.docx, 124K
http://genomebiology.com/imedia/1606511482107963/supp5.docx
Additional file 6: 1732670654987478_add6.xlsx, 2118K
http://genomebiology.com/imedia/1841882703107963/supp6.xlsx
Additional file 7: 1732670654987478_add7.xlsx, 368K
http://genomebiology.com/imedia/2125593161107963/supp7.xlsx
Additional file 8: 1732670654987478_add8.pdf, 12211K
http://genomebiology.com/imedia/5942374671079634/supp8.pdf
Additional file 9: 1732670654987478_add9.pdf, 364K
http://genomebiology.com/imedia/5033776471079634/supp9.pdf
Additional file 10: 1732670654987478_add10.pdf, 474K
http://genomebiology.com/imedia/1962635871107963/supp10.pdf
Additional file 11: 1732670654987478_add11.xlsx, 17K
http://genomebiology.com/imedia/2513916061079634/supp11.xlsx

Leukemia Research 36 (2012) 895–899

Contents lists available at SciVerse ScienceDirect

Leukemia Research

journal homepage: www.elsevier.com/locate/leukres

Epigenetic regulation of *PRAME* in acute myeloid leukemia is different compared to CD34+ cells from healthy donors: Effect of 5-AZA treatment

S. Gutierrez-Cosío[a], L. de la Rica[b], E. Ballestar[b], C. Santamaría[a], L.I. Sánchez-Abarca[a,c], T. Caballero-Velazquez[a,c], B. Blanco[a], C. Calderón[c], C. Herrero-Sánchez[a], S. Carrancio[a], L. Ciudad[b], C. Cañizo[a], J.F. San Miguel[a], J.A. Pérez-Simón[a,c,*]

[a] *Hematology Unit, Hospital Universitario de Salamanca, Salamanca, Spain*
[b] *Bellvitge Biomedical Research Institute (IDIBELL), Barcelona, Spain*
[c] *Hematology Unit, Hospital Universitario Virgen del Rocío and Instituto de Biomedicina de Sevilla (IBIS), Sevilla, Spain*

ARTICLE INFO

Article history:
Received 31 October 2011
Received in revised form 21 February 2012
Accepted 27 February 2012
Available online 13 April 2012

Keywords:
Preferentially expressed antigen of melanoma (PRAME)
5-Azacytidine (5-aza)
Tumor associated antigen (TAA)
Myeloid leukemia
Epigenetic regulation

ABSTRACT

PRAME is a tumor associated antigen (TAA) of particular interest since it is widely expressed by lymphoid and myeloid malignancies. Several studies have associated high *PRAME* RNA levels with good prognosis in acute myeloid leukemia (AML). *PRAME* expression is regulated at the epigenetic level. For this reason inhibitors of DNA methylation, such as 5-azacytidine, can modulate the expression of this TAAs. In the current study we analyzed the effect of 5-azaC on the expression of *PRAME* in blasts versus CD34+ cells from healthy donors in an attempt to increase its expression, thus inducing a potential target for therapeutic strategies.

1. Introduction

Preferentially expressed antigen of melanoma (*PRAME*) was first isolated as a human melanoma antigen by cDNA expression cloning using melanoma-reactive cytotoxic Tcells (CTL) [1,2]. *PRAME* is a tumor associated antigen (TAA) of particular interest since it is widely expressed by lymphoid and myeloid malignancies [3,4] and solid tumors, including melanomas, sarcomas, head and neck cancers, small-cell lung carcinomas and renal cell cancers [1,2]. In normal tissues, *PRAME* expression has been reported in testis and low levels are found in endometrium, ovaries and adrenals [1,2].

Regarding myeloid leukemias, it has been reported that *PRAME* is overexpressed in acute myeloid leukemia (AML) compared with blood and bone marrow healthy donors [4–8]. These studies have associated high *PRAME* RNA levels with good prognosis in some AML subtypes, especially those with favorable cytogenetics, i.e., *t*(8;21) [7] or *t*(15;17) [5] and normal karyotype [8]. In addition, several authors have suggested that *PRAME* could be used as a target for anticancer T-cell therapy [9,10]. However, Steinbach et al. have shown that *PRAME* gene is also expressed, although at a lower intensity, in CD34+ stem cells from healthy donors, which might constitute a problem for its application as a target in tumor immunotherapy [11].

Epigenetic events represent the main mechanism regulating the expression of *PRAME* including DNA methylation of several promoter regions [12,13]. In fact, a correlation between hypomethylated CpG dinucleotides in TAA promoters (i.e. *MAGE*, *GAGE* or *PRAME*) and their overexpression has been found in neoplastic cell lines and tissues [13–15] however, no information on both selected cells from patients with AML and CD34+ cells from healthy donors are available. Inhibitors of DNA methylation, such as 5-azacytidine [16,17] can reverse this epigenetic event suggesting a potential use in cancer therapy by inducing TAAs expression. Perhaps you can introduce here a comment on the use of DNA methyltransferase inhibitors in the treatment of myeloid malignances and its implications in relation with *PRAME*.

In the current study, we have compared the expression and methylation pattern of *PRAME* in blast cells from AML patients versus selected CD34+ stem cells from healthy donors. We found that the promoter region is highly methylated in normal CD34+ cells compared to blasts and this pattern correlates with higher expression of *PRAME* in blasts. Further, after treating cells with

* Corresponding author at: Hematology Unit, Instituto de Biomedicina de Sevilla (IBIS) Hospital Universitario Virgen del Rocío/CSIC/Universidad de Sevilla, Avda Manuel Siurot s/n 41013, Sevilla, Spain. Tel.: +34 955 013260; fax: +34 955 013265.
E-mail address: josea.perez.simon.sspa@juntadeandalucia.es (J.A. Pérez-Simón).

0145-2126/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
doi:10.1016/j.leukres.2012.02.030

5-azaC we observed that the level of *PRAME* methylation was reduced in AML patients which correlated with an increase in the expression of *PRAME*. By contrast, in CD34+ normal cells the effect of 5-azaC on the methylation pattern of the promoter of *PRAME* was significantly lower.

2. Patients and methods

2.1. Patients and healthy stem cell donors

Eight samples from healthy donors and eleven bone marrow samples from AML patients were obtained, after written consent of patients and donors.

2.2. Cell cultures

CD34+ stem cells from healthy donors were isolated using the AutoMACS system (Miltenyi Biotec, bergisch Gladbach, Germany) according to the manufacturer's instructions. The purity of the stem cells was >95% as determined on a FACSCalibur flow cytometer (Becton Dickinson Bioscience).

Mononuclear cells from AML bone marrow samples were obtained by Ficoll-Hypaque density gradient. Mononuclear cells were stained with a four-color combination of monoclonal antibodies (MoAbs) conjugated with the following fluorochromes: fluorescein isothiocyanate (FITC), phycoerithrin (PE) and allophycocyanin (APC). Specific antibodies were purchased from Becton Dickinson Bioscience (BDB) Pharmingen (San Jose, CA). The following combinations were used: CD45FITC/AnexinaPE/7AAD/CD33APC or CD45FITC/AnexinaPE/7AAD/CD34APC, depending on blasts phenotype. After this procedure, leukemic cells were isolated with a flow cytometer, equipped with its accompanying software (FACSAria and FACSDiva, respectively, Becton Dickinson Biosciences).

The purity of the isolated cell populations was evaluated after acquiring blasts corresponding to each FACS-sorted cell fraction (FACSAria flow cytometer) and it was higher than >95% in all cases.

CD34+ stem cells and blasts were washed and cultured for four days under standard conditions in RPMI 1640 L-glutamine (2 mM), penicillin (100 UI/mL) and streptomycin (10 mg/mL) plus 10% human AB serum (Sigma). 1 μM 5-azacytidine (Sigma) was added at day 1 and then every 24 h during the four days of culture [16].

2.3. RNA extraction, retro-transcription and quantitative PCR

Total RNA was isolated using the DNA/RNA Micro Kit (Qiagen, Valencia, CA) following the protocol for animal and human cells. One μg of RNA was retrotranscribed by using the High Capacity cDNA Reverse Transcription Kits (Applied Biosystems). The quantification of *PRAME* expression was performed using the Step One Plus Real-Time PCR System and TaqMan® Gene Expression Assays (Applied Biosystems) according to the manufacturer's instructions. Relative quantification was calculated using the equation $2^{-\Delta\Delta Ct}$ where $\Delta Ct = Ctgen - Ct\mathit{ABL1}$ and $\Delta\Delta Ct = \Delta Ct$ sample $- \Delta Ct$ control at 0 h of incubation [18].

2.4. Analysis of gene promoter methylation: bisulfite sequencing

The CpG island DNA methylation status was determined by sequencing bisulfite-modified genomic DNA (Applied Biosystems 3730 DNA Analyzer). Bisulfite modification of genomic DNA was carried out as described by Herman et al. [19]. For the CpG island present at PRAME promoter, primers were designed using the Methyl Primer Express v1.0 program (Applied Biosystems). (Primers F: ATTTTTTTAGAGGGTTTGGGAG R: TTCCCAAAACTTTCTAAAACCC).

2.5. Statistical analysis

Statistical analyses were performed using the SPSS software program (SPSS 15.0, Chicago, IL). The Wilcoxon two-sample paired signed rank test was used to compare different gene expressions between 5-azaC treated or untreated samples at different time-points. *p*-Values less than 0.05 were considered significant.

3. Results and discussion

3.1. PRAME expression and methylation status in AML blasts and CD34+ cells from healthy donors

First, we compared *PRAME* mRNA expression of blast cells from 11 AML patients at diagnosis versus CD34+ stem cells from 8 healthy donors by RT-PCR. As shown in Fig. 1, *PRAME* is significantly overexpressed in blasts compared with normal CD34+ cells (700 ± 1102 vs. 1.8 ± 2.5 $p = 0.002$).

While few studies have been reported analyzing *PRAME* expression in cell lines, the information using primary AML cells is scanty.

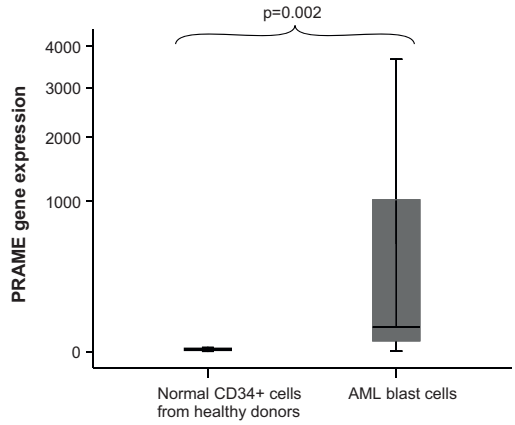


Fig. 1. *PRAME* gene expression in CD34+ cells from healthy donors and in blast cells from AML patients.

On the other hand, Steinbach et al. [11] showed that *PRAME* is expressed by CD34+ stem cells, which might constitute a problem for the use of this antigen as a target in immunotherapy. In contrast, Greiner et al. [20] reported that CD34+ cells do not express *PRAME*. These inconsistent results could be explained since authors used different qualitative PCR protocols. In the current study we show that *PRAME* is overexpressed in leukemic cells as compared to CD34+ cells so that it could be used as a therapeutic target. Strictly speaking, qRT-PCR provides information about the levels of mRNA (not even about transcription rate/activity). A western blot showing the levels of PRAME would be necessary.

Next, in order to test whether the methylation status of *PRAME* correlates with *PRAME* mRNA levels, we analyzed *PRAME* promoter from blast cells and CD34+ cells from healthy donors. We observed an inverse correlation between *PRAME* mRNA levels and methylation status as shown in Fig. 2A. Using a 500-fold the mean of *PRAME* expression in healthy donors as cut-off, patients with low *PRAME* gene expression levels (patients 1–4) showed a trend towards a higher percentage of methylated CpG (median 37.5%, range 29–44.5) as compared to cases with high *PRAME* expression (patients 5–7, median of methylated CpG 6.0%, range 5.6–34.0, $p = 0.08$). Of note, we found a statistically significant association between *PRAME* expression and the degree of methylation in the promoter of *PRAME* among both AML samples and healthy donors ($r = -0.77$, $p = 0.010$; Spearman correlation; Fig. 3).

Interestingly, CpG 13 (marked with an asterisk in Fig. 2A) showed a very good correlation between methylation status and *PRAME* gene expression levels. Thus, we confirmed that there is a good correlation between *PRAME* methylation promoter, particularly CpG13, and *PRAME* expression. In this regard, Ortmann et al. [14] have previously reported a good correlation between *PRAME* intron1 hypomethylation status and *PRAME* overexpression in AML patients and Roman-Gomez et al. [13] have described an association between low methylation of *PRAME* exon2 and high gene expression. By contrast, there is only one study in which methylation status has been evaluated in the *PRAME* promoter region in 4 childhood AML samples, peripheral blood samples from healthy donors and K-562, U-937 and HL-60 cell lines [15]. In the latter study it was observed that *PRAME* promoter is methylated in peripheral blood from healthy donors, in AML patients and in the cell line U-937 in which *PRAME* expression is negative, whereas this region is demethylated in high *PRAME* expressers cell lines (K-562 and HL-60) [15]. However, the methylation status of the

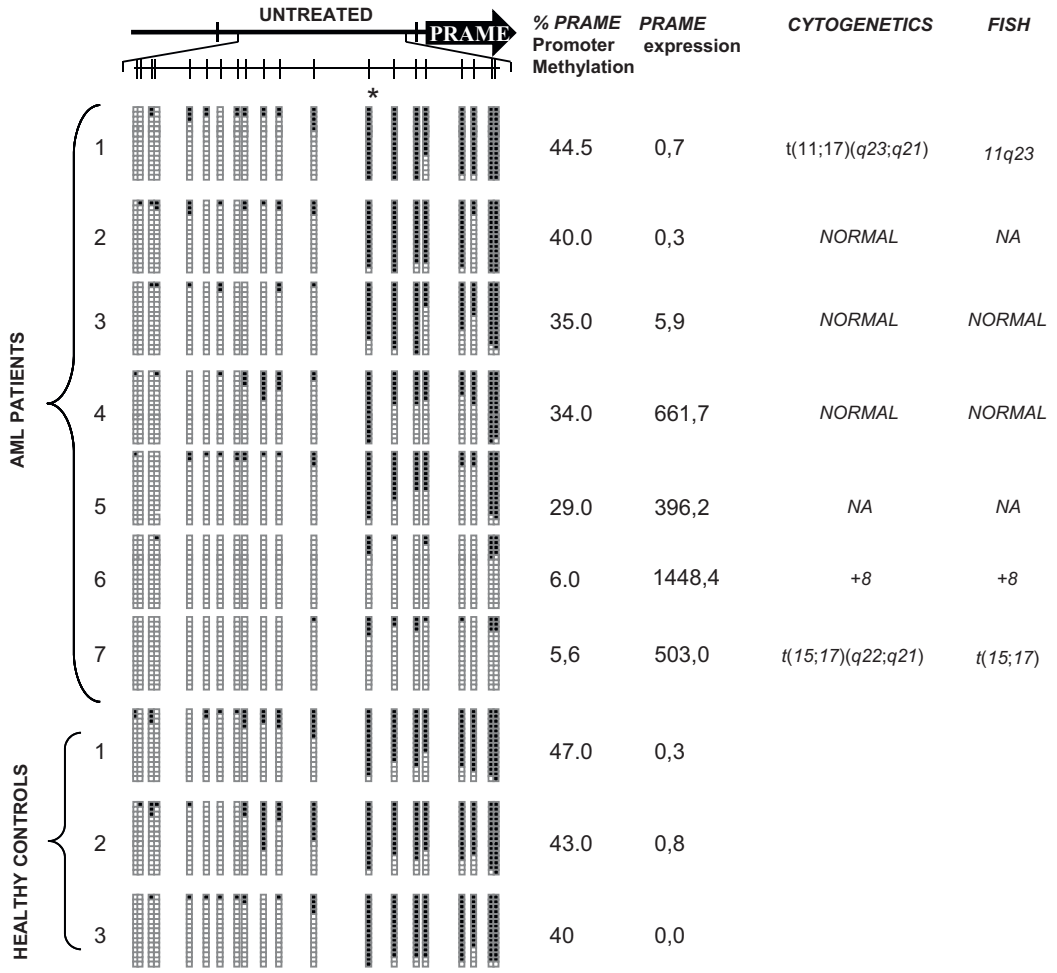


Fig. 2. (a) *PRAME* promoter methylation status in AML patients and in healthy donors. (b) *PRAME* gene expression and methylation status in AML patients and healthy donors cells treated and non-treated with 1 μM 5-azacytidine.

promoter region was not evaluated neither in selected blasts from AML patients nor in CD34+ cells from healthy donors. In the present study we show that *PRAME* promoter is demethylated specifically in the same CpG site as reported by Schenk et al. [15] in leukemic cells from *PRAME*-positive patients. Additionally, in leukemic cells from *PRAME*-negative patients and in CD34+ cells from healthy donors this region is methylated and *PRAME* expression is lower. This latter finding has not been previously reported and explains the lower expression of *PRAME* in normal cells compared to blasts and supports the use of *PRAME* as a therapeutic target in AML patients.

3.2. Treatment of blast cells from AML patients and selected CD34+ stem cells from healthy donors with 5-azaC

To evaluate the effect of a hypomethylating agent (such as 5-azaC) on *PRAME* gene expression, we treated blast cells and CD34+ cells with 5-azaC. We observed a trend towards a reduction in methylated CpGs in cells treated with 5-azaC (percentage of methylation, median 21.6%, range 5–37.8) (Fig. 2B) compared to non-treated control (percentage of methylation, median 34%, range 5.6–44.5, $p = 0.09$) (Fig. 2A). Interestingly, those cases with a lower expression of *PRAME*, i.e. those, with a higher percentage of methylation in *PRAME* promoter, showed the higher differences in methylation pattern prior to and after treatment with 5-azaC (27.2% vs. 37.5%, $p = 0.068$). By contrast, in healthy donor cells, no differences were observed between 5-azaC-treated and non-treated cells (median 45 vs. 43.5%, $p = 0.6$). These results could suggest that the effect of 5-azaC could be leukemia-specific, which a higher effect on AML blasts in comparison to healthy cells. Thus, this study demonstrated for the first time that *PRAME* is overexpressed in primary leukemic cells but not in CD34+ normal cells after treatment with 5-azaC 1 μM. Although further studies are needed to explain this selective effect, the induction of *PRAME* overexpression in leukemic cells without affecting CD34+ cells provides a potential selective target for immunotherapy. Furthermore, considering the role of leukemic stem cells on relapse even after allogeneic transplantation [21], it would be interesting to evaluate the expression of *PRAME* in this cell subset and the effect of 5-azaC in order to improve the efficacy of the immune response especially in the transplant setting.

In conclusion we show that 5-azaC induces overexpression of *PRAME* in blast cells from AML patients with no effect on CD34+

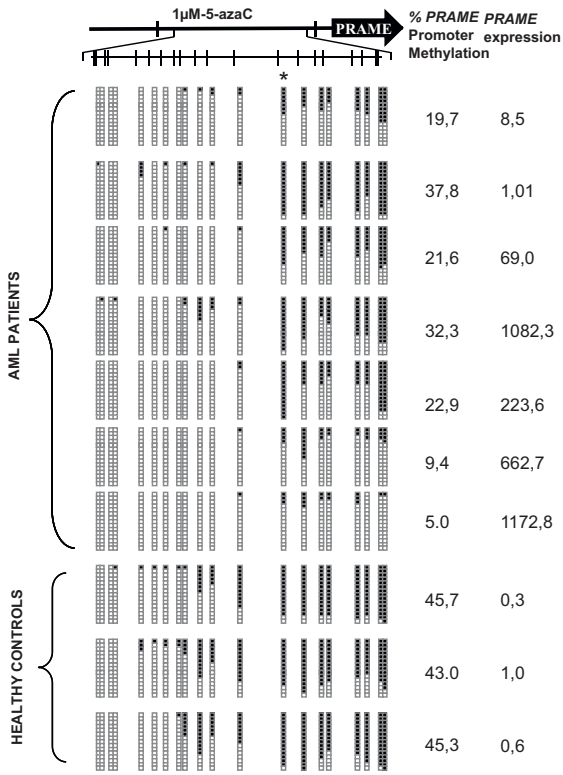


Fig. 2. (a) *PRAME* promoter methylation status in AML patients and in healthy donors. (b) *PRAME* gene expression and methylation status in AML patients and healthy donors cells treated and non-treated with 1 μM 5-azacytidine.

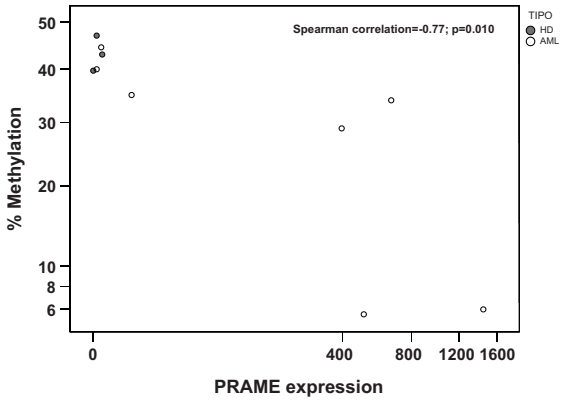


Fig. 3. Correlation between *PRAME* methylation status and *PRAME* expression in AML patients and in healthy donors.

cells from healthy donors, suggesting that *PRAME* could be used as a target in tumor immunotherapy.

Conflict of interest statement

The authors declare no competing financial interests.

Acknowledgements

Role of the funding source: This work has been supported by Fondo de Investigación Sanitaria, Instituto de Salud Carlos III ref: PI080047.

Contributors: S.G.-C. Performed cell cultures and molecular assays. E.B. performed arrays assays. C.S., C.H-S. performed molecular assays. L.I.S-A., C.C. performed cell cultures and developed experiments. B.B. performed cell cultures and critically reviewed the manuscript. T.C-V., S.C. performed cell cultures. L.C. performed genomic DNA methylation assays. C.C., J.F. San M. critically reviewed the research Project. J.A.P-S. developed experimental designs and the Research Project.

References

[1] Ikeda H, Lethe B, Lehmann F, van Baren N, Baurain JF, de Smet C, et al. Characterization of an antigen that is recognized on a melanoma showing partial HLA loss by CTL expressing an NK inhibitory receptor. Immunity 1997;6:199–208.
[2] Epping MT, Bernards R. A causal role for the human tumor antigen preferentially expressed antigen of melanoma in cancer. Cancer Res 2006;66:10639–42.
[3] van Baren N, Chambost H, Ferrant A, Michaux L, Ikeda H, Millard I, et al. PRAME, a gene encoding an antigen recognized on a human melanoma by cytolytic T cells, is expressed in acute leukaemia cells. Br J Haematol 1998;102:1376–9.
[4] Paydas S, Tanriverdi K, Yavuz S, Disel U, Baslamisli F, Burgut R. PRAME mRNA levels in cases with acute leukemia: clinical importance and future prospects. Am J Hematol 2005;79:257–61.
[5] Santamaria C, Chillon MC, Garcia-Sanz R, Balanzategui A, Sarasquete ME, Alcoceba M, et al. The relevance of preferentially expressed antigen of melanoma (PRAME) as a marker of disease activity and prognosis in acute promyelocytic leukemia. Haematologica 2008;93:1797–805.

[6] Steinbach D, Schramm A, Eggert A, Onda M, Dawczynski K, Rump A, et al. Identification of a set of seven genes for the monitoring of minimal residual disease in pediatric acute myeloid leukemia. Clin Cancer Res 2006;12:2434–41.
[7] Greiner J, Schmitt M, Li L, Giannopoulos K, Bosch K, Schmitt A, et al. Expression of tumor-associated antigens in acute myeloid leukemia: implications for specific immunotherapeutic approaches. Blood 2006;108:4109–17.
[8] Santamaria CM, Chillon MC, Garcia-Sanz R, Perez C, Caballero MD, Ramos F, et al. Molecular stratification model for prognosis in cytogenetically normal acute myeloid leukemia. Blood 2009;114:148–52.
[9] Griffioen M, Kessler JH, Borghi M, van Soest RA, van der Minne CE, Nouta J, et al. Detection and functional analysis of CD8+ T cells specific for PRAME: a target for T-cell therapy. Clin Cancer Res 2006;12:3130–6.
[10] Rezvani K, Yong AS, Tawab A, Jafarpour B, Eniafe R, Mielke S, et al. Ex vivo characterization of polyclonal memory CD8+ T-cell responses to PRAME-specific peptides in patients with acute lymphoblastic leukemia and acute and chronic myeloid leukemia. Blood 2009;113:2245–55.
[11] Steinbach D, Hermann J, Viehmann S, Zintl F, Gruhn B. Clinical implications of PRAME gene expression in childhood acute myeloid leukemia. Cancer Genet Cytogenet 2002;133:118–23.
[12] Esteller M. Epigenetics in cancer. N Engl J Med 2008;358:1148–59.
[13] Roman-Gomez J, Jimenez-Velasco A, Agirre X, Castillejo JA, Navarro G, Jose-Eneriz ES, et al. Epigenetic regulation of PRAME gene in chronic myeloid leukemia. Leuk Res 2007;31:1521–8.
[14] Ortmann CA, Eisele L, Nuckel H, Klein-Hitpass L, Fuhrer A, Duhrsen U, et al. Aberrant hypomethylation of the cancer-testis antigen PRAME correlates with PRAME expression in acute myeloid leukemia. Ann Hematol 2008;87: 809–18.
[15] Schenk T, Stengel S, Goellner S, Steinbach D, Saluz HP. Hypomethylation of PRAME is responsible for its aberrant overexpression in human malignancies. Genes Chromosomes Cancer 2007;46:796–804.
[16] Sanchez-Abarca LI, Gutierrez-Cosio S, Santamaria C, Caballero-Velazquez T, Blanco B, Herrero-Sanchez C, et al. Immunomodulatory effect of 5-azacytidine (5-azaC): potential role in the transplantation setting. Blood 2010;115: 107–21.
[17] Sigalotti L, Fratta E, Coral S, Tanzarella S, Danielli R, Colizzi F, et al. Intratumor heterogeneity of cancer/testis antigens expression in human cutaneous melanoma is methylation-regulated and functionally reverted by 5-aza-2′-deoxycytidine. Cancer Res 2004;64:9167–71.
[18] Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2−-Delta Delta C(T)) method. Methods 2001;25:402–8.
[19] Herman JG, Graff JR, Myohanen S, Nelkin BD, Baylin SB. Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands. Proc Natl Acad Sci U S A 1996;93:9821–6.
[20] Greiner J, Ringhoffer M, Simikopinko O, Szmaragowska A, Huebsch S, Maurer U, et al. Simultaneous expression of different immunogenic antigens in acute myeloid leukemia. Exp Hematol 2000;28:1413–22.
[21] Gerber JM, Smith BD, Ngwang B, et al. A clinically relevant population of leukemic CD34+CD38- cells in AML. Blood 2012 [Epub ahead of print].