Oncogene (2005) 24, 3358–3368
© 2005 Nature Publishing Group All rights reserved 0950-9232/05 $30.00
www.nature.com/onc

Molecular signature of retinoic acid treatment in acute promyelocytic leukemia

Natalia Meani[1,2], Simone Minardi[1], Silvia Licciulli[1,2], Vania Gelmetti[3], Francesco Lo Coco[4], Clara Nervi[3], Pier Giuseppe Pelicci[1,2], Heiko Müller[1,2] and Myriam Alcalay*[,1,2]

[1]*IFOM – Institute of Molecular Oncology of the Italian Foundation for Cancer Research, Via Adamello 16, 20139 Milan, Italy;* [2]*European Institute of Oncology, Via Ripamonti 435, 20141 Milan, Italy;* [3]*San Raffaele Bio-medical Science Park of Rome, Via di Castel Romano 100, 00128 Rome, Italy;* [4]*Department of Biopathology, University of Tor Vergata, Rome, Italy*

Acute promyelocytic leukemia (APL) is a distinct subtype of acute myeloid leukemia characterized by a block of differentiation at the promyelocytic stage. APL patients respond to pharmacological concentrations of all-trans retinoic acid (RA) and disease remission correlates with terminal differentiation of leukemic blasts. The PML/RAR oncogenic transcription factor is responsible for both the pathogenesis of APL and for its sensitivity to RA. In order to identify physiological targets of RA therapy, we analysed gene expression profiles of RA-treated APL blasts and found 1056 common target genes. Comparing these results to those obtained in RA-treated U937 cell lines revealed that transcriptional response to RA is largely dependent on the expression of PML/RAR. Several genes involved in the control of differentiation and stem cell renewal are early targets of RA regulation, and may be important effectors of RA response. Modulation of chromatin modifying genes was also observed, suggesting that specific structural changes in local chromatin domains may be required to promote RA-mediated differentiation. Computational analysis of upstream genomic regions in RA target genes revealed nonrandom distribution of transcription factor binding sites, indicating that specific transcriptional regulatory complexes may be involved in determining RA response.
Oncogene (2005) **24,** 3358–3368. doi:10.1038/sj.onc.1208498
Published online 21 February 2005

Keywords: differentiation therapy; acute myeloid leukemia; transcriptional regulation; expression profiling; transcription factor binding sites

Introduction

Acute promyelocytic leukemia (APL) is characterized by the 15;17 translocation, which leads to the expression of the PML/RAR fusion protein, an aberrant transcription factor that interferes with the processes of myeloid differentiation (Alcalay *et al*., 2001). Characteristic of APL is a marked sensitivity to differentiation induced by all-trans retinoic acid (RA) both *in vitro* and *in vivo* (Lin *et al*., 1999). Treatment of APL with RA induces disease remission, and represents the first successful attempt of differentiation therapy that specifically targets the aberrant protein underlying disease onset (Huang *et al*., 1988; Lin *et al*., 1999; Minucci and Pelicci, 1999). A peculiarity of APL is the dose-dependent, dual response to RA: PML/RAR expressing cells do not respond to low doses, but have an enhanced sensitivity to therapeutic doses of RA (Grignani *et al*., 1993).

The interaction of PML/RAR with co-repressor/HDAC complexes is responsible for both transcriptional repression of target genes and sensitivity to RA (Minucci and Pelicci, 1999). PML/RAR, in the absence of RA, recruits the nuclear co-repressor (N-CoR)-HDAC complex through the RAR CoR box, binds to target promoters, and represses transcription (Grignani *et al*., 1998). High doses of RA release HDAC activity from PML/RAR, thus permitting transcription of target genes.

PML/RAR functions by deregulating RA target genes that are critical to myeloid differentiation, and are thought to represent the downstream effectors of its oncogenic potential. The consequent aberrant transcriptional pattern may result in signaling networks that directly interfere with the normal differentiation program of myeloid precursor cells. RA is thought to act by antagonizing PML/RAR-dependent gene regulation, thereby favoring terminal differentiation. Analysis of the regulatory pathways impaired during leukemogenesis and reactivated during RA-induced differentiation may contribute to the identification of new molecular targets for leukemia therapy.

Previous studies describe gene expression modifications induced by RA in cell lines (Tamayo *et al*., 1999; Liu *et al*., 2000; Lee *et al*., 2002; Park *et al*., 2003). We here present a detailed analysis of transcriptional regulation in blasts derived from APL patients prior to and after RA treatment. We show that groups of genes encoding for specific functions, including vast networks of differentiation regulators and chromatin modifiers, are early targets of regulation by RA. Exploiting a cell line model, we show that RA transcriptional response is

*Correspondence: M Alcalay;
E-mail: myriam.alcalay@ifom-ieo-campus.it
Received 11 October 2004; revised 17 December 2004; accepted 10 January 2005; published online 21 February 2005

largely PML/RAR-dependent. Analysis of the upstream genomic regions of RA regulated genes reveals a nonrandom distribution of transcription factor binding sites (TFBS), suggesting the involvement of specific regulatory complexes in RA-response. Taken together, our data provide a detailed description of the transcriptional response to RA treatment in APL blasts, and describe the molecular basis of differentiation therapy.

Results

Identification of 1056 genes that are common targets of transcriptional regulation during RA treatment of APL blasts

In order to define the molecular basis of RA response in APL, we analysed gene expression profiles in two experimental models. First, to define genes that are physiological targets of RA-treatment in APL, we used blasts from three APL patients expressing PML/RAR, before and after treatment with 10^{-6} M RA *in vitro* for 4 h. Next, to identify genes that are RA targets in a PML/RAR-dependent manner, we studied a U937 clone conditionally expressing PML/RAR (U937-PR) (Grignani *et al*., 1993; Alcalay *et al*., 2003), and compared the gene expression profile prior to and after 4 h of treatment with 10^{-6} M RA, to that obtained from a cell line bearing an empty vector (U937-Mt).

For each sample, biotinylated cRNA targets were synthesized starting from 5 μg of total RNA, and hybridized to the complete set of HG-U133 Affymetrix oligonucleotide chips, which explores the expression of approximately 45 000 human transcripts. Results were analysed using MASv5 and further elaborated with the GenePicker software (Alcalay *et al*., 2003; Finocchiaro *et al*., 2004). GeneChip probe sets regulated by RA in each sample were clustered into nonredundant regulated genes according to UniGene release Hs.166. The total of RA regulated genes was 2942 in APL patient 1 (APL#1), 3335 in APL patient 2 (APL#2), 2850 in APL patient 3 (APL#3), homogeneously distributed between induced and repressed genes in each case (Figure 1). Cross comparison of results identified 1056 genes concordantly regulated in all three patients, which represent 32–37% of the total regulated genes in each sample. Our screens, therefore, identified a large cohort of genes that are targets of regulation during RA treatment in APL blasts derived from different individuals.

RA regulates the expression of 2353 genes in U937-PR cells and of 721 genes in control U937-Mt cells (Figure 1). Of these, 268 are in common and concordantly regulated between the two cell lines. These results suggest that a large proportion of RA-target genes in U937 cells depend on PML/RAR expression (89% of genes regulated in the U937-PR clone are not regulated by RA in control cells). All regulated GeneChip Probe sets identified in our screens are shown in Supplementary Table 1, and conversion to non-redundant regulated genes according to UniGene release Hs.166 can be seen in Supplementary Table 2.

RA Regulated Genes

		Total	*Increased*	*Decreased*
	APL#1	2942	1345	1597
	APL#2	3335	1609	1726
	APL#3	2850	1592	1258
	U937-PR+RA	2353	1259	1094
	U937-Mt+RA	721	336	385
Common	All APLs	1056	515	541
	APLs & U937-PR + RA	546	310	236
	APLs & U937-Mt +RA	112	77	35

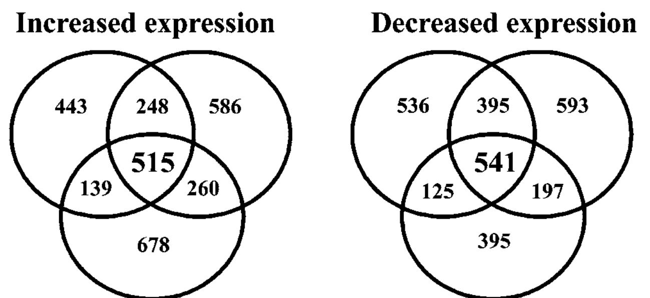

Figure 1 Schematic representation of GeneChip results. Expression profiles of leukemic blasts from three APL patients (APL#1, 2 and 3 respectively) and of U937 cell lines expressing PML/RAR (U937-PR) or bearing an empty vector (U937-Mt) were analysed prior to and after 4 h of treatment with RA *in vitro* using Affymetrix GeneChips (HG-U133set). Numbers of common target genes are reported, and suggest a high degree of homogeneity in samples derived from APL patients. Approximately 50% of target genes identified in APL blasts are also regulated in U937-PR, whereas less than 10% of genes are also regulated in U937-Mt cells. Venn diagrams represent overlaps in regulated genes among the three APL samples

Hierarchical clustering and comparison of global gene expression profiles reveals PML/RAR-dependent transcriptional regulation during RA treatment

In order to better evaluate the gene expression profiles obtained from the two experimental model systems, we applied a series of statistical tests. We first performed an unsupervised hierarchical clustering (Pearson correlation) of 7282 probe sets that showed a two-fold variation with respect to the median of all absolute expression values in at least one APL patient sample. Strikingly, the samples clustered according to treatment rather than to patient samples, indicating a strong molecular signature determined by RA (Figure 2a). We next performed unsupervised clustering (Pearson correlation) of the same cohort of probe sets using the comparative expression values across all conditions, with the aim of identifying similarities in the patterns of gene regulation. Genes regulated by RA treatment in U937-PR cells partially overlap with those regulated by RA treatment of APL blasts, whereas control U937-Mt cells appear notably different (Figure 2b).

In order to eliminate the influence of individual genetic differences, we considered for further studies the regulated genes that are common to the three patient

samples. The 1056 common RA target genes (515 induced and 541 repressed), and the corresponding 1621 probe sets are reported in Supplementary Tables 3 and 4, respectively. We then compared the behavior of these RA target genes to the group of genes regulated by PML/RAR expression in the U937 system (Figure 2c) (Alcalay *et al*., 2003). We found that 209 genes that are induced upon RA treatment were targets of repression by PML/RAR, and 135 genes repressed after RA treatment were instead induced by PML/RAR.

Interestingly, 546 of the 1056 common RA-target genes (52%) were also regulated by RA in U937-PR cells, but 80% of these do not show variations in expression levels in RA-treated control U937-Mt cells, further suggesting that the transcriptional response to RA in APL blasts and U937 cells is largely dependent on the presence of PML/RAR (Figure 2c). Of these 546 RA-common targets, 174 genes activated upon RA treatment were targets of repression by PML/RAR (Alcalay *et al*., 2003), whereas 84 genes repressed by RA treatment were activated in U937 cells expressing PML/RAR.

Comparison of identified RA targets to previously reported studies

Different studies of gene expression profiling that use cell lines, either NB4 or U937, as models to study RA-induced differentiation in APL have already been published (Tamayo *et al*., 1999; Liu *et al*., 2000; Lee *et al*., 2002; Park *et al*., 2003). In order to compare our results with those of the previously published reports, we first re-annotated the published results according to UniGene release Hs.166, starting from the GeneChip probe set ID (where available), or, alternatively, from the GeneBank accession number. We then calculated the number of reported regulated genes in each study that are present on the HG-U133 chip set (see Table 1, ‘comparable genes’), and searched for overlaps with our results, both in the U937-PR system and in APL patients. As shown in Table 1, the number of common target genes identified is variable due to the use of technical approaches that are not fully comparable, and/or to the biological differences inherent in the experimental systems under analysis. A full list of the commonly identified genes is reported in Supplementary Table 5. The functional categories that are recurrently identified by all studies include regulators of differentiation and genes involved in the control of cellular proliferation and apoptosis.

Table 1 Comparison of RA target genes identified in this study with results presented in previous reports

	Tamayo	*Liu*	*Lee*	*Park*
Comparable genes	498	167	93	33
Concordant with U937-PRRA	39	29	5	27
Concordant with APL blasts + RA	61	26	3	21

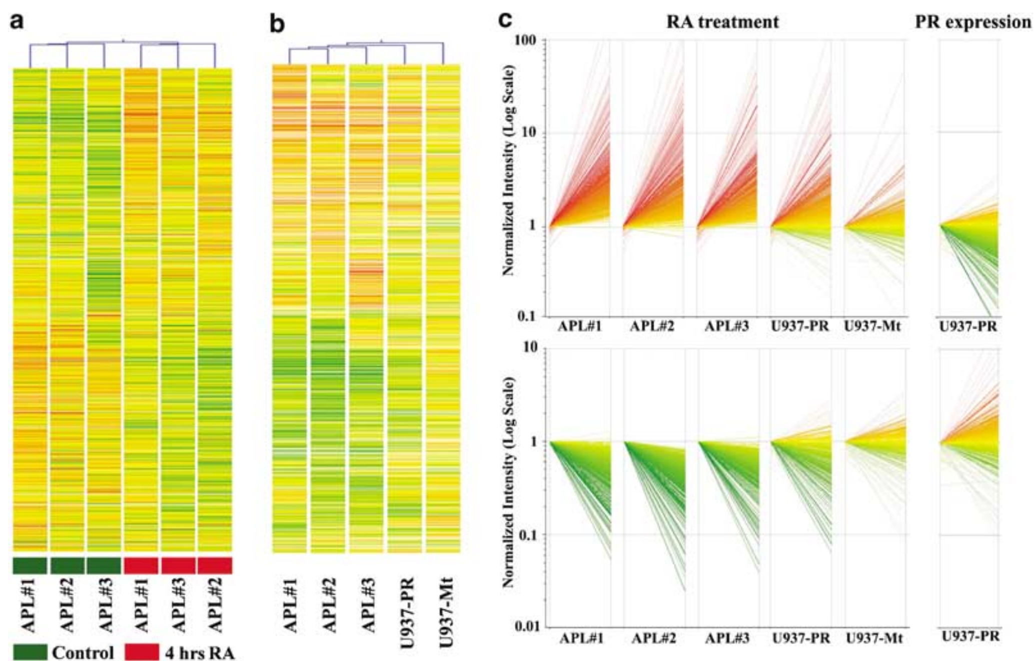

Figure 2 Analysis of RA-dependent transcriptional regulation. (**a**) Unsupervised clustering (Pearson correlation) of 7282 probe sets that showed variation in at least one APL patient was performed using absolute expression values of all samples derived from APL blasts. APL samples cluster according to RA treatment. (**b**) Unsupervised clustering (Pearson correlation) of 7282 probe sets that showed variance in at least one APL patient was performed using comparative expression values of all samples. In each condition, the expression value of each gene in cells treated with RA for 4 h was compared to the level of expression in the sample prior to RA treatment. (**c**) Comparative expression levels of 1056 genes (represented by 1621 Affymetrix probe sets) that are induced or repressed in all three APL patients is shown in all experimental conditions described in the study. The profile of regulation is clearly dependent on the presence of PML/RAR, since the pattern in the U937-Mt cells is notably different from all other conditions

RA treatment of APL blasts regulates genes involved in hemopoietic differentiation

We next classified the 1056 common target genes according to function, by collecting annotations and keywords from various Web-based sources, as previously described (Alcalay *et al*., 2003). We retrieved useful information for 669 genes, which were classified into 14 major functions. Figure 3 indicates the number of genes belonging to each major functional category, which can be seen in detail in Supplementary Table 6. We also performed a comparative analysis of Gene Ontology categories (GO biological process) using EASE software (Hosack *et al*., 2003) (http://apps1.niaid.nih.gov/david/upload.jsp) (see Supplementary Table 7A for complete results, B for enriched categories and C and D for graphical representation of enriched GO biological processes in RA induced and repressed gene lists). Interestingly, among the functions that are over-represented in genes induced by RA treatment, there are specialized functions of mature neutrophils (inflammatory and immune response) and negative regulators of transcription (Supplementary Table 7C). Among the repressed genes, instead, there is an over-representation of positive regulators of cell proliferation and of genes involved in programmed cell death (Supplementary Table 7D).

We first focused our attention on the 55 genes classified as regulators of differentiation and/or stem cell maintenance (see Supplementary Table 8). Within this group, eight genes are involved in Notch signaling and another eight genes in Wnt signaling, with two genes (*TLE1*, *TLE3*) in common to both signaling pathways. Regulation of gene expression levels of these 55 targets appears to be PML/RAR dependent. In fact, these genes show similar modifications in expression levels in APL blasts and in U937-PR cells after RA treatment, whereas their expression appears virtually unmodified in RA-treated U937-Mt control cells (Figure 4a).

We assessed the expression levels of 30 of the newly identified target genes belonging to this category by RT–qPCR and obtained full concordance between Affymetrix predictions and expression results in APL patients (Figure 4b). Numerical data and direct comparison between GeneChip and RT–qPCR results can be seen in Supplementary Table 9.

RA treatment of APL blasts induces modifications in the expression levels of genes that regulate chromatin function

It has been demonstrated that PML/RAR recruits co-repressor complexes and DNA-methyltransferases to its target promoters (Grignani *et al*., 1998; He *et al*., 1998; Lin *et al*., 1999; Di Croce *et al*., 2002). Moreover, it was recently shown that PML/RAR has a stronger transcriptional repressor activity than the wild-type RARα nuclear receptor, and imposes a more condensed chromatin conformation, which is less accessible even in the presence of RA (Segalla *et al*., 2003). This finding suggests that multiple repressive pathways might be involved in oncogenic transcriptional control by PML/RAR. To date, it is not known whether PML/RAR, aside from recruiting chromatin modifiers, also regulates their expression levels after RA treatment. Among the 669 classified RA target genes, 32 have functions that are tightly connected to chromatin (see Supplementary Table 10). Expression levels of these genes also appear to be regulated in a PML/RAR-dependent manner (Figure 5a).

Histone lysine methylation has been associated with transcriptional regulation: methylated H3 Lys9 is present in silenced chromatin, while H3 Lys4 methylation is enriched in transcriptionally active chromatin. We found that RA treatment of APL blasts is associated

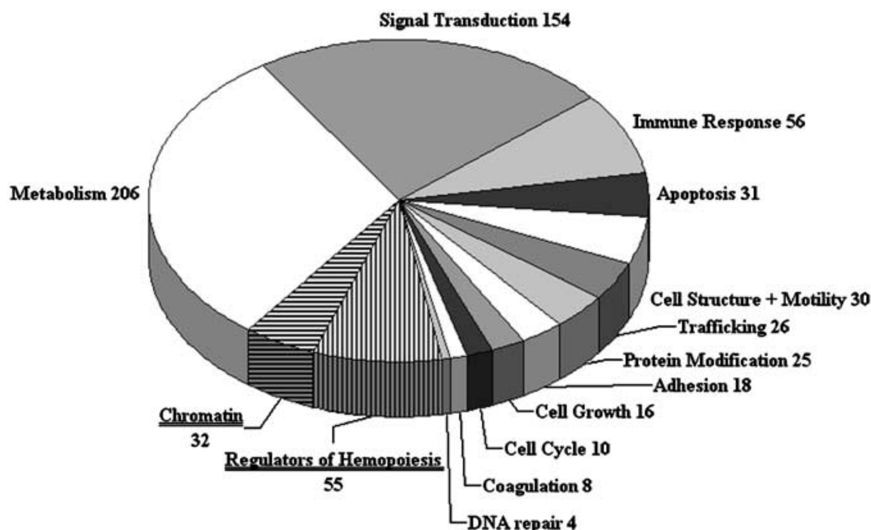

Figure 3 Functional classification of common target genes. A total of 669 common targets was grouped according to known functions: the 14 major categories and the number of genes in each one are shown. Functional classes further discussed in the text are underlined

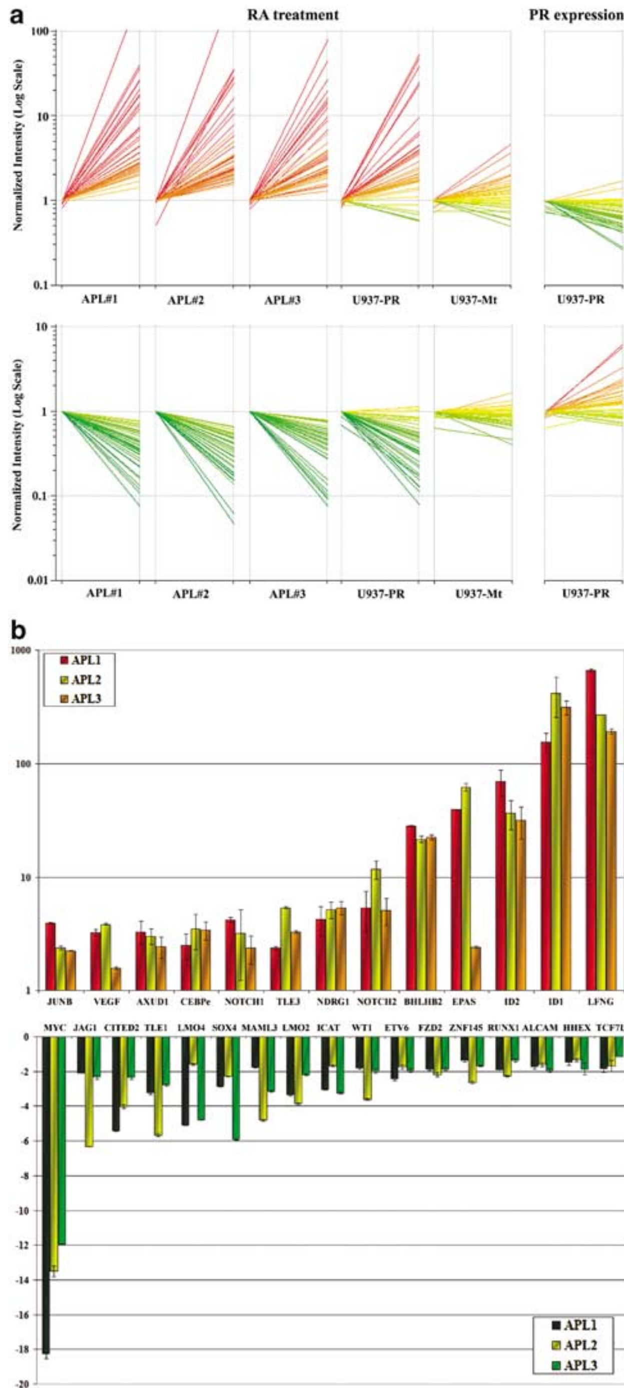

Figure 4 RA treatment regulates expression of genes involved in the control of hemopoiesis. (**a**) Graphic representation of the expression patterns of 55 genes classified as regulators of hemopoiesis in the experimental conditions studies on Affymetrix chips. The genes show similar behavior in APL blasts and in U937-PR cells after RA treatment, whereas their expression appears unmodified in RA-treated U937-Mt cells or reversed in U937-PR cells. (**b**) RT–qPCR of 30 common target genes belonging to this functional category. Comparative expression levels in blasts derived from three APL patients prior to and after 4 h of treatment with 10^{-6} M RA are shown

to decreased expression of the G9A (BAT8) and the SET7/9 methyltransferases, and increased expression of the DOT1-like (DOT1L) methyltransferase (Supplementary Table 10). We assessed the expression levels of the three histone methyltransferase genes by RT–qPCR and obtained full concordance between Affymetrix predictions and expression results in APL patients (Figure 5b and Supplementary Table 9). G9A is responsible for mono- and di-methylation of H3 Lys9 and marks silent chromatin domains within euchromatin, a function that has been associated to repression of developmental genes in embryonic systems (Tachibana *et al*., 2002). No global changes in histone H3 Lys9 di-methylation were detectable in total protein lysates derived from APL blasts treated with RA (data not shown). Local changes in the degree of methylation might, however, occur.

RA target genes in APL blasts display specific patterns of TFBS in the upstream genomic sequences and first exons

We next investigated the presence of TFBS in the regulatory regions of the 1056 genes identified as common targets of RA treatment in APL blasts. The genomic sequence corresponding to 1000 bp upstream, plus the first exon, for 920 genes that have unambiguous LocusLink entries were obtained using the TRASER retrieval system (http://genome-www6.stanford.edu/cgi-bin/Traser/traser). We retrieved only the regulatory sequences of the most 5′ exon 1; genes with multiple promoters and alternative exon 1 usage were, therefore, not completely studied.

The presence of RAR responsive elements (RAREs) in these sequences was assessed using MatInspector Professional (http://www.genomatix.de/), querying for the V$RARF matrix family (Quandt *et al*., 1995). We thus identified 234 occurrences of the selected matrix in the regulatory regions of 202 nonredundant genes (22% of total analysed), 103 repressed and 99 induced by RA treatment (Supplementary Table 11A). Since the V$RARF family is composed of two distinct matrices, V$RAR.01 and V$RTR.01, which identify binding sites for RAR and GCNF/RTR (retinoid receptor-related testis-associated receptor), respectively, we restricted further analysis to those genes that presented one or more copies of the V$RAR.01 sequence in their regulatory regions. We thus identified a group of 72 induced and 66 repressed genes (Supplementary Table 11B). Therefore, 138 genes (15% of total) contain putative RAREs, equally distributed between induced and repressed genes. This proportion is similar to the frequency of the V$RAR.01 matrix in a random group of nonregulated genes (data not shown).

A recent report by Kamashev *et al*. (2004) describes the identification of high-affinity binding sites for PML/RAR, composed of two AGGTCA motifs separated by spacers of variable length. Notably, the length of the spacer can range from 1 to 13 base pairs. The AGGTCA consensus, derived from an alignment of 172 individual high-affinity binding sites, can be found as a direct repeat (AGGTCAn(1–13)AGGTCA), as an inverted repeat (AGGTCAn(1–13)TGACCT), or as an everted

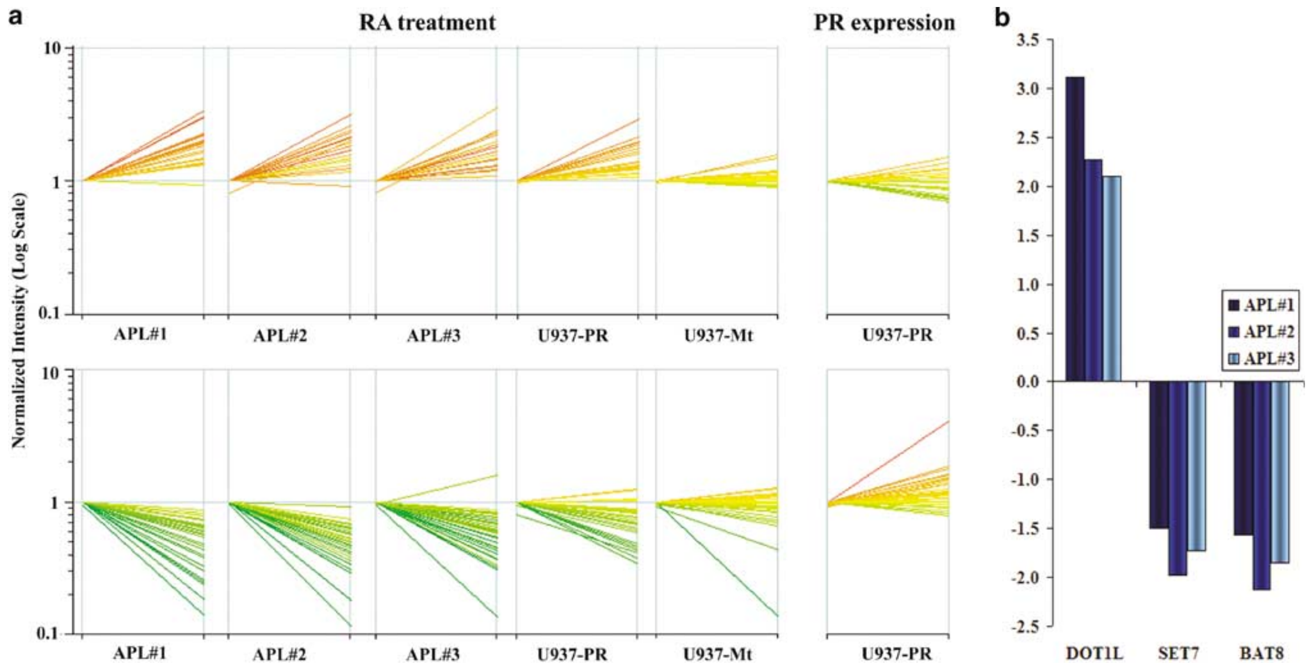

Figure 5 RA treatment modulates genes involved in chromatin function. (**a**) Graphic representation of the 32 genes, whose function is connected to chromatin modifications. Their behavior is very similar in APL blasts and in U937-PR cells after RA treatment, whereas their expression appears unmodified in RA-treated U937-Mt cells or opposite in U937-PR cells. (**b**) Expression levels of three histone methyltransferase encoding genes (*DOLT1L*, *SET7* and *BAT8\G9A*) assayed by RT–qPCR. Comparative expression levels in blasts derived from three APL patients prior to and after 4 h of treatment with 10^{-6} M RA are shown

repeat (TGACCTn(1–13) AGGTCA). We tested all of these motifs explicitly for enrichment in the promoter regions of the set of RA regulated genes, utilizing position frequency matrices.

The promoter regions and alignments of promoter regions of human and mouse were downloaded from the UCSC genome browser. No enrichment of any of the tested motifs was found when the unaligned promoter regions were tested. However, when the search was performed on the conserved regions of human and mouse promoters, nearly all of the motifs were found to be slightly enriched in the promoters of RA upregulated as well as in the RA down regulated genes (*P*-values given in Supplementary Table 12A). This enrichment is weak, probably because of the large variety of spacer lengths. We therefore tested the possibility that genes whose promoters contain any one or a combination of different motifs in their promoter regions might be found preferentially in our list of RA regulated genes. Genes were ranked by the number of binding sites for the tested motifs and Kolmogorow–Smirnow–Scanning was performed as described (Hollander and Wolfe, 1999) and is shown in Supplementary Table 12B. We observed highly significant enrichment of RA target genes (P-value<0.00001) by performing the promoter search on the alignment of human and mouse promoters but not on the unaligned sequences. Importantly, the enrichment of RA target genes at the top of the ranked list becomes more pronounced as the stringency of the motif search is increased from 0.7 to 0.85. The list of RA target genes containing these motifs in their promoter regions can be seen in Supplementary Table 12C.

The entire set of promoters of RA induced and repressed genes (I and D, respectively, in Table 2 and Figure 6), as well as the subset of genes containing the V$RAR.01 sequence (RARE I are the 72 RA induced genes and RARE D are the 66 repressed genes containing the V$RAR.01 sequence, in Table 2 and Figure 6) were studied for the occurrence of TFBS by scanning genomic sequences with position weight matrices obtained from TRANSFAC (http://www.biobase.de/pages/products/transfac.html). As controls, we analysed 50 groups of sequences, each consisting of *n* upstream genomic regions of genes randomly selected from those present in the Affymetrix HG-U133 set ($n=444$ for induced genes; $n=476$ for repressed genes). In parallel to the lists of regulated genes with putative RAREs, we studied a group of 50 genes containing the V$RAR.01 matrix that were not regulated by RA in our experimental conditions (RARE in Table 2 and Figure 6).

The affinity (score) of a transcription factor for a potential binding site was calculated as previously described (Kel *et al*., 2001; Vernell *et al*., 2003). We thus identified 25 matrices (corresponding to 24 transcription factors), whose distribution in the series of genes under analysis is different from that of the random group (Table 2). Among these, binding sites for few generic transcription factors present a nonrandom distribution in regulated genes: the SP3 binding site, characteristic of GC-rich regions, is over-represented, particularly in sequences containing putative RAREs, whereas the TATA binding site resulted under-represented in all genes induced by RA treatment (Figure 6a and b).

Table 2 TFBS that present a nonrandom distribution in RA target genes

TRANSFAC accession	*Matrix identifier*	*Transciption factor*	*I*	*D*	*RARE*	*RARE I*	*RARE D*
Generic TFs							
M00177	V$CREB_Q2	CREB	−		−	−	+
M00178	V$CREB_Q4	CREB	−		−	−	+
M00665	V$SP3_Q3	Sp3	+	+		+	+
M00216	V$TATA_C	TBP	−		−	−	
HIF							
M00236	V$ARNT_01	AhR (HIF1-β)	+	+		+	+
M00466	V$HIF1_Q5	HIF-1	+	+		+	+
EGR family							
M00243	V$EGR1_0	Egr-1	+	+		+	+
M00246	V$EGR2_01	Egr-2	+	+			+
M00244	V$NGFIC_01	NGFI-C	+	+		+	+
NF-κB							
M00051	V$NFKAPPAB50_01	NF-κB (P50)	+			+	+
M00052	V$NFKAPPAB65_01	NF-κB (P65)	+				+
Homeobox TFs							
M00102	V$CDP_02	CDP	−			−	−
M00437	V$CHX10_01	CHX10	−			−	−
M00206	V$HNF1_C	HNF-1	−			−	−
M00639	V$HNF6_Q6	HNF-6	−			−	
M00725	V$HP1SITEFACTOR_Q6	HP1	−			−	
M00241	V$NKX25_02	Nkx2-5	−	−			
M00485	V$NKX22_01	NKX2B	−	−		−	
M00424	V$NKX61_01	NKX6-1	−	−		−	
M00162	V$OCT1_06	POU2F1 (OCT1)	−		+	−	
M00145	V$BRN2_01	POU3F2	−			−	
M00465	V$POU6F1_01	POU6F1	−	−			
Other							
M00037	V$NFE2_01	NF-E2	+			+	
M00257	V$RREB1_01	RREB-1	+		−	+	+
M00532	V$RP58_01	RP58		+			+

Columns I and D indicate the result obtained for each matrix in RA induced and repressed genes, respectively; RARE corresponds to a random group of nonregulated genes containing putative RAREs in their regulatory regions: RARE I and RARE D are RA induced and repressed genes, respectively, that contains putative RAREs in their regulatory regions (see text). − indicates TFBS that are under-represented in the gene lists, whereas + indicates TFBS that are over-represented

Interestingly, many transcription factors whose functions have been associated to specific developmental and differentiation processes, including 11 homeobox-family TFBS, are under-represented in induced genes, particularly in those containing putative RAREs (Table 2). In other cases, such as the NF-E2 transcription factor, binding sites for proteins that favor differentiation are over-represented in promoters of induced genes, especially in those predicted to contain RAREs (Figure 6c). There is less specificity for TFBS in the promoters of repressed genes. We could only detect the RP58 binding site as specifically over-represented in repressed genes containing RAREs (Figure 6d). Interestingly, RP58 has been described as a heterochromatin-associated transcriptional repressor (Meng *et al*., 2000).

A new functional RARE has recently been identified in the promoter of the *BLR1* gene (Wang and Yen, 2004). RA treatment induces binding of a specific transcriptional complex, containing RARα, RXR, NFATc-3, Oct-1 (POU2F1) and CREB2, to this RARE. RA sensitivity, however, depends on the presence of functional binding motifs for POU2F1, NFATc-3 and CREB2. Our analysis reveals under-representation of both CREB and POU2F1 binding sites in the RA induced genes, suggesting that formation of this newly identified multiprotein complex is not directly involved in RA response in PML/RAR-expressing cells (Figure 6e and f). Taken together, analysis of upstream genomic regions of RA target genes in APL models reveals a nonrandom distribution of specific TFBS, particularly in genes induced by RA treatment.

Discussion

RA treatment causes differentiation of APL cells through PML/RAR-mediated transcriptional regulation. We analysed global gene expression modifications at an early time point of RA treatment in two models of PML/RAR expressing cells. First, we assessed gene expression profiles of RA-treated APL blasts with the aim of studying the physiological response to RA therapy, and found significant overlap of results

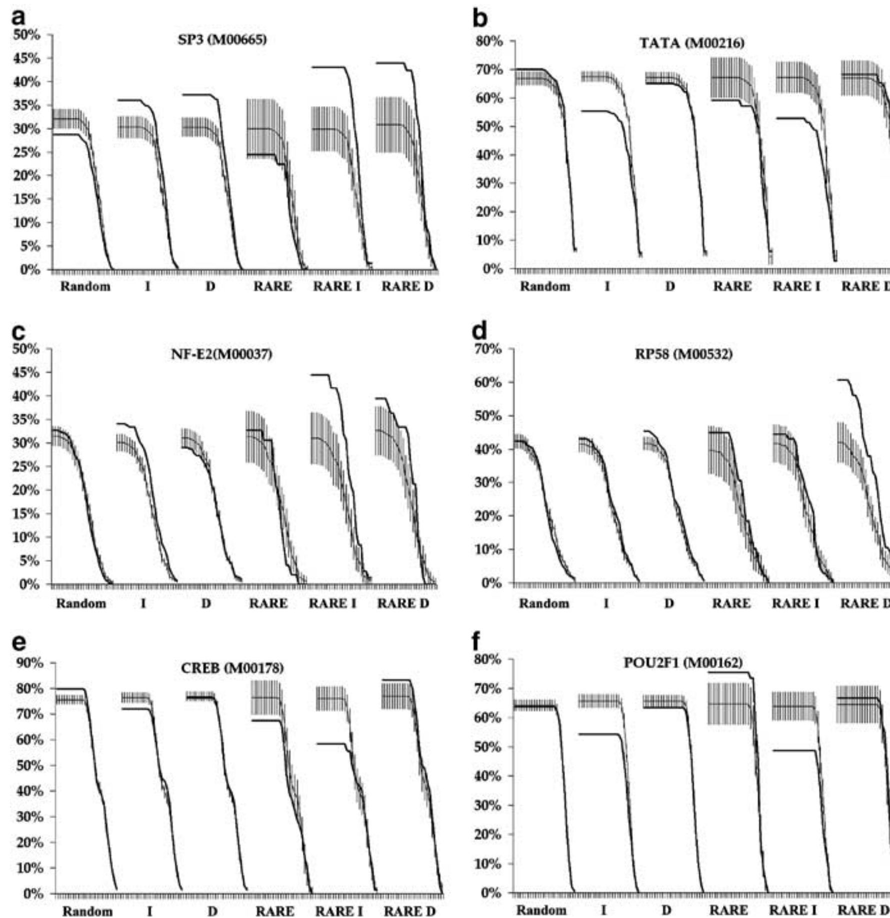

Figure 6 Distribution of specific TFBS predicted in the regulatory regions of common RA targets. The percentage of genes predicted to contain each TFBS is shown on the *y*-axis. Stringencies of the position weight matrix search for each of the groups of co-regulated genes are shown on the *x*-axis. Thin lines with error bars correspond to the distribution of the TFBS in a random series. Standard deviations are calculated as described in Materials and methods. Thicker lines correspond to specific groups of genes: a single random set is shown on the left side of each graph, followed by the distribution in the following groups of genes: I and D indicate all RA induced and repressed genes, respectively; RARE corresponds to a random group of nonregulated genes that contain putative RAREs in their regulatory regions, RARE I and RARE D are RA induced and repressed genes, respectively, that contains putative RAREs in their regulatory regions

from three different individuals. We then assessed PML/RAR-dependence of transcriptional modulation induced by RA by studying gene expression regulation in the U937 cell system (U937-PR and control U937-Mt cells, prior to and after RA treatment) and found that the vast majority of the identified target genes is modulated by RA only in the presence of PML/RAR. This result is in agreement with the finding that PML/RAR regulates transcription in response to RA through mechanisms that differ from its native counterpart RARα (Segalla *et al*., 2003).

Functional classification of common target genes in APL blasts showed that induction of many specific regulators of differentiation, repression of genes involved in stem cell renewal and/or inhibition of cell differentiation are early events in RA-dependent maturation. Some of the repressed genes classified in this category, including *RUNX1* (*AML1*) (Lacaud *et al*., 2002), *HHEX* (Guo *et al*., 2003), *ALCAM* (Ohneda *et al*., 2001), *WT1* (Alberta *et al*., 2003), *ETV6* (Wang *et al*., 1998), are involved in multiple hemopoietic stem cell (HSC) functions such as development, maintenance and homing. Many RA target genes have been found mutated and/or aberrantly expressed in malignant hemopoietic disorders.

Specific signaling pathways, responsible for the maintenance of stem cell renewal, are active in APL blasts, and need to be modulated by RA to achieve differentiation. We previously showed that AML fusion proteins induce expression of the Notch ligand Jagged-1 and repress the *LFNG* gene encoding for Notch-specific glycosyltransferase Lunatic Fringe, and that RA treatment causes opposite regulation of both genes (Alcalay *et al*., 2003). We now found that also *NOTCH1* and *NOTCH2* expression is activated by RA. Several studies suggest the relevance of Notch signaling in the maintenance of HSC, but other lines of evidence indicate its possible role in promoting commitment and differentiation (Radtke *et al*., 2002). The functional consequence of these modulations in the expression levels of

components of the Notch signaling pathway is currently being investigated.

It has recently been reported that AML fusion proteins activate Wnt signaling, thus enhancing self-renewal of leukemic stem cells (Muller-Tidow *et al*., 2004; Zheng *et al*., 2004). In accordance with these results our data indicate a repression of Wnt signalling during RA treatment.

Recruitment of chromatin modifiers by PML/RAR has previously been described, and is thought to play a relevant role in determining its oncogenic potential and in RA response (Grignani *et al*., 1998; He *et al*., 1998; Lin *et al*., 1999; Di Croce *et al*., 2002). Several genes involved in the regulation of chromatin function are modulated upon RA treatment of APL blasts. The availability of chromatin modifiers may represent an important mechanism in causing structural changes in local chromatin domains: in fact, competition among different regulatory complexes with equal binding affinity for specific loci might favor the most abundant one.

RA treatment determines a complex combination of different histone modifications that influence the final state of the chromatin. We have found deregulated expression of genes that encode for proteins involved in histone methylation and deacetylation, DNA methylation, and subunits of co-repressor complexes. In particular we found that the expression of three different histone methyltransferases (*G9A, DOT1L, SET7*) is modulated upon RA treatment. We suggest that a decrease of G9A at an early stage of RA treatment might be required for re-activation of genes involved in differentiation.

DNA methylation has been demonstrated to play a key role in PML/RAR mediated transcriptional repression (Di Croce *et al*., 2002). A decrease in DNA methyltransferase expression upon RA treatment may be of relevance to restore normal gene transcription (Fazi F *et al*., 2005), a finding that is in agreement with the reported efficacy of methyltransferase inhibitors in the treatment of hematological malignancies (Leone *et al*., 2003).

Mutual cross-talk between transcriptional regulatory pathways is essential for differentiation. We performed a study of the abundance of TFBS in the upstream genomic regions and/or the first exons of RA target genes. Surprisingly, putative RAREs are not enriched compared to a random set of genes, and are equally represented in induced and repressed genes. Instead, the high-affinity PML/RAR binding motifs identified by Kamashev *et al*. are strongly enriched in the set promoters of RA target genes. This enrichment can be identified only in alignments of human and mouse promoter regions, but is too weak to be observed using traditional methods in unaligned sequences, probably because of the large variability in spacer lengths. These findings may reflect a poor characterization of the canonic RAR binding consensus.

We found that other specific TFBS display a nonrandom distribution in the regulatory regions of RA target genes. Some enriched TFBS are bound by transcription factors that are, themselves, encoded by RA target genes. For example, in accordance with other studies (Park *et al*., 2003; Witcher *et al*., 2003), we found that RA regulates a number of genes involved in the TNF signaling pathway. RA enhances TNF-induced NF-κB binding to specific responsive elements (Witcher *et al*., 2003). Interestingly, we see an enrichment of NF-κB binding sites in the promoters of RA target genes, further suggesting a role for this pathway in regulating cell survival in response to RA.

Several TFBS are over or under-represented predominantly in the subset of sequences that contain putative RAREs. This finding suggests that the cis-regulatory capacity of PML/RAR in response to RA may depend on the presence of other TFBS. These data may help identify the existence of novel regulatory complexes involved in PML/RAR-dependent RA response.

Through the analysis of RA target genes we have identified regulation of functions that are crucial for reprogramming differentiation in APL cells, and revealed the distribution of specific TFBS in target gene promoters. Further analysis of these results will help clarify regulatory networks that may help improve leukemia therapy, especially in RA-resistant AML.

Materials and methods

Cell lines and patient samples

APL blasts were obtained at disease onset from peripheral blood of three patients that showed $\geqslant 75\%$ leukemic infiltration. Blasts were isolated by centrifugation on a Ficoll-Hypaque gradient (Lo Coco *et al*., 1992; Benedetti *et al*., 1996), and treated *in vitro* for 4 h with 10^{-6} M RA (Sigma-Aldrich, St Louis, MO, USA) prior to RNA extraction. The three patients resulted sensitive to RA therapy. The U937 PML/RAR#9 clone (PR9) was generated as previously described (Grignani *et al*., 1993; Alcalay *et al*., 2003); a bulk population of U937 cells transfected with the empty pSG-MtNEO vector (Mt) was used as the control. PR9 and Mt cells were treated for 12 h with 100 μM ZnSO4, then 10^{-6} M RA was added to the culture for 4 h prior to RNA extraction. APL blasts and U937 cell lines were maintained in RPMI-1640 supplemented with 50 U/ml penicillin, 50 μg/ml streptomycin, and 10% FCS at 37°C in a humidified atmosphere containing 5% CO_2.

RNA extraction

Total RNA was extracted using TRIzol reagent (Invitrogen, Carlsbad, California, USA) method, followed by clean up on RNeasy Mini/Midi Kit (QIAGEN, Valencia, CA, USA). For each of the U937 cell lines (PML/RAR and Mt), three independent vials were thawed, and the $ZnSO_4$ inductions, RA treatment and RNA extractions were performed separately. An equal quantity of each of the three RNA preparations was then mixed to generate an RNA pool for each sample.

Affymetrix GeneChip hybridization

Biotin-labeled cRNA targets were obtained from 5 μg of total RNA derived from samples described above. cDNA synthesis was performed with Gibco SuperScript Custom cDNA Synthesis Kit, and biotin-labelled antisense RNA was transcribed using

the *In vitro* Transcription System (Ambion Inc., Austin, TX, USA) including Bio-11-UTP and Bio-11-CTP in the reaction (NEN Life Sciences, PerkinElmer Inc., Boston, MA, USA). GeneChip hybridization, washing, staining and scanning were performed according to Affymetrix protocols (Santa Clara, CA, USA). Two copies of the entire HG-U133 chip set (HG-U133A, HG-133B) were hybridized with each target. Full details of microarray methods are described in Supplementary Methods 1.

Microarray data analysis

Absolute and comparative analyses were performed with Affymetrix Microarray Suite version 5 (MASv5) software, scaling all images to a value of 500. Results were further elaborated using the GenePicker software (Finocchiaro *et al*., 2004) and exported to GeneSpring software version 6.1 (Silicon Genetics, San Carlos, CA, USA) for visualization. Primary data are deposited in the ArrayExpress database (http://www.ebi.ac.uk/arrayexpress), accession no. E-MEXP-149, and elaborated data are available at Oncogene's website.

Promoter analysis

Genomic sequences of the first exons along with 1 kb of upstream regulatory regions were retrieved for 920 genes from TRASER (http://genome-www6.stanford.edu/cgi-bin/Traser/traser) using the following criteria: sequence to retrieve = upstream + first exon; length to retrieve = 1000 bp (of upstream sequence); restrict retrieval to transcript with most 5′ exon 1. RAREs were searched using Genomatix MatInspector software using optimized thresholds (Quandt *et al*., 1995). Over/under representation of a specific transcription factor binding site in a set of co-regulated genes was assessed by position weight matrix scanning as described (Kel *et al*., 2001; Vernell *et al*., 2003). Briefly, the percentage of genes predicted to contain a given binding site at a given search stringency is determined for search stringencies between 0.5 and 1.0, where 0.5 corresponds to over prediction (many false positives) and 1.0 corresponds to under prediction (many false negatives). The expected percentage of genes to contain the site if a list of genes was picked randomly from the genome was determined by averaging, for each search stringency, the percentage obtained from 50 randomly selected groups of genes where each group of genes contained the same number of genes as the list of co-regulated genes. The standard deviation of these values was calculated (for each stringency) and a site was considered over/under represented if the observed percentage of genes differed by at least three standard deviations for at least three search stringencies.

Real-time RT–PCR

PCR reactions were performed with either TaqMan or SYBR Green chemistries. SYBR Green chemistry: each 25 μl reaction contained: 0.5 μM of each primer (sequences selected with Primer Express software; PE Biosystems, Foster City, CA, USA), 12.5 μl of SYBR Green PCR Master MIX, 20 ng of template. Thermal Cycling Parameters were: 2 min at 50°C, followed by 10 min at 95°C, followed by 40 cycles of 15 s at 95°C, and 1 min at 60°C. The sequence of primers used can be found in Supplementary Materials 2. TaqMan chemistry: Applied Biosystems Assay-on-Demand products, consisting of pre-designed TaqMan probes, were used (http://events-na.appliedbiosystems.com/mk/get/MFC_LANDING?isource = fr_E_Pg_AB_Home_081803). Each sample was run in triplicate. The mean value of the replicates for each sample was calculated and expressed as previously described (Alcalay *et al*., 2003). GAPDH was used as endogenous control to calibrate the amount of mRNA target in different samples.

Acknowledgements
We are grateful to Roberta Bergomas, Elisa Venturini and Laura Tizzoni for excellent technical support and to Elisabetta Citterio and Lucilla Luzi for helpful discussions. This work was supported by grants from AIRC (Associazione Italiana per la Ricerca sul Cancro) to M Alcalay and PG Pelicci. N Meani and S Licciulli were supported by fellowships from Fondazione Italiana per la Ricerca sul Cancro.

References

Alberta JA, Springett GM, Rayburn H, Natoli TA, Loring J, Kreidberg JA and Housman D. (2003). *Blood*, **101,** 2570–2574.
Alcalay M, Meani N, Gelmetti V, Fantozzi A, Fagioli M, Orleth A, Riganelli D, Sebastiani C, Cappelli E, Casciari C, Sciurpi MT, Mariano AR, Minardi SP, Luzi L, Muller H, Di Fiore PP, Frosina G and Pelicci PG. (2003). *J. Clin. Invest.*, **112,** 1751–1761.
Alcalay M, Orleth A, Sebastiani C, Meani N, Chiaradonna F, Casciari C, Sciurpi MT, Gelmetti V, Riganelli D, Minucci S, Fagioli M and Pelicci PG. (2001). *Oncogene*, **20,** 5680–5694.
Benedetti L, Grignani F, Scicchitano BM, Jetten AM, Diverio D, Lo Coco F, Avvisati G, Gambacorti-Passerini C, Adamo S, Levin AA, Pelicci PG and Nervi C. (1996). *Blood*, **87,** 1939–1950.
Di Croce L, Raker VA, Corsaro M, Fazi F, Fanelli M, Faretta M, Fuks F, Lo Coco F, Kouzarides T, Nervi C, Minucci S and Pelicci PG. (2002). *Science*, **295,** 1079–1082.
Fazi F, Travaglini L, Carotti D, Palitti F, Diverio D, Alcalay M, McNamara S, Miller WH, Lo Coco F, Pelicci PG and Nervi C. (2005). *Ongogene*, in press.
Finocchiaro G, Parise P, Minardi SP, Alcalay M and Muller H. (2004). *Bioinformatics*, **20,** 3670–3672.
Grignani F, De Matteis S, Nervi C, Tomassoni L, Gelmetti V, Cioce M, Fanelli M, Ruthardt M, Ferrara FF, Zamir I, Seiser C, Lazar MA, Minucci S and Pelicci PG. (1998). *Nature*, **391,** 815–818.
Grignani F, Ferrucci PF, Testa U, Talamo G, Fagioli M, Alcalay M, Mencarelli A, Grignani F, Peschle C, Nicoletti I and Pelicci PG. (1993). *Cell*, **74,** 423–431.
Guo Y, Chan R, Ramsey H, Li W, Xie X, Shelley WC, Martinez-Barbera JP, Bort B, Zaret K, Yoder M and Hromas R. (2003). *Blood*, **102,** 2428–2435.
He LZ, Guidez F, Tribioli C, Peruzzi D, Ruthardt M, Zelent A and Pandolfi PP. (1998). *Nat. Genet.*, **18,** 126–135.
Hollander M and Wolfe D. (1999). *Nonparametric Statistical Methods*, 2nd edn. Wiley: New York.
Hosack DA, Dennis Jr G, Sherman BT, Lane HC and Lempicki RA. (2003). *Genome Biol.*, **4,** R70.
Huang ME, Ye YC, Chen SR, Chai JR, Lu JX, Zhoa L, Gu LJ and Wang ZY. (1988). *Blood*, **72,** 567–572.
Kamashev D, Vitoux D and De The H. (2004). *J. Exp. Med.*, **199,** 1163–1174.
Kel AE, Kel-Margoulis OV, Farnham PJ, Bartley SM, Wingender E and Zhang MQ. (2001). *J. Mol. Biol.*, **309,** 99–120.

Lacaud G, Gore L, Kennedy M, Kouskoff V, Kingsley P, Hogan C, Carlsson L, Speck N, Palis J and Keller G. (2002). *Blood*, **100,** 458–466.

Lee KH, Chang MY, Ahn JI, Yu DH, Jung SS, Choi JH, Noh YH, Lee YS and Ahn MJ. (2002). *Biochem. Biophys. Res. Commun.*, **296,** 1125–1133.

Leone G, Voso MT, Teofili L and Lubbert M. (2003). *Clin. Immunol.*, **109,** 89–102.

Lin RJ, Egan DA and Evans RM. (1999). *Trends Genet.*, **15,** 179–184.

Liu TX, Zhang JW, Tao J, Zhang RB, Zhang QH, Zhao CJ, Tong JH, Lanotte M, Waxman S, Chen SJ, Mao M, Hu GX, Zhu L and Chen Z. (2000). *Blood*, **96,** 1496–1504.

Lo Coco F, Diverio D, Pandolfi PP, Biondi A, Rossi V, Avvisati G, Rambaldi A, Arcese W, Petti MC, Meloni G, Mandelli F, Grignani F, Masera G, Barbui T and Pelicci PG. (1992). *Lancet*, **340,** 1437–1438.

Meng G, Inazawa J, Ishida R, Tokura K, Nakahara K, Aoki K and Kasai M. (2000). *Gene*, **242,** 59–64.

Minucci S and Pelicci PG. (1999). *Semin. Cell Dev. Biol.*, **10,** 215–225.

Muller-Tidow C, Steffen B, Cauvet T, Tickenbrock L, Ji P, Diederichs S, Sargin B, Kohler G, Stelljes M, Puccetti E, Ruthardt M, deVos S, Hiebert SW, Koeffler HP, Berdel WE and Serve H. (2004). *Mol. Cell Biol.*, **24,** 2890–2904.

Ohneda O, Ohneda K, Arai F, Lee J, Miyamoto T, Fukushima Y, Dowbenko D, Lasky LA and Suda T. (2001). *Blood*, **98,** 2134–2142.

Park DJ, Vuong PT, de Vos S, Douer D and Koeffler HP. (2003). *Blood*, **102,** 3727–3736.

Quandt K, Frech K, Karas H, Wingender E and Werner T. (1995). *Nucleic Acids Res.*, **23,** 4878–4884.

Radtke F, Wilson A, Ernst B and MacDonald HR. (2002). *Immunol. Rev.*, **187,** 65–74.

Segalla S, Rinaldi L, Kilstrup-Nielsen C, Badaracco G, Minucci S, Pelicci PG and Landsberger N. (2003). *Mol. Cell. Biol.*, **23,** 8795–8808.

Tachibana M, Sugimoto K, Nozaki M, Ueda J, Ohta T, Ohki M, Fukuda M, Takeda N, Niida H, Kato H and Shinkai Y. (2002). *Genes Dev.*, **16,** 1779–1791.

Tamayo P, Slonim D, Mesirov J, Zhu Q, Kitareewan S, Dmitrovsky E, Lander ES and Golub TR. (1999). *Proc. Natl. Acad. Sci. USA*, **96,** 2907–2912.

Vernell R, Helin K and Muller H. (2003). *J. Biol. Chem.*, **278,** 46124–46137.

Wang J and Yen A. (2004). *Mol. Cell. Biol.*, **24,** 2423–2443.

Wang LC, Swat W, Fujiwara Y, Davidson L, Visvader J, Kuo F, Alt FW, Gilliland DG, Golub TR and Orkin SH. (1998). *Genes Dev.*, **12,** 2392–2402.

Witcher M, Ross DT, Rousseau C, Deluca L and Miller Jr WH. (2003). *Blood*, **102,** 237–245.

Zheng X, Beissert T, Kukoc-Zivojnov N, Puccetti E, Altschmied J, Strolz C, Boehrer S, Gul H, Schneider O, Ottmann OG, Hoelzer D, Henschler R and Ruthardt M. (2004). *Blood*, **103,** 3535–3543.

Supplementary Information accompanies the paper on Oncogene website (http://www.nature.com/onc)