# Allele-specific DNA methylation capacitates *PEAR1* enhancer activity

Benedetta Izzi,[1] Mariaelena Pistoni,[2] Katrien Cludts,[1] Pinar Akkor,[1] Diether Lambrechts,[3] Catherine Verfaillie,[2] Peter Verhamme,[1] Kathleen Freson,[1] Marc F. Hoylaerts[1]

[1]Department of Cardiovascular Sciences, Center for Molecular and Vascular Biology, University of Leuven, Leuven, Belgium

[2]Department Development and Regeneration, Stem Cell Institute, University of Leuven, Leuven, Belgium

[3]Laboratory for Translational Genetics, Vesalius Research Center (VRC), Department of Oncology, VIB, University of Leuven, Leuven, Belgium

**Correspondence:**

Marc F. Hoylaerts

Center for Molecular and Vascular Biology

Department of Cardiovascular Sciences

University of Leuven

Herestraat 49, 3000, Leuven, Belgium.

E-mail: marc.hoylaerts@med.kuleuven.be

**Text word count:** 3993

**Abstract word count:** 235

**Number of Figures:** 6

**Number of Tables:** 0

**Number of References:** 60

**Key Points**

- **Rs12041331 is the first functional CpG-SNP related to platelet function whose regulatory mechanism depends on DNA methylation**
- **Rs12041331 marks allele-specific methylation at the CpG island encompassing the first untranslated exon during megakaryopoiesis**

**Abstract**

Genetic variation in the *PEAR1* locus is linked to platelet reactivity and cardiovascular disease. The major *G*-allele of rs12041331, an intronic CpG-SNP, is associated with higher *PEAR1* expression in platelets and endothelial cells than the minor *A*-allele. The molecular mechanism underlying this difference remains elusive. We have characterized the histone modification profiles of the intronic region surrounding rs12041331 and identified H3K4Me1 enhancer-specific enrichment for the region that covers the CpG-SNP. Interestingly, methylation studies revealed that the CpG site is fully methylated in leukocytes of *GG* carriers. Nuclear protein extracts from megakaryocytes, endothelial cells *vs.* control HEK-293 cells show a 3-fold higher affinity for the methylated *G*-allele, compared to non-methylated *G* or *A* alleles in a gel EMSA. To understand the positive relationship between methylation and gene expression, we studied DNA methylation at 4 different loci of *PEAR1* during *in vitro* megakaryopoiesis. During differentiation the CpG-SNP remained fully methylated, while we observed rapid methylation increases at the CpG-island overlapping the first UTR exon, paralleling the increased *PEAR1* expression. In the same region, *A*-allele carriers of rs12041331 showed significantly lower DNA methylation at CGI1 compared to *GG* homozygote. This CpG-island contains binding sites for the methylation-sensitive transcription factor CTCF, whose binding is known to play a role in enhancer activation and/or repression. In conclusion, we report the molecular characterization of the first platelet function-related CpG-SNP, a genetic predisposition that capacitates *PEAR1* enhancer activity through allele-specific DNA methylation.

## Introduction

*PEAR1* encodes the Platelet-endothelial aggregation receptor 1 (PEAR1), a transmembrane tyrosine kinase receptor predominantly expressed in platelets and endothelial cells.[1] In platelets, PEAR1 is involved in cell activation through its engagement in sustaining activation of the platelet integrin $\alpha_{IIb}\beta_{3,}$[2] and platelet FcεR1α has been identified as one of its activating ligands.[3] *PEAR1* is progressively more actively transcribed towards the later stages of megakaryocyte (MK) specification,[4] following a similar pattern as that observed for GATA1.[5,6] *PEAR1* knock-down in $CD34^+$ hematopoietic stem cells enhances megakaryopoietic progenitor proliferation,[4] while in endothelial cells it is implicated in the regulation of neo-angiogenesis.[7] Both processes occur in a PI3K/PTEN-dependent manner.

Several epidemiological studies have identified *PEAR1* as one potential candidate gene to explain platelet function variability both in healthy and pathological conditions.[8-19]
The non-coding *PEAR1* SNP rs12041331 A-allele was associated with lower *PEAR1* expression[9] and platelet response to agonists in different reports,[8,9,13,14,17,19] while among patients with coronary artery disease taking antiplatelet agents, rs12041331 *A*-allele carriers surprisingly appeared to experience more adverse cardiovascular outcomes and had higher death rates than *GG* homozygotes.[14] Rs12041331 seems to play a role in driving endothelial cell function too.[20] Rs12041331 is a polymorphism that removes a CpG-site in the DNA sequence, and is defined as a "CpG-SNP". CpG-sites in mammals influence gene regulation, dependent on their methylation status. DNA methylation has been extensively described in the control of cell specification, cancer, and other human diseases[21] and co-operates with histone modifications in rendering the chromatin more or less accessible for transcription factors, thereby influencing gene

expression.[22] Despite reported associations of CpG-SNPs with multiple diseases,[23-29] the role of CpG-SNPs in regulating platelet phenotypes was never investigated.

Two studies have already shown that *PEAR1* rs12041331 G/A substitution leads to lower platelet PEAR1 expression[9] and reduces endothelial cell migration in carriers of the A allele.[20] In this study we have addressed the transcriptional importance of CpG-SNP rs12041331 in the context of the *PEAR1* gene and have identified, for the first time, allele-specific methylation as a regulatory mechanism capable of modulating gene expression at more than one level. We found that this type of epigenetic regulation contributes to the fine-tuning of *PEAR1* expression and provides a molecular explanation for population variability in PEAR1-dependent platelet function.

## Methods

All human samples involved in the study were obtained from healthy volunteers after giving written informed consent. The study was approved by the Medical Ethical Committee of the University of Leuven (Reference number: S56976).

### Chromatin immunoprecipitation (ChIP) assay

ChIP experiments have been performed using the megakaryocytic CHRF cell line. The detailed protocol, adapted from Pistoni *et al.*,[30] is reported in **Supplemental methods**. We designed 10 amplicons P1-P10, (characteristics and genomic location are reported in **Table S1**) The ChIP-qPCR data relative to histones are presented as % input/tot H3: 2^(-(CT Histone modification – CT 10% input))/2^(-(CT Total H3 – CT 10% input)).

### Electrophoretic Mobility Shift Assays (EMSAs)

Nuclear proteins were extracted from CHRF-288-11, MEG01, EA.hy926 and HEK-293 cells and their total protein concentration was determined by Bradford analysis. Single stranded 33 bp 5`Biotin-labeled oligonucleotides were synthesized by Integrated DNA Technology (Integrated DNA Technologies, Leuven, Belgium) and then annealed to generate double stranded probes. Three probe combinations were generated containing either the A or the G allele with or without methylation at its preceding C, at the CpG-SNP position. Specific details on probes are reported in **Table S2**.

**Luciferase gene-reporter assay**

The PGL3 promoter vector and assay reagents were bought from Promega Corporation (Fitchburg, WI, USA). A 223-bp DNA fragment containing the CpG-SNP was amplified (chr1:156899830-156900052, GRCh38/hg38 Assembly) by PCR with DNA from two subjects homozygous for the rs1204331 A/G polymorphism (G/G and A/A, respectively) and inserted in a PGL3 promoter vector (further details on protocol are reported in the **Supplemental methods**). The PGL3 plasmids and GFP control vector were cotransfected into HEK-293 cells using Polyplus transfection kit (Polyplus, Illkirch, France). Enhancer activity was expressed by relative light units of luciferase normalized to the luciferase measured in empty pGL3-promoter vector. One-way ANOVA was used to determine the significance of differences in luciferase expression among PGL3 promoter, with a significance level defined at $p<0.05$.

**Human CD34$^+$ stem cell differentiation *in vitro***

Human CD34$^+$ hematopoietic stem cells (HSCs) were separated by magnetic cell sorting from buffy coats isolated from healthy donor peripheral blood (Milteny Biotech, Auburn, CA) using a protocol modified from[4] and described in **Supplemental methods**.
The cultured CD34$^+$ cells were harvested on days 0, 7 and 14, washed in Dulbecco phosphate-buffered saline (PBS) before the cell pellets were either snap-frozen in liquid

nitrogen and stored at -80°C for further DNA and RNA extraction or processed to detect CD34, CD41 and *PEAR1* expression by flow cytometry using APC-labeled CD34, FITC-labeled CD41 (Becton Dickinson) and anti-hPEAR1 PE (clone: FAB45278, R&D systems). *PEAR1* and *GAPDH* expression analysis by quantitative RT-PCR was performed as described in the **Supplemental methods**.

**DNA methylation analysis**

CpG-SNP methylation was assessed by bisulfite-sequencing or bisulfite cloning and sequencing, using primers reported in **Table S3**.

*PEAR1* (CGI1, CGI2 and intron 1 region) methylation was evaluated using the Sequenom EpiTYPER MassARRAY platform as described.[31,32] Details for the protocol used are described in the **Supplemental methods**. Statistical analysis was performed using One-way ANOVA with Friedman test, with a significance level of $p<0.05$ to study the progression of methylation during MK differentiation and unpaired t-test, with a significance level of $p<0.05$ to investigate differences between G- and A-allele carriers.

**Results**

**Transcriptional activity control by epigenetic modification of rs12041331**

*PEAR1* is encoded by 23 exons located on chromosome 1q23.1 (**Figure 1**) with two CpG islands covering 1) the first untranslated exon region (CGI1) and 2) the region comprising exons 8-9 (CGI2), both with a potential regulatory role in the expression of *PEAR1*. Several transcripts arise from the *PEAR1* promoter but only 2 of them are translated into a full receptor (**Figure 1**). The G to A polymorphic variant rs12041331 (CpG-SNP) is located within a non-coding region of about 10kb that links the first untranslated exon to the transcription start site (TSS) of *PEAR1*, located downstream the CpG-SNP **(Figure 1).** We examined the histone modification profile of this region obtained from ChIP-seq data in megakaryoblastic K562 and HUVEC endothelial cell lines

available in ENCODE. In general, gene activity is linked to high levels of H3K4Me3 and H3K27Ac.[33-37] Enhancer regions are typically enriched for H3K4Me1 marking while poised enhancers harbor H3K4me1 and H3K27Me3[12], and active enhancers are co-marked by H3K4Me1 and H3K27Ac.[38-40] From this first analysis, the *PEAR1* intronic region where the CpG-SNP is located shows enrichment for histone modifications H3K4Me1 and H3K27Ac, with particular higher deposition in HUVECs (light blue peaks in **Figure 1**), and a DNAse Hypersensitivity I-dependent open chromatin conformation (**Figure 1**).

*PEAR1* is mostly expressed in megakaryocytes, platelets and endothelial cells.[1] To be able to integrate and validate ENCODE data relative to histone modifications within the CpG-SNP region, we have used 4 different hematopoietic cell lines (erythroblastoid K562, myeloid MEG01 and megakaryocytic DAMI and CHRF) to identify the one with the highest expression of *PEAR1*. Via quantitative RT-PCR, we observed that the CHRF cells express 8 times higher *PEAR1* levels compared to K562, while MEG01 and DAMI show an intermediate *PEAR1* expression (**Figure S1**), CHRF cells were, therefore, used to perform a series of Chromatin Immunoprecipitation (ChIP) experiments. We have designed 10 different amplicons to cover the full *PEAR1* intron region, including the specific CpG-SNP position (amplicons P5-6) (**Figure 2A**). A GAPDH promoter-specific amplicon was used as positive or negative control to evaluate enrichment of 4 different histone modifications: H3K27Ac, H3K4Me3, H3K27Me3 and H3K4Me1. H3K27Ac and H3K4Me3 were strongly enriched at the GAPDH promoter amplicon, while a mild enrichment was observed for *PEAR1* amplicons P2 and P3, indicating that the *PEAR1* intron1 region in CHRF cells does not contain promoter activity (**Figure 2B**). On the contrary, H3K4Me1, a known histone modification often associated to enhancers, was also enriched in amplicons P1-P6, including the CpG-SNP region (P5 and P6) over the GAPDH promoter signal for the same histone, while it was detectable in amplicons P7 to P10 (just upstream to the ATG starting codon) in a comparable way as the GAPDH promoter amplicon (**Figure 2B**). To further characterize the enhancer features of the *PEAR1* intron1 region, we analyzed the H3K27Me3 deposition in the same CHRF samples. H3K27Me3 in combination with

H3K4Me1 defines closed or poised enhancers.[41] We found that its enrichment was not significantly different from the signal for the GAPDH promoter region, thus suggesting that the *PEAR1* intron1, including the CpG-SNP region, contains a primed enhancer with potential binding sites for transcription factors that can regulate *PEAR1* expression in the megakaryocytic specific CHRF cell line. The CpG site generated by the major G allele of *PEAR1* CpG-SNP is a unique event in a region of around 200bp downstream in the same locus, representing a relevant functional change that could influence gene regulation. We have genotyped leukocyte DNA from a total of 90 healthy individuals for the CpG-SNP, 60 European Caucasians and 30 East or South Asians. We identified 44 G/G (73.3), 16 G/A (26.7) in the Caucasian group and found 12 G/G (40.0), 14 G/A (46.7) and 4 A/A (13.3) individuals in the Asian group. Allelic frequency in both groups was in line with the Minor Allele Frequency of rs12041331 in Caucasian European (MAF = 0.9) and in East and South Asians (MAF = 0.46 and 0.37, respectively). We next investigated the methylation status of the CpG-SNP in 5 G/G, 3 G/A and 2 A/A individuals via bisulfite-cloning and sequencing. This analysis revealed a fully methylated status of the C preceding the G in all G-allele carriers (**Figure S2**). Full methylation for the C preceding the G was also observed in the different cell lines. All this evidence indicated that the G allele is always characterized by a fully methylated C.

To evaluate the potential transcriptional activity of the G *vs.* the A allele, we have used a luciferase gene reporter assay to investigate if a given DNA sequence is able to produce gene transcription as enhancer region. For this purpose we have used the pGL3-promoter vector in which a sequence of about 200bp flanking the CpG-SNP has been inserted. In agreement with previously reported findings,[9] the G allele was associated with considerably higher luciferase transcription than the A allele (**Figure S3**), demonstrating a sequence-dependent role of this CpG-SNP in driving differential gene expression. Even the vectors used for the luciferase-gene reporter assay, after they had been purified from competent bacteria, showed a constantly full methylation status for the G-allele, directly linking this specific genetic-epigenetic combination to the higher transcriptional activity in the reporter assay.

To investigate the influence of methylation on the nuclear protein binding properties for G allele, we performed a classical EMSA using 5`-biotinylated probes corresponding to location chr1:156899906-156899938 (GRCh38/hg38 Assembly) and including the CpG-SNP. More specifically, we have used 3 different double-stranded probes containing either the G allele with or without methylation at the C preceding the G (referred to as Gmeth and G, respectively) or the A allele, and incubated these with nuclear extracts from CHRF, MEG01, EA.hy926, but also HEK-293 cells. In all conditions, the Gmeth resulted in significantly increased recruitment of nuclear proteins compared to the G-allele, non-methylated at the preceding C, or the A allele (**Figure 3**), with different DNA-protein complexes formed for the various genotypes analyzed (**Figure 3**). Additional controls with scrambled 5`-biotinylated probes confirmed that increased nuclear protein binding only occurred when methylation was indeed present at the CpG-SNP position and not at another cytosine in the scrambled probe (not shown).

The formation of different DNA-protein complexes between the Gmeth versus A probes are suggestive for genotype-specific and methylation-dependent nuclear protein binding. To speculate whether the CpG-SNP disrupts possible transcription factor binding sites (TFBS), we have investigated the putative CpG-SNP region (25 bp) using the PROMO software with the presence of the G versus A-allele.[42,43] PROMO uses TRANSFAC databases to search for putative TFBS. PROMO analysis generated a prediction of at least 10 different interacting transcription factors for each of the alleles, but with 2 predicted to bind exclusively to the *G* –allele, i.e. ATF3 and c-Jun, while YY1 only binds only to the A-allele (**Figure 3**). However, this software does not account for the impact of methylation status for the G allele on nuclear protein binding.

**Role of CpG-SNP and DNA methylation in primary CD34$^+$ hematopoietic stem cell proliferation and differentiation to megakaryocytes**

Our group has previously reported a significant increase of PEAR1 expression in $CD34^+$ hematopoietic stem cells (HSC) during their differentiation into megakaryocytes (MK).[4]

To elucidate whether increase in *PEAR1* expression during MK differentiation would also be associated with the CpG-SNP, we have isolated $CD34^+$ cells from peripheral blood of 11 healthy donors (8 G/G, 2 A/G and 1 A/A), and have induced *in vitro* MK differentiation for 14 days. Cells have been harvested at day 0, day 7 (proliferation phase) and day 14 (maturation phase). During this time frame, MK differentiation (CD34 and CD41a) and PEAR1 expression were monitored via FACS analysis (**Figure 4A**, left panel**)**. The analysis shows that CD41a expression is intermediate at day 7, whereas PEAR1 expression already reached maximal expression levels at this time point. Analysis of the median fluorescence intensity at this time point for all individuals revealed further and progressive differentiation for CD41a from day 7 to 14, whereas PEAR1 remained unchanged during MK maturation (**Figure 4A**, right panel).

To evaluate a possible influence of the CpG-SNP on MK progenitors during the first 7 differentiation days, we performed a clonogenic culture assay of $CD34^+$ HSC isolated from 3 G/G, 2 A/G and 1 A/A carriers (all Asians). Total CFU-MK colonies were counted in a blinded manner at day 8 after isolation and while the total number of colonies was not different according to the genotype, we observed significant differences in the size of single MK colonies. The A-allele carriers had significantly larger CFU-MK colonies compared to G/G carriers that mainly have small MK colonies (**Figure 4B,** left panel). We did not obtain enough cells to study MK ploidy but genetic defects in MK maturation, ploidy and proplatelet formation (e.g. GATA1,[44,45] DIAPH1[46]) typically have enlarged and dense MK colonies. In addition, we have shown before that human *PEAR1* knock-down megakaryocyte progenitors proliferate more than progenitors that can upregulate *PEAR1* during differentiation.[4]

Quantitative RT-PCR (n=4) also confirmed that *PEAR1* expression increases during differentiation, as reported before[4] (data not shown). To investigate whether the PEAR1 expression would be different in cells committed towards MK maturation, we investigated the median fluorescence intensity of PEAR1 in CD41a positive cells

(selection shown in **Figure S4**). We were not able to identify significant differences in the CD41a-normalized PEAR1 expression between G/G and A-allele carriers (**Figure 4B, right panel**) at day 7 or day 14 of MK progenitor differentiation.

During HSC differentiation to MKs, the C preceding the G variant in the G-carriers was constantly fully methylated. To identify whether other methylation-sensitive loci would play a role in *PEAR1* expression, we selected 3 other regions within the *PEAR1* locus (**Figure 5A)** and performed methylation estimate measurements using the Sequenom EpiTYPER. This methylation analysis involves both CpG islands ("$P_{CGI1}$" and "$P_{CGI2}$" in **Figure 5A**) and the P3 intronic region where the strongest enrichment of histone H3K27Me3 and H3K4Me1 was observed in previous ChIP experiments (indicated as "$P_{intron1}$" in **Figure 5A**) for a total of 67 CpG sites. Significantly increasing methylation was only observed during MK progenitor differentiation for the CGI1 region, that co-localizes with the first untranslated exon of *PEAR1* but also is located close to the predicted enhancer region (**Figure 2**). We evaluated total methylation as an average of 32 CpG sites and observed that during MK differentiation it significantly increased from 3±1% at day 0 to 14±4% at day 14 (One-way ANOVA with Friedman test, p=0.0011) (**Figure 5B**). Specific individual CpG-site analysis showed that CpG7-8 and CpG19-22 significantly increased methylation from day 0 to day 7 and further mildly increased from day 7 to day 14 (**Figure 5C,** top panel). $P_{intron1}$ and $P_{CGI2}$ remained highly methylated throughout HSC to MK differentiation (**Figure 5B**).

Though overall *PEAR1* methylation was not different at $P_{Intron}$ and $P_{CGI2}$ between G/G and A-allele carriers, a significantly higher methylation for the CpG sites 7-8, 11-13, 19-22 and 31-32 in the $P_{CGI1}$ region was observed for G/G carriers at the end of differentiation day 14 (unpaired t-test, $p<0.05$) (**Figure 5D**).

**Discussion**

PEAR1 is highly expressed in platelets and megakaryocytes and participates in both platelet activation and megakaryocyte progenitor proliferation.[2,4] The minor A-allele of

rs12041331 was found to be associated with lower PEAR1 expression in platelets[9] and associations with inter-individual variability in the responses to antiplatelet drugs has been shown. A recent genome-wide study also associated rs12566888 (the only SNP in high linkage disequilibrium (LD) with rs12041331 in the all *PEAR1* gene region) with platelet aggregation. Despite this evidence, in a study involving two independent populations, while the minor A-allele was still linked to lower platelet function on aspirin, it was also found to be a risk factor for cardiovascular events.[14] All the evidence produced together with this apparently divergent result fostered our present investigations focusing on the molecular mechanisms by which this variant would modulate *PEAR1* expression. We hypothesized that such a mechanism of fine-tuning expression, able to interfere with residual platelet activation variability, as shown in the different epidemiological studies, could be mediated by a combination of genetic predisposition and epigenetic mechanisms.

*PEAR1* variants associated with platelet function and individual variability are located both in coding and in non-coding regions of the *PEAR1* gene and 4 (rs2768759, rs12041331, rs11264579 and rs3737224) out of 8 are CpG-SNPs. Three of these CpG-SNPs are positioned either upstream the CGI1 (rs2768759) or in intron1 of the *PEAR1* gene (rs12041331 and rs11264579). The only non CpG-SNP present in the same region (rs12566888) is in high LD with rs12041331.

Association studies between CpG-SNPs and certain phenotypes are increasing,[23-29] but the molecular characterization of their mechanism of action still remains a challenge.

Most of the CpG-SNPs functionally characterized are located in the 5'-UTR or promoter regions,[48-51] while *PEAR1* rs12041331 is located in a non-coding intronic region, yet with specific enhancer features. Independently of their genomic location, CpG-SNPs collectively appear to act by impeding or facilitating the binding of specific factors that define their functional role in gene expression. In the case of rs12041331, the latter is suggested by the enhanced DNA-protein binding for the major G allele, which we observed in EMSA coupled to DNA methylation. Our study is in line with previous reports on the impact of rs12041331 on platelet[9] and endothelial cell[20] function and it

adds epigenetic and allele-specific methylation as new regulatory mechanism for *PEAR1* regulation.

Three different transcription factors are predicted to bind to the G- or the A-allele: ATF3 and c-Jun and YY1, respectively. Even if the prediction of these factors does not take in account the influence of DNA methylation, it is interesting to know that both YY1 and c-Jun have been found to be highly enriched in enhancer regions in a number of cell lines including HUVECs.[47] In addition to that, a recent study defines the binding profile of TFs to the DNA Methyl Transferases (DNMTs), the enzymes responsible for the maintenance or the *de novo* deposition of DNA methylation.[52] ATF3, c-Jun and YY1 have all been found to interact with DNMT3L, the enzyme that participates in *de novo* DNA methylation by interacting with DNMT3A and DNMT3B. This interaction could be both activating or inhibiting the action of DNMT3L representing a possible mechanism by which the G-methylated allele contributes to the CGI1 *de novo* methylation during MK differentiation (**Figure 6**).

Methylation at rs12041331 is associated with a higher PEAR1 expression in platelets[9] and differential function in both platelets[9-11,14,16-18] and endothelial cells.[20] While an inverse correlation between DNA methylation and gene expression is very well known, the opposite is not yet well understood although reported in some large-scale studies[53] and highly dependent on the genomic context.[54-56] To better understand this unusual relation between DNA methylation and gene expression, we have studied *PEAR1* DNA methylation, during early and late megakaryopoiesis. We found that parallel progressive elevation of *PEAR1* CGI1 methylation and PEAR1 expression occurs in early differentiation. At the same time, we also observed that rs12041331 was constitutively methylated throughout differentiation. CpG-SNP allele-specific methylation occurs at the same time also at the CGI1 region of the gene, where A-allele carriers present with significantly lower methylated CpG sites. Taken together with the different methylation at CGI1, rs12041331 contributes to drive and finetune PEAR1 expression in platelets. This interpretation is in line with the reported platelet PEAR1 protein differences

between individuals with the two genotypes[9] as well as with the luciferase gene-reporter assay, which measured the isolated effect of the methylated G allele for rs12041331. Although we could not detect significant differences in MK PEAR1 expression at the end of differentiation, we hypothesize that its expression might be more important for proliferation at the very early stages of differentiation, where we observed differences in the size of the CFU-MK between A-allele carriers and *GG* homozygotes (**Figure 4B**, left panel). In fact PEAR1 reaches maximal expression levels already at day 7 of differentiation suggesting that a still earlier time point might better reveal differences in PEAR1 expression between the *GG* and *A*-allele carriers. We could not include such analysis in our present study, because of the limited sample availability (comprising a single *AA* carrier only) and to the necessity to already determine membrane protein expression studies on these samples (FACS), CFU-MK colony assays and DNA methylation studies. We did analyze PEAR1 expression as a function of the Forward Scatter, an index of cellular size, and hence, indirectly of megakaryocyte maturation, during which process the higher ploidy results in larger cell masses.

The CpG sites of the CGI1 regions that undergo the most significant methylation changes identified in the MK progenitor-specific analysis, reside within predicted CTCF binding sites. CTCF is an important player regulating complex processes including transcription, imprinting, long-range chromatin interactions and subnuclear localization in health and disease. Its binding has been shown to be highly influenced by DNA methylation both *in vitro* and *in vivo*, and, in combination with other regulatory mechanisms, drives spatial and temporal-dependent gene expression. ENCODE CTCF annotations for HUVECs that express high levels of PEAR1, show binding of CTCF at the CGI1 region, and moreover, specifically located in that region investigated by us in the methylation study (**Figure 1**). At the same region, also specific active enhancer-interacting promoter histone marks[47] are present (H3K27Ac and H3K4Me3), while H3K4Me1, typically defining enhancers (such as the rs12031441 region), is not comparably enriched. This region is highly conserved across vertebrates (**Figure 1, last track**) suggesting its possible relevance during evolution. CTCF binding may play an important role in mediating interactions

between the CGI1 region and the gene TSS, located approximately 10kb downstream. CTCF binding depends on higher CpG content[57] and DNA methylation restricting enhancer-mediated transcriptional inductions or working as a chromatin barrier separating active and repressive domains at gene promoters or enhancers.[58] CTCF binding at *PEAR1* CGI1 would then be prevented by elevated methylation, allowing the downstream locus to act as an active enhancer in the *PEAR1* transcription in *G*-carriers. This model is in line with the higher G-allele dependent expression and methylation observed. At the same time, *PEAR1* CGI1 could be the enhancer of more distant promoters along chromosome 1 influencing transcription of other genes[59].

In conclusion, we have mechanistically characterized the role of the CpG-SNP rs12041331 as genetic predisposition that capacitates *PEAR1* enhancer activity through allele-specific methylation. Future studies are needed to investigate whether this mechanism of gene regulation is highly specific for *PEAR1* exclusively or also applies to more platelet function- or megakaryopoiesis-related genes.

**Acknowledgments**

This work was supported by the "Fonds voor Wetenschappelijk Onderzoek (FWO) Vlaanderen" grant G0A6514N, by the "Programmafinanciering KU Leuven (PF/10/014)". B.I. and M.P. are both FWO post-doctoral fellows (12M2715N and 1288714N, respectively), P.V. holds a FWO senior clinical investigator grant (1801414N). We thank Ploi Yibmantasiri and Antonio Passarella for graphical assistance.

**Authorship Contributions**

B.I. designed and performed experiments, analyzed data and wrote the manuscript; M.P. performed experiments and advised on ChIP research methodology; K.C. and P.A. performed experiments; D.L. advised on Sequenom methodology and reviewed the manuscript; C.V., P.V. reviewed the manuscript; K.F. and M.F.H. designed experiments

and wrote the manuscript.

**Disclosure of Conflicts of Interest**

The authors declare no competing financial interests.

**References**

1. Nanda N, Bao M, Lin H, et al. Platelet endothelial aggregation receptor 1 (PEAR1), a novel epidermal growth factor repeat-containing transmembrane receptor, participates in platelet contact-induced activation. *J Biol Chem*. 2005;280(26):24680-24689.
2. Kauskot A, Di Michele M, Loyen S, Freson K, Verhamme P, Hoylaerts MF. A novel mechanism of sustained platelet alphaIIbbeta3 activation via PEAR1. *Blood*. 2012;119(17):4056-4065.
3. Sun Y, Vandenbriele C, Kauskot A, Verhamme P, Hoylaerts MF, Wright GJ. A human platelet receptor protein microarray identifies FcepsilonR1alpha as an activating PEAR1 ligand. *Mol Cell Proteomics*. 2015.
4. Kauskot A, Vandenbriele C, Louwette S, et al. PEAR1 attenuates megakaryopoiesis via control of the PI3K/PTEN pathway. *Blood*. 2013;121(26):5208-5217.
5. Crispino JD. GATA1 in normal and malignant hematopoiesis. *Semin Cell Dev Biol*. 2005;16(1):137-147.
6. Ferreira R, Ohneda K, Yamamoto M, Philipsen S. GATA1 function, a paradigm for transcription factors in hematopoiesis. *Mol Cell Biol*. 2005;25(4):1215-1227.
7. Vandenbriele C, Kauskot A, Vandersmissen I, et al. Platelet endothelial aggregation receptor-1: a novel modifier of neoangiogenesis. *Cardiovasc Res*. 2015.
8. Eicher JD, Xue L, Ben-Shlomo Y, Beswick AD, Johnson AD. Replication and hematological characterization of human platelet reactivity genetic associations in men from the Caerphilly Prospective Study (CaPS). *J Thromb Thrombolysis*. 2016;41(2):343-350.
9. Faraday N, Yanek LR, Yang XP, et al. Identification of a specific intronic PEAR1 gene variant associated with greater platelet aggregability and protein expression. *Blood*. 2011;118(12):3367-3375.
10. Herrera-Galeano JE, Becker DM, Wilson AF, et al. A novel variant in the platelet endothelial aggregation receptor-1 gene is associated with increased platelet aggregability. *Arterioscler Thromb Vasc Biol*. 2008;28(8):1484-1490.
11. Johnson AD, Yanek LR, Chen MH, et al. Genome-wide meta-analyses identifies seven loci associated with platelet aggregation in response to agonists. *Nat Genet*. 2010;42(7):608-613.
12. Jones CI, Bray S, Garner SF, et al. A functional genomics approach reveals novel quantitative trait loci associated with platelet signaling pathways. *Blood*. 2009;114(7):1405-1416.
13. Kim Y, Suktitipat B, Yanek LR, et al. Targeted deep resequencing identifies coding variants in the PEAR1 gene that play a role in platelet aggregation. *PLoS One*. 2013;8(5):e64179.
14. Lewis JP, Ryan K, O'Connell JR, et al. Genetic variation in PEAR1 is associated with platelet aggregation and cardiovascular outcomes. *Circ Cardiovasc Genet*. 2013;6(2):184-192.

15. Qayyum R, Becker LC, Becker DM, et al. Genome-wide association study of platelet aggregation in African Americans. *BMC Genet*. 2015;16:58.
16. Voora D, Horton J, Shah SH, Shaw LK, Newby LK. Polymorphisms associated with in vitro aspirin resistance are not associated with clinical outcomes in patients with coronary artery disease who report regular aspirin use. *Am Heart J*. 2011;162(1):166-172 e161.
17. Wurtz M, Nissen PH, Grove EL, Kristensen SD, Hvas AM. Genetic determinants of on-aspirin platelet reactivity: focus on the influence of PEAR1. *PLoS One*. 2014;9(10):e111816.
18. Xiang Q, Cui Y, Zhao X, Zhao N. Identification of PEAR1 SNPs and their influences on the variation in prasugrel pharmacodynamics. *Pharmacogenomics*. 2013;14(10):1179-1189.
19. Yao Y, Tang XF, Zhang JH, et al. Association of PEAR1 genetic variants with platelet reactivity in response to dual antiplatelet therapy with aspirin and clopidogrel in the Chinese patient population after percutaneous coronary intervention. *Thromb Res*. 2016;141:28-34.
20. Fisch AS, Yerges-Armstrong LM, Backman JD, et al. Genetic Variation in the Platelet Endothelial Aggregation Receptor 1 Gene Results in Endothelial Dysfunction. *PLoS One*. 2015;10(9):e0138795.
21. Breiling A, Lyko F. Epigenetic regulatory functions of DNA modifications: 5-methylcytosine and beyond. *Epigenetics Chromatin*. 2015;8:24.
22. Santoro SW, Dulac C. Histone variants and cellular plasticity. *Trends Genet*. 2015;31(9):516-527.
23. Harlid S, Ivarsson MI, Butt S, et al. A candidate CpG SNP approach identifies a breast cancer associated ESR1-SNP. *Int J Cancer*. 2011;129(7):1689-1698.
24. Taqi MM, Bazov I, Watanabe H, et al. Prodynorphin CpG-SNPs associated with alcohol dependence: elevated methylation in the brain of human alcoholics. *Addict Biol*. 2011;16(3):499-509.
25. Polsinelli G, Zai CC, Strauss J, Kennedy JL, De Luca V. Association and CpG SNP analysis of HTR4 polymorphisms with suicidal behavior in subjects with schizophrenia. *J Neural Transm*. 2013;120(2):253-258.
26. Dayeh TA, Olsson AH, Volkov P, Almgren P, Ronn T, Ling C. Identification of CpG-SNPs associated with type 2 diabetes and differential DNA methylation in human pancreatic islets. *Diabetologia*. 2013;56(5):1036-1046.
27. Ye H, Zhou A, Hong Q, et al. Association of seven thrombotic pathway gene CpG-SNPs with coronary heart disease. *Biomed Pharmacother*. 2015;72:98-102.
28. Mansego ML, Milagro FI, Zulet MA, Martinez JA. SH2B1 CpG-SNP is associated with body weight reduction in obese subjects following a dietary restriction program. *Ann Nutr Metab*. 2015;66(1):1-9.
29. Koestler DC, Chalise P, Cicek MS, et al. Integrative genomic analysis identifies epigenetic marks that mediate genetic risk for epithelial ovarian cancer. *BMC Med Genomics*. 2014;7:8.

30. Pistoni M, Verrecchia A, Doni M, Guccione E, Amati B. Chromatin association and regulation of rDNA transcription by the Ras-family protein RasL11a. *EMBO J*. 2010;29(7):1215-1224.
31. Izzi B, Decallonne B, Devriendt K, et al. A new approach to imprinting mutation detection in GNAS by Sequenom EpiTYPER system. *Clin Chim Acta*. 2010;411(23-24):2033-2039.
32. Izzi B, Francois I, Labarque V, et al. Methylation defect in imprinted genes detected in patients with an Albright's hereditary osteodystrophy like phenotype and platelet Gs hypofunction. *PLoS One*. 2012;7(6):e38579.
33. Pokholok DK, Harbison CT, Levine S, et al. Genome-wide map of nucleosome acetylation and methylation in yeast. *Cell*. 2005;122(4):517-527.
34. Schubeler D, MacAlpine DM, Scalzo D, et al. The histone modification pattern of active genes revealed through genome-wide chromatin analysis of a higher eukaryote. *Genes Dev*. 2004;18(11):1263-1271.
35. Schneider R, Bannister AJ, Myers FA, Thorne AW, Crane-Robinson C, Kouzarides T. Histone H3 lysine 4 methylation patterns in higher eukaryotic genes. *Nat Cell Biol*. 2004;6(1):73-77.
36. Heintzman ND, Ren B. Finding distal regulatory elements in the human genome. *Curr Opin Genet Dev*. 2009;19(6):541-549.
37. Heintzman ND, Hon GC, Hawkins RD, et al. Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature*. 2009;459(7243):108-112.
38. Calo E, Wysocka J. Modification of enhancer chromatin: what, how, and why? *Mol Cell*. 2013;49(5):825-837.
39. Creyghton MP, Cheng AW, Welstead GG, et al. Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc Natl Acad Sci U S A*. 2010;107(50):21931-21936.
40. Pekowska A, Benoukraf T, Zacarias-Cabeza J, et al. H3K4 tri-methylation provides an epigenetic signature of active enhancers. *EMBO J*. 2011;30(20):4198-4210.
41. Shlyueva D, Stampfel G, Stark A. Transcriptional enhancers: from properties to genome-wide predictions. *Nat Rev Genet*. 2014;15(4):272-286.
42. Messeguer X, Escudero R, Farre D, Nunez O, Martinez J, Alba MM. PROMO: detection of known transcription regulatory elements using species-tailored searches. *Bioinformatics*. 2002;18(2):333-334.
43. Farre D, Roset R, Huerta M, et al. Identification of patterns in biological sequences at the ALGGEN server: PROMO and MALGEN. *Nucleic Acids Res*. 2003;31(13):3651-3653.
44. Li Z, Godinho FJ, Klusmann JH, Garriga-Canut M, Yu C, Orkin SH. Developmental stage-selective effect of somatically mutated leukemogenic transcription factor GATA1. *Nat Genet*. 2005;37(6):613-619.
45. de Waele L, Freson K, Louwette S, et al. Severe gastrointestinal bleeding and thrombocytopenia in a child with an anti-GATA1 autoantibody. *Pediatr Res*. 2010;67(3):314-319.

46. Stritt S, Nurden P, Turro E, et al. A gain-of-function variant in DIAPH1 causes dominant macrothrombocytopenia and hearing loss. *Blood*. 2016.
47. Whalen S, Truty RM, Pollard KS. Enhancer-promoter interactions are encoded by complex genomic signatures on looping chromatin. *Nat Genet*. 2016.
48. Moser D, Ekawardhani S, Kumsta R, et al. Functional analysis of a potassium-chloride co-transporter 3 (SLC12A6) promoter polymorphism leading to an additional DNA methylation site. *Neuropsychopharmacology*. 2009;34(2):458-467.
49. Reynard LN, Bui C, Syddall CM, Loughlin J. CpG methylation regulates allelic expression of GDF5 by modulating binding of SP1 and SP3 repressor proteins to the osteoarthritis susceptibility SNP rs143383. *Hum Genet*. 2014;133(8):1059-1073.
50. Reynard LN, Bui C, Canty-Laird EG, Young DA, Loughlin J. Expression of the osteoarthritis-associated gene GDF5 is modulated epigenetically by DNA methylation. *Hum Mol Genet*. 2011;20(17):3450-3460.
51. Oertel BG, Doehring A, Roskam B, et al. Genetic-epigenetic interaction modulates mu-opioid receptor regulation. *Hum Mol Genet*. 2012;21(21):4751-4760.
52. Pacaud R, Sery Q, Oliver L, Vallette FM, Tost J, Cartron PF. DNMT3L interacts with transcription factors to target DNMT3L/DNMT3B to specific DNA sequences: role of the DNMT3L/DNMT3B/p65-NFkappaB complex in the (de-)methylation of TRAF1. *Biochimie*. 2014;104:36-49.
53. Shen Y, Takahashi M, Byun HM, et al. Boswellic acid induces epigenetic alterations by modulating DNA methylation in colorectal cancer cells. *Cancer Biol Ther*. 2012;13(7):542-552.
54. Xu J, Pope SD, Jazirehi AR, et al. Pioneer factor interactions and unmethylated CpG dinucleotides mark silent tissue-specific enhancers in embryonic stem cells. *Proc Natl Acad Sci U S A*. 2007;104(30):12377-12382.
55. Uhm TG, Lee SK, Kim BS, et al. CpG methylation at GATA elements in the regulatory region of CCR3 positively correlates with CCR3 transcription. *Exp Mol Med*. 2012;44(4):268-280.
56. Unoki M, Nakamura Y. Methylation at CpG islands in intron 1 of EGR2 confers enhancer-like activity. *FEBS Lett*. 2003;554(1-2):67-72.
57. Fang R, Wang C, Skogerbo G, Zhang Z. Functional diversity of CTCFs is encoded in their binding motifs. *BMC Genomics*. 2015;16:649.
58. Lee BK, Iyer VR. Genome-wide studies of CCCTC-binding factor (CTCF) and cohesin provide insight into chromatin structure and regulation. *J Biol Chem*. 2012;287(37):30906-30913.
59. Sanyal A, Lajoie BR, Jain G, Dekker J. The long-range interaction landscape of gene promoters. *Nature*. 2012;489(7414):109-113.
60. Gardiner-Garden M, Frommer M. CpG islands in vertebrate genomes. *J Mol Biol*. 1987;196(2):261-282.

**Legends to figures**

**Figure 1. *PEAR1* gene structure (chr1:156893731-156916434, GRCh38/hg38 Assembly).** Exons are indicated by full blue boxes. The ATG codon is indicated at exon1. The major *PEAR1* transcript (ENST00000292357), translated into the full PEAR1 receptor, is indicated as "X" (in magenta). Two CpG islands are located in the gene (CGI1 and CGI2, depicted as green boxes on the left and right side, respectively. Location chromosome 1:156892430-156893919 and 156907978-156908857, GRCh38/hg38 Assembly). Rs12041331 (chr1:156899922, GRCh38/hg38 Assembly) is indicated in red. CpG islands are defined based on the formula described by Gardiner-Garden).[60] H3K4Me1, H3K4Me3 and H3K27Ac profiles in HUVEC and K562 cell lines are displayed as colored overlayed histograms (light blue for HUVEC and purple for K562) in "auto-scale to data view" mode that takes the highest signal in the selected region as the 100% of the intensity and display all other signals accordingly (data produced by the Bernstein Lab at the Broad Institute and the UCSC and part of the ENCODE database). DNase Hypersensitivity 1 regions are displayed as gray to black boxes (from less to more open chromatin conformation). CTCF ChIP-seq data from HUVECs and K562 cell lines are depicted as colored histograms in full visualization and are normalized for IgG signals (data produced by ENCODE). *PEAR1* conservation across vertebrates is displayed as blue histograms at the bottom of the figure using the Vertebrate Multiz Alignment & Conservation (100 Species) UCSC track. "$P_{CGI1}$" (also depicted as black box) refers to the region we analyzed for DNA methylation estimates in $CD34^+$ hematopoietic stem cells into megakaryocyte precursors and better described in figure 5. A yellow box (discontinue line) highlights the $P_{CGI1}$ region that co-localizes with CTCF binding in HUVECs and shows conservation across different vertebrates. Exons 15 and 16 are indicated in brown and indicate the region used for real-time q-PCR quantification of *PEAR1* expression in the present study.

**Figure 2. Histone modifications profiling of *PEAR1* intron1 region via Chromatin Immuno-precipitation (ChIP)**

**(A)** Graphical representation of PEAR1 intron1 region. Primers were randomly selected to cover the full intronic region, including rs12041331 (amplicons P5 and P6 in correspondence of rs12041331 indicated above the gene line) and are depicted as black boxes. **(B)** Levels of CHRF PEAR1 intron1 (P1-P10) and GAPDH amplicons immunoprecipitated using H3K27me3, H3K4me3, H3K27ac and H3K4me1 antibodies measured following polymerase chain reaction (PCR) amplification of DNA fragments. Signals from IgG control antibody used for each ChIP were subtracted from each of the histone modification IPs signals and the fold enrichment ratios of ChIP enriched versus total input DNA were represented (n=4 for each histone, average ± SEM).

**Figure 3. Transcriptional activity of rs12041331 and enhancing role of methylation at the G allele.**

EMSA analysis to characterize the methylated G-allele *vs.* the G-allele and A-allele binding properties to nuclear proteins. EMSA were performed using 5`biotinylated probes corresponding to location chr1:156869698-156869731 (GRCh38/hg38 Assembly) and including rs12041331. **(A,C)** Biotinylated double stranded probes containing the G allele with ('Gmeth') or without ('G') methylation, or the A allele ('A') were incubated in the presence of a 10X amount of cold probe (lanes 1-3), or alone with a fixed amount of nuclear extract from CHRF (**A**) or MEG01, EA.hy926 or HEK-293 cells (**C**). (lanes 4-6). Lanes 7-9 show unbound biotinylated probes in all cases. Methylation significantly enhanced binding of the probe to nuclear extract proteins as shown by the increased normalized intensity of the signal in lane 4 *vs.* that in lane 5 and 6, quantified for 3 replicates of CHRF cells (A, mean ± SD) and for the other 3 cell lines investigated (**C**), expressed as lane Integrated Density. (D) PROMO based putative binding sites search for CpG-SNP region (25bp). The G-allele (top panel) produced specific binding prediction for c-Jun (indicated as number 7) and ATF3 (indicated as number 8), while for the A-allele (bottom panel) YY1 (indicated as number 3) is predicted to bind the SNP site.

**Figure 4. Influence of Rs12041331 on MK progenitor proliferation and PEAR1 expression during megakaryocyte differentiation**

**(A)** Left panel: histogram for CD41a and PEAR1 during differentiation of $CD34^+$ cells from day 0 to day 14, as indicated. Right panel: CD34, CD41a and PEAR1 expression measured by FACS analysis and expressed as Mean Fluorescence Intensity during differentiation of $CD34^+$ cells into MK precursors harvested at day 0 (D0), day 7 (D7) and day 14 (D14). **(B**) Left panel: Number of CFU-MK, defined as small (3-20 cells), medium (20-80 cells) and large (>80 cells) colonies, counted in a blind fashion at day 8 after differentiation in *GG* (n=3) *vs. AG+AA* (n=3) individuals. *p<0.05, multiple unpaired t-test with post Holm-SIdak for multiple comparisons. Right panel: *GG* (n=2) *vs. AG+AA* (n=3) PEAR1/CD41a ratios measured on CD41a positive events at day 7 of differentiation. SP1, SP2, SP3 are CD41a positive subpopulations defined by marker positivity and size (Forward Scatter, FSC) as defined in figure S4.

**Figure 5. Increased PEAR1 expression during megakaryopoiesis parallels DNA methylation enrichment at CGI1.**

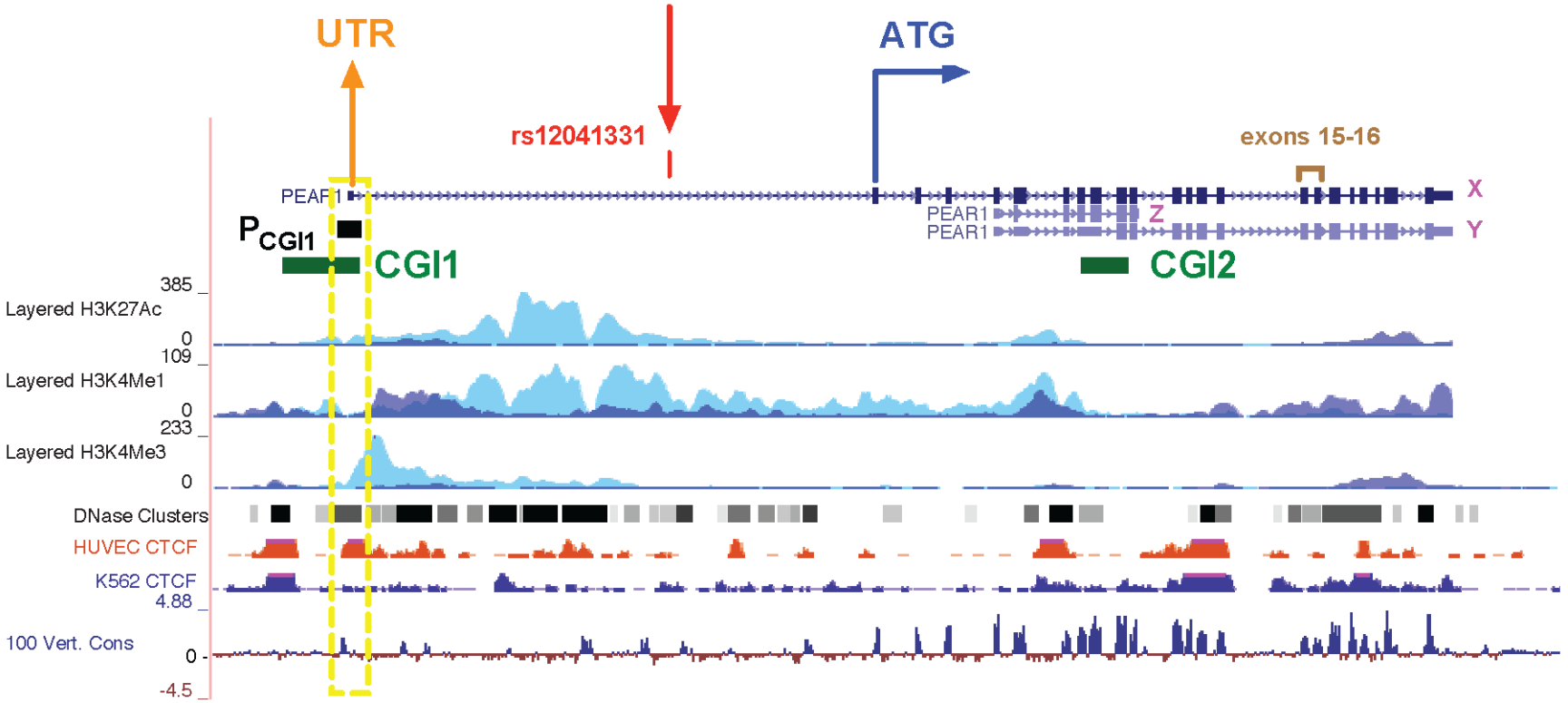
**(A)** Schematic representation of PEAR1 regions investigated in the methylation study indicated as black vertical boxes ('$P_{CGI1}$' in CpG-island 1 (CGI1), '$P_{intron1}$' in intron1 of PEAR1, 'rs12041331', '$P_{CGI2}$' in CpG-island 2 (CGI2)). **(B)** Sequenom methylation estimates (% average) for $P_{CGI1,}$ $P_{intron1}$ and $P_{CGI2}$ (bar representation, average+/-SD, n=8). **(C)** Single CpG (dot line representation), during differentiation of $CD34^+$ cells into MK precursors harvested at day 0 (D0), day 7 (D7) and day 14 (D14) and (**D**) of *GG*(n=5) *vs. AG+AA*(n=3) methylation at day 14. *p<0.05 one-way ANOVA.

**Figure 6. Epigenetic-dependent regulatory mechanism of rs12041331 in PEAR1 expression**

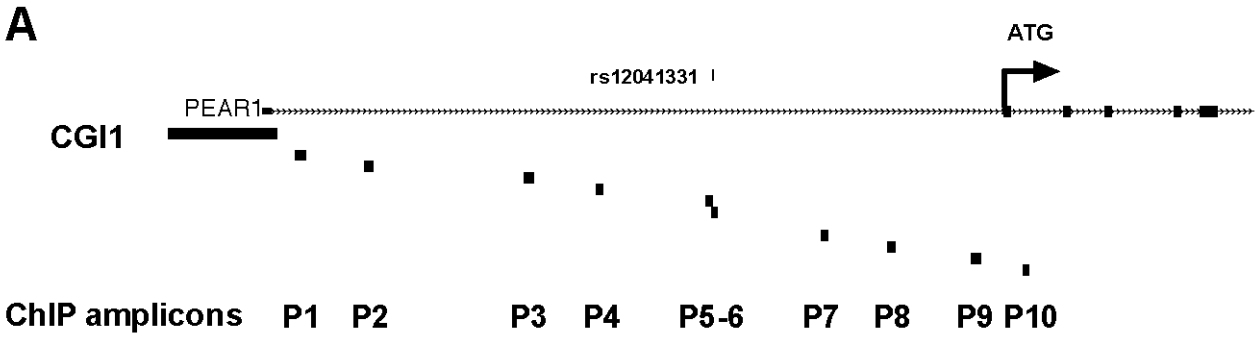
The major G-allele of the CpG-SNP, introduces a CpG-site in the enhancer region that binds with higher affinity to several nuclear proteins than the A allele. At the same time,

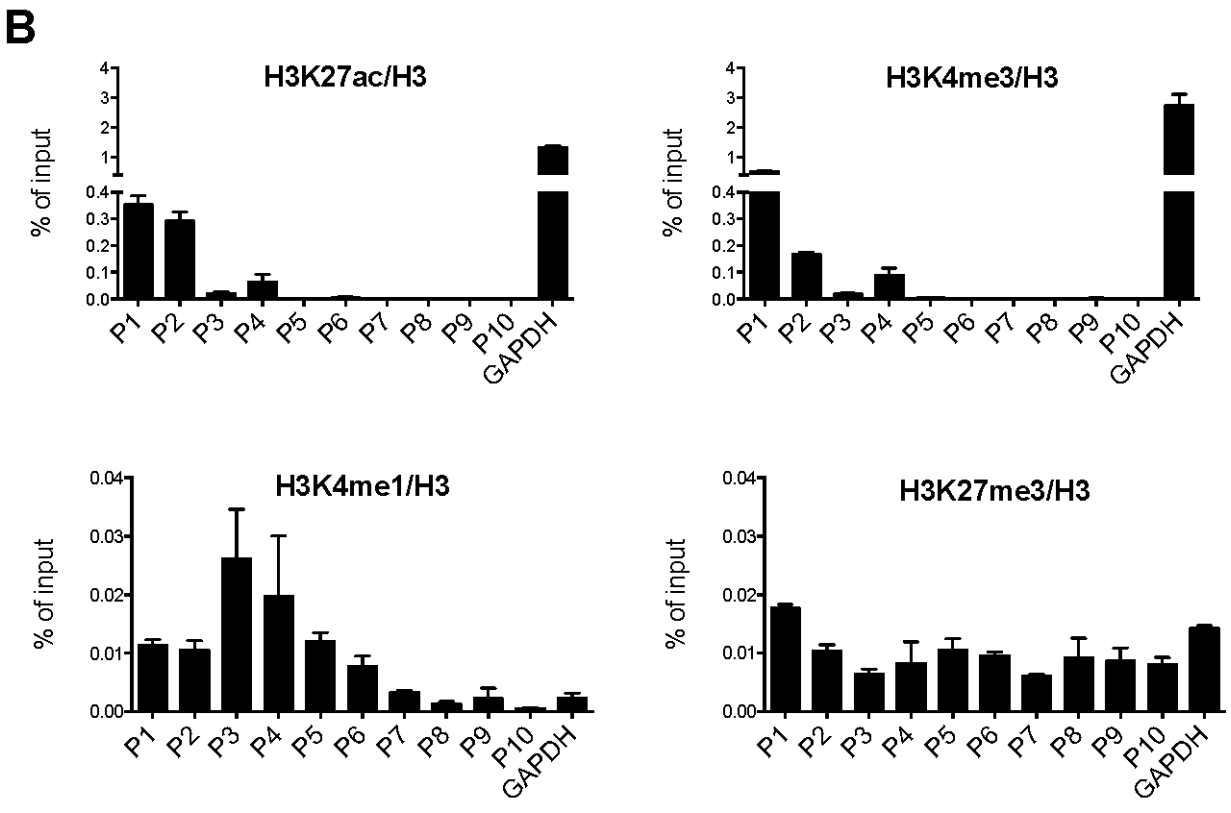
the G-methylated allele, marks higher methylation at CGI1, the CpG island located at the UTR region of PEAR1 that contains several binding sites for CTCF, important for chromosome looping and architecture. When higher methylation is present, CTCF binds with lower affinity to the CGI1 region, liberating the enhancer activity of intron 1 where the CpG-SNP is also located. On the contrary, lower methylation at CGI1 is associated with the minor A-allele of the CpG-SNP, reinforcing CTCF binding and partially inhibiting gene expression. This model advances a positive effect of methylation on gene expression, and is in accordance with the association between increasing methylation and PEAR1 expression during early MK progenitor development.

# Figure 1


UTR
ATG
rs12041331
exons 15-16
PEAR1
PEAR1
PEAR1
X
Z
Y
$P_{CGI1}$
CGI1
CGI2
385
Layered H3K27Ac
0
109
Layered H3K4Me1
0
233
Layered H3K4Me3
0
DNase Clusters
HUVEC CTCF
K562 CTCF
4.88
100 Vert. Cons
0
-4.5

# Figure 2

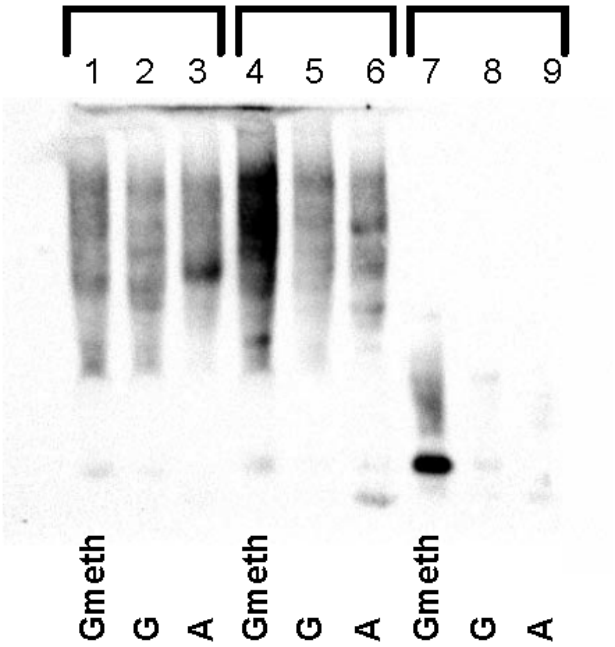

A
ATG
rs12041331
PEAR1
CGI1
ChIP amplicons
P1
P2
P3
P4
P5-6
P7
P8
P9
P10


B
H3K27ac/H3
H3K4me3/H3
H3K4me1/H3
H3K27me3/H3
% of input
P1
P2
P3
P4
P5
P6
P7
P8
P9
P10
GAPDH

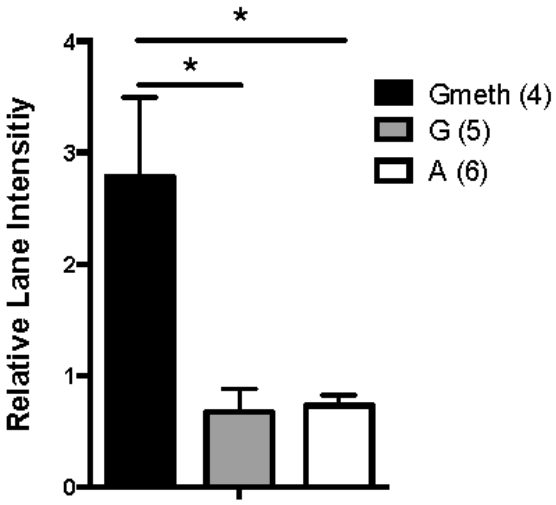
# Figure 3

**A**

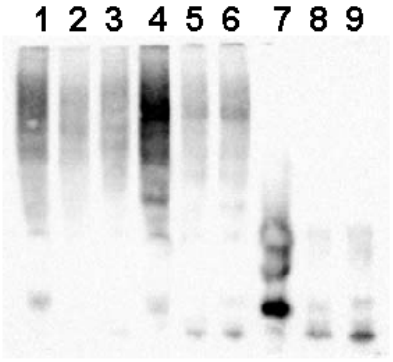
| | | | |
|---|---|---|---|
| + | + | - | Nuclear extracts |
| + | + | + | biotinylated probe |
| + | - | - | cold probe |



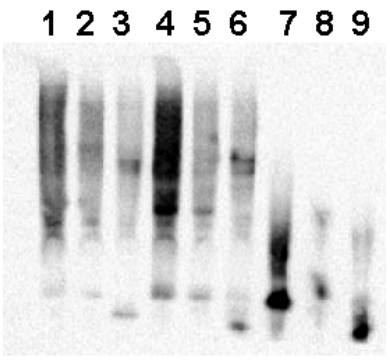

**B**



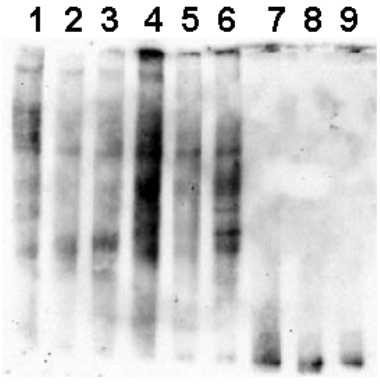

**C**



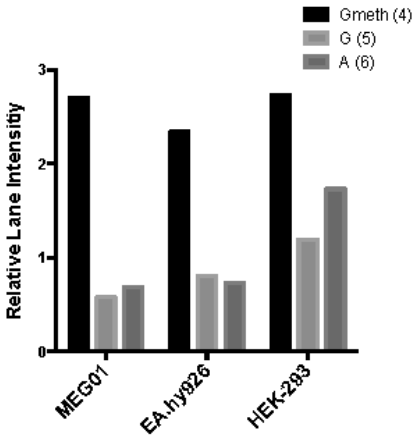

MEG01



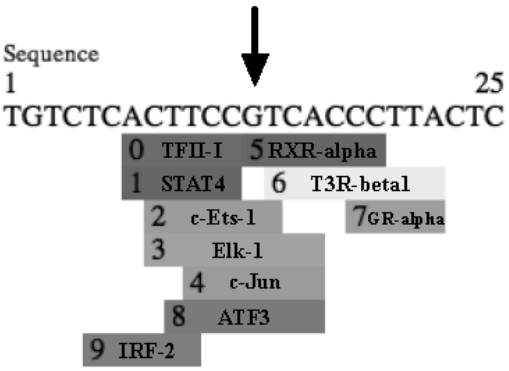

EA.hy926

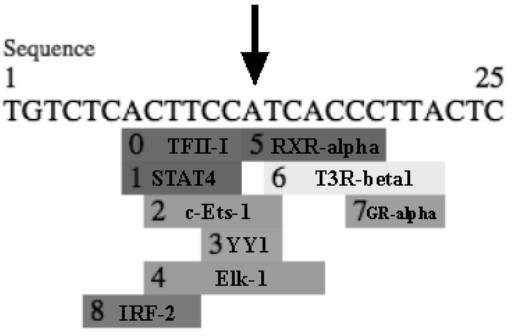



HEK-293



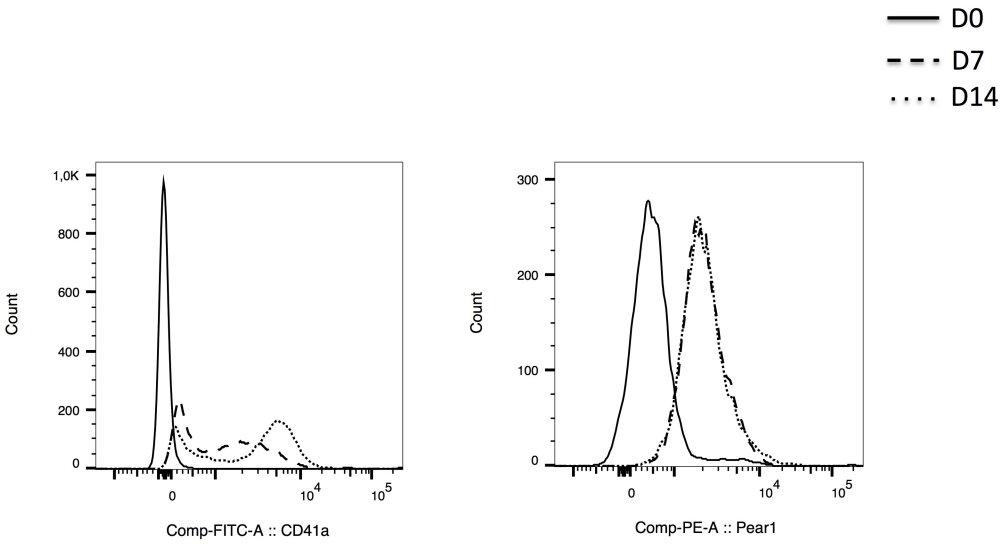

**D**

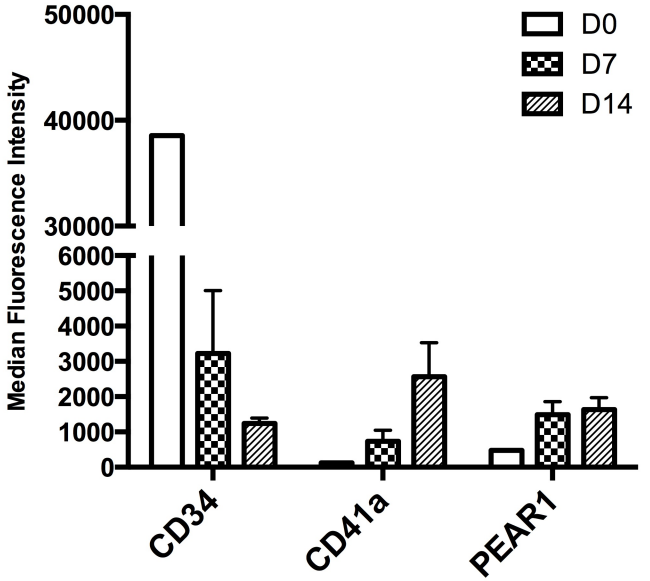


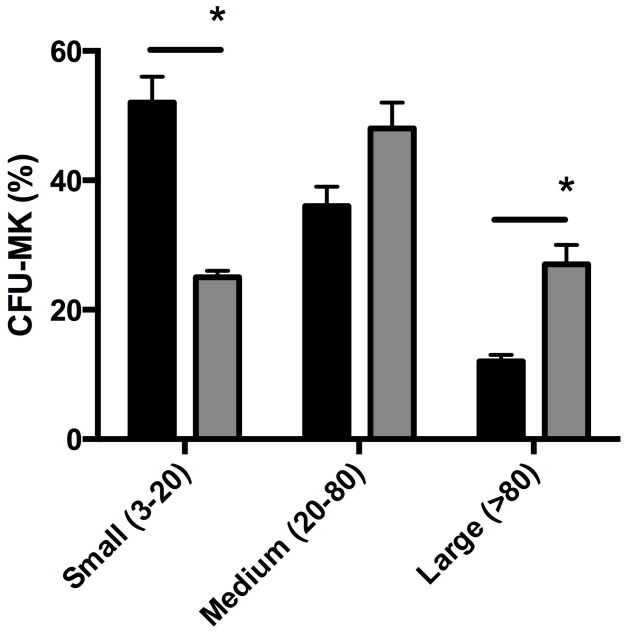
# Figure 4

**A**


D0
D7
D14
Count
Comp-FITC-A :: CD41a
Comp-PE-A :: Pear1


D0
D7
D14
Median Fluorescence Intensity
CD34
CD41a
PEAR1

**B**


CFU-MK (%)
Small (3-20)
Medium (20-80)
Large (>80)

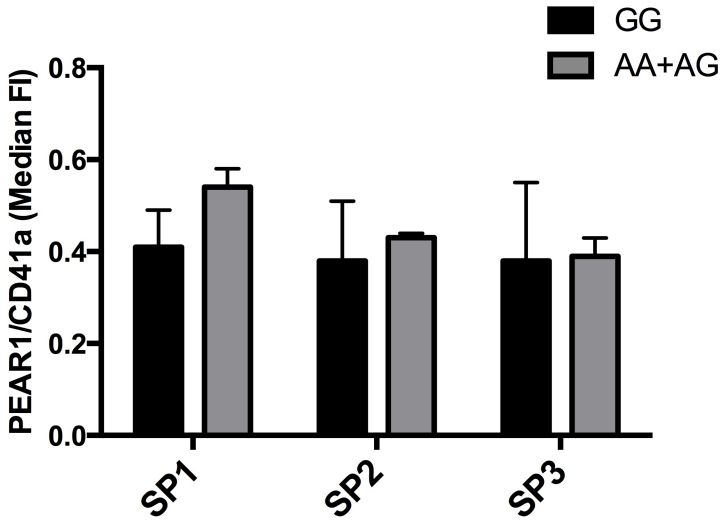

GG
AA+AG
PEAR1/CD41a (Median FI)
SP1
SP2
SP3

Figure 5

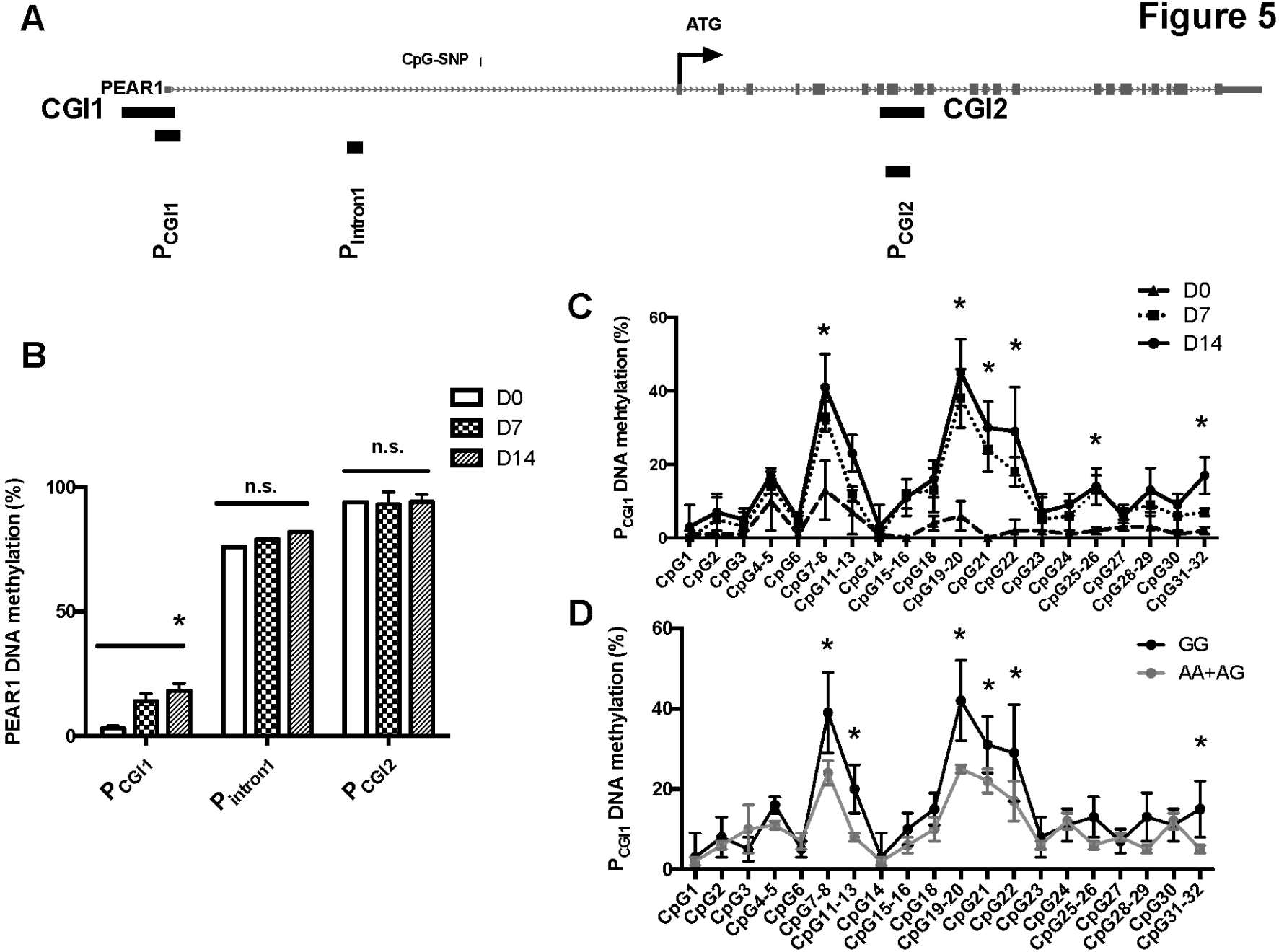
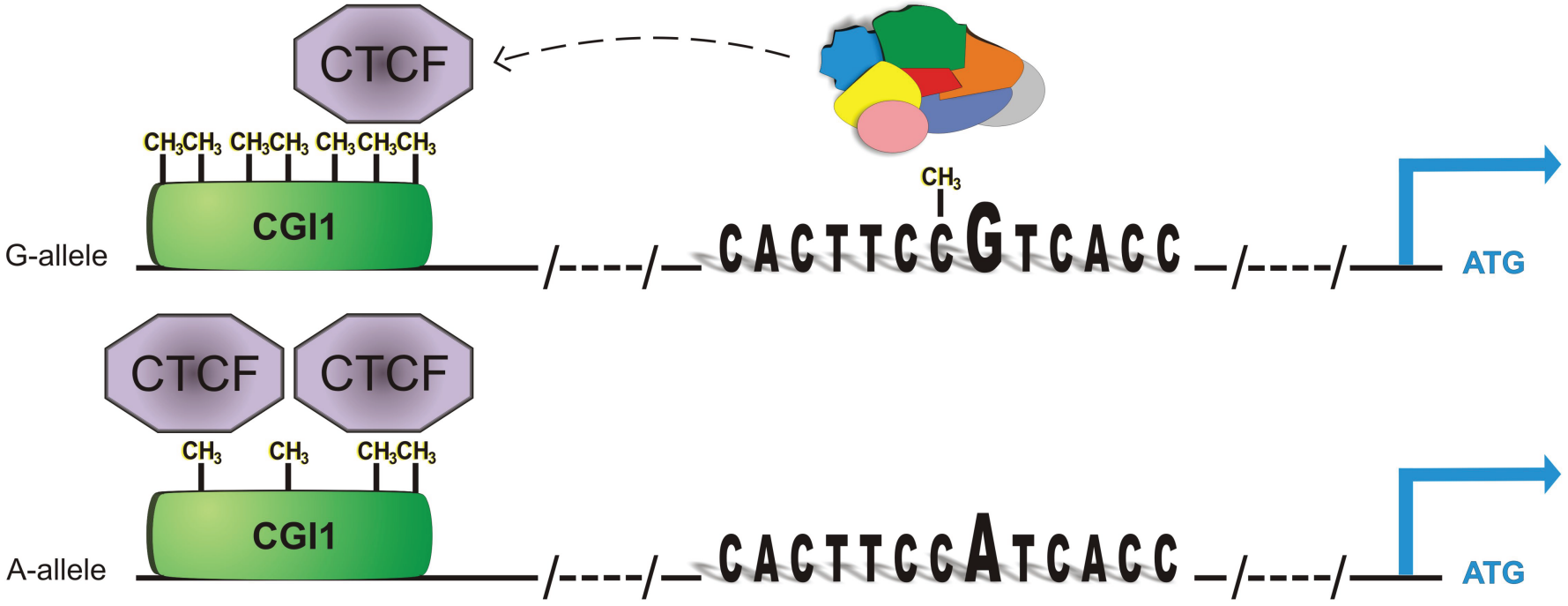

A
ATG
CpG-SNP
PEAR1
CGI1
CGI2
P CGI1
P Intron1
P CGI2
B
D0
D7
D14
n.s.
n.s.
PEAR1 DNA methylation (%)
100
50
0
* 
P CGI1
P Intron1
P CGI2
C
D0
D7
D14
P CGI1 DNA mehtylation (%)
60
40
20
0
CpG1
CpG2
CpG3
CpG4-5
CpG6
CpG7-8
CpG11-13
CpG14
CpG15-16
CpG18
CpG19-20
CpG21
CpG22
CpG23
CpG24
CpG25-26
CpG27
CpG28-29
CpG30
CpG31-32
D
GG
AA+AG
P CGI1 DNA methylation (%)
60
40
20
0
CpG1
CpG2
CpG3
CpG4-5
CpG6
CpG7-8
CpG11-13
CpG14
CpG15-16
CpG18
CpG19-20
CpG21
CpG22
CpG23
CpG24
CpG25-26
CpG27
CpG28-29
CpG30
CpG31-32

Figure 6


CTCF
CH3CH3 CH3CH3 CH3CH3CH3
CGI1
G-allele
CH3
CACTTCCGTCACC
ATG
CTCF
CTCF
CH3 CH3 CH3CH3
CGI1
A-allele
CACTTCCATCACC
ATG